

# A Comparative Analysis of Transformer and LSTM Models for Detecting Suicidal Ideation on Reddit

BIBLIOGRAPHIE ET ETUDE DES CAS, RAPPORT FINAL, 19 AVRIL 2025

Superviseur Académique : Dr. HO Tuong Vinh

1<sup>st</sup> LUTALA LUSHULI David

*option : Systèmes Intelligents & Multimédia*  
*Institut Francophone International*  
Hanoï, Vietnam  
Auteur du rapport

1<sup>st</sup> Khalid Hasan

*Department of Computer Science*  
*Missouri State University*  
Springfield, Missouri, USA  
Auteur de l'article

3<sup>rd</sup> Jamil Saquer

*Department of Computer Science*  
*Missouri State University*  
Springfield, Missouri, USA  
Auteur de l'article

**Résumé**—Ce rapport s'inscrit dans le cadre du module de Bibliographie et Étude des Cas, qui vise à développer chez l'étudiant la capacité à réaliser une analyse critique et approfondie d'un article scientifique. Le travail présenté porte sur la détection automatisée des idées suicidaires exprimées sur les réseaux sociaux, un enjeu majeur de santé publique à l'ère numérique.

Face à la croissance exponentielle des publications en ligne et à l'impossibilité d'un suivi humain exhaustif, l'intégration de technologies d'intelligence artificielle (IA), et plus particulièrement du traitement automatique du langage naturel (NLP), offre des perspectives prometteuses pour la prévention du suicide.

L'article principal analysé compare rigoureusement deux grandes familles de modèles de classification de texte — les architectures Transformers (BERT, RoBERTa, etc.) et les modèles LSTM — appliquées à des messages extraits de Reddit. Les résultats expérimentaux montrent une nette supériorité des Transformers, notamment RoBERTa, qui atteint un F1-score de 93,14 %, confirmant leur efficacité pour cette tâche sensible.

L'étude met également en lumière plusieurs défis cruciaux, tels que le biais des données, l'absence de validation clinique, et les implications éthiques liées à la vie privée et à la transparence algorithmique. Pour approfondir cette réflexion, trois travaux connexes ont été étudiés. Le premier propose une approche semi-supervisée avec Su-RoBERTa et une augmentation de données par GPT-2 pour pallier le manque d'annotations. Le deuxième introduit une infrastructure Big Data capable de traiter en temps réel des flux de messages sociaux, démontrant ainsi la faisabilité d'un système de détection à grande échelle. Le troisième met l'accent sur l'explicabilité des modèles à travers MentalRoBERTa, afin de rendre les prédictions compréhensibles et exploitables dans un cadre clinique.

Ces travaux mettent en évidence que les différentes approches étudiées sont complémentaires : certaines sont légères et peu coûteuses en calcul, d'autres sont capables de traiter de très grands volumes de données, et certaines offrent une meilleure compréhension des résultats. Cela montre qu'aucune méthode seule ne suffit. Il devient donc essentiel d'adopter une approche mixte, qui allie à la fois de bonnes performances techniques, une fiabilité des résultats, une transparence dans les décisions prises par l'IA, et un respect des principes éthiques.

Ainsi, ce rapport conclut que la détection des idées suicidaires à travers les réseaux sociaux doit s'appuyer sur une collaboration entre techniques avancées et expertise humaine, afin de répondre aux exigences des différents contextes d'utilisation.

**Index Terms**—Intelligence artificielle, traitement automatique du langage naturel, modèles Transformers, LSTM (Long Short-Term Memory), prévention du suicide, réseaux sociaux, idées suicidaires, RoBERTa, Su-RoBERTa, détection automatique, santé mentale, NLP explicable, Big Data.

## I. INTRODUCTION

### A. Contexte

À l'ère du numérique, les réseaux sociaux sont devenus des espaces privilégiés d'expression, où de nombreuses personnes partagent librement leurs pensées, leurs émotions, et parfois, leurs détresses profondes. Dans ce flux constant d'informations, des signaux avant-coureurs de mal-être psychologique, voire d'idées suicidaires, peuvent émerger. La détection précoce de ces signes est devenue un enjeu majeur de santé publique, susceptible de sauver des vies.

Cependant, face à l'immensité et à la rapidité de diffusion des contenus, une surveillance manuelle s'avère non seulement inefficace, mais également irréalisable. Ce constat a motivé le recours à des technologies avancées telles que le traitement automatique du langage naturel (NLP) et l'intelligence artificielle, capables d'analyser automatiquement les messages publiés sur les plateformes sociales afin d'identifier des signaux de détresse.

