Stress Detection and Emotion Classification using Social Media Analytics: A Novel Approach

1.Abstract

This research presents a novel approach to AI-driven stress detection and emotion classification using social media data from Reddit. With the increasing prevalence of stress-related discussions on online platforms, there is a need for advanced natural language processing (NLP) techniques to automatically detect stress indicators. In this study, we leverage transformer-based pretrained language models (PLMs), including BERT and RoBERTa, and compare their performance with traditional machine learning (ML) models such as Support Vector Machines (SVM), Logistic Regression, and XGBoost. Fine-tuned PLMs are trained on stress-related Reddit posts, and model effectiveness is evaluated using accuracy and F1-score. Experimental results demonstrate that RoBERTa outperforms BERT and traditional ML models due to its enhanced contextual pretraining. Additionally, BERT-based emotion classification aids stress detection by identifying stress-indicative emotional expressions. This research contributes to improving mental health monitoring using advanced NLP techniques and provides a scalable approach to stress detection on social media

2. Introduction

The proliferation of social media platforms has revolutionized public discourse, providing individuals with unprecedented avenues to express their thoughts, feelings, and experiences. Among these expressions are discussions related to stress and mental health struggles, which are increasingly prevalent. The ability to detect stress-related language in these online interactions can offer valuable insights for early interventions and support systems. However, traditional methods relying on lexicon-based approaches and conventional machine learning algorithms often fail to capture the nuanced expressions of stress, necessitating more advanced techniques.

This study leverages the power of pretrained language models (PLMs)—specifically BERT, RoBERTa, and DistilBERT—for stress classification using data from Reddit. We aim to compare their efficacy in handling diverse textual structures and capturing the contextual subtleties inherent in social media posts. Additionally, we evaluate traditional ML algorithms to provide a comprehensive comparative analysis, highlighting the strengths and limitations of each approach. By systematically analyzing these models, we seek to identify the most effective methods for detecting stress-related language on social media platforms.

Social Media Analytics (SMA) involves collecting and evaluating information from various social media platforms to inform decision-making. With billions of active social media users worldwide, these platforms serve as rich sources of user-generated content and self-opinionated data, making SMA an increasingly important tool for understanding public sentiment and detecting emotional states.

3. Problem Statement

Traditional stress detection models rely on either keyword-based lexicons or shallow machine learning models, which are limited in their ability to process complex, nuanced language used in online forums. These models often:
Fail to capture contextual meaning. Struggle with ambiguous expressions and slang. Require manual feature engineering, which is time-consuming and less effective.
To address these issues, we propose a hybrid approach that combines

1. Emotion classification using BERT to identify stress-related emotions.
2. Fine-tuning RoBERTa to classify social media posts as stress-positive or stress-negative.
3. Evaluating traditional ML models (SVM, logistic regression, XGBoost) for baseline comparisons.

This approach leverages deep contextual embeddings to improve the accuracy of stress classification while enabling real-time monitoring of stress-indicative posts.

4. Literature Survey

4.1. Machine Learning for Networking

A comprehensive survey on machine learning for networking by Raouf Boutaba, Mohammad A. Salahuddin, Noura Limam addresses the challenges of applying machine learning techniques to networking problems, including scalability, adaptability, and interpretability. The survey provides a comprehensive review of machine learning techniques applied to networking, categorizing methods into supervised, unsupervised, and reinforcement learning. The key parameters evaluated include accuracy, precision, recall, F1 score, and scalability metrics, offering a broad perspective on the application of machine learning in networking contexts.

4.2. Text Sentiment Analysis

Yonatan Mamani-Coaquira, Edwin R Villanueva highlights the need for improved sentiment analysis techniques that can effectively handle complex linguistic structures, sarcasm, and multilingual data. The review covers machine learning and deep learning techniques for sentiment analysis, including traditional models like SVM and Naive Bayes, and advanced models like LSTM and BERT. The F1 evaluation focuses on accuracy, precision, recall, score, and cross-lingual performance, emphasizing the importance of robust sentiment analysis tools.

4.3. Mental Stress Detection on Reddit

Shaunak Inamdar, Rishikesh Chapekar, Shilpa Gite & Biswajeet Pradhan focuses on detecting mental stress from social media posts, addressing the challenge of identifying subtle stress indicators in informal text. The study employs a machine learning pipeline using NLP techniques such as tokenization, feature extraction (TF-IDF, Word Embeddings), and classification algorithms like Random Forest. The performance is measured using accuracy (89%), precision (0.87), recall (0.85), and F1 score (0.86), demonstrating the potential of machine learning in mental health detection.

