الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للاعلام الالي 8 ماي 1945
École Supérieure en Informatique
8 Mai 1945 Sidi Bel Abbès

MÉMOIRE

En vue de l'obtention du diplôme de **Master**
Filière: **Informatique**
Spécialité: **Ingénierie des Systèmes Informatiques (ISI)**

Thème

BRAINWASHING 2.0: BIAS AND DOUBLE STANDARD IN WESTERN SOCIAL MEDIA AND NEWS OUTLETS.

Présenté par:
Abdellatif Abderaouf
Khemkham Mohamed

Soutenu le : **03 Juillet 2023**   devant le jury composé de :

Dr Badia Klouche	ESI-SBA	Président
Dr Oussama Serhane	ESI-SBA	Examinateur
Dr Sid Ahmed Benabderrahmane	NYU Abu Dhabi	Encadrant
Pr Sidi Mohammed Benslimane	ESI-SBA	Co-Encadrant

Année Universitaire : 2022-2023

ABSTRACT

In this study, we leverage techniques from machine learning, deep learning, natural language processing and sentiment analysis to investigate the presence and impact of biases in the coverage of the Western news platforms and social media sources.

Our focus centers around detecting and analyzing biases, differentiating between intentional and accidental misinformation, and interpreting how these biases may influence public perception, followed by sentiment analysis to study the variation of sentiments over time.

This research report delves into the intersection of Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), sentiment analysis, and media bias detection. The study elucidates how these innovative technologies can be leveraged to scrutinize and interpret media content, ensuring unbiased information dissemination.

RÉSUMÉ

Dans cette étude, nous exploitons des techniques issues de l'apprentissage automatique, de l'apprentissage profond, du traitement du langage naturel et de l'analyse des sentiments pour examiner la présence et l'impact des biais dans la couverture des plateformes d'information occidentales et des sources de médias sociaux.

Notre attention se concentre sur la détection et l'analyse des biais, en différenciant les informations erronées intentionnelles et accidentelles, et en interprétant comment ces biais peuvent influencer la perception du public, suivis par une analyse des sentiments pour étudier la variation des sentiments dans le temps.

Ce mémoire se plonge dans l'intersection de l'Apprentissage Automatique (AA), de l'Apprentissage Profond (AP), du Traitement du Langage Naturel (TLN), de l'analyse des sentiments et de la détection des biais médiatiques. L'étude éclaire comment ces technologies innovantes peuvent être exploitées pour scruter et interpréter le contenu médiatique, assurant une diffusion de l'information sans biais.

ملخص

في هذه الدراسة، نستفيد من التقنيات المستمدة من التعلم الآلي، التعلم العميق، معالجة اللغة الطبيعية وتحليل المشاعر للتحقيق في وجود وتأثير التحيز في تغطية وسائل الإعلام الغربية ومصادر وسائل التواصل الاجتماعي. تتمحور اهتماماتنا حول الكشف عن التحيز وتحليله، التفريق بين المعلومات المضللة العمدية والغير عمدية، وتفسير كيفية تأثير هذه التحيزات قد تؤثر على ادراك الجمهور، متبوعة بتحليل المشاعر لدراسة تغير العواطف على مر الزمن.

يغوص تقرير البحث هذا في تقاطع التعلم الآلي، التعلم العميق، معالجة اللغة الطبيعية، تحليل المشاعر، وكشف التحيز في الإعلام. توضح الدراسة كيف يمكن استغلال هذه التقنيات الابتكارية لفحص وتفسير محتوى الإعلام، مع ضمان نشر المعلومات بدون تحيز.

DEDICATIONS

We dedicate this work:

We would like to dedicate this graduation report to our parents, whose unwavering love and support have been instrumental throughout our academic journey.

We are deeply grateful to our mentors, who have provided us with invaluable guidance and wisdom, shaping our research and helping us navigate through challenges.

Our heartfelt appreciation goes to our friends and family, who have constantly encouraged and understood us during the countless late nights and long weekends spent working on this report.

We extend our sincere gratitude to the individuals and communities who graciously participated in our research, sharing their time and experiences, which enriched our work.

Lastly, we would like to acknowledge the faculty and staff of the ESI SBA for their unwavering dedication to teaching and mentoring, which has played a significant role in our growth and development.

- **AI**: Artificial Intelligence

- **ML**: Machine Learning

- **DL**: Deep Learning

- **ANNs**: Artificial Neural Networks

- **NLP**: Natural Language Processing

- **SVMs**: Support Vector Machines

- **MSE**: Mean Square Error

- **LASSO**: Least Absolute Shrinkage and Selection Operator

- **L1**: L1 norms

- **L2**: L2 norms

- **CBOW**: Continuous Bag of Words

- **GloVe**: Global Vectors for Word Representation

- **Cox Regression**: Proportional Hazards Regression

- **K-NN**: K-Nearest Neighbors

- **PCA**: Principal Component Analysis

- **t-SNE**: t-Distributed Stochastic Neighbor Embedding

- **LSTM**: Long Short-Term Memory

- **GRUs**: Gated Recurrent Units

- **RBFNs**: Radial Basis Function Networks

- **GANs**: Generative Adversarial Networks

- **SOMs**: Self-Organizing Maps

- **TF**: Term Frequency

- **IDF**: Inverse Document Frequency

- **TF-IDF**: Term Frequency-Inverse Document Frequency

- **ASR**: Automatic Speech Recognition

- **CRF**: Conditional Random Fields

- **BiLSTM**: Bidirectional Long Short-Term Memory

- **SQuAD**: Stanford Question Answering Dataset

- **QA**: Question Answering

- **NER**: Named Entity Recognition

- **DUC**: Document Understanding Conference