Néanmoins, la mise en place de systèmes automatisés de détection soulève d'importants défis, aussi bien techniques qu'éthiques, qui doivent être minutieusement adressés pour garantir une analyse fiable, responsable et respectueuse de la vie privée.

## B. Problématique

La complexité de la détection des idées suicidaires sur les réseaux sociaux réside dans trois catégories principales de défis :

### 1) Défis liés à l'extraction des données :

- Accès limité aux données (restrictions des API des plateformes)
- Formats de données variés et hétérogènes (abréviations, emojis, etc.).
- Contraintes liées au manque des données annotées de manière fiable.

### 2) Défis liés au volume de données :

- Traitement de millions de publications par jour
- Gestion de stockage des données massives
- Besoin de mécanismes de filtrage intelligents pour éviter les faux positifs

### 3) Défis liés au traitement en temps réel :

- Nécessité d'une détection instantanée des signaux de détresse
- Minimisation de la latence dans le traitement des données
- Adaptation continue aux évolutions du langage et des tendances en ligne

Face à ces enjeux, il est impératif de développer des modèles NLP performants qui répondent à ces exigences pour une prise en charge rapide et efficace des situations à risque.

## C. Objectifs

Pour relever ces défis, l'objectif est de concevoir des solutions basées sur le NLP capable de :

- Détecter avec précision les idées suicidaires dans les publications sur les réseaux sociaux, en tenant compte des spécificités du langage (abréviations, emojis, ironie,...).
- Garantir une classification robuste, minimisant les erreurs pour éviter des signalements injustifiés ou des omissions critiques.
- Optimiser l'intervention, en réduisant le délai entre la détection d'un signal d'alerte et l'activation d'assistance afin de prévenir efficacement les passages à l'acte.

## II. PRÉSENTATION DE L'ARTICLE PRINCIPAL

### A. Contexte, problématique et objectifs spécifiques

L'article intitulé « *A Comparative Analysis of Transformer and LSTM Models for Detecting Suicidal Ideation on Reddit* » s'inscrit dans le contexte de la prévention du suicide à travers l'analyse de publications sur les réseaux sociaux. Reddit, en particulier la communauté `r/SuicideWatch`, constitue une source précieuse pour identifier les signaux de détresse psychologique.

Face à la complexité du langage et au volume massif des données en ligne, l'étude vise à comparer deux grandes familles de modèles de classification : les Transformers (BERT, RoBERTa, etc.) et les LSTM, avec diverses techniques d'embedding, pour détecter automatiquement les publications à caractère suicidaire. Deux questions principales sont posées :

- 1) Les modèles Transformers sont-ils plus performants pour détecter les idées suicidaires ?
- 2) Comment les embeddings influencent-ils les performances des modèles LSTM ?

### B. Méthodologie et techniques utilisées

L'article adopte une méthodologie rigoureuse en plusieurs étapes, allant de la constitution du jeu de données à l'évaluation comparative de différents modèles de classification.

**a) Constitution et annotation du corpus** Les données ont été collectées via l'API Pushshift sur Reddit entre octobre et décembre 2022. Un total de 37 821 publications a été extrait de plusieurs sous-forums, dont `r/SuicideWatch`, `r/socialanxiety`, `r/TrueOffMyChest`, `r/bipolar`, `r/confidence` et `r/geopolitics`. Les publications provenant de `r/SuicideWatch`, forum explicitement dédié aux pensées suicidaires, ont été automatiquement étiquetées comme « suicidaires », les autres comme « non-suicidaires ».

Pour valider cette annotation automatique, les auteurs ont eu recours à deux techniques :

- *Modélisation thématique (LDA)* : pour vérifier la nature des contenus sur `r/SuicideWatch`.
- *Annotation humaine* : sur un échantillon de 756 posts, avec un coefficient de Cohen's Kappa supérieur à 0.85, indiquant une fiabilité quasi parfaite.

### b) Prétraitement des données

Les textes ont été nettoyés et normalisés (suppression des caractères spéciaux, emojis, URLs, etc.). Ensuite, une lemmatisation a été appliquée pour simplifier l'analyse linguistique. Plusieurs versions du corpus ont été utilisées selon les tâches (POS-tagging, topic modeling, extraction de phrases suicidaires via TextRank, etc.).

