

Nom et prénom	Etablissement	Grade	Qualité
TAIRI Hamid	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Président
BERTELLE Cyrille	Université Le Havre Normandie, France	PES	Rapporteur & Examineur
TROCAN Maria	Institut Supérieur d'Électronique de Paris, France	PES	Rapporteuse & Examinatrice
SABRI My Abdelouahed	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Rapporteur & Examineur
BENNANI Younes	Université Sorbonne Paris Nord, France	PES	Examineur
MAHRAZ Mohamed Adnane	Faculté des Sciences Dhar EL Mahraz, Fès	MCH	Examineur
RAMADAN Hiba	Faculté des Sciences Dhar EL Mahraz, Fès	MCH	Examinatrice
ROGOVSCHI Nicoleta	Université Paris Cité, France	MCH	Co-directrice de thèse
MALEK Maria	CY Cergy Paris Université	MCH	Directrice de thèse
YAHYAOUY Ali	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Directeur de thèse



Résumé :

À l'ère de la numérisation généralisée, les données sont non seulement abondantes, mais également intrinsèquement multimodales, hétérogènes et relationnelles, englobant les comportements utilisateurs, les réseaux sociaux, les contenus visuels, les séquences temporelles et les attributs structurés. Concevoir des systèmes intelligents capables d'apprendre à partir de cette complexité soulève des défis majeurs pour les modèles de recommandation modernes, en particulier en termes de passage à l'échelle, de robustesse et d'explicabilité. Cette thèse propose un cadre unifié qui intègre l'apprentissage de représentations multimodales, la modélisation de graphes hétérogènes, le raisonnement topologique et la modélisation séquentielle du comportement utilisateur, le tout dans une approche fondée sur l'apprentissage sur graphes.

Les graphes sont omniprésents pour représenter les données du monde réel, faisant de l'apprentissage de représentations sur graphes et des réseaux de neurones graphes (GNNs) des piliers des avancées récentes en intelligence artificielle. Cependant, malgré leur efficacité démontrée, les GNNs présentent encore des limites lorsqu'il s'agit de gérer les déséquilibres entre modalités, les décalages sémantiques, l'hétérogénéité structurelle, ou encore les effets de sur-lissage. Cette thèse développe une stratégie de recherche cohérente pour surmonter ces obstacles en s'appuyant sur des architectures de modèles innovantes et l'intégration de la topologie algébrique (TDA).

La première contribution est un modèle appelé MM-HGNN (Multimodal Heterogeneous Graph Neural Network), qui introduit un mécanisme d'attention hiérarchique au niveau des nœuds et des arêtes, ainsi qu'une fonction de transférabilité de modalité permettant d'aligner dynamiquement les caractéristiques hétérogènes issues de différentes modalités. Ce modèle capture les dépendances intra-type et inter-type et conserve l'intégrité sémantique via des connexions résiduelles multi-niveaux, renforçant ainsi la robustesse tout en atténuant le sur-lissage typique des GNNs profonds.

La deuxième contribution de cette thèse est GMSRec, un cadre de recommandation séquentielle multimodale basé sur les graphes, qui combine des transformations guidées par des méta-chemins (meta-paths) préservant les relations sémantiques de haut niveau entre entités hétérogènes, avec un alignement probabiliste des modalités via la divergence de Kullback-Leibler. L'évolution des préférences des utilisateurs est modélisée à l'aide d'encodeurs GRU traitant les séquences d'interactions, permettant des recommandations personnalisées, sensibles au contexte et robustes face aux situations de cold-start.

La troisième contribution introduit TDA-MMRec, un cadre de recommandation multimodale conscient de la topologie, qui exploite l'homologie persistante pour extraire des descripteurs topologiques stables et invariants à l'échelle à partir de graphes multimodaux. Ces signatures structurelles globales sont intégrées dans le processus d'apprentissage des représentations et alimentent une stratégie de pruning guidée par la topologie. En identifiant et supprimant les nœuds structurellement redondants sans compromettre les caractéristiques topologiques essentielles, TDA-MMRec améliore la scalabilité, renforce la robustesse face au bruit et à la rareté des données, et permet une meilleure interprétabilité des graphes appris.