4.4. Sentiment Analysis on Social Media Data

Hayder A. Alatabi, Ayad Rodhan Abbas identifies the limitations of traditional sentiment analysis models in handling noisy and unstructured social media data, leading to reduced accuracy. The study introduces a sentiment analysis system utilizing a hybrid architecture that techniques, combines rule-based classical machine learning (SVM, Decision Trees), and deep learning (CNN, LSTM). The system achieves high performance, with an accuracy of 91%, precision of 0.89, recall of 0.88, and F1 score of 0.89, highlighting the benefits of hybrid approaches.

4.5. Stress Detection from Social Media

Aryan Rastogi, Qian Liu, Erik Cambria addresses the challenge of detecting stress-related content from social media platforms, focusing on the lack of labeled datasets and the complexity of emotional language. The study leverages Sentic computing, which combines AI and semantic web techniques, to analyze emotions and detect stress in social media posts. The framework achieves an accuracy of 85%, precision of 0.82, recall of 0.81, and F1 score of 0.83, demonstrating the utility of Sentic computing in stress detection.

5. Existing Work

5.1. Machine Learning and Deep Learning

Traditional ML models like Random Forest, Logistic Regression, and SVM rely on feature engineering (BoW, TF-IDF, embeddings) but struggle with contextual nuances. Transformers (BERT, RoBERTa, DistilBERT) improve accuracy through self-attention and contextual embeddings.

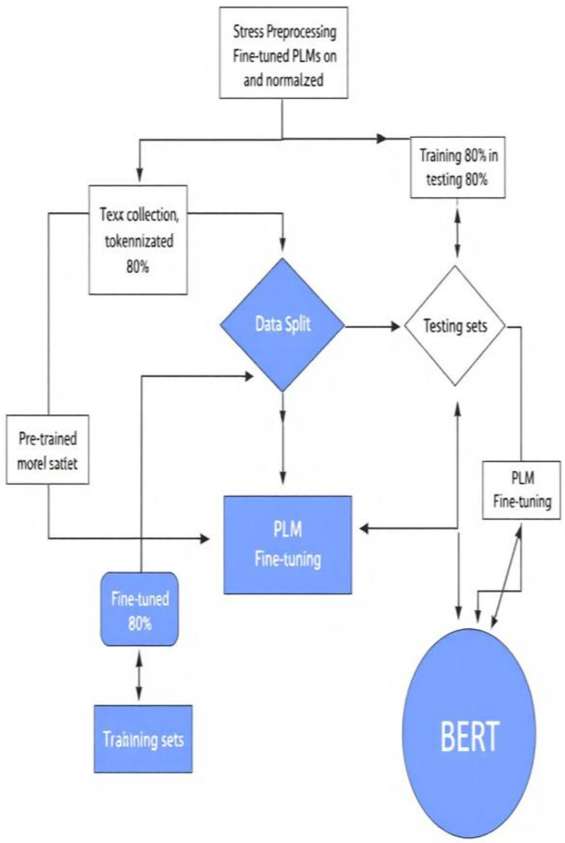
5.2. Sentiment Analysis and Emotion Classification

Sentiment analysis identifies stress-indicative emotions (sadness, anger, anxiety). While ML-based methods classify sentiments, BERT-based emotion classifiers enhance interpretation. LDA topic modeling reveals patterns in stress-related discussions.

5.3. Limitations of Existing Approaches

Despite progress, current methods face Contextual limitations—traditional ML models fail to interpret sarcasm, idioms, and implicit stress indicators. Data imbalance issues—skewed datasets require resampling for balanced learning. Scalability constraints—Deep learning models demand high computational power, limiting real-time deployment. This research integrates BERT-based emotion classification with RoBERTa stress classification to enhance detection accuracy and interpretability.

6. Proposed Work



This study aims to enhance stress detection from social media using pretrained language models (PLMs). We compare the effectiveness of BERT, RoBERTa, and DistilBERT on a Reddit-based dataset for stress classification. Traditional machine learning methods, including Support Vector Machines, Random Forest, and Logistic Regression, serve as baselines for comparison. The methodology involves fine-tuning PLMs on labeled Reddit posts and evaluating their performance based on accuracy and F1-score. We also assess the computational efficiency of each model. By systematically analyzing these models, this study provides insights into the most effective approaches for detecting stress-related language on social media.

