

École Nationale Supérieure d'Informatique et d'Analyse des Systèmes
Centre d'Études Doctorales en Sciences des Technologies de l'Information et de l'Ingénieur

AVIS DE SOUTENANCE DE THÈSE DE DOCTORAT

Madame Sofia Bourhim

Soutiendra publiquement sa thèse de Doctorat en Informatique

Le Jeudi 21 Décembre à 10h00 au Grand Amphi à l'ENSIAS de Rabat

Intitulé de la thèse

**DGCF: A Graph-Based Deep Collaborative Filtering Approach For
Recommender Systems**

Président :

Pr. Karim Baina, PES, ENSIAS, Université Mohammed V de Rabat

Directeurs de thèse :

Pr. Lamia Ben Hiba, PH, ENSIAS, Université Mohammed V de Rabat

Co-Encadrant :

Pr. Mohamed Abdou Janati Idrissi, Ex PES, Expert

Rapporteurs :

Pr. Mostafa Bellafkih, PES, Institut National des postes et Télécommunications, Rabat

Pr. Houda Benbrahim, PES, ENSIAS, Université Mohammed V de Rabat

Pr. Mohamed Kissi, PES, FSTM, Université Hassan II, Casablanca

Examineur :

Pr. Laila Benhlime, PES, EMI, Université Mohammed V de Rabat

Invité :

Dr. Fatima Zohra Daha, Expert, Microsoft, USA

Abstract: In an era characterized by an overwhelming abundance of information, online service providers, including prominent platforms like Amazon, Netflix, and YouTube, are constantly engaged in a competitive struggle to capture users' attention. Hence, providing personalized experiences is crucial for their e-market platforms. By facilitating tailored user experiences, Recommendation Systems (RS) contribute to the enhancement of customer loyalty, satisfaction, and lifetime value. Extensive research efforts have been dedicated to the study of recommender systems for e-commerce applications over the years. However, there remains scope for further advancement in the implementation and investigation of recommender system approaches. Within the context of this thesis, we aim to contribute to the refinement of RS models with the overarching objective of effectively extracting and recommending the most precise items to users by leveraging graph representation techniques.

In detail, we propose a Deep Graph-based Collaborative Filtering recommendation system (DGCF), which incorporates the concept of community profiling and leverages the power of Graph Neural Networks (GNN). DGCF utilizes multiple graphs to exploit different information from the user interactions. It extracts the overlapping communities from the homophily user-user graph and also integrates the high-order information from the user-item bipartite graph. DGCF contains three main key components. The first layer is a Community Encoding layer (CE) that encodes latent information based on the user-user similarities graph by extracting overlapping sub-structures of the graph. The second is a Bipartite Graph Convolutional Networks encoder (EB-GCN) that generates embeddings of users and items and captures the collaborative signal in the bipartite graph of user-item interaction. The outputs of both layers are federated in an information fusion layer (IF) that aggregates the embeddings from different perspectives.

We assess our model on two main applications: movie recommendations and drug discovery. In the first application, we use three main datasets with different sizes: MovieLens Datasets (ML-100K and ML-1M) and Douban dataset. For the second application, we use DrugBank and FoodDB datasets. We base our evaluation on the effectiveness of the RS and its performance compared to the latest benchmark reference models and we investigate the viability and benefit from incorporating the CE layer into our model. DGCF excels in accuracy, outperforming other recommendation models and enhancing user recommendations.

Keywords: Graph Neural Networks, Graph Representation Learning, Recommendation Systems, Collaborative Filtering, Community Profiling, Link Prediction, Drug Discovery, PyTorch.

Résumé: À l'ère de la surcharge d'informations, les fournisseurs de services en ligne tels qu'Amazon, Netflix, YouTube, etc. se retrouvent en compétition constante pour capter l'attention des utilisateurs. Offrir à leurs clients une expérience personnalisée grâce à des systèmes de recommandation (RS) est donc devenu un élément essentiel. Depuis plusieurs années, des efforts considérables ont été déployés pour la recherche sur les systèmes de recommandation pour les applications e-commerce. Malgré cela, il existe encore des possibilités d'amélioration dans la mise en œuvre et la recherche de nouvelles approches de systèmes de recommandation.

Dans cette thèse, nous présentons notre contribution pour améliorer les modèles de RS dans le but général d'extraire et de recommander à l'utilisateur un contenu plus précis en utilisant la représentation par graphe. Cette prévalence est due au pouvoir de représenter les interactions (liens) entre différentes entités (noeuds).

En détail, nous proposons une nouvelle méthode de recommandation par filtrage collaboratif basé sur les graphes (DGCF), qui intègre le concept de profilage de la communauté et exploite la puissance des réseaux de neurones spécifique aux graphes (GNN). DGCF extrait les communautés qui se chevauchent du graphe user-user et intègre également les informations du graphe biparti user-item. DGCF contient trois principaux composants clés. La première couche est une couche d'encodage communautaire (CE) qui extrait les sous-structures chevauchantes du graphe. La seconde est un encodeur de réseaux convolutifs de graphes biparties (EB-GCN) qui capture le signal de collaboration dans le graphe user-item. Les sorties des deux couches sont fédérées dans une couche de fusion d'informations (IF) qui agrège les embeddings de différentes perspectives.

Nous évaluons notre modèle sur deux applications principales: les recommandations de films et la découverte de médicaments. Dans la première application, nous utilisons trois datasets de tailles différentes: MovieLens (ML-100K et ML-1M) et Douban, et pour la deuxième application, nous utilisons DrugBank et FoodDB. En outre, nous effectuons une évaluation approfondie de la qualité



جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

des communautés dans le composant CE et nous étudions la viabilité d'inclure la couche CE dans notre modèle. En particulier, nous montrons que DGCF surpasse tous les autres modèles de recommandation en termes de précision et améliore ainsi les recommandations aux utilisateurs.

Mots-clés: Réseaux de neurones spécifique aux graphes, Apprentissage de la représentation sous forme de Graphe, Systèmes de recommandation, Filtrage collaboratif, Profilage communautaire, Prédiction des liens, Découverte de médicaments, PyTorch.