Outre ces contributions principales, cette thèse présente également deux études exploratoires indépendantes, portant sur l'optimisation des stations de recharge pour véhicules électriques



dans des environnements urbains dynamiques, en utilisant respectivement les approches d'apprentissage par renforcement profond et Multi-Agent Deep Deterministic Policy Gradient. Intégrées en annexe, ces études prolongent la thématique de la décision intelligente, en montrant comment des stratégies de contrôle par apprentissage renforcé peuvent améliorer la coordination, l'adaptabilité et l'efficacité dans les systèmes de mobilité intelligente.

Grâce à un développement théorique rigoureux et à une validation empirique étendue sur plusieurs jeux de données à grande échelle, cette thèse apporte des innovations méthodologiques et des contributions pratiques qui consolident les fondations de l'intelligence artificielle fondée sur les graphes. Elle démontre comment l'interaction entre signaux multimodaux, structures de graphe et géométrie topologique permet de concevoir des systèmes de recommandation plus robustes, évolutifs et interprétables. Ce travail ouvre la voie à de futures recherches en recommandation centrée sur l'humain, informée par la topologie et ancrée dans des représentations sémantiques riches et fiables.

Mots clés :

Systèmes de recommandation, Apprentissage multimodal, Graphes hétérogènes, Réseaux de neurones de graphes, Analyse topologique des données (TDA), Modélisation séquentielle, Apprentissage par renforcement profond, Mobilité intelligente.



Multimodal Representation Learning and Topological Graph Neural Architectures for Recommendation

Abstract :

In the era of pervasive digitalization, data is not only abundant but also inherently multimodal, heterogeneous, and relational spanning user behaviors, social networks, visual content, temporal sequences, and structured attributes. Building intelligent systems that can learn from such complexity poses significant challenges for modern recommendation models, especially in achieving scalability, robustness, and explainability. This thesis addresses these challenges by proposing a unified framework that integrates multimodal representation learning, heterogeneous graph modeling, topological reasoning, and sequential behavior modeling all under the broader umbrella of graph-based machine learning.

Graphs are ubiquitous in representing real-world data, making graph representation learning and graph neural networks (GNNs) central to recent machine learning advances. However, while GNNs have demonstrated powerful capabilities for learning on graph-structured data, their performance is often hindered by limitations in handling modality imbalance, semantic misalignment, structural heterogeneity, and over-smoothing. This thesis proposes a cohesive research agenda that overcomes these barriers through novel model architectures and the integration of topological data analysis (TDA).

The first core contribution is a Multimodal Heterogeneous Graph Neural Network (MM-HGNN) that introduces hierarchical node- and edge-level attention, along with a novel Modality Transferability Function to dynamically weigh and align multimodal heterogeneous features.

This model captures both intra-type and inter-type dependencies and preserves semantic integrity via multi-level residual connection, enhancing robustness and mitigating over-smoothing in GNN propagation.

The second contribution of this thesis is a graph-based multimodal sequential recommendation (GMSRec) framework that unifies meta-path-based graph transformations preserving high-order semantic relationships among heterogeneous entities with probabilistic multimodal alignment via KL divergence. To model evolving user interests, GRU-based encoders process user interaction sequences, enabling personalized and context-aware video recommendations, especially under cold-start conditions.

The third core contribution is a topology-aware multimodal recommendation (TDA-MMRec) framework that incorporates persistent homology to extract stable and scale-invariant topological descriptors from multimodal graphs. These global structural signatures are integrated into the representation learning process and inform a topology-guided graph pruning strategy. By identifying and removing structurally redundant nodes without compromising critical topological features, TDA-MMRec enhances scalability, improves robustness to noise and sparsity, and offers interpretable graph structures.

In addition to these core contributions, this dissertation presents two independent research explorations that apply reinforcement learning specifically Deep Q-Network and



Multi-Agent Deep Deterministic Policy Gradient to the optimization of electric vehicle charging station selection in dynamic urban environments. These studies, included as appendices, extend the broader theme of intelligent decision-making by demonstrating how learning-based control strategies can enhance scalability, adaptability, and coordination in smart mobility systems.

Through theoretical development and extensive empirical validation across multiple large-scale datasets, this dissertation offers methodological innovations and empirical insights that strengthen the foundations of graph-based artificial intelligence. It demonstrates how the interplay of multimodal signals, graph structure, and topological geometry can lead to more robust, scalable, and explainable recommendation systems. The presented work lays a strong foundation for future research in human-centered, topology-informed, and semantically grounded recommendation technologies.

Key Words :

Recommender System, Multimodal Learning, Heterogeneous Graph Neural Networks, Topological Data Analysis, Sequential Modeling, Deep Reinforcement Learning, Smart Mobility