

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Informatique

École doctorale : Information, Structures, Systèmes

Unité de recherche : Laboratoire d'Informatique, de Robotique et de Microélectronique de
Montpellier (LIRMM), France

**Prédiction d'évènements cliniques inattendus : représentation
latente du parcours patient par graphes temporels et
visualisation de trajectoires**

Présentée par Hugo LE BAHER
Le 18 Novembre 2024

Sous la direction de Jérôme AZÉ,
Sandra BRINGAY et Pascal PONCELET

Devant le jury composé de

Thomas GUYET, Chargé de recherche, INRIA Lyon

Lina SOUALMIA, Professeure des universités, LITIS, Université de Rouen

Philippe LENCA, Professeur, IMT Atlantique

Pascal PONCELET, Professeur des universités, LIRMM, Université de Montpellier

Sandra BRINGAY, Professeure des universités, LIRMM, Université Paul Valéry Montpellier 3

Jérôme AZÉ, Professeur des universités, LIRMM, Université de Montpellier

Rapporteur

Rapporteuse

Examineur

Directeur

Co-directrice

Co-encadrant



UNIVERSITÉ DE
MONTPELLIER

Résumé

La numérisation des systèmes d'information hospitaliers se généralise, permettant l'intégration croissante des méthodes de visualisation interactive et d'apprentissage automatique dans les systèmes d'aide à la décision. Cette évolution facilite l'anticipation des risques critiques chez les patients sous surveillance et contribue à réduire la charge de travail des soignants. Cependant, les Dossiers Patients Informatisés (DPI) contiennent des informations volumineuses, hétérogènes et temporelles. L'application de ces méthodes à des données aussi complexes pose deux principaux défis : (1) Concevoir des modèles capables de s'adapter à ces caractéristiques spécifiques ; (2) Fournir des outils permettant de comprendre ces données complexes et les résultats produits par ces modèles.

Dans cette thèse, nous introduisons GLIMPSE-Med, une interface de visualisation des événements temporels et hétérogènes du DPI sous forme de chronologie interactive, intégrant un score généré par un modèle prédictif. L'évaluation de cette interface, réalisée auprès de professionnels de santé, a confirmé qu'elle répondait à deux besoins essentiels : (1) Évaluer la qualité des données collectées dans un dossier médical informatisé ; (2) Estimer l'évolution de l'état du patient au fil du temps.

Nous présentons également une modélisation originale basée sur les avancées récentes en convolution de graphes. Dans notre approche, le parcours patient est modélisé sous forme de graphe où chaque sommet représente un événement, et où les proximités temporelles sont représentées par des arcs pondérés. Cette modélisation a été testée à l'aide de méthodes d'apprentissage par convolution de graphes pour prédire le décès dans les 24 heures, en appliquant une tâche de classification de graphes sur un ensemble de données réelles. Les résultats obtenus ont été comparés avec succès aux performances des méthodes de l'état de l'art.

Les travaux de cette thèse sont organisés en deux parties :

(1) La première partie se concentre sur la visualisation. Elle propose un état de l'art sur la visualisation des chronologies multivariées, tant dans le domaine médical que non médical, et inclut la présentation de GLIMPSE-Med ainsi que son évaluation par des professionnels de santé ; (2) La seconde partie aborde l'utilisation des méthodes d'apprentissage automatique. Elle commence par un état de l'art des méthodes d'apprentissage par convolution de graphes, suivi de la présentation d'une nouvelle méthode de représentation du parcours patient sous forme de graphes. Cette partie se termine par une étude des représentations latentes générées, visant à évaluer leur capacité à révéler des caractéristiques communes chez les patients.

Mots-clés – Dossier Patient Informatisé, Apprentissage automatique, Visualisation, Réseaux de neurones pour graphes

Abstract

The digitization of hospital information systems is becoming widespread, enabling the increasing integration of interactive visualization methods and machine learning into decision support systems. This development facilitates the anticipation of critical risks in monitored patients and helps reduce the workload of healthcare providers. However, Electronic Health Records (EHRs) contain large, heterogeneous, and temporal data. Applying these methods to such complex data presents two main challenges : (1) Designing models that can adapt to these specific characteristics ; (2) Providing tools to understand these complex data and the results produced by these models.

In this thesis, we introduce GLIMPSE-Med, an interactive timeline-based visualization interface for temporal and heterogeneous events in the EHR, incorporating a score generated by a predictive model. The evaluation of this interface, conducted with healthcare professionals, confirmed that it meets two essential needs : (1) Assess the quality of data collected in an electronic health record ; (2) Estimate the patient's condition over time.

We also present an original modeling approach based on recent advancements in graph convolution. In our approach, the patient journey is modeled as a graph where each node represents an event, and temporal proximities are represented by weighted edges. This modeling has been tested using graph convolution-based learning methods to predict death within 24 hours, employing a graph classification task on a real dataset. The obtained results were successfully compared with state-of-the-art methods.

The work of this thesis is organized into two parts :

(1) The first part focuses on visualization. It provides a state-of-the-art review on the visualization of multivariate timelines, both in medical and non-medical fields, and includes the presentation of GLIMPSE-Med along with its evaluation by healthcare professionals ; (2) The second part addresses the use of machine learning methods. It begins with a state-of-the-art review of graph convolution methods, followed by the presentation of a new method for representing patient journeys as graphs. This part concludes with a study of the generated latent representations, aiming to assess their ability to reveal common characteristics among patients.

Keywords – Electronic Health Records, Machine Learning, Visualization, Graph Neural Networks

Cette thèse a été financée par une bourse CIFRE, par la société 5 DEGRÉS et réalisée en collaboration avec le LIRMM-UM-CNRS et le CHU de Montpellier.

Tout d'abord, je souhaite remercier les membres de mon jury. Thomas Guyet et Lina Soualmia pour l'évaluation de cette thèse, leurs remarques et questions inspirantes. De la même manière, je remercie Philippe Lenca d'avoir accepté l'invitation à faire partie du jury. Je remercie aussi les membres de mon comité de suivi individuel Marc Chaumont et Adrien Coulet pour leur soutien et leurs conseils au long de ces années.

Je remercie l'entreprise 5 DEGRÉS pour le financement de mes travaux. En particulier, je remercie Antonin Goude pour son aide au lancement de ce projet.

Je tiens à remercier toutes les personnes rencontrées au Centre Hospitalier Universitaire de Montpellier. En premier lieu, Boris Jung à l'initiative scientifique du projet. Merci à Caroline Dunoyer pour sa disponibilité, ses indications et les discussions riches sans lesquelles ma compréhension des dynamiques et des problématiques hospitalières aurait été impossible. Je remercie aussi Maxime Fosset. Son expérience précieuse des services de réanimation ainsi que sa curiosité et ses compétences en apprentissage automatique ont été un terrain fertile pour les expérimentations préparatoires. Je remercie aussi l'ensemble des personnes que j'ai pu côtoyer au Service de Santé Publique pour leur générosité et le partage de leurs connaissances. En particulier, merci à Dominique Delseny et David Pellecuer pour leur enthousiasme et leur aide sur la technique. Merci aussi à l'équipe ERIOS pour les échanges et leurs retours à propos des systèmes d'aide à la décision. Enfin, je remercie l'ensemble des personnes qui se sont prêtés à l'évaluation de mon interface pour le temps qu'ils m'ont accordé, leur intérêt et leurs retours sincères.

Tout particulièrement, j'adresse un remerciement à Jérôme Azé, Sandra Bringay et Pascal Poncelet pour l'opportunité qu'ils m'ont offerte. Je les remercie sincèrement pour leur suivi infaillible, leur patience et leurs innombrables conseils scientifiques, méthodologiques, techniques et humains pendant l'encadrement de cette thèse. Chacun à leur manière, ils ont contribué à mettre en place un environnement sain qui a permis ma progression. Merci à tous les collègues rencontrés au Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier, à l'Université de Montpellier, à l'Université Paul-Valéry-Montpellier-III et à l'Institut Universitaire de Technologie de Béziers. Je remercie les membres passés et actuels des équipes Advanse et Web3 avec qui j'ai pu échanger, ponctuellement ou quotidiennement, sur la science et le reste. Merci à Nancy Rodriguez pour son accompagnement au lancement du projet et à Arnaud Sallaberry pour son aide d'une grande valeur sur la visualisation. Merci à Vincent Raveneau pour sa générosité et pour le soutien indéfectible qu'il a pu me fournir, quel que soit le sujet. Enfin, je remercie Leonardo Moros pour l'entraide et la complicité constantes tout au long de ces années. Il est toujours rassurant d'avoir à ses côtés un camarade qui connaît des succès et des épreuves semblables.

Merci, à mes amis et ma famille. De Charente, de Bordeaux, de Nantes, de Montpellier, ou de contrées plus exotiques. À ceux qui ont tendu leurs oreilles dans les instants de peine. À ceux qui ont ri dans les moments de joie.

À ma mère Sylvie Le Baher et mon père Hubert Le Baher, qui m'ont toujours encouragé et manifesté leur confiance. Je suis reconnaissant des sacrifices auxquels vous avez consenti pour que je poursuive mes études sereinement. L'aboutissement de ce travail est aussi le vôtre.

À la mémoire de Hubert Le Baer.

TABLE DES MATIÈRES

1	Introduction	1
1.1	Contexte	2
1.2	Spécificités liées aux données cliniques du Dossier Patient Informatisé	3
1.2.1	Contenu du Dossier Patient Informatisé	3
1.2.2	Les données cliniques de la thèse	4
1.3	Challenges liés à la visualisation et à la prédiction à partir des données du DPI	5
1.3.1	Challenges liés à la visualisation du DPI	5
1.3.2	Challenges liés à la prédiction d'évènements cliniques	6
1.4	Objectifs	9
1.5	Contributions	9
1.6	Organisation de la thèse	13
2	Visualisation d'enregistrements cliniques	15
2.1	Visualisation de séries temporelles multivariées d'un individu : une étude exploratoire de la littérature	16
2.1.1	Motivations et état de l'art	18
2.1.2	Méthode	19
2.1.2.1	Sources d'information et définition de la requête	19
2.1.2.2	Pré-sélection semi-automatique des articles	20
2.1.2.3	Recherche manuelle des articles par citations	21
2.1.2.4	Analyse par un évaluateur	21
2.1.2.5	Extraction des informations dans les articles	22
2.1.3	Résultats	24
2.1.4	Discussion	24
2.2	GLIMPSE-Med	27
2.2.1	Analyse des besoins	27
2.2.2	Conception de l'interface	30
2.2.3	Étude expérimentale	35
2.2.4	Résultats	37
2.2.4.1	Analyse quantitative	37
2.2.4.2	Analyse qualitative	38
2.2.5	Discussion	40
2.3	Conclusions et Perspectives	42
3	Graphe temporel pour la modélisation du parcours patient	45
3.1	Définitions	46
3.1.1	Graphe	46

3.1.2	Apprentissage de représentations de graphes	47
3.2	Paradigme encodeur-décodeur	49
3.3	Techniques d'apprentissage de représentations de graphes	52
3.3.1	Approches historiques	52
3.3.2	Marches aléatoires	53
3.3.3	Premiers réseaux de neurones pour les graphes	54
3.3.4	Réseaux convolutifs	54
3.3.5	Extensions de la convolution	56
3.3.6	Synthèse de l'historique	57
3.4	Graphes appliqués au médical	58
3.5	Graphes temporels de patients	59
3.5.1	Le dossier patient en tant que graphe temporel	60
3.5.1.1	Méthode	60
3.5.1.2	Fonction d'oubli	62
3.5.1.3	Expérimentations et résultats	63
3.5.2	Représentation latente des évènements médicaux	66
3.5.3	Limitations	67
3.6	Conclusion et Perspectives	68
4	Conclusion	71
4.1	Résumé des contributions	72
4.1.1	Visualisation de données cliniques	72
4.1.2	Apprentissage de représentations de parcours patients	72
4.2	Perspectives	73
4.2.1	Visualisation de parcours cliniques	73
4.2.2	Apprentissage de représentations de parcours cliniques	75
4.2.3	Visualisation et apprentissage conjoints	79
	Bibliographie	83
	Annexe	105

TABLE DES FIGURES

1.1	Illustration de la mise en place d'outils d'aide à la décision pour la surveillance de l'état des patients.	6
1.2	Frise temporelle des approches de visualisation de données d'un patient.	7
1.3	Agrégation d'un parcours patient sous forme de vecteur.	8
1.4	Agrégation d'un parcours patient sous forme de matrice.	8
1.5	Aperçu de GLIMPSE-Med.	11
1.6	Résumé du processus de modélisation du DPI sous forme de graphe temporel pour la prédiction d'un évènement clinique.	12
1.7	Illustration du processus d'apprentissage et de validation d'un modèle supervisé sur des données cliniques.	12
2.1	Chronologie récapitulative des 21 interfaces identifiées dans l'étude.	17
2.2	Diagramme de flux décrivant le volume de documents obtenu à l'issue de chaque étape du processus PRISMA-ScR.	19
2.3	Capture d'écran de GLIMPSE-Med.	31
2.4	Illustration de l'interaction de zoom.	33
2.5	Représentation symbolique des enregistrements.	34
2.6	Composants du symbole représentant un enregistrement.	34
2.7	Illustration d'une infobulle apparaissant lors du survol d'enregistrement.	36
2.8	Preuve de concept de la vue compacte de plusieurs séries sous forme de diagramme de densité en violon non lissé.	41
3.1	Exemples de graphes aux caractéristiques variées et impact de leurs caractéristiques sur la représentation visuelle et sur les structures de données.	48
3.2	Présentation visuelle du principe d'apprentissage de représentations de graphes sous forme d'encodeur-décodeur.	51
3.3	Illustration du principe d'apprentissage avec des marches aléatoires.	53
3.4	Illustration d'un encodeur GCN.	57
3.5	Chronologie récapitulative des méthodes présentées.	58
3.6	Illustration du processus de transformation d'un ensemble d'enregistrement d'un patient en un graphe temporel.	61
3.7	Exemple de graphe obtenu avec une fonction de proximité à décroissance rapide.	64
3.8	Exemple de graphe obtenu avec une fonction de proximité à décroissance lente.	64
3.9	Quantification de l'importance des hyperparamètres sur les performances.	66
3.10	Représentation extraite du modèle GCN par rapport à quatre cas cliniques.	67

LISTE DES TABLEAUX

2.1	Mots-clés de recherche par thème utilisé dans le moteur PubMed. . . .	20
2.2	Synthèse du contenu des 21 articles identifiés à propos des fonctionnalités et des interactions.	25
2.3	Synthèse du contenu des 21 articles identifiés à propos des méthodes d'évaluation.	26
2.4	Comparaison de la réponse aux besoins de 21 approches identifiées dans la littérature qui abordent la représentation de séries temporelles multivariées d'un seul individu.	29
2.5	Questionnaire à choix multiples utilisé lors de l'évaluation par résolution de tâches.	37
2.6	Questionnaire F-SUS.	38
2.7	Récapitulatif des remarques données par les utilisateurs lors des échanges oraux.	39
3.1	Caractéristiques des données extraites de MIMIC-III.	65
3.2	Comparaison des résultats sur la prédiction de décès dans les 24h. . . .	65

INTRODUCTION

Sommaire

1.1	Contexte	2
1.2	Spécificités liées aux données cliniques du Dossier Patient Informatisé	3
1.2.1	Contenu du Dossier Patient Informatisé	3
1.2.2	Les données cliniques de la thèse	4
1.3	Challenges liés à la visualisation et à la prédiction à partir des données du DPI	5
1.3.1	Challenges liés à la visualisation du DPI	5
1.3.2	Challenges liés à la prédiction d'évènements cliniques	6
1.4	Objectifs	9
1.5	Contributions	9
1.6	Organisation de la thèse	13

1.1 Contexte

La numérisation des systèmes de santé a connu récemment un développement très important. Elle promet, entre autres, de réduire les erreurs de diagnostics et de traitements, d'éviter les tests redondants et de guider une allocation plus efficace des ressources de santé, tout en favorisant l'innovation dans les approches préventives et thérapeutiques [SRS20]. Cette numérisation est associée à des enjeux majeurs comme la protection des données personnelles¹, l'exploitation de ces données par de multiples professionnels de santé via le développement du Dossier Médical Partagé (DMP)² lancé en 2018 ou de l'espace numérique de santé³ lancé en 2022, ou encore l'exploitation de ces données numérisées par des intelligences artificielles [Car+23], etc.

Dans cette thèse, nous nous intéressons particulièrement au Dossier Patient Informatisé (DPI). Les dossiers papiers anciennement administratifs, médicaux et de soins cohabitent désormais dans des DPIs déployés au sein des structures de soins, suite à des bouleversements organisationnels, techniques et culturels. Selon la Haute Autorité de Santé (HAS), dès 2020, 97% des établissements ont déclaré un projet d'informatisation du dossier médical achevé ou en cours⁴. Plus qu'un passage du support papier au numérique, les DPIs ont profondément modifié les pratiques professionnelles [Tor22]. Ce dossier centralise au sein d'un établissement de santé l'ensemble des données de santé des patients pris en charge.

Les professionnels de santé les utilisent quotidiennement pour prendre des décisions critiques et surveiller les effets des traitements ou des actes médicaux, dans des environnements généralement sous pression. Ces dossiers améliorent la continuité et la pertinence des soins tout en facilitant la communication et la coordination entre les patients et les professionnels de santé. Les chercheurs les exploitent pour extraire des données médicales, par exemple pour améliorer le taux d'inclusion des patients dans les essais cliniques, tandis que les ingénieurs de la donnée les utilisent pour corriger des erreurs, etc.

Si les principales limites du DPI sont désormais en cours de résolution, comme son intégration au sein d'un même système d'informations et son interopérabilité, les implémentations actuelles du DPI peuvent encore être améliorées. Dans ce travail, nous allons nous focaliser sur des outils visuels et prédictifs qui intègrent les systèmes d'aide à la décision en s'appuyant sur le DPI pour assister les utilisateurs. Pour ce faire, nous proposons l'étude de deux directions de recherche complémentaires :

- **Direction de recherche 1 (DR1) : développer des visualisations de dossiers patients** pour améliorer la prise de décision clinique en offrant aux professionnels de santé une compréhension rapide et claire de données complexes et hétérogènes. Cela permet de gagner du temps et d'améliorer l'efficacité dans la recherche d'informations grâce à des outils avancés d'exploration, de recherche, et de synthèse ;

1. <https://www.cnil.fr/fr/le-rgpd-applique-au-secteur-de-la-sante> consulté le 25/06/24.
2. <https://www.dmp.fr> consulté le 25/06/24.
3. <https://www.monespacesante.fr> consulté le 25/06/24.
4. https://www.has-sante.fr/upload/docs/application/pdf/2023-09/fiche_pedagogique_dpa.pdf consulté le 25/06/24.

- **Direction de recherche 2 (DR2)** : concevoir des **outils prédictifs** pour détecter les événements cliniques, personnaliser les traitements en fonction des caractéristiques des patients, proposer des actions de prévention, et optimiser l'allocation des ressources médicales.

Dans ce chapitre, nous allons tout d'abord motiver nos travaux en décrivant la nature spécifique des données cliniques contenus dans le DPI dans la section 1.2 et les challenges associés dans la section 1.3. Puis, nous fixerons les objectifs de la thèse dans la section 1.4, listerons nos contributions dans la section 1.5 et donnerons l'organisation du manuscrit dans la section 1.6.

1.2 Spécificités liées aux données cliniques du Dossier Patient Informatisé

1.2.1 Contenu du Dossier Patient Informatisé

Quelque soit la structure de santé, les enregistrements disponibles dans le DPI sont hétérogènes à la fois par les réalités différentes qu'ils représentent et par les formes variées dans lesquelles on les retrouve. Pour être enregistrés, les données, les mesures, les connaissances et les concepts doivent être traduits dans des formats compréhensibles par la machine. La nature des données a aussi un impact sur les techniques d'apprentissage utilisées pour la prédiction automatique. Nous distinguons principalement les données suivantes :

- **Informations démographiques**, comme l'âge et le sexe.
- **Mesures physiologiques**, mesurées ponctuellement ou à intervalles réguliers sur le patient, souvent par le biais d'un appareil de mesure, qui servent à surveiller le fonctionnement de son organisme. Ces mesures sont principalement numériques, associées à des unités de mesure physique comme les Battements Par Minute (BPM) pour le rythme cardiaque ou en degrés Celsius pour la température.
- **Biologie**, l'ensemble des prélèvements biologiques qui visent à quantifier des dosages au sein du sang ou de l'urine par exemple. Les résultats sont majoritairement numériques exprimés en doses, accompagnés de plages adaptées au profil du patient qui décrivent les intervalles dans lesquelles les valeurs sont attendues.
- **Microbiologie**, l'ensemble des prélèvements biologiques qui visent à détecter la présence de micro-organismes chez un patient, principalement pour identifier les infections. Le résultat est binaire (présence ou absence).
- **Médicaments**, prescrits au patient durant son séjour. Cette donnée est complexe puisque plusieurs informations sont enregistrées dans le système comme la prescription, la délivrance et la prise. De plus, un seul enregistrement est composé de la molécule prescrite, d'un dosage, d'une fréquence et de la durée de traitement.
- **Diagnosics**, qui décrivent les maladies ou affections identifiées chez un patient. Dans de nombreux pays dont la France avec la tarification à l'activité (T2A), les hôpitaux sont rémunérés en échange de l'inventaire de leurs activités. Ainsi, les

diagnostics sont obtenus dans un cadre administratif par codage médical, qui consiste à identifier les diagnostics par séjours selon la Classification Internationale des Maladies (CIM⁵), une codification standardisée à l'échelle internationale.

- **Actes**, qui décrivent les interventions médicales réalisées sur le patient. De la même manière et pour le même objectif que pour les diagnostics, les actes sont annotés selon des systèmes de codes standards comme la *Current Procedural Terminology* (CPT⁶) aux États-Unis.
- **Organisation**, c'est-à-dire la description de la structure hospitalière à laquelle les données sont associées lors de la trajectoire du patient. Cela inclut les informations de ses **séjours**, du **type d'unités** médicales visitées et les **transferts** entre ces unités.

Dans le cadre de nos travaux, nous utilisons les données décrites précédemment. Toutefois, le DPI contient aussi d'autres types de données. Par exemple, les **notes cliniques** sous forme de textes, qui regroupent les comptes rendus de sortie ou les transmissions infirmières, sont traitées par le biais du traitement du langage naturel [Bea+23; Zho+24]. L'**imagerie médicale** [Lit+17; Cel+23] regroupe l'ensemble des images acquises par des dispositifs comme la radiographie, l'échographie, l'imagerie par résonance magnétique (IRM) ou la magnétoencéphalographie (MEG), etc. Les **données génétiques** incluent les résultats des tests génétiques comme les analyses ADN, les mutations génétiques identifiées, ou encore les prédispositions héréditaires à certaines maladies.

1.2.2 Les données cliniques de la thèse

Ces travaux s'inscrivent dans le cadre d'une thèse CIFRE, financée par la société INGIMA/5 Degrés, qui s'est déroulée entre le LIRMM et le CHU de Montpellier. Le CHU est un établissement public situé dans le Sud de la France en région Occitanie. Il réunit 8 établissements de soins principaux pour un total de 2 070 lits pour 242 314 hospitalisations en 2022⁷. Les données cliniques sont réparties au sein du vaste système d'information du CHU. Le Service de Santé Publique, anciennement Département d'Information Médicale (DIM), est chargé d'extraire, d'uniformiser et de valider ces données à froid dans un entrepôt pour différentes applications, c'est-à-dire sans qu'elles soient en cours d'utilisation dans les services. Cela implique des processus de validation tels que la vérification de l'intégrité des données, le contrôle des erreurs ou des anomalies, et la mise en conformité avec les standards de qualité des données, le tout sans interférer avec les opérations quotidiennes ou les flux de données en temps réel. Parmi les missions de ce service, on trouve notamment : (1) le codage des actes et des diagnostics pour des activités administratives comme le financement de l'hôpital dans le cadre de la tarification à l'activité (T2A) ; (2) Le support aux activités de recherche

5. <https://www.who.int/fr/news/item/11-02-2022-icd-11-2022-release> consulté le 14/06/24.

6. <https://www.ama-assn.org/amaone/cpt-current-procedural-terminology> consulté le 27/06/24.

7. Issu des chiffres clés de 2022 : <https://www.chu-montpellier.fr/fr/a-propos-du-chu/decouvrir/chiffres-cles/> consulté le 17/06/24.

des cliniciens dans la constitution de cohortes et d'études; (3) La veille de l'activité du centre hospitalier pour les orientations stratégiques et la répartition du budget.

Les données du CHU de Montpellier ont servi de base pour l'établissement des directions de recherche (DR1 et DR2) décrites dans la section 1.1 et l'identification des cas à traiter. Cependant, dans l'objectif de nous comparer à des approches existantes de la littérature, nous avons fait le choix de présenter les résultats de recherche sur la base MIMIC-III. Il s'agit une base de données mise à disposition de la communauté scientifique, publiquement et gratuitement sur demande. Elle contient les données cliniques des patients admis au Beth Israel Deaconess Medical Center à Boston dans le Massachusetts de 2001 à 2012. Son principal atout réside dans le travail d'anonymisation des données. Les données sont disponibles sur la plateforme Physionet [Gol+00] sous le nom *MIMIC-III Clinical Database* [JPM15; Joh+16]. La base répertorie 38 597 patients adultes pour un total de 53 423 séjours. Cette base contient des données très similaires à celles du CHU de Montpellier (données démographiques, diagnostics, paramètres vitaux, comptes rendus de sorties, notes infirmières, etc.). Toutefois, ces données sont en anglais et sont restreintes à celles d'un service de soins critiques, ce qui limite les enregistrements disponibles et les cas d'usages.

1.3 Challenges liés à la visualisation et à la prédiction à partir des données du DPI

La numérisation des processus de soins a conduit à de grands volumes de données patient numérisées, permettant de nouveaux usages et nécessitant de nouveaux développements tels que décrits dans la section 1.1. Dans ce manuscrit, nous allons nous focaliser sur les challenges liés à la visualisation et à la prédiction à partir des données cliniques présentes dans le DPI, tel que décrit dans la figure 1.1.

1.3.1 Challenges liés à la visualisation du DPI

L'évolution de l'état de santé d'un patient est généralement suivie en examinant l'évolution de plusieurs variables hétérogènes au fil du temps. Appréhender cette évolution par le biais d'une interface spécialisée peut être complexe, même pour les professionnels de santé avec de l'expérience. La temporalité des données, la diversité des sources et l'hétérogénéité des enregistrements stockés dans des bases de données volumineuses, constituent autant d'obstacles à une compréhension globale du parcours d'un patient. Bien que les systèmes interactifs soient indispensables et largement utilisés dans les établissements de santé, ils provoquent parfois des expériences utilisateur insatisfaisantes. Par exemple, des recherches ont révélé des problèmes tels que la fatigue due aux clics répétés et l'interruption des interactions avec les patients [Col18]. La présentation claire de ces données complexes et volumineuses dans un espace restreint sans surcharger visuellement l'utilisateur au sein de systèmes d'aide à la décision représente un véritable défi.

La littérature contient de nombreuses revues sur les interfaces de visualisation des données médicales. Dans cette thèse, nous avons choisi de nous focaliser sur la

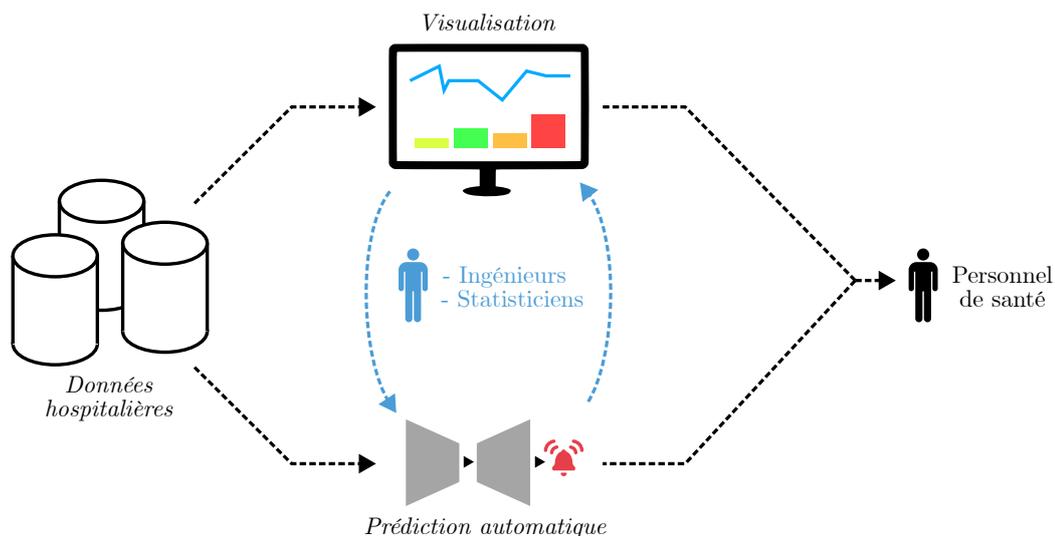


FIGURE 1.1 – Illustration de la mise en place d’outils d’aide à la décision pour la surveillance de l’état des patients. En parallèle, les données sont utilisées dans des systèmes de visualisation et de prédiction de risques à destination du personnel soignant. En bleu, nous montrons que les ingénieurs ou les statisticiens chargés du développement de ces interfaces, de la mise en place des modèles prédictifs ou de la qualité des données peuvent utiliser ces systèmes pour les enrichir mutuellement.

visualisation des données d’un seul patient sous la forme d’une frise temporelle et notons l’absence d’une synthèse spécifique dédiée à cette problématique particulière. La Figure 1.2 illustre deux interfaces identifiées dans le chapitre 2 positionnées dans une frise chronologique. L’analyse de la littérature montre que de tels systèmes existaient déjà en 1998 avec *LifeLines* [Alo+98] et que cette problématique a continué d’être développée. On constate que cette interface présente uniquement des données temporelles catégorielles. De plus, l’axe temporel affiche l’ensemble des données, ce qui n’est pas adapté pour des parcours s’étendant sur plusieurs années. *ClinicalPath* [Lin+23] montre des données hétérogènes via une synthèse jour par jour. Cependant, il n’est pas possible d’observer l’évolution des données à une échelle inférieure. Par ailleurs, on observe qu’aucune de ces interfaces n’intègre des résultats de modèles de prédiction.

1.3.2 Challenges liés à la prédiction d’évènements cliniques

La littérature présente de nombreuses tâches prédictives réalisées sur des données du DPI. On observe des tâches très représentées et qui ont un lien fort avec la dégradation de l’état du patient comme la prédiction de la mortalité [JPM17; Pur+17; Raj+18; Dub+18; Liu+19; El+20; RBH22], de la réadmission à l’hôpital sous une certaine durée [Pur+17; Raj+18; Dub+18; HAR19; Liu+19; Bar+20; COP23], de la durée du séjour [Raj+18; Mek+20] et de diagnostics qui s’apparente au codage médical [Pur+17; Raj+18; Luc+21]. Il existe aussi un nombre important de travaux sur des diagnostics ou évènements médicaux particuliers comme, par exemple, l’hypertension [Sil+22], les conséquences d’un AVC [Heo+19] ou les effets de médicaments [DOW23]. Dans cette thèse, même si nos propositions peuvent être appliquées à des tâches variées, nous nous concentrons plus particulièrement sur la tâche de prédiction de la mortalité.

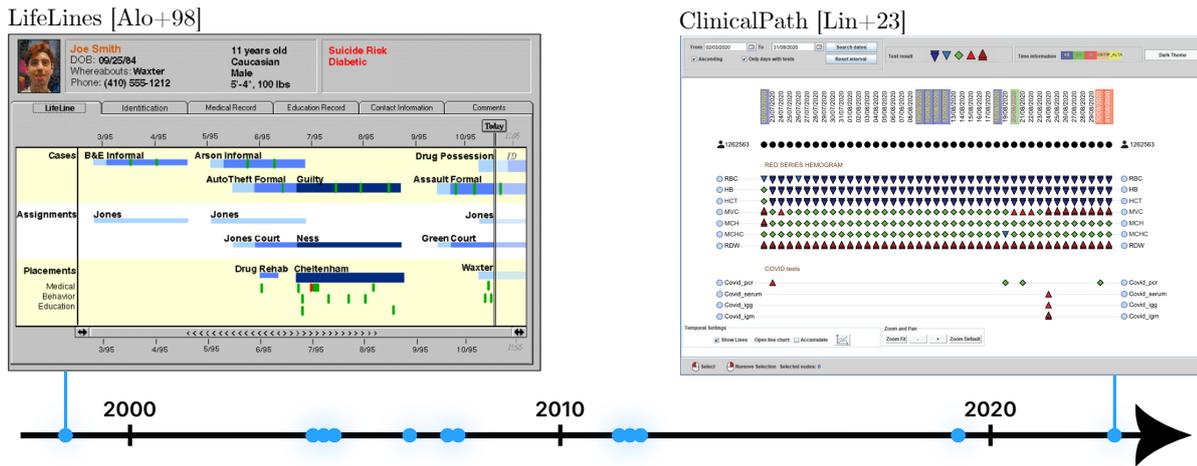


FIGURE 1.2 – Frise temporelle des approches de visualisation de données d’un patient identifiées dans le chapitre 2. Deux interfaces sont montrées en détail, la plus ancienne et la plus récente.

Indépendamment de la tâche ou du modèle utilisé, il est difficile de trouver une représentation numérique du parcours d’un patient qui synthétise les informations observées. Ce défi est particulièrement complexe lorsque les données possèdent de nombreuses dimensions [SRS20] et évoluent dans le temps. Paradoxalement, la quantité élevée de modalités peut également constituer un obstacle. Ce problème est couramment désigné sous le terme de la malédiction de la dimension (*curse of dimensionality*), et il est particulièrement amplifié dans le contexte des séries temporelles [VF05]. Parmi les nombreux effets observés, la comparaison entre patients devient difficile car chaque patient est décrit par un ensemble unique de modalités (prélèvements, mesures physiologiques, actes, diagnostics, etc.), qui est rarement identique à celui d’un autre patient. Des techniques de pré-traitement, telles que la réduction de dimension ou l’imputation des données manquantes, sont nécessaires pour exploiter efficacement ces données complexes. En fonction de l’architecture du modèle utilisée, un modèle génère des représentations intermédiaires plus ou moins informatives quant à la tâche visée. Dans tous les cas, la représentation en entrée du modèle conditionne la qualité de l’apprentissage et la performance prédictive du modèle généré.

Afin d’obtenir une représentation numérique d’un parcours patient, les sections suivantes présentent deux méthodes d’agrégation des parcours, respectivement pour des modèles traditionnels ou des modèles séquentiels.

Parcours patient agrégé La méthode la plus simple consiste à agréger l’ensemble des données de chaque modalité dans un canal dédié, que ce soit pour un patient ou pour un séjour. Par exemple, comme illustré sur la figure 1.3, les données disponibles à gauche sont agrégées en utilisant trois fonctions : la valeur minimale, moyenne et maximale. Cet ensemble de valeurs décrit le parcours du patient et est ensuite utilisé comme entrée pour un modèle d’apprentissage afin de produire une prédiction \tilde{Y} . Ainsi, de nombreuses études utilisent cette approche pour résoudre des tâches variées en appliquant différents algorithmes d’apprentissage. Cela peut être des modèles traditionnels comme les Forêts aléatoires [Bre01] utilisées dans [Kar+12; Gru+18; DPA18;

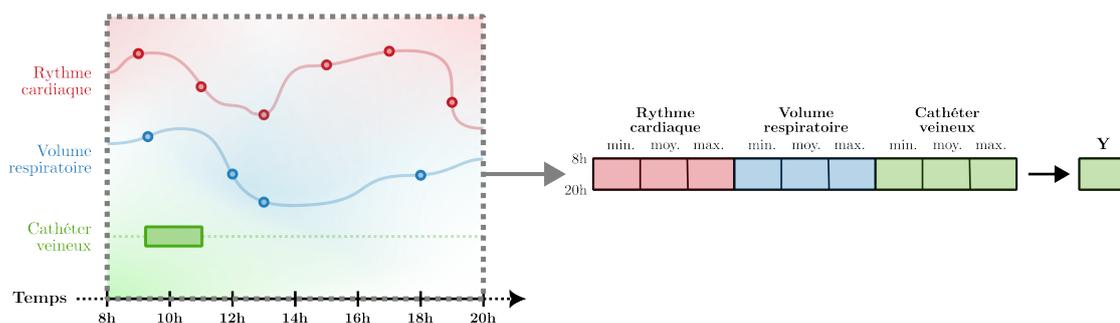


FIGURE 1.3 – À gauche, un exemple de parcours patient décrit par 3 modalités différentes. À droite, les données sont agrégées selon 3 fonctions d'agrégation pour être passées dans un modèle d'apprentissage afin de prédire une étiquette Y pour le patient en question.

[El+20; Bha+23] ou les Machines à vecteurs de support (*Support Vector Machines*) [CV95] utilisées dans [Him+09; Kar+12; Dai+15; DPA18; Bha+23]. En outre, des réseaux de neurones profonds sont également exploités, comme dans [KK17; Pur+17; Raj+18; DPA18; Bha+23]. La principale limite de cette approche réside dans l'importante agrégation des données. Par exemple, une moyenne peut difficilement capturer les évolutions soudaines des facteurs vitaux. La qualité de l'agrégation dépend donc de l'expressivité des fonctions d'agrégation utilisées.

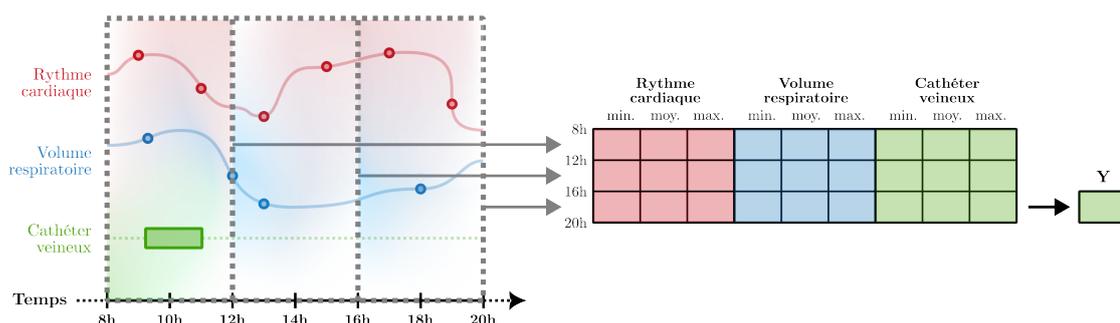


FIGURE 1.4 – À gauche, un exemple de parcours patient décrit par 3 modalités différentes. À droite, les données sont agrégées selon 3 fenêtres temporelles successives de 2 heures. Dans chaque fenêtre temporelle les données sont agrégées selon 3 fonctions d'agrégation pour être passées dans un modèle d'apprentissage afin de prédire une seule étiquette Y pour le patient en question.

Modèles séquentiels Une extension de la méthode décrite ci-dessus consiste à extraire des représentations sous forme de séquences pour observer les évolutions dans le temps. Ce principe est illustré à la Figure 1.4, où plusieurs synthèses sont réalisées de manière similaire à la Figure 1.3. Cependant, au lieu d'agréger l'ensemble des données disponibles, l'agrégation se fait sur des fenêtres temporelles successives. Pour traiter ces synthèses bidimensionnelles, qui prennent la forme de matrices, il est nécessaire d'utiliser des modèles adaptés, tels que les réseaux de neurones récurrents, comme le *Long Short-Term Memory* [HS97] (LSTM), utilisés dans [Li+20b; Ash+19; Har+19; Liu+22; RBH22] et les modèles à base de convolution [Kat+18; Wan+19a; Wan+24].

Bien que ces modèles puissent offrir de bonnes performances prédictives, la construc-

tion des séquences nécessite de discrétiser la dimension temporelle continue. Cela implique de découper les données en fenêtres successives selon une unité temporelle uniforme, ce qui peut entraîner une perte d'informations importantes. Par exemple, avec des fenêtres temporelles de taille fixe, il est difficile de représenter sur une même échelle temporelle des prescriptions modifiées quotidiennement et des mesures de rythme cardiaque, enregistrées chaque minute. Bien que cette technique conserve la séquentialité, c'est-à-dire l'ordre des événements, elle peut altérer les relations temporelles entre les enregistrements. Dans cette thèse, nous avons choisi d'explorer l'utilisation de réseaux de neurones basés sur des convolutions de graphes pour surmonter ces limitations.

1.4 Objectifs

À travers cette thèse, nous visons à améliorer l'utilisation et la réutilisation des données cliniques contenues dans les dossiers patients informatisés en explorant les deux directions de recherche (DR1 et DR2) décrites dans la section 1.1.

Objectif 1 : construire une interface de visualisation de la trajectoire d'un patient sous la forme d'une frise temporelle. Notre objectif est de développer des visualisations offrant à la fois une vue d'ensemble et des détails précis des données temporelles hétérogènes et multivariées pour simplifier la compréhension des parcours individuels des patients. Ces visualisations s'adressent à des professionnels de santé manipulant le DPI pour le suivi de patients ou la recherche médicale et des ingénieurs ou statisticiens responsables de la qualité des données.

Objectif 2 : proposer un modèle prédictif. Notre objectif est de développer un modèle capable de synthétiser des données multivariées et temporelles pour effectuer des prédictions sur des événements cliniques telles que le décès inattendu, l'arrêt cardiaque ou le succès de la réintubation, sans avoir besoin de définir *a priori* un découpage temporel.

Objectif 3 : intégrer le modèle prédictif dans l'interface. Le développement conjoint de l'interface et du modèle prédictif permet alors un enrichissement de ceux-ci. L'utilisation de la visualisation permet aux ingénieurs ou aux statisticiens d'explorer en détail les parcours patient et d'émettre des hypothèses sur la sélection ou les pré-traitements à réaliser sur les données afin de développer de meilleurs modèles. Par ailleurs, les modèles prédictifs identifient les données les plus significatives en extrayant des motifs et des récurrences, ce qui aide l'utilisateur lorsque ces informations sont traduites par un encodage visuel. Ce processus d'enrichissement mutuel entre l'interface et le modèle prédictif est illustré par la boucle bleue sur la Figure 1.1.

1.5 Contributions

Les contributions de cette thèse s'articulent autour de deux axes de recherche complémentaires (DR1 et DR2) décrits dans la section 1.1, et répondent aux trois objectifs définis dans la section 1.4. Elles se focalisent sur deux domaines étroitement liés : la visualisation des données et l'apprentissage automatique via des réseaux de

représentation de graphes.

- **Contribution 1 : état de l’art des visualisations de données cliniques.** Nous présentons une analyse détaillée de visualisations qui portent sur des données similaires à celles du DPI. En particulier, nous étudions les visualisations de séries temporelles multivariées d’un seul individu au sein de 21 interfaces. Cette analyse de la littérature exploratoire est réalisée selon le protocole PRISMA-ScR et permet d’identifier les fonctionnalités et les interactions qui aident ces interfaces à répondre aux besoins des cliniciens. Elle met également en lumière les limites et lacunes de ces approches, en les comparant à des interfaces utilisées dans des domaines autres que le médical.
- **Contribution 2 : développement et évaluation d’une interface de frise temporelle des données d’un patient.** Nous présentons GLIMPSE-Med, une interface visuelle pour l’affichage de données du DPI destinée à deux types d’utilisateurs : les professionnels de la santé et les ingénieurs de données/statisticiens. Un aperçu de cette interface est illustré dans la Figure 1.5, et ses fonctionnalités sont détaillées dans le chapitre 2. L’originalité de cette interface réside dans plusieurs aspects : (1) la compression visuelle de l’axe temporel par une distorsion qui s’adapte aux données et à l’interaction de l’utilisateur ; (2) la compression visuelle de l’axe temporel et de mesures discrètes ou continues à travers un affichage uniformisé, dont la navigation est réalisée au sein d’une seule fenêtre par le biais d’un menu hiérarchique interactif ; (3) l’intégration visuelle de la prédiction issue de modèles prédictifs qui évoluent dans le temps en mettant en évidence les facteurs d’importance. Nous avons mené une évaluation expérimentale de cette interface auprès de 14 professionnels de santé du CHU de Montpellier. Cette étude réalisée par le biais de résolution de tâches, d’un questionnaire fermé et de questions ouvertes a permis de recueillir des résultats quantitatifs et qualitatifs complets sur son utilisation.
- **Contribution 3 : modélisation de parcours patients sous forme de graphes temporels.** Nous proposons une modélisation originale du parcours patient sous forme de graphes temporels où chaque parcours est représenté par un graphe dont les sommets correspondent à des événements cliniques. Les relations temporelles entre ces événements sont exprimées par des arcs pondérés, définis à l’aide d’une fonction paramétrique de proximité temporelle. Chaque patient est ainsi représenté par une composante connexe unique.

