

Proposition de sujet de thèse

Explicabilité des décisions d'un GNN, application à la chémoinformatique

Résumé

Les laboratoires de recherche GREYC(informatique) et CERMN (pharmacologie) travaillent depuis de nombreuses années au sein d'un groupe commun pour développer des méthodes informatiques innovantes afin de traiter des données liées aux médicaments et notamment d'essayer de prédire et de comprendre le mode d'action d'une nouvelle molécule à partir de connaissances extraites sur un ensemble de molécules.

Les molécules pouvant être modélisées par des graphes, il est possible d'apprendre à un GNN (Graph neural network) à classifier ou prédire l'action de molécules, mais actuellement, aucune explication sur la prise de décision du réseau n'est donnée. L'explicabilité des réseaux est fondamentale dans la prédiction des propriétés thérapeutiques de molécules. En effet, celle-ci permet :

1. de valider les prédictions avant d'engager des efforts et des moyens sur les synthèses d'une série moléculaire
2. d'obtenir une intuition sur les propriétés physico-chimiques clés que doit posséder une molécule pour avoir une action biologique ciblée.

L'objectif de cette thèse est de proposer des solutions pour expliquer les décisions que prend un réseau de neurones opérant sur des graphes en vue de l'appliquer sur des données des molécules chimiques. Aucune connaissance en chimie thérapeutique n'est requise.

Financement : 2100 euros brut, voir <https://www.enseignementsup-recherche.gouv.fr/fr/le-financement-doctoral-46472>

Localisation : Université de Caen Normandie, laboratoire GREYC, équipe CODAG,

Contact : jean-luc.lamotte@at@unicaen.fr

Date limite de candidature : 13 mai 2024

Sujet détaillée : <https://lamotte.users.greyc.fr/files/theseCODAG-1.pdf>

Le dossier de candidature CV, relevé de notes de L3, M1 et M2 (ou niveau équivalent), une lettre de motivation qui précisera l'adéquation entre votre candidature et le sujet, et les coordonnées d'une ou plusieurs personnes de référence pouvant être contactées.

Profil recherché la personne candidate doit être inscrit en dernière année d'un Master ou d'un diplôme d'ingénieur, ou être titulaire d'un tel diplôme, dans un domaine lié à l'informatique ou aux mathématiques appliquées, et posséder de solides compétences en programmation. Une expérience en informatique pour la Science des Données, l'apprentissage profond, ... sera un plus. La personne doit avoir des capacités à rédiger des rapports scientifiques et à

communiquer des résultats de recherche lors de conférences en anglais.

The GREYC and CERMN are two research laboratories working in computer science and pharmacology, respectively. For many years, a common group from both labs has been developing innovative computer methods to process data related to drugs. Their focus is to predict and understand the mode of action of new molecules based on knowledge extracted from a dataset of molecules.

As molecules can be represented by graphs, we can train a Graph Neural Network (GNN) to classify or predict their behaviour. However, the network's decision-making process is currently not explained. The explainability of networks is crucial in predicting the therapeutic properties of molecules. This enables :

- Validate predictions before committing resources to the synthesis of a molecular series.
- Obtaining insight into the key physicochemical properties required for a molecule to have a targeted biological action.

The objective of this thesis is to propose solutions to explain the decisions made by a neural network operating on graphs, with the aim of applying it to chemical molecule data. No knowledge in therapeutic chemistry is required.

Funding : follows current legislation (2100 euros for gross salary), see <https://www.enseignementsup-recherche.gouv.fr/fr/le-financement-doctoral-46472>

Location : University of Caen Normandy, GREYC laboratory, CODAG team

Contact : jean-luc.lamotte@at@unicaen.fr

Application deadline : May 13, 2024

Full subject : <https://lamotte.users.greyc.fr/files/theseCODAG-1.pdf>

Application form : CV, transcripts of L3, M1, and M2 (or equivalent level), a letter of motivation specifying the fit between your application and the topic, and the contact information of one or more references who can be contacted.

Profile sought : The candidate must be enrolled in the final year of a Master's degree or an engineering degree, or hold such a degree, in a field related to computer science or applied mathematics, and possess strong programming skills. Experience in computer science for Data Science, deep learning, ... will be a plus. The individual must have the ability to write scientific reports and communicate research results at conferences in English.

