

Irina Illina
Maître de conférences en Informatique
Habilité à Diriger les Recherches
Université de Lorraine
Équipe Multispeech
LORIA-INRIA Lorraine
BP 239, 54506 Vandœuvre-lès-Nancy, France
tél : + 33 3 54 95 84 90, mail : illina@loria.fr

Nancy, le 20 novembre 2017

**Rapport sur le mémoire de thèse de
Monsieur Mohamed Bouaziz**

*Réseaux de neurones récurrents pour la classification de séquences
dans des flux audiovisuels parallèles*
Université d'Avignon et des Pays de Vaucluse

Introduction

Le travail de thèse de monsieur Mohamed Bouaziz s'inscrit dans le cadre d'une Convention Industrielle de Formation par la Recherche (CIFRE), signée en partenariat avec l'Association Nationale de la Recherche et de Technologie -ANRT) et avec l'Européenne de Données (EDD). La thèse se situe dans le domaine de la classification automatique de données séquentielles provenant des flux parallèles. Mr Bouaziz s'intéresse plus particulièrement au cadre de flux audiovisuels et à la tâche de prédiction du genre d'émission TV. La thèse est effectuée sous la direction de Georges Linarès, Richard Dufour et Mohamed Morchid.

Le document de thèse est rédigé en français, organisé en 2 parties et contient 153 pages. Le style de présentation est clair, pédagogique et agréable à lire. A la fin de chaque chapitre l'auteur fait le bilan de ce qui rend la compréhension de ce document plus aisée.

Description du document

Une courte introduction donne un panorama bref mais clair du contexte général de la thèse, de l'entreprise EDD, de la problématique et de la structure du document.

Partie 1

Le chapitre 2 de la thèse est consacré à l'état de l'art dans le domaine de traitement du contenu télévisuel et de l'apprentissage supervisé. Cette partie est décrite de façon classique et ne représente pas de contribution importante. L'auteur a choisi de faire une description assez synthétique et de ne présenter

Campus scientifique



UNIVERSITÉ
DE LORRAINE

Inria
INVENTEURS DU MONDE NUMÉRIQUE

que l'essentiel de ces deux domaines. Dans les sections, consacrées au traitement automatique du contenu TV, la discussion consacrée aux différentes taxonomies des genres TV me paraît bien intéressante. En effet, une taxonomie peut dépendre du cadre d'étude, des besoins, du pays, etc. Cela peut rendre la comparaison de résultats de différentes taxonomies bien difficile. Dans la section décrivant les taxonomies réduites pour des traitements automatiques, quelques précisions supplémentaires et références seraient utiles car c'est ce type de taxonomie qui est utilisée tout au long de cette thèse. La section consacrée à la classification en genres d'émissions représente le cœur de la problématique de cette thèse. L'auteur passe en revue des différentes méthodes de classification et prépare la suite du document. Je ferais juste un regret que l'auteur ne parle des données séquentielles que dans la conclusion de ce chapitre. Quelques références seraient utiles ici. Les deux sections suivantes présentent l'état de l'art de la structuration et du résumé automatique du contenu télévisuel. A mon avis dans ce chapitre, la discussion sur la classification en genre d'émission, le centre de cette thèse, pourraient être élargie, et permettrait ainsi de mieux comprendre la problématique de cette thèse et ses contributions.

Le chapitre 3 présente les bases théoriques d'apprentissage, donne des définitions nécessaires et quelques illustrations. La présentation est claire. L'auteur a choisi de présenter les méthodes classiques de classification, puis les méthodes adaptées aux séquences et enfin les réseaux de neurones. Je regrette que dans cette partie ne figure pas un aperçu rapide des Machines à Vecteurs de Support (SVM) et des réseaux de neurones de type Perceptron (MLP). Pourtant ces deux classifieurs sont utilisés et évalués dans les chapitres 5 et 6.