Nous avons évalué cette modélisation en utilisant des méthodes d’apprentissage par convolution de graphes pour prédire le décès dans un délai de 24 heures, en nous basant sur une tâche de classification de graphes appliquée à un jeu de données réel issu de MIMIC-III. Nous proposons un processus complet pour apprendre un modèle complexe, intégré au sein du système d’information, qui utilise l’ensemble des descripteurs disponibles. Ce processus peut être appliqué à de nouveaux événements cliniques à prédire et permet de prendre en compte les spécificités de l’hôpital dans lequel il serait appliqué. La Figure 1.7 résume le processus d’apprentissage supervisé divisé en 5 étapes :

1. Les données hospitalières sont souvent réparties à travers des modèles

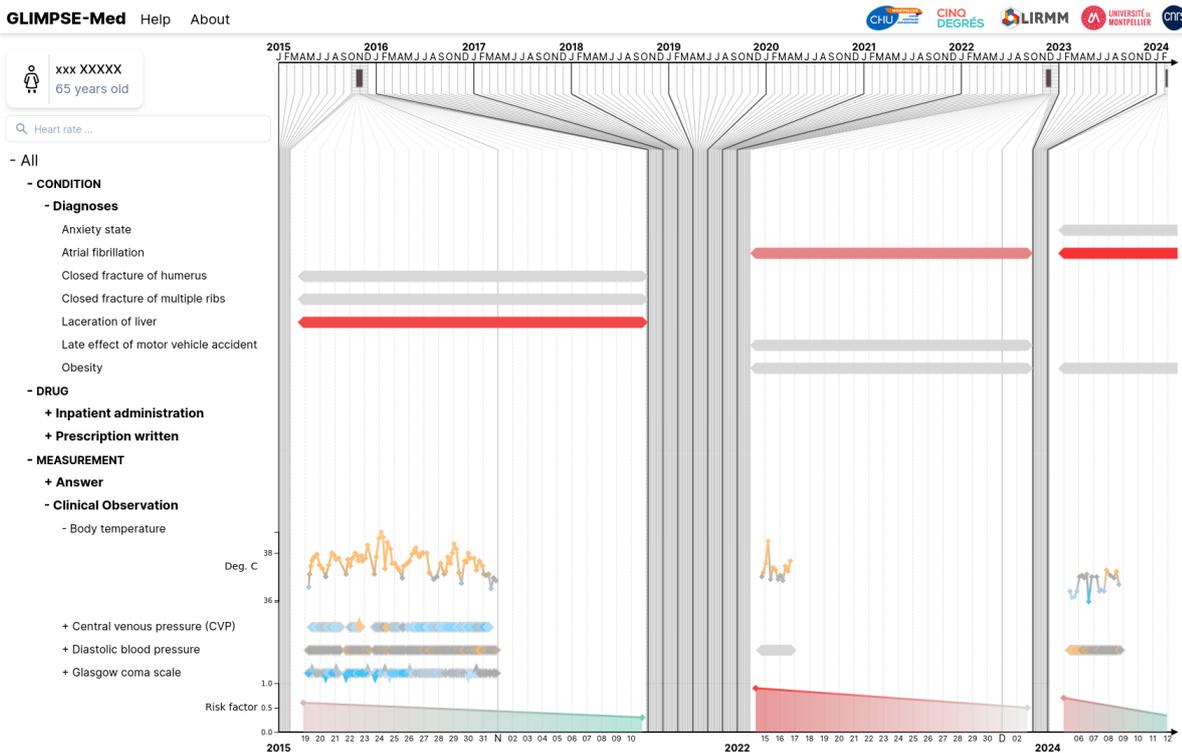


FIGURE 1.5 – Aperçu de GLIMPSE-Med, notre interface de visualisation d'enregistrements cliniques temporels et hétérogènes.

relationnels complexes voire distribuées. Il est donc nécessaire de rassembler et d'unifier les informations d'un patient à partir de ces sources hétérogènes.

2. Ces enregistrements doivent être associés à une étiquette de supervision représentant la réalité que nous cherchons à prédire. Cette étiquette peut être extraite directement des enregistrements disponibles dans le système d'information ou déterminée manuellement par des cliniciens. Le décès ou la survenue d'un arrêt cardiaque en sont des exemples.
3. Les enregistrements bruts sont transformés, ré-organisés ou agrégés pour être mis en entrée de modèles prédictifs. Dans notre cas, nous avons choisi d'utiliser une modélisation des parcours patients sous forme de graphes qui sera détaillée dans le chapitre 3.
4. Les données transformées sont ensuite utilisées pour entraîner le modèle, qui ajuste ses poids ou ses règles internes afin de générer de nouvelles prédictions d'étiquettes.
5. Les étiquettes prédites sont comparées aux étiquettes réelles pour générer des scores de performance établis au préalable dans le processus d'apprentissage. Atteindre des performances prédictives satisfaisantes, mesurées par des critères tels que la précision ou le rappel, est une condition essentielle pour le déploiement des systèmes prédictifs.

Nous avons comparé cette approche avec succès aux résultats des modèles sé-

quentiels de l'état de l'art. L'étude de l'impact des hyperparamètres permet de montrer l'impact de la modélisation sur les résultats. Un aperçu visuel de la modélisation et du processus d'apprentissage est présenté dans la Figure 1.6. Le processus complet est décrit en détail dans le chapitre 3. Ces travaux ont donné lieu à deux publications [LeB+23a ; LeB+23b].

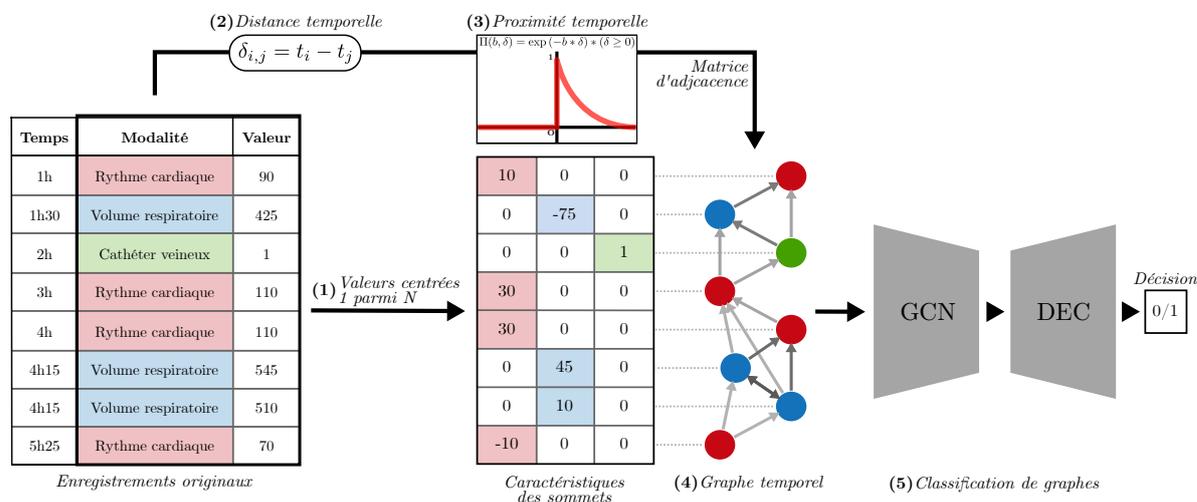


FIGURE 1.6 – Résumé du processus complet de modélisation du DPI sous forme de graphe temporel et d'apprentissage dans un réseau de neurones à convolution de graphes pour la production d'une prédiction d'un évènement clinique.

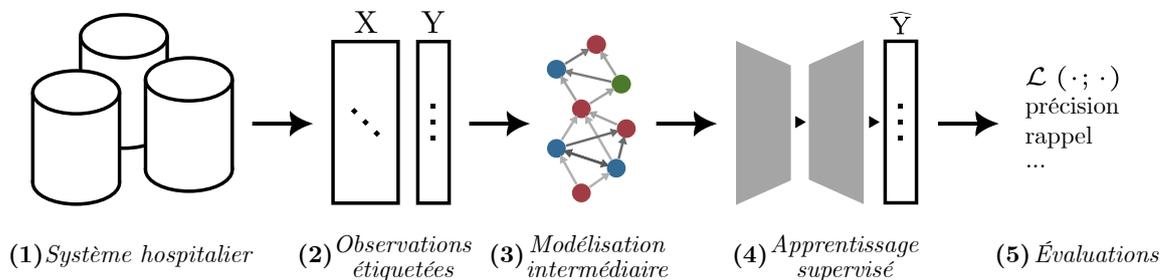


FIGURE 1.7 – Illustration du processus d'apprentissage et de validation d'un modèle supervisé sur des données cliniques en 5 étapes.

- **Contribution 4 : exploration de l'espace latent.** Nous cherchons à évaluer la qualité de l'espace latent généré par les réseaux à convolution en utilisant notre modélisation basée sur les graphes. L'objectif est d'examiner la capacité du modèle à apprendre une représentation riche et utile pour de nouvelles tâches et pour la construction de cohortes. Pour cela, nous avons analysé la séparabilité de la représentation intermédiaire par rapport à trois nouveaux cas cliniques différents de l'étiquette d'apprentissage, qui est le décès à 24 heures. Ces cas incluent la lésion cérébrale anoxique, l'anévrisme cérébral et l'hypertension. Nous montrons la limite de l'usage de la classification à l'échelle du graphe dans ce contexte. Ces travaux ont été présentés dans l'atelier SIMPA [LeB+23c].

1.6 Organisation de la thèse

Cette thèse s'organise dans deux chapitres principaux qui correspondent chacun aux deux directions de recherche (DR1 et DR2). Ces chapitres contiennent deux états de l'art spécifiques aux méthodes étudiées.

- **Visualisation de données** : dans le chapitre 2, nous présentons d'abord un état de l'art des techniques de visualisation des enregistrements cliniques. Nous nous concentrons particulièrement sur l'analyse des fonctionnalités et des interactions utilisées dans les interfaces de visualisation de séries temporelles multivariées pour un seul individu. Cette revue de la littérature a été réalisée selon le protocole standard PRISMA-ScR, et nous avons identifié 21 études tant dans le domaine médical qu'en dehors de celui-ci.

Les conclusions de cette revue de la littérature ont conduit au développement de GLIMPSE-Med, une interface conçue pour répondre aux besoins des professionnels de santé et des ingénieurs de données/statisticiens. Ce chapitre détaille les fonctionnalités de cette interface basée sur une frise temporelle et présente une étude expérimentale réalisée auprès de ses utilisateurs finaux : professionnels de santé, ingénieurs de données et statisticiens. Cette visualisation permet également d'afficher un score de risque, comme décrit dans le chapitre suivant.

- **Apprentissage automatique** : nous présentons dans le chapitre 3 un inventaire des méthodes d'apprentissage de représentation de graphes sous le prisme du paradigme encodeur-décodeur. À travers cet état de l'art, nous évaluons la pertinence de l'usage de ces techniques comme alternatives aux techniques séquentielles pour modéliser les parcours patients.

Dans ce chapitre, nous introduisons une nouvelle modélisation des parcours patients sous forme de graphes temporels. Nous présentons des résultats expérimentaux sur l'utilisation de cette modélisation avec des réseaux de neurones à convolution de graphes pour la prédiction de la mortalité. Nous comparons ces résultats avec ceux obtenus à l'aide de modèles séquentiels présents dans la littérature et examinons la représentation latente générée par nos modèles pour évaluer leur capacité à généraliser pour de nouvelles tâches.

Dans le chapitre 4, nous concluons cette thèse en présentant les perspectives offertes par les travaux que nous avons présentés. Ces perspectives s'articulent en trois parties : (1) la visualisation de parcours cliniques ; (2) la modélisation de parcours patients sous formes de graphes pour l'apprentissage de représentation d'enregistrements cliniques et (3) l'usage et le développement de modèles en combinant des systèmes visuels appropriés.

VISUALISATION D'ENREGISTREMENTS CLINIQUES

Sommaire

2.1	Visualisation de séries temporelles multivariées d'un individu : une étude exploratoire de la littérature	16
2.1.1	Motivations et état de l'art	18
2.1.2	Méthode	19
2.1.2.1	Sources d'information et définition de la requête	19
2.1.2.2	Pré-sélection semi-automatique des articles	20
2.1.2.3	Recherche manuelle des articles par citations	21
2.1.2.4	Analyse par un évaluateur	21
2.1.2.5	Extraction des informations dans les articles	22
2.1.3	Résultats	24
2.1.4	Discussion	24
2.2	GLIMPSE-Med	27
2.2.1	Analyse des besoins	27
2.2.2	Conception de l'interface	30
2.2.3	Étude expérimentale	35
2.2.4	Résultats	37
2.2.4.1	Analyse quantitative	37
2.2.4.2	Analyse qualitative	38
2.2.5	Discussion	40
2.3	Conclusions et Perspectives	42

Ce chapitre détaille nos travaux inscrits dans la direction de recherche 1 (DR1) sur le développement de visualisations de dossiers patients. En particulier, l'usage des systèmes existants soulève des problèmes. L'évaluation de l'état d'un patient à travers une interface dédiée est parfois difficile même pour un professionnel de santé expérimenté. La temporalité des données, la multiplicité des sources et l'hétérogénéité des enregistrements conservés dans des bases volumineuses sont autant d'obstacles à la compréhension globale du parcours d'un patient par les différents corps de métiers d'un établissement de soins. Même si les systèmes interactifs sont reconnus comme indispensables, leur utilisation entraîne aussi potentiellement des expériences utilisateur inconfortables. De plus, l'intégration progressive de modèles prédictifs au sein de ses systèmes pour guider la décision des soignants est parfois opaque. Il est donc indispensable de proposer des visualisations qui répondent à leurs besoins et adaptées au quotidien des professionnels de santé.

Dans ce chapitre, en réponse à l'objectif 1 présenté dans la section 1.4, nous proposons GLIMPSE-Med (*Graphical Longitudinal Interface for Multimodal Pathways Search and Expertise for Medicine*). Cette nouvelle interface de visualisation vise à simplifier la compréhension des parcours individuels des patients en compressant les séries temporelles multivariées dans une seule fenêtre. En réponse à l'objectif 3, cette interface intègre visuellement un modèle de prédiction.

Ce chapitre est organisé de la manière suivante : la section 2.1 présente une revue de la littérature sur la visualisation des séries temporelles multivariées, réalisée selon le protocole standard PRISMA-ScR. Ensuite, la section 2.2 décrit la conception et l'évaluation de GLIMPSE-Med notre interface de visualisation. Enfin, la section 2.3 vient clore le chapitre en offrant des perspectives d'amélioration.

2.1 Visualisation de séries temporelles multivariées d'un individu : une étude exploratoire de la littérature

Généralement, l'évaluation globale de l'état d'un patient se fait par l'étude des variations au cours du temps d'une ou de plusieurs variables hétérogènes. Par exemple, si le nombre de globule blancs mesurés et la température d'un patient augmentent conjointement, on suspecte une infection en cours. De même, si l'hémoglobine et le nombre de plaquettes décroissent rapidement, on suspecte alors une hémorragie chez le patient. L'objectif de cette section est d'identifier les fonctionnalités visuelles pour représenter des données temporelles et multivariées, c'est-à-dire impliquant plus de deux variables distinctes évoluant dans le temps. Plus précisément, notre objectif est d'explorer des visualisations qui offrent une vision globale et détaillée des données temporelles hétérogènes multivariées. Pour cela, nous avons mené une étude exploratoire de la littérature.

Nous présentons une chronologie récapitulative des interfaces de visualisation de données temporelles multivariées d'un individu, illustrée dans la Figure 2.1. Cette figure montre 7 interfaces représentatives de l'évolution de ces méthodes, proches de nos travaux, et sélectionnées selon le protocole décrit dans la section 2.1.2.

2.1. Visualisation de séries temporelles multivariées d'un individu : une étude exploratoire de la littérature

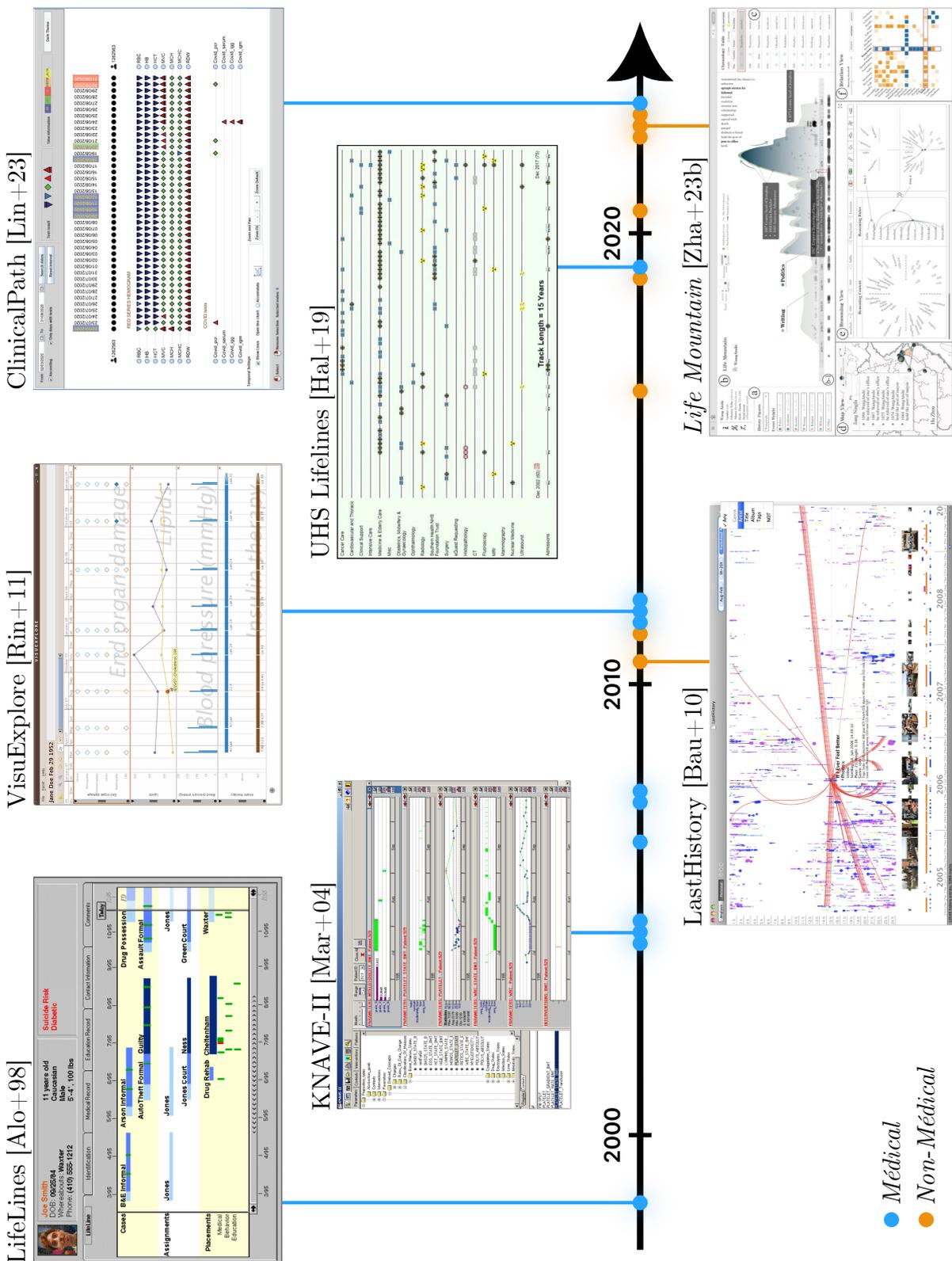


FIGURE 2.1 – Chronologie récapitulative des 21 interfaces identifiées dans cette étude. Chaque interface est représentée par un point le long de l'axe temporel, avec en bleu les applications au domaine médical et en orange les applications hors du domaine médical. Nous avons choisi de mettre en valeur 7 interfaces pour illustrer l'évolution au cours du temps des techniques de représentation de séries temporelles multivariées d'un seul individu.

Les motivations ainsi que l'état de l'art des analyses réalisées sur des sujets similaires sont abordés dans la section 2.1.1. Le protocole et les critères utilisés pour l'analyse sont détaillés dans la section 2.1.2. Les résultats obtenus lors de l'annotation manuelle sont exposés dans la section 2.1.3. Enfin, la section 2.1.4 synthétise ces résultats et conclut cette partie par une discussion.

2.1.1 Motivations et état de l'art

L'analyse visuelle de données temporelles multivariées est appliquée dans de nombreux domaines comme, par exemple, l'analyse financière [Yue+19], l'histoire [Zha+23], ou l'aide à la narration [Shi+23]. Dans le domaine médical, en particulier, la visualisation des données issues du DPI a fait l'objet de nombreuses recherches. Des études ont synthétisé les caractéristiques communes de ces visualisations ainsi que les fonctionnalités associées, qui serviront de base à cette revue de littérature.

Au cours des années 2010, plusieurs synthèses ont été conduites sur la représentation visuelle des données temporelles et multivariées dans le domaine médical. En particulier, le paragraphe suivant se concentre sur les synthèses réalisées à propos de la visualisation du DPI.

[CKS10] ont établi une taxonomie des méthodes d'affichage de ces données, en précisant si elles représentent un individu unique, tel qu'un patient, ou une cohorte entière. [Aig+11]¹ offrent une étude détaillée des approches de représentation des données temporelles pour un ou plusieurs individus. La portée de l'étude s'étend au-delà du seul domaine médical. [Rin+13] se concentrent sur 14 approches spécifiques appliquées au domaine médical et distinguent également le nombre d'individus. Enfin, [WBH15] utilisent le protocole PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses* [Moh+09]) pour une revue systématique des approches de visualisation de données du DPI, mettant en évidence une augmentation des publications dans ce domaine jusqu'en 2012, quand ces travaux ont été réalisés. Les auteurs soulignent la nécessité de recherches se concentrant sur la représentation de données volumineuses et multivariées dans un seul écran et sur la représentation de données manquantes, qui restent des problématiques d'actualité.

Depuis 2019, de nouvelles synthèses ont été proposées. [Sch+22] ont utilisé une extension du protocole PRISMA, PRISMA-ScR (*PRISMA extension for Scoping Reviews*) [Tri+18], pour décrire 22 approches axées sur des cohortes, ce qui diffère de l'approche centrée sur un patient unique, qui est l'objet de ce chapitre. [WL22] proposent une taxonomie détaillée de 51 travaux limités au domaine médical. Les auteurs soulignent que l'inclusion de modèles d'apprentissage automatique est récent. Ces travaux ne distinguent pas les approches pour les patients individuels de celles des cohortes. L'étude systématique de [Tur+19] se concentre sur des visualisations à destination des patients, notant le manque actuel de solutions. Une taxonomie des données affichées, des encodages visuels et des méthodes d'évaluation compare 39 approches, dont 80 % représentent les données longitudinales par des graphiques linéaires.

L'examen des différentes synthèses révèle l'absence d'une synthèse spécifiquement

1. Voir <https://browser.timeviz.net/>, consulté le 08/01/24, pour le répertoire complet des techniques.

dédiée à la visualisation de la trajectoire d'un patient, caractérisée par des données temporelles multivariées. Dans la suite de ce chapitre, nous proposons une telle synthèse en appliquant le protocole PRISMA-ScR et en proposant une nouvelle taxonomie pour comparer les outils disponibles détaillée dans la section 2.1.2.5.

2.1.2 Méthode

Nous avons réalisé une étude exploratoire de la littérature en nous concentrant sur les visualisations qui représentent un seul individu. Nous avons étendu le protocole de [Sch+22] qui se base sur le protocole PRISMA-ScR [Tri+18]. Nous avons inclus des travaux issus d'autres domaines que la santé pour étudier la représentation des enregistrements multivariés d'un seul individu sur une frise temporelle ou *timeline*.

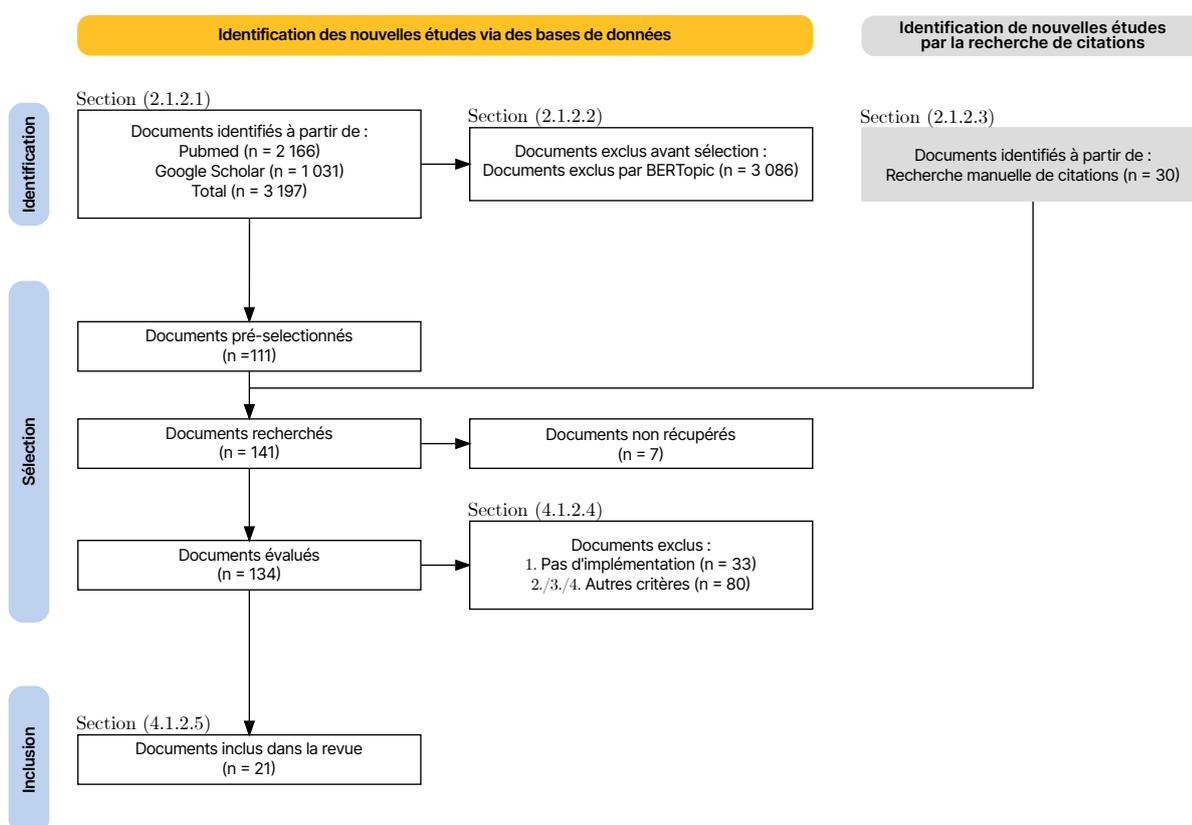


FIGURE 2.2 – Diagramme de flux décrivant le volume de documents obtenu à l'issue de chaque étape du processus PRISMA-ScR. Diagramme adapté des travaux de [Had+22].

Le protocole complet est représenté par un diagramme de flux adapté de [Had+22] et illustré dans la Figure 2.2. Le diagramme précise les étapes successives du protocole décrites dans les sections suivantes.

2.1.2.1 Sources d'information et définition de la requête

Pour obtenir des articles pertinents de tous les domaines d'application, nous avons consulté la base de données spécialisée dans le domaine de la santé, PubMed², ainsi

2. <https://pubmed.ncbi.nlm.nih.gov/>, consulté le 04/05/2024.

que la base de données généraliste Google Scholar³. Cette recherche a été menée le 13 novembre 2023 à l'aide des bibliothèques Python Pymed⁴ et Scholarly⁵.

La requête utilisée dans [Sch+22] englobe les quatre concepts suivants : temps, visualisation, données et médecine. Afin d'identifier les approches concernant les visualisations de données temporelles à travers tous les domaines, nous avons seulement utilisé les mots-clés des trois concepts « temps », « visualisation » et « données ». La requête est visible dans le Tableau 2.1.

TABLEAU 2.1 – Mots-clés de recherche par thème utilisé dans le moteur PubMed. Du fait des limitations du moteur de recherche Google Scholar, le suffixe « [tiab] », qui limite la recherche au contenu du titre et du résumé seulement, y a été enlevé à chaque mot-clé. Aussi, les mots-clés factorisés par le biais de l'expression « * » ont été développés (par exemple « timeframe* » devient « timeframe OR timeframes »).

Temps

(« temporal data »[tiab] OR « temporal sequence* »[tiab] OR « temporal pattern* »[tiab] OR « temporal abstraction* »[tiab] OR « temporal event* »[tiab] OR « time sequence* »[tiab] OR « time series »[tiab] OR « time period* »[tiab] OR « time frame* »[tiab] OR « timeframe* »[tiab] OR timeline*[tiab] OR time-oriented[tiab] OR (time[tiab] AND events[tiab])) AND

Visualisation

(visuali*[tiab] OR « visual analy* »[tiab]) AND

Données

(data[tiab] OR information[tiab])

À l'issue de cette étape, nous avons obtenu 3 197 documents.

2.1.2.2 Pré-sélection semi-automatique des articles

Parmi les résultats de la requête, nous avons obtenu de nombreux articles ne portant pas sur les interfaces de visualisation de données temporelles. L'annotation manuelle de milliers de documents pour la sélection des articles pertinents est un processus coûteux, réalisable uniquement par une large équipe d'évaluateurs. Pour cette raison, nous avons décidé de réaliser une première sélection grâce à des outils semi-automatiques.

Nous avons appliqué un algorithme d'extraction des thèmes appelé BERTopic [Gro22]. BERTopic est issu de BERT (*Bidirectional Encoder Representations from Transformers*) qui est un modèle profond de représentation bidirectionnelle non supervisé du langage, ayant donné de bons résultats dans l'extraction de sujets [Dev+19]. BERTopic se base sur une variation du TF-IDF pour extraire les sujets pertinents de textes et regrouper ces textes selon ces sujets. En appliquant BERTopic sur le titre, le résumé et les mots-clés des documents obtenus à l'étape précédente, les documents ont été classés selon 11 partitions qui correspondent chacune à des thèmes scientifiques. Pour s'assurer de la correspondance du partitionnement automatique avec nos critères de

3. <https://scholar.google.com/>, consulté le 04/05/2024.

4. <https://github.com/gijswobben/pymed>, consulté le 04/05/2024.

5. <https://github.com/scholarly-python-package/scholarly>, consulté le 04/05/2024.

pré-sélection, nous avons inclus 19 articles mentionnés dans les travaux de synthèse de notre domaine cible⁶. Il s'agit des synthèses qui ont été décrites dans la section 2.1.1 et qui traitent toutes de la visualisation de données médicales. Parmi ces 19 articles, 14 se sont retrouvés au sein de la même partition que nous avons retenue pour la suite de l'étude. Cette partition contient au total 111 documents. Cette partition est associée aux mots-clés tels que « data », « visualization », « visual ».

À l'issue de cette étape, nous avons retenu 111 documents.

2.1.2.3 Recherche manuelle des articles par citations

En complément de l'approche semi-automatique et pour que la sélection de documents soit la plus exhaustive possible, nous avons extrait manuellement les citations au sein des travaux de synthèse dédiés à la représentation de données temporelles dans le domaine médical [CKS10; Aig+11; Rin+13; WBH15; Tur+19; Sch+22]. Les documents identifiés au sein de ces revues de la littérature et qui n'auraient pas été retrouvés à l'issue de l'étape précédente sont alors retenus.

Cette recherche supplémentaire a permis d'identifier 30 documents supplémentaires, portant ainsi le nombre total de documents à 141.

2.1.2.4 Analyse par un évaluateur

De manière à garantir que les articles sélectionnés correspondaient bien à notre contexte, les articles ont été analysés par un évaluateur qui est un spécialiste du développement d'interface pour le domaine médical. Parmi les documents identifiés à l'issue des étapes précédentes, 7 n'ont pas pu être récupérés, aboutissant à un total final de 134 documents. Les articles sont retenus s'ils remplissent les quatre critères suivants :

1. Ils présentent une application implémentée, testée sur des données réelles ou fictives. Cela exclut les articles présentant une preuve de concept, un protocole d'évaluation, l'évaluation d'un aspect visuel ou technique ou une revue de l'état de l'art.
2. Ils proposent la vue d'un individu ciblé par l'utilisateur, extrait d'un jeu de données. Cela exclut les articles qui présentent uniquement une vue synthétique de l'ensemble ou d'un sous-ensemble du jeu de données.
3. Ils projettent les données sur un axe temporel unique ou synchronisé à travers différentes vues.
4. Ils présentent des données multivariées ou multimodales, reliées à l'individu cible.

Parmi les 134 documents évalués, 33 ont été exclus parce qu'ils ne correspondaient pas au premier critère et 80 ne correspondaient pas aux trois autres critères. À l'issue de cette étape, nous avons retenu 21 documents.

6. L'identification de ces 19 articles a été réalisée manuellement à partir de l'étude de citations et de revues de la littérature à échelle française et internationale.

2.1.2.5 Extraction des informations dans les articles

Pour l'extraction des informations, nous avons développé un questionnaire affiné au cours de deux itérations successives pour caractériser le contenu des articles. Ces critères sont présentés dans les deux paragraphes suivants : d'abord les fonctionnalités puis les méthodes d'évaluation utilisées. En particulier, les fonctionnalités décrivent les encodages visuels choisis pour représenter les données ou les interactions qui permettent à l'utilisateur de changer cet affichage. Dans tous les cas, un critère est satisfait si l'affirmation énoncée est applicable au contenu de l'article annoté. L'affirmation doit être confirmée explicitement au sein du texte de l'article ou dans les illustrations.

Fonctionnalités Les critères suivants, décrivent la capacité de l'interface à montrer l'intégralité des données présentes dans le parcours, en détail et de manière lisible :

- **Contexte et focalisation** : l'utilisateur peut décider d'afficher en détail une zone ou un élément. Il doit en permanence être en mesure d'estimer le contexte dans lequel la zone ciblée est englobée.
- **Extension et réduction** : l'interface offre le choix entre une vue compacte ou détaillée d'un élément.
- **Granularité minimale** : l'interface montre la représentation la plus fine des données possible.

Les critères suivants décrivent le procédé visuel utilisé pour représenter les enregistrements du parcours patient.

Deux critères décrivent l'aspect hétérogène ou structuré des enregistrements :

- **Multimodales** : l'interface affiche des données numériques ou catégorielles en conjonction de données de natures différentes comme des textes, des images, etc.
- **Hiérarchie** : l'interface affiche des données dont la nature est organisée en hiérarchie, de deux niveaux ou plus. Cette structure doit être visible.

Deux autres critères décrivent l'aspect temporel des enregistrements :

- **Distorsion temporelle** : l'interface affiche l'axe temporel selon une échelle non-linéaire, dans le but de grossir certaines zones.
- **Séries temporelles** : l'interface affiche l'évolution de plusieurs variables continues selon un axe temporel.

Quatre critères décrivent la capacité de l'interface à mettre en valeur certains éléments importants pour la prise de décision ou la compréhension des tendances et de la qualité des données. La mise en valeur visuelle peut être réalisée selon :

- **Importance** : l'interface attire l'attention sur les valeurs importantes. L'importance peut être issue de critères variables.
- **Plages standard** : l'interface affiche différemment les mesures dont les valeurs attendues sortent des plages standard.
- **Variations** : l'interface synthétise les variations de valeurs.

- **Valeurs manquantes** : l'interface met en valeur les types de données ou les points de données qui pourraient manquer dans les données présentées. L'interface montre également l'absence de valeurs numériques dans des séries.

Deux critères décrivent la capacité de l'interface à exploiter un modèle de prédictions sous-jacent pour ajouter des informations supplémentaires :

- **Carte de proximité** : l'interface montre dans un encart les relations ou la proximité de l'individu ciblé avec les autres individus. Ces relations sont issues d'un modèle prédictif ou des données directement.
- **Prédiction** : l'interface intègre visuellement les prédictions obtenues par un modèle sous-jacent dans la présentation des données.

Deux critères complètent les critères déjà énoncés :

- **Environnement ouvert** : l'interface affiche des données issues d'un modèle de données ouvert ou permet à l'utilisateur de charger ses données dans un format interopérable.
- **Informations individuelles générales** : l'interface affiche dans un encart les informations générales sur l'individu cible.

Évaluation Les sept critères suivants décrivent le processus d'évaluation de l'interface :

- **Précision** : les participants doivent résoudre des tâches fermées en interaction avec l'interface et le nombre de succès pour résoudre ces tâches est mesuré.
- **Temps de résolution** : de la même manière, le temps nécessaire à la résolution de chaque tâche est mesuré.
- **Réflexions à haute-voix** : au cours de l'interaction avec l'interface, les participants sont encouragés à expliciter à haute-voix leur processus d'exploration et d'interaction, ainsi que tout ce qui pourrait les aider à résoudre leur problème.
- **Questions graduées** : les participants sont invités à répondre à des questions orientées sur des aspects particuliers de leur utilisation de l'interface en suivant une échelle prédéfinie. Cela peut être deux options « Oui » / « Non », une échelle numérique comme l'échelle de Likert [Lik32] de 1 à 5, une échelle de criticité, etc.
- **Questions ouvertes** : de la même manière, les participants sont invités à répondre à des questions sans contraintes sur leurs réponses.
- **Retour d'expérience ouvert** : à la suite de l'utilisation de l'interface, les participants sont invités à se livrer mais sans contraintes sur leur expérience.
- **Étude de cas** : le scénario d'usage de l'interface par un utilisateur est retranscrit dans le détail. Son utilisation de l'interface doit viser à la résolution d'une tâche identifiée en amont.

2.1.3 Résultats

Les résultats sont présentés dans les Tableaux 2.2 et 2.3, ordonnés par date de publication. Nous distinguons les travaux issus du domaine médical des autres.

Les observations suivantes sont réalisées à partir des données obtenues dans le Tableau 2.2, qui traite des fonctionnalités et des interactions. D'après nos observations, aucune approche ne permet d'identifier facilement la présence de valeurs manquantes. La plupart des approches affichent les données temporelles multivariées sous forme de séries temporelles parallèles (16 parmi 21). Peu d'approches sont compatibles avec des environnements ouverts (3 parmi 21). Peu d'approches ont recours à la distorsion temporelle (4 parmi 21). Toutes les approches ne montrent pas les données brutes avec la granularité la plus fine (13 parmi 21 le montrent). Peu d'approches montrent des données multimodales (5 parmi 21). Aucune approche du domaine médical ne propose une exploration de l'espace latent via des prédictions ou la projection d'individus similaires.

Les observations suivantes sont réalisées à partir des données obtenues dans le Tableau 2.3, qui traite des méthodes d'évaluation des besoins auxquelles doivent répondre les solutions visuelles. Certaines approches font de la conception centrée utilisateurs (9 parmi 21) mais toutes les approches répondent à des besoins utilisateurs identifiés. Aucune différence majeure n'apparaît du point de vue de l'expression ou de l'évaluation des besoins entre les domaines médicaux et non médicaux. Les évaluations qualitatives et quantitatives sont combinées dans les approches les plus récentes uniquement. Aucune des approches identifiées dans cette étude ne déclare collecter les réflexions à haute-voix des participants pour compléter les résultats qualitatifs. Les méthodes de collecte des besoins centrées autour des utilisateurs sont difficiles à classer puisque chacun des travaux identifiés expose un protocole différent.

2.1.4 Discussion

La méthodologie présentée dans cette section, même si elle repose sur les principes de la méthode PRISMA-ScR, présente des limites. Tout d'abord, avec la requête définie dans la section 2.1.2.1, nous avons obtenu des documents hors de la problématique étudiée dans ce chapitre. Une nouvelle combinaison de mots-clés réduirait le nombre de documents hors-sujet. Par ailleurs, le succès de l'utilisation d'une approche semi-automatique, telle que décrite dans la section 2.1.2.2, pour filtrer les approches similaires est dépendante de la qualité de l'algorithme de partitionnement. Nous avons retenu dans nos travaux uniquement l'algorithme BERTopic [Gro22] et nous pourrions le comparer à d'autres approches, en particulier LDA [BNJ03]. L'enrichissement de la représentation des articles par l'utilisation du texte complet et des illustrations dans des modèles multimodaux [Lop09] est une autre piste d'amélioration de la qualité du partitionnement. Pour finir, le choix des sources de données, Pubmed et Google Scholar, a été motivé par la couverture respective de ces sources du domaine médical et général. D'autres sources auraient pu être choisies e. g. Web of Science, DBLP,

L'analyse de la littérature exploratoire révèle des lacunes dans le développement d'interfaces de visualisation dédiées à un seul individu. Notamment, il y a peu d'ap-

TABEAU 2.2 – Synthèse du contenu des 21 articles identifiés à propos des fonctionnalités et des interactions, selon les critères développés dans la section 2.1.2.5.

	Fonctionnalités														
	Contexte + Focalisation	Extension + Réduction	Granularité minimale	Multimodales	Hierarchie	Distorsion temporelle	Séries temporelles	Mise en valeur de l'importance	Plages standard	Mise en valeur des variations	Valeurs manquantes	Carte de proximité	Prédiction	Environnement ouvert	Information individuelles générales
Médical															
LifeLines [Alo+98]			✓		✓										✓
KNAVE-II [Mar+04]			✓		✓		✓								
MIDGARD [BSM04]	✓					✓	✓		✓	✓					✓
Caregiver [Bro+05]	✓					✓	✓	✓		✓					
CareVis [AM06]	✓	✓	✓			✓			✓	✓					
Lung Transplant [PFH07]	✓						✓			✓					✓
TimeLine [BAK07]			✓	✓	✓		✓							✓	✓
MIVA [FN11]			✓	✓			✓		✓						✓
VisuExplore [Rin+11]			✓		✓		✓								✓
CareCruiser [Gsc+11]			✓		✓		✓			✓					
UHS Lifelines [Hal+19]				✓	✓										
Tumor Board [Ste+20]					✓				✓						✓
ClinicalPath [Lin+23]	✓		✓		✓		✓	✓	✓	✓					
Non-Médical															
LastHistory [Bau+10]			✓	✓	✓										
ChronoLenses [Zha+11]	✓		✓			✓	✓								✓
Temporal MDS [Jäc+16]			✓				✓	✓							✓
BitExtract [Yue+19]	✓						✓					✓			
PromotionLens [Zha+22]	✓		✓				✓					✓	✓		✓
Roslingfier [Shi+23]			✓				✓			✓		✓			
Life Mountain [Zha+23]					✓		✓	✓		✓		✓			✓
LiveRetro [Wu+23]				✓	✓		✓					✓	✓		
Total :	8	1	13	5	11	4	16	4	5	8	0	5	2	3	9

TABLEAU 2.3 – Synthèse du contenu des 21 articles identifiés à propos des méthodes d'évaluation, selon les critères développés dans la section 2.1.2.5.