1 Contexte scientifique de la thèse

L'apprentissage profond a révolutionné de nombreux domaines tels que l'analyse d'images ou le traitement automatique des langues. Pourtant, l'incapacité de ces systèmes à justifier leurs décisions est très vite apparue comme une limite forte. Des méthodes permettant de mesurer la saillance de chaque donnée d'entrée ont donc été définies. Ces méthodes peuvent se décomposer en deux grandes familles : Les méthodes dites "boîte blanche" qui supposent le réseau connu et se basent donc sur toutes les informations de celui-ci : topologie, poids appris et gradient calculé par une rétropropagation. Inversement, les méthodes dites "boîte noire", considèrent le réseau comme une fonction inconnue et mesurent l'importance de chaque élément en masquant aléatoirement des parties du signal d'entrée. L'importance de chaque

élément et alors définie comme la moyenne des réponses du système sur l'ensemble des masques sur lesquels l'élément n'est pas masqué.

La chémoinformatique, domaine relatif au traitement de l'information chimique, a été fortement impacté par l'arrivée de l'apprentissage profond et des réseaux convolutionnels (CNN). Un Graph Neural Network (GNN) est l'équivalent d'un CNN opérant sur des graphes. La convolution y est remplacée par une convolution sur graphe et des opérations de pooling peuvent également être définies. Une différence notable entre CNN et GNN se situe au niveau du pooling. Dans les CNN, le pooling opère sur des topologies fixes et utilise des fonctions de réduction généralement non apprises comme max pooling ou average pooling. Dans les GNN, le pooling apprend à la fois des regroupements de sommets de topologie arbitraire et des fonctions de réduction sur ces regroupements.

L'action thérapeutique d'une molécule, dépend de la capacité de celle-ci à se fixer sur les protéines ciblées. Cette capacité de la molécule est déterminée par ses conformations les plus probables ainsi que par la présence de différents groupes fonctionnels, appelés aussi caractéristiques pharmacophoriques, qui vont permettre à la molécule de se fixer. On peut ainsi distinguer, deux types de représentations de molécules pour ce type de prédiction : le graphe squelettique (sommets et arêtes représentent respectivement les atomes et leurs liaisons covalentes) et le graphe des caractéristiques pharmacophoriques où chaque sommet code un groupe fonctionnel de la molécule et la distance (topologique ou topographique) entre deux caractéristiques est codée par une arête. Dans ce cas, le graphe est un un graphe complet avec beaucoup moins de sommets que dans la représentation squelettique [13, 7].

L'explicabilité des réseaux est fondamentale dans la prédiction des propriétés thérapeutiques de molécules. En effet, celle-ci permet :

1. de valider les prédictions avant d'engager des efforts et des moyens sur les synthèses d'une série moléculaire
2. d'obtenir une intuition sur les propriétés physico-chimiques clés que doit posséder une molécule pour avoir une action biologique ciblée.

Les méthodes expliquant l'action des réseaux dans le cadre des graphes se situent dans la prolongation des méthodes conçues pour les CNN. On peut notamment appliquer la distinction boîte noire [19, 23, 15, 15, 8]/boîte blanche [16, 4, 6] mais également subdiviser les approches en méthodes locales (telles que les méthodes précédemment citées), qui fournissent une explication de la prédiction de chaque entrée et les méthodes globales [22, 12, 20] qui tentent de mettre en avant une explication globale des motifs clés pour la prédiction. Notons que les deux approches ne sont pas orthogonales. Par exemple [1] agglomère les explications locales sur un jeu de données pour fournir une explication globale. D'autres critères de décomposition sont également possibles. On peut par exemple décomposer les approches en méthodes factuelles, par exemple [4], qui cherchent des motifs (noeuds, arêtes, sous-graphes) expliquant au mieux la prédiction et des modèles contre-factuels [11, 5] cherchant au contraire des modifications minimales des données d'entrées qui changent la prédiction. Parmi les méthodes factuelles, on peut également distinguer les méthodes qui intègrent l'extraction de motif à la prédiction. Ces méthodes sont donc auto-interprétables [14, 21] et les méthodes qui combinent le modèle de prédiction à une autre méthode pour générer des explications (par ex. [4]). Le lecteur intéressé peut se référer aux surveys suivants pour davantage d'information [10, 9].