Partie 2

La deuxième partie du document de thèse concerne les contributions de l'auteur pour répondre aux problèmes soulevés par cette thèse.

Le chapitre 4 présente les corpus de données utilisés, les protocoles d'évaluation adoptés et la taxonomie proposée. L'auteur utilise le corpus d'historique d'émissions TV de 2013-2015 des chaînes *M6*, *TF1*, *France 5* et *TV5 Monde*. Il aurait été intéressant de savoir si la taxonomie utilisée tout au long de cette thèse a été discutée et approuvée par l'entreprise EDD, utilisateur direct des travaux de cette thèse. Également, l'auteur de ce document dit que le séquençement de genres d'émission suit un modèle prédéfini qui reste stable pendant plusieurs semaines. Ce passage méritera d'être un peu plus discuté et justifié expérimentalement pour mieux motiver la problématique de cette thèse.

Le chapitre 5 propose une évaluation de différentes méthodes de classification sur le corpus de la chaîne M6. Les modèles SVM et MLP sont étudiés et considérés comme « baseline ». Parmi les modèles adaptés aux séquences, n-grammes et LSTM sont retenus par l'auteur et évalués. La « combinaison » originale de LSTM et SVM (SE-SVM) est proposée dans cette partie. Il est dommage que cette « combinaison » soit « vite adoptée » sans discuter les autres possibilités de combinaison. Également, des détails plus techniques des classifieurs (entrées/sorties, nombre de paramètres, le temps de calcul, etc.) auraient été bienvenus dans ce chapitre. Certains résultats sont donnés en termes de taux d'erreurs, les autres en termes de F-mesure. A mon avis, il serait

judicieux de donner les deux pour chaque classifieur étudié. Cette remarque est valable pour toutes les approches évaluées dans cette thèse. A la fin du chapitre, un tableau synthétique des résultats obtenus facilitera la comparaison des méthodes. Il serait intéressant de voir si les méthodes évaluées donnent les performances similaires sur la tâche de classification d'autres chaînes.

Le chapitre 6 est un chapitre central de cette thèse. Il est consacré à la classification au moyen de flux parallèles. L'auteur propose d'utiliser simultanément plusieurs flux de données provenant de chaînes différentes pour prédire le genre de l'émission d'une chaîne de test. Plusieurs modèles novateurs et intéressants sont proposés et évalués : LSTM parallèle (PLSTM), une combinaison de LSTM avec le SVM multiflux (MSE-SVM) et le MSE-SVM contextuel (AMSE-SVM). Ces trois modèles ont donné lieu aux publications dans les conférences internationales. Une formulation théorique, suivie par une évaluation expérimentale importante, est proposée pour chaque modèle. Les propositions de modèles sont originales, bien motivées et cohérentes. Pour chaque modèle proposé l'auteur obtient une amélioration par rapport au modèle précédent. Il serait intéressant de voir si les améliorations obtenues sont significatives. Une petite remarque sur la forme de la présentation : un tableau final comparant les résultats de différents modèles aurait pu aider le lecteur à mieux se retrouver. Ce chapitre montre un travail conséquent de recherche et de très bon niveau effectué par l'auteur.

L'auteur conclut son mémoire en résumant ses différentes contributions et en donnant quelques perspectives.

Conclusion

Monsieur Mohamed Bouaziz a fait la démonstration dans cette thèse de sa bonne connaissance du domaine. Il a proposé des solutions originales, validées par les publications dans des conférences internationales. Le mémoire de thèse est facile à lire, bien structuré, synthétique, avec un travail important d'implémentation et de validation. Malgré les critiques faites, dont plusieurs peuvent être levées dans la version définitive du document, mon appréciation générale est que le travail effectué est intéressant et important. Je suis donc très favorable à sa soutenance en vue d'obtention du titre de Docteur d'Université.

Irina Illina

*Maître de conférences en Informatique
Habilité à Diriger les Recherches
Université de Lorraine*