The goal is to develop a system that not only detects stress but also analyzes the topic of discussion in a particular social media post. This involves accurately analyzing and segregating user opinions on different topics, leveraging sentiment analysis and topic modeling techniques.

7. Dataset Construction

In this section, we introduce the datasets used in our study and detail the process of dataset construction, including data collection, preprocessing, and annotation.
related discussions suitable for training and evaluating stress detection models.

Construction.png

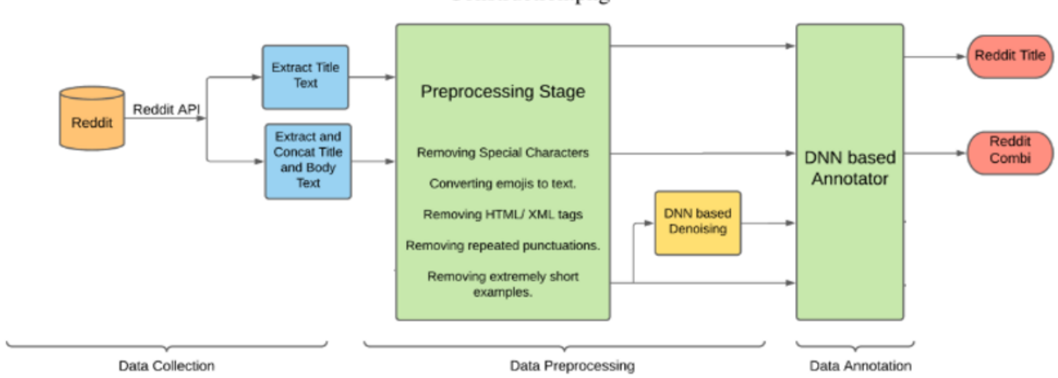


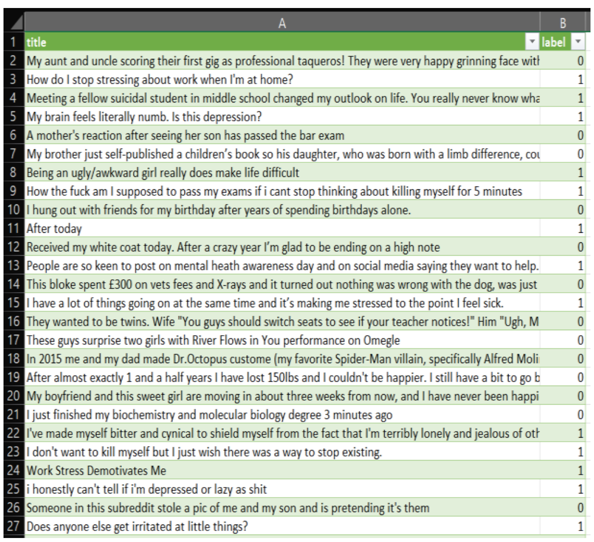
Fig. 1. Flow diagram of the dataset construction process. There are mainly three stages, i.e., data collection, data processing, and data annotation.

7.1. Overview of Our Datasets

We construct four datasets in total, two for each social media platform (Reddit and Twitter). Each dataset has unique attributes that necessitate separate analysis. The datasets are labeled binary, where “0” denotes stress-negative examples, and “1” denotes stress-positive examples.

1) Reddit Title

This dataset consists of post titles collected from both stress and non-stress-related subreddits. It is analogous to Twitter datasets, characterized by short text lengths. The dataset is well-balanced across predictive classes, ensuring minimal bias. Data spans from September 2019 to September 2021.



