

L'intelligence artificielle pour l'analyse de la situation sécuritaire du Burkina Faso

Abdoul Fataoh Kabore
*Centre d'Excellence Interdisciplinaire
en Intelligence Artificielle
pour le Développement (CITADEL)
Université Virtuelle du Burkina (UVBF)*
Ouagadougou, Burkina Faso
abdoufataoh@gmail.com

Maïmouna Ouattara
*Centre d'Excellence Interdisciplinaire
en Intelligence Artificielle
pour le Développement (CITADEL)
Université Joseph KI-ZERBO (UJKZ)*
Ouagadougou, Burkina Faso
ouattaramouna80@gmail.com

Adama Soro
Université Norbert ZONGO (UNZ)
Koudougou, Burkina Faso
adamsoro0321@gmail.com

Rodrigue Kafando
*Centre d'Excellence Interdisciplinaire
en Intelligence Artificielle
pour le Développement (CITADEL)
Université Virtuelle du Burkina (UVBF)*

Abdoul Kader Kabore
*Centre d'Excellence Interdisciplinaire
en Intelligence Artificielle
pour le Développement (CITADEL)
Université du Luxembourg*

Résumé—Face à la situation d'insécurité due au terrorisme que vit le Burkina Faso depuis 2015, la population ne manque de manifester ses sentiments. Les diverses réactions de la population se font par les commentaires sur les différentes plateformes sociales créant ainsi une grande quantité de données. L'analyse de ces opinions pourrait être une aide à la prise de décision appliquée à la sécurité. Cette analyse pourrait se faire grâce aux techniques et méthodes offertes par l'intelligence artificielle (IA). Pour nos travaux dans cet article, nous proposons un outil de web scraping pour collecter des données. Puis, nous utilisons des modèles d'analyse des sentiments et de détection de propos haineux basés sur les transformers [2]. A travers ces recherches, nos contributions sont les suivantes : la mise en place d'une base de données textuelles ou un corpus sur la situation sécuritaire, la proposition d'une approche méthodologique basée sur l'IA pour l'analyse de ces données et la proposition d'un tableau de bord pour la visualisation des résultats d'analyse.

Nos travaux sont disponibles sur le dépôt suivant <https://github.com/abdoufataoh/security-situation-analysis>

Index Terms—situation sécuritaire, intelligence artificielle, transformers, web scraping .

I. INTRODUCTION

Depuis Août 2015, le Burkina Faso fait l'objet d'attaques terroristes visant aussi bien les Forces de Défense et de Sécurité (FDS), que les populations civiles. Cette situation d'insécurité impacte le pays dans tous les domaines. Sur le plan social, le rapport de situation de OCHA [1] dénombre à la date du 30 avril 2022, 1 520 012 de personnes déplacées internes dont 59,13% sont des enfants, et 4 258 écoles fermées. La population exprime diverses opinions et sentiments sur cette situation sécuritaire que vit le pays. De nos jours, avec la vulgarisation des technologies de l'information et de la communication (TIC), les réactions de la majeure partie de la population se manifeste le plus à travers les médias sociaux.

L'analyse de ces opinions pourrait être une aide à la prise de décision appliquée à la sécurité. L'intelligence artificielle offre de nombreuses méthodes et techniques pour la collecte et l'analyse des données. Elle permet de faire l'analyse de sentiment, la détection de propos haineux, l'extraction d'entités nommées, etc. Ainsi, l'intelligence artificielle pourrait être un apport dans le défi sécuritaire au Burkina Faso. Cependant, l'application des méthodes de l'intelligence artificielle nécessite une quantité importante de données. Pour ce faire, nous orientons nos questions de recherche sur les points suivants : comment obtenir une quantité de données suffisante pour l'application des méthodes d'IA ? Quelles peuvent être les analyses pertinentes à effectuer sur la base des données pour une meilleure prise de décision ?

L'objectif de notre recherche est d'acquérir des données de diverses sources sur la situation sécuritaire du Burkina Faso et de les analyser à travers les méthodes et techniques offertes par l'intelligence artificielle. A travers ces travaux, nous faisons les contributions suivantes : la mise en place d'une base de données textuelle ou un corpus sur la situation sécuritaire, la proposition d'une approche méthodologique basée sur l'IA pour l'analyse de ces données et la proposition d'un tableau de bord pour la visualisation des résultats d'analyse.

Cet article comporte cinq sections, la première étant cette introduction. Dans la deuxième section, nous présentons l'approche méthodologique que nous proposons. La section 3 présente l'application de notre approche. Dans la section 4, nous présentons les différents résultats obtenus et leur interprétation. La section 5 concerne une conclusion ainsi que les perspectives de nos travaux.