### c) Modèles évalués

Deux grandes familles de modèles sont comparées :

- *Transformers* : BERT, RoBERTa, DistilBERT, ALBERT, ELECTRA.
- *LSTM* : avec ou sans attention, bidirectionnels ou non, utilisant les embeddings Word2Vec, GloVe ou BERT.

Tous les modèles ont été entraînés à l'aide d'une validation croisée stratifiée à 5 plis. Les hyperparamètres ont été optimisés via Ray-Tune.

### C. Résultats expérimentaux et performances

Les résultats montrent que l'ensemble des modèles Transformers obtiennent d'excellentes performances, avec une précision et un F1-score supérieurs à 91%.

- Le modèle RoBERTa est le plus performant avec une précision de 93.22% et un F1-score de 93.14%.
- Les modèles ELECTRA, BERT et ALBERT suivent de très près, dans une marge d'environ 1%.
- DistilBERT, bien que plus léger, atteint un F1-score de 91.87% et réduit considérablement le temps d'entraînement.

Concernant les modèles LSTM :

- Ceux utilisant les embeddings BERT atteignent des scores comparables aux Transformers (jusqu'à 92.69% de F1-score).
- Les modèles avec GloVe ou Word2Vec chutent considérablement en performance (souvent en dessous de 75% de F1-score).

Les résultats expérimentaux montrent que les modèles Transformers surpassent les LSTM en termes de précision et de capacité à détecter les signaux suicidaires, notamment grâce à leur aptitude à analyser le contexte global d'un message. Cependant, l'étude met en avant certains défis, notamment la sensibilité aux données d'entraînement et le risque de faux positifs.

### D. Analyse critique et limites

Malgré des résultats solides, plusieurs limites sont à souligner :

- *Biais de plateforme* : L'intégralité des données provient de Reddit et des subreddits américains, ce qui peut poser des problèmes de généralisation à d'autres réseaux sociaux (Twitter, TikTok, etc.) ou à d'autres contextes culturels.
- *Méthode d'annotation simplifiée* : Le choix d'étiqueter automatiquement tous les messages de r/SuicideWatch comme suicidaires, bien que pragmatique, pourrait inclure certains faux positifs (ex. : des messages de soutien peuvent être classés comme suicidaires).
- *Absence de validation clinique* : Aucun professionnel de santé mentale n'a été impliqué dans l'évaluation des résultats.
- *Peu de traitement des enjeux éthiques* : La question de la confidentialité, du consentement, et de l'utilisation réelle de ces systèmes n'est abordée que superficiellement.

### E. Apports de l'étude

Hormis ses limites, l'étude constitue une avancée importante dans le domaine de la détection automatique des idées suicidaires, en offrant des contributions importantes telles que :

- Elle fournit un jeu de données Reddit annoté précieux pour la communauté scientifique.

- Elle propose une évaluation comparative rigoureuse de plusieurs modèles NLP récents.
- Elle confirme que les Transformers sont les mieux adaptés à cette tâche sensible.
- Elle montre que les LSTM restent compétitifs lorsqu'ils sont associés à des embeddings contextuels comme ceux de BERT.

Ces apports offrent des perspectives concrètes pour le développement de systèmes de veille automatisée à visée préventive dans le domaine de la santé mentale.

## III. TRAVAUX CONNEXES

Dans le prolongement de l'article principal, plusieurs travaux récents explorent des approches complémentaires pour la détection des idées suicidaires sur les réseaux sociaux. Ces recherches se distinguent tant par leurs méthodologies que par les types de modèles mobilisés, l'architecture technique déployée, ou encore la manière dont elles abordent les enjeux d'explicabilité.

Trois contributions majeures méritent d'être présentées ici :

- une approche semi-supervisée légère avec Su-RoBERTa
- une architecture Big Data orientée traitement en temps réel
- un modèle explicable destiné à l'aide à la décision clinique.

### A. ARTICLE 1 : « SU-RoBERTa : A Semi-supervised Approach to Predicting Suicide Risk through Social Media using Base Language Models »

Dans cet article, les auteurs proposent une stratégie alternative à l'apprentissage supervisé classique en combinant un modèle pré-entraîné (RoBERTa) avec une approche semi-supervisée. L'objectif est de pallier le manque de données annotées en exploitant des données non labellisées, tout en renforçant les performances du modèle à travers une augmentation de données générée par GPT-2.