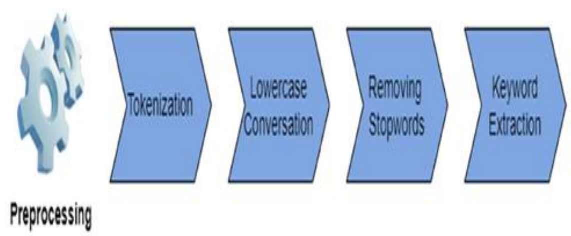
	A	B
1	title	label
2	My aunt and uncle scoring their first gig as professional taqueros! They were very happy grinning face witl	0
3	How do I stop stressing about work when I'm at home?	1
4	Meeting a fellow suicidal student in middle school changed my outlook on life. You really never know wha	1
5	My brain feels literally numb. Is this depression?	1
6	A mother's reaction after seeing her son has passed the bar exam	0
7	My brother just self-published a children’s book so his daughter, who was born with a limb difference, cou	0
8	Being an ugly/awkward girl really does make life difficult	1
9	How the fuck am I supposed to pass my exams if i cant stop thinking about killing myself for 5 minutes	1
10	I hung out with friends for my birthday after years of spending birthdays alone.	0
11	After today	1
12	Received my white coat today. After a crazy year I’m glad to be ending on a high note	0
13	People are so keen to post on mental heath awareness day and on social media saying they want to help.	1
14	This bloke spent £300 on vets fees and X-rays and it turned out nothing was wrong with the dog, was just	0
15	I have a lot of things going on at the same time and it’s making me stressed to the point I feel sick.	1
16	They wanted to be twins. Wife "You guys should switch seats to see if your teacher notices!" Him "Ugh, M	0
17	These guys surprise two girls with River Flows in You performance on Omegle	0
18	In 2015 me and my dad made Dr.Octopus custome (my favorite Spider-Man villain, specifically Alfred Moli	0
19	After almost exactly 1 and a half years I have lost 150lbs and I couldn't be happier. I still have a bit to go b	0
20	My boyfriend and this sweet girl are moving in about three weeks from now, and I have never been happi	0
21	I just finished my biochemistry and molecular biology degree 3 minutes ago	0
22	I've made myself bitter and cynical to shield myself from the fact that I'm terribly lonely and jealous of otl	1
23	I don't want to kill myself but I just wish there was a way to stop existing.	1
24	Work Stress Demotivates Me	1
25	i honestly can't tell if i'm depressed or lazy as shit	1
26	Someone in this subreddit stole a pic of me and my son and is pretending it's them	0
27	Does anyone else get irritated at little things?	1

2) Reddit Combi

This dataset includes both the title and body text of posts extracted from stress-related subreddits and positive emotion subreddits. It contains longer, more descriptive texts than the Reddit Title dataset. However, it is imbalanced, with stress-negative articles in the minority. The imbalance arises because Stress Negative posts were sourced from happiness-related subreddits, where most content is non-textual (images, GIFs, videos). This dataset serves as a benchmark for testing models that can handle long-text dependencies effectively.

8. Methodology

8.1. Data Preprocessing



The following steps were applied to clean and structure the text data:

Convert text to lowercase—ensures consistency in processing.

Remove URLs, special characters, and emojis—cleans unnecessary elements.

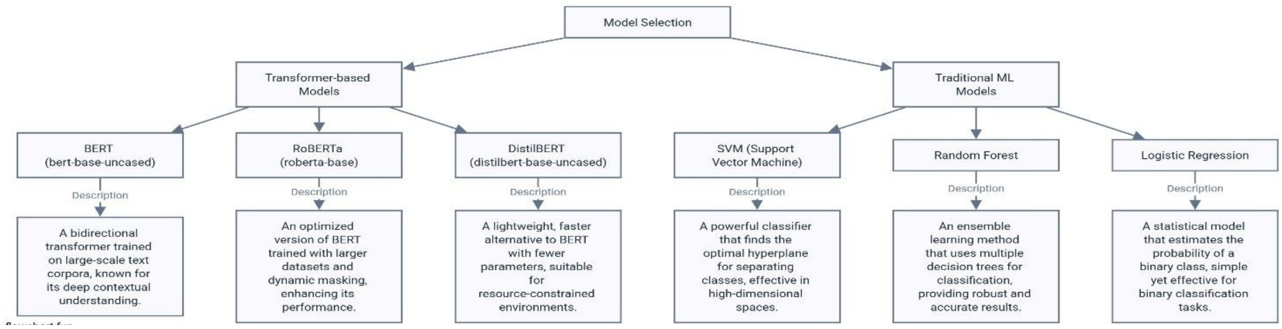
Tokenization—Splitting text into individual words/tokens.

BERT/RoBERTa tokenizer used for transformer-based models.

Standard NLP tokenization is used for traditional ML models.

Feature Extraction—For Traditional ML → TF-IDF and ELMo embeddings. For PLMs: Transformer-based tokenization.