	Évaluation						
	Précision	Temps de résolution	Questions graduées	Questions ouvertes	Retour d'expérience ouvert	Réflexions à haute-voix	Étude de cas
Médical							
LifeLines [Alo+98]	✓	✓	✓				
KNAVE-II [Mar+04]	✓	✓	✓				
MIDGARD [BSM04]							✓
Caregiver [Bro+05]							✓
CareVis [AM06]				✓			
Lung Transplant [PFH07]	✓	✓					
TimeLine [BAK07]							
MIVA [FN11]	✓	✓	✓				
VisuExplore [Rin+11]				✓			✓
CareCruiser [Gsc+11]							✓
UHS Lifelines [Hal+19]							
Tumor Board [Ste+20]			✓				
ClinicalPath [Lin+23]	✓	✓	✓		✓		✓
Non-Médical							
LastHistory [Bau+10]				✓	✓		✓
ChronoLenses [Zha+11]							
Temporal MDS [Jäc+16]							✓
BitExtract [Yue+19]							✓
PromotionLens [Zha+22]			✓		✓		✓
Roslingfier [Shi+23]			✓		✓		✓
Life Mountain [Zha+23]	✓		✓	✓			✓
LiveRetro [Wu+23]							✓
Total :	6	5	8	4	4	0	1

proches exploitant la distorsion de l'axe temporel pour examiner la granularité minimale des données. De plus, le nombre d'approches interopérables est limité. Par ailleurs, aucune approche n'intègre visuellement les prédictions de modèles, une fonctionnalité pourtant présente dans des approches non médicales comme [Zha+22; Wu+23]. La section suivante présente une interface que nous avons conçue pour répondre à ces problématiques ainsi qu'un protocole d'évaluation pour valider la pertinence de notre proposition auprès de professionnels de santé.

2.2 GLIMPSE-Med

Dans cette section, nous présentons une nouvelle interface de visualisation interactive : GLIMPSE-Med, pour *Graphical Longitudinal Interface for Multimodal Pathways Search and Expertise for Medicine*. Cette interface, conçue pour un seul écran, propose une frise temporelle (ou *timeline*) qui affiche les informations relatives aux trajectoires des patients à partir des données de leur DPI. Elle offre une vue globale, des vues détaillées de ses enregistrements, ainsi que des prédictions de risque sous forme de score évolutif déduites de ses mesures.

Cette interface est dédiée à deux types d'utilisateurs :

1. Les professionnels de santé, tels que les médecins ou les infirmiers, qui doivent appréhender rapidement l'état d'un patient, ses mesures et les prédictions.
2. Les ingénieurs de données ou les statisticiens chargés de vérifier l'exhaustivité et la qualité des enregistrements d'un patient.

Cette section est organisée comme suit : les besoins identifiés en collaboration avec les utilisateurs sont présentés dans la section 2.2.1. Ensuite, nous détaillons la conception de notre interface dans la section 2.2.2. Le protocole expérimental pour évaluer GLIMPSE-Med est décrit dans la section 2.2.3 et les résultats obtenus dans la section 2.2.4. Enfin, une discussion vient conclure cette partie dans la section 2.2.5.

2.2.1 Analyse des besoins

Dans la pratique, afin de respecter les recommandations de Munzner [Mun09] et le processus de conception proposé par Sedlmair et al. [SMM12], la conception de systèmes visuels est menée par itérations successives impliquant : (1) la définition des besoins et des structures de données appropriées, (2) la proposition d'encodages visuels et de fonctionnalités interactives répondant à ces besoins, et (3) la présentation des résultats aux utilisateurs pour recueillir leurs retours et affiner les besoins.

Pour développer GLIMPSE-Med, nous avons établi une liste de besoins en nous appuyant sur la littérature existante et grâce à une collaboration avec des professionnels de santé et des ingénieurs de données. Les experts impliqués étaient affiliés au Centre Hospitalier Universitaire de Montpellier. À la suite de la description complète de chaque besoin, nous présentons un inventaire des fonctionnalités qui permettent d'y répondre. Cela permet d'identifier dans un deuxième temps la réponse des approches existantes à ces besoins. Ces fonctionnalités sont issues de la section 2.1.2.5 et mises en valeur

en gras dans le texte. Les besoins identifiés, numérotés de 1 à 6, et les fonctionnalités associées sont présentés dans la liste suivante :

- (B1) **Vue globale et détaillée** : pour faciliter la compréhension et permettre une prise de décision éclairée, l'interface doit offrir une visualisation à la fois synthétique et détaillée de l'état du patient. Cela nécessite de fournir un résumé global du contexte tout en offrant un accès facile aux détails spécifiques sur demande. De plus, afin d'être adaptée à une utilisation dans les services de santé, il est crucial que la densité d'information puisse être affichée sur un écran de taille réaliste, tel qu'un ordinateur de bureau.

La conjonction de ces deux visions côte à côte correspond à la fonctionnalité **contexte + focalisation**. De plus, l'adaptation de l'axe du temps en utilisant la **distorsion temporelle** et l'ajout d'une interaction **extension + réduction** qui contrôle le niveau de détail des éléments affichés complètent la réponse à ce besoin.

- (B2) **Trajectoire** : l'interface doit rendre compte des changements de l'état d'un patient au fil du temps en affichant de manière cohérente toutes les mesures individuelles. En offrant une séquence chronologique de l'évolution des mesures, les professionnels de santé observent les tendances et les variations de l'état de santé du patient.

Cela se traduit par la prise en compte des **séries temporelles** et de la **granularité minimale** des valeurs. Afin d'aider l'utilisateur, l'évolution des valeurs peut être accentuée par la **mise en évidence des variations**.

- (B3) **Hétérogénéité** : l'interface doit être capable de gérer une variété de types de données liées à la santé du patient, englobant notamment des valeurs numériques ou catégorielles, issues de sources variées souvent structurées.

La représentation des données doit montrer la **multimodalité** des données représentées et leur **hiérarchie** structurelle.

- (B4) **Qualité** : les utilisateurs doivent pouvoir évaluer l'exhaustivité, la validité, la cohérence et la conformité des enregistrements affichés.

La qualité des données doit être mise en évidence par l'accentuation visuelle des **plages standard** de valeurs et la représentation des **valeurs manquantes**.

- (B5) **Importance** : l'interface doit être capable de mettre en évidence des enregistrements spécifiques en fonction de critères prédéfinis. En mettant en évidence des points de données essentiels, l'interface permet aux professionnels de santé d'identifier et d'interpréter rapidement des informations critiques même au sein d'un grand volume de données.

Les enregistrements avec une forte **importance** doivent être visuellement accentués, ce qui peut s'accompagner de la mise en relief des valeurs par rapport aux **plages standard**.

- (B6) **Prédiction** : ce besoin implique l'intégration d'un modèle prédictif tiers dans l'interface. Les prévisions concernant la santé du patient doivent être présen-

tées visuellement aux côtés des enregistrements traditionnels et pouvoir être distingués de ces derniers.

L'interface doit pouvoir afficher les sorties d'un modèle de **prédiction**.

Si tous ces besoins visent les deux publics que nous avons identifiés, (B4) répond principalement aux problématiques des ingénieurs de données et des statisticiens.

TABLEAU 2.4 – Comparaison de la réponse aux besoins de 21 approches identifiées dans la littérature qui abordent la représentation de séries temporelles multivariées d'un seul individu. Une coche verte (✓) indique que l'approche satisfait le besoin en question et une coche orange (✓) indique une conformité partielle. Les noms des approches présentées en italique n'ont pas été cités comme tels par leurs auteurs respectifs.

	B1	B2	B3	B4	B5	B6
	Global/Détaillé	Trajectoire	Hétérogénéité	Qualité	Importance	Prédiction
Médical						
LifeLines [Alo+98]		✓	✓			
KNAVE-II [Mar+04]		✓	✓			
MIDGARD [BSM04]	✓	✓		✓	✓	
Caregiver [Bro+05]	✓	✓			✓	
CareVis [AM06]	✓	✓		✓	✓	
<i>Lung Transplant</i> [PFH07]	✓	✓			✓	
TimeLine [BAK07]		✓	✓			
MIVA [FN11]		✓	✓	✓	✓	
VisuExplore [Rin+11]		✓	✓			
CareCruiser [Gsc+11]		✓	✓		✓	
UHS Lifelines [Hal+19]			✓			
<i>Tumor Board</i> [Ste+20]			✓	✓	✓	
ClinicalPath [Lin+23]	✓	✓	✓	✓	✓	
Non-Médical						
LastHistory [Bau+10]		✓	✓			
ChronoLenses [Zha+11]	✓	✓				
Temporal MDS [Jäc+16]		✓			✓	
BitExtract [Yue+19]	✓	✓				
PromotionLens [Zha+22]	✓	✓				✓
Roslingfier [Shi+23]		✓			✓	
<i>Life Mountain</i> [Zha+23]		✓	✓		✓	
LiveRetro [Wu+23]		✓	✓			✓
Total (✓✓) :	8	19	12	5	11	2
Total (✓) :	1	3	4	0	1	2

La comparaison des approches recensées dans la section 2.1 d'après l'intégration de fonctionnalités qui répondent aux besoins est présentée dans le Tableau 2.4. Chaque besoin correspond à plusieurs fonctionnalités et une fonctionnalité peut satisfaire plusieurs besoins. Ce tableau constitue donc une synthèse où, pour une approche donnée, la présence de toutes les fonctionnalités qui correspondent à un besoin donné est indiquée par une coche verte (✓), signifiant une satisfaction complète du besoin correspondant. Si au moins une fonctionnalité est absente, la coche est orange (✓), indiquant une satisfaction partielle du besoin. En revanche, si aucune fonctionnalité

n'est présente, la cellule correspondante est vide, indiquant que le besoin n'est pas satisfait par l'approche en question.

Concrètement on constate par exemple dans ce tableau que pour le premier besoin (**B1**), CareVis [AM06] est la seule interface à satisfaire complètement ce besoin (✓) parce qu'elle intègre les trois critères reliés à celui-ci, c'est-à-dire le contexte + focalisation, la distorsion temporelle et l'extension + réduction. En revanche, PromotionLens [Zha+22], BitExtract [Yue+19] et Lung Transplant [PFH07] intègrent uniquement le contexte + focalisation. Ces approches et quatre autres ne satisfont que partiellement (✓) le besoin (**B1**). Enfin, treize approches dont LiveRetro [Wu+23] ou Tumor Board [Ste+20] n'intègrent aucun des critères et ne satisfont donc pas du tout le besoin (**B1**).

Ainsi, à notre connaissance, aucune approche existante n'intègre visuellement la qualité des données ou les sorties de modèles prédictifs à des séries temporelles multivariées sur un seul écran. Dans le contexte médical, aucune interface existante ne se concentre sur la compression à la fois des axes temporel et des modalités. La caractéristique principale de l'interface que nous proposons est la distorsion dynamique de l'axe temporel, qui offre une vue compacte et détaillée des mesures. À travers cette interaction les professionnels de santé, les ingénieurs de données ou les statisticiens peuvent alors accomplir des tâches variées.

2.2.2 Conception de l'interface

Dans cette section, nous présentons notre interface, GLIMPSE-Med, dont la conception a priorisé la compression des axes temporels et des modalités. D'après la taxonomie de [Aig+11], notre interface utilise un cadre de référence abstrait, des variables multivariées, un agencement du temps linéaire, des primitives temporelles à la fois instantanées et sous forme d'intervalles, une correspondance visuelle statique et une dimensionnalité bidimensionnelle.

L'interface est illustrée dans la Figure 2.3, où elle présente des données d'un patient spécifique provenant de MIMIC-III, une base de données présentée dans la section 1.2.2 (ici, ces données anonymisées sont celles d'une femme de 65 ans). L'interface se compose de trois parties principales : un panneau affichant des informations générales sur le patient (Fig. 2.3.a), un menu déroulant pour naviguer entre les différentes modalités (Fig. 2.3.b) et un axe temporel (Fig. 2.3.c). En termes d'implémentation, l'interface exploite React⁷, D3⁸ et Tailwindcss⁹. Les paragraphes suivants fournissent une exploration approfondie des fonctionnalités disponibles et des encodages visuels.

Informations générales Un panneau, illustré dans la Figure 2.3.a, affiche les informations statiques sur le patient, ce qui englobe les données dépourvues de dynamique temporelle et nécessitant un accès rapide. Ce panneau sert d'espace dédié à l'examen individuel du patient (**B1**), facilitant un accès rapide aux détails démographiques (**B2**) pour une surveillance efficace.

7. <https://react.dev/> consulté le 21/03/24.

8. <https://d3js.org/> consulté le 21/03/24.

9. <https://tailwindcss.com/> consulté le 21/03/24.

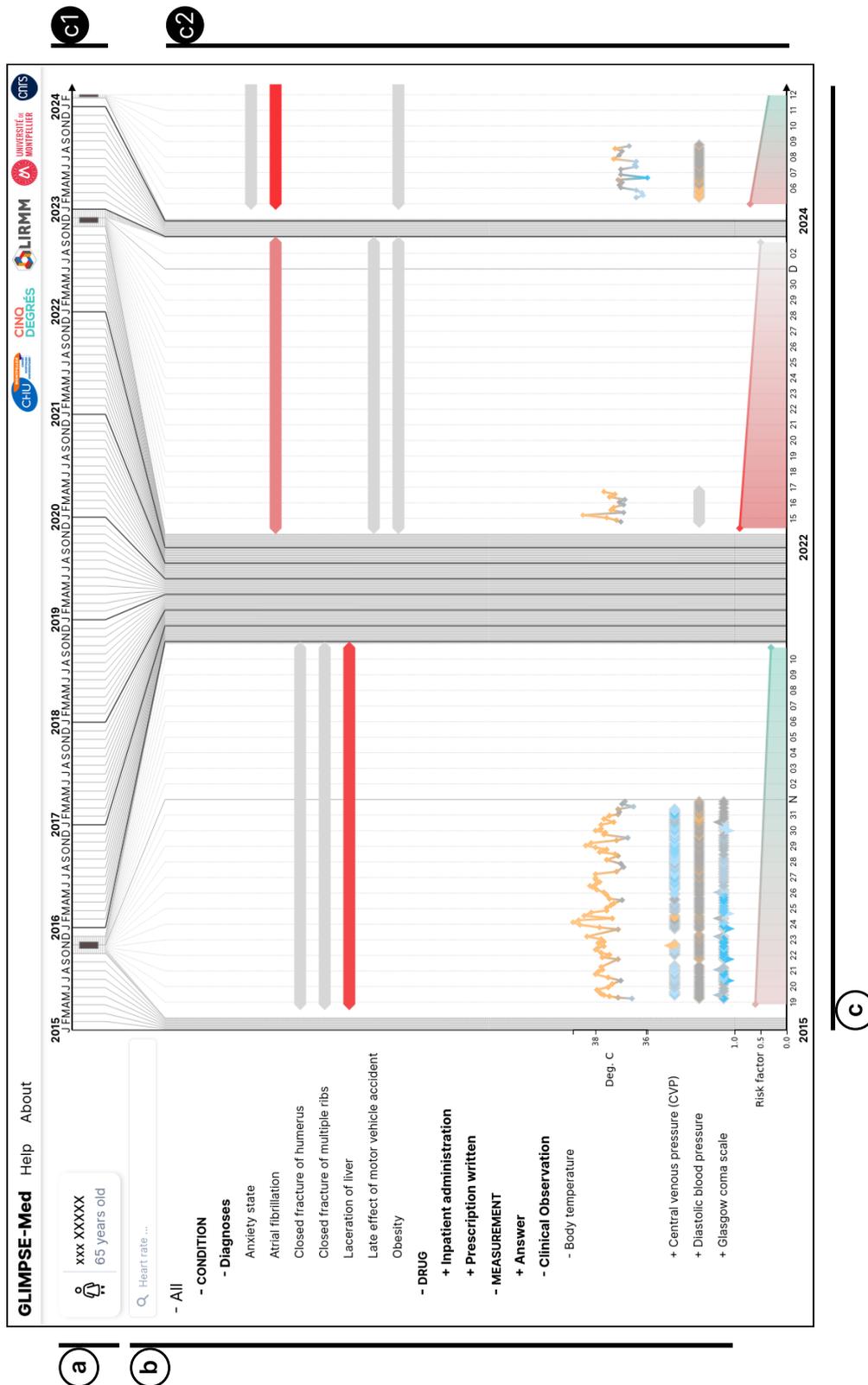


FIGURE 2.3 – Capture d'écran de GLIMPSE-Med, présentant une patiente dont le parcours contient 7 885 points de données couvrant trois séjours hospitaliers. (a) Panneau d'informations générales sur le patient (anonymisé). (b) Menu déroulant affichant les modalités mesurées regroupées par catégories et sous-catégories. Une barre de recherche permet aux utilisateurs de filtrer les modalités affichées. (c) Affichage principal des données sous forme de frise. (c1) Axe temporel linéaire. (c2) Axe temporel déformé.

Modalités Les modalités (e. g. les diagnostics, les prescriptions médicamenteuses, les mesures de la pancarte ...) sont présentées via un menu arborescent dépliant (Figure 2.3.b), où chaque catégorie peut être développée ou réduite d'un simple clic. Les feuilles de la structure hiérarchique correspondent aux modalités des enregistrements stockés dans les dossiers patients présentés. Seules les modalités mesurées chez un patient et leurs ancêtres dans la hiérarchie des modalités sont affichées dans le menu. En cliquant sur un nœud feuille, une vue détaillée de la modalité correspondante est activée. L'organisation hiérarchique des modalités en catégories, obtenue à partir de taxonomies médicales telles que la Classification Internationale des Maladies (CIM) pour les diagnostics, n'est pas un prérequis obligatoire pour le bon fonctionnement de l'interface. En effet, en l'absence de hiérarchie, il suffit d'assigner tous les éléments à un nœud parent commun, à savoir la racine. L'utilisateur est capable d'explorer la structure globale et la nature des mesures affichées (**B3**). L'utilisateur est également capable d'identifier d'éventuelles erreurs ou mauvaises classifications de certains types de données (**B4**).

Barre de recherche Une barre de recherche (Figure 2.3.b) est disponible au-dessus du menu des modalités. Les utilisateurs peuvent filtrer les titres des modalités ou des catégories ne contenant pas le terme saisi dans la barre, insensiblement à la casse. L'utilisateur est capable de trouver rapidement un élément désiré parmi une structure dense (**B1**) et d'identifier les types de données qui n'auraient pas été stockés et affichés (**B4**).

Axe temporel Les données temporelles sont représentées le long d'une frise, qui s'étend des mesures initiales jusqu'aux plus récentes, avec des marges pour englober la plus large plage de granularité temporelle. Les subdivisions temporelles sont représentées par des traits verticaux qui ont des nuances de gris et des largeurs variables, pour distinguer les intervalles intermédiaires telles que les années et les mois, comme illustré dans la Figure 2.3.

La frise est divisée en trois parties : un axe linéaire (Figure 2.3.c1), un axe déformé (Figure 2.3.c2) et une région de transition reliant les deux. La déformation est conçue pour condenser l'espace horizontal vide où aucun enregistrement n'a été capturé. Cela a pour effet de grossir l'axe temporel autour des séjours du patient, ce qui indique à quels instants les données ont été collectées. Ensuite, l'utilisateur est capable de naviguer à travers toutes les mesures disponibles dans un contexte temporel complet (**B2**) tout en maintenant une conscience des relations temporelles à une échelle globale (**B1**). Les portions de l'axe qui doivent être élargies sont identifiées de manière récursive par un parcours en profondeur, en utilisant un algorithme personnalisé où l'axe temporel est subdivisé dynamiquement en arbre. Les nœuds de cet arbre sont ensuite basés sur des intervalles calendaires ajustés à l'étendue temporelle requise. L'utilisation optimale de l'espace à l'écran est assurée en adaptant les données affichées dans une portion pré-définie de l'axe, contrôlée par un paramètre.

Séjours hospitaliers et zoom Chaque séjour au sein du parcours du patient est affiché sur l'axe linéaire avec un symbole rectangulaire (Figure 2.3.c1), avec les côtés

gauche et droite du symbole correspondant respectivement à la date d'admission et à la date de sortie. Comme illustré dans la Figure 2.4, un clic sur un symbole élargit l'affichage sur l'admission correspondante, en maintenant la portion pré-définie de l'affichage. L'utilisateur contrôle le niveau de détail désiré en grossissant l'affichage sur une admission (B1).

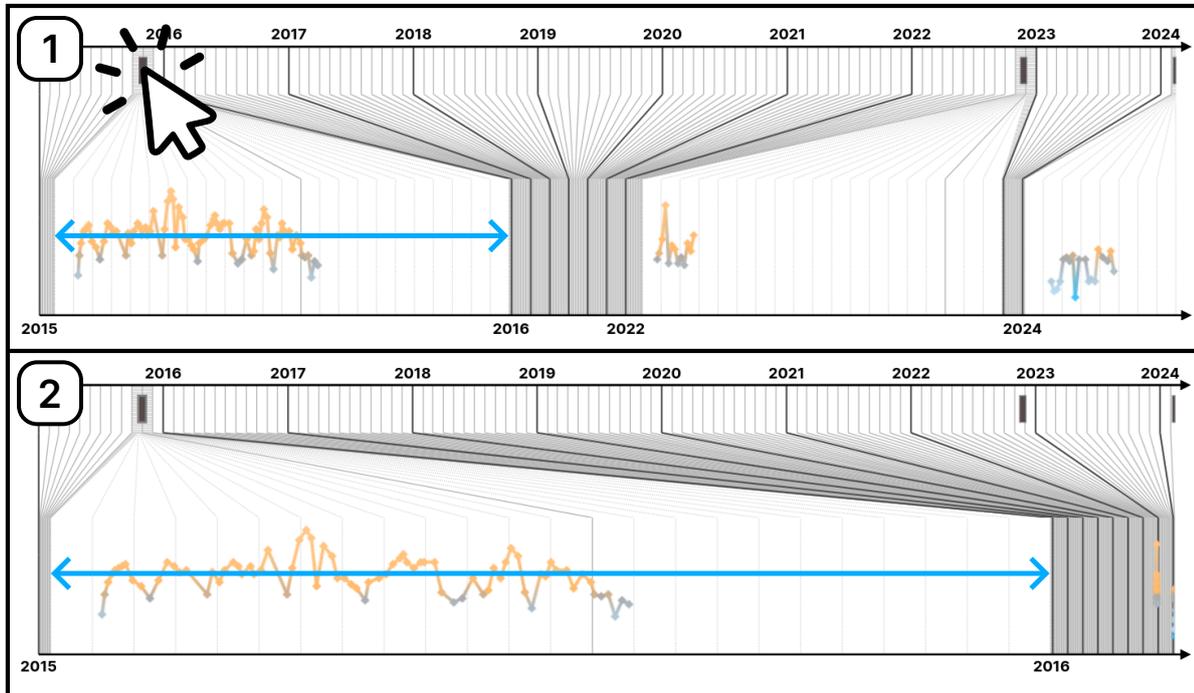


FIGURE 2.4 – Illustration de l'interaction de zoom. (1) La figure supérieure montre la vue par défaut, où tous les séjours hospitaliers sont représentés sur la même échelle au sein de l'axe déformé. (2) La figure inférieure montre l'affichage après le clic sur le premier symbole de séjour. L'axe déformé est alors concentré sur la plage temporelle de cette admission.

Enregistrements Sur la frise, chaque enregistrement est représenté par un symbole en forme de losange. Pour les enregistrements instantanés, indiquant des durées nulles, le coin supérieur du losange est aligné sur la coordonnée temporelle correspondante (Fig. 2.5.a1). Pour les périodes temporelles, le losange s'étend horizontalement, avec les pics gauche et droit du losange allongé indiquant respectivement les coordonnées temporelles de début et de fin (Fig. 2.5.a3). Pour chaque modalité enregistrée, un arrangement séquentiel des enregistrements correspondants est affiché sur l'interface sous forme de série. Les utilisateurs choisissent d'afficher une série dans un format compact ou détaillé. Le clic sur la série correspondante agit comme un commutateur entre les deux vues. Les caractéristiques et fonctionnalités de ces deux vues sont exposées dans les sections suivantes.

Vue compacte d'une série Lorsque les données sont affichées selon la vue compacte, elles sont réparties le long d'une ligne horizontale. Si l'enregistrement englobe une valeur numérique, un petit losange concentrique est inséré à l'intérieur (Fig. 2.6.a2). De plus, pour indiquer une évolution significative de la valeur numérique de la série

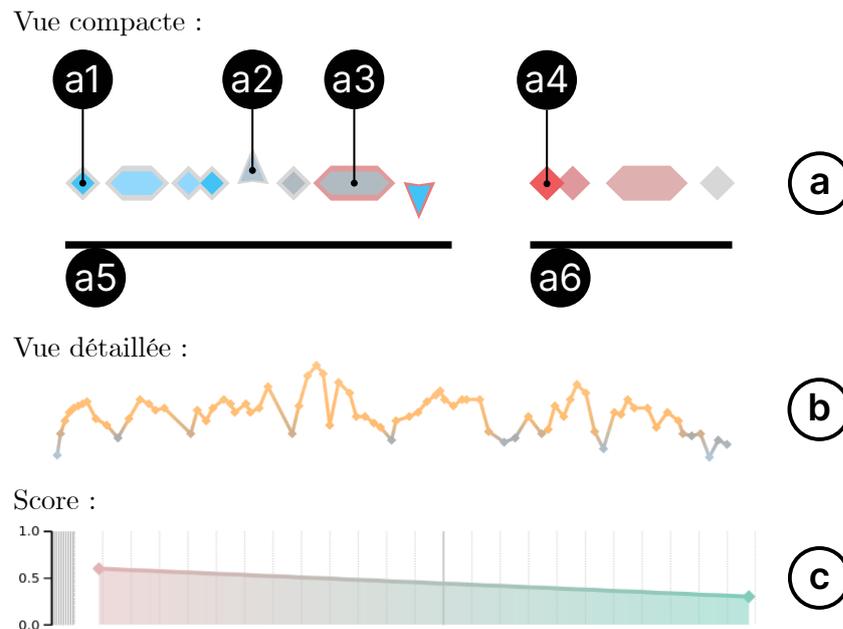


FIGURE 2.5 – Représentation symbolique des enregistrements. (a) Vue compacte. (a1) Exemple d'enregistrement instantané. (a2) Exemple d'augmentation significative. (a3) Exemple d'intervalle. (a4) Exemple de donnée importante. (a5) Enregistrements numériques. (a6) Enregistrements non numériques. (b) Vue détaillée d'une série numérique. (c) Score de risque.

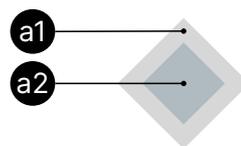


FIGURE 2.6 – Composants extérieur (a1) et intérieur (a2) du symbole représentant un enregistrement. Le composant intérieur est affiché uniquement si l'enregistrement présente une valeur numérique.

par rapport à la valeur précédente, la forme du losange est transformée en un triangle, avec l'orientation du triangle (vers le haut ou vers le bas) indiquant une augmentation ou une diminution, respectivement (Fig. 2.5.a2). Dans notre contexte, une évolution significative correspond à un changement supérieur à l'écart-type de l'ensemble de données pour la modalité considérée. Les utilisateurs discernent donc rapidement la nature (**B3**) d'une série affichée et peuvent ainsi identifier d'éventuelles anomalies telles que des valeurs manquantes, des doublons ou des valeurs aberrantes (**B4**). De plus, les utilisateurs capturent la dynamique globale de la série (**B2**) en un coup d'œil.

Vue détaillée d'une série Lors de la visualisation détaillée d'une série contenant des enregistrements numériques, les données sont affichées sous forme de graphique linéaire. La coordonnée spatiale Y de chaque symbole correspond à la valeur numérique de l'enregistrement associé, si celle-ci est disponible. Les points successifs sont reliés par des segments de droite (Fig. 2.5.b). Cependant, seuls les points appartenant à la même admission sont connectés, afin d'éviter l'interpolation entre des périodes de collecte de données distinctes. Cette méthode permet aux utilisateurs d'analyser en détail la dynamique des mesures (**B2**) et de repérer des anomalies (**B4**) et des

variations (**B5**), tout en conservant une vue d'ensemble contextuelle (**B1**).

Mise en valeur d'enregistrements Le symbole pour un enregistrement est composé de deux parties (Fig. 2.6) : le contour et la partie intérieure. La partie intérieure n'est visible que pour une valeur numérique (Fig. 2.6.a2). Chaque composant est associé à une palette de couleurs indépendante : (1) Les losanges intérieurs sont colorés en fonction des valeurs standard. Étant donné les sémantiques variées des modalités capturées dans les bases de données médicales, établir des hypothèses sur la signification de chaque valeur est irréaliste. Par exemple, tandis que les valeurs maximales de certaines mesures telles que le Score de Glasgow indiquent un état stable, pour d'autres comme le rythme cardiaque, c'est la proximité de la valeur avec la moyenne qui indiquent un état sain du patient. Par conséquent, une palette de couleurs a été sélectionnée avec un minimum de biais sémantique envers un impact positif ou négatif pour la santé du patient : bleu pour les valeurs en dessous de la norme, orange pour celles au-dessus et gris pour les valeurs à l'intérieur de la plage. Inspiré des couleurs traditionnellement associées à la température, ce choix tire parti de la familiarité des utilisateurs avec les interprétations contextuelles du chaud et du froid. (2) Les composants extérieurs (Fig. 2.6.a1) sont saturés vers le rouge pour indiquer l'importance (Fig. 2.5.a4). Cette couleur a été choisie pour son immédiateté dans la transmission des contributions aux facteurs négatifs (**B6**). Bien que le critère d'importance soit agnostique du modèle de prédiction et puisse être adapté pour répondre aux exigences contextuelles, en général, les enregistrements identifiés par le modèle comme contribuant à une détérioration de l'état du patient devraient être mis en évidence par cette palette.

Score prédit En conjonction de l'affichage des enregistrements, une estimation du risque du patient est affichée en bas de l'interface (Fig. 2.5.c) (**B6**). Ces prédictions sont indépendantes des modèles mais devraient être générées soit par des estimateurs statistiques, soit par des modèles d'apprentissage automatique capables de prédictions temporelles. L'étiquette de prédiction est adaptable pour correspondre au contexte applicatif, ce qui nécessite alors l'ajustement de la palette de couleurs et de sa sémantique. Ici, un dégradé de couleur binaire utilisant des teintes de rouge et de vert a été adopté, où le score représenté signifie une estimation généralisée de la détérioration du patient (par exemple, risque de mortalité, insuffisance cardiaque ou réintubation).

Infobulle Lorsque le curseur de la souris survole un élément, une infobulle (*tooltip*) est révélée pour fournir des informations complètes à propos de celui-ci. Cette fonctionnalité permet aux utilisateurs d'évaluer le contenu d'un enregistrement en particulier (**B1**).

2.2.3 Étude expérimentale

Pour évaluer l'efficacité de notre interface, nous avons mené une évaluation utilisateur impliquant 14 participants sur site. Ce panel diversifié comprenait des professionnels de santé du Centre Hospitalier Universitaire de Montpellier, notamment des médecins, des statisticiens et des ingénieurs de données, composé de neuf femmes et cinq hommes. Les participants n'ont reçu aucun avantage financier ou privilège. Cette



FIGURE 2.7 – Illustration d'une infobulle apparaissant lors du survol d'enregistrement. De plus, lors du survol, une série est mise en surbrillance avec un fond gris clair.

expérience a été approuvée par le comité d'éthique de l'Université de Montpellier, avis du comité consultatif numéro UM 2023-041bis.

Les expériences ont été menées avec des trajectoires de patients issues du monde réel provenant du jeu de données MIMIC-III [JPM15; Joh+16] dont les caractéristiques sont détaillées dans la section 1.2.2. Les informations comprennent les données démographiques, les mesures des signes vitaux, les analyses de laboratoire, les procédures, les diagnostics, les médicaments et diverses notes.

Une quatrième version de la base de données [Joh+23] est disponible depuis 2023 et régulièrement mise à jour. La mise à jour vise à étendre le type de données disponibles et propose une plage temporelle plus large, allant de 2008 à 2019. Puisque la littérature et nos expérimentations étaient déjà largement basées sur la version 3, qui offre une diversité et des volumes de données suffisants, nous n'avons pas utilisé cette nouvelle version.

Chaque session d'évaluation, d'une durée d'environ 40 minutes, a débuté par une présentation complète de l'étude, fournissant un aperçu des fonctionnalités de l'interface. Les participants ont eu accès à deux dossiers de patients distincts, accessibles via des onglets de navigateur différents. L'expérience comprenait trois étapes :

Étape 1 : Les participants ont répondu à des questions fermées à travers un questionnaire à choix multiples de 15 questions, en interaction avec l'interface. Les questions attribuées aux participants pour la complétion de tâches sont accessibles dans le Tableau 2.5. Ces questions ont été conçues pour inciter les participants à explorer toutes les fonctionnalités. Les questions nécessitaient parfois des connaissances médicales avancées (questions 14 et 15). Bien que les participants aient répondu sans aide, ils avaient la possibilité de demander des clarifications à l'évaluateur. Aussi, les participants étaient encouragés à verbaliser leurs pensées, partageant toute réflexion ou défi rencontré lors de la navigation dans l'interface.

Étape 2 : Les participants ont rempli le questionnaire SUS (*System Usability Scale*) [Bro96]. Étant donné que tous les participants étaient francophones, nous avons utilisé la version française du questionnaire, F-SUS [GB21]. Composé de 10 questions, visibles dans le Tableau 2.6, ce questionnaire standard vise à évaluer l'efficacité, l'efficience et la satisfaction de l'interface. Les participants ont répondu à chaque question sur une échelle de Likert en 5 points. Le questionnaire génère un score allant de 1 à 100.

Étape 3 : Les participants ont pris part à des discussions ouvertes, répondant oralement aux questions posées par l'évaluateur. Ces questions étaient conçues pour

recueillir des commentaires sur les aspects généraux ou spécifiques de l'interface. Elles étaient l'occasion pour les participants d'exprimer leurs pensées ouvertement.

2.2.4 Résultats

Dans cette section, nous analysons à la fois quantitativement et qualitativement les résultats des expériences menées.

2.2.4.1 Analyse quantitative

TABEAU 2.5 – Questionnaire à choix multiples utilisé lors de l'évaluation par résolution de tâches. La précision moyenne est calculée à partir des réponses des 14 participants.

N ^o	Questions	Préc.
1.	Quel est l'âge du patient ?	100 %
2.	Combien d'années sont montrées à l'écran ?	100 %
3.	Combien de séjours sont montrés sur l'interface ?	100 %
4.	Quelle est la modalité de décharge pour la deuxième admission ?	93 %
5.	Parmi ces choix, quel diagnostic a été identifié dans deux séjours différents ?	100 %
6.	Quelle est la dernière valeur mesurée du « Heart Rate » ?	100 %
7.	Parmi ces choix, quelle « Clinical Observation » ne varie pas lors de la dernière admission ?	93 %
8.	Quel séjour présente le pic de risque ?	93 %
9.	Quelle est la modalité la plus importante pour l'estimation du risque ?	79 %
10.	Quelle est la date de la valeur « Heart Rate » montrée comme importante ?	93 %
11.	Combien de valeurs ont une erreur de codage dans « Fluid output miscellaneous route » ?	100 %
12.	Quelle modalité a des erreurs de codage uniquement lors de la première admission ?	93 %
13.	Quel est le niveau de risque actuel du patient ?	100 %
14.	Est-ce que le diagnostic « Obesity » pourrait être codé sur le séjour du patient ?	79 %
15.	Est-ce que le diagnostic « Essential hypertension » pourrait être codé sur le séjour du patient ?	14 %
Précision moyenne globale :		89 %

Comme le montre le Tableau 2.5, les participants ont obtenu une précision globale de 89 % dans la résolution des questions fermées à l'aide de l'interface. La précision est calculée comme la proportion du nombre total de réponses réussies à une question donnée par rapport au nombre total de participants. Comme anticipé, les résultats étaient élevés, en accord avec l'accessibilité intentionnelle des questions, qui servaient comme un moyen pour les participants d'explorer l'interface avec des objectifs clairs en tête. Comme prévu, les participants ont montré des performances plus faibles sur les deux dernières questions (respectivement 79 % et 14 % pour les questions 14 et 15), qui nécessitaient des connaissances médicales avancées et une habitude de codage médical. Seul un participant avait une habitude du codage médical. Il a réussi à répondre aux questions en rapport avec cette problématique. Par ailleurs, il est important de noter que certains participants ont négligé des enregistrements d'importance élevée, entraînant des performances inférieures à la moyenne pour la question 9. De plus, les résultats des questionnaires fermés indiquent que les participants ont réussi à accomplir diverses tâches et étaient satisfaits de leur utilisation de l'interface.

En se basant sur le score SUS, les participants ont exprimé leur satisfaction quant à

TABLEAU 2.6 – Questionnaire F-SUS. Scores calculés à partir des directives originales de [Bro96] et moyennés sur 14 participants.

N ^o	Questions	Score
1.	Je voudrais utiliser cette interface fréquemment.	7,67
2.	Cette interface est inutilement complexe.	7,50
3.	Cette interface est facile à utiliser.	7,85
4.	J'aurais besoin du soutien d'un technicien pour être capable d'utiliser cette interface.	8,75
5.	Les différentes fonctionnalités de cette interface sont bien intégrées.	7,67
6.	Il y a trop d'incohérences dans cette interface.	8,22
7.	La plupart des gens apprendront à utiliser cette interface très rapidement.	8,02
8.	Cette interface est très lourde à utiliser.	8,02
9.	Je me suis senti-e très en confiance en utilisant cette interface.	7,67
10.	J'ai eu besoin d'apprendre beaucoup de choses avant de pouvoir utiliser cette interface.	8,22
Score total :		79,62 /100

leur expérience d'utilisation de l'interface. Comme le montre le Tableau 2.6, le score moyen est de 79,62, le positionnant dans le quartile le plus élevé en termes de scores absolus. En comparaison avec d'autres études [BKM08], notre interface se situe également dans le quartile le plus élevé, recevant des évaluations entre « bon » et « excellent » sur l'échelle des adjectifs.

2.2.4.2 Analyse qualitative

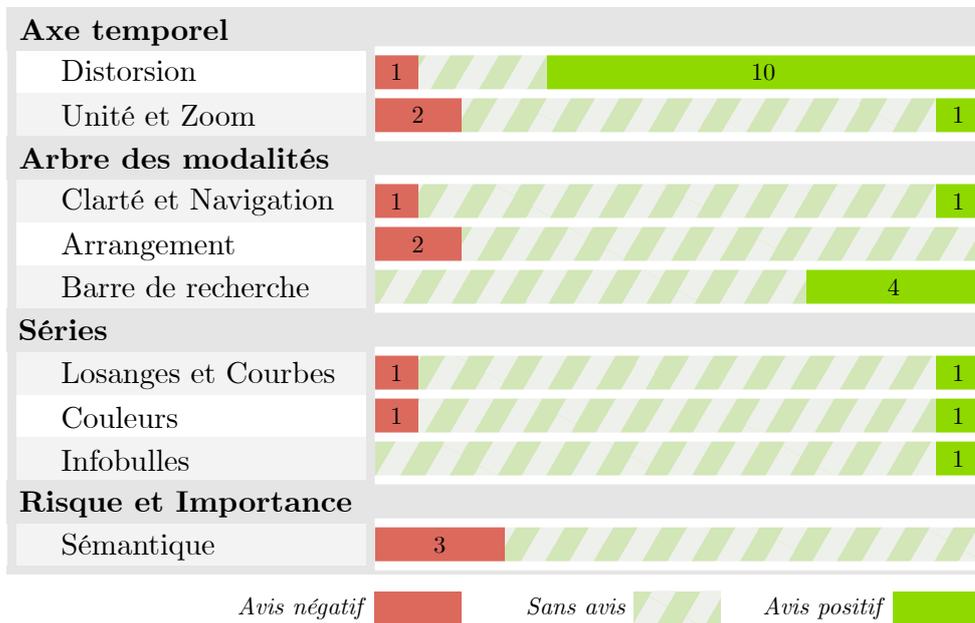
En plus de l'évaluation quantitative, les participants ont été encouragés à partager des commentaires complets en exprimant leurs pensées à haute voix lorsqu'ils ont abordé les tâches et répondu aux questions orales ouvertes. Cette section vise à offrir une synthèse de ces retours.

Fonctionnalités visuelles De nombreux commentaires sur les éléments de l'interface ont été extraits, en particulier à propos de la principale contribution de notre approche : la compression des deux axes de représentation pour les mesures temporelles et multi-variées - l'axe temporel et le menu de modalités.

Une extraction détaillée de ces commentaires est disponible en [Annexe](#) tandis qu'une synthèse est disponible dans le Tableau 2.7. Dans cette synthèse, nous avons choisi d'extraire les avis positifs et négatifs et de les représenter visuellement en vert et en rouge respectivement. Les participants étaient encouragés à exprimer à haute voix les obstacles qu'ils auraient pu rencontrer et leur avis sur les fonctionnalités qui leur déplaisaient. Les résultats montrent qu'ils ne semblent pas avoir hésité à le faire. Nous avons donc représenté l'absence d'avis comme plutôt positif, avec une hachure verte et grise, sous l'hypothèse que les participants aient validé les fonctionnalités à propos desquelles ils n'ont pas émis d'avis négatif. La synthèse présentée résume uniquement les avis à propos des fonctionnalités principales.

On observe que la majorité des participants ont apprécié la distorsion de l'axe temporel (Fig. 2.3.c2) puisque 10 ont exprimé des sentiments positifs. Tous les participants semblent confirmer la lisibilité de cette fonctionnalité. La réserve émise exprime le

TABLEAU 2.7 – Récapitulatif des remarques données par les utilisateurs lors des échanges oraux groupés par fonctionnalités sous forme de diagramme à bandes horizontales. Les avis négatifs sont montrés à gauche et en rouge et les avis positifs à droite et en vert. Les avis non sont montrés en vert hachuré au centre. Au total, un commentaire par participant soit 14 commentaires sont attendus par fonctionnalités.



fait que dans certains cas, il y aurait un avantage à conserver une échelle temporelle linéaire, puisque les espaces entre les mesures auraient une sémantique importante.

Les participants ont apprécié la présence et le fonctionnement de la barre de recherche. Ils ont utilisé avec succès les infobulles (Fig. 2.7). Les autres fonctionnalités sont plus mitigées comme l’affichage des unités, l’interaction du zoom (Fig. 2.4), la navigation dans l’arbre des modalités (Fig. 2.3.b) et la représentation visuelle des séries (Fig. 2.5) même si ces fonctionnalités semblent en majorité avoir été lisibles et compréhensibles.

L’élément le plus contesté est la sémantique de la courbe de risque (Fig. 2.5.c) et des valeurs « importantes » (Fig. 2.5.a4). Par exemple, trois participants n’ont pas pu comprendre le lien entre les deux, affirmant ainsi qu’ils étaient incapables de voir quels éléments sont impliqués dans le calcul du risque et qu’ils n’étaient pas en mesure d’identifier si une valeur rouge donnée concerne le séjour actuel ou le suivant.

Usage L’interface a été conçue pour deux objectifs principaux : (1) la consultation des dossiers médicaux électroniques des patients en milieu clinique et (2) la capacité à évaluer la qualité des données au sein de ces dossiers. Ensuite, nous avons interrogé les participants sur l’utilisation de l’interface en fonction de leur expérience. Nous avons cherché à confirmer la réponse de l’interface aux exigences propres à ces tâches, et également à rechercher de nouvelles applications de l’interface, que nous avons pas anticipées.

Comme prévu, deux médecins ont exprimé leur intention d’utiliser l’interface dans un environnement clinique. Par exemple, à leur arrivée dans le service, ils souhaitent

observer les signes vitaux de la journée précédente pour prendre contact avec le patient. De plus, durant le service, ils pourraient surveiller l'évolution d'un paramètre spécifique. En outre, ils ont souligné son utilité dans un scénario de consultation, affirmant qu'elle aiderait à trouver des résultats d'analyses, des rapports ou à déterminer si un test particulier a été effectué.