Cet article explore l'efficacité d'un modèle NLP compact et peu coûteux, Su-RoBERTa, dans un contexte de données limitées. Face à la rareté des données annotées, les auteurs proposent une méthode semi-supervisée qui repose sur un ensemble initial de 500 publications étiquetées, enrichi ensuite par 1500 publications non-étiquetées traitées via un mécanisme de pseudo-labellisation progressive.

Afin de pallier le déséquilibre des classes, une augmentation de données est réalisée à l'aide de GPT-2, qui génère des textes synthétiques reflétant les différents niveaux de risque suicidaire (de l'indication simple jusqu'à la tentative).

### **Méthodologie :**

- Utilisation de Su-RoBERTa, un modèle basé sur RoBERTa contenant 355 millions de paramètres, plus léger et plus rapide que GPT-3.5.
- Mise en œuvre d'un apprentissage semi-supervisé pour entraîner le modèle sur des données partiellement annotées.
- Génération de données synthétiques avec **GPT-2** afin d'augmenter la diversité des exemples et équilibrer les classes.

### **Résultats :**

Le modèle atteint un F1-score pondéré de 69,84%. Malgré des ressources limitées, cette méthode permet d'obtenir des performances satisfaisantes pour une application pratique.

### **Analyse critique :**

L'approche se distingue par son pragmatisme : en partant d'un faible volume de données annotées, les auteurs parviennent à entraîner un modèle stable et performant. L'usage de GPT-2 pour générer des exemples est pertinent pour enrichir le corpus, et l'architecture compacte de Su-RoBERTa est adaptée à des déploiements sur des infrastructures légères.

### **B. ARTICLE 2 : «Big Data Analytics for Suicide Prevention : Analyzing Online Discussions and Social Media Data»**

Ce deuxième article adopte une approche orientée temps réel et scalabilité à travers l'utilisation d'une architecture Big Data. Contrairement aux approches classiques basées sur un traitement en batch, ce travail met l'accent sur le traitement continu des messages provenant de Twitter et Reddit afin de prédire les risques suicidaires de manière instantanée.

Pour cela, les auteurs mettent en œuvre une architecture combinant Apache Spark (pour le traitement distribué) et Apache Kafka (pour la gestion des flux).

### **Méthodologie :**

- Architecture technique fondée sur Apache Spark pour le traitement distribué et Apache Kafka pour la gestion des flux de données en temps réel.
- Évaluation de plusieurs modèles de machine learning : Naïve Bayes, Régression Logistique, Arbre de Décision, MLP, SVM.
- Enrichissement des textes avec des outils de traitement linguistique pour extraire les signaux clés.

### **Résultats :**

Plusieurs modèles classiques sont comparés (SVM, Random Forest, etc.), mais c'est le perceptron multicouche (MLP) qui affiche les meilleures performances en termes de précision et de temps de réponse, avec une précision de 93,47% sur les données en lot. Le système peut produire des prédictions en moins de deux secondes, ce qui en fait une option robuste pour des déploiements à grande échelle.

### **Analyse critique :**

Cette approche démontre la faisabilité d'un traitement en temps réel de signaux suicidaires sur les réseaux sociaux. Le principal inconvénient réside dans l'infrastructure nécessaire, complexe et coûteuse à maintenir. De plus, l'article ne traite pas suffisamment la dimension sémantique des messages, ce qui pourrait limiter la précision dans des contextes plus ambigus ou émotionnellement chargés.

### **C. ARTICLE 3 : «Conceptualizing Suicidal Behavior : Utilizing Explanations of Predicted Outcomes to Analyze Longitudinal Social Media Data»**

Cet article s'intéresse à l'interprétabilité des modèles utilisés pour la détection de messages suicidaires. Contrairement aux modèles « boîtes noires » comme GPT ou BERT en version brute, les auteurs plaident pour des approches explicables qui permettent de comprendre les décisions prises par l'algorithme, notamment pour favoriser leur adoption dans un cadre clinique.

### **Méthodologie :**

- Utilisation de MentalRoBERTa, un modèle pré-entraîné sur des corpus spécialisés en santé mentale.
- Intégration de techniques d'explicabilité : Layer Integrated Gradients (LIG), LIME et SHAP. des méthodes permettant d'expliquer les décisions du modèle en mettant en évidence les éléments clés ayant influencé une prédiction.
- Complémentarité avec des méthodes d'analyse classiques comme le TF-IDF pour renforcer la pertinence des signaux détectés.