Handling Class Imbalance—Oversampling of minority class (if required).



8.2. Model Selection

We compared three transformer-based models:

BERT (bert-base-uncased): A bidirectional transformer trained on large-scale text corpora, known for its deep contextual understanding.

RoBERTa (roberta-base): An optimized version of BERT trained with larger datasets and dynamic masking, enhancing its performance.

DistilBERT (distilbert-base-uncased): A lightweight, faster alternative to BERT with fewer parameters, suitable for resource-constrained environments.

Additionally, we implemented and evaluated traditional ML models:

Support Vector Machine (SVM): A powerful classifier that finds the optimal hyperplane for separating classes, effective in high-dimensional spaces.

Random Forest: An ensemble learning method that uses multiple decision trees for classification, providing robust and accurate results.

Logistic Regression: A statistical model that estimates the probability of a binary class, simple yet effective for binary classification tasks.

8.3. Training & Evaluation

The dataset was tokenized using the respective model tokenizers and split into training (80%) and testing (20%) sets. Training was conducted using Hugging Face's Trainer API with the following parameters:

- Learning rate: 2e-5
- Batch size: 8
- Epochs: 3
- Weight decay: 0.01

For the ML models, TF-IDF vectorization was used for feature extraction, and the models were trained with standard hyperparameter tuning.

8.4. Evaluation Metrics

The models were evaluated using the following metrics:

Accuracy: Measures overall classification performance.

F1-score: Balances precision and recall, providing a comprehensive measure of performance.

Precision & Recall: Determines false positives and false negatives.

