



UNIVERSITÉ DE TOURS

ÉCOLE DOCTORALE MIPTIS

LIFAT (EA 6300)

THÈSE présentée par :

Vanessa HAYKAL

Soutenance : 02/12/2019

pour obtenir le grade de : Docteur de l'Université de Tours

Discipline/ Spécialité : INFORMATIQUE

Modélisation des séries temporelles par apprentissage profond

THÈSE DIRIGÉE PAR :

CARDOT Hubert

RAGOT Nicolas

Professeur, Université de Tours

Maître de Conférences, HDR, Université de Tours

RAPPORTEURS :

LOUBES Jean-Michel

SALOTTI Jean-Marc

Professeur, Université Toulouse Paul Sabatier

Professeur, Institut Polytechnique de Bordeaux

JURY :

BONÉ Romuald

CARDOT Hubert

LOUBES Jean-Michel

RAGOT Nicolas

SALOTTI Jean-Marc

Professeur, INSA Strasbourg

Professeur, Université de Tours

Professeur, Université Toulouse Paul Sabatier

Maître de Conférences, HDR, Université de Tours

Professeur, Institut Polytechnique de Bordeaux

1.3 Contributions de la thèse

Les modèles de prévision des séries temporelles ont une importance fondamentale dans divers domaines applicatifs. Il est souhaitable que ces méthodes puissent apprendre des dépendances non linéaires et avoir une résistance au bruit élevée. Dans cette thèse, on s'est intéressé aux méthodes issues de l'apprentissage profond. Suite à l'état de l'art de ces modèles, voici les contributions de notre travail :

- En premier lieu, on montre l'intérêt des architectures composites ou multiples pour améliorer les performances des prédicteurs, notamment avec le principe de réduction du bruit. L'étude de l'état de l'art nous a permis de constater l'importance des segments d'erreurs commis dans le passé pour la prévision des valeurs futures. Nous proposons une nouvelle architecture de réduction de bruit hybride (*Hybrid Noise Reduction Architecture*, HNRA) qui utilise des segments d'erreurs récurrents pour l'apprentissage et l'ajustement des prévisions. La solution proposée est basée sur une fusion simultanée entre les sorties d'un réseau de neurones convolutifs (CNN) et d'un réseau récurrent à mémoire (LSTM). D'une part, nous utilisons des couches convolutives sur les entrées pour extraire les caractéristiques locales. Et d'autre part, en calculant les segments d'erreurs récurrents, nous empilons des couches de mémoire pour des dépendances à long et court termes. Ce nouveau modèle, nommé HNRA-CNN-LSTM, est appris par une optimisation globale (la partie prédictive et corrective sont apprises conjointement). Il est capable de capturer les différentes propriétés de la série, mieux qu'un modèle unique.
- En deuxième lieu, nous présentons un nouveau modèle pour les séries temporelles avec l'apprentissage profond basé sur le changement de régimes. Dans une optique de diviser pour régner, les approches locales détectent les changements d'états dans la série afin d'y associer différents prédicteurs HNRA-CNN-LSTM. Le modèle proposé est capable de générer des observations relatives à des comportements multiples et structurels. Cette technique est basée sur les propriétés des modèles de Markov cachés (HMM). Il suppose que les observations sont générées par plusieurs modèles locaux et apprend les caractéristiques de chaque modèle local. Le HMM a également la fonctionnalité de basculer dynamiquement entre les différents modèles locaux. Nous proposons une modification dans l'algorithme d'apprentissage de ces deux modèles séquentiels pour permettre leur bon fonctionnement.
- En troisième lieu, le bruit causé par divers facteurs externes altère souvent les séries temporelles et influe par conséquent sur les performances du modèle. Pour une prévision plus précise, nous utilisons une technique récente en multi-résolution. Le VMD (décomposition en mode variationnel) est une méthode adaptative pour l'analyse des propriétés non linéaires des séries temporelles. La série originale est décomposée en plusieurs modes de hautes et basses fréquences. Ces sous-séries utilisées dans notre architecture proposée tendent à confirmer l'intérêt des méthodes de pré-traitements conjointement aux approches profondes proposées.