II. APPROCHE MÉTHODOLOGIQUE

Pour ce travail, nous proposons une approche subdivisée en deux grandes parties : la collecte des données et l'analyse de ces données par les modèles d'intelligence artificielle. La Figure 1 présente l'architecture générale du système. La première partie consiste en la mise en place d'outils pour la collecte et le stockage des données. Quant à la deuxième partie, elle consiste à la mise en place de module pour l'exploitation des modèles intelligence artificielle. Ces deux parties sont connectées entre elles afin de garantir une analyse fiable basée sur des informations rafraîchies.

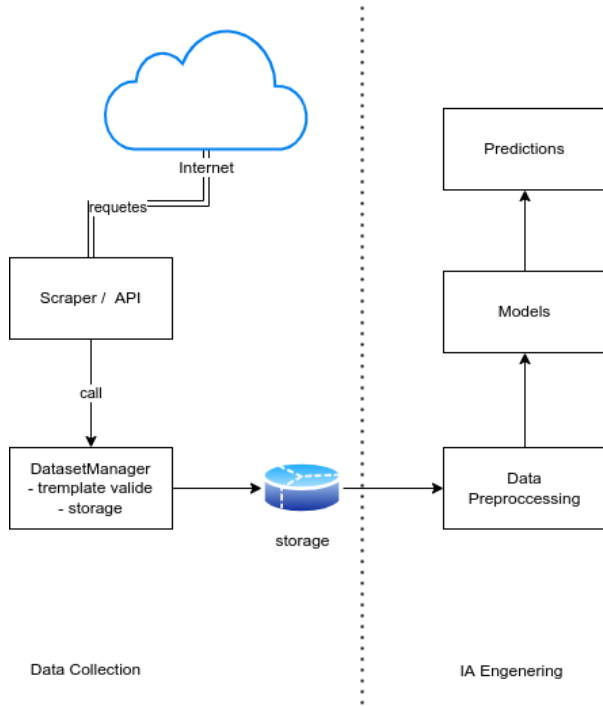


FIGURE 1. Architecture

III. APPLICATION MÉTHODOLOGIQUE

A. Collecte et prétraitement des données

Pour collecter et stocker les données des publications(ou articles) de l'actualité de la situation sécuritaire du burkina sur le site web *lefaso.net*¹, nous avons :

- implémenté un module asynchrone de web scraping nommé *lefaso.py*. Ce module permet de récupérer pour chaque rubrique, tous les articles ainsi que les commentaires associés à ces articles.Ce module permet de récupérer toutes les données pour la rubrique *attaque terroristes*
- mis en place un gestionnaire du jeu de données. Il assure la validation de la structure des données. Chaque donnée est stockée au format *json* dans une liste. La description de ces clés est énumérée dans le Tableau 1.

1. www.lefaso.net

TABLE I
LISTE DES CLÉS D'UN ARTICLE

Clé	Description
<i>article_type</i>	le type de article (rapport, article, etc.)
<i>article_title</i>	le titre de l'article
<i>published_date</i>	la date de publication de l'article
<i>origin</i>	la provenance de l'article
<i>url</i>	l'url d'accès de l'article
<i>content</i>	le contenu de l'article
<i>comments</i>	la liste des commentaires sur l'article
<i>comments_number</i>	le nombre de commentaires

Nous avons collecté 1080 articles avec 12975 commentaires au total. la distribution du nombre de commentaires par article est présente sur la Figure 2.

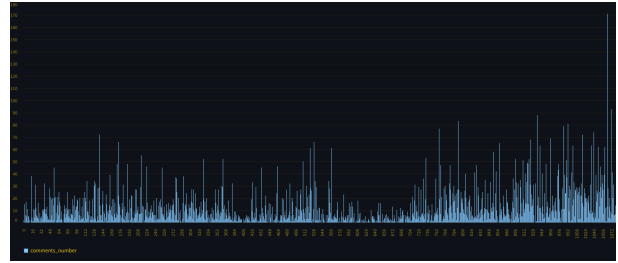


FIGURE 2. Distribution du nombre de commentaires par article

Le module de prétraitement des données a pour but de transformer les commentaires(données textuelles) en minuscule, de supprimer les caractères non alpha numérique et de supprimer les accents sur les lettres accentuées.

