VRIJE UNIVERSITEIT BRUSSEL

The faculty of Engineering of the Vrije Universiteit Brussel invites you to attend the public defense leading to the degree of

DOCTOR OF ENGINEERING SCIENCES

of **Xiangyu Yang**

The public defense will take place on **Tuesday 15th October 2024 at 4:00 pm** in room **D.0.05** (Building D, VUB Main Campus)

To join the digital defense, please click here

LEVERAGING DEEP LEARNING MODELS FOR BIG DATA ANALYTICS

BOARD OF EXAMINERS

Prof. dr. ir. John Lataire

Prof. dr. ir. Wendy Meulebroeck

Prof. dr. ir. Bruno da Silva Gomes

Dr. Do Huu

Prof. dr. Katrien Beuls

PROMOTORS

Prof. dr. ir. Nikolaos Deligiannis

VRIJE UNIVERSITEIT BRUSSEL

Abstract of the PhD research

With the exponential growth of data generated daily from social media, e-commerce, and various digital interactions, the necessity to effectively harness and leverage this vast expanse of information is more critical than ever. In this context, deep learning, a subfield of Artificial Intelligence (AI), has emerged as a transformative force, offering unprecedented capabilities in data analysis, pattern recognition, and predictive modeling. Deep learning takes large amounts of available data as fuel to train itself, and significantly impacts various fields ranging from healthcare to finance, enabling advanced applications in natural language processing (NLP), computer vision (CV), and recommender systems (RS).

This thesis delves into the essential role of AI in leveraging big data, focusing on information extraction from social media, deep learning model explainability, and the development of explainable recommender systems. With the vast, ever-growing volume of data, extracting meaningful insights from unstructured social media becomes increasingly complex, necessitating cutting-edge AI solutions. Concurrently, the reliance on deep learning models for critical decisions brings explainability to the forefront, emphasizing the importance of developing transparent methods that ensure user trust. Furthermore, the demand for recommender systems that provide understandable textual explanations has surged, highlighting the need for explainable systems that align with user preferences and decision-making processes.

This thesis advances the field through three key contributions. Initially, we establish two traffic-related datasets from social media, annotated for comprehensive traffic event detection. Employing BERT-based models, we tackle this detection problem via text classification and slot filling, proving these models' efficacy in parsing social media for traffic-related information. Our second contribution introduces LRP-based methods to explain deep conditional random fields, with successful applications in fake news detection and image segmentation. Lastly, we present an innovative personalized explainable recommender system that integrates user and item context into a language model, producing textual explanations that enhance system transparency.