NavaJyoti, International Journal of Multi-Disciplinary Research

Volume 9, Issue 1, June 2025

Detecting Suicidal Thoughts on Social Media using Machine Learning and Deep Learning Models

Deepthi Shalini D[1], Shilpa Abhang[2]

[1]Student, Department of Computer Science(PG), Jyoti Nivas College Autonomous, Bangalore, India
[2] Associate Professor, Department of Computer Science(PG), Jyoti Nivas College Autonomous, Bangalore, India

Abstract

Suicidal ideation is a significant public health concern, with many individuals expressing their thoughts and emotions on social media. Early detection of such ideation is crucial for timely intervention and prevention. Traditional methods for identifying suicidal thoughts rely on human observation and are often limited in scale. To overcome this challenge, this research proposes a machine learning and deep learning-based system for detecting suicidal tendencies in social media posts. The study leverages a dataset of user-generated text, undergoing multiple preprocessing techniques to improve data quality. Two models, CNN-BiLSTM and XGBoost, are employed to classify posts as suicidal or non-suicidal. CNN-BiLSTM integrates feature extraction and sequential learning for better accuracy, while XGBoost utilizes TF-IDF for classification. The models are validated using precision, recall, and F1-score metrics. Results indicate that CNN-BiLSTM outperforms XGBoost, achieving an accuracy of 95.57% compared to 90.89%. This research demonstrates the potential of machine learning in mental health monitoring and provides a foundation for automated suicide prevention systems.

Keywords: Suicide Detection, Machine Learning, Deep Learning, Natural Language Processing, Social Media.

1. Introduction

Suicide is one of the leading causes of death worldwide, with millions of individuals struggling with suicidal thoughts. Many express their distress through social media, providing an opportunity to identify at-risk individuals using computational methods. Traditional mental health interventions are often reactive; however, machine learning and deep learning models can enable proactive detection and support. The increasing availability of social media data allows for scalable, automated solutions to detect distress signals. The integration of artificial intelligence into mental health monitoring presents an innovative approach that can complement traditional psychological assessments, enabling faster response times and broader accessibility. This research aims to develop an efficient suicidal ideation detection system using text analysis and classification techniques, ultimately contributing to improved mental health surveillance and intervention strategies.

2. Literature Review

Several studies have explored the use of natural language processing (NLP) and machine learning for mental health detection. Sentiment analysis and text classification methods have been employed to identify distress patterns in online content. Prior research has applied support vector machines (SVM), logistic regression, and recurrent neural networks (RNNs) for suicide detection. However, these approaches often fail to capture contextual meaning and

dependencies within text. Recent advances in deep learning, particularly convolutional neural networks (CNNs) and bidirectional long short-term memory networks (BiLSTMs), have improved classification accuracy.

Research findings suggest that deep learning models are more effective in understanding complex human emotions embedded in textual data, which is critical for accurately detecting suicidal ideation. This study builds upon previous research by implementing a hybrid CNN-BiLSTM model and comparing its performance with the XGBoost classifier, thereby demonstrating the advantages of deep learning over traditional approaches.

3. Methodology

The methodology consists of multiple stages, including data collection, preprocessing, feature extraction, model training, and evaluation.

3.1 Data Collection

The dataset used in this study consists of social media posts labeled as suicidal or non-suicidal. The dataset comprises approximately 50,000 posts, with an equal distribution between suicidal and non-suicidal content to ensure balance in training and evaluation. Data was sourced from publicly available repositories to ensure real-world applicability. The dataset is preprocessed to remove noise and irrelevant information. To maintain data integrity and reduce bias, the dataset includes a diverse range of posts from multiple social media platforms. Ethical considerations were also taken into account to ensure user privacy and data confidentiality.

	text	class
2	Ex Wife Threatening SuicideRecently I left my wife for good because she has cheated on me twice and lied	suicide
3	Am I weird I don't get affected by compliments if it's coming from someone I know irl but I feel really good	non-suicide
4	Finally 2020 is almost over... So I can never hear "2020 has been a bad year" ever again. I swear to fucking G	non-suicide
8	i need helpjust help me im crying so hard	suicide
9	Iâ€™m so lostHello, my name is Adam (16) and Iâ€™ve been struggling for years and Iâ€™m afraid.	suicide
11	Honetly idkI dont know what im even doing here. I just feel like there is nothing and nowhere for me. All	suicide
12	[Trigger warning] Excuse for self inflicted burns*I do know the crisis line and used it after when I was	suicide
13	It ends tonight.I canâ€™t do it anymore.	suicide
16	Everyone wants to be "edgy" and it's making me self conscious I feel like I don't stand out. I can draw yes	non-suicide
18	My life is over at 20 years oldHello all. I am a 20 year old balding male. My hairline is trash and to make ma	suicide
19	I took the rest of my sleeping pills and my painkillersI canâ€™t wait for it to end, Iâ€™ve struggled for the	suicide
20	Can you imagine getting old? Me neither.Wrinkles, weight gain, hair loss, messed up teeth and bones,	suicide
21	Do you think getting hit by a train would be painful?Guns are hard to come by in my country but trains are	suicide
22	death, continuedI posted here before and saw something interesting. I asked for information. You know	suicide
23	Been arrested - feeling suicidalEdit	suicide
24	Fuck the verizon smart family app I canâ€™t even watch porn privately anymore wtf why is that a feature	non-suicide
25	Iâ€™m scared. Everything just seems to be getting worse and worse. Iâ€™m young and I think Iâ€™m	suicide
26	Well, Im screwed. I locked myself in the school toilet, and can't get out. For now.	non-suicide
27	I'm fucked assignment is due tomorrow and I haven't even started yet.	non-suicide
29	yeaputting a knife to my wrist didn't give me any hesitation like how it used to, i am free from that, free	suicide
30	I am ending my life today, goodbye everyone.I am 36 almost 37, I am on disability for PTSD and	suicide
31	Me: I know I have a really toxic house and I do my best to cope with with it by going to school, etc Rona: ha	non-suicide

Fig.1 Dataset

3.2 Data Preprocessing

To enhance model performance, several preprocessing techniques were applied:

- Removing punctuation, special characters, and numbers to reduce noise in the data.
- Converting text to lowercase for uniformity and avoiding duplicate variations of the same words.
- Removing stop words to eliminate common words that do not contribute meaningfully to the classification process.