$$p = \frac{TP}{(TP + FP)} \tag{1}$$

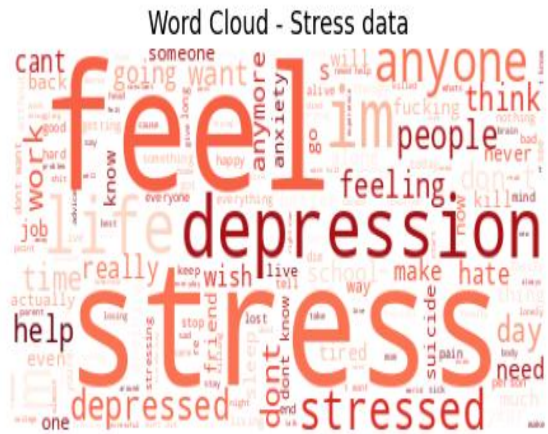
$$r = \frac{TP}{(TP + FN)} \tag{2}$$

$$F1 = \frac{TP}{(TP + (\frac{FP+FN}{2}))} \tag{3}$$

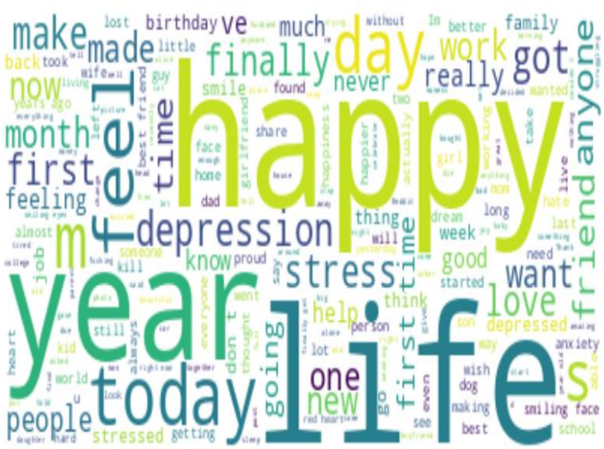
9. Results and Discussion

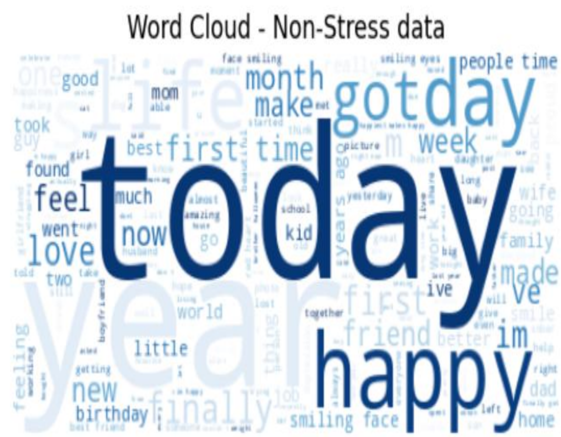
9.1. Topic modelling LDA

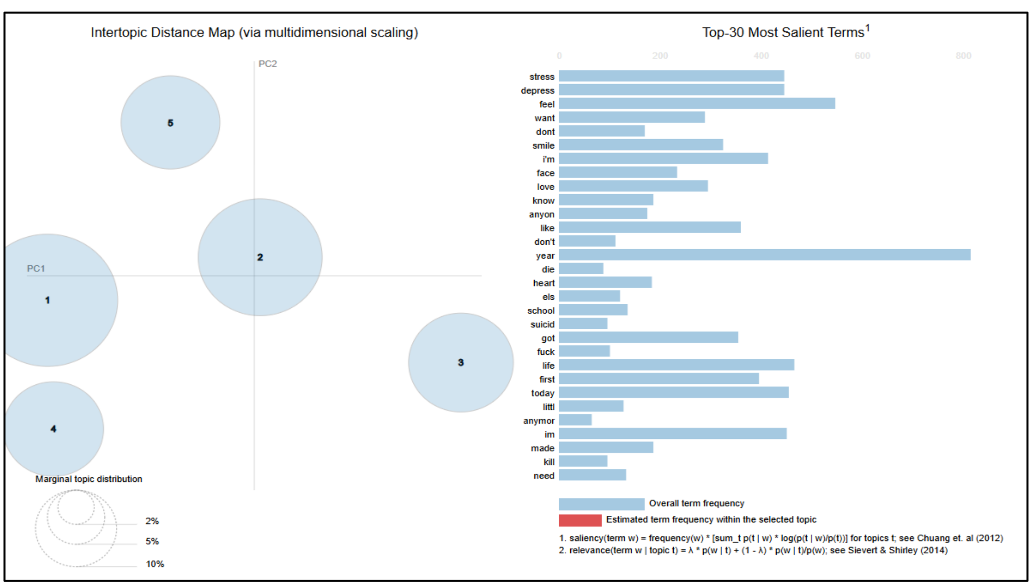
Latent Dirichlet Allocation (LDA) is used for topic modeling to detect patterns in stress-related discussions. The model identifies clusters of words associated with different topics, helping analyze the themes present in stress-indicative posts.