Comme prévu également, quatre ingénieurs de données ou statisticiens ont exprimé leur intention d'utiliser l'interface pour évaluer la qualité des données, pour des applications telles que la préparation d'une étude lorsqu'ils travaillent avec des cliniciens ou la vérification de leur travail. Par exemple, ils ont mentionné qu'ils pourraient mener des recherches dans des entrepôts de santé, observer des cas de patients typiques pour évaluer la disponibilité des données et identifier d'éventuelles incohérences. À des fins de contrôle qualité, ils pourraient l'utiliser pour vérifier le chargement correct des dossiers et en inspectant les données des patients.

Nous avons également identifié une troisième utilisation inattendue de l'interface lors de l'évaluation : six participants utiliseraient l'interface pour des études rétrospectives, où les chercheurs doivent explorer *a posteriori* les trajectoires individuelles des patients. Selon les participants, l'interface serait utile pour développer des hypothèses de recherche, en observant par exemple des corrélations, pour planifier l'ensemble de variables à analyser et comment, pour interpréter les données de laboratoire ou pour déterminer si une variante génétique est causale. Un participant a argumenté que cela serait plus rapide que d'aller dans le logiciel hospitalier déjà installé pour ces tâches. Un participant a recommandé d'utiliser l'interface pour le codage médical. Un autre participant a exprimé l'intention de l'utiliser pour faire de l'apprentissage automatique, mettant en avant la capacité de l'interface à montrer et faire comprendre un score prédictif.

2.2.5 Discussion

Le questionnaire ouvert et les retours à haute voix durant l'expérience permettent d'identifier les composants nécessaires aux utilisateurs pour réaliser les tâches. Nous avons aussi pu obtenir des enseignements précieux quant aux utilisations potentielles de l'interface. Ces retours de participants complètent donc l'analyse quantitative et sont indispensables pour identifier précisément les axes d'amélioration de l'interface évaluée. Suite à cette évaluation, l'esthétique générale a été retravaillée, et certaines fonctionnalités ont été modifiées pour prendre en compte les retours des participants.

L'évaluation a montré que l'interface est utilisable comme une alternative plus rapide et plus simple que le logiciel actuel disponible à l'hôpital pour des études *a posteriori*, ou pour évaluer la qualité des données d'un dossier. La présentation sur un seul écran permet à l'utilisateur d'explorer les données avec moins de clics et nécessite moins de connaissances préalables sur le logiciel. Il faudrait réaliser une étude comparative basée sur des tâches pour vérifier la différence dans le nombre de clics.

À notre connaissance, et sans se limiter au domaine médical, GLIMPSE-Med est la seule interface qui intègre visuellement la qualité des données ou les sorties de

prédiction avec des séries temporelles multivariées sur un seul écran. Dans le domaine médical, cette approche est la première à mettre en avant la compression à la fois des axes temporel et de modalité de manière contextuelle/focale.

Limitations Il semble que certains participants aient trouvé difficile d'évaluer l'interface indépendamment du choix des données affichées ou du risque prédit, utilisés à des fins de démonstration. La différence entre le cadre expérimental et la configuration quotidienne des utilisateurs a potentiellement influencé leur capacité à envisager une utilisation régulière de l'interface. De plus, les participants étaient francophones, mais les libellés et les modalités de l'interface étaient rédigés en anglais. Aucun participant n'a déclaré de difficultés linguistiques, mais cet aspect pourrait également avoir un impact sur les résultats.

En raison de la composition du panel et des tâches à résoudre lors de l'évaluation, nous ne pouvons pas valider son efficacité dans un contexte clinique. Cependant, les retours des participants sont positifs et ouvrent des perspectives de développement de l'interface. Une nouvelle phase d'évaluation devra donc être conduite intégrant de nouvelles fonctionnalités comme l'exportation de données ou l'intégration du contenu textuel. Les transmissions ou les rapports, devront être inclus en tant qu'enregistrements au sein de la frise et une nouvelle interaction devra être ajoutée afin de visualiser les données textuelles brutes dans une fenêtre dédiée. De plus, l'efficacité de l'interface dans un contexte clinique devra être étudiée, avec un panel composé de personnels de santé (médecins, infirmiers) travaillant dans les services au quotidien.

Le principal défaut identifié lors de l'évaluation était l'exploration du menu de modalités (Fig. 2.3.b) : les participants ne pouvaient identifier les sous-éléments d'une catégorie qu'en cliquant dessus. Ainsi, nous prévoyons de mettre en œuvre une vue compacte de plusieurs séries. Les sous-éléments d'un nœud d'arborescence replié seraient synthétisés sous forme de diagramme de densité en violon non lissé : les enregistrements à l'intérieur d'un sous-arbre replié sont agrégés par unité du découpage temporel. Chaque agrégation est représentée par un rectangle, dont la hauteur verticale indique le volume des mesures. Dans le cas illustré, l'importance moyenne est mise en évidence à l'aide d'un dégradé de couleurs rouge. De plus, les détails du contenu agrégé sont visibles au survol.

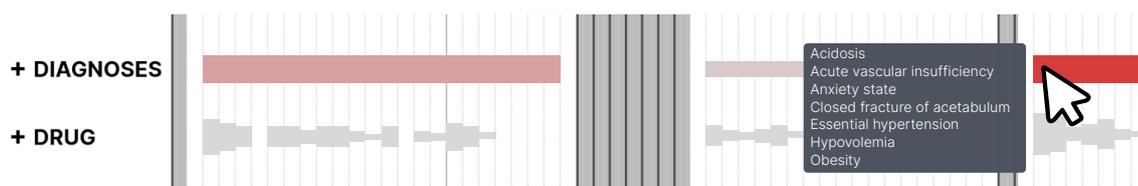


FIGURE 2.8 – Preuve de concept de la vue compacte de plusieurs séries qui prend la forme d'un diagramme de densité en violon non lissé : les enregistrements à l'intérieur d'un sous-arbre replié sont agrégés par unité du découpage temporel. Chaque agrégation est représentée par un rectangle, dont la hauteur verticale indique le volume des mesures. Dans le cas illustré, l'importance moyenne est mise en évidence à l'aide d'un dégradé de couleurs rouge. De plus, les détails du contenu agrégé sont visibles au survol.

Dans les prochaines versions, l'utilisateur pourra épingler les séries désirées dans la partie supérieure de l'axe des ordonnées. L'affichage constant d'un ensemble de séries, aux choix de l'utilisateur, permettra un ordonnancement personnalisé et une comparaison avec le reste des séries. Nous prévoyons également d'incorporer l'exportation des données vers des fichiers externes.

Enfin, si la lisibilité du score de risque (Fig. 2.5.c) et des valeurs importantes (Fig. 2.5.a4) indépendamment l'une de l'autre a été validée, la représentation visuelle de leur relation devra être étudiée plus en détail. Par exemple, lorsque l'utilisateur survole une prédiction, les valeurs importantes pour cette prédiction spécifique survolée devront être mises en évidence dynamiquement à travers une nouvelle palette de couleurs, une infobulle étendue ou des liens spatiaux.

2.3 Conclusions et Perspectives

Dans ce chapitre, nous avons dans un premier temps étudié 21 visualisations de séries temporelles multivariées d'un seul individu selon le protocole PRISMA-ScR. Cette analyse de la littérature exploratoire révèle que cette problématique gagne de l'intérêt récemment mais qu'il y a des limitations dans les solutions apportées. Notamment, il y a peu d'approches qui utilisent la distorsion de l'axe temporel pour permettre à l'utilisateur d'examiner les enregistrements avec la granularité minimale. De plus, aucune approche n'intègre visuellement les prédictions des modèles dans le domaine médical, une fonctionnalité pourtant présente dans des approches non médicales comme [Zha+22; Wu+23] et qui aurait de la valeur dans le cadre de la prise en charge anticipée des patients. Enfin, le nombre d'approches interoperables est limité.

Dans un second temps, nous avons proposé GLIMPSE-Med, une interface visuelle ciblant deux publics : les professionnels de santé et les ingénieurs de données/statisticiens. À travers des encodages visuels innovants et des fonctionnalités interactives, l'interface a facilité la présentation de dossiers médicaux volumineux et multivariés dans une seule vue à l'écran. Notamment, ses capacités de distorsion temporelle et la possibilité de basculer entre des vues compactes et détaillées des séries de données ont offert aux utilisateurs une expérience de navigation simplifiée. Une évaluation empirique a été menée auprès de 14 professionnels de la santé affiliés au CHU de Montpellier. Les résultats ont indiqué un taux de précision de 89 % dans la résolution de tâches, tandis que les évaluations utilisant une échelle d'usabilité (SUS) ont positionné l'interface dans le plus haut quartile pour l'usabilité. De plus, les retours qualitatifs obtenus à partir de questions ouvertes ont permis d'identifier des améliorations potentielles. À travers tous les domaines, GLIMPSE-Med est la première interface, à notre connaissance, qui intègre visuellement la qualité des données ou les sorties de prédiction avec des séries temporelles multivariées dans une seule vue à l'écran. Dans le domaine médical, cette interface est la première à offrir la compression à la fois de l'axe temporel et de l'axe des modalités. Bien que testé uniquement dans un contexte médical, GLIMPSE-Med n'est pas limité à ce cas d'usage. L'évaluation a montré que GLIMPSE-Med répondait à la problématique, cependant les retours qualitatifs et les échanges avec les utilisateurs ont mis en évidence des pistes d'améliorations pour l'interface.

L'évaluation a montré l'efficacité de la barre de recherche. En revanche, en l'état, elle est limitée à une correspondance exacte dans les libellés des catégories. La recherche devrait être étendue à l'ensemble des données présentées dans l'interface, comme les valeurs des enregistrements ou le contenu des rapports textuels, ce qui implique d'interpréter des mesures non-textuelles et de calculer une proximité. De plus, la recherche serait plus intuitive et plus efficace si elle prenait en compte une proximité

floue avec les éléments filtrés. Nous envisageons l'usage d'une recherche basée sur la distance d'édition [Lev65] ou sur la proximité sémantique entre les termes [Guo+22], qui permet par exemple de trouver les synonymes. Cette proximité est obtenue par exemple par l'apprentissage de co-occurrences entre termes [Mik+13b].

L'objectif principal de l'interface est de mettre en valeur la trajectoire du patient. Puisque les données que nous présentons sont issues uniquement de services de soins intensifs, notre interface ne permet pas d'observer les transferts entre services ni les types de services visités. La succession des types de séjours visités et leurs temporalités doivent donc être représentés sur la frise temporelle. Cette donnée peut être considérée comme un enregistrement catégoriel puisque des coordonnées de début et de fin sont associées à chaque visite et visualisée comme celles des autres séries. Un code couleur guiderait l'utilisateur pour leur reconnaissance. Cependant, les types de séjours sont nombreux et variés avec par exemple, les services d'urgences ou de soins critiques et toutes les spécialités comme l'oncologie, les soins palliatifs, la pneumologie, etc. Or, il faut utiliser un nombre limité de couleurs pour que l'utilisateur soit en mesure de les distinguer. Notamment [HB03] préconisent l'usage de 7 couleurs maximum pour des cartes choroplèthes, ce qui semble trop peu pour la diversité des services à représenter. Une alternative serait l'usage d'un système de symboles dont la sémantique évoque directement le service représenté [Lam+08].

L'ordre alphabétique des catégories dans le menu de modalités (Fig. 2.3.b) a semblé peu naturel pour les utilisateurs. Cependant, regrouper de manière fonctionnelle un vaste ensemble d'étiquettes en langage naturel n'est pas trivial. Cela nécessite un tri et un regroupement effectués soit par un travail manuel coûteux par des experts, soit par des modèles linguistiques ou statistiques automatisés, soit par la récupération de relations sémantiques dans des bases de connaissances [Ran+23; MF20]. De plus, le choix optimal diffère selon le service hospitalier ou la discipline ciblée.

La visualisation proposée offre une vue individuelle des patients. Or, la gestion et la planification des lits est un domaine critique pour réduire les coûts, la charge de travail des soignants, l'attente et les risques chez les patients [Tor04; Ajm+22]. [How+08; VC20] montrent l'impact d'une gestion active des lits, qui peut être aidée par des systèmes automatisés. Ces méthodes seraient facilement complétées par un système visuel d'aide à la décision. Une solution possible serait de relier la vue actuelle à une méta-vue, d'un niveau supérieur, donnant une synthèse en temps réel de tous les patients présents et de la disponibilité des lits au sein d'un service donné, voire de l'hôpital. Chaque patient serait représenté par un ensemble de symboles et de couleurs qui synthétisent son état actuel et prédit. Le score de risque et les facteurs importants ainsi que le modèle développé dans le chapitre suivant seraient alors réutilisés dans cette vue. Cependant une étude des meilleures variables visuelles à considérer pour les vues et la méta-vue est indispensable et n'a pas été réalisée à ce stade. Nous reviendrons sur cette perspective en conclusion du mémoire.

Guidés par les retours des professionnels de santé, nous avons travaillé sur une nouvelle méthode pour prédire des événements médicaux que nous allons décrire dans le chapitre suivant. En particulier, l'originalité de cette approche consiste à utiliser des méthodes à base de convolutions de graphes pour considérer l'intégralité des

enregistrements hétérogènes et temporels d'un parcours patient sous forme de graphe temporel pour l'estimation du score par le modèle de prédiction d'un évènement clinique comme le décès à 24h.

GRAPHE TEMPOREL POUR LA MODÉLISATION DU PARCOURS PATIENT

Sommaire

3.1 Définitions	46
3.1.1 Graphe	46
3.1.2 Apprentissage de représentations de graphes	47
3.2 Paradigme encodeur-décodeur	49
3.3 Techniques d'apprentissage de représentations de graphes	52
3.3.1 Approches historiques	52
3.3.2 Marches aléatoires	53
3.3.3 Premiers réseaux de neurones pour les graphes	54
3.3.4 Réseaux convolutifs	54
3.3.5 Extensions de la convolution	56
3.3.6 Synthèse de l'historique	57
3.4 Graphes appliqués au médical	58
3.5 Graphes temporels de patients	59
3.5.1 Le dossier patient en tant que graphe temporel	60
3.5.1.1 Méthode	60
3.5.1.2 Fonction d'oubli	62
3.5.1.3 Expérimentations et résultats	63
3.5.2 Représentation latente des évènements médicaux	66
3.5.3 Limitations	67
3.6 Conclusion et Perspectives	68

Ce chapitre est consacré au développement de la direction de recherche 2 (DR2), telle que présentée dans la section 1.1. Cette direction vise à l'élaboration d'outils prédictifs pour détecter la survenue d'évènements cliniques en général. Ces méthodes permettent de suggérer des mesures préventives aux soignants, de personnaliser les traitements qui sont adaptés aux caractéristiques des patients, et d'optimiser les ressources médicales.

En particulier, l'objectif 2 que nous avons énoncé dans la section 1.4 consiste à proposer un modèle capable de réaliser des prédictions d'évènements tels que le décès inattendu, l'arrêt cardiaque ou le succès de réintubation, à partir de données temporelles et hétérogènes issues du DPI. Comme observé dans le chapitre 1, l'aspect temporel des enregistrements cliniques est un défi important. En effet, les modèles d'apprentissage séquentiels nécessitent un découpage des données selon des fenêtres discrètes. Ce découpage nécessite un niveau de connaissance minimal quant à la tâche d'apprentissage et aux caractéristiques des données. Par exemple, un découpage selon une granularité détaillée engendre des séquences très volumineuses tout en diminuant la proportion d'informations réellement mesurées. Nous avons aussi remarqué dans le chapitre 2 que le DPI présente des données multivariées et hétérogènes. Cette variété complexifie les pré-traitements manuels en rendant les parcours médicaux complexes à appréhender dans leur entièreté ou les adaptations semi-automatiques coûteuses par l'explosion combinatoire de la quantité d'ajustements possibles. Ce chapitre étudie l'utilisation de structures de graphes en réponse à cette problématique.

Ce chapitre est organisé de la manière suivante : la section 3.1 présente les définitions du graphe en tant que structure de données et de l'apprentissage de représentations de graphes. La section 3.2 détaille le paradigme encodeur-décodeur sur lequel nous nous appuyons pour l'argumentation dans ce chapitre. La section 3.3 décrit un historique des techniques d'apprentissage sur les graphes. La section 3.4 liste les applications existantes de ces méthodes dans le cas du DPI. La section 3.5 est le cœur de nos travaux et décrit notre modélisation du parcours patient sous forme de graphe temporel ainsi que les expérimentations menées, comparant essentiellement des approches séquentielles basées sur le LSTM et notre approche basée sur la convolution de graphes. Enfin, nous concluons dans la section 3.6 et proposons des perspectives.

3.1 Définitions

La piste que nous avons envisagée pour l'apprentissage sur les enregistrements cliniques repose sur une structure de données traditionnelle : **les graphes**.

3.1.1 Graphe

Le graphe est un modèle utilisé en mathématique et en informatique pour décrire des réseaux d'objets. Dans sa description la plus élémentaire, un graphe est un couple d'ensembles, usuellement notés V et E tels que $G = (V, E)$, décrivant respectivement l'ensemble des sommets (V) et des arêtes (E). Ce cadre théorique permet de décrire des réalités complexes et s'adapte à des applications variées. En complément de cette

définition, des propriétés peuvent être ajoutées selon les besoins théoriques ou applicatifs. D'un point de vue computationnel, ce modèle théorique est flexible et représenté par diverses structures de données, telles que des matrices ou des listes d'adjacence, par exemple. Chacune de ces structures présente des avantages et des inconvénients selon les algorithmes étudiés, les applications développées ou les caractéristiques des graphes prises en compte. La Figure 3.1 montre quatre graphes aux caractéristiques variées. La structure du graphe est décrite par une liste de sommets et une matrice d'adjacence. Cette figure illustre l'incidence des propriétés des structures de données sur la représentation visuelle des graphes.

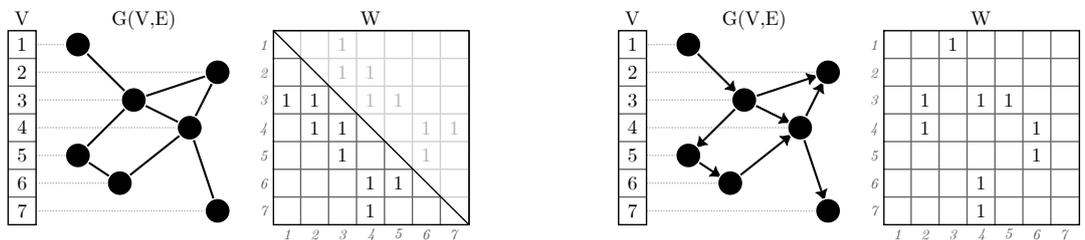
Les sommets, aussi appelés nœuds, sont l'unité du graphe, c'est-à-dire la plus petite partie indécomposable de cet objet. Les sommets représentent des entités variées selon le domaine d'application comme en bio-informatique où une molécule est représentée comme un réseau d'atomes [Mah+21], en détection de fraudes dans des réseaux où les sommets sont des utilisateurs qui interagissent entre eux [Liu+21] ou en urbanisme pour la prédiction de trafic dans un réseau routier où les sommets sont des radars placés à des endroits fixes [Sha+23]. Suivant le cas d'application, les sommets comportent des propriétés intrinsèques, comme un type ou des valeurs numériques. Par exemple, les sommets de la Figure 3.1(d) contiennent des caractéristiques décrites dans la matrice X de dimension 3. Cette représentation prend la forme d'un encodage 1 parmi n (*one-hot encoding*) assimilé à un typage des sommets. Sur la représentation visuelle, chaque dimension de X est représentée par un canal du code couleur rouge-vert-bleu et chaque sommet est coloré en conséquence.

Les arêtes, aussi appelées liens, décrivent les relations entre les sommets du graphe. Chaque arête décrit l'existence d'une relation entre deux sommets donnés, usuellement nommés u et v . Ces deux sommets (u, v) sont dits voisins ou adjacents. La relation induite par la présence d'une arête est symétrique dans le cas d'une arête non orientée, asymétrique si l'arête est orientée. Ainsi, dans le cas d'un graphe orienté on parle d'arcs. La Figure 3.1(a) montre un graphe non-orienté dont la structure est décrite par une matrice d'adjacence W qui est symétrique, c'est-à-dire $W_{uv} = W_{vu}$. Au contraire, la Figure 3.1(b) montre un graphe orienté où la matrice W est asymétrique. Un couple de sommets peut être lié par plusieurs arêtes dans le cas d'un multigraphe, en opposition à un graphe simple. Enfin, les arêtes peuvent contenir des propriétés, comme un poids. Un exemple de pondération est visible dans la Figure 3.1(c) où la matrice W est définie sur $\mathbb{R}^{|V| \times |V|}$, ce qui est représenté visuellement par des arêtes d'opacité variable. Les autres exemples de graphes des Figures 3.1(a), 3.1(b) et 3.1(d) comportent des arêtes binaires décrites par des matrices d'adjacence $W \in \{0, 1\}^{|V| \times |V|}$.

3.1.2 Apprentissage de représentations de graphes

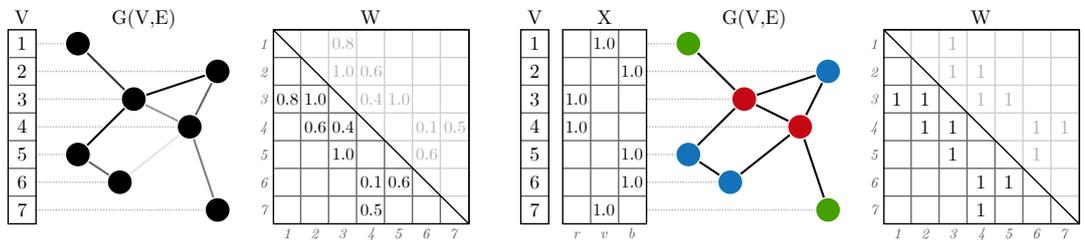
L'apprentissage de représentations de graphes est une technique qui consiste à trouver une fonction d'un graphe vers un espace continu. En particulier, l'objectif consiste à représenter la structure et les attributs des graphes dans un espace latent. Trouver une représentation de la structure n'est pas trivial puisque les sommets et les arêtes n'ont pas d'ordre, contrairement au texte où les mots sont organisés de manière séquentielle dans une phrase ou bien aux images où les pixels sont disposés sur une

FIGURE 3.1 – Quatre exemples de graphes aux caractéristiques variées et impact de leurs caractéristiques sur la représentation visuelle et sur les structures de données. Les cellules vides des matrices représentent une valeur nulle, soit 0, pour la lisibilité. Le masque appliqué sur les matrices d'adjacence symétriques des sous figures (a, c et d) occulte les informations redondantes.



(a) Structure de graphe G , décrite par une liste de sommets V et une matrice d'adjacence symétrique W .

(b) Structure de graphe orienté G , décrite par une liste de sommets V et une matrice d'adjacence asymétrique W .



(c) Structure de graphe pondéré G , décrite par une liste de sommets V et une matrice d'adjacence symétrique W composée de réels.

(d) Structure de graphe G , décrite par une liste de sommets V , les caractéristiques de ces sommets X et une matrice d'adjacence symétrique W . La représentation visuelle du graphe montre les 3 dimensions de X à travers les 3 canaux du système de couleur rouge, vert et bleu.

grille ordonnée [Hoa+23].

Formellement, à partir d'un graphe $G = (V, E)$ qui est une structure discrète, le but est d'apprendre une représentation vectorielle $Z \in \mathbb{R}^{|V| \times d}$ des sommets V du graphe, des arêtes E ou du graphe G lui-même. d est la taille fixe de l'espace latent pour le plongement, dont la valeur est choisie. Cette représentation vectorielle doit préserver les propriétés du graphe original. Ces propriétés diffèrent selon les méthodes utilisées et les objectifs théoriques ou applicatifs à remplir. En vue de nos objectifs, nous pouvons observer que l'avantage principal est l'apprentissage de représentations à partir d'un ensemble de sommets et d'arêtes de taille non-fixe. Ces plongements dans des espaces latents alimentent ensuite des systèmes prédictifs pour des tâches variées telles que la classification de sommets, le regroupement ou la prédiction de liens [HYL18].

L'apprentissage de représentations de graphes est un processus complexe décomposable en de nombreuses étapes comme, par exemple, l'expression de la tâche d'apprentissage, les typologies de graphes prises en compte ou les propriétés théoriques du processus de convergence vers une solution, etc. L'inventaire de ces méthodes a été réalisé à plusieurs reprises dans la littérature [GF18; Lee+19; Zho+20; Cha+22; Hoa+23; KA24]. L'apparition de nouvelles méthodes ou taxonomies a été l'occasion pour les auteurs de synthétiser et classer les techniques selon différents paradigmes qui se sont succédés dans le temps. Le choix d'une intuition ou d'un vocabulaire pour l'expression des méthodes oriente les choix scientifiques ou techniques. La section suivante 3.2 décrit le paradigme *encodeur-décodeur* qui a orienté la rédaction de ce chapitre. La section 3.3 montre l'évolution dans le temps de ces techniques d'apprentissage. La section 3.4 présente des applications de ces techniques au domaine médical.

3.2 Paradigme encodeur-décodeur

Principe général Les techniques d'apprentissage de représentations de graphes sous forme d'encodeur et de décodeur ont d'abord été introduites par [HYL17] puis étendues dans [Cha+22]. Formellement, d'après ce dernier, on considère le graphe $G = (V, E)$ et $X \in \mathbb{R}^{|V| \times d_0}$ les caractéristiques initiales de dimension d_0 des sommets et $W \in \mathbb{R}^{|V| \times |V|}$ la matrice d'adjacence, pour décrire les caractéristiques de ces deux ensembles. Dans un cadre non supervisé, c'est-à-dire sans étiquette à prédire, un modèle d'apprentissage est constitué de deux éléments :

- Un **encodeur de graphe** $\text{ENC}_{\Theta^E} : \mathbb{R}^{|V| \times |V|} \times \mathbb{R}^{|V| \times d_0} \rightarrow \mathbb{R}^{|V| \times d}$, avec pour paramètres Θ^E , qui combine la structure et les caractéristiques des sommets pour produire une matrice latente de plongements de sommets $Z \in \mathbb{R}^{|V| \times d}$, telle que :

$$Z = \text{ENC}(W, X; \Theta^E)$$

- Un **décodeur de graphe** $\text{DEC}_{\Theta^D} : \mathbb{R}^{|V| \times d} \rightarrow \mathbb{R}^{|V| \times |V|}$, avec pour paramètres Θ^D , qui utilise la matrice latente des sommets Z pour calculer la similarité de chaque paire de sommets sous forme de matrice $\widehat{W} \in \mathbb{R}^{|V| \times |V|}$ de même forme que la matrice d'adjacence, telle que :

$$\widehat{W} = \text{DEC}(Z; \Theta^D)$$

Le modèle peut alors être décrit comme une fonction de reconstruction d'une matrice de similarité \widehat{W} à partir de la matrice d'adjacence W , telle que :

$$\widehat{W} = \text{DEC} \left(\text{ENC} \left(W, X; \Theta^E \right); \Theta^D \right)$$

La fonction est représentée visuellement dans la Figure 3.2(a).

Dans un cadre supervisé, l'objectif est de prédire l'étiquette des sommets $Y \in \mathcal{Y}^{|V|}$, où \mathcal{Y} est l'espace des étiquettes. Une alternative au décodeur de graphe est alors choisie :

- Un **réseau de classification** $\text{DEC}_{\Theta^S} : \mathbb{R}^{|V| \times d} \rightarrow \mathbb{R}^{|V| \times |\mathcal{Y}|}$ qui a pour paramètres Θ^S . La sortie est une distribution \widehat{Y} sur les étiquettes, à partir de la matrice latente des sommets, telle que :

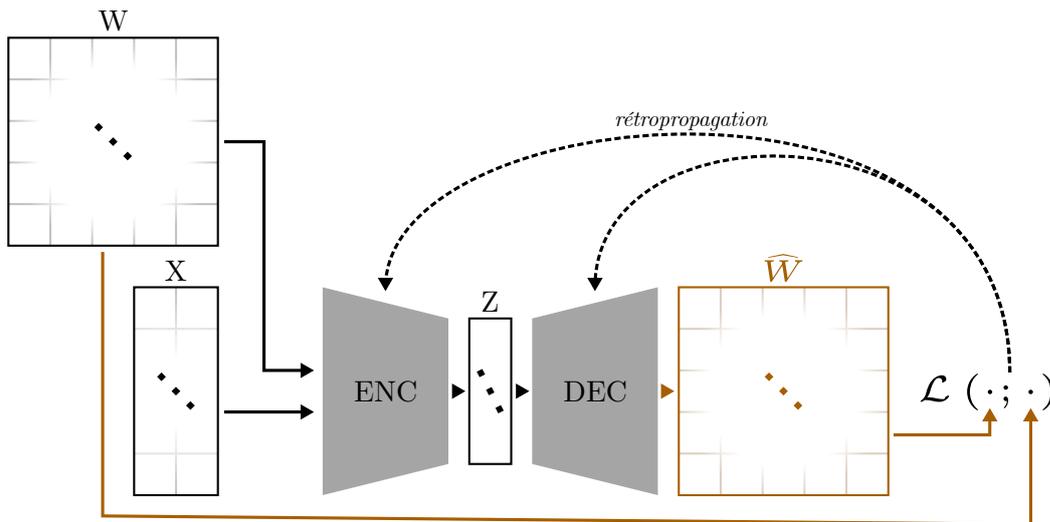
$$\widehat{Y} = \text{DEC} \left(Z; \Theta^S \right)$$

$$\widehat{Y} = \text{DEC} \left(\text{ENC} \left(W, X; \Theta^E \right); \Theta^S \right)$$

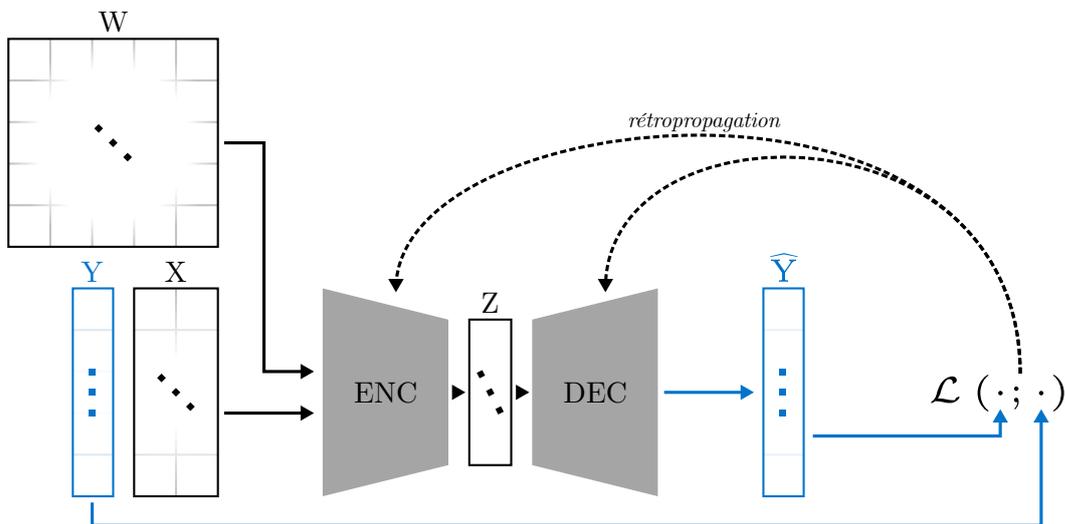
La fonction est représentée visuellement dans la Figure 3.2(b).

On observe sur les Figures 3.2(a) et 3.2(b) que l'encodeur ne varie pas suivant la tâche ou le contexte d'apprentissage et qu'il faut uniquement adapter le décodeur et la fonction de perte en conséquence. De plus, la présentation des techniques d'apprentissage de représentations de graphes sous forme d'encodeur et de décodeur permet de décrire un contexte d'entraînement de manière modulaire et permet de dissocier la description du fonctionnement intrinsèque d'un modèle d'un côté et de l'autre ses modalités d'entraînement.

Informations disponibles La distinction entre apprentissage transductif et inductif décrit la capacité d'un modèle à être utilisé sur des données qu'il n'a pas encore observées. La distinction est différente de l'opposition entre apprentissage supervisé et non-supervisé. Dans un contexte transductif, l'entière des sommets est observé durant l'apprentissage, par exemple dans le cas d'un seul graphe fixe. Un découpage en jeu d'entraînement et en jeu de test est réalisé en séparant les sommets d'un même graphe. Cette méthode est utilisée pour découvrir de nouvelles relations ou de nouvelles étiquettes dans une structure déjà connue. Dans le contexte inductif au contraire, le modèle est utilisé en test sur des graphes qui n'ont pas été observés pendant l'entraînement [Cha+22], ce qui est très utile, en particulier pour des tâches à grande échelle [YCS16]. L'apprentissage transductif est généralement plus performant que l'apprentissage inductif car l'apprentissage transductif exploite les données de test non étiquetées [YCS16]. L'apprentissage inductif est une tâche plus difficile à résoudre. Historiquement, les approches ont d'abord été développées pour le contexte transductif, comme les approches purement spectrales [BN03; Bru+14; KW17], puis ont été étendues pour l'usage en contexte inductif [HYL18]. Cependant, la classification des méthodes d'après la dichotomie entre transductif et inductif est non pertinente puisque [SR20] a montré que des méthodes développées pour des contextes spécifiques pouvaient fonctionner dans d'autres contextes. Il est donc important d'évaluer les performances de ces méthodes dans des contextes variés.



(a) Présentation du contexte non-supervisé dans le cas de prédiction de liens. Le décodeur produit à partir de la représentation latente Z une distance pour chaque couple de sommets, sous la forme d'une nouvelle matrice d'adjacence reconstruite. La fonction de perte mesure la différence entre la matrice d'adjacence et la matrice de distance entre sommets reconstruite pour mettre à jour l'encodeur et le décodeur du modèle.



(b) Présentation du contexte supervisé dans le cas de la classification de sommets. Le décodeur produit à partir de la représentation latente Z des étiquettes prédites pour l'ensemble des sommets. La fonction de perte mesure la différence entre les étiquettes de supervision Y et les étiquettes prédites pour mettre à jour l'encodeur et le décodeur du modèle.

FIGURE 3.2 – Présentation visuelle du principe d'apprentissage de représentations de graphes sous forme d'encodeur-décodeur [Cha+22]. La figure est décomposée en deux parties qui présentent le contexte non-supervisé et supervisé. Les symboles en noir montrent les parties communes aux deux contextes : les structures de données qui décrivent le graphe, soit une matrice d'adjacence quelconque W et une matrice de caractéristiques de sommets X , sont entrées dans un encodeur pour produire une représentation latente quelconque Z . Ensuite, un décodeur utilise la représentation latente pour produire en sortie des probabilités de classes ou une matrice de distance, par exemple, en fonction de la tâche visée et du contexte d'apprentissage pour calculer la fonction de perte. La fonction de perte est utilisée pour ajuster les poids des deux parties du modèle, l'encodeur et le décodeur, communément par rétropropagation du gradient. Les différences entre les deux contextes sont montrées en couleur, en orange pour le contexte non-supervisé et en bleu pour le contexte supervisé.

Sortie du modèle La description ci-dessus se concentre sur des tâches d'apprentissage au niveau des sommets. Selon les besoins, la cible de l'apprentissage peut être les arêtes ou les graphes. Parmi quelques options, la régression ou la classification sur les arêtes prend la forme de prédiction d'un poids continu, d'un type prédéfini ou de la probabilité de l'existence d'une arête [Zho+20]. Dans un réseau social par exemple, la recommandation d'amis pertinents consiste à prédire de nouveaux liens dans un graphe composé d'utilisateurs [BL11]. Pour l'apprentissage à l'échelle du graphe, le modèle doit capturer grâce à des plongements latents les caractéristiques du graphe, qui est une structure complexe, dans un vecteur plat ou un score, pour de la classification ou de la régression.

3.3 Techniques d'apprentissage de représentations de graphes

Comme évoqué dans la section 3.1.2, une représentation latente pertinente des éléments du graphe doit préserver la structure et les connexions entre les sommets. Le défi principal consiste à choisir la propriété du graphe que le plongement dans un espace latent doit préserver. Étant donné la multitude de propriétés définies pour les graphes ou de métriques de distance par exemple, ce choix est difficile et les performances dépendent de l'application [GF18]. De nombreuses méthodes ont essayé de présenter des propriétés pour l'apprentissage de représentations de graphes ainsi que des techniques pour tenter de résoudre ces problèmes que nous décrivons ci-après.

3.3.1 Approches historiques

L'apprentissage de représentation sur les graphes n'est pas un domaine nouveau. Si des développements récents sont apparus avec l'apparition de modèles neuronaux profonds, il existait déjà des méthodes basées sur la théorie des graphes.

Noyaux de graphes Les modèles d'apprentissage de représentations de graphes les plus anciens sont les noyaux de graphes (*graphs kernels*, apparus en 1999 avec [Hau99]). L'idée derrière cette méthode est l'utilisation de fonction de noyaux pour mesurer la similarité entre des sous-parties caractéristiques de graphes à la manière du test d'isomorphisme de Weisfeiler-Lehman [Hoa+23].

Factorisation de matrices Une autre approche, dès 2001 [BN01], est la factorisation de matrices inspirée des méthodes de compression de l'information structurelle. L'objectif est de réduire la taille de la matrice d'adjacence pour permettre de travailler sur des graphes avec un grand nombre de sommets, à une époque où les ressources computationnelles étaient plus limitées. Par exemple, certaines de ces méthodes utilisent les *Laplacian eigenmaps* pour représenter les sommets avec leurs valeurs propres [BN03].

D'après les travaux réalisés dans [Hoa+23], le nombre d'études sur l'apprentissage de représentations de graphes a considérablement augmenté au cours des dernières

années. Durant cette période, les études basées sur les développements en apprentissage profond ont connu une croissance significative. En revanche, le nombre d'études sur les noyaux de graphes et les modèles basés sur la factorisation de matrices n'a que légèrement augmenté. La principale limitation de ces approches réside dans leur forte complexité computationnelle.

3.3.2 Marches aléatoires

En parallèle de ces développements directement appliqués à l'apprentissage de représentations de graphes, des systèmes de recommandations se sont développés. Assimilable à une tâche de prédiction de liens dans un graphe, l'algorithme PageRank [BP98], fondateur de Google, est basé sur l'idée de marche aléatoire (*random walks*, *random surfer* dans le texte original). L'intuition derrière cette technique est de pouvoir bénéficier des approches d'autres domaines qui fonctionnent sur des données vectorielles ou séquentielles en extrayant des séquences de sommets au sein d'un graphe. Ces séquences de sommets sont choisies aléatoirement en parcourant les arêtes du graphes, générant ainsi des séquences représentatives des voisinages.

Ce type de méthode a connu de nombreux développements dans les années 2010 suite au gain de popularité des approches d'apprentissage profonds liée à l'augmentation des capacités de calcul. En particulier, les progrès dans l'apprentissage de texte comme les modèles Word2Vec avec CBOW et SkipGram [Mik+13a; Mik+13b] ont encouragé l'utilisation de marches aléatoires pour appliquer ces techniques sur les graphes. Ainsi des travaux comme [BL11; PAS14; GL16] appliquent ces marches selon différentes stratégies, pour faire de la prédiction de liens ou du plongement latent de sommets.

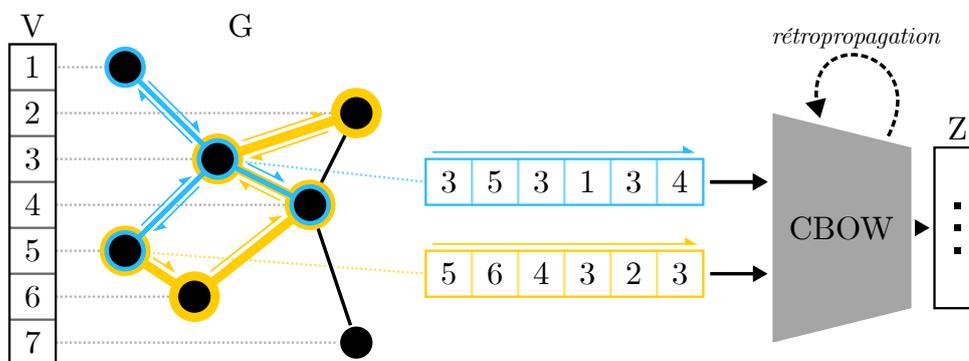


FIGURE 3.3 – Illustration du principe d'apprentissage sur deux marches aléatoires tirées au sein d'un graphe par l'application de *node2vec* [GL16].

La Figure 3.3 présente un exemple d'application d'une méthode basée sur les marches aléatoires pour l'apprentissage de représentation des sommets. En particulier, l'illustration montre le principe de *node2vec* [GL16]. Deux marches aléatoires sont extraites du graphe illustrant les deux stratégies évoquées dans ces travaux. Le choix de la stratégie est contrôlé par un paramètre de retour en arrière. La marche bleue a pour origine le sommet 3 et visite successivement les sommets 5, puis 1, puis 4. Arbitrairement, la probabilité de retour en arrière étant élevée pour cet exemple, la

marche revient au sommet 3 pour réaliser un parcours en largeur, i. e. concentré dans le voisinage du sommet de départ. À l'inverse, le parcours jaune a une faible probabilité de retour en arrière et est un parcours en profondeur, avec pour sommet de départ le sommet 5, puis les sommets 6, 4, 3, 2 et 3. Ces deux marches aléatoires forment chacune une séquence de sommets de taille fixe. Ces séquences viennent nourrir un modèle CBOW [Mik+13a] où chaque sommet distinct est représenté par un jeton (*token*), i. e. un identifiant unique. Si on fait une analogie avec l'apprentissage textuel, une marche aléatoire pourrait être assimilée à une phrase, où chaque sommet successif représente un mot de cette phrase.

3.3.3 Premiers réseaux de neurones pour les graphes

Comme mentionné dans les sections précédentes, les méthodes traditionnelles traitent les graphes à travers une série de pré-traitements qui les transforment en représentations plus simples, comme des vecteurs ou des séquences. Ces techniques montrent parfois des performances satisfaisantes. Cependant, des informations sont perdues lors de ces pré-traitements et, par conséquent, elles impactent les modèles. Le GNN (*Graph Neural Networks*), ou réseau de neurones pour graphes, est une nouvelle formulation [GMS05] qui a été étendue dans [Sca+09]. Ce modèle voit la supervision d'un graphe comme un mécanisme de diffusion où les sommets envoient des informations à leurs voisins jusqu'à ce qu'un point d'équilibre soit atteint [Cha+22].

Ces approches qui parcourent la structure du graphe de manière récursive sont une extension des marches aléatoires et sont inspirées par des modèles performants sur des données séquentielles comme les chaînes de Markov. Rétrospectivement, ce nouveau paradigme a unifié les méthodes les plus anciennes pour la transformation des graphes tout en étant précurseur du paradigme de la transmission de messages [ZCZ22] que nous évoquerons dans la section suivante 3.3.4.

3.3.4 Réseaux convolutifs

Introduits dans [KW17], les GCN (*Graph Convolutional Networks*) ou réseaux de neurones convolutifs pour graphes utilisent la convolution en tant que couche de propagation à la place du parcours récursif utilisé auparavant dans les GNN. L'usage de la convolution a fait ses preuves dans l'application des réseaux de neurones sur les images, un domaine spatial, mais son adaptation à une structure non euclidienne n'est pas triviale. En effet, le voisinage d'un sommet n'a ni ordre, ni taille fixe.

Convolution spectrale Dans un premier temps, l'opération de convolution du domaine spatial a été adaptée aux graphes par l'utilisation du domaine spectral [Cha+22]. Ces travaux appliquent des filtres de convolution sur les vecteurs propres de la matrice Laplacienne [Bru+14] ou leurs approximations [DBV16; KW17] calculées à partir d'une matrice d'adjacence composée de valeurs réelles. Cependant, ces algorithmes nécessitent la matrice Laplacienne complète pendant l'entraînement, ce qui limite le passage à l'échelle sur de grands graphes [HYL18] et contraint l'usage de ces méthodes à un contexte transductif.