### **Résultats :**

MentalRoBERTa-base a obtenu une précision de 63.05%. L'approche permet une interprétation visuelle et un contrôle accru sur les prédictions, ouvrant la voie à des outils d'aide à la décision plus fiables en contexte clinique.

### **Analyse critique :**

L'article offre une perspective novatrice sur l'importance de la transparence dans les systèmes de détection automatique. Il met en évidence la nécessité de proposer des outils compréhensibles et auditables, en particulier dans des domaines sensibles comme la santé mentale, où les professionnels doivent pouvoir justifier toute décision prise sur la base d'un modèle prédictif.

La performance globale du modèle reste en retrait par rapport aux standards actuels. De plus, l'intégration de couches explicatives, bien que pertinente, introduit un surcoût computationnel et peut limiter l'usage en temps réel. L'étude gagnerait également à proposer des cas d'usage concrets avec des professionnels de santé.

## IV. ETUDE COMPARATIVE DES ARTICLES

Critères	Article principal (Transformers vs LSTM)	Article Connexe 1 : SU_RoBERTa (semi-supervisé)	Article Connexe 2 : Big Data (Spark + Kafka)	Article Connexe 3 : EXPLICA (explicable)
<b>Approche</b>	Supervisée	Semi-supervisée	Traitement en flux / temps réel	Supervisée + interprétable
<b>Type de données</b>	Reddit (~38k messages)	Reddit (2k messages)	Reddit + Twitter (flux)	Reddit (UMD SuicideWatch)
<b>Modèle principal</b>	RoBERTa, LSTM, BERT	Su-RoBERTa + GPT-2	MLP, SVM, LR, RF	MentalRoBERTa + LIG
<b>F1-score / Précision</b>	93,14 % (RoBERTa)	69,84 %	93,47 % (MLP)	63,05 %
<b>Infrastructure</b>	Standard (GPU)	Léger, modèle compact	Lourde (cluster distribué)	Moyenne (GPU + interprétation)
<b>Forces</b>	Très bonnes performances, rigueur expérimentale	Adapté aux faibles ressources, simple à entraîner	Scalabilité, rapidité, déploiement temps réel	Explicabilité, utilité clinique
<b>Limites</b>	Données centrées Reddit, aspects éthiques peu abordés	Données limitées, dépendance au contenu synthétique	Infrastructure coûteuse, difficilement généralisable	Performance modeste, coût computationnel plus élevé
<b>Cas d'usage idéal</b>	Analyse comparative en environnement académique	Scénarios à données réduites ou budget limité	Système de monitoring automatique à grande échelle	Outil d'aide à la décision pour professionnels de santé

Table 1 : Comparaison entre l'article principal et les travaux connexes

## V. PERSPECTIVES ET AMÉLIORATIONS POSSIBLES

L'analyse comparative des différentes approches de détection de l'idéation suicidaire sur les réseaux sociaux met en lumière plusieurs pistes d'amélioration, tant au niveau des modèles que des pratiques de mise en œuvre. Ces perspectives concernent à la fois l'optimisation des performances techniques, la prise en compte de la dimension humaine, et l'adaptation aux contraintes éthiques et sociétales.

### *Améliorations technologiques*

Tout d'abord, les performances des modèles pourraient être renforcées par l'intégration d'approches multimodales. À ce jour, la majorité des études se concentrent uniquement sur les données textuelles, alors que des informations précieuses peuvent être extraites des images, vidéos ou métadonnées associées aux publications. La combinaison texte + image, notamment sur des plateformes comme Instagram ou TikTok, permettrait d'augmenter la sensibilité des systèmes sans pour autant compromettre la précision.

Par ailleurs, les modèles de type Transformer continuent d'évoluer. L'adoption de versions plus récentes ou spécialisées — comme DeBERTa, BioBERT (dans un contexte santé) ou LLaMA — pourrait offrir de meilleures capacités de contextualisation, notamment dans les langues moins représentées ou les contextes culturels spécifiques.

### *Enrichissement des données et robustesse*

Un axe essentiel réside dans la qualité des données utilisées pour l'entraînement. Il serait pertinent d'étendre les corpus au-delà de Reddit pour inclure d'autres sources (commentaires YouTube, Twitter, etc.). L'ajout de données issues de plusieurs pays, langues et cultures permettrait également de concevoir des modèles plus inclusifs et représentatifs.

L'utilisation de techniques de data augmentation contrôlée, combinée à des méthodes de pseudo-labellisation supervisée par experts, pourrait également améliorer la robustesse des modèles tout en limitant les coûts humains de l'annotation.