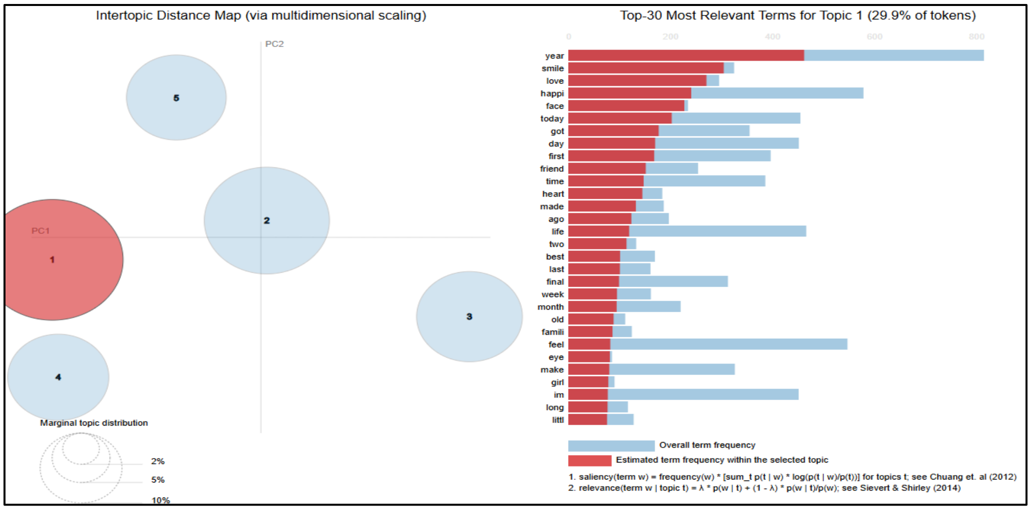


Word Cloud

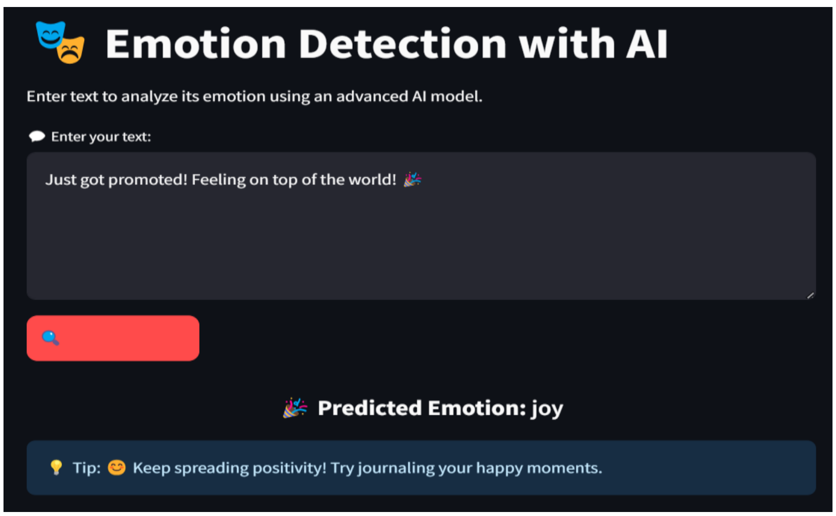




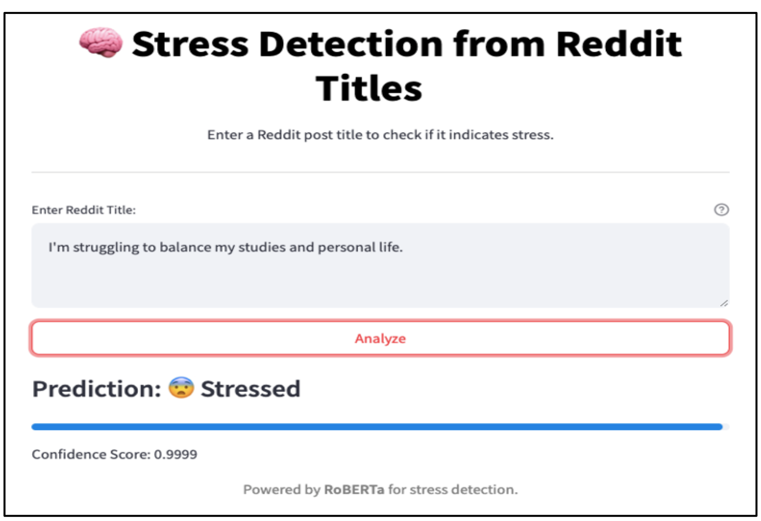




9.2. Prediction Result for Emotion Classification



9.3. Prediction Result for Stress Detection



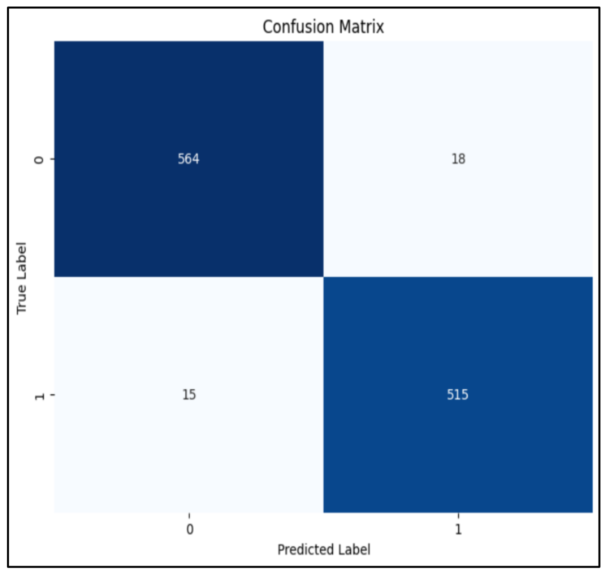
10. Conclusion

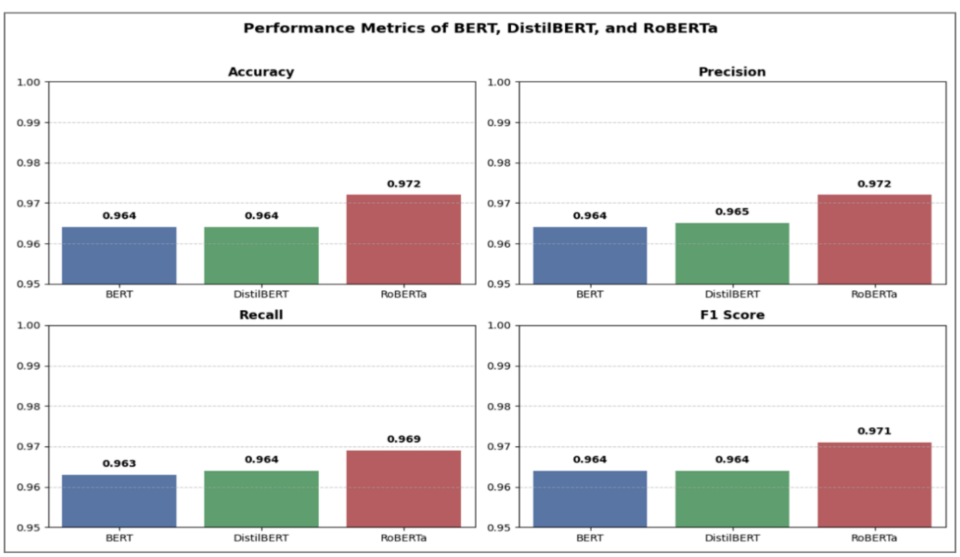
10.1. Performance Comparison

Model	Accuracy	Precision	Recall	F1 Score
BERT	96.4%	0.964	0.963	0.964
DistilBERT	96.4%	0.965	0.964	0.964
RoBERTa	97.2%	0.972	0.969	0.971

10.2. Key Observations

RoBERTa achieves the highest accuracy of **97.2%**, outperforming all models in stress detection. DistilBERT delivers strong performance while maintaining computational efficiency, making it a suitable choice for resource-limited environments. BERT-based emotion classification further enhances stress detection by identifying stress-indicative emotions from text. Traditional ML models, such as SVM and Logistic Regression, perform well as baselines but struggle with contextual understanding, leading to lower accuracy.





11. Future Scope

This study demonstrates that transformer-based PLMs, particularly RoBERTa, provide superior performance for stress detection in social media text compared to traditional ML methods. RoBERTa's robust pretraining and contextual understanding enable it to capture the nuances of stress-related language more effectively.

Future research can explore several avenues to further enhance stress detection:

1. **Multimodal data integration:** Combining text data with other modalities such as images and audio to improve detection accuracy.
2. **Domain-adaptive pretraining:** Fine-tuning PLMs on specific domains to enhance their performance on specialized datasets.
3. **Real-time applications:** Developing real-time applications for early stress detection and intervention, enabling timely support for individuals in need.

