AVIS DE SOUTENANCE DE THESE

Le Doyen de la Faculté des Sciences Dhar El Mahraz –Fès – annonce que

Mr **BARRY Khalidou Abdoulaye**
Soutiendra : **le Jeudi 10/07/2025** à **10H00**
Lieu : **FSDM – *Centre Visioconférence***

Une thèse intitulée :

« Big Data Analytics in angina pectoris prediction (Heart Attack) »

En vue d'obtenir le ***Doctorat***

FD : ***Sciences et Techniques***
Spécialité : ***Sciences des Données***

Devant le jury composé comme suit :

Nom et prénom	Etablissement	Grade	Qualité
ZINEDINE AHMED	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Président
MERRAS MOSTA	Ecole Supérieure de Technologie, UMI, Meknès	MCH	Rapporteur
JARRAR Abdessamad	Ecole Nationale des Sciences Appliquées, Berrechid, Université Hassan 1er	MCH	Rapporteur
EL GAROUANI SAID	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Rapporteur
BENNANI Mohamed TAJ	Faculté des Sciences Dhar EL Mahraz, Fès	MCH	Examinateur
EL FAZAZY Khalid	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Examinateur
RAMADAN Hiba	Faculté des Sciences Dhar EL Mahraz, Fès	MCH	Invitée
EL FAR Mohamed	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Directeur de thèse

Résumé :

Les maladies cardiaques demeurent l'une des principales causes de mortalité dans le monde. Cette thèse présente cinq contributions majeures exploitant l'apprentissage automatique afin d'améliorer la prédiction, la classification et la gestion des maladies cardiovasculaires (MCV).

La première contribution consiste à améliorer la technique de la forêt aléatoire en y intégrant des règles d'association.
La deuxième associe l'optimisation par essaims particulaires (PSO) à une version modifiée de la forêt aléatoire (MRF), dans le but d'optimiser les performances du modèle classique.
La troisième propose un modèle pondéré de k-plus proches voisins (k-NN), dans lequel les scores des caractéristiques sont utilisés comme poids. Ce modèle a surpassé l'algorithme k-NN traditionnel, atteignant une précision élevée sur cinq ensembles de données agrégés.

La quatrième contribution porte sur le développement d'un ensemble de modèles k-NN, chacun étant entraîné sur les attributs les plus pertinents d'un jeu de données spécifique. Par ailleurs, une application web conviviale intégrant un chatbot a été développée afin de sensibiliser le public et de faciliter le dépistage des maladies cardiaques.

Enfin, la cinquième contribution met en œuvre une architecture composée d'une couche de convolution unidimensionnelle, d'une couche de mémoire à long terme (Long Short-Term Memory), ainsi que d'un mécanisme d'attention pour la classification binaire des battements de cœur normaux et anormaux à partir de la base de données MIT-BIH Arrhythmia. Cette approche a permis d'atteindre des niveaux remarquables de précision et de fiabilité selon diverses métriques d'évaluation. Des règles d'association ont également été employées pour interpréter les caractéristiques extraites de cette base de données.

En somme, ces contributions font progresser l'application de l'apprentissage automatique dans le domaine de la prédiction et du diagnostic des MCV, offrant ainsi aux professionnels de santé des outils innovants pour une détection précoce et une meilleure prise en charge.

Mots clés :
Maladie cardiovasculaire ; Règles d'association ; Apprentissage automatique; Classification ; Sélection de caractéristiques ; Signal ECG ; Chatbot.

BIG DATA ANALYTICS IN ANGINA PECTORIS PREDICTION (HEART ATTACK)

Abstract:

Heart diseases remain one of the leading causes of mortality worldwide.
This thesis presents five major contributions leveraging machine learning to enhance the prediction, classification, and management of cardiovascular diseases (CVDs).

The first contribution improves the random forest technique by integrating association rules. The second combines particle swarm optimization (PSO) with a modified random forest (MRF) to optimize the performance of the standard random forest model.
The third proposes a weighted k-nearest neighbors (k-NN) model, where feature scores are used as weights. This model outperformed the traditional k-NN algorithm, achieving high accuracy across five aggregated datasets.

The fourth contribution involves the development of an ensemble of k-NN models, each focusing on the most relevant attributes of a specific dataset. Additionally, a user-friendly web application incorporating a chatbot was developed to raise awareness and facilitate the screening of heart diseases.

Finally, the fifth contribution applies an architecture consisting of a one-dimensional convolutional layer, a Long Short-Term Memory (LSTM) layer, and an attention mechanism for binary classification of normal and abnormal heartbeats using the MIT-BIH Arrhythmia dataset. This approach yielded outstanding accuracy and reliability across various evaluation metrics. Furthermore, association rules were used to interpret the features extracted from this dataset.

In summary, these contributions advance the application of machine learning in the prediction and diagnosis of CVDs, providing healthcare professionals with innovative tools for early detection and improved treatment.

Key Words:
Cardiovascular disease; Association rules; Machine Learning; Classification; Feature Selection; ECG signal; Chatbot;