Convolutions spatiales D'autres approches ont donc adapté la convolution en restant dans le domaine spatial présentant les avantages suivants dans les graphes : (1) la convolution extrait les informations locales à partir des voisins de chaque sommet ; (2) le partage des paramètres prend en compte les caractéristiques de tous les sommets dans l'encodeur et (3) l'utilisation de plusieurs couches de convolution successives étend le nombre de localités explorées [KA24].

Les approches spatiales définissent les convolutions directement sur le graphe en se basant sur sa topologie. Le principal défi est de définir l'opération de convolution avec des voisinages de tailles différentes et de maintenir l'invariance locale des réseaux de neurones convolutifs. GraphSAGE [HYL18] contourne la contrainte du voisinage de taille variable en ajoutant un échantillonnage de taille fixe lors de la propagation. Cette approche évite aussi la contrainte d'ordre en utilisant une fonction d'agrégation du voisinage invariante à l'ordre, ou en tirant des ordres aléatoires. Par ailleurs, elle étend les possibilités d'apprentissage au contexte inductif, généralisable à de nouveaux sommets et de nouveaux graphes.

Transmission de messages Dans une tentative d'unifier les travaux évoqués précédemment, une autre formulation de ces méthodes consiste à voir la convolution comme une transmission d'informations entre les sommets à travers la structure du graphe. Ce paradigme, appelé transmission de messages ou passage de messages (*message passing*), a d'abord été exprimé dans [Gil+17]. Les méthodes de transmission de messages peuvent être considérées comme des convolutions spatiales et réciproquement, car les messages sont calculés sur les voisinages locaux des sommets [HYL18].

D'après ce paradigme, l'encodeur est décrit par la succession de trois types principaux de couches [Zho+20] :

- La **couche de propagation** propage les informations à travers la structure du voisinage d'un sommet et agrège ces informations pour en capturer les propriétés.
- La **couche d'échantillonnage** sélectionne un sous ensemble de sommets.
- La **couche de synthèse ou regroupement** (*pooling* ou *readout*) regroupe et capture les informations d'un sous-graphe, soit plusieurs sommets, dans une nouvelle représentation de taille fixe. Cela réduit la dimensionalité de la représentation et l'aplatit à l'échelle d'un sous-graphe ou d'un graphe, par exemple, pour l'entraînement d'un décodeur placé à la suite.

Typiquement, un encodeur est constitué d'une succession de couches de propagation, chacune pouvant être accompagnée d'une couche d'échantillonnage. Dans le cas de la supervision de graphe, on place en sortie une couche de synthèse.

Un exemple d'encodeur est présenté dans la Figure 3.4. L'illustration représente visuellement l'application d'une couche de propagation à la manière du GCN [KW17].

1. En entrée de l'encodeur, le graphe non-orienté G est décrit par une matrice d'adjacence symétrique W et une matrice de caractéristiques de sommets X de dimension 3. Visuellement, les sommets sont représentés selon le système de couleurs rouge-vert-bleu comme pour la Figure 3.1(d) associé aux trois dimensions

de la matrice de caractéristiques.

2. Au sein de l'encodeur, la couche de convolution correspond à la transmission des caractéristiques de tous les sommets à leurs voisins. La matrice d'adjacence W est additionnée à la matrice identité I de même dimension. Cette matrice définit chaque sommet comme son propre voisin, sous forme de boucle. La convolution permet alors de combiner la représentation des sommets X , le voisinage d'ordre 0 à travers I , avec le voisinage d'ordre 1 à travers W . La transmission d'informations est représentée visuellement par les flèches entre les sommets. La structure du graphe disparaît mais sert de canal pour l'échange d'informations entre les sommets. Dans cet exemple, l'encodeur est constitué d'une seule couche de convolution. À la suite de la transmission d'informations, une couche de normalisation réduit l'amplitude des vecteurs à une norme commune. À l'issue de ces étapes, la représentation d'un sommet est alors la moyenne de sa représentation et de celle de ses voisins.
3. À la sortie de l'encodeur, on observe que la représentation latente de sommets Z est une synthèse à la fois des caractéristiques originales X et de la structure W . Par exemple, les sommets 1 et 7 sont tous les deux de couleur verte et ont pour voisinage un seul sommet de couleur rouge. Respectivement les sommets 3 et 4, ont le même vecteur caractéristique et sont donc de même couleur dans la représentation latente. Cette représentation est la synthèse des vecteurs décrivant les couleurs verte et rouge des sommets du graphe original. On observe que les caractéristiques extraites correspondent au voisinage d'ordre 1, puisque l'encodeur est constitué d'une seule couche de convolution.

Ainsi, de nombreuses approches développées au fil du temps peuvent être décrites comme des variantes du GCN avec différentes agrégations du voisinage et différentes synthèses du graphe [Xu+19].

3.3.5 Extensions de la convolution

Les paradigmes présentés, soit l'encodeur-décodeur et la convolution ou la transmission de messages, servent à décrire des modèles et des modalités d'entraînement complexes. De nombreuses améliorations visent à obtenir de meilleures performances de prédiction.

L'efficacité des GCN repose sur une hypothèse fondamentale : l'**homophilie**, un principe observable dans de nombreux réseaux du monde réel selon lequel les sommets reliés entre eux appartiennent souvent à la même classe ou ont des caractéristiques similaires [MSC01] et leurs plongements dans l'espace latent doivent avoir une distance faible [GL16]. [Zhu+20] montrent que le plongement concaténé d'un sommet avec celui de ses voisins, plutôt que l'agrégation de tous, obtient de meilleures performances sur des réseaux hétérophiles, i. e. qui ne satisferaient pas l'hypothèse d'homophilie.

Des travaux ont émergé pour incorporer les mécanismes d'attention aux graphes, une technique améliorant les méthodes d'apprentissage profond en général. Si ce mécanisme adopte des formes variées, l'objectif est que le modèle accorde une plus grande priorité aux éléments les plus pertinents. Par exemple, la convolution avec

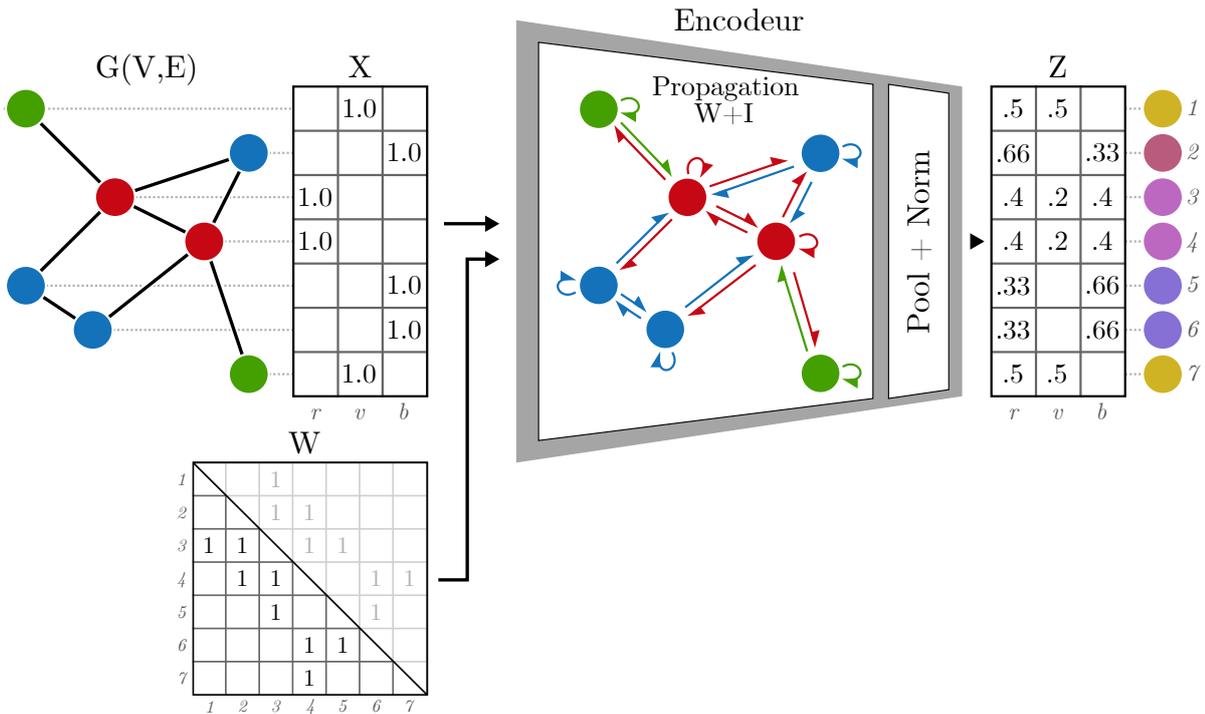


FIGURE 3.4 – Illustration d'un encodeur GCN [KW17] composé d'une couche de propagation, d'une couche de regroupement et d'une normalisation. Dans cet exemple, la normalisation utilisée pour produire le vecteur Z est la normalisation ℓ_1 qui utilise la norme absolue telle que $\|v\|_1 = \sum_{i=1}^n |v_i|$.

attention correspond à une somme pondérée de la représentation de ses voisins, où le poids est une fonction des vecteurs caractéristiques de l'ego et d'un voisin donné [Vel+18]. Les bénéfices sont les suivants [Lee+19], par le biais de l'attention, le modèle : (1) masque les parties du graphe qui apportent du bruit, (2) met en évidence les éléments contenant les informations les plus pertinentes pour la tâche et (3) fournit également un moyen de rendre les résultats d'un modèle plus interprétables, en mettant en valeur les principaux facteurs contribuant à une décision en particulier.

3.3.6 Synthèse de l'historique

Nous présentons une chronologie récapitulative des méthodes d'apprentissage de représentations de graphes que nous illustrons dans la Figure 3.5. Cette illustration montre les méthodes décrites dans les paragraphes précédents avec en particulier les méthodes à base de marches aléatoires en bleu et les réseaux de neurones pour graphes en orange. On observe sur cette chronologie que les marches aléatoires ont été utilisées tôt, dès 1998 avec [BP98] et ont connu des développements dans les années 2010 [BL11; PAS14; GL16]. Les développements les plus récents font partie de la famille des réseaux de neurones pour graphes apparus dans les années 2000 avec [GMS05; Sca+09]. La formulation de ces modèles sous forme d'encodeur-décodeur avec GraphEDM [Cha+22] est très récente à l'échelle du domaine.

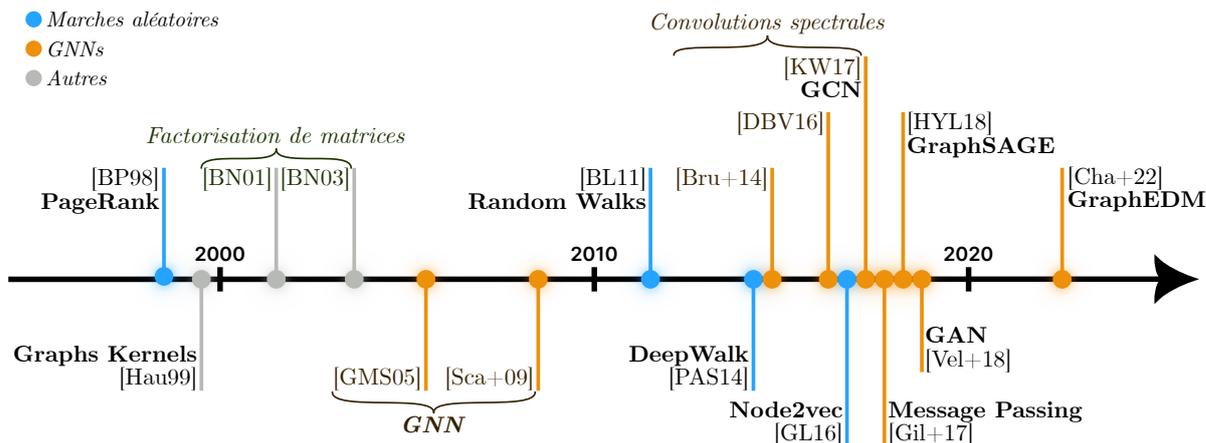


FIGURE 3.5 – Chronologie récapitulative des méthodes présentées dans ce chapitre. Chaque méthode est représentée par un point le long de l'axe temporel, avec les méthodes à base de marches aléatoires en bleu, les réseaux de neurones pour graphes en orange et les autres en gris.

3.4 Graphes appliqués au médical

Les modèles d'apprentissage de représentations de graphes sont utilisés dans les domaines qui reposent naturellement sur des structures à base de réseaux. Ainsi, dans les domaines clinique ou biomédical, les GNN ont été utilisés dans des exemples variés comme l'histopathologie [Ahm+22], la modélisation d'interactions de protéines [PSH22]), l'analyse de réseaux cérébraux [Cha+21], l'étude des effets secondaires de plusieurs médicaments [ZAL18] ou l'estimation de la survie au cancer à partir de l'expression des gènes [Gao+22].

Dans le cas spécifique du DPI, les données sont temporelles et multivariées ce qui rend l'utilisation de graphes non triviale et représentent les deux défis principaux étudiés dans cette thèse. Aussi, les représentations apprises par les méthodes d'apprentissage profond devraient être alignées avec les connaissances médicales [Cho+16] qui existent déjà sous forme structurée. Les exemples suivants dans cette section tentent de résoudre des tâches prédictives à partir du DPI par l'utilisation de réseaux de neurones à base de graphes.

De nombreuses approches réutilisent le découpage d'un parcours patient en visites successives pour modéliser l'évolution de son état et constituer le graphe. [Liu+20] combinent l'apprentissage sur des séjours successifs par un RNN avec un GCN appliqué sur une matrice de co-occurrence des enregistrements. [Gu+22] produisent une similarité entre séjours de patients par l'entraînement de réseaux siamois, c'est-à-dire par l'ajustement de la structure d'un graphe décrivant les séjours d'un patient par paire de séjours. [Cho+16] complètent les enregistrements du DPI avec des informations hiérarchiques inhérentes aux ontologies médicales. Basée sur la structure de l'ontologie, cette approche représente un concept médical dans l'espace latent comme une combinaison de ses ancêtres dans l'ontologie via un mécanisme d'attention. Cette méthode consiste donc à enrichir le plongement des codes diagnostics associés aux vecteurs aplatis des séjours. La prédiction est alors réalisée par un RNN. [Li+20a] apprennent une succession de codes associés aux visites. Ces codes décrivent les diagnostics, les

actes et les médicaments. Cette méthode enrichit cette représentation avec la structure hiérarchique commune à tous les parcours en tant que matrice d'adjacence. Le modèle utilisé est une convolution sur deux axes : spatial par rapport à la structure du graphe et temporel par rapport à l'axe temporel discrétisé en séjours. On observe alors que si ces méthodes appliquent des réseaux à base de graphes pour l'enrichissement des concepts présents dans les DPI à partir de connaissances structurées ou apprises, elles se basent principalement sur des modèles séquentiels pour leurs prédictions.

Certaines approches utilisent le graphe comme un réseau de similarité entre patients. Ainsi, les connaissances extraites sur les patients avec des caractéristiques communes pourraient s'enrichir mutuellement. Par exemple, [Cho+22] réalisent de la prédiction à l'échelle du parcours patient grâce à la concaténation de 3 représentations : (1) un LSTM traite les données temporelles, (2) un MLP traite les données démographiques et (3) un GNN est appliqué sur un graphe dont la matrice d'adjacence est une similarité entre patients. La similarité est calculée à partir du nombre de diagnostics en commun.

L'utilisation de ces méthodes pourrait répondre aux deux problématiques des données du DPI : la temporalité et l'hétérogénéité. Récemment, les approches ont de plus en plus tiré parti des informations multimodales, hétérogènes ou multivariées du DPI. Les événements médicaux seraient traduits dans un graphe hétérogène afin de produire un réseau dense d'interactions et capturer la complexité de l'état de santé d'un patient. De plus, les méthodes d'apprentissage seraient alors appliquées à des paramètres cliniques et à des sources de données plus variées [Bol+24]. Étant donné que la performance de prédiction dans l'estimation de l'apparition d'évènements cliniques serait favorisée par la prise en compte des aspects temporels du DPI, la représentation des enregistrements sous forme de graphe intègre cet aspect. Une méthode consiste à former une série de graphes qui correspondent à différents moments pour un patient donné. Une autre utilise un graphe dynamique, où les sommets et les arêtes changent au fil du temps pour correspondre à la réalité observée [Bol+24]. Ces deux méthodes s'appuient sur des modèles prédictifs utilisant la convolution ou la récurrence et considèrent le temps de manière séquentielle. Or le temps est intrinsèquement continu et sa discrétisation constitue une perte d'information.

3.5 Graphes temporels de patients

Dans les sections précédentes, nous avons décrit la capacité des modèles à base de graphe et de leur formulation sous forme d'encodeur-décodeur à exprimer des problématiques et des contextes d'apprentissage variés. Cette modularité offre l'utilisation de ces formulations hors de leur champ d'étude, comme pour la prédiction d'enregistrements au sein de parcours cliniques. Le graphe est une structure de données qui peut être utilisée comme un support pour transmettre des informations sémantiques à travers ses sommets par l'utilisation de modèles à base de convolutions de graphes. Cette technique simplifie l'uniformisation de données hétérogènes vers un espace latent unique.

Les contributions de cette section sont les suivantes :

- Nous proposons une nouvelle modélisation originale des données temporelles et

multivariées sous la forme d'un graphe orienté où chaque sommet correspond à une mesure et où les proximités temporelles sont représentées via des arcs pondérés.

- Nous utilisons cette représentation en entrée d'un GCN que nous évaluons sur le jeu de données réelles MIMIC-III décrit par [Har+19], pour prédire, à partir des données recueillies durant leur parcours, le décès des patients à 24h. Nous montrons que notre approche offre des résultats comparables à ceux de l'état de l'art.
- Nous étudions la représentation latente obtenue au sein de ces modèles pour évaluer la capacité du modèle à apprendre une représentation enrichie par rapport à trois nouveaux cas.

3.5.1 Le dossier patient en tant que graphe temporel

3.5.1.1 Méthode

Dans la continuité de l'objectif 2 mentionné dans le chapitre 1, nous cherchons à prédire l'occurrence d'un évènement médical à partir du dossier patient informatisé, contenant des données temporelles et multivariées. Dans cette section, nous proposons une modélisation du parcours patient sous forme de graphe simple orienté, dans le but d'apprendre un modèle permettant de prédire un futur évènement à partir des enregistrements médicaux d'un patient. En particulier ici, nous considérons un ensemble d'enregistrements ponctuels pour une tâche de classification. La structure de données que nous proposons est constituée de sommets représentant les enregistrements et d'arcs représentant les liens temporels entre ces enregistrements. La modélisation proposée est récapitulée dans la Figure 3.6, à travers ses différentes étapes.

Nous proposons une modélisation des données extraites des dossiers de patients sous forme de graphes. La principale motivation de cette modélisation est d'éviter le découpage temporel des données en amont de l'apprentissage.

Formellement, les enregistrements d'un patient donné sont représentés sous forme d'un graphe $G = (E, P, X)$ où l'ensemble des sommets $E = \{e_1, e_2, \dots, e_n\}$ représente les évènements, $P = \{(e_i, e_j, p)\}, 1 \leq i \leq n, 1 \leq j \leq n, p \in \mathbb{R}^+$ représente l'ensemble des arcs pondérés avec la proximité temporelle p entre deux évènements e_i et e_j , et $X \in \mathbb{R}^{n \times m}$ représente les caractéristiques des évènements avec m le nombre de modalités distinctes mesurées. Notre objectif est de prédire un évènement médical, par exemple le décès à 24 heures. Il s'agit d'un problème de classification binaire, à savoir trouver une correspondance f telle que $f : G \rightarrow Y, Y \in \{0, 1\}$, Y correspondant à l'étiquette binaire.

En particulier, les caractéristiques des sommets $X \in \mathbb{R}^{n \times m}$ sont représentées d'après l'encodage 1 parmi n (*one-hot encoding*). Puisque les données d'origine sont variées, chaque colonne de cette matrice est vue comme un canal indépendant qui sert à décrire une modalité en particulier. Chaque modalité a sa distribution et correspond à un type de mesure observé dans le monde réel. Par exemple, dans la Figure 3.6, $m = 4$ les colonnes de la matrice correspondent dans l'ordre au rythme cardiaque, au volume

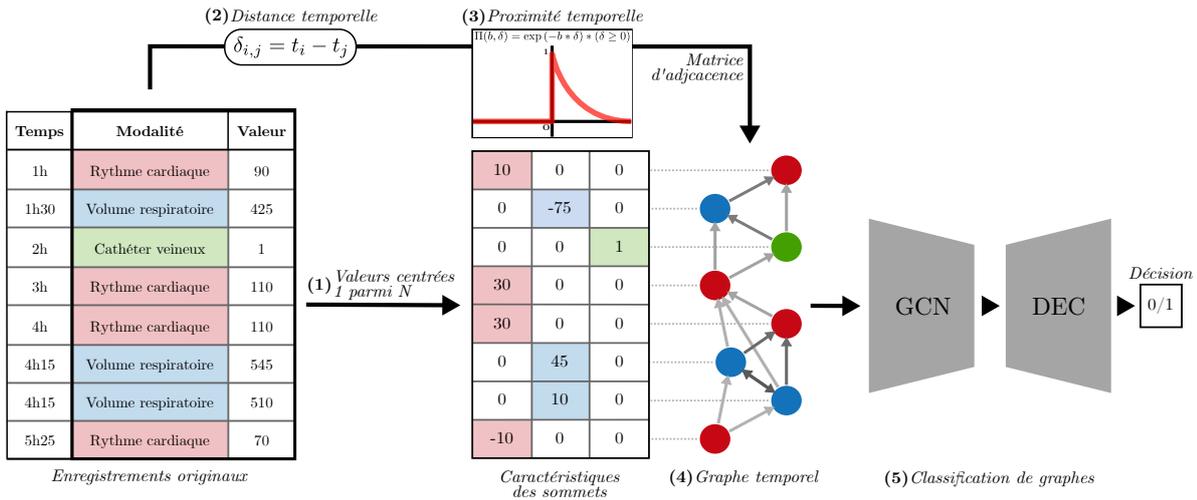


FIGURE 3.6 – Illustration du processus de transformation d'un ensemble d'enregistrement d'un patient en un graphe temporel. Illustration basée sur des données fictives. (1) Transformation des valeurs à travers un encodage 1 parmi n . Les valeurs sont centrées par rapport à la moyenne des valeurs de la modalité correspondante dans la base de données. (2) Distance par paire des coordonnées temporelles. (3) Calcul de la proximité temporelle. (4) Représentation du patient : Caractéristiques des sommets et matrice d'adjacence qui décrivent le graphe. (5) Modèle de classification de graphe.

respiratoire et à la pose d'un cathéter veineux. Chaque modalité est aussi décrite par une unité de mesure implicite. Dans un cas réel, on pose l'hypothèse que tous les enregistrements d'une même modalité sont mesurés avec la même unité et sur la même échelle. Cette problématique rejoint le concept de qualité des données évoqué dans la section 2.2.1. Par exemple, dans la Figure 3.6, le rythme cardiaque est décrit en BPM (battements par minute), le volume respiratoire en millilitres et la pose du cathéter veineux n'a pas d'unité (présence ou non, respectivement 1 ou 0). Cette représentation implique le remplissage d'une seule valeur non nulle par ligne, par vecteur de sommet. Les cellules qui ne donnent pas d'information sont alors assignées à une valeur nulle 0. Nous avons décidé de centrer les valeurs de l'encodage 1 parmi n afin que le 0 qui décrit une valeur nulle corresponde aussi à la moyenne de la modalité. L'opération de centrage est réalisée en soustrayant la valeur donnée par la moyenne de la distribution de la modalité correspondante à l'échelle de toute la base de données. Dans la Figure 3.6, on observe que la première mesure du rythme cardiaque est $10 = 90 - 80$, 80 correspond à la moyenne du rythme cardiaque de tous les patients dans cet exemple.

La particularité de la modélisation réside dans la constitution des arcs. Au sein du modèle utilisé dans nos expérimentations, l'information des sommets est propagée à travers le processus de convolution. L'hypothèse proposée est que les événements proches dans le temps ont des propriétés comparables ce qui correspond à l'hypothèse d'homophilie. Cela se traduit dans le graphe en calculant une proximité temporelle p pour chaque couple (e_i, e_j) à partir de leurs coordonnées temporelles $T = \{t_1, t_2, \dots, t_n\}$ qui agit comme un poids pour les arcs. Ainsi, plus deux événements sont proches dans le temps, plus le poids de l'arc qui les relie est grand. La diminution graduelle du poids lorsque la distance temporelle augmente est proche de travaux provenant des sciences cognitives quant à l'étude de l'oubli. Nous avons comparé plusieurs fonctions d'oubli

par expérimentation et retenu la fonction développée dans [Whi01], notée Π , qui est définie par :

$$\delta_{i,j} = t_i - t_j$$

$$b > 0, \Pi(b, \delta_{i,j}) = \begin{cases} \exp(-b * \delta_{i,j}) & \text{si } \delta_{i,j} \geq 0 \\ 0 & \text{sinon.} \end{cases} \quad (3.1)$$

δ correspond à la distance temporelle calculée pour chaque paire d'évènements. b est un paramètre de modélisation, qui influence l'intensité de la diminution dans la fonction. Les motivations et le principe de cette fonction sont détaillés dans la section suivante.

Les différentes étapes de la modélisation sont récapitulées dans la figure 3.6. La modélisation utilise les données pour produire un graphe de proximité temporelle : d'une part, la valeur centrée et la modalité de chaque mesure sont transformées en une matrice où chaque colonne correspond à une modalité mesurée. (1) Chaque ligne correspond à la représentation vectorielle d'un sommet donné. D'autre part, les coordonnées temporelles des évènements sont transformées en distances ; (2) par paires de sommets. Les distances sont ensuite transformées en proximité par l'application de la fonction Π (voir l'équation 3.1) ; (3) Le résultat est une matrice d'adjacence qui décrit le poids des arcs du graphe final. La matrice d'adjacence et le vecteur de représentation des sommets décrivent alors le graphe du parcours d'un patient ; (4) mis en entrée d'un modèle supervisé pour prédire l'étiquette binaire. L'étiquette est prédite au niveau du graphe, on parle alors de classification de graphe.

3.5.1.2 Fonction d'oubli

La particularité de cette méthode est la fonction d'oubli Π utilisée pour constituer les arcs du graphe temporel. L'objectif de cette fonction est de transformer une distance temporelle $\delta_{i,j} = t_i - t_j$ définie telle que $\delta_{i,j} \in]-\infty; +\infty[$ en une mesure de proximité $\Pi \in [0; 1]$. La valeur 1 correspond à une proximité maximale dans le graphe temporel final, c'est-à-dire l'existence d'un arc tandis que la valeur 0 correspond à une dissimilarité maximale, soit l'absence d'un arc. Les coordonnées et la distance temporelles sont ici exprimées en heures.

La fonction Π , définie dans l'équation 3.1, prend des valeurs non nulles uniquement quand $\delta_{i,j} \geq 0$, c'est-à-dire quand $t_i \geq t_j$. Cela permet de rendre la fonction asymétrique soit $\delta_{i,j} \neq \delta_{j,i}$ si $t_i \neq t_j$. Ainsi, le graphe généré sera « directionnel », c'est-à-dire que les arcs suivront le cours du temps dans une même direction, des enregistrements les plus récents vers les plus anciens. Nous avons choisi cette orientation plutôt que des plus anciens vers les plus récents puisque les enregistrements antérieurs servent de contexte à un enregistrement présent. Chaque enregistrement doit donc être « enrichi » par les enregistrements passés.

Le paramètre $b > 0$ qui contrôle la décroissance de la fonction est inspiré des travaux de [Whi01]. Ce paramètre doit être choisi au moment de l'apprentissage pour la modélisation mise en entrée du modèle d'apprentissage, donc en tant qu'hyperparamètre du processus d'apprentissage complet. Or, la compréhension de ce paramètre est abstraite pour un humain. Pour interpréter cette fonction, nous sommes plutôt intéressés par sa

valeur d'oubli, c'est-à-dire à partir de quelle distance temporelle la fonction prend elle une valeur nulle à la fin de la décroissance. Cependant, la fonction n'a pas de racine sur $\delta_{j,i} > 0$. La valeur de la fonction est calculable pour une valeur très faible ϵ , quasi nulle telle que $\Pi(b, \delta_{i,j}) = \epsilon$. Cette racine approximative \tilde{r} est obtenue par l'analyse de :

$$\begin{aligned} \Pi(b, \delta_{i,j}) &= \epsilon \\ \exp(-b * \delta_{i,j}) * 1 &= \epsilon \\ -b * x &= \ln(\epsilon) \\ x &= \frac{-\ln(\epsilon)}{b} \end{aligned} \tag{3.2}$$

Ainsi, pour les valeurs d'illustration utilisées dans les Figures 3.7 et 3.8, soit $b = 0,6$ et $b = 0,2$, la racine vaut respectivement $\tilde{r} \approx 8h$ et $\tilde{r} \approx 23h$ avec $\epsilon = 0,01$ par exemple.

Une projection des graphes générés par cette modélisation selon un algorithme de force est illustrée dans les Figures 3.7 et 3.8. Nous observons que la valeur de ce paramètre a un effet direct sur la structure des graphes générés. Une valeur basse du paramètre d'oubli b cause une décroissance faible de la fonction f . Ainsi, les poids entre les arcs sont élevés ce qui produit une structure plus compacte. Au contraire, une valeur élevée produit des proximités plus faible et une structure éparse.

3.5.1.3 Expérimentations et résultats

Pour évaluer notre approche, nous nous sommes concentrés sur la tâche de prédiction du décès à 24 heures décrite dans Harutyunyan et al. [Har+19]. Les données originales sont disponibles sur Physionet [Gol+00] dans la base de données clinique MIMIC-III [Joh+16; JPM15]. Cette base de données publique est souvent utilisée dans la littérature pour comparer les approches pour de nombreuses tâches de supervision médicale car elle est volumineuse, hétérogène et anonymisée. Seuls les parcours de patients d'une durée d'au moins 4 heures sont sélectionnés. Pour chacun d'eux, nous cherchons à prédire, pour chaque tranche d'une heure, si un patient décède dans les 24 heures suivantes, à partir des données antérieures à l'observation. Pour leurs expériences, Harutyunyan et al. [Har+19] sélectionnent 17 descripteurs (par exemple, la fréquence cardiaque ou la pression artérielle) selon des critères médicaux. Les valeurs sont filtrées, prétraitées et les données catégorielles ordonnées sont transformées en données continues. Les volumes de données sont décrits dans le Tableau 3.1. Pour obtenir une comparaison équitable avec nos propres expériences, nous nous appuyerons exactement sur ces données transformées, qui sont disponibles publiquement.

Le modèle que nous présentons est choisi après la recherche des hyperparamètres et de la structure optimaux, réalisée grâce à la bibliothèque Optuna [Aki+19]. Les résultats présentés à travers l'application de notre méthode proviennent d'une validation croisée avec 5 plis, découpée dans l'ensemble d'entraînement puis évaluée sur l'ensemble de test. Le modèle appelé "Convolution de Graphes" dans le Tableau 3.2 se compose d'une succession des couches suivantes : Trois couches de GraphSAGE [HYL18] de la bibliothèque DGL (DeepGraph Library [Zhe+21]¹) avec la moyenne comme fonction

1. <https://www.dgl.ai/> consulté le 04/02/24.

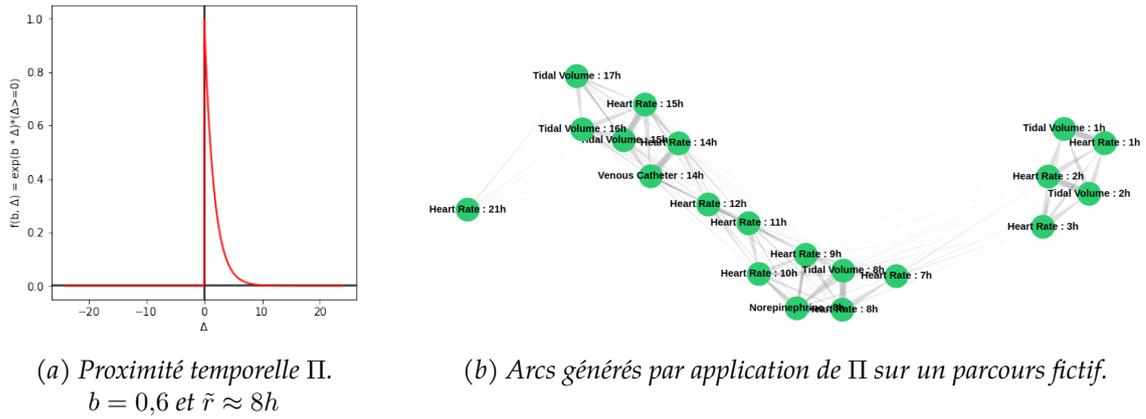


FIGURE 3.7 – Exemple de graphe obtenu avec une fonction de proximité à décroissance rapide qui génère un graphe orienté épars.

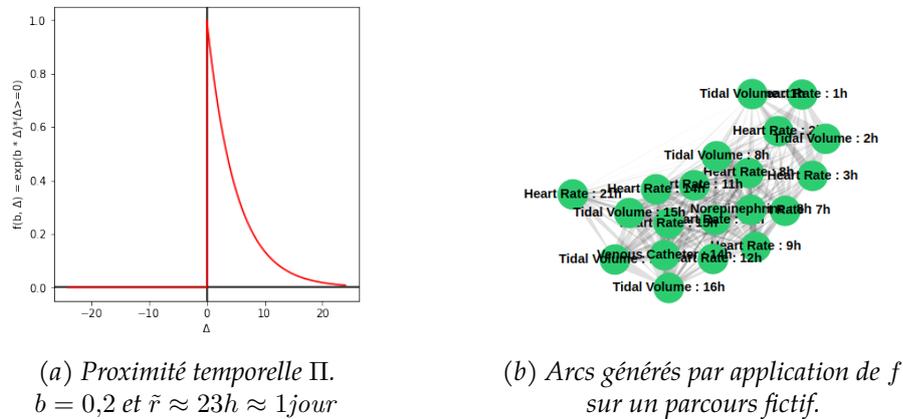


FIGURE 3.8 – Exemple de graphe obtenu avec une fonction de proximité à décroissance lente qui génère un graphe orienté compact.

d'agrégation, puis la concaténation de trois fonctions d'agrégation *min*, *max* et *average* sur la représentation de tous les sommets (*readout_nodes*), deux couches linéaires et une sortie *softmax*. Chaque couche entraînée est suivie d'une fonction d'activation *tanh*, exceptée la dernière. Nous avons choisi l'optimiseur *Adam*. À l'issue de l'optimisation des hyperparamètres, nous avons obtenu les valeurs suivantes : le taux d'apprentissage fixé à 0,003 et la décroissance de poids (*weight decay*) à 0,05. Le paramètre de proximité temporelle b est 4,5 et la largeur des couches à 15. L'apprentissage est calculé sur une machine équipée de 40 cœurs de calcul, 126 Go de RAM et une carte graphique NVIDIA GeForce GTX 1080 avec 11 Go de VRAM.

Pour comparer les approches, nous mesurons l'aire sous la courbe sensibilité/spécificité, appelée AUC (*Area Under ROC Curve* ou *AUC ROC*), sur l'ensemble de test. Ce score est adapté à l'évaluation pour des ensembles de données avec des proportions d'étiquettes déséquilibrées comme notre tâche. Comme le montre le Tableau 3.2, notre modélisation offre des résultats similaires à ceux de l'état de l'art, observés dans Harutyunyan et al. [Har+19], à savoir la régression logistique et les différentes variantes de LSTM. Notre objectif dans ce chapitre est de valider la représentation des données des patients sous forme de graphe. Le résultat obtenu par notre méthode dépasse

TABLEAU 3.1 – Caractéristiques des données extraites de MIMIC-III.

	Quantité
Nombre de patients du jeu d'entraînement	28 620
Nombre de patients du jeu de test	5 058
Pourcentage de patients de la classe minoritaire	11,0 %
Durée moyenne des séjours	86,6h
Durée maximale des séjours	2 103h

TABLEAU 3.2 – Comparaison des résultats sur la prédiction de décès dans les 24h. Les différentes versions du LSTM sont extraites du papier original avec pour significations respectives « S » : Standard, « C » : Channel-wise et « DS » : Deep Supervision.

Modèle	AUC
[Har+19] :	
Régression Logistique	0,870
LSTM - S	0,892
LSTM - C	0,906
LSTM - C + DS	0,911
Notre méthode :	
Convolution de Graphes	0,897

certaines versions de LSTM. De plus, la valeur de l'hyperparamètre obtenue après optimisation est $b = 4,5$ qui a une racine approximative $\tilde{r} \approx 1h$. On observe donc que notre modélisation est comparable aux modèles séquentiels, puisque sur les tâches de prédictions d'évènements médicaux, les LSTM auxquels nous avons comparé notre approche utilisent une fenêtre de découpage de taille 1h.

Une différence clé entre notre modélisation sous forme de graphes temporels et les modèles séquentiels réside dans l'efficacité du pré-traitement des données. Dans les modèles séquentiels, chaque ajustement de la valeur de découpage temporel nécessite la création d'un nouveau jeu de données séquentiel complet, ce qui peut entraîner une complexité et des temps de calcul considérables. De plus, la taille de la fenêtre temporelle influence directement la dimension de la matrice d'entrée du modèle, engendrant des temps d'entraînement et d'inférence variables en fonction du découpage choisi. En revanche, notre approche par graphes temporels simplifie ce processus en nécessitant seulement le calcul d'une matrice de proximité temporelle pour chaque nouvelle valeur du paramètre b . Chaque ajustement de b génère ainsi un graphe de taille fixe, dont la complexité peut être anticipée, permettant une optimisation plus efficace des ressources computationnelles.

On peut se demander dans quelle mesure les prédictions du modèle présenté sont influencées par la propagation de l'information à partir des GNN à travers les arcs comme prévu, grâce à notre modélisation, plutôt que par d'autres effets. Le Tableau

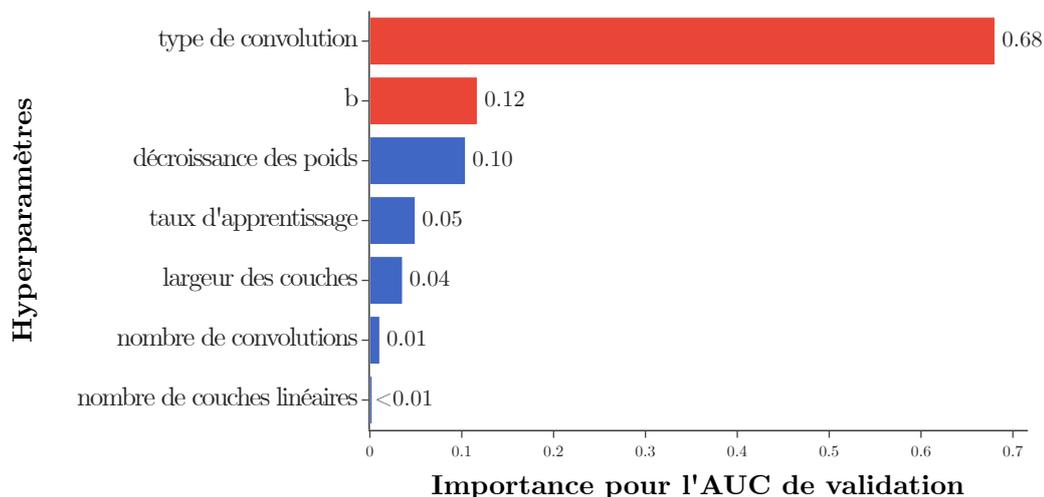


FIGURE 3.9 – Importance des hyperparamètres sur les performances, suite à l’optimisation par Optuna pour la convolution de graphes.

3.2 montre la valeur contributive des différents hyperparamètres lors de l’optimisation par Optuna. Les résultats proviennent d’une analyse fonctionnelle de la variance, dont le nom est raccourci en fANOVA [HHL14]. Nous observons sur ces résultats que les hyperparamètres les plus importants sont le type de convolution et le paramètre b de la modélisation. Nous en déduisons que la constitution des arcs et la manière dont l’information est propagée à l’intérieur du graphe dans le modèle ont un impact fort sur les performances mesurées.

3.5.2 Représentation latente des évènements médicaux

Dans cette section, nous nous intéressons plus particulièrement à la représentation latente apprise au cours de l’apprentissage. Nous souhaitons étudier si cette représentation met en évidence des patients aux caractéristiques communes. L’objectif est d’étudier si le processus mis en place pour une tâche spécifique est utilisable sur de nouvelles tâches, avec des applications comme le sous-typage de patients [Bay+17] ou de maladies [Kan+21], la construction de cohortes [WHM20] ou la définition de modèles prédictifs personnalisés [Hat+19]. En d’autres termes, est-ce que le modèle appris pour prédire des décès à 24h, peut être utilisé pour prédire d’autres évènements médicaux ?

Notre modèle étant basé sur une architecture à base de GCN, nous pouvons extraire directement les informations issues des sorties des convolutions. La Figure 3.10 illustre la représentation extraite d’une couche intermédiaire sur laquelle nous appliquons une Analyse en Composantes Principales (ACP) pour réduire les dimensions et les visualiser. Par la suite, nous sélectionnons 5760 patients pour lesquels nous identifions la présence ou l’absence d’un diagnostic en particulier. Nous considérons quatre diagnostics : décès à 24h, lésions cérébrales anoxiques, anévrisme cérébral non rompu et hypertension (représentés sur la figure de gauche à droite). Pour la prédiction du décès à 24h, nous pouvons constater un regroupement très localisé dans l’espace (points verts) ce que confirme la forte valeur de l’AUC (0,89). De même, les lésions cérébrales

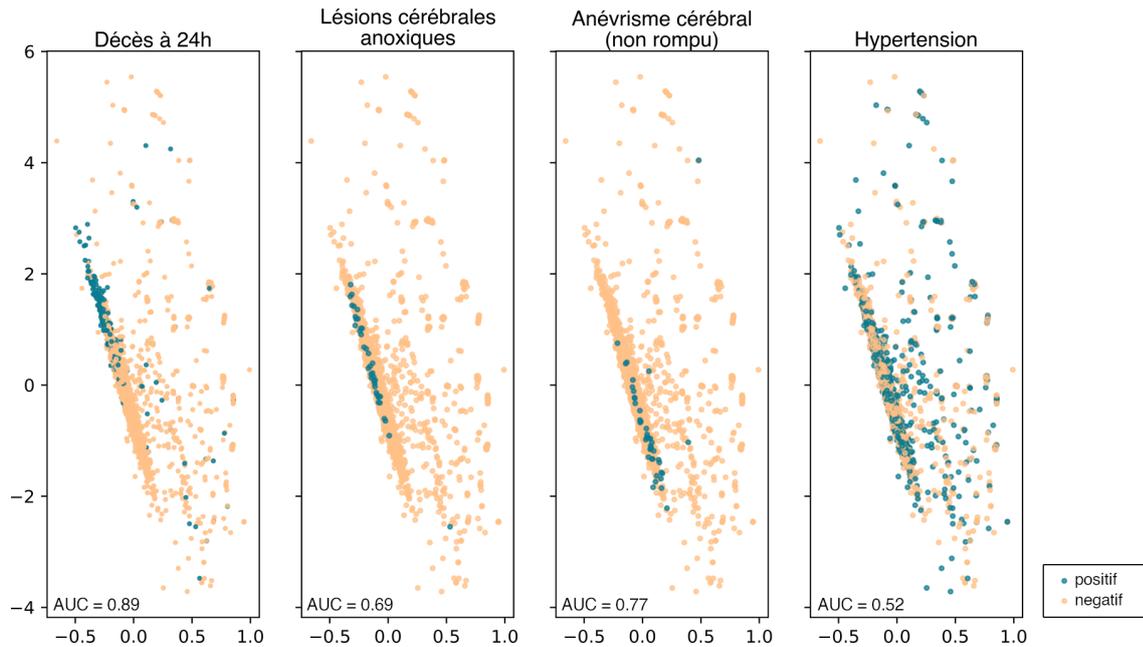


FIGURE 3.10 – Représentation extraite du modèle GCN par rapport à quatre cas cliniques. Chaque parcours est représenté par un point sur la figure dont la couleur correspond à la présence ou l’absence de l’enregistrement d’intérêt du cas étudié. Le plongement en deux dimensions est obtenu par l’application d’une ACP sur les plongements de graphes. Pour chaque cas, un séparateur linéaire est appliqué sur les plongements de graphes afin d’évaluer l’AUC pour le cas donné. Cela estime la séparabilité du problème en question par rapport à cette représentation.

anoxiques et l’anévrisme cérébral sont aussi localisés. Au contraire, l’hypertension ne semble pas posséder de localité dans cette représentation latente.