B. L'analyse de sentiments

Nous utilisons *4-sentiments-with-flaubert* [3] qui est un modèle pré-entraîné sur des données de la langue française et possédant une tâche pour l'analyse de sentiments. En entrée, il prend un texte et classifie ce texte en suivant quatre classes :

- OBJECTIVE
- POSITIVE
- NEGATIVE
- MIXED

C. La détection de propos haineux

Dans cette analyse, nous cherchons à détecter les commentaires à caractère de haine. Pour cela, nous utilisons *hubert-mono-french*, modèle spécialisé dans cette tâche. il classifie si un texte est à caractères haineux pas avec les labels :

- Hate
- No_Hate

IV. RÉSULTATS

A. L'analyse de sentiments

En utilisant le modèle *4-sentiments-with-flaubert* [3], nous avons réussi à déterminer les sentiments des internautes grâce à leur commentaires sur la situation sécuritaire du Burkina. La Figure 3 présente l'évolution des différents sentiments avec une granularité mensuelle. La courbe avec le label *negative*

(couleur bleue) est généralement au-dessus des autres autres courbes. Nous constatons sur cette courbe deux grands pics témoignant un sentiment négatif très élevé : Le premier au mois de *novembre 2019* donc dans l’histoire du Burkina à la date du *6 novembre 2019* le pays à déplorer des perte humaines élevé suite à l’attaque de la *Semafo de Boungou*. Quant au second pic, il est enregistré au mois de *juin 2019*. A la date du 5 juin 2021 le village de Solhan a été attaqué avec au moins 160 morts [5].

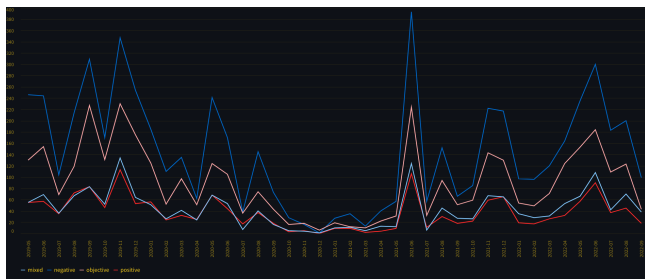


FIGURE 3. Courbe d’évolution des sentiments en fonction de la mensualité

B. Détection de propos haineux

L’analyse des commentaires avec le modèle hubert-mono-french [4] nous a permis de tracer la Figure 4. Elle présente l’évolution du nombre des commentaires à caractère haineux en fonction de la mensualité. Nous constatons une forte corrélation entre la courbe des commentaires à caractères haineux et celle des sentiments négatifs. Il y a donc des propos haineux dans les commentaires après chaque grande attaque terroriste.

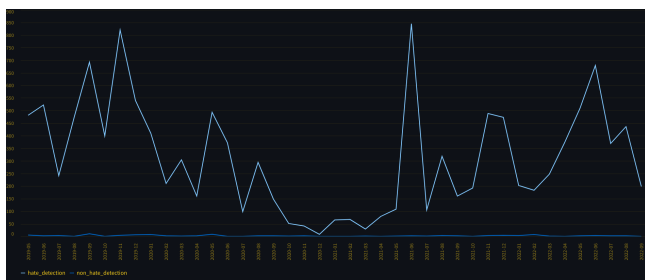


FIGURE 4. Courbe d’évolution du nombre de commentaires haineux en fonction de la mensualité

V. CONCLUSION

Le but de ces travaux était d’acquérir et d’analyser des données sur la situation sécuritaire du Burkina Faso. Nous avons collecté des données à travers les techniques de web scraping. Ces données ont été utilisées pour entraîner des modèles transformers pour l’analyse des sentiments et la détection des propos haineux. Les résultats obtenus à travers les visualisations nous présentent une hausse de sentiments négatifs et des propos haineux dans les périodes où il y a eu des attaques terroristes. Ces travaux permettent de mettre un système qui devrait aider à la prise de décision sur la situation sécuritaire du Burkina Faso. Les perspectives pour nos travaux sont d’une part, la collecte de données hétérogènes. Ainsi, nous prévoyons d’autres sources de données telles que d’autres médias sociaux. D’autre part, il va d’extraire des entités nommées spatiales sur les données obtenues afin de visualiser la relation entre ces entités nommées et les différentes analyses précédentes.

REMERCIEMENT

Nous remercions le Centre d’Excellence Interdisciplinaire en Intelligence Artificielle pour le Développement (CITADEL) pour le cadre de travail et l’accompagnement reçus. Nous remercions également les réviseurs anonymes pour leurs précieux commentaires et suggestions.

RÉFÉRENCES

- [1] <https://reports.unocha.org/fr/country/burkina-faso>
- [2] LIN, Tianyang, WANG, Yuxin, LIU, Xiangyang, et al. A survey of transformers. arXiv preprint arXiv :2106.04554, 2021.
- [3] Jérémy Demange, Four sentiments with FlauBERT, (2021), Hugging Face repository, <https://huggingface.co/DemangeJeremy/4-sentiments-with-flaubert>
- [4] Aluru, Sai Saket and Mathew, Binny and Saha, Punyajoy and Mukherjee, Animesh Deep Learning Models for Multilingual Hate Speech Detection, 2020
- [5] https://fr.wikipedia.org/wiki/Massacre_de_Solhan