



Pôle des Etudes Doctorales  
Centre des Etudes Doctorales  
Sciences et Techniques et Sciences Médicales

## AVIS DE SOUTENANCE DE THESE DE DOCTORAT

**Monsieur NDAMA Oussama**  
**Présentera ses travaux de recherche en vue de l'obtention du**  
**Doctorat**



**Formation Doctorale : Sciences et Techniques de l'Ingénieur**  
**Discipline : Informatique**  
**Spécialité : Informatique et Intelligence Artificielle**

**Le 27/09/2025 à 10H00 à la Salle de Conférence à la Faculté des  
Science et Techniques de Tanger, UAE**

**Sous le thème**

**Classification multi-domaines avancée par fusion contextuelle et séquentielle  
: Framework hybride unifié basé sur le Deep Learning et les Transformers  
pour la détection de fraude, l'analyse de sentiment financier et la  
classification de textes biomédicaux**

**Devant le jury composé de :**

Nom et Prénom	Etablissement	Qualité
Pr. FENNAN Abdelhadi	FST de Tanger, UAE	Président
Pr. HADDOUCH Khalid	ENSA de Fès, USMBA	Rapporteur
Pr. ELAACHAK Lotfi	FST de Tanger, UAE	Rapporteur
Pr. ZOUHAIR Abdelhamid	FST de Tanger, UAE	Rapporteur
Pr. GHADI Abderrahim	FST de Tanger, UAE	Examinateur
Pr. BOUHORMA Mohammed	FST de Tanger, UAE	Examinateur
Pr. EN-NAIMI El Mokhtar	FST de Tanger, UAE	Directeur

*Structure de recherche : Data Science, Artificial Intelligence and Smart Systems (DSAI2S) Research Team,  
Computer Science and Smart Systems (C3S) Laboratory, FST of Tangier, Abdelmalek Essaâdi University,  
Tetouan, Morocco.*

## Résumé



Dans de nombreux secteurs où la décision repose sur l'analyse de données textuelles et numériques en constante expansion, la finance et la biomédecine se distinguent par des besoins particulièrement exigeants. Les institutions financières, confrontées à la croissance rapide des transactions numériques, peinent à débusquer des fraudes rares et souvent dissimulées dans des volumes massifs. Parallèlement, l'analyse des sentiments financiers requiert un discernement subtil pour repérer les nuances de ton dans des articles de presse et anticiper les réactions de marché.

Du côté biomédical, l'avalanche de publications scientifiques, rédigées dans un langage hautement spécialisé, complique la recherche et la classification automatiques de textes pourtant essentiels pour la recherche clinique et la pratique médicale. Malgré leurs spécificités, ces défis partagent une problématique commune : comment exploiter conjointement la richesse contextuelle et séquentielle des données afin de mieux détecter des anomalies, saisir les nuances émotionnelles ou classifier précisément des textes complexes ?

Pour aborder cette double contrainte, cette thèse propose un framework analytique hybride, articulé autour de BERT et de LSTM. BERT fournit des représentations contextuelles riches capables d'interpréter les subtilités lexicales et sémantiques présentes dans les transactions financières, les actualités économiques ou les résumés biomédicaux. De son côté, LSTM excelle dans l'analyse séquentielle nécessaire pour traiter les séries temporelles bancaires ou suivre le fil narratif des articles scientifiques. En intégrant ces deux approches via un ajustement fin adapté à chaque tâche, ce framework unifié obtient des résultats probants : de la détection efficace de transactions suspectes malgré des jeux de données très déséquilibrés, à l'identification précise des nuances tonales en analyse sentimentale financière, en passant par la segmentation pertinente des textes biomédicaux en sections clés. Les résultats empiriques soulignent l'intérêt de fusionner ces deux approches : non seulement le rappel peut atteindre un niveau quasi parfait en détection de fraude, mais la précision dans l'analyse du sentiment neutre ou la segmentation de sections narratives s'améliore sensiblement par rapport aux architectures traditionnelles. Au-delà de ces trois applications, l'architecture BERT-LSTM brille par sa modularité : l'intégration de données multimodales (textuelles, numériques, catégorielles) et l'ajustement du modèle en fonction d'exigences spécifiques en font un outil polyvalent et robuste. Les limites identifiées, telles que le chevauchement sémantique entre certaines catégories biomédicales ou la charge computationnelle élevée, ouvrent la voie à des pistes d'amélioration : mécanismes d'attention affinés, modèles plus légers pour faciliter le déploiement, ou encore extension à d'autres domaines. Cette thèse démontre ainsi qu'une approche hybride, croisant la force contextuelle de BERT et la puissance séquentielle de LSTM, constitue une solution de référence pour des problématiques variées, alliant haute performance, évolutivité et capacité de généralisation à grande échelle.

**Mots clés :** Framework BERT-LSTM, Approche hybride, Représentations contextuelles, Modélisation séquentielle, Détection de fraude par carte de crédit, Analyse de sentiment financier, Classification de textes biomédicaux, Fine-tuning, Apprentissage profond, Traitement automatique du langage (Natural Language Processing : NLP), Données multimodales, Jeux de données déséquilibrés, Architecture modulaire.