Les lésions cérébrales anoxiques et l’anévrisme cérébral concernent respectivement la perte d’oxygène dans le cerveau et la dilatation de la paroi d’une artère. La détection de ces diagnostics est fortement liée à la mortalité du patient et la représentation latente apprise reste adaptée. L’hypertension est un diagnostic identifié chez des patients à gravité très variable ce qui peut expliquer l’absence de localité bien définie par rapport à la représentation latente apprise. Ces expérimentations montrent que pour une tâche particulière, la représentation latente est intéressante et met en évidence des caractéristiques communes. En revanche, pour des diagnostics qui ne semblent pas liés au décès, on constate qu’elle n’est pas adaptée.

3.5.3 Limitations

Dans le processus d’apprentissage présenté, la tâche de prédiction consiste à estimer la valeur d’une étiquette finale par rapport au graphe entier. Or, d’après [Dwi+20], les réseaux de neurones appliqués aux graphes, obtiennent des résultats significativement meilleurs que des réseaux classiques pour des tâches comme la prédiction de liens ou la classification de sommets alors que les gains de performance restent limités pour la classification de graphe. D’après l’étude des représentations latentes, on observe que le modèle capture uniquement des caractéristiques en lien avec l’étiquette de supervision.

Cet effet pourrait être accru par l'utilisation d'une couche de synthèse sur l'ensemble du graphe afin de produire la prédiction finale à l'échelle du graphe, ce qui a pour effet d'écraser toutes les caractéristiques structurelles pour la prédiction. Le changement de formulation de l'apprentissage pourrait encourager le modèle à apprendre une représentation enrichie des données et augmenter les performances prédictives.

Les modèles que nous comparons sont des classifieurs binaires, nous avons donc choisi l'AUC ROC parce que cette métrique est une alternative à la précision dans le cas de classes déséquilibrées [YY22]. De plus, dans la littérature, ce critère sert de référence pour l'évaluation de modèles de prédictions d'évènements cliniques, comme les travaux avec lesquels nous comparons nos résultats [Har+19]. D'autres critères complémentaires pourraient être envisagés comme l'aire sous la courbe précision-rappel (*Precision-Recall Curve* ou PRC), le F1-score ou le coefficient de corrélation de Matthews [CJ23].

La conception de notre modélisation a été directement influencée par les lourdes exigences en termes de pré-traitement des données lors de l'utilisation de modèles séquentiels, notamment en ce qui concerne le découpage temporel. Comparer les temps de pré-traitement, d'entraînement et d'inférence entre les deux types de méthodes pourrait introduire un nouveau critère essentiel dans l'évaluation de ces approches. Par exemple, la réduction du temps d'apprentissage d'un modèle est cruciale pour faciliter l'optimisation des hyperparamètres en conditions réelles où la rapidité et l'efficacité sont primordiales. D'après les expériences que nous avons menées, notre méthode, qui n'a pas recours à des pré-traitements, offre un gain de temps important.

3.6 Conclusion et Perspectives

Dans ce chapitre, nous avons proposé un état de l'art des paradigmes et des techniques d'apprentissage de représentations de graphes puis une nouvelle modélisation des données patients sous la forme de graphes temporels. Dans cette modélisation, nous utilisons les arcs pondérés pour représenter la proximité temporelle entre deux évènements. Cette proximité temporelle est déterminée à l'aide d'une fonction d'oubli paramétrable, qui remplit un rôle similaire au découpage des séquences en fenêtres temporelles successives dans l'apprentissage séquentiel. Contrairement à ce découpage, qui est discret et nécessite un pré-traitement distinct pour chaque valeur d'hyperparamètre temporel, notre méthode préserve la continuité des données originales. Elle se contente d'ajuster la matrice d'adjacence pour chaque nouvelle valeur d'hyperparamètre temporel. Nous avons expérimenté cette modélisation en tirant parti des convolutions de graphes et montré que nos résultats approchent ceux de l'état de l'art. Nous proposons donc une représentation alternative aux représentations séquentielles des données pour la prise en compte du temps. Du fait de la modularité offerte par les structures de graphes et des nombreux apports récents dans le domaine des GNN, il est possible d'imaginer divers axes d'améliorations.

La modélisation sous forme de graphes présente l'avantage principal d'intégrer des données hétérogènes issues de sources différentes. Actuellement, les arcs de notre modélisation décrivent uniquement des relations temporelles entre enregistrements.

L'avènement du web sémantique a permis un échange de connaissances médicales interopérables [Haq+22] qui enrichiraient notre modélisation. Par exemple, l'ajout de hiérarchies déjà présentes dans MIMIC-III, comme la Classification Internationale des Maladies² (CIM, ICD en anglais) pour les diagnostics [Wu+24], permettrait au modèle d'associer des diagnostics de la même famille par exemple. Il existe des ontologies variées pour identifier et classer les données cliniques comme la *SNOMED CT*³ pour les concepts généraux du DPI, *RxNorm*⁴ pour les médicaments, *Foundational Model of Anatomy*⁵ pour décrire l'anatomie, etc. Outre des taxonomies hiérarchiques, des bases de données décrivent aussi les interactions entre ces éléments comme par exemple un réseau d'interaction chimiques à propos des protéines, des gènes ou des médicaments [Oug+19]. L'apprentissage sur des graphes avec différents types de liens par des GNN nécessite cependant des méthodes adaptées à ces graphes hétérogènes [Zha+19; Hu+20]. Une autre condition réside dans l'alignement [ABM15] de ces ontologies, c'est-à-dire la capacité à croiser les intitulés des données disponibles, de faible qualité, avec les ontologies disponibles, voire même les ontologies entre elles.

La performance des modèles basés sur des représentations latentes dépend principalement de leur capacité à placer judicieusement les éléments mis en entrée dans un espace donné, afin d'obtenir les distances les plus représentatives entre deux éléments donnés pour la tâche finale. Ce processus est conditionné par de nombreux éléments comme la fonction de coût, l'expressivité du modèle ou les données observées. Cependant, on observe que la plupart des méthodes de GNN s'appuient sur une géométrie euclidienne. Or, des variantes de ces méthodes ont été développées pour incorporer des géométries non-euclidiennes qui présentent différents avantages. Les géométries hyperboliques, sous de nombreuses variantes, présentent une courbure dont un des avantages est une meilleure représentation de structures en arbre comme des hiérarchies sémantiques [NK18]. [Sun+22] modélisent les évolutions temporelles d'un graphe dynamique en variant la courbure d'un espace riemannien.

Dans la modélisation présentée, la distance δ en entrée de la fonction de proximité considère les enregistrements comme instantanés, c'est-à-dire sans durée, puisque elle ne prend en compte que les deux horodatages de début. En plus de leur coordonnée de début, de nombreux enregistrements disponibles dans les DPI ont aussi une coordonnée de fin, comme la prescription de médicaments par exemple. Il faudrait donc adapter le calcul de la distance temporelle à des intervalles plutôt qu'à des coordonnées ponctuelles. L'algèbre des intervalles [All83] identifie 13 relations logiques comme le chevauchement, l'antériorité ou l'inclusion. Cette algèbre qualitative a été étendue dans les *Temporal Constraints Networks* [DMP91; Mei91] pour générer des graphes d'enregistrements avec des contraintes temporelles représentées par des arcs pour effectuer des analyses quantitatives. Cependant, ces modélisations sont dédiées principalement à l'application de méthodes traditionnelles de la théorie des graphes. L'unification avec les GNN ne semble pas triviale. Parmi d'autres incompatibilités, les arcs des réseaux de contraintes sont étiquetés par des intervalles. Un rare exemple

2. CIM version 11, <https://icd.who.int/en> consulté le 05/06/24.

3. <https://www.snomed.org/use-snomed-ct> consulté le 05/06/24.

4. <https://www.nlm.nih.gov/research/umls/rxnorm/index.html> consulté le 05/06/24.

5. <http://si.washington.edu/projects/fma> consulté le 05/06/24.

est [WLG22] qui appliquent un GNN sur le réseau minimal de distances issue de l'application de l'algorithme Floyd Warshall sur le réseau de contraintes temporelles pour de la planification en robotique.

En opposition au contexte purement supervisé que nous avons présenté où l'étiquette de supervision est exogène au graphe, le décès pourrait être représenté comme un sommet inclus directement au sein de celui-ci. Désormais, le processus d'apprentissage consisterait à prédire l'existence du nouveau sommet étiquette ou l'existence d'un arc entre ce sommet et les autres mesures du patient. Ce principe est comparable à l'auto-encodeur masqué [He+21]. Le principal avantage de cette approche en auto-supervision serait que le modèle n'aurait besoin que d'informations intrinsèques pour l'apprentissage. De plus, tous les sommets du graphe pourraient être considérés comme étiquettes d'intérêt, indépendamment des tâches à résoudre. Cela aurait pour effet de mutualiser le coût de l'apprentissage à travers des utilisations finales variées grâce à une représentation latente riche. L'auto-supervision sur les graphes a déjà été expérimentée sous des formes variées comme [Ron+20; FEK21; Sun+22].

CONCLUSION

Sommaire

4.1	Résumé des contributions	72
4.1.1	Visualisation de données cliniques	72
4.1.2	Apprentissage de représentations de parcours patients	72
4.2	Perspectives	73
4.2.1	Visualisation de parcours cliniques	73
4.2.2	Apprentissage de représentations de parcours cliniques	75
4.2.3	Visualisation et apprentissage conjoints	79

Dans ce mémoire, nous avons réalisé deux contributions principales pour la tâche de prédiction et la visualisation d'évènements cliniques, à l'intersection de trois domaines de recherche très actifs : l'informatique médicale, l'apprentissage automatique et la visualisation de données. Dans ce chapitre, nous résumons les contributions dans la section 4.1 et donnons des perspectives dans la section 4.2.

4.1 Résumé des contributions

Dans ce mémoire, nous avons présenté des contributions selon deux axes distincts et complémentaires pour la mise en place de systèmes d'aide à la décision au sein des services de soins : la représentation visuelle de données temporelles hétérogènes et l'apprentissage automatique de modèles de prédiction d'évènements cliniques.

4.1.1 Visualisation de données cliniques

Nous avons réalisé une analyse approfondie des visualisations de séries temporelles multivariées d'un seul individu dans 21 interfaces, en utilisant le protocole PRISMA-ScR. Cette analyse met en lumière les fonctionnalités et interactions communes à ces interfaces en les comparant dans le domaine médical et dans d'autres contextes.

Nous avons conçu GLIMPSE-Med, une interface visuelle dédiée aux données du DPI, destinée tant aux professionnels de la santé qu'aux ingénieurs et statisticiens. Ses principales innovations incluent la compression visuelle et adaptative de l'axe temporel, la représentation uniformisée de séries temporelles et de mesures discrètes/continues avec une navigation hiérarchique et l'intégration des prédictions de modèles évoluant dans le temps ainsi que des facteurs contributifs de la décision.

Une évaluation expérimentale de GLIMPSE-Med auprès de 14 professionnels du CHU de Montpellier a fourni des résultats quantitatifs et qualitatifs.

4.1.2 Apprentissage de représentations de parcours patients

Nous avons proposé une modélisation des parcours patients sous forme de graphes temporels, où chaque sommet représente un évènement et où les relations temporelles sont modélisées par des arcs pondérés. Cette approche a été évaluée avec des méthodes d'apprentissage par convolution de graphes pour prédire le décès dans les 24 heures à partir des données MIMIC-III, avec des résultats comparables à ceux de l'état de l'art.

Nous avons également évalué la qualité de l'espace latent généré par notre modélisation en examinant sa capacité à apprendre des représentations pertinentes pour de nouvelles tâches. La séparabilité de cette représentation a été testée sur trois nouveaux cas cliniques : lésion cérébrale anoxique, anévrisme cérébral et hypertension. Cette évaluation a mis en évidence les limites de la richesse de la représentation latente dans le cadre de la classification à l'échelle du graphe.

4.2 Perspectives

Nous avons déjà abordé les perspectives spécifiques à chaque domaine dans la section 2.3 pour la visualisation et dans la section 3.6 pour l'apprentissage. Cependant, comme le souligne le chapitre 1, le développement de ces deux types de méthodes se renforce mutuellement. Nous proposons de nouveaux axes de développement à moyen et long terme pour chaque méthode dans les sections 4.2.1 et 4.2.2, tout en explorant leur amélioration conjointe dans la section 4.2.3.

4.2.1 Visualisation de parcours cliniques

Multimodalité

La multimodalité fait référence à la diversité des types de données intégrés dans un système. Dans le cadre de GLIMPSE-Med, nous avons déjà abordé l'intégration de différentes modalités telles que les données catégorielles et numériques, avec une temporalité discrète ou continue. Le DPI contient deux autres modalités : les notes cliniques et l'imagerie médicale. Pour maintenir la cohérence visuelle et bénéficier des deux axes de compression de notre interface, l'ajout de ces modalités doit s'adapter à la présentation sous forme de séries temporelles parallèles. De plus, puisque notre interface repose sur la compression visuelle des enregistrements tout en permettant à l'utilisateur de les voir entièrement à la demande, il est nécessaire de trouver une représentation compressée et une vue détaillée pour chaque modalité. Parmi les approches identifiées dans le chapitre 2, seuls [BAK07; Wu+23] intègrent le texte et l'image dans leurs interfaces.

Une **approche simpliste** consisterait à traiter les images et documents affichés dans l'aperçu comme de simples enregistrements, sans prendre en compte leur contenu. Les collections de documents ou d'images seraient alors représentées sous forme de séries temporelles catégorielles, comme illustré dans la Figure 2.5.a. L'utilisation de symboles adaptés, tels qu'une icône de courrier ou d'appareil d'imagerie, permettrait de mettre en valeur la nature de ces enregistrements. Un clic sur un symbole pourrait révéler le contenu brut du texte ou de l'imagerie dans une sous-fenêtre. Bien que cette solution s'éloigne de notre principe d'affichage dans une fenêtre unique, elle reste intuitive. Toutefois, bien qu'elle soit facile à mettre en œuvre, cette approche fournit peu d'informations à l'utilisateur. Il est donc essentiel que les représentations compactes ou détaillées offrent un maximum d'informations, afin d'éviter à l'utilisateur de devoir lire ou visionner l'intégralité du contenu pour trouver l'information recherchée.

L'utilisation de techniques d'**extraction d'informations**, notamment pour extraire des données catégorielles et numériques à partir des textes et des images, pourrait permettre de générer des représentations plus riches et informatives. Ces nouvelles données pourraient être visualisées sous forme de séries temporelles catégorielles ou numériques sans nécessiter de modification majeure de notre interface. Dans le cas d'un affichage détaillé via une sous-fenêtre, la représentation compacte de ces informations dans un espace visuel restreint pourrait être obtenue par des techniques de **résumé** ou de **synthèse**. Cependant, l'extraction d'informations et la génération automatique de résumés pour les textes ou les images restent des défis complexes qui nécessitent des

recherches approfondies. Les paragraphes suivants explorent des pistes pour aborder ces enjeux.

Les **textes** médicaux sont généralement structurés en paragraphes introduits par des titres, lesquels peuvent être exploités pour définir des règles métier ou orienter l'extraction d'informations pertinentes. Toutefois, cette structure varie d'un établissement à l'autre et évolue au fil du temps. Pour surmonter ces variations, les méthodes automatiques se concentrent principalement sur la reconnaissance d'entités nommées et l'extraction d'informations, comme le montrent les revues de la littérature proposées par [Wan+18; PDE20; NJM21]. Un texte peut également être résumé à travers des mots-clés représentatifs des thèmes abordés. Cette approche peut être mise en œuvre via l'extraction d'entités nommées, mentionnée précédemment, ou par des modélisations thématiques (*topic modeling*) [VK20]. Par exemple, [Wu+23] présentent l'évolution des commentaires d'un flux vidéo en ligne en affichant des mots-clés représentatifs et en projetant l'évolution sémantique issue du modèle de langue BERT [Dev+19] visualisée sous forme de couleurs.

Dans le domaine de l'**imagerie**, l'extraction d'informations est plus complexe que celle du texte. Les modèles sont souvent conçus pour un type spécifique d'imagerie, comme l'IRM, le scanner CT, le TEP scan ou les rayons X [JK20]. Une méthode couramment utilisée est la segmentation, qui consiste à classer chaque pixel d'une image afin de détecter des formes caractéristiques [Qur+23]. Cette technique permet d'isoler les sous-régions les plus pertinentes de l'image offrant ainsi un aperçu informatif qui pourrait être utile dans le cadre de notre visualisation.

Patients et séjours multiples

L'interface que nous avons développée est spécifiquement conçue pour faciliter la compréhension des trajectoires au sein du DPI et évaluer la gravité de l'état d'un patient. La visualisation d'un seul parcours à la fois répond efficacement à ces problématiques. Toutefois, de nouveaux usages, tels que la création de cohortes ou la comparaison de parcours patients, nécessitent la capacité de visualiser plusieurs trajectoires simultanément. Comme discuté dans le chapitre 2, cette question, est abordée par [Sch+22]. Les auteurs représentent visuellement des corrélations, des distributions et des flux. Bien que cette approche puisse sembler incompatible avec les fonctionnalités développées pour des données temporelles et hétérogènes non-agrégées, elle pourrait néanmoins enrichir l'interface via ces ajouts :

1. **Une méta-vue pour la navigation entre patients** : une méta-vue permettant de naviguer entre les patients pourrait être intégrée en s'inspirant des interfaces existantes dédiées à l'étude de cohortes, comme celles proposées dans [GW12; Mag+19]. Dans ce cadre, l'interaction de l'utilisateur avec un flux ou une distribution déclenche l'ouverture de la vue individuelle d'un patient sur GLIMPSE-Med. Les éléments caractéristiques identifiés dans la méta-vue de la cohorte sont ensuite mis en valeur dans la vue individuelle.
2. **L'intégration des propriétés de la cohorte** : les propriétés de la cohorte à laquelle le patient appartient peuvent être affichées directement dans la vue individuelle, afin de situer les données non-agrégées dans un contexte plus large de pratiques

et d'observations cliniques générales. Cette technique correspond aux approches classées comme *single-to-multiple* et *single-to-cohort* dans [Sch+22].

Des contextes comme la recherche, l'évaluation de pratiques cliniques ou l'enseignement auprès de futurs cliniciens pourraient bénéficier de la comparaison de séjours. Actuellement, notre vue affiche plusieurs séjours de manière chronologique, en les concaténant horizontalement le long de l'axe temporel avec une temporalité absolue. Cependant, l'étude de multiples séjours s'inscrit également dans la problématique de la représentation des cohortes évoquée précédemment. Les fonctionnalités visuelles que nous avons développées pourraient ainsi être adaptées pour comparer un nombre limité de séjours entre eux, qu'il s'agisse de séjours d'un même patient ou de patients différents. Pour permettre cette comparaison, il est nécessaire d'aligner les séjours par rapport à un référentiel commun, en utilisant une temporalité relative. Dans un contexte général, le début du séjour, c'est-à-dire l'admission dans le service, constitue un bon référentiel. Dans un cadre d'étude spécifique, l'alignement pourrait se faire par rapport à un événement d'intérêt, tel qu'un arrêt cardiaque, une admission en soins intensifs, ou un acte chirurgical, afin de comparer les causes et les conséquences de cet événement sur les différents séjours. L'affichage de plusieurs parcours qui partagent le même axe temporel relatif peut alors se faire par deux moyens :

1. **Juxtaposition** : chaque séjour est affiché dans une sous-fenêtre indépendante [Wan+09; BBL15]. Cette méthode permet une comparaison visuelle claire entre les séjours tout en préservant leur individualité.
2. **Superposition** : les données de tous les parcours sont présentées dans une vue unique [Wil+18; Pol+18] avec une mise en évidence des données du patient ciblé. Une différenciation par couleur pour chaque patient ou pour distinguer un patient de la cohorte semble adaptée à cette méthode, bien que cela puisse entrer en conflit avec les palettes de couleurs déjà utilisées dans notre vue.

4.2.2 Apprentissage de représentations de parcours cliniques

Hypergraphe hospitalier

Notre modélisation du parcours patient sous forme de graphes engendre une composante connexe distincte par patient. Les implémentations de la convolution de graphes reposent sur des matrices d'adjacence. Notre méthode permet donc d'apprendre un modèle à base de graphes à partir d'un ensemble volumineux de patients en isolant des sous-graphes indépendants pour chacun d'eux. Cela produit des matrices d'adjacence dont la taille est, au maximum, le carré du nombre d'enregistrements d'un patient.

L'objectif de nos modèles est d'utiliser les connaissances de patients déjà observés pour prédire les événements cliniques chez d'autres patients. Ce partage pourrait s'illustrer directement dans les graphes que nous générons selon des relations issues de bases sémantiques ou de la structure hospitalière. Cependant, la création d'une composante connexe unique, par exemple pour un service ou un établissement, peut entraîner des problèmes pour le passage à l'échelle. La taille de la matrice d'adjacence, utilisée dans la convolution, devient trop grande et obstacle nécessite l'usage de stratégies

d'échantillonnage. De plus, l'apprentissage des poids du modèle qui sont communs à toutes les observations, servent déjà à généraliser à de futurs individus.

Cependant, comme décrit dans la section 3.4, on observe dans la littérature de nombreux modèles [Cho+16; Liu+20; Li+20a; Gu+22; Cho+22; Bol+24] qui construisent des graphes de patients où chaque patient ou séjour est représenté par un unique sommet. Ces modèles cherchent à déterminer la meilleure distance entre les patients ou à utiliser des relations sémantiques pour classer les patients. Une approche intéressante serait de combiner ces méthodes avec notre modélisation en développant une structure de graphes à deux niveaux : un premier niveau à l'échelle de l'hôpital où chaque sommet représente un patient, et un second niveau où chaque sommet correspond au détail du parcours de ce patient sous forme de graphe temporel grâce à notre modélisation.

Extensions de la fonction d'oubli

Nouvelles fonctions Nous avons observé l'importance de la fonction d'oubli dans le processus d'apprentissage. La fonction que nous avons expérimentée est strictement décroissante et asymétrique, avec des valeurs positives quand la distance temporelle est positive et avec une valeur maximale quand la distance est nulle. Cela engendre des relations asymétriques et unidirectionnelles dans le graphe. Il pourrait être intéressant d'expérimenter des fonctions d'oublis avec des propriétés différentes comme d'autres allures décroissantes, un pic à retardement ou des valeurs positives pour les distances positives et négatives. Le défaut principal de la fonction d'oubli actuel est son interprétabilité. En effet, le recours à une racine approximative pour estimer la densité de la fonction engendrée par un paramètre b , sans relation linéaire entre ces deux éléments est difficile à appréhender. L'usage du modèle serait simplifié si la racine de la fonction était contrôlée par un paramètre plus intuitif que celui actuel.

Vers l'attention La fonction d'oubli permet de pondérer la convolution effectuée au sein du modèle en fonction du temps, via une fonction paramétrée manuellement par l'utilisateur. Cependant, une approche plus récente consiste à réaliser des convolutions dont les poids sont déterminés par descente de gradient plutôt qu'à travers les paramètres de notre fonction. Dans le domaine de l'apprentissage profond, l'introduction des mécanismes d'attention [BCB16] a conduit au développement de méthodes auto-attentives ou transformeurs (*transformers*) [Vas+17], qui ont ensuite été adaptés aux GNNs [Zhu+19; BAY22; NNP22]. Appliqués aux graphes, ces couches d'attention attribuent un poids à chaque voisin d'un sommet, en suivant les arêtes du graphe, afin de réaliser une moyenne pondérée du voisinage. Cette moyenne pondérée remplace alors l'agrégation uniforme comme la moyenne utilisée dans GraphSAGE [HYL18] ou GCN [KW17]. Le poids est issu d'une fonction de score d'importance calculée à partir des caractéristiques du sommet et de son voisin.

Attention temporelle En utilisant des approches existantes comme GAT [Vel+18] ou GATv2 [BAY22] sur notre modélisation, il n'est pas garanti que les performances prédictives soient meilleures que celles du modèle que nous avons expérimenté. En effet, l'attention dans ces modèles applique un masque sur les arêtes, c'est-à-dire sur la

structure existante du graphe. Or, notre structure elle-même est déjà un masque parce qu'elle est issue de l'application de la fonction d'oubli sur les relations temporelles entre tous les sommets. Ainsi, notre modélisation peut alors être considérée comme une forme d'attention appliquée sur le graphe complet et qui considère les coordonnées de chaque paire de sommet pour le calcul du score d'importance. Il serait donc pertinent de fusionner notre approche avec le mécanisme d'attention pour pondérer les relations temporelles entre sommets en fonction de leurs caractéristiques et de leurs coordonnées temporelles. Cette nouvelle forme d'attention temporelle devrait alors être appliquée à l'ensemble des paires de sommets, c'est-à-dire sur le graphe complet, afin de générer une structure temporelle optimisée.

Cette nouvelle méthode a trois avantages principaux :

1. La structure temporelle est calculée au sein de la couche de convolution, ce qui évite le recours à une phase de pré-traitements en amont de l'apprentissage. La modélisation est alors intégrée directement dans l'architecture convolutionnelle.
2. La recherche des paramètres optimaux de la fonction d'oubli est déléguée à l'apprentissage du modèle par descente de gradient. Cela restreint le nombre d'hyperparamètres de la méthode.
3. Les approches à base d'attention bénéficient de l'utilisation de mécanismes multi-têtes (*multi-head attention*) : au lieu d'une seule couche d'attention, plusieurs couches indépendantes sont combinées en parallèle pour obtenir une représentation plus riche. Appliqué à notre contexte, il serait possible d'accumuler plusieurs couches d'attention temporelle pour exploiter simultanément différentes relations temporelles. La forme de la fonction d'oubli pourrait d'ailleurs varier d'une tête à l'autre, selon les nouvelles fonctions décrites ci-dessus. Le modèle serait alors en capacité d'exploiter des relations temporelles différentes entre les enregistrements à la manière des réseaux récurrents bidirectionnels comme le BiLSTM [Cor+16] par exemple.

Cette méthode s'éloigne du cadre classique des graphes, car le graphe représentant l'individu en entrée se réduit à un simple ensemble de sommets. En effet, la structure du graphe complet n'est plus intrinsèquement liée aux spécificités de l'individu. Par conséquent, cette formulation pourrait être exprimée en dehors du vocabulaire des graphes et intégrée dans les approches basées sur les transformeurs auto-attentifs. [Shi+22], par exemple, proposent un transformeur multimodal qui considère chaque enregistrement comme un token, à la manière de l'apprentissage sur les textes, pour la prédiction de multiples événements cliniques. Comme dans notre méthode, les auteurs soulignent la capacité de leur modèle à traiter un parcours de manière parallèle sur toute sa durée plutôt que d'avoir recours à un découpage en pas de temps pour les modèles séquentiels. Cette approche récente est prometteuse pour l'apprentissage sur des enregistrements temporels et hétérogènes et pourrait être améliorée par nos travaux sur la fonction d'oubli pour une modélisation temporelle plus riche et explicite.

Apprentissage fédéré

Les modèles d'apprentissage automatique reposent sur l'observation de données existantes et visent à extraire des motifs au sein de la distribution observée. Actuellement, chaque hôpital entraîne son propre modèle sur ses données internes, ce qui limite la capacité de ces modèles à reconnaître certains motifs comme des pathologies rares qui ne sont pas présentes, ou en trop faible quantité, dans les données locales. Parallèlement, les biais inhérents aux pratiques cliniques, aux soins ou aux contextes socio-démographiques spécifiques à un établissement sont susceptibles d'être reproduits par ces modèles. Pour surmonter ces limitations et développer des modèles plus généralistes capables de détecter une plus grande diversité de cas cliniques, il est essentiel de mutualiser l'apprentissage des modèles à une échelle plus large, qu'elle soit régionale, nationale, voire internationale.

Le partage des données brutes entre plusieurs établissements présente un risque évident pour la sécurité et la confidentialité des données des patients. Au delà de l'anonymisation des textes et de l'offuscation d'identifiants démographiques qui sont des techniques nécessaires, la chronologie des séjours hospitaliers peut encore permettre d'identifier un individu. Pour contrer cet effet, notre modélisation convertit les dates en une temporalité relative, propre à chaque patient.

Une alternative à l'apprentissage centralisé traditionnel, réalisé en un seul lieu avec un accès complet à l'ensemble des données, est l'apprentissage fédéré. Dans ce paradigme, les modèles sont entraînés de manière distribuée et sécurisée. Les participants construisent ensemble un modèle commun sans jamais divulguer leurs données locales, en échangeant uniquement des paramètres chiffrés [Zha+21]. Ce nouveau domaine pose alors de nouvelles questions de recherche comme la sécurisation des modèles, l'aboutissement à un consensus ou l'instauration de récompenses pour la participation à l'entraînement [Mam21; Mot+21]. Des travaux récents mettent en place des approches fédérées pour des applications cliniques [Xu+21]. Par exemple, l'étude menée par [Day+21] a réuni 20 instituts pour développer un modèle prédictif du besoin en oxygène chez les patients atteints de COVID-19.

Complexité

L'objectif principal dans la prédiction automatique d'évènements cliniques est de développer des modèles aussi fiables que possible afin de maximiser les bénéfices sociaux et économiques, tels que sauver des vies ou détecter précocement des maladies graves. Cependant, l'entraînement de modèles de plus en plus complexes requiert des ressources computationnelles de plus en plus élevées. Par exemple, l'entraînement de modèles est désormais réalisée sur des baies de serveurs équipés de matériel dédié. La croissance de l'énergie et du matériel nécessaires augmentent l'impact négatif environnemental, social et économique [SGM19]. [Lig+22] montrent le manque de travaux qui étudient l'impact de l'apprentissage profond sur l'environnement. En revanche, on sait que pour maximiser l'impact positif de cette balance il faut donc produire des modèles les plus légers et performants possibles. La performance prédictive des modèles est directement dépendante du volume des données d'apprentissage. Dans le cas des GNN, la plupart des graphes issus du monde réel sont vastes et contiennent

des millions de sommets et d'arêtes. Pour les cas cliniques, les hôpitaux accueillent quotidiennement des centaines de patients et recueillent des milliers de mesures. Les méthodes d'apprentissage de représentations doivent alors être capables de les prendre en considération [GF18]. Au delà des seules considérations théoriques, l'étude des limites computationnelles des modèles et de leurs implémentations assure l'inclusion du plus grand volume de données possible tout en réduisant l'impact négatif de ces calculs.

Les paradigmes théoriques décrits dans le chapitre 3 comme la transmission de messages [Gil+17] ou l'encodeur-décodeur [Cha+22] impactent aussi les implémentations des modèles utilisés en pratique. Par exemple, les deux bibliothèques les plus populaires pour les GNN, *Pytorch Geometric* (PyG) [FL19] et *Deep Graph Library* (DGL) [Wan+19b], sont toutes les deux conçues suivant le paradigme de transmission de messages. L'avantage principal de ce paradigme, utilisé pour nos expérimentations, est la capacité à apprendre des modèles sur des données volumineuses par l'exploitation de structures creuses et par échantillonnage. Cependant, l'implémentation des architectures nécessite un code complexe et très différent du contenu des articles dans lesquels ils sont décrits. L'apparition de nouveaux algorithmes de multiplication de matrices creuses (*sparse matrix*) [YBO18] permet l'adaptation de l'algèbre linéaire sur processeur graphique avec des matrices creuses pour l'implémentation de réseaux de neurones à base de graphes. Cette approche permet d'utiliser des syntaxes qui imitent l'algèbre linéaire comme celui écrit dans les articles et ainsi rendre les implémentations plus accessibles. Une matrice creuse est un format de stockage des valeurs non nulles d'une matrice sous forme d'indices qui permet d'obtenir un stockage et des opérations plus efficaces si cette matrice contient beaucoup de valeurs nulles. Ainsi, cela permet d'étendre l'efficacité sur les structures creuses aux représentations des nœuds. Typiquement, l'encodage de données catégorielles, souvent réalisé avec la technique 1 parmi n (*one-hot*), est majoritairement vide. Cette nouvelle approche est récente, prometteuse et doit être expérimentée et comparée à la transmission de messages. Certains outils existants doivent être adaptés comme l'échantillonnage qui est à la base de méthodes comme GraphSAGE [HYL18].

En plus de l'architecture et de l'implémentation qui influencent la quantité de ressources computationnelles requises, [SGM19] montrent que le processus le plus coûteux est de loin la recherche des hyperparamètres optimaux. Il faut développer des méthodes avec un espace d'hyperparamètres réduit. Par exemple, notre modélisation repose sur un paramètre unique qui est la pente de la fonction d'oubli.

4.2.3 Visualisation et apprentissage conjoints

La boucle d'enrichissement entre visualisation des données et entraînements de modèles décrite dans le chapitre 1 s'applique à moyen ou long terme. L'objectif est d'affiner progressivement l'architecture des modèles et la représentation visuelle des enregistrements et des prédictions suite à une meilleure compréhension des données. Cependant, l'enrichissement mutuel entre la visualisation et le modèle pourrait être accéléré et réalisé sur une échelle de temps plus courte par le biais d'un système interactif plus complet. Cette section décrit un ensemble de fonctionnalités qui pourraient

permettre à l'utilisateur de mieux comprendre les données patients et la décision des modèles par de nouvelles interactions et représentations visuelles.

Modèles explicatifs

Outre l'amélioration des performances de prédiction, la valeur explicative d'un modèle est un aspect crucial pour son adoption par les professionnels de santé. Notre objectif est de développer des extensions qui apportent des dimensions complémentaires, telles que l'explicabilité ou la découverte de nouvelles connaissances, au-delà des seules performances. Par exemple, nous cherchons à comprendre comment un ensemble de maladies et d'interventions interagissent pour aboutir à un événement clinique spécifique. Les méthodes d'intelligence artificielle explicatives ou *explainable artificial intelligence* (XAI) se distinguent en deux catégories [Spe22] :

1. Les méthodes *post-hoc* qui génèrent des explications à partir des prédictions de modèles déjà entraînés.
2. L'explication *ante-hoc* qui utilise directement les éléments du modèle.

Dans notre cas, l'usage de graphe est un avantage significatif puisque cette structure de données est à la fois compréhensible par les utilisateurs et facilement visualisable en deux dimensions [HMM00]. Le développement d'une attention temporelle, évoquée dans la section 4.2.2, pourrait mettre en évidence des relations *ante-hoc* qui pourraient être exploitées par les ingénieurs de données pour comprendre et améliorer les modèles et par le personnel soignant pour saisir les raisons sous-jacentes des décisions prises par le modèle. De plus, on observe l'apparition de méthodes complémentaires *post-hoc* qui extraient des informations des GNN dans le but d'étayer les prédictions. Par exemple, [Yin+19] simplifient le graphe d'entrée, en sélectionnant les sommets et arcs qui contribuent le plus à la prédiction. Toutes ces connaissances pourraient être mises en valeur dans la visualisation.

Carte d'individus

Dans le cadre de l'extension de notre interface, nous avons déjà suggéré dans la section 2.3 l'ajout d'une méta-vue qui permettrait de naviguer entre les patients d'un même service par exemple. D'une autre manière, l'ajout d'un encart visuel qui montre la représentation latente du patient parmi tous les autres pourrait répondre à deux enjeux complémentaires :

1. Nos travaux ont principalement porté sur l'apprentissage d'un modèle et l'intégration visuelle d'un score pour la prédiction d'évènements cliniques, que ce soit sous la forme d'une tâche de classification ou de régression. Cependant, la représentation latente intermédiaire des parcours cliniques pourrait offrir une valeur ajoutée significative pour la navigation entre différents parcours. En effet, de nombreuses approches dans la littérature se focalisent sur l'élaboration de mesures de similarité entre patients [Par+18; Jia+20; Gu+22]. Ces représentations latentes pourraient ainsi servir de base pour améliorer l'exploration et la comparaison des parcours cliniques, impliquant une meilleure compréhension, un diagnostic plus précis et l'anticipation des risques graves.

2. L'apprentissage actif vise à optimiser les performances prédictives tout en réduisant le besoin d'annotations manuelles [Ren+21], ce qui est particulièrement crucial pour les cas cliniques rares ou non encore annotés. Annoter des données hétérogènes est une tâche complexe, et GLIMPSE-Med pourrait jouer un rôle clé dans cet étiquetage manuel. Outre la compréhension individuelle d'un parcours pour produire une étiquette fiable, plusieurs systèmes visuels ont été proposés pour surveiller le bon déroulement de l'apprentissage comme [IHG13; Wan+21]. Cependant, afin d'évaluer la qualité globale de la représentation obtenue, la visualisation doit représenter un ensemble d'individus plutôt qu'un seul ce qui constitue un défi différent de ceux abordés dans nos travaux.

L'encart peut prendre la forme d'une carte centrée autour du patient visualisé et qui montre la projection de tous les individus en deux dimensions. À partir de notre analyse de la littérature réalisée dans le chapitre 2, nous avons identifié 5 exemples d'approches qui intègrent de telles vues. Par exemple, *BitExTract* [Yue+19] montre un réseau d'interactions entre les marchés de cryptomonnaies et *LiveRetro* affiche l'enchaînement des segments d'une diffusion vidéo les plus récents pour montrer l'évolution de la performance du diffuseur. Dans notre cas, l'espace latent issu du modèle de prédiction de d'évènements pourrait être utilisé visuellement pour trouver des patients avec des caractéristiques proches, constituer des cohortes et évaluer les frontières de décision. Cela permettrait aux concepteurs de modèles d'évaluer la représentation latente obtenue et aux professionnels de santé d'explorer des patients avec des caractéristiques proches et de guider le processus d'annotation.

Explications contrefactuelles

Pour mieux comprendre un modèle et ses prédictions, on peut aussi se demander ce qu'il se passe si on change les données d'un patient existant. Par exemple, on pourrait se demander si l'état du patient pourrait être amélioré en l'absence d'une intervention chirurgicale ou d'une prescription de médicament. L'usage de tels scénarios fictifs, qui se basent sur la modification de quelques aspects de parcours existants, rejoint le domaine de l'explication contrefactuelle. Adapté sous forme de visualisation, cela peut prendre deux formes :

1. L'interface permet à l'utilisateur d'ajouter, de supprimer ou de modifier des points de données avec une mise à jour instantanée du score en fonction des modifications apportées. Cette fonctionnalité permet à l'utilisateur de tester des hypothèses spécifiques, facilitant ainsi l'exploration de scénarios alternatifs qui peuvent éclairer ses décisions. Par exemple, l'approche décrite dans [Wex+20] illustre une implémentation pratique de ces simulations interactives.
2. Dans le contexte d'une classification binaire, le système identifie automatiquement les enregistrements dont la modification pourrait entraîner un changement de la prédiction d'une classe à l'autre. Il vise à déterminer l'ensemble minimal de modifications nécessaires pour altérer la décision de classification. Ce processus d'optimisation explore un vaste espace d'hypothèses pour isoler les éléments les plus critiques, comme le décrit [Gui22].

La prise en compte de l'hétérogénéité et du caractère temporel des données du DPI, ainsi que la compréhension des besoins et pratiques des professionnels de santé, représentent des défis majeurs qui incitent à développer des systèmes d'aide à la décision complexes. Les contributions et perspectives que nous avons présentées encouragent une approche conjointe entre modèles d'apprentissage et visualisations de données dans un cadre clinique. L'intégration visuelle des prédictions, le développement de modèles auto-supervisés pour prédire des événements cliniques, ou encore la création d'interfaces interactives entre utilisateurs et modèles, sont autant de domaines qui pourraient transformer les pratiques cliniques.

BIBLIOGRAPHIE

- [ABM15] Fatima ARDJANI, Djelloul BOUCHIHA, and Mimoun MALKI. "Ontology-alignment techniques: survey and analysis". In: *International Journal of Modern Education and Computer Science* 7.11 (2015), p. 67 (cit. on p. 69).
- [Ahm+22] David AHMEDT-ARISTIZABAL, Mohammad Ali ARMIN, Simon DENMAN, Clinton FOOKES, and Lars PETERSSON. "A survey on graph-based deep learning for computational histopathology". In: *Computerized Medical Imaging and Graphics* 95 (2022), p. 102027. ISSN: 0895-6111. DOI: <https://doi.org/10.1016/j.compmedimag.2021.102027>. URL: <https://www.sciencedirect.com/science/article/pii/S0895611121001762> (cit. on p. 58).
- [Aig+11] Wolfgang AIGNER, Silvia MIKSCH, Heidrun SCHUMANN, and Christian TOMINSKI. *Visualization of Time-Oriented Data*. en. Human-Computer Interaction Series. London: Springer, 2011. ISBN: 978-0-85729-078-6 978-0-85729-079-3. DOI: [10.1007/978-0-85729-079-3](https://doi.org/10.1007/978-0-85729-079-3) (cit. on pp. 18, 21, 30).
- [Ajm+22] Faiza AJMI, Faten AJMI, Sarah BEN OTHMAN, Hayfa ZGAYA, Gregoire SMITH, Jean-Marie RENARD, and Slim HAMMADI. "Impact of the Automation of Inpatient Bed Management to Reduce the Emergency Service Waiting Time". In: *Studies in Health Technology and Informatics* 290 (June 2022), pp. 942–946. ISSN: 1879-8365. DOI: [10.3233/SHTI220219](https://doi.org/10.3233/SHTI220219) (cit. on p. 43).
- [Aki+19] Takuya AKIBA, Shotaro SANO, Toshihiko YANASE, Takeru OHTA, and Masanori KOYAMA. "Optuna: A Next-generation Hyperparameter Optimization Framework". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2019 (cit. on p. 63).
- [All83] James F. ALLEN. "Maintaining Knowledge about Temporal Intervals". In: *Communications of the ACM* 26.11 (Nov. 1983), pp. 832–843. ISSN: 0001-0782. DOI: [10.1145/182.358434](https://doi.org/10.1145/182.358434) (cit. on p. 69).
- [Alo+98] Diane Lindwarm ALONSO, Anne ROSE, Catherine PLAISANT, and Kent L. NORMAN. "Viewing Personal History Records: A Comparison of Tabular Format and Graphical Presentation Using LifeLines". In: *Behaviour & Information Technology* 17.5 (Jan. 1998), pp. 249–262. ISSN: 0144-929X. DOI: [10.1080/014492998119328](https://doi.org/10.1080/014492998119328) (cit. on pp. 6, 25, 26, 29).
- [AM06] Wolfgang AIGNER and Silvia MIKSCH. "CareVis: integrated visualization of computerized protocols and temporal patient data". eng. In: *Artificial Intelligence in Medicine* 37.3 (July 2006), pp. 203–218. ISSN: 0933-3657. DOI: [10.1016/j.artmed.2006.04.002](https://doi.org/10.1016/j.artmed.2006.04.002) (cit. on pp. 25, 26, 29, 30).