Notons toutefois que les méthodes type boîte blanche se sont focalisées sur des réseaux de type GNN sans pooling. L'utilisation du pooling pour l'explicabilité n'a à notre connaissance pas encore été exploré.

2 Sujet et déroulé de la thèse

Le travail commencera par une étude et une comparaison des méthodes permettant d'expliquer les GNNs sur des jeux de données standards et également sur nos jeux d'interaction protéine/ligand à la fois sur les graphes complets et sur les graphes squelettiques. Notons à ce propos que plusieurs critères [9] tels que l'accuracy, l'aire sous la courbe, la fidélité, la parcimonie, . . . , ont été définis.

Sur la base de ces comparaisons, nous pensons proposer nos propres méthodes et notamment explorer l'utilisation de l'étape de pooling [18, 17] pour l'explicabilité. En effet, un pooling appris doit regrouper des sommets importants pour la propriété à prédire ce qui constitue une forme d'expliquabilité encore peu explorée jusqu'à présent. Notre objectif est d'obtenir pour chaque molécule une mesure de l'importance de ses groupes fonctionnels et de leurs distances. Cette étape sera suivie d'une étape de fouille sur l'ensemble de la base, pour identifier les groupes fonctionnels importants situés à des distances similaires.

Ces groupes définiront ce que nous appelons des pharmacophores. Cette étude sera étendue aux graphes squelettiques. Ceci nous permettra de prendre en compte l'ensemble de la molécule et d'apprendre les groupes fonctionnels actifs pour une action biologique ciblée. On étudiera dans ce cadre plus particulièrement la hiérarchie de pooling.

3 Équipe d'encadrement

Cette thèse sera co-dorignée par Jean-Luc Lamotte (équipe CODAG - GREYC) et Luc Brun (équipe Image - GREYC) et co-encadrée par Pierre Héroux (équipe Apprentissage - LITIS). La pertinence de cette équipe d'encadrement pour le sujet proposé est développée plus bas.

4 Profil du candidat

Le candidat doit être inscrit en dernière année d'un Master ou d'un diplôme d'ingénieur, ou être titulaire d'un tel diplôme, dans un domaine lié à l'informatique ou aux mathématiques appliquées, et posséder de solides compétences en programmation. Une expérience en informatique pour la Science des Données, l'apprentissage profond, ... sera un plus. Le candidat doit avoir des capacités à rédiger des rapports scientifiques et à communiquer des résultats de recherche lors de conférences en anglais.

5 Conditions de la thèse

La thèse débutera à l'automne 2024 – début septembre ou début octobre. Le travail se déroulera principalement au GREYC dans les équipes CODAG et IMAGE avec une forte interaction avec le laboratoire LITIS de Rouen sur la partie apprentissage profond et explicabilité et avec le Laboratoire CERMN à Caen pour l'analyse des résultats.

6 Environnement de la thèse

Le GREYC mène une collaboration soutenue avec le laboratoire CERMN autour de la fouille de données appliquée aux molécules thérapeutiques.

Le GREYC entretient par ailleurs une collaboration avec le LITIS sur la thématique de l'apprentissage profond sur graphes. Cette collaboration s'est ces dernières

années concrétisée à travers le projet AGAC (2018-2019) soutenu par la Région Normandie. Le travail commun et complémentaire des deux laboratoires sur la thématique de l'apprentissage profond sur graphes a également été reconnu par la sélection et le financement par l'ANR du projet CoDeGNN (2022-2025) et des publications de très bon niveau qui en ont découlé [3, 2]. Notons que cette thèse pourra renforcer ce projet de recherche commun sur l'aspect chémoinformatique.

La chémoinformatique offre également pour le LITIS un cadre applicatif support de problématiques scientifiques qui ont émané des collaborations qu'il entretient notamment avec le laboratoire de chimie COBRA (UMR-6014). Ainsi, les méthodes développées dans le cadre du présent projet pourront également trouver des applications, par exemple, dans l'identification des caractéristiques moléculaires que doivent observer les polymères (à synthétiser) devant disposer de certaines propriétés physico-chimiques.