### *Vers des modèles plus explicables*

L'intégration de mécanismes d'explicabilité comme LIME, SHAP ou les gradients intégrés est une priorité dans les contextes sensibles tels que la santé mentale. Ces outils doivent devenir la norme afin de garantir la transparence des décisions algorithmiques et de permettre un travail collaboratif entre l'IA et les professionnels de santé.

Des efforts supplémentaires sont à fournir pour rendre ces explications non seulement accessibles mais aussi exploitables par des personnes non-expertes. Des visualisations intuitives ou des résumés automatisés pourraient par exemple accompagner chaque prédiction.

### *Défis éthiques et réglementaires*

Enfin, au-delà de la performance, les perspectives futures doivent intégrer une réflexion approfondie sur les questions éthiques et juridiques. Le respect de la vie privée, la gestion du consentement, et la prévention des usages malveillants des technologies de détection constituent des enjeux majeurs.

Une collaboration plus étroite entre chercheurs en IA, juristes, cliniciens et psychologues permettra d'élaborer des cadres de déploiement responsables (comme le RGPD).

### *Vers des systèmes hybrides*

À terme, la conception de systèmes hybrides combinant des modèles performants avec des capacités explicatives, une architecture flexible pour un traitement en lot ou en temps réel et une supervision humaine représente une direction prometteuse. Ces systèmes pourraient être utilisés dans des plateformes d'assistance psychologique.

## VI. CONCLUSION GÉNÉRALE

La détection des idées suicidaires sur les réseaux sociaux constitue un enjeu crucial à l'heure où ces plateformes deviennent des espaces d'expression privilégiés pour les personnes en détresse. Face au volume croissant et à la diversité des contenus publiés en ligne, l'automatisation de cette tâche apparaît comme une nécessité, mais également comme un défi technologique, méthodologique et éthique de grande ampleur.

À travers l'analyse de l'article principal, nous avons observé que les modèles Transformers, et en particulier RoBERTa, surpassent les approches classiques de type LSTM, offrant des performances élevées en classification binaire sur des messages Reddit. La méthodologie rigoureuse, l'annotation validée et l'évaluation comparative apportent une contribution solide à l'état de l'art.

Les travaux connexes étudiés permettent de nuancer et d'enrichir cette perspective. L'approche semi-supervisée de Su-RoBERTa montre qu'il est possible de concevoir des solutions efficaces avec peu de données annotées, tandis que l'architecture Big Data démontre la faisabilité d'un traitement en temps réel à l'échelle industrielle. Enfin, l'approche explicable d'EXPLICA rappelle la nécessité de concevoir des systèmes interprétables, notamment lorsqu'ils sont utilisés dans des contextes cliniques sensibles.

Cette étude souligne qu'aucune approche ne peut répondre seule à tous les besoins. Le futur de la détection des signaux suicidaires en ligne réside probablement dans une intégration intelligente de ces différentes dimensions : performance, explicabilité, scalabilité et éthique. La recherche dans ce domaine doit continuer à évoluer dans un esprit interdisciplinaire, en s'appuyant sur la complémentarité des approches techniques, médicales et humaines.

## RÉFÉRENCES

- [1] Khalid Hasan, et Jamil Saquer, "A Comparative Analysis of Transformer and LSTM Models for Detecting Suicidal Ideation on Reddit" IEEE ICMLA, vol. 1, Page 1-7, Novembre 2024.
- [2] Chayan Tank, Shaina Mehta, Sarthak Pol, Vinayak Katoch, Avinash Anand, Raj Jaiswal et Rajiv Ratn Shah, "Su-RoBERTa : A Semi-supervised Approach to Predicting Suicide Risk through Social Media using Base Language Models", IEEE Big Data, vol. 2 Page 1-8, Décembre 2024.
- [3] Mohamed A. Allayla et Serkan Ayvaz, "A Big Data Analytics System for Predicting Suicidal Ideation in Real-Time Based on Social Media Streaming Data" IEEE Computation and Language, vol. 3 Page 1-23, Mars 2024.
- [4] Van Minh Nguyen, Nasheen Nur, William Stern, Thomas Mercer, Chiradeep Sen, Siddhartha Bhattacharyya, Victor Tumbiolo et Seng Jhing Goh, "Conceptualizing Suicidal Behavior : Utilizing Explanations of Predicted Outcomes to Analyze Longitudinal Social Media Data", IEEE ICMLA, vol. 2 Page 1-8, Décembre 2023.