- [Ash+19] Awais ASHFAQ, Anita SANT'ANNA, Markus LINGMAN, and Sławomir NOWACZYK. "Readmission prediction using deep learning on electronic health records". en. In: *Journal of Biomedical Informatics* 97 (Sept. 2019), p. 103256. ISSN: 1532-0464. DOI: [10.1016/j.jbi.2019.103256](https://doi.org/10.1016/j.jbi.2019.103256) (cit. on p. 8).
- [BAK07] Alex A. T. BUI, Denise R. ABERLE, and Hooshang KANGARLOO. "Time-Line: Visualizing Integrated Patient Records". In: *IEEE Transactions on Information Technology in Biomedicine* 11.4 (July 2007), pp. 462–473. ISSN: 1558-0032. DOI: [10.1109/TITB.2006.884365](https://doi.org/10.1109/TITB.2006.884365) (cit. on pp. 25, 26, 29, 73).
- [Bar+20] Sebastiano BARBIERI, James KEMP, Oscar PEREZ-CONCHA, Sradha KOTWAL, Martin GALLAGHER, Angus RITCHIE, and Louisa JORM. "Benchmarking Deep Learning Architectures for Predicting Readmission to the ICU and Describing Patients-at-Risk". In: *Scientific Reports* 10.1 (Jan. 2020), p. 1111. ISSN: 2045-2322. DOI: [10.1038/s41598-020-58053-z](https://doi.org/10.1038/s41598-020-58053-z) (cit. on p. 6).
- [Bau+10] Dominikus BAUR, Frederik SEIFFERT, Michael SEDLMAIR, and Sebastian BORING. "The streams of our lives: visualizing listening histories in context". eng. In: *IEEE transactions on visualization and computer graphics* 16.6 (2010), pp. 1119–1128. ISSN: 1077-2626. DOI: [10.1109/TVCG.2010.206](https://doi.org/10.1109/TVCG.2010.206) (cit. on pp. 25, 26, 29).
- [Bay+17] Inci M. BAYTAS, Cao XIAO, Xi ZHANG, Fei WANG, Anil K. JAIN, and Jiayu ZHOU. "Patient Subtyping via Time-Aware LSTM Networks". In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, Aug. 2017. DOI: [10.1145/3097983.3097997](https://doi.org/10.1145/3097983.3097997) (cit. on p. 66).
- [BAY22] Shaked BRODY, Uri ALON, and Eran YAHAV. *How Attentive Are Graph Attention Networks?* Tech. rep. arXiv:2105.14491. arXiv, Jan. 2022. arXiv: [2105.14491 \[cs\]](https://arxiv.org/abs/2105.14491) (cit. on p. 76).
- [BBL15] Sara H. BROWNE, Yashar BEHZADI, and Gwen LITTLEWORT. "Let Visuals Tell the Story: Medication Adherence in Patients with Type II Diabetes Captured by a Novel Ingestion Sensor Platform". In: *JMIR mHealth and uHealth* 3.4 (Dec. 2015), e108. ISSN: 2291-5222. DOI: [10.2196/mhealth.4292](https://doi.org/10.2196/mhealth.4292) (cit. on p. 75).
- [BCB16] Dzmitry BAHDANAU, Kyunghyun CHO, and Yoshua BENGIO. *Neural Machine Translation by Jointly Learning to Align and Translate*. 2016. arXiv: [1409.0473 \[cs.CL\]](https://arxiv.org/abs/1409.0473) (cit. on p. 76).
- [Bea+23] Daniel M. BEAN, Zeljko KRALJEVIC, Anthony SHEK, James TEO, and Richard J. B. DOBSON. "Hospital-Wide Natural Language Processing Summarising the Health Data of 1 Million Patients". In: *PLOS Digital Health* 2.5 (May 2023), e0000218. ISSN: 2767-3170. DOI: [10.1371/journal.pdig.0000218](https://doi.org/10.1371/journal.pdig.0000218) (cit. on p. 4).
- [Bha+23] Nandita BHASKHAR, Wui IP, Jonathan H. CHEN, and Daniel L. RUBIN. "Clinical Outcome Prediction Using Observational Supervision with Electronic Health Records and Audit Logs". In: *Journal of Biomedical Informatics* 147 (Nov. 2023), p. 104522. ISSN: 1532-0464. DOI: [10.1016/j.jbi.2023.104522](https://doi.org/10.1016/j.jbi.2023.104522) (cit. on p. 8).

-
- [BKM08] Aaron BANGOR, Philip T. KORTUM, and James T. MILLER. “An empirical evaluation of the system usability scale”. In: *International Journal of Human–Computer Interaction* 24.6 (July 2008), pp. 574–594. ISSN: 1044-7318. DOI: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776) (cit. on p. 38).
- [BL11] Lars BACKSTROM and Jure LESKOVEC. “Supervised Random Walks: Predicting and Recommending Links in Social Networks”. In: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*. (WSDM ’11). Association for Computing Machinery, Feb. 2011, pp. 635–644. ISBN: 978-1-4503-0493-1. DOI: [10.1145/1935826.1935914](https://doi.org/10.1145/1935826.1935914) (cit. on pp. 52, 53, 57).
- [BN01] Mikhail BELKIN and Partha NIYOGI. “Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering”. In: *Advances in Neural Information Processing Systems*. Vol. 14. MIT Press, 2001 (cit. on p. 52).
- [BN03] Mikhail BELKIN and Partha NIYOGI. “Laplacian Eigenmaps for Dimensionality Reduction and Data Representation”. In: *Neural Computation* 15.6 (June 2003), pp. 1373–1396. ISSN: 0899-7667. DOI: [10.1162/08997660321780317](https://doi.org/10.1162/08997660321780317) (cit. on pp. 50, 52).
- [BNJ03] David M. BLEI, Andrew Y. NG, and Michael I. JORDAN. “Latent dirichlet allocation”. In: *J. Mach. Learn. Res.* 3 (2003), pp. 993–1022. ISSN: 1532-4435. DOI: <http://dx.doi.org/10.1162/jmlr.2003.3.4-5.993> (cit. on p. 24).
- [Bol+24] Heloísa OSS BOLL, Ali AMIRAHMADI, Mirfarid Musavian GHAZANI, Wagner Ourique de MORAIS, Edison Pignaton de FREITAS, Amira SOLIMAN, Farzaneh ETMINANI, Stefan BYTTNER, and Mariana RECAMONDE-MENDOZA. “Graph neural networks for clinical risk prediction based on electronic health records: A survey”. In: *Journal of Biomedical Informatics* 151 (2024), p. 104616. ISSN: 1532-0464. DOI: <https://doi.org/10.1016/j.jbi.2024.104616> (cit. on pp. 59, 76).
- [BP98] Sergey BRIN and Lawrence PAGE. “The Anatomy of a Large-Scale Hypertextual Web Search Engine”. In: *Computer Networks and ISDN Systems*. Proceedings of the Seventh International World Wide Web Conference 30.1 (Apr. 1998), pp. 107–117. ISSN: 0169-7552. DOI: [10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X) (cit. on pp. 53, 57).
- [Bre01] Leo BREIMAN. “Random Forests”. English. In: *Machine Learning* 45.1 (2001), pp. 5–32. ISSN: 0885-6125. DOI: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324) (cit. on p. 7).
- [Bro+05] Dominique BRODBECK, Roland GASSER, Markus DEGEN, Serge REICHLIN, and Jürg LUTHIGER. “Enabling Large-Scale Telemedical Disease Management through Interactive Visualization”. In: (Jan. 2005) (cit. on pp. 25, 26, 29).
- [Bro96] John BROOKE. “SUS - A quick and dirty usability scale”. en. In: *CRC Press* (1996) (cit. on pp. 36, 38).
- [Bru+14] Joan BRUNA, Wojciech ZAREMBA, Arthur SZLAM, and Yann LECUN. *Spectral Networks and Locally Connected Networks on Graphs*. 2014. arXiv: [1312.6203](https://arxiv.org/abs/1312.6203) [cs.LG] (cit. on pp. 50, 54).

- [BSM04] Ragnar BADE, Stefan SCHLECHTWEG, and Silvia MIKSCHE. “Connecting time-oriented data and information to a coherent interactive visualization”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (CHI '04). New York, NY, USA: Association for Computing Machinery, 2004, pp. 105–112. ISBN: 978-1-58113-702-6. DOI: [10.1145/985692.985706](https://doi.org/10.1145/985692.985706) (cit. on pp. 25, 26, 29).
- [Car+23] Lucía A. CARRASCO-RIBELLES, José LLANES-JURADO, Carlos GALLEGO-MOLL, Margarita CABRERA-BEAN, Mònica MONTEAGUDO-ZARAGOZA, Concepción VIOLÁN, and Edurne ZABALETA-DEL-OLMO. “Prediction models using artificial intelligence and longitudinal data from electronic health records: a systematic methodological review”. In: *Journal of the American Medical Informatics Association* 30.12 (Sept. 2023), pp. 2072–2082. ISSN: 1527-974X. DOI: [10.1093/jamia/ocad168](https://doi.org/10.1093/jamia/ocad168) (cit. on p. 2).
- [Cel+23] Pedro CELARD, Eva LORENZO IGLESIAS, José Manuel SORRIBES-FDEZ, Rubén ROMERO, A. Seara VIEIRA, and Lourdes BORRAJO. “A Survey on Deep Learning Applied to Medical Images: From Simple Artificial Neural Networks to Generative Models”. In: *Neural Computing and Applications* 35.3 (Jan. 2023), pp. 2291–2323. ISSN: 1433-3058. DOI: [10.1007/s00521-022-07953-4](https://doi.org/10.1007/s00521-022-07953-4) (cit. on p. 4).
- [Cha+21] Qi CHANG, Cancheng LI, Qing TIAN, Qijing BO, Jicong ZHANG, Yanbing XIONG, and Chuanyue WANG. “Classification of First-Episode Schizophrenia, Chronic Schizophrenia and Healthy Control Based on Brain Network of Mismatch Negativity by Graph Neural Network”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021), pp. 1784–1794. ISSN: 1558-0210. DOI: [10.1109/TNSRE.2021.3105669](https://doi.org/10.1109/TNSRE.2021.3105669) (cit. on p. 58).
- [Cha+22] Ines CHAMI, Sami ABU-EL-HAJJA, Bryan PEROZZI, Christopher RÉ, and Kevin MURPHY. “Machine Learning on Graphs: A Model and Comprehensive Taxonomy”. In: *Journal of Machine Learning Research* 23.89 (2022), pp. 1–64. ISSN: 1533-7928 (cit. on pp. 49–51, 54, 57, 79).
- [Cho+16] Edward CHOI, Mohammad Taha BAHADORI, Le SONG, Walter F. STEWART, and Jimeng SUN. “GRAM: Graph-based Attention Model for Healthcare Representation Learning”. In: *CoRR abs/1611.07012* (2016). arXiv: [1611.07012](https://arxiv.org/abs/1611.07012) (cit. on pp. 58, 76).
- [Cho+22] Ha Na CHO, Imjin AHN, Hansle GWON, Hee Jun KANG, Yunha KIM, Hyeram SEO, Heejung CHOI, Minkyung KIM, Jiye HAN, Gaeun KEE, Tae Joon JUN, and Young-Hak KIM. “Heterogeneous graph construction and HinSAGE learning from electronic medical records”. en. In: *Scientific Reports* 12.1 (Dec. 2022), p. 21152. ISSN: 2045-2322. DOI: [10.1038/s41598-022-25693-2](https://doi.org/10.1038/s41598-022-25693-2) (cit. on pp. 59, 76).
- [CJ23] Davide CHICCO et Giuseppe JURMAN. « The Matthews Correlation Coefficient (MCC) Should Replace the ROC AUC as the Standard Metric for Assessing Binary Classification ». In : *BioData Mining* 16.1 (fév. 2023), p. 4. ISSN : 1756-0381. DOI : [10.1186/s13040-023-00322-4](https://doi.org/10.1186/s13040-023-00322-4). (Visité le 05/09/2024) (cf. page 68).
- [CKS10] Carlo COMBI, Elpidia KERAVNOU-PAPAILIOU, and Yuval SHAHAR. *Temporal Information Systems in Medicine*. en. Boston, MA: Springer US, 2010. ISBN:

-
- 978-1-4419-6542-4 978-1-4419-6543-1. DOI: [10.1007/978-1-4419-6543-1](https://doi.org/10.1007/978-1-4419-6543-1) (cit. on pp. 18, 21).
- [Col18] Roger COLLIER. “Rethinking EHR interfaces to reduce click fatigue and physician burnout”. In: *CMAJ : Canadian Medical Association Journal* 190.33 (Aug. 2018), E994–E995. ISSN: 0820-3946. DOI: [10.1503/cmaj.109-5644](https://doi.org/10.1503/cmaj.109-5644) (cit. on p. 5).
- [COP23] Ricardo M. S. CARVALHO, Daniela OLIVEIRA, and Catia PESQUITA. “Knowledge Graph Embeddings for ICU Readmission Prediction”. In: *BMC medical informatics and decision making* 23.1 (Jan. 2023), p. 12. ISSN: 1472-6947. DOI: [10.1186/s12911-022-02070-7](https://doi.org/10.1186/s12911-022-02070-7) (cit. on p. 6).
- [Cor+16] Savelie CORNEGRUTA, Robert BAKEWELL, Samuel WITHEY, and Giovanni MONTANA. *Modelling Radiological Language with Bidirectional Long Short-Term Memory Networks*. Sept. 2016. DOI: [10.48550/arXiv.1609.08409](https://doi.org/10.48550/arXiv.1609.08409). arXiv: [1609.08409](https://arxiv.org/abs/1609.08409) [cs, stat] (cit. on p. 77).
- [CV95] Corinna CORTES and Vladimir N VAPNIK. “Support-vector networks”. In: *Machine learning* 20.3 (1995), pp. 273–297 (cit. on p. 8).
- [Dai+15] Wuyang DAI, Theodora S. BRISIMI, William G. ADAMS, Theofanie MELA, Venkatesh SALIGRAMA, and Ioannis Ch. PASCHALIDIS. “Prediction of Hospitalization Due to Heart Diseases by Supervised Learning Methods”. In: *International journal of medical informatics* 84.3 (Mar. 2015), pp. 189–197. ISSN: 1386-5056. DOI: [10.1016/j.ijmedinf.2014.10.002](https://doi.org/10.1016/j.ijmedinf.2014.10.002) (cit. on p. 8).
- [Day+21] Ittai DAYAN et al. “Federated Learning for Predicting Clinical Outcomes in Patients with COVID-19”. In: *Nature Medicine* 27.10 (Oct. 2021), pp. 1735–1743. ISSN: 1546-170X. DOI: [10.1038/s41591-021-01506-3](https://doi.org/10.1038/s41591-021-01506-3) (cit. on p. 78).
- [DBV16] Michaël DEFFERRARD, Xavier BRESSON, and Pierre VANDERGHEYNST. “Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering”. In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*. (NIPS ’16). Curran Associates Inc., Dec. 2016, pp. 3844–3852. ISBN: 978-1-5108-3881-9 (cit. on p. 54).
- [Dev+19] Jacob DEVLIN, Ming-Wei CHANG, Kenton LEE, and Kristina TOUTANOVA. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, June 2019, pp. 4171–4186. DOI: [10.18653/v1/N19-1423](https://doi.org/10.18653/v1/N19-1423) (cit. on pp. 20, 74).
- [DMP91] Rina DECHTER, Itay MEIRI, and Judea PEARL. “Temporal Constraint Networks”. In: *Artificial Intelligence* 49.1 (May 1991), pp. 61–95. ISSN: 0004-3702. DOI: [10.1016/0004-3702\(91\)90006-6](https://doi.org/10.1016/0004-3702(91)90006-6) (cit. on p. 69).
- [DOW23] Jonas DENCK, Elif OZKIRIMLI, and Ken WANG. “Machine-Learning-Based Adverse Drug Event Prediction from Observational Health Data: A Review”. In: *Drug Discovery Today* 28.9 (Sept. 2023), p. 103715. ISSN: 1359-6446. DOI: [10.1016/j.drudis.2023.103715](https://doi.org/10.1016/j.drudis.2023.103715) (cit. on p. 6).
- [DPA18] Hugo DE OLIVEIRA, Martin PRODEL, and Vincent AUGUSTO. “Binary Classification on French Hospital Data: Benchmark of 7 Machine Learning Algorithms”. In: *2018 IEEE International Conference on Systems, Man, and*

- Cybernetics (SMC)*. Oct. 2018, pp. 1743–1748. DOI: [10.1109/SMC.2018.00301](https://doi.org/10.1109/SMC.2018.00301) (cit. on pp. 7, 8).
- [Dub+18] Sebastien DUBOIS, Nathanael ROMANO, David C. KALE, Nigam SHAH, and Kenneth JUNG. “Effective Representations of Clinical Notes”. In: *arXiv:1705.07025 [cs, stat]* (Aug. 2018). arXiv: [1705.07025](https://arxiv.org/abs/1705.07025) [cs, stat] (cit. on p. 6).
- [Dwi+20] Vijay Prakash DWIVEDI, Chaitanya K. JOSHI, Thomas LAURENT, Yoshua BENGIO, and Xavier BRESSON. “Benchmarking Graph Neural Networks”. In: *arXiv preprint arXiv:2003.00982* (2020) (cit. on p. 67).
- [El-+20] Nora EL-RASHIDY, Shaker EL-SAPPAGH, Tamer ABUHMED, Samir ABDELRAZEK, and Hazem M. EL-BAKRY. “Intensive care unit mortality prediction: An improved patient-specific stacking ensemble model”. In: *IEEE Access* 8 (2020), pp. 133541–133564. ISSN: 2169-3536. DOI: [10.1109/ACCESS.2020.3010556](https://doi.org/10.1109/ACCESS.2020.3010556) (cit. on pp. 6, 8).
- [FEK21] Bahare FATEMI, Layla EL ASRI, and Seyed Mehran KAZEMI. “SLAPS: Self-Supervision Improves Structure Learning for Graph Neural Networks”. In: *Advances in Neural Information Processing Systems*. Vol. 34. Curran Associates, Inc., 2021, pp. 22667–22681 (cit. on p. 70).
- [FL19] Matthias FEY and Jan Eric LENSSEN. “Fast Graph Representation Learning with PyTorch Geometric”. In: *CoRR abs/1903.02428* (2019). arXiv: [1903.02428](https://arxiv.org/abs/1903.02428) (cit. on p. 79).
- [FN11] Anthony FAIOLA and Chris NEWLON. “Advancing Critical Care in the ICU: A Human-Centered Biomedical Data Visualization Systems”. en. In: *Ergonomics and Health Aspects of Work with Computers*. Ed. by Michelle M. ROBERTSON. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2011, pp. 119–128. ISBN: 978-3-642-21716-6. DOI: [10.1007/978-3-642-21716-6_13](https://doi.org/10.1007/978-3-642-21716-6_13) (cit. on pp. 25, 26, 29).
- [Gao+22] Jianliang GAO, Tengfei LYU, Fan XIONG, Jianxin WANG, Weimao KE, and Zhao LI. “Predicting the survival of cancer patients with multimodal graph neural network”. In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 19.2 (Mar. 2022), pp. 699–709. ISSN: 1557-9964. DOI: [10.1109/TCBB.2021.3083566](https://doi.org/10.1109/TCBB.2021.3083566) (cit. on p. 58).
- [GB21] Guillaume GRONIER and Alexandre BAUDET. “Psychometric evaluation of the F-SUS: Creation and validation of the french version of the system usability scale”. In: *International Journal of Human-Computer Interaction* 37.16 (Oct. 2021), pp. 1571–1582. ISSN: 1044-7318. DOI: [10.1080/10447318.2021.1898828](https://doi.org/10.1080/10447318.2021.1898828) (cit. on p. 36).
- [GF18] Palash GOYAL and Emilio FERRARA. “Graph Embedding Techniques, Applications, and Performance: A Survey”. In: *Knowledge-Based Systems* 151 (July 2018), pp. 78–94. ISSN: 09507051. DOI: [10.1016/j.knosys.2018.03.022](https://doi.org/10.1016/j.knosys.2018.03.022). arXiv: [1705.02801](https://arxiv.org/abs/1705.02801) (cit. on pp. 49, 52, 79).
- [Gil+17] Justin GILMER, Samuel S. SCHOENHOLZ, Patrick F. RILEY, Oriol VINYALS, and George E. DAHL. *Neural Message Passing for Quantum Chemistry*. June 2017. DOI: [10.48550/arXiv.1704.01212](https://doi.org/10.48550/arXiv.1704.01212). arXiv: [1704.01212](https://arxiv.org/abs/1704.01212) [cs] (cit. on pp. 55, 79).

-
- [GL16] Aditya GROVER and Jure LESKOVEC. *Node2vec: Scalable Feature Learning for Networks*. July 2016. DOI: [10.48550/arXiv.1607.00653](https://doi.org/10.48550/arXiv.1607.00653). arXiv: [1607.00653](https://arxiv.org/abs/1607.00653) [cs, stat] (cit. on pp. 53, 56, 57).
- [GMS05] Marco GORI, Gabriele MONFARDINI, and Franco SCARSELLI. "A New Model for Learning in Graph Domains". In: *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005*. Vol. 2. July 2005, 729–734 vol. 2. DOI: [10.1109/IJCNN.2005.1555942](https://doi.org/10.1109/IJCNN.2005.1555942) (cit. on pp. 54, 57).
- [Gol+00] Ary L. GOLDBERGER, Luis A. N. AMARAL, Leon GLASS, Jeffrey M. HAUSDORFF, Plamen C. IVANOV, Roger G. MARK, Joseph E. MIETUS, George B. MOODY, Chung-Kang PENG, and Eugene STANLEY. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals." In: *Circulation* 101.23 (June 2000), E215–220. ISSN: 1524-4539 0009-7322. DOI: [10.1161/01.cir.101.23.e215](https://doi.org/10.1161/01.cir.101.23.e215) (cit. on pp. 5, 63).
- [Gro22] Maarten GROOTENDORST. "BERTopic: Neural topic modeling with a class-based TF-IDF procedure". In: *arXiv preprint arXiv:2203.05794* (2022) (cit. on pp. 20, 24).
- [Gru+18] Robert W GRUNDMEIER, Rui XIAO, Rachael K ROSS, Mark J RAMOS, Dean J KARAVITE, Jeremy J MICHEL, Jeffrey S GERBER, and Susan E COFFIN. "Identifying Surgical Site Infections in Electronic Health Data Using Predictive Models". In: *Journal of the American Medical Informatics Association : JAMIA* 25.9 (June 2018), pp. 1160–1166. ISSN: 1067-5027. DOI: [10.1093/jamia/ocy075](https://doi.org/10.1093/jamia/ocy075) (cit. on p. 7).
- [Gsc+11] Theresia GSCHWANDTNER, Wolfgang AIGNER, Katharina KAISER, Silvia MIK-SCH, and Andreas SEYFANG. "CareCruiser: Exploring and visualizing plans, events, and effects interactively". In: *2011 IEEE Pacific Visualization Symposium*. Mar. 2011, pp. 43–50. DOI: [10.1109/PACIFICVIS.2011.5742371](https://doi.org/10.1109/PACIFICVIS.2011.5742371) (cit. on pp. 25, 26, 29).
- [Gu+22] Yifan GU, Xuebing YANG, Lei TIAN, Hongyu YANG, Jicheng LV, Chao YANG, Jinwei WANG, Jianing XI, Guilan KONG, and Wensheng ZHANG. "Structure-aware siamese graph neural networks for encounter-level patient similarity learning". In: *Journal of Biomedical Informatics* 127 (Mar. 2022), p. 104027. ISSN: 1532-0464. DOI: [10.1016/j.jbi.2022.104027](https://doi.org/10.1016/j.jbi.2022.104027) (cit. on pp. 58, 76, 80).
- [Gui22] Riccardo GUIDOTTI. "Counterfactual Explanations and How to Find Them: Literature Review and Benchmarking". In: *Data Mining and Knowledge Discovery* (Apr. 2022). ISSN: 1573-756X. DOI: [10.1007/s10618-022-00831-6](https://doi.org/10.1007/s10618-022-00831-6) (cit. on p. 81).
- [Guo+22] Jiafeng GUO, Yinqiong CAI, Yixing FAN, Fei SUN, Ruqing ZHANG, and Xueqi CHENG. "Semantic models for the first-stage retrieval: A comprehensive review". In: *ACM Transactions on Information Systems* 40.4 (Mar. 2022), 66:1–66:42. ISSN: 1046-8188. DOI: [10.1145/3486250](https://doi.org/10.1145/3486250) (cit. on p. 43).
- [GW12] David GOTZ and Krist WONGSUPHASAWAT. "Interactive Intervention Analysis". In: *AMIA Annual Symposium proceedings*. 2012 (2012), pp. 274–280. ISSN: 1942-597X (cit. on p. 74).

- [Had+22] Neal R. HADDAWAY, Matthew J. PAGE, Chris C. PRITCHARD, and Luke A. McGUINNESS. “PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis”. en. In: *Campbell Systematic Reviews* 18.2 (2022), e1230. ISSN: 1891-1803. DOI: [10.1002/c12.1230](https://doi.org/10.1002/c12.1230) (cit. on p. 19).
- [Hal+19] Alan Arthur HALES, David CABLE, Eleanor CROSSLEY, Callum FINDLAY, and David Anthony REW. “Design and implementation of the stacked, synchronised and iconographic timeline-structured electronic patient record in a UK NHS Global Digital Exemplar hospital”. In: *BMJ Health & Care Informatics* 26.1 (Dec. 2019), e100025. ISSN: 2632-1009. DOI: [10.1136/bmjhci-2019-100025](https://doi.org/10.1136/bmjhci-2019-100025) (cit. on pp. 25, 26, 29).
- [Haq+22] A. K. M. Bahalul HAQUE, B. M. ARIFUZZAMAN, Sayed Abu Noman SIDDIK, Abul KALAM, Tabassum Sadia SHAHJAHAN, T. S. SALEENA, Morshed ALAM, Md. Rabiul ISLAM, Foyez AHMED, and Md. Jamal HOSSAIN. “Semantic Web in Healthcare: A Systematic Literature Review of Application, Research Gap, and Future Research Avenues”. In: *International Journal of Clinical Practice* 2022 (Oct. 2022), p. 6807484. ISSN: 1368-5031. DOI: [10.1155/2022/6807484](https://doi.org/10.1155/2022/6807484) (cit. on p. 69).
- [Har+19] Hrayr HARUTYUNYAN, Hrant KHACHATRIAN, David C. KALE, Greg VER STEEG, and Aram GALSTYAN. “Multitask Learning and Benchmarking with Clinical Time Series Data”. In: *Scientific Data* 6.1 (Dec. 2019), p. 96. ISSN: 2052-4463. DOI: [10.1038/s41597-019-0103-9](https://doi.org/10.1038/s41597-019-0103-9) (cit. on pp. 8, 60, 63–65, 68).
- [HAR19] Kexin HUANG, Jaan ALTOSAAR, and Rajesh RANGANATH. “ClinicalBERT: Modeling Clinical Notes and Predicting Hospital Readmission”. In: *arXiv:1904.05342 [cs]* (Apr. 2019). arXiv: [1904.05342 \[cs\]](https://arxiv.org/abs/1904.05342) (cit. on p. 6).
- [Hat+19] Christopher M. HATTON, Lewis W. PATON, Dean McMILLAN, James CUSSENS, Simon GILBODY, and Paul A. TIFFIN. “Predicting persistent depressive symptoms in older adults: A machine learning approach to personalised mental healthcare”. In: *Journal of Affective Disorders* 246 (Mar. 2019), pp. 857–860. DOI: [10.1016/j.jad.2018.12.095](https://doi.org/10.1016/j.jad.2018.12.095) (cit. on p. 66).
- [Hau99] David HAUSSLER. “Convolution Kernels on Discrete Structures UCSC CRL”. In: 1999 (cit. on p. 52).
- [HB03] Mark HARROWER and Cynthia A. BREWER. “ColorBrewer.Org: An Online Tool for Selecting Colour Schemes for Maps”. In: *The Cartographic Journal* 40.1 (June 2003), pp. 27–37. ISSN: 0008-7041. DOI: [10.1179/000870403235002042](https://doi.org/10.1179/000870403235002042) (cit. on p. 43).
- [He+21] Kaiming HE, Xinlei CHEN, Saining XIE, Yanghao LI, Piotr DOLLÁR, and Ross GIRSHICK. *Masked Autoencoders Are Scalable Vision Learners*. Tech. rep. arXiv:2111.06377. arXiv, Dec. 2021. arXiv: [2111.06377 \[cs\]](https://arxiv.org/abs/2111.06377) (cit. on p. 70).
- [Heo+19] JoonNyung HEO, Jihoon G. YOON, Hyungjong PARK, Young Dae KIM, Hyo Suk NAM, and Ji Hoe HEO. “Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke”. In: *Stroke* 50.5 (May 2019),

-
- pp. 1263–1265. ISSN: 1524-4628. DOI: [10.1161/STROKEAHA.118.024293](https://doi.org/10.1161/STROKEAHA.118.024293) (cit. on p. 6).
- [HHL14] Frank HUTTER, Holger HOOS, and Kevin LEYTON-BROWN. “An Efficient Approach for Assessing Hyperparameter Importance”. In: *Proceedings of the 31st International Conference on Machine Learning*. PMLR, Jan. 2014, pp. 754–762 (cit. on p. 66).
- [Him+09] Blanca E. HIMES, Yi DAI, Isaac S. KOHANE, Scott T. WEISS, and Marco F. RAMONI. “Prediction of Chronic Obstructive Pulmonary Disease (COPD) in Asthma Patients Using Electronic Medical Records”. en. In: *Journal of the American Medical Informatics Association* 16.3 (May 2009), pp. 371–379. ISSN: 1067-5027. DOI: [10.1197/jamia.M2846](https://doi.org/10.1197/jamia.M2846) (cit. on p. 8).
- [HMM00] Ivan HERMAN, Guy MELANCON, and M. Scott MARSHALL. “Graph Visualization and Navigation in Information Visualization: A Survey”. In: *IEEE Transactions on Visualization and Computer Graphics* 6.1 (Jan. 2000), pp. 24–43. ISSN: 1941-0506. DOI: [10.1109/2945.841119](https://doi.org/10.1109/2945.841119) (cit. on p. 80).
- [Hoa+23] Van Thuy HOANG, Hyeon-Ju JEON, Eun-Soon YOU, Yoewon YOON, Sungyeop JUNG, and O.-Joun LEE. “Graph Representation Learning and Its Applications: A Survey”. In: *Sensors* 23.8 (Jan. 2023), p. 4168. ISSN: 1424-8220. DOI: [10.3390/s23084168](https://doi.org/10.3390/s23084168) (cit. on pp. 49, 52).
- [How+08] Eric HOWELL, Edward BESSMAN, Steven KRAVET, Ken KOLODNER, Robert MARSHALL, and Scott WRIGHT. “Active Bed Management by Hospitalists and Emergency Department Throughput”. In: *Annals of Internal Medicine* 149.11 (Dec. 2008), pp. 804–810. ISSN: 0003-4819. DOI: [10.7326/0003-4819-149-11-200812020-00006](https://doi.org/10.7326/0003-4819-149-11-200812020-00006) (cit. on p. 43).
- [HS97] Sepp HOCHREITER and Jürgen SCHMIDHUBER. “Long Short-Term Memory”. In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. ISSN: 0899-7667. DOI: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735) (cit. on p. 8).
- [Hu+20] Ziniu HU, Yuxiao DONG, Kuansan WANG, and Yizhou SUN. “Heterogeneous Graph Transformer”. In: *CoRR* abs/2003.01332 (2020). arXiv: [2003.01332](https://arxiv.org/abs/2003.01332) (cit. on p. 69).
- [HYL17] William L. HAMILTON, Rex YING, and Jure LESKOVEC. “Representation Learning on Graphs: Methods and Applications”. In: *IEEE Data Eng. Bull.* 40.3 (2017), pp. 52–74 (cit. on p. 49).
- [HYL18] William L. HAMILTON, Rex YING, and Jure LESKOVEC. “Inductive Representation Learning on Large Graphs”. In: *arXiv:1706.02216 [cs, stat]* (Sept. 2018). arXiv: [1706.02216 \[cs, stat\]](https://arxiv.org/abs/1706.02216) (cit. on pp. 49, 50, 54, 55, 63, 76, 79).
- [IHG13] Tomoharu IWATA, Neil HOULSBY, and Zoubin GHAHRAMANI. “Active Learning for Interactive Visualization”. In: *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*. PMLR, Apr. 2013, pp. 342–350 (cit. on p. 81).
- [Jäc+16] Dominik JÄCKLE, Fabian FISCHER, Tobias SCHRECK, and Daniel A. KEIM. “Temporal MDS Plots for Analysis of Multivariate Data”. In: *IEEE Transactions on Visualization and Computer Graphics* 22.1 (Jan. 2016), pp. 141–150. ISSN: 1941-0506. DOI: [10.1109/TVCG.2015.2467553](https://doi.org/10.1109/TVCG.2015.2467553) (cit. on pp. 25, 26, 29).

- [Jia+20] Zheng JIA, Xian ZENG, Huilong DUAN, Xudong LU, and Haomin LI. “A Patient-Similarity-Based Model for Diagnostic Prediction”. In: *International Journal of Medical Informatics* 135 (Mar. 2020), p. 104073. ISSN: 1386-5056. DOI: [10.1016/j.ijmedinf.2019.104073](https://doi.org/10.1016/j.ijmedinf.2019.104073) (cit. on p. 80).
- [JK20] Vatika JALALI and Dapinder KAUR. “A Study of Classification and Feature Extraction Techniques for Brain Tumor Detection”. In: *International Journal of Multimedia Information Retrieval* 9.4 (Dec. 2020), pp. 271–290. ISSN: 2192-662X. DOI: [10.1007/s13735-020-00199-7](https://doi.org/10.1007/s13735-020-00199-7) (cit. on p. 74).
- [Joh+16] Alistair E. W. JOHNSON, Tom J. POLLARD, Lu SHEN, Li-Wei H. LEHMAN, Mengling FENG, Mohammad GHASSEMI, Benjamin MOODY, Peter SZOLOVITS, Leo Anthony CELI, and Roger G. MARK. “MIMIC-III, a freely accessible critical care database”. eng. In: *Scientific Data* 3 (May 2016), p. 160035. ISSN: 2052-4463. DOI: [10.1038/sdata.2016.35](https://doi.org/10.1038/sdata.2016.35) (cit. on pp. 5, 36, 63).
- [Joh+23] Alistair E. W. JOHNSON, Lucas BULGARELLI, Lu SHEN, Alvin GAYLES, Ayad SHAMMOUT, Steven HORNG, Tom J. POLLARD, Sicheng HAO, Benjamin MOODY, Brian GOW, Li-wei H. LEHMAN, Leo A. CELI, and Roger G. MARK. “MIMIC-IV, a Freely Accessible Electronic Health Record Dataset”. In: *Scientific Data* 10.1 (Jan. 2023), p. 1. ISSN: 2052-4463. DOI: [10.1038/s41597-022-01899-x](https://doi.org/10.1038/s41597-022-01899-x) (cit. on p. 36).
- [JPM15] Alistair E. W. JOHNSON, Tom J. POLLARD, and Roger G. MARK. *MIMIC-III Clinical Database*. 2015. DOI: [10.13026/C2XW26](https://doi.org/10.13026/C2XW26) (cit. on pp. 5, 36, 63).
- [JPM17] Alistair E. W. JOHNSON, Tom J. POLLARD, and Roger G. MARK. “Reproducibility in Critical Care: A Mortality Prediction Case Study”. In: *Proceedings of the 2nd Machine Learning for Healthcare Conference*. PMLR, Nov. 2017, pp. 361–376 (cit. on p. 6).
- [KA24] Shima KHOSHRAFTAR and Aijun AN. “A Survey on Graph Representation Learning Methods”. In: *ACM Transactions on Intelligent Systems and Technology* 15.1 (Jan. 2024), 19:1–19:55. ISSN: 2157-6904. DOI: [10.1145/3633518](https://doi.org/10.1145/3633518) (cit. on pp. 49, 55).
- [Kan+21] Hiroki KANEKO, Hironobu UMAKOSHI, Masatoshi OGATA, Norio WADA, Norifusa IWAHASHI, Tazuru FUKUMOTO, Maki YOKOMOTO-UMAKOSHI, Yui NAKANO, Yayoi MATSUDA, Takashi MIYAZAWA, Ryuichi SAKAMOTO, and Yoshihiro OGAWA. “Machine learning based models for prediction of sub-type diagnosis of primary aldosteronism using blood test”. In: *Scientific Reports* 11.1 (May 2021). DOI: [10.1038/s41598-021-88712-8](https://doi.org/10.1038/s41598-021-88712-8) (cit. on p. 66).
- [Kar+12] Shreyas KARNIK, Sin Lam TAN, Bess BERG, Ingrid GLURICH, Jinfeng ZHANG, Humberto J. VIDAILLET, C. David PAGE, and Rajesh CHOWDHARY. “Predicting Atrial Fibrillation and Flutter Using Electronic Health Records”. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2012* (2012), pp. 5562–5565. ISSN: 2694-0604. DOI: [10.1109/EMBC.2012.6347254](https://doi.org/10.1109/EMBC.2012.6347254) (cit. on pp. 7, 8).
- [Kat+18] Takayuki KATSUKI, Masaki ONO, Akira KOSEKI, Michiharu KUDO, Kyoichi HAIDA, Jun KURODA, Masaki MAKINO, Ryosuke YANAGIYA, and Atsushi

- SUZUKI. “Risk Prediction of Diabetic Nephropathy via Interpretable Feature Extraction from EHR Using Convolutional Autoencoder”. eng. In: *Studies in Health Technology and Informatics* 247 (2018), pp. 106–110. issn: 1879-8365 (cit. on p. 8).
- [KK17] Hye Jin KAM and Ha Young KIM. “Learning Representations for the Early Detection of Sepsis with Deep Neural Networks”. In: *Computers in Biology and Medicine* 89 (Oct. 2017), pp. 248–255. issn: 0010-4825. doi: [10.1016/j.compbiomed.2017.08.015](https://doi.org/10.1016/j.compbiomed.2017.08.015) (cit. on p. 8).
- [KW17] Thomas N. KIPF and Max WELLING. “Semi-Supervised Classification with Graph Convolutional Networks”. In: *arXiv:1609.02907 [cs, stat]* (Feb. 2017). arXiv: [1609.02907](https://arxiv.org/abs/1609.02907) [cs, stat] (cit. on pp. 50, 54, 55, 57, 76).
- [Lam+08] Jean-Baptiste LAMY, Catherine DUCLOS, Avner BAR-HEN, Patrick OUVARD, and Alain VENOT. “An iconic language for the graphical representation of medical concepts”. In: *BMC Medical Informatics and Decision Making* 8.1 (Apr. 2008), p. 16. issn: 1472-6947. doi: [10.1186/1472-6947-8-16](https://doi.org/10.1186/1472-6947-8-16) (cit. on p. 43).
- [LeB+23a] Hugo LE BAHER, Jérôme AZÉ, Sandra BRINGAY, Pascal PONCELET, Nancy RODRIGUEZ et Caroline DUNOYER. « Modélisation de parcours patients : graphes temporels pour la supervision médicale ». In : *EGC 2023 - Extraction et Gestion des Connaissances*. RNTI-E-39. Jan. 2023, p. 321-328 (cf. page 12).
- [LeB+23b] Hugo LE BAHER, Jérôme AZÉ, Sandra BRINGAY, Pascal PONCELET, Nancy RODRIGUEZ, and Caroline DUNOYER. “Patient Electronic Health Record as Temporal Graphs for Health Monitoring”. In: *Studies in health technology and informatics* 302 (May 2023), pp. 561–565. issn: 0926-9630. doi: [10.3233/shti230205](https://doi.org/10.3233/shti230205) (cit. on p. 12).
- [LeB+23c] Hugo LE BAHER, Jérôme AZÉ, Sandra BRINGAY, Pascal PONCELET, Nancy RODRIGUEZ et Caroline DUNOYER. « Exploration de la représentation latente d’un modèle profond de patient sous forme de graphes ». In : *Actes de la journée d’étude sur la Similarité entre Patients, SimPa 2023* (2023), p. 6 (cf. page 12).
- [Lee+19] John Boaz LEE, Ryan A. ROSSI, Sungchul KIM, Nesreen K. AHMED, and Eunyee KOH. “Attention Models in Graphs: A Survey”. In: *ACM Transactions on Knowledge Discovery from Data* 13.6 (Dec. 2019), pp. 1–25. issn: 1556-4681, 1556-472X. doi: [10.1145/3363574](https://doi.org/10.1145/3363574) (cit. on pp. 49, 57).
- [Lev65] V. LEVENSHTAIN. “Binary Codes Capable of Correcting Deletions, Insertions, and Reversals”. In: *Soviet physics. Doklady* (1965) (cit. on p. 43).
- [Li+20a] Yang LI, Buyue QIAN, Xianli ZHANG, and Hui LIU. “Graph Neural Network-Based Diagnosis Prediction”. In: *Big Data* 8.5 (2020), pp. 379–390. doi: [10.1089/big.2020.0070](https://doi.org/10.1089/big.2020.0070) (cit. on pp. 58, 76).
- [Li+20b] Yikuan LI, Shishir RAO, José Roberto Ayala SOLARES, Abdelaali HASSAINE, Rema RAMAKRISHNAN, Dexter CANOY, Yajie ZHU, Kazem RAHIMI, and Gholamreza SALIMI-KHORSHIDI. “BEHRT: Transformer for Electronic Health Records”. In: *Scientific Reports* 10.1 (Apr. 2020), p. 7155. issn: 2045-2322. doi: [10.1038/s41598-020-62922-y](https://doi.org/10.1038/s41598-020-62922-y). (Visited on 09/02/2024) (cit. on p. 8).