- Tokenization and sequence padding for deep learning models, ensuring that text data is transformed into a structured format suitable for training.
- Using word embeddings to convert textual data into numerical representations that capture semantic relationships between words.

3.3 Model Implementation

- **CNN-BiLSTM Model:** This hybrid model integrates CNN for feature extraction and BiLSTM for sequential pattern recognition. CNN captures local patterns, while BiLSTM learns long-range dependencies within text data. This combination allows for a robust understanding of suicidal ideation cues present in user-generated content.
- **XGBoost Model:** A gradient boosting-based classifier that uses TF-IDF vectorization for text representation. It is known for its efficiency in structured data classification, although it lacks the contextual depth that deep learning models provide.

3.4 Model Training and Evaluation

- The dataset is split into training (80%) and testing (20%) subsets to ensure the models generalize well to unseen data.
- Models are trained using supervised learning techniques with hyperparameter tuning to optimize performance.
- Evaluation metrics include precision, recall, F1-score, and accuracy to comprehensively assess model effectiveness.
- CNN-BiLSTM achieved an accuracy of 95.57%, while XGBoost attained 90.89%, demonstrating the superiority of deep learning in analyzing text-based mental health data.

4. Results and Discussion

The results indicate that deep learning models, particularly CNN-BiLSTM, perform significantly better in capturing contextual relationships in suicidal ideation text. Compared to XGBoost, which relies on TF-IDF features, CNN-BiLSTM effectively understands emotional cues and linguistic patterns associated with suicidal thoughts. The performance metrics highlight the importance of feature extraction and sequential learning in improving classification accuracy. Additionally, the model demonstrated strong generalizability when tested on new data, reinforcing its reliability in real-world applications. The findings suggest that advanced NLP techniques can enhance the detection of mental health risks on social media, paving the way for AI-driven mental health monitoring.

5. Conclusion

This study emphasizes the success of machine learning and deep learning methods in identifying suicidal thoughts expressed in social media content. The CNN-BiLSTM model excels in understanding contextual nuances, whereas XGBoost serves as a robust benchmark for evaluation. Future investigations could focus on incorporating more linguistic and psychological elements to enhance classification precision. The suggested system aids in real-time mental health surveillance and establishes a foundation for proactive intervention approaches in suicide prevention. This research underscores the promise of AI-driven mental health tools to complement conventional psychological support frameworks.

Future improvements to this research may concentrate on broadening the dataset to encompass various social media platforms, thereby enhancing generalizability. Additionally, the implementation of real-time suicide prevention systems that incorporate chatbot-based mental health support could be beneficial. Investigating transformer-based models such as BERT or GPT may also yield advancements in natural language understanding. Collaborating with

mental health professionals is essential for validating the model's effectiveness in practical settings.

Creating an interactive and explainable AI system that offers rationales for its classifications can further bolster trust in its predictions, rendering it a more dependable resource for mental health applications. Moreover, adopting multimodal strategies that integrate speech and image analysis for detecting suicidal behavior could improve the system's predictive accuracy. The development of mobile applications or web platforms that facilitate real-time monitoring and crisis intervention services could enhance the accessibility and impact of suicide prevention efforts. Finally, incorporating cultural and linguistic diversity in the training data could enhance the model's adaptability across various populations, ensuring a more inclusive and effective detection system.

References

1. Ji, Shaoxiong, et al. "Suicidal ideation detection: A review of machine learning methods and applications." IEEE Transactions on Computational Social Systems 8.1 (2020): 214-226.

2. Coppersmith, G., Ngo, K., Leary, R., & Wood, A. (2016). Exploratory analysis of social media prior to a suicide attempt. *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*, 106–117.

3. Aldhyani, T. H., Alsubari, S. N., Alshebami, A. S., Alkahtani, H., & Ahmed, Z. A. (2022). Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models. International journal of environmental research and public health, 19(19), 12635.

4. Tadesse, M.M., Lin, H., Xu, B., & Yang, L. (2019). Detection of depression-related posts in Reddit social media forum. *IEEE Access, 7*, 44883–44893.

5. Haque, Rezaul, Naimul Islam, Maidul Islam, and Md Manjurul Ahsan. "A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning." Technologies 10, no. 3 (2022): 57.

6. Sawhney, Ramit, et al. "Exploring and learning suicidal ideation connotations on social media with deep learning." Proceedings of the 9th workshop on computational approaches to subjectivity, sentiment and social media analysis. 2018.

7. Renjith, S., Abraham, A., Jyothi, S. B., Chandran, L., & Thomson, J. (2022). An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms. Journal of King Saud University-Computer and Information Sciences, 34(10), 9564-9575.
8. Sawhney, Ramit, Prachi Manchanda, Puneet Mathur, Rajiv Shah, and Raj Singh. "Exploring and learning suicidal ideation connotations on social media with deep learning." In Proceedings of the 9th workshop on computational approaches to subjectivity, sentiment and social media analysis, pp. 167-175. 2018.

9. Haque, Rezaul, Naimul Islam, Maidul Islam, and Md Manjurul Ahsan. "A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning." Technologies 10, no. 3 (2022): 57.