المملكة المغربية
وزارة التعليم العالي
و البحث العلمي و الإبتكار

جامعة ابن طفيل
Ibn Tofaïl University

AVIS DE SOUTENANCE D'UNE THESE DE DOCTORAT

Le Directeur de l'Ecole Nationale des Sciences Appliquées a le plaisir d'informer le public qu'une soutenance de thèse de Doctorat en

«Sciences et ingénierie»

aura lieu le 13/09/2025 à l'ENSA de Kénitra

La Thèse sera présentée par Mr TOUARSI MOUAD

Sous le thème :

Conception des réseaux de neurones hybrides pour la classification et l'implémentation de données mixtes

Devant le jury composé de :

Nom et Prénom	Titre	Etablissement
BOUJIHA TARIK	Président	ENSA, Kénitra
SBAA MOHAMMED	Rapporteur	CRMEF, Rabat - Salé - Kénitra
LAARIBI AZIZ	Rapporteur	Faculté Polydisciplinaire, Béni Mellal
GAHI YOUSSEF	Rapporteur	ENSA, Kénitra
TAKI REGRAGUI	Examinateur	FSJES, Salé
ERRACHID MOHAMED	Examinateur	CRMEF, Rabat - Salé - Kénitra
EL OUADI ABDELMAJID	Co-Directeur de thèse	ENSA, Kénitra
GRETETE DRISS	Directeur de thèse	ENSA, Kénitra

UNIVERSITE IBN TOFAIL
PÔLE DES ETUDES DOCTORALES
KENITRA

رئاسة جامعة ابن طفيل، المركب الجامعي، ص.ب 242 - القنيطرة - 14000
Presidency Ibn Tofail University, University Campus, BP. 242 Kénitra 14000
(+212) 537 32 92 00 (+212) 79 45 21 30 05 UIT.ac.ma

ⵜⴰⴳⵍⴷⵉⵜ ⵏ ⵍⵎⵖⵔⵉⴱ
ⵜⴰⵎⴰⵡⴰⵙⵜ ⵏ ⵓⵙⵙⵍⵎⴷ ⴰⵏⴰⴼⵍⵍⴰ
ⴷ ⵓⵔⵣⵣⵓ ⴰⵎⴰⵙⵙⴰⵏ ⴷ ⵓⵙⵏⴼⵍⵓⵍ

المملكة المغربية
وزارة التعليم العالي
و البحث العلمي و الإبتكار

جامعة ابن طفيل
ⵜⴰⵙⴷⴰⵡⵉⵜ ⵉⴱⵏ ⵟⵓⴼⴰⵢⵍ
Ibn Tofaïl University

Nom et Prénom : TOUARSI MOUAD
Date de soutenance : 13/09/2025
Directeur de Thèse : GRETETE DRISS

Sujet de thèse :

Conception des réseaux de neurones hybrides pour la classification et l'implémentation de données mixtes

Résumé:

L´objectif de la thèse est de développer des réseaux de neurones adaptées pour les taches de classification et traitement des données manquantes.

Cette recherche propose une méthode novatrice pour optimiser les cartes auto-organisatrices (SOM) de Kohonen afin de classer des données mixtes (numériques, intervalle et catégorielles). Traditionnellement, les SOM traitent principalement des données numériques, ce qui limite leur efficacité sur des ensembles de données hétérogènes. L'approche suggérée améliore les SOM en intégrant l'algorithme ASAICC (Adaptive Selection Algorithm for Initial Cluster Centers) pour une initialisation optimale des vecteurs de référence, en tenant compte de la densité et de la dissimilarité. De plus, la méthode utilise plusieurs vecteurs de référence par cluster et la computation de distances partielles pour mieux gérer les données mixtes. La méthode développée ne nécessite pas un encodage des données.

L'efficacité de cette méthode est validée par des ensembles de données de référence, démontrant des performances supérieures aux méthodes traditionnelles en termes de précision et de robustesse contre le bruit et les valeurs aberrantes. Cette étude ouvre des perspectives prometteuses pour le regroupement des données mixtes, en soulignant le potentiel des modèles de réseaux neuronaux avancés et des techniques de clustering hybrides.

Nous avons également mené une étude comparative des méthodes d'imputation couramment utilisées et développé une version améliorée des réseaux de neurones adaptatifs. Ces approches ont été implémentées dans l'environnement R.

Abstract:

This thesis aims to develop adaptive neural networks specifically designed for classification tasks and missing data handling

We introduced a novel method to optimize Kohonen's Self-Organizing Maps (SOM) for clustering mixed data types, including numerical, interval, and categorical attributes. Traditional SOMs are primarily designed for numerical data, limiting their effectiveness with heterogeneous datasets. The proposed approach enhances SOMs by integrating the Adaptive Selection Algorithm for Initial Cluster Centers (ASAICC) to initialize referent vectors optimally. ASAICC selects initial vectors based on density and dissimilarity, ensuring a robust starting point for clustering.

The effectiveness of this enhanced SOM is validated using well-known benchmark datasets, demonstrating superior performance compared to traditional methods. The results show improved clustering accuracy and robustness against noise and outliers.

The study highlights the advantages of this approach, emphasizing its potential for various applications in data clustering. Future research directions include integrating advanced neural network models, developing ensemble methods, and exploring scalability with distributed computing frameworks. This research offers a promising direction for mixed data clustering, leveraging the strengths of SOMs and advanced initialization techniques.

We also compared commonly used imputation methods and developed an enhanced version of adaptive neural networks. These approaches were implemented in the R environment.