

AVIS DE SOUTENANCE THESE DE DOCTORAT

Présentée par

Mlle: HARCHLI FIDAE

Spécialité : Recherche opérationnelle et informatique

Sujet de la thèse : Méthodes probabilistes et estimation des paramètres, contribution à la sélection de modèles neuronaux de type : cartes topologiques probabilistes et réseaux multicouches, application aux domaines réels.

Formation Doctorale : Sciences de l'ingénieur Sciences Physiques, Mathématiques et Informatique.

Thèse présentée et soutenue le mardi 04 juillet 2017 à 14h au centre de conférence devant le jury composé de :

Nom Prénom	Titre	Etablissement	
Ahmed EL HILALI ALAOU	PES	Faculté des Sciences et Techniques de Fès	Président
Mohamed BAHAJ	PES	Faculté des Sciences et Techniques de Settat	Rapporteur
Abdellatif EL AFIA	PES	ENSIAS de Rabat	Rapporteur
Fatima EZZAKI	PES	Faculté des Sciences et Techniques de Fès	Rapporteur
Youssef GHANOU	PH	Ecole Supérieure de Technologie de Meknès	Examineur
Mohammed ETTAOUIL Mohamed MESRAR	PES PES	Faculté des Sciences et Techniques de Fès Faculté des Sciences et Techniques de Fès	Directeurs de thèse

Laboratoire d'accueil : Modélisation et calcul scientifique.

Etablissement : Faculté des Sciences et Techniques de Fès.

Titre de la thèse : Méthodes probabilistes et estimation des paramètres, contribution à la sélection de modèles neuronaux de type : cartes topologiques probabilistes et réseaux multicouches, application aux domaines réels.

Nom du candidat : Fidae HARCHLI

Spécialité : Mathématiques appliquées

Résumé de la thèse

Dans la dernière décennie, l'Internet a été envahi par une quantité immense de services. Par conséquent, plusieurs information et produits sont créés, distribués, évalués et placés en ligne. Ainsi, l'accès à l'information désirée dans un temps convenable est une tâche assez fastidieuse. La fouille de données est la discipline la plus répandue pour extraire l'information pertinente dissimulée dans les bases de données. Plusieurs méthodes dédiées à cette discipline souffrent du problème de choix du modèle. Attiré par le problème de sélection des hyper-paramètres des modèles neuronaux de types PMC, PrSOM et de la méthode de k-means, nous avons développé de nouvelles méthodes capables de surmonter cette problématique. Les solutions proposées concernent le développement de nouvelles approches à base de modèles d'apprentissage artificiel dans un formalisme de sélection des hyper-paramètres. Le premier modèle décrit une nouvelle méthode dédiée à la sélection d'architecture neuronale du réseau probabiliste PrSOM. Cette approche s'inspire des méthodes hiérarchiques de sélection de la partition optimale d'un ensemble de données. Le second modèle traite le même problème de sélection d'architecture pour les réseaux neuronaux de type PMC. Dans ce travail, nous modélisons le problème sous forme d'un modèle non linéaire à variables mixtes résolu par les algorithmes génétiques. La troisième et quatrième contributions décrivent un nouveau formalisme de sélection de paramètres de l'algorithme de k-means pour deux objectifs différents. Deux variantes sont développées : une approche hiérarchique de décomposition de l'ensemble de données pour la sélection de l'ensemble d'apprentissage du PMC et une approche de partitionnement pour la sélection de variables et la réduction de dimensionnalité. A l'aide de ces nouvelles méthodes, nous avons traité deux thématiques de recherche dans le domaine des applications réelles qui font face, à leur tour, à l'explosion de la masse de données : la compression de la parole et la réduction de la dimensionnalité.

La diversité des méthodes proposées pour la sélection de paramètres, représentent des points forts qui caractérisent notre travail de recherche. Les résultats obtenus après la validation des approches proposées dans cette étude sont encourageants et prometteurs pour la classification et la modélisation.

Mots clés : Extraction de connaissance, Apprentissage artificiel, Clustering, Classification, Réseaux connexionistes probabilistes, Estimation de paramètres, Sélection d'architectures neuronales, Nombre de classes, Sélection de meilleure partition, Réseau PrSOM, k-means, Réseau PMC, Compression de la parole, Réduction de la dimensionnalité.