- [Lig+22] Anne-Laure LIGOZAT, Julien LEFEVRE, Aurélie BUGEAU, and Jacques COMBAZ. “Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions”. In: *Sustainability* 14.9 (Jan. 2022), p. 5172. ISSN: 2071-1050. DOI: [10.3390/su14095172](https://doi.org/10.3390/su14095172) (cit. on p. 78).
- [Lik32] Rensis LIKERT. “A Technique for the Measurement of Attitudes”. In: *Archives of Psychology* 22 140 (1932), pp. 55–55 (cit. on p. 23).
- [Lin+23] Claudio D. G. LINHARES, Daniel M. LIMA, Jean R. PONCIANO, Mauro M. OLIVATTO, Marco A. GUTIERREZ, Jorge POCO, Caetano TRAINA, and Agma J. M. TRAINA. “ClinicalPath: A visualization tool to improve the evaluation of electronic health records in clinical decision-making”. In: *IEEE Transactions on Visualization and Computer Graphics* 29.10 (Oct. 2023), pp. 4031–4046. ISSN: 1941-0506. DOI: [10.1109/TVCG.2022.3175626](https://doi.org/10.1109/TVCG.2022.3175626) (cit. on pp. 6, 25, 26, 29).
- [Lit+17] Geert LITJENS, Thijs KOOL, Babak Ehteshami BEJNORDI, Arnaud Arindra Adiyoso SETIO, Francesco CIOMPI, Mohsen GHAFORIAN, Jeroen A. W. M. VAN DER LAAK, Bram VAN GINNEKEN, and Clara I. SÁNCHEZ. “A Survey on Deep Learning in Medical Image Analysis”. In: *Medical Image Analysis* 42 (Dec. 2017), pp. 60–88. ISSN: 1361-8415. DOI: [10.1016/j.media.2017.07.005](https://doi.org/10.1016/j.media.2017.07.005) (cit. on p. 4).
- [Liu+19] Luchen LIU, Haoran LI, Zhiting HU, Haoran SHI, Zichang WANG, Jian TANG, and Ming ZHANG. “Learning hierarchical representations of electronic health records for clinical outcome prediction”. eng. In: *AMIA ... Annual Symposium proceedings. AMIA Symposium 2019* (2019), pp. 597–606. ISSN: 1942-597X (cit. on p. 6).
- [Liu+20] Sicen LIU, Tao LI, Haoyang DING, Buzhou TANG, Xiaolong WANG, Qingcai CHEN, Jun YAN, and Yi ZHOU. “A hybrid method of recurrent neural network and graph neural network for next-period prescription prediction”. en. In: *International Journal of Machine Learning and Cybernetics* 11.12 (Dec. 2020), pp. 2849–2856. ISSN: 1868-8071, 1868-808X. DOI: [10.1007/s13042-020-01155-x](https://doi.org/10.1007/s13042-020-01155-x) (cit. on pp. 58, 76).
- [Liu+21] Yang LIU, Xiang AO, Zidi QIN, Jianfeng CHI, Jinghua FENG, Hao YANG, and Qing HE. “Pick and Choose: A GNN-based Imbalanced Learning Approach for Fraud Detection”. In: *Proceedings of the Web Conference 2021*. (WWW '21). Association for Computing Machinery, June 2021, pp. 3168–3177. ISBN: 978-1-4503-8312-7. DOI: [10.1145/3442381.3449989](https://doi.org/10.1145/3442381.3449989) (cit. on p. 47).
- [Liu+22] Siru LIU, Joseph J. SCHLESINGER, Allison B. MCCOY, Thomas J. REESE, Bryan STEITZ, Elise RUSSO, Brian KOH, and Adam WRIGHT. “New Onset Delirium Prediction Using Machine Learning and Long Short-Term Memory (LSTM) in Electronic Health Record”. In: *Journal of the American Medical Informatics Association: JAMIA* 30.1 (Dec. 2022), pp. 120–131. ISSN: 1527-974X. DOI: [10.1093/jamia/ocac210](https://doi.org/10.1093/jamia/ocac210) (cit. on p. 8).
- [Lop09] Patrice LOPEZ. “GROBID: Combining Automatic Bibliographic Data Recognition and Term Extraction for Scholarship Publications”. In: *Re-*

-
- search and Advanced Technology for Digital Libraries*. Springer Berlin Heidelberg, 2009, pp. 473–474. ISBN: 978-3-642-04346-8 (cit. on p. 24).
- [Luc+21] Christine A. LUCAS, Emily HADLEY, Robert CHEW, Jason NANCE, Peter BAUMGARTNER, Rita THISSEN, David M. PLOTNER, Christine CARR, and Aerian TATUM. “Machine Learning for Medical Coding in Healthcare Surveys”. In: *Vital and Health Statistics. Ser. 1, Programs and Collection Procedures* 189 (Oct. 2021), pp. 1–29. ISSN: 2332-8401 (cit. on p. 6).
- [Mag+19] Jessica MAGALLANES, Lindsey van GEMEREN, Steven WOOD, and Maria-Cruz VILLA-URIOL. “Analyzing Time Attributes in Temporal Event Sequences”. In: *CoRR abs/1908.00903* (2019). arXiv: 1908.00903 (cit. on p. 74).
- [Mah+21] Omar MAHMOOD, Elman MANSIMOV, Richard BONNEAU, and Kyunghyun CHO. “Masked Graph Modeling for Molecule Generation”. In: *Nature Communications* 12.1 (May 2021), p. 3156. ISSN: 2041-1723. DOI: 10.1038/s41467-021-23415-2 (cit. on p. 47).
- [Mam21] Priyanka Mary MAMMEN. *Federated Learning: Opportunities and Challenges*. Jan. 2021. DOI: 10.48550/arXiv.2101.05428. arXiv: 2101.05428 [cs] (cit. on p. 78).
- [Mar+04] Susana B. MARTINS, Yuval SHAHAR, Maya GALPERIN, Herbert KAIZER, Dina GOREN-BAR, Deborah McNAUGHTON, Lawrence V. BASSO, and Mary K. GOLDSTEIN. “Evaluation of KNAVE-II: a tool for intelligent query and exploration of patient data”. eng. In: *Studies in Health Technology and Informatics* 107.Pt 1 (2004), pp. 648–652. ISSN: 0926-9630 (cit. on pp. 25, 26, 29).
- [Mei91] Itay MEIRI. “Combining Qualitative and Quantitative Constraints in Temporal Reasoning”. In: *Artificial Intelligence* 87.1 (Nov. 1991), pp. 343–385. ISSN: 0004-3702. DOI: 10.1016/0004-3702(95)00109-3 (cit. on p. 69).
- [Mek+20] Rachda Naila MEKHALDI, Patrice CAULIER, Sondes CHAABANE, Abdelahad CHRAIBI, and Sylvain PIECHOWIAK. “Using Machine Learning Models to Predict the Length of Stay in a Hospital Setting”. In: *Trends and Innovations in Information Systems and Technologies. Advances in Intelligent Systems and Computing*. Springer International Publishing, 2020, pp. 202–211. ISBN: 978-3-030-45688-7. DOI: 10.1007/978-3-030-45688-7_21 (cit. on p. 6).
- [MF20] Yuqing MAO and Kin Wah FUNG. “Use of word and graph embedding to measure semantic relatedness between Unified Medical Language System concepts”. eng. In: *Journal of the American Medical Informatics Association: JAMIA* 27.10 (Oct. 2020), pp. 1538–1546. ISSN: 1527-974X. DOI: 10.1093/jamia/ocaa136 (cit. on p. 43).
- [Mik+13a] Tomas MIKOLOV, Kai CHEN, Gregory S. CORRADO, and Jeffrey DEAN. “Efficient Estimation of Word Representations in Vector Space”. In: *International Conference on Learning Representations*. 2013 (cit. on pp. 53, 54).
- [Mik+13b] Tomas MIKOLOV, Ilya SUTSKEVER, Kai CHEN, Greg S CORRADO, and Jeff DEAN. “Distributed Representations of Words and Phrases and Their

- Compositionality". In: *Advances in Neural Information Processing Systems*. Vol. 26. Curran Associates, Inc., 2013 (cit. on pp. 43, 53).
- [Moh+09] David MOHER, Alessandro LIBERATI, Jennifer TETZLAFF, and Douglas G. ALTMAN. "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement". In: *PLoS Medicine* 6.7 (July 2009), e1000097. ISSN: 1549-1277. DOI: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097) (cit. on p. 18).
- [Mot+21] Virraji MOTHUKURI, Reza M. PARIZI, Seyedamin POURIYEH, Yan HUANG, Ali DEGHANTANHA et Gautam SRIVASTAVA. « A Survey on Security and Privacy of Federated Learning ». In : *Future Generation Computer Systems* 115 (fév. 2021), p. 619-640. ISSN : 0167-739X. DOI : [10.1016/j.future.2020.10.007](https://doi.org/10.1016/j.future.2020.10.007). (Visité le 04/09/2024) (cf. page 78).
- [MSC01] Miller McPHERSON, Lynn SMITH-LOVIN, and James M. COOK. "Birds of a Feather: Homophily in Social Networks". In: *Annual Review of Sociology* 27 (2001), pp. 415–444. ISSN: 0360-0572. JSTOR: 2678628 (cit. on p. 56).
- [Mun09] Tamara MUNZNER. "A Nested Model for Visualization Design and Validation". In: *IEEE Transactions on Visualization and Computer Graphics* 15.6 (2009), pp. 921–928. DOI: [10.1109/TVCG.2009.111](https://doi.org/10.1109/TVCG.2009.111) (cit. on p. 27).
- [NJM21] Zara NASAR, Syed Waqar JAFFRY, and Muhammad Kamran MALIK. "Named Entity Recognition and Relation Extraction: State-of-the-Art". In: *ACM Computing Surveys* 54.1 (Feb. 2021), 20:1–20:39. ISSN: 0360-0300. DOI: [10.1145/34445965](https://doi.org/10.1145/34445965) (cit. on p. 74).
- [NK18] Maximilian NICKEL and Douwe KIELA. *Learning Continuous Hierarchies in the Lorentz Model of Hyperbolic Geometry*. July 2018. DOI: [10.48550/arXiv.1806.03417](https://doi.org/10.48550/arXiv.1806.03417). arXiv: [1806.03417](https://arxiv.org/abs/1806.03417) [cs, stat] (cit. on p. 69).
- [NNP22] Dai QUOC NGUYEN, Tu Dinh NGUYEN, and Dinh PHUNG. *Universal Graph Transformer Self-Attention Networks*. Mar. 2022. DOI: [10.48550/arXiv.1909.11855](https://doi.org/10.48550/arXiv.1909.11855). arXiv: [1909.11855](https://arxiv.org/abs/1909.11855) [cs, stat] (cit. on p. 76).
- [Oug+19] Rose OUGHTRED et al. "The BioGRID Interaction Database: 2019 Update". In: *Nucleic Acids Research* 47.D1 (Jan. 2019), pp. D529–D541. ISSN: 0305-1048. DOI: [10.1093/nar/gky1079](https://doi.org/10.1093/nar/gky1079) (cit. on p. 69).
- [Par+18] E. PARIMBELLI, S. MARINI, L. SACCHI, and R. BELLAZZI. "Patient Similarity for Precision Medicine: A Systematic Review". In: *Journal of Biomedical Informatics* 83 (July 2018), pp. 87–96. ISSN: 1532-0464. DOI: [10.1016/j.jbi.2018.06.001](https://doi.org/10.1016/j.jbi.2018.06.001) (cit. on p. 80).
- [PAS14] Bryan PEROZZI, Rami AL-RFOU, and Steven SKIENA. "DeepWalk: Online Learning of Social Representations". In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. (KDD '14). Association for Computing Machinery, Aug. 2014, pp. 701–710. ISBN: 978-1-4503-2956-9. DOI: [10.1145/2623330.2623732](https://doi.org/10.1145/2623330.2623732) (cit. on pp. 53, 57).
- [PDE20] Nadeesha PERERA, Matthias DEHMER, and Frank EMMERT-STREIB. "Named Entity Recognition and Relation Detection for Biomedical Information Extraction". In: *Frontiers in Cell and Developmental Biology* 8 (Aug. 2020). ISSN: 2296-634X. DOI: [10.3389/fcell.2020.00673](https://doi.org/10.3389/fcell.2020.00673) (cit. on p. 74).

-
- [PFH07] David S. PIECZKIEWICZ, Stanley M. FINKELSTEIN, and Marshall I. HERTZ. “Design and Evaluation of a Web-Based Interactive Visualization System for Lung Transplant Home Monitoring Data”. In: *AMIA Annual Symposium Proceedings 2007 (2007)*, pp. 598–602. ISSN: 1942-597X (cit. on pp. 25, 26, 29, 30).
- [Pol+18] Peter J. POLACK, Shang-Tse CHAN, Minsik KAHNG, Kaya DE BARBARO, Rahul BASOLE, Moushumi SHARMIN, and Duen Horng CHAU. “Chronodes: Interactive Multifocus Exploration of Event Sequences”. In: *ACM transactions on interactive intelligent systems* 8.1 (Feb. 2018), p. 2. ISSN: 2160-6455. DOI: [10.1145/3152888](https://doi.org/10.1145/3152888) (cit. on p. 75).
- [PSH22] Bastian PFEIFER, Anna SARANTI, and Andreas HOLZINGER. “GNN-SubNet: disease subnetwork detection with explainable graph neural networks”. In: *Bioinformatics* 38.Supplement_2 (Sept. 2022), pp. ii120–ii126. ISSN: 1367-4803. DOI: [10.1093/bioinformatics/btac478](https://doi.org/10.1093/bioinformatics/btac478) (cit. on p. 58).
- [Pur+17] Sanjay PURUSHOTHAM, Chuizheng MENG, Zhengping CHE, and Yan LIU. “Benchmark of Deep Learning Models on Large Healthcare MIMIC Datasets”. In: *arXiv:1710.08531 [cs, stat]* (Oct. 2017). arXiv: [1710.08531](https://arxiv.org/abs/1710.08531) [cs, stat] (cit. on pp. 6, 8).
- [Qur+23] Imran QURESHI, Junhua YAN, Qaisar ABBAS, Kashif SHAHEED, Awais Bin RIAZ, Abdul WAHID, Muhammad Waseem Jan KHAN, and Piotr SZCZUKO. “Medical Image Segmentation Using Deep Semantic-Based Methods: A Review of Techniques, Applications and Emerging Trends”. In: *Information Fusion* 90 (Feb. 2023), pp. 316–352. ISSN: 1566-2535. DOI: [10.1016/j.inffus.2022.09.031](https://doi.org/10.1016/j.inffus.2022.09.031) (cit. on p. 74).
- [Raj+18] Alvin RAJKOMAR et al. “Scalable and Accurate Deep Learning with Electronic Health Records”. In: *npj Digital Medicine* 1.1 (May 2018), pp. 1–10. ISSN: 2398-6352. DOI: [10.1038/s41746-018-0029-1](https://doi.org/10.1038/s41746-018-0029-1) (cit. on pp. 6, 8).
- [Ran+23] Xingcheng RAN, Yue XI, Yonggang LU, Xiangwen WANG, and Zhenyu LU. “Comprehensive survey on hierarchical clustering algorithms and the recent developments”. en. In: *Artificial Intelligence Review* 56.8 (Aug. 2023), pp. 8219–8264. ISSN: 1573-7462. DOI: [10.1007/s10462-022-10366-3](https://doi.org/10.1007/s10462-022-10366-3) (cit. on p. 43).
- [RBH22] Eliane RÖÖSLI, Selen BOZKURT, and Tina HERNANDEZ-BOUSSARD. “Peeking into a Black Box, the Fairness and Generalizability of a MIMIC-III Benchmarking Model”. In: *Scientific Data* 9.1 (Dec. 2022), p. 24. ISSN: 2052-4463. DOI: [10.1038/s41597-021-01110-7](https://doi.org/10.1038/s41597-021-01110-7) (cit. on pp. 6, 8).
- [Ren+21] Pengzhen REN, Yun XIAO, Xiaojun CHANG, Po-Yao HUANG, Zhihui LI, Brij B. GUPTA, Xiaojiang CHEN, and Xin WANG. “A Survey of Deep Active Learning”. In: *ACM Computing Surveys* 54.9 (Oct. 2021), 180:1–180:40. ISSN: 0360-0300. DOI: [10.1145/3472291](https://doi.org/10.1145/3472291) (cit. on p. 81).
- [Rin+11] Alexander RIND, Wolfgang AIGNER, Silvia MIKSCH, Sylvia WILTNER, Margit POHL, Thomas TURIC, and Felix DREXLER. “Visual exploration of time-oriented patient data for chronic diseases: design study and evaluation”. en. In: *Information Quality in e-Health*. Ed. by Andreas HOLZINGER and Klaus-Martin SIMONIC. Lecture Notes in Computer Science. Berlin,

- Heidelberg: Springer, 2011, pp. 301–320. ISBN: 978-3-642-25364-5. DOI: [10.1007/978-3-642-25364-5_22](https://doi.org/10.1007/978-3-642-25364-5_22) (cit. on pp. 25, 26, 29).
- [Rin+13] Alexander RIND, Taowei David WANG, Wolfgang AIGNER, Silvia MIKSCH, Krist WONGSUPHASAWAT, Catherine PLAISANT, and Ben SHNEIDERMAN. “Interactive Information Visualization to Explore and Query Electronic Health Records”. English. In: *Foundations and Trends® in Human–Computer Interaction* 5.3 (Feb. 2013), pp. 207–298. ISSN: 1551-3955, 1551-3963. DOI: [10.1561/11000000039](https://doi.org/10.1561/11000000039) (cit. on pp. 18, 21).
- [Ron+20] Yu RONG, Yatao BIAN, Tingyang XU, Weiyang XIE, Ying WEI, Wenbing HUANG, and Junzhou HUANG. “Self-Supervised Graph Transformer on Large-Scale Molecular Data”. In: *Advances in Neural Information Processing Systems*. Vol. 33. Curran Associates, Inc., 2020, pp. 12559–12571 (cit. on p. 70).
- [Sca+09] Franco SCARSELLI, Marco GORI, Ah Chung TSOI, Markus HAGENBUCHNER, and Gabriele MONFARDINI. “The Graph Neural Network Model”. In: *IEEE Transactions on Neural Networks* 20.1 (2009), pp. 61–80. DOI: [10.1109/TNN.2008.2005605](https://doi.org/10.1109/TNN.2008.2005605) (cit. on pp. 54, 57).
- [Sch+22] Jan SCHEER, Alisa VOLKERT, Nicolas BRICH, Lina WEINERT, Nandhini SANTHANAM, Michael KRONE, Thomas GANSLANDT, Martin BOEKER, and Till NAGEL. “Visualization Techniques of Time-Oriented Data for the Comparison of Single Patients With Multiple Patients or Cohorts: Scoping Review”. EN. In: *Journal of Medical Internet Research* 24.10 (Oct. 2022), e38041. DOI: [10.2196/38041](https://doi.org/10.2196/38041) (cit. on pp. 18–21, 74, 75, 106).
- [SGM19] Emma STRUBELL, Ananya GANESH, and Andrew McCALLUM. “Energy and Policy Considerations for Deep Learning in NLP”. In: *CoRR abs/1906.02-243* (2019). arXiv: [1906.02243](https://arxiv.org/abs/1906.02243) (cit. on pp. 78, 79).
- [Sha+23] Amit SHARMA, Ashutosh SHARMA, Polina NIKASHINA, Vadim GAVRILENKO, Alexey TSELYKH, Alexander BOZHENYUK, Mehedi MASUD, and Hossam MESHREF. “A Graph Neural Network (GNN)-Based Approach for Real-Time Estimation of Traffic Speed in Sustainable Smart Cities”. In: *Sustainability* 15.15 (Jan. 2023), p. 11893. ISSN: 2071-1050. DOI: [10.3390/su151511893](https://doi.org/10.3390/su151511893) (cit. on p. 47).
- [Shi+22] Benjamin SHICKEL, Brandon SILVA, Tezcan OZRAZGAT-BASLANTI, Yuanfang REN, Kia KHEZELI, Ziyuan GUAN, Patrick J. TIGHE, AZYA BIHORAC, and Parisa RASHIDI. “Multi-Dimensional Patient Acuity Estimation with Longitudinal EHR Tokenization and Flexible Transformer Networks”. In: *Frontiers in Digital Health* 4 (Nov. 2022), p. 1029191. ISSN: 2673-253X. DOI: [10.3389/fdgth.2022.1029191](https://doi.org/10.3389/fdgth.2022.1029191) (cit. on p. 77).
- [Shi+23] Minjeong SHIN, Joohee KIM, Yunha HAN, Lexing XIE, Mitchell WHITELAW, Bum Chul KWON, Sungahn KO, and Niklas ELMQVIST. “Roslingifier: Semi-Automated Storytelling for Animated Scatterplots”. In: *IEEE Transactions on Visualization and Computer Graphics* 29.6 (June 2023), pp. 2980–2995. ISSN: 1941-0506. DOI: [10.1109/TVCG.2022.3146329](https://doi.org/10.1109/TVCG.2022.3146329) (cit. on pp. 18, 25, 26, 29).
- [Sil+22] Gabriel F. S. SILVA, Thales P. FAGUNDES, Bruno C. TEIXEIRA, and Alexandre D. P. CHIAVEGATTO FILHO. “Machine Learning for Hypertension Predic-

-
- tion: A Systematic Review”. In: *Current Hypertension Reports* 24.11 (Nov. 2022), pp. 523–533. ISSN: 1534-3111. DOI: [10.1007/s11906-022-01212-6](https://doi.org/10.1007/s11906-022-01212-6) (cit. on p. 6).
- [SMM12] Michael SEDLMAIR, Miriah MEYER, and Tamara MUNZNER. “Design Study Methodology: Reflections from the Trenches and the Stacks”. In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), pp. 2431–2440. DOI: [10.1109/TVCG.2012.213](https://doi.org/10.1109/TVCG.2012.213) (cit. on p. 27).
- [Spe22] Timo SPEITH. “A Review of Taxonomies of Explainable Artificial Intelligence (XAI) Methods”. In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. (FAccT ’22). Association for Computing Machinery, June 2022, pp. 2239–2250. ISBN: 978-1-4503-9352-2. DOI: [10.1145/3531146.3534639](https://doi.org/10.1145/3531146.3534639) (cit. on p. 80).
- [SR20] Balasubramaniam SRINIVASAN and Bruno RIBEIRO. *On the Equivalence between Positional Node Embeddings and Structural Graph Representations*. Sept. 2020. DOI: [10.48550/arXiv.1910.00452](https://doi.org/10.48550/arXiv.1910.00452). arXiv: [1910.00452](https://arxiv.org/abs/1910.00452) [cs, stat] (cit. on p. 50).
- [SRS20] Smadar SHILO, Hagai ROSSMAN, and Eran SEGAL. “Axes of a revolution: challenges and promises of big data in healthcare”. en. In: *Nature Medicine* 26.1 (Jan. 2020), pp. 29–38. ISSN: 1078-8956, 1546-170X. DOI: [10.1038/s41591-019-0727-5](https://doi.org/10.1038/s41591-019-0727-5) (cit. on pp. 2, 7).
- [Ste+20] Nastasja STEINHAEUER, Marc HÖRBRUGGER, Andreas Dominik BRAUN, Thomas TÜTING, Steffen OELTZE-JAFRA, and Juliane MÜLLER. *Comprehensive Visualization of Longitudinal Patient Data for the Dermatological Oncological Tumor Board*. en. The Eurographics Association, 2020. ISBN: 978-3-03868-106-9. DOI: [10.2312/evs.20201067](https://doi.org/10.2312/evs.20201067) (cit. on pp. 25, 26, 29, 30).
- [Sun+22] Li SUN, Junda YE, Hao PENG, and Philip S. YU. “A Self-supervised Riemannian GNN with Time Varying Curvature for Temporal Graph Learning”. In: *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. (CIKM ’22). Association for Computing Machinery, Oct. 2022, pp. 1827–1836. ISBN: 978-1-4503-9236-5. DOI: [10.1145/3511808.3557222](https://doi.org/10.1145/3511808.3557222) (cit. on pp. 69, 70).
- [Tor04] William TORMEY. “The Impact of Inpatient Bed Mismanagement on the Irish Health Service”. In: *Irish Journal of Medical Science* 173.3 (July 2004), pp. 124–124. ISSN: 1863-4362. DOI: [10.1007/BF03167922](https://doi.org/10.1007/BF03167922) (cit. on p. 43).
- [Tor22] Anissa TORKI. « Impact du Dossier Patient Informatisé sur la qualité des soins. L’expérience d’un centre hospitalier au Luxembourg ». In : *Projectics / Proy ctica / Projectique Hors S rie (HS 2022)*, p. 57-79. ISSN : 9782807399501. DOI : [10.3917/proj.hs03.0057](https://doi.org/10.3917/proj.hs03.0057) (cf. page 2).
- [Tri+18] Andrea C. TRICCO et al. “PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation”. In: *Annals of Internal Medicine* 169.7 (Oct. 2018), pp. 467–473. ISSN: 0003-4819. DOI: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850) (cit. on pp. 18, 19).
- [Tur+19] Meghan Reading TURCHIOE, Annie MYERS, Samuel ISAAC, Dawon BAIK, Lisa V. GROSSMAN, Jessica S. ANCKER, and Ruth Masterson CREBER. “A Systematic Review of Patient-Facing Visualizations of Personal Health

- Data". eng. In: *Applied Clinical Informatics* 10.4 (Aug. 2019), pp. 751–770. ISSN: 1869-0327. DOI: [10.1055/s-0039-1697592](https://doi.org/10.1055/s-0039-1697592) (cit. on pp. 18, 21).
- [Vas+17] Ashish VASWANI, Noam SHAZEER, Niki PARMAR, Jakob USZKOREIT, Llion JONES, Aidan N. GOMEZ, Lukasz KAISER, and Illia POLOSUKHIN. "Attention Is All You Need". In: *arXiv:1706.03762 [cs]* (Dec. 2017). arXiv: [1706.03762 \[cs\]](https://arxiv.org/abs/1706.03762) (cit. on p. 76).
- [VC20] Josep Lluís VEGA and Ricard CASADEVALL. "MALO (More Activity, Less Overcrowding): A Bed Management Model as a High-Impact Solution to the Problem of Hospital Overcrowding". In: *Management in Healthcare* 4.4 (Jan. 2020), pp. 331–339 (cit. on p. 43).
- [Vel+18] Petar VELIČKOVIĆ, Guillem CUCURULL, Arantxa CASANOVA, Adriana ROMERO, Pietro LIÒ, and Yoshua BENGIO. *Graph Attention Networks*. Tech. rep. arXiv:1710.10903. arXiv, Feb. 2018. arXiv: [1710.10903 \[cs, stat\]](https://arxiv.org/abs/1710.10903) (cit. on pp. 57, 76).
- [VF05] Michel VERLEYSSEN and Damien FRANÇOIS. "The Curse of Dimensionality in Data Mining and Time Series Prediction". In: *Computational Intelligence and Bioinspired Systems*. Springer, 2005, pp. 758–770. ISBN: 978-3-540-32106-4. DOI: [10.1007/11494669_93](https://doi.org/10.1007/11494669_93) (cit. on p. 7).
- [VK20] Ike VAYANSKY and Sathish A. P. KUMAR. "A Review of Topic Modeling Methods". In: *Information Systems* 94 (Dec. 2020), p. 101582. ISSN: 0306-4379. DOI: [10.1016/j.is.2020.101582](https://doi.org/10.1016/j.is.2020.101582) (cit. on p. 74).
- [Wan+09] Taowei David WANG, Catherine PLAISANT, Ben SHNEIDERMAN, Neil SPRING, David ROSEMAN, Greg MARCHAND, Vikramjit MUKHERJEE, and Mark SMITH. "Temporal Summaries: Supporting Temporal Categorical Searching, Aggregation and Comparison". In: *IEEE Transactions on Visualization and Computer Graphics* 15.6 (Nov. 2009), pp. 1049–1056. ISSN: 1941-0506. DOI: [10.1109/TVCG.2009.187](https://doi.org/10.1109/TVCG.2009.187) (cit. on p. 75).
- [Wan+18] Yanshan WANG, Liwei WANG, Majid RASTEGAR-MOJARAD, Sungrim MOON, Feichen SHEN, Naveed AFZAL, Sijia LIU, Yuqun ZENG, Saeed MEHRABI, Sunghwan SOHN, and Hongfang LIU. "Clinical Information Extraction Applications: A Literature Review". In: *Journal of Biomedical Informatics* 77 (Jan. 2018), pp. 34–49. ISSN: 1532-0480. DOI: [10.1016/j.jbi.2017.11.011](https://doi.org/10.1016/j.jbi.2017.11.011) (cit. on p. 74).
- [Wan+19a] Yu-Hsiang WANG, Phung-Anh NGUYEN, Md Mohaimenul ISLAM, Yu-Chuan LI, and Hsuan-Chia YANG. "Development of Deep Learning Algorithm for Detection of Colorectal Cancer in EHR Data". eng. In: *Studies in Health Technology and Informatics* 264 (Aug. 2019), pp. 438–441. ISSN: 1879-8365. DOI: [10.3233/SHTI190259](https://doi.org/10.3233/SHTI190259) (cit. on p. 8).
- [Wan+19b] Minjie WANG et al. "Deep Graph Library: Towards Efficient and Scalable Deep Learning on Graphs". In: *CoRR abs/1909.01315* (2019). arXiv: [1909.01315](https://arxiv.org/abs/1909.01315) (cit. on p. 79).
- [Wan+21] Zihan WANG, Jialin LU, Oliver SNOW, and Martin ESTER. "An Interactive Visualization Tool for Understanding Active Learning". In: *CoRR abs/2111.-04936* (2021) (cit. on p. 81).
- [Wan+24] Zekai WANG, Suhao CHEN, Tieming LIU, and Bing YAO. "Multi-Branching Temporal Convolutional Network With Tensor Data Completion for

-
- Diabetic Retinopathy Prediction". In: *IEEE Journal of Biomedical and Health Informatics* 28.3 (Mar. 2024), pp. 1704–1715. ISSN: 2168-2208. DOI: [10.1109/JBHI.2024.3351949](https://doi.org/10.1109/JBHI.2024.3351949) (cit. on p. 8).
- [WBH15] Vivian L. WEST, David BORLAND, and W Ed HAMMOND. "Innovative information visualization of electronic health record data: a systematic review". In: *Journal of the American Medical Informatics Association : JAMIA* 22.2 (Mar. 2015), pp. 330–339. ISSN: 1067-5027. DOI: [10.1136/amiajnl-2014-002955](https://doi.org/10.1136/amiajnl-2014-002955) (cit. on pp. 18, 21).
- [Wex+20] James WEXLER, Mahima PUSHKARNA, Tolga BOLUKBASI, Martin WATTENBERG, Fernanda VIÉGAS, and Jimbo WILSON. "The What-If Tool: Interactive Probing of Machine Learning Models". In: *IEEE Transactions on Visualization and Computer Graphics* 26.1 (Jan. 2020), pp. 56–65. ISSN: 1941-0506. DOI: [10.1109/TVCG.2019.2934619](https://doi.org/10.1109/TVCG.2019.2934619) (cit. on p. 81).
- [Whi01] K. Geoffrey WHITE. "Forgetting Functions". In: *Animal Learning & Behavior* 29.3 (Aug. 2001), pp. 193–207. ISSN: 0090-4996, 1532-5830. DOI: [10.3758/BF03192887](https://doi.org/10.3758/BF03192887) (cit. on p. 62).
- [WHM20] Lucas WIRBKA, Walter E. HAEFELI, and Andreas D. MEID. "A framework to build similarity-based cohorts for personalized treatment advice – a standardized, but flexible workflow with the R package SimBaCo". In: *PLOS ONE* 15.5 (May 2020), e0233686. DOI: [10.1371/journal.pone.0233686](https://doi.org/10.1371/journal.pone.0233686) (cit. on p. 66).
- [Wil+18] Jeremy WILDFIRE, Ryan BAILEY, Rebecca Z. KROUSE, Spencer CHILDRESS, Britt SIKORA, Nathan BRYANT, Shane ROSANBALM, Emily WILSON, and Jack G. MODELL. "The Safety Explorer Suite: Interactive Safety Monitoring for Clinical Trials". In: *Therapeutic innovation & regulatory science* 52.6 (Nov. 2018), pp. 696–700. ISSN: 2168-4790. DOI: [10.1177/2168479018754846](https://doi.org/10.1177/2168479018754846) (cit. on p. 75).
- [WL22] Qiru WANG and Robert S. LARAMEE. "EHR STAR: The State-Of-the-Art in Interactive EHR Visualization". en. In: *Computer Graphics Forum* 41.1 (2022), pp. 69–105. ISSN: 1467-8659. DOI: [10.1111/cgf.14424](https://doi.org/10.1111/cgf.14424) (cit. on p. 18).
- [WLG22] Zheyuan WANG, Chen LIU, and Matthew GOMBOLAY. "Heterogeneous Graph Attention Networks for Scalable Multi-Robot Scheduling with Temporospacial Constraints". In: *Autonomous Robots* 46.1 (Jan. 2022), pp. 249–268. ISSN: 1573-7527. DOI: [10.1007/s10514-021-09997-2](https://doi.org/10.1007/s10514-021-09997-2) (cit. on p. 70).
- [Wu+23] Yuchen WU, Yuansong XU, Shenghan GAO, Xingbo WANG, Wenkai SONG, Zhiheng NIE, Xiaomeng FAN, and Quan LI. *LiveRetro: Visual Analytics for Strategic Retrospect in Livestream E-Commerce*. Aug. 2023. DOI: [10.48550/arXiv.2307.12213](https://doi.org/10.48550/arXiv.2307.12213) (cit. on pp. 25–27, 29, 30, 42, 73, 74).
- [Wu+24] Yuzhou WU, Xuechen CHEN, Xin YAO, Yongang YU, and Zhigang CHEN. "Hyperbolic Graph Convolutional Neural Network with Contrastive Learning for Automated ICD Coding". In: *Computers in Biology and Medicine* 168 (Jan. 2024), p. 107797. ISSN: 0010-4825. DOI: [10.1016/j.combiomed.2023.107797](https://doi.org/10.1016/j.combiomed.2023.107797) (cit. on p. 69).

- [Xu+19] Keyulu XU, Weihua HU, Jure LESKOVEC, and Stefanie JEGELKA. “How Powerful Are Graph Neural Networks?” In: *arXiv:1810.00826 [cs, stat]* (Feb. 2019). arXiv: [1810.00826 \[cs, stat\]](https://arxiv.org/abs/1810.00826) (cit. on p. 56).
- [Xu+21] Jie XU, Benjamin S. GLICKSBERG, Chang SU, Peter WALKER, Jiang BIAN, and Fei WANG. “Federated Learning for Healthcare Informatics”. In: *Journal of Healthcare Informatics Research* 5.1 (Mar. 2021), pp. 1–19. ISSN: 2509-498X. DOI: [10.1007/s41666-020-00082-4](https://doi.org/10.1007/s41666-020-00082-4) (cit. on p. 78).
- [YBO18] Carl YANG, Aydin BULUÇ, and John D. OWENS. “Design Principles for Sparse Matrix Multiplication on the GPU”. In: *CoRR abs/1803.08601* (2018). arXiv: [1803.08601](https://arxiv.org/abs/1803.08601) (cit. on p. 79).
- [YCS16] Zhilin YANG, William W. COHEN, and Ruslan SALAKHUTDINOV. “Revisiting Semi-Supervised Learning with Graph Embeddings”. In: *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48. (ICML '16)*. JMLR.org, June 2016, pp. 40–48 (cit. on p. 50).
- [Yin+19] Rex YING, Dylan BOURGEOIS, Jiaxuan YOU, Marinka ZITNIK, and Jure LESKOVEC. “GNN Explainer: A Tool for Post-hoc Explanation of Graph Neural Networks”. In: *CoRR abs/1903.03894* (2019). arXiv: [1903.03894](https://arxiv.org/abs/1903.03894) (cit. on p. 80).
- [Yue+19] Xuanwu YUE, Xinhuan SHU, Xinyu ZHU, Xinnan DU, Zheqing YU, Dimitrios PAPADOPOULOS, and Siyuan LIU. “BitExTract: Interactive Visualization for Extracting Bitcoin Exchange Intelligence”. In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (Jan. 2019), pp. 162–171. ISSN: 1941-0506. DOI: [10.1109/TVCG.2018.2864814](https://doi.org/10.1109/TVCG.2018.2864814) (cit. on pp. 18, 25, 26, 29, 30, 81).
- [YY22] Tianbao YANG et Yiming YING. « AUC Maximization in the Era of Big Data and AI : A Survey ». In : *ACM Computing Surveys* 55.8 (déc. 2022). ISSN : 0360-0300. DOI : [10.1145/3554729](https://doi.org/10.1145/3554729). URL : <https://doi.org/10.1145/3554729> (cf. page 68).
- [ZAL18] Marinka ZITNIK, Monica AGRAWAL, and Jure LESKOVEC. “Modeling polypharmacy side effects with graph convolutional networks”. en. In: *Bioinformatics* 34.13 (July 2018), pp. i457–i466. ISSN: 1367-4803, 1460-2059. DOI: [10.1093/bioinformatics/bty294](https://doi.org/10.1093/bioinformatics/bty294) (cit. on p. 58).
- [ZCZ22] Ziwei ZHANG, Peng CUI, and Wenwu ZHU. “Deep Learning on Graphs: A Survey”. In: *IEEE Transactions on Knowledge and Data Engineering* 34.01 (Jan. 2022), pp. 249–270. ISSN: 1558-2191. DOI: [10.1109/TKDE.2020.2981333](https://doi.org/10.1109/TKDE.2020.2981333) (cit. on p. 54).
- [Zha+11] Jian ZHAO, Fanny CHEVALIER, Emmanuel PIETRIGA, and Ravin BALAKRISHNAN. “Exploratory Analysis of Time-Series with ChronoLenses”. In: *IEEE Transactions on Visualization and Computer Graphics* 17.12 (Dec. 2011), pp. 2422–2431. ISSN: 1941-0506. DOI: [10.1109/TVCG.2011.195](https://doi.org/10.1109/TVCG.2011.195) (cit. on pp. 25, 26, 29).
- [Zha+19] Chuxu ZHANG, Dongjin SONG, Chao HUANG, Ananthram SWAMI, and Nitesh V. CHAWLA. “Heterogeneous Graph Neural Network”. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. (KDD '19)*. Association for Computing Ma-

-
- chinery, July 2019, pp. 793–803. ISBN: 978-1-4503-6201-6. DOI: [10.1145/3292500.3330961](https://doi.org/10.1145/3292500.3330961) (cit. on p. 69).
- [Zha+21] Chen ZHANG, Yu XIE, Hang BAI, Bin YU, Weihong LI, and Yuan GAO. “A Survey on Federated Learning”. In: *Knowledge-Based Systems* 216 (Mar. 2021), p. 106775. ISSN: 0950-7051. DOI: [10.1016/j.knosys.2021.106775](https://doi.org/10.1016/j.knosys.2021.106775) (cit. on p. 78).
- [Zha+22] Chenyang ZHANG, Xiyuan WANG, Chuyi ZHAO, Yijing REN, Tianyu ZHANG, Zhenhui PENG, Xiaomeng FAN, and Quan LI. *PromotionLens: Inspecting Promotion Strategies of Online E-commerce via Visual Analytics*. Aug. 2022. DOI: [10.48550/arXiv.2208.01404](https://doi.org/10.48550/arXiv.2208.01404) (cit. on pp. 25–27, 29, 30, 42).
- [Zha+23] Wei ZHANG, Siwei TAN, Siming CHEN, Linghao MENG, Tianye ZHANG, Rongchen ZHU, and Wei CHEN. “Visual Reasoning for Uncertainty in Spatio-Temporal Events of Historical Figures”. In: *IEEE Transactions on Visualization and Computer Graphics* 29.6 (June 2023), pp. 3009–3023. ISSN: 1941-0506. DOI: [10.1109/TVCG.2022.3146508](https://doi.org/10.1109/TVCG.2022.3146508) (cit. on pp. 18, 25, 26, 29).
- [Zhe+21] Da ZHENG, Minjie WANG, Quan GAN, Xiang SONG, Zheng ZHANG, and George KARYPIS. “Scalable Graph Neural Networks with Deep Graph Library”. In: *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. (WSDM ’21). Association for Computing Machinery, Mar. 2021, pp. 1141–1142. ISBN: 978-1-4503-8297-7. DOI: [10.1145/3437963.3441663](https://doi.org/10.1145/3437963.3441663) (cit. on p. 63).
- [Zho+20] Jie ZHOU, Ganqu CUI, Shengding HU, Zhengyan ZHANG, Cheng YANG, Zhiyuan LIU, Lifeng WANG, Changcheng LI, and Maosong SUN. “Graph Neural Networks: A Review of Methods and Applications”. In: *AI Open* 1 (Jan. 2020), pp. 57–81. ISSN: 2666-6510. DOI: [10.1016/j.aiopen.2021.01.001](https://doi.org/10.1016/j.aiopen.2021.01.001) (cit. on pp. 49, 52, 55).
- [Zho+24] Binggui ZHOU, Guanghua YANG, Zheng SHI, and Shaodan MA. “Natural Language Processing for Smart Healthcare”. In: *IEEE Reviews in Biomedical Engineering* 17 (2024), pp. 4–18. ISSN: 1941-1189. DOI: [10.1109/RBME.2022.3210270](https://doi.org/10.1109/RBME.2022.3210270) (cit. on p. 4).
- [Zhu+19] Jie ZHU, Junhui LI, Muhua ZHU, Longhua QIAN, Min ZHANG, and Guodong ZHOU. “Modeling Graph Structure in Transformer for Better AMR-to-Text Generation”. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Association for Computational Linguistics, Nov. 2019, pp. 5459–5468. DOI: [10.18653/v1/D19-1548](https://doi.org/10.18653/v1/D19-1548) (cit. on p. 76).
- [Zhu+20] Jiong ZHU, Yujun YAN, Lingxiao ZHAO, Mark HEIMANN, Leman AKOGLU, and Danai KOUTRA. “Beyond Homophily in Graph Neural Networks: Current Limitations and Effective Designs”. In: *Advances in Neural Information Processing Systems*. Vol. 33. Curran Associates, Inc., 2020, pp. 7793–7804 (cit. on p. 56).

ANNEXE

Comme nous l'avons présenté dans le Chapitre 2, l'évaluation de GLIMPSE-Med auprès de 14 professionnels du CHU de Montpellier a donné lieu à des retours qualitatifs en réponse à des questions ouvertes. Cette section expose donc une synthèse complète des résultats obtenus.

Analyse qualitative de GLIMPSE-Med

La majorité des participants a apprécié la clarté de la distorsion de l'axe temporel (Fig. 2.3.c2) (10 ont exprimé des sentiments positifs, 4 des sentiments neutres sur 14) ainsi que l'interaction du clic sur une admission pour effectuer un zoom (Fig. 2.4). Parmi les opinions positives, 6 participants ont admis que la distorsion nécessitait un court moment d'adaptation. Un participant a exprimé une réserve : dans certains cas, il y aurait un avantage à conserver une échelle temporelle linéaire, puisque les espaces entre les mesures auraient une sémantique importante. Un autre participant a exprimé des difficultés à identifier la durée d'un séjour patient. Enfin, un participant aurait aimé pouvoir agrandir la vue à une granularité plus fine. 2 participants ont déclaré qu'ils préféreraient que l'interface montre le type d'admission ou l'unité médicale du séjour sur le calendrier linéaire (Fig. 2.3.c1), par le biais d'un code couleur, pour mettre en évidence leur parcours.

Les opinions sur le menu des modalités (Fig. 2.3.b) étaient mitigées. Les participants ont dû effectuer un grand nombre de clics pour accomplir les tâches demandées. Seul un participant a exprimé des sentiments négatifs à propos de l'arborescence, qui contient trop d'éléments. Un participant a apprécié la présence des diagnostics mesurés de tous les séjours, même en effectuant un zoom sur un autre. 2 participants n'ont pas aimé l'arrangement alphabétique des feuilles de l'arborescence et préfèrent une organisation basée sur des thèmes cliniques ou des fonctions médicales, où la « pression diastolique » et la « pression systolique » seraient au même endroit, par exemple. Cependant, les participants ont apprécié la barre de recherche, 4 ayant exprimé une opinion positive et aucune opinion négative ou neutre n'a été exprimée. 2 participants auraient aimé pouvoir sélectionner des modalités pour les visualiser côte à côte.

Aucun participant n'a exprimé d'opinion négative sur la lisibilité des données ou des courbes, un participant a dit qu'il aimait la représentation sous forme de losanges compacts et allongés (Fig. 2.5.a1 et a3). Un participant atteint d'astigmatisme a déclaré qu'il avait du mal à associer les losanges allongés à leur modalité, par exemple, pour trouver des diagnostics. Un participant a souligné et apprécié la différence entre les données numériques (Fig. 2.5.a5) et non numériques (Fig. 2.5.a6).

Un participant a exprimé qu'il aurait préféré un résumé des valeurs élevées et importantes à côté de la modalité, un autre a demandé un résumé statistique par séjour et un autre a demandé un bouton pour afficher tous les valeurs aberrantes. Un participant a contesté le choix des couleurs en affirmant que le vert, le jaune et le rouge sont des choix de couleurs plus courants, mais a reconnu que la charge sémantique était forte et pourrait être inappropriée, et a suggéré d'éclaircir ces couleurs. Aucun participant n'a émis d'objection quant à la lisibilité des couleurs. Un participant atteint d'un daltonisme non identifié a déclaré qu'il pouvait facilement discerner les couleurs. Les valeurs manquantes (Fig. 2.5.a6) sont visibles, mais semblent avoir perturbé certains participants.

Tous les participants ont utilisé avec succès les infobulles (Fig. 2.7). Leur absence pour le score de risque a été regrettée par 2 autres participants, soulignant leur importance. Un participant a exprimé le désir de pouvoir cliquer sur un point pour le rendre persistant, afin de le comparer avec d'autres.

La lisibilité du score de risque (Fig. 2.5.c) et des valeurs importantes (Fig. 2.5.a4) n'a pas été remise en question. Cependant, 2 participants n'ont pas pu comprendre le lien entre les 2, affirmant ainsi qu'ils étaient incapables de voir quels éléments sont impliqués dans le calcul du risque et qu'ils n'étaient pas en mesure d'identifier si une valeur rouge donnée concerne le séjour actuel ou le suivant.

Un participant a déclaré que la mise en place d'une fonction d'exportation de données était essentielle pour une utilisation de l'interface en condition réelle. 4 participants ont exprimé le besoin de visualiser et appréhender des cohortes de patients. Nous n'avons pas approfondi cette question, car la visualisation de trajectoires de patients multiples ou de cohortes constitue un axe de recherche différent. Récemment, ce défi a été étudié dans [Sch+22].