Ces collaborations ont permis d'obtenir des résultats de portée internationale. L'apprentissage profond sur graphes fait actuellement l'objet de nombreuses recherches au niveau international.

Cette thèse permettra un rapprochement des équipes du GREYC et du LITIS travaillant sur les GNN en développant un nouvel axe de recherche sur l'explicabilité des décisions des GNN. Cette collaboration GREYC/LITIS, permettra de renforcer l'axe transverse Graphes de NormaSTIC et plus généralement les collaborations promues par la fédération. Pour le CERMN, le travail permettra d'acquérir une nouvelle expertise dans les technologies de Deep Learning, de manière efficace et pleinement adaptée à leurs objets d'étude.

Personnes impliquées dans le suivi du projet :

- GREYC : Jean-Luc Lamotte (co-directeur), Luc Brun (co-directeur), Bertrand Cuissart
- LITIS : Pierre Héroux (co-encadrant), Paul Honeine
- CERMN : Ronan Bureau, Alban Lepailleur

Financement : 2100 euros brut, voir <https://www.enseignementsup-recherche.gouv.fr/fr/le-financement-doctoral-46472>

Localisation : Université de Caen Normandie, laboratoire GREYC, équipe CODAG,

Contact : jean-luc.lamotte@at.unicaen.fr

Date limite de candidature : 13 mai 2024

7 Dossier de candidature

Il est composé d'un curriculum vitae, d'un relevé de notes de L3, M1 et M2 (ou niveau équivalent), d'une lettre de motivation qui précisera l'adéquation entre votre candidature et le sujet, et des coordonnées d'une ou plusieurs personnes de référence pouvant être contactées.

Références

- [1] Steve Azzolin, Antonio Longa, Pietro Barbiero, Pietro Liò, and Andrea Passerini. Global explainability of gnn's via logic combination of learned concepts. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023.

- [2] Muhammet Balcilar, Pierre Héroux, Benoit Gaüzère, Pascal Vasseur, Sébastien Adam, and Paul Honeine. Breaking the Limits of Message Passing Graph Neural Networks. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning (ICML)*, volume 139 of *Proceedings of Machine Learning Research*, pages 599–608, Vienna, Austria, July 2021.
- [3] Muhammet Balcilar, Guillaume Renton, Pierre Héroux, Benoit Gaüzère, Sébastien Adam, and Paul Honeine. Analyzing the expressive power of graph neural networks in a spectral perspective. In *International Conference on Learning Representations*, 2021.
- [4] Federico Baldassarre and Hossein Azizpour. Explainability techniques for graph convolutional networks. *CoRR*, abs/1905.13686, 2019.
- [5] Ziheng Chen, Fabrizio Silvestri, Jia Wang, Yongfeng Zhang, Zhenhua Huang, Hongshik Ahn, and Gabriele Tolomei. GREASE : generate factual and counterfactual explanations for gnn-based recommendations. *CoRR*, abs/2208.04222, 2022.
- [6] Qizhang Feng, Ninghao Liu, Fan Yang, Ruixiang Tang, Mengnan Du, and Xia Hu. DEGREE : decomposition based explanation for graph neural networks. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022.
- [7] Damien Geslin, Alban Lepailleur, Jean-Luc Manguin, Nhat-Vinh Vo, Jean Luc Lamotte, Bertrand Cuissart, and Ronan Bureau. Deciphering a pharmacophore network : A case study using BCR-ABL data. *J. Chem. Inf. Model.*, 62(3) :678–691, 2022.
- [8] Qiang Huang, Makoto Yamada, Yuan Tian, Dinesh Singh, and Yi Chang. Graphlime : Local interpretable model explanations for graph neural networks. *IEEE Trans. Knowl. Data Eng.*, 35(7) :6968–6972, 2023.
- [9] Jaykumar Kakkad, Jaspal Jannu, Kartik Sharma, Charu C. Aggarwal, and Sourav Medya. A survey on explainability of graph neural networks. *CoRR*, abs/2306.01958, 2023.
- [10] Yiqiao Li, Jianlong Zhou, Sunny Verma, and Fang Chen. A survey of explainable graph neural networks : Taxonomy and evaluation metrics. *CoRR*, abs/2207.12599, 2022.
- [11] Ana Lucic, Maartje A. ter Hoeve, Gabriele Tolomei, Maarten de Rijke, and Fabrizio Silvestri. Cf-gnnexplainer : Counterfactual explanations for graph neural networks. In Gustau Camps-Valls, Francisco J. R. Ruiz, and Isabel Valera, editors, *International Conference on Artificial Intelligence and Statistics, AISTATS 2022, 28-30 March 2022, Virtual Event*, volume 151 of *Proceedings of Machine Learning Research*, pages 4499–4511. PMLR, 2022.
- [12] Lucie Charlotte Magister, Dmitry Kazhdan, Vikash Singh, and Pietro Liò. Gcexplainer : Human-in-the-loop concept-based explanations for graph neural networks. *CoRR*, abs/2107.11889, 2021.
- [13] Jean-Philippe Métivier, Bertrand Cuissart, Ronan Bureau, and Alban Lepailleur. The pharmacophore network : A computational method for exploring structure–activity relationships from a large chemical data set. *Journal of Medicinal Chemistry*, 61(8) :3551–3564, 04 2018.
- [14] Siqi Miao, Mia Liu, and Pan Li. Interpretable and generalizable graph learning via stochastic attention mechanism. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato, editors, *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore*,

- Maryland, USA, volume 162 of *Proceedings of Machine Learning Research*, pages 15524–15543. PMLR, 2022.
- [15] Tamara A. Pereira, Erik Nascimento, Lucas E. Resck, Diego Mesquita, and Amauri A. Souza. Distill n’ explain : explaining graph neural networks using simple surrogates. In Francisco J. R. Ruiz, Jennifer G. Dy, and Jan-Willem van de Meent, editors, *International Conference on Artificial Intelligence and Statistics, 25-27 April 2023, Palau de Congressos, Valencia, Spain*, volume 206 of *Proceedings of Machine Learning Research*, pages 6199–6214. PMLR, 2023.
- [16] Phillip E. Pope, Soheil Kolouri, Mohammad Rostami, Charles E. Martin, and Heiko Hoffmann. Explainability methods for graph convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 10772–10781. Computer Vision Foundation / IEEE, 2019.
- [17] Stevan Stanovic, Benoit Ga"uz'ere, and Luc Brun. Maximal Independent Vertex Set applied to Graph Pooling. In *Structural and Syntactic Pattern Recognition (SSPR)*, Montr'eal, Canada, August 2022.
- [18] Stevan Stanovic, Benoit Ga"uz'ere, and Luc Brun. Maximal independent sets for pooling in graph neural networks. In Mario Vento, Pasquale Foggia, Donatello Conte, and Vincenzo Carletti, editors, *Graph-Based Representations in Pattern Recognition - 13th IAPR-TC-15 International Workshop, GBRPR 2023, Vietri sul Mare, Italy, September 6-8, 2023, Proceedings*, volume 14121 of *Lecture Notes in Computer Science*, pages 113–124. Springer, 2023.
- [19] Minh N. Vu and My T. Thai. Pgm-explainer : Probabilistic graphical model explanations for graph neural networks. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors, *Advances in Neural Information Processing Systems 33 : Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020.
- [20] Han Xuanyuan, Pietro Barbiero, Dobrik Georgiev, Lucie Charlotte Magister, and Pietro Liò. Global concept-based interpretability for graph neural networks via neuron analysis. In Brian Williams, Yiling Chen, and Jennifer Neville, editors, *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023*, pages 10675–10683. AAAI Press, 2023.
- [21] Junchi Yu, Tingyang Xu, Yu Rong, Yatao Bian, Junzhou Huang, and Ran He. Graph information bottleneck for subgraph recognition. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021.
- [22] Hao Yuan, Jiliang Tang, Xia Hu, and Shuiwang Ji. XGNN : towards model-level explanations of graph neural networks. In Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash, editors, *KDD '20 : The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, pages 430–438. ACM, 2020.
- [23] Yue Zhang, David DeFazio, and Arti Ramesh. Relex : A model-agnostic relational model explainer. In Marion Fourcade, Benjamin Kuipers, Seth Lazar, and Deirdre K. Mulligan, editors, *AIES '21 : AAAI/ACM Conference on AI, Ethics, and Society, Virtual Event, USA, May 19-21, 2021*, pages 1042–1049. ACM, 2021.