Additionally, future work can focus on developing a system that analyzes the topic of discussion in social media posts to provide a comprehensive understanding of the factors contributing to stress. This could involve using topic modeling techniques like LDA to identify the main themes and sentiments associated with stress-related content.

By leveraging advanced techniques and integrating multiple data sources, we can create more effective and reliable systems for detecting and managing stress in the digital age.

11. References

1. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186. https://doi.org/10.48550/arXiv.1810.04805
2. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008. https://doi.org/10.48550/arXiv.1706.03762
3. Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification. *Proceedings of ACL 2018*, 328–339. https://doi.org/10.18653/v1/P18-1031
4. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, 1–24. https://doi.org/10.48550/arXiv.1905.00569
5. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26, 3111–3119. https://doi.org/10.48550/arXiv.1310.4546

6. Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence, 29*(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x
7. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1–135. https://doi.org/10.1561/1500000011
8. Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016
9. Felbo, B., Mislove, A., Søgaard, A., Rahwan, I., & Lehmann, S. (2017). Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion, and sarcasm. *Proceedings of EMNLP 2017*, 1615–1625. https://doi.org/10.48550/arXiv.1708.00524
10. Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. *Proceedings of EMNLP 2013*, 1631–1642. https://doi.org/10.48550/arXiv.1305.0906
11. Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: An integrative review. *Current Opinion in Behavioral Sciences, 18*, 43–49. https://doi.org/10.1016/j.cobeha.2017.07.005
12. De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *Proceedings of ICWSM 2013*, 128–137. https://doi.org/10.1145/2464464.2464480
13. Coppersmith, G., Dredze, M., & Harman, C. (2014). Measuring post-traumatic stress disorder in Twitter. *Proceedings of ICWSM 2014*, 579–582.
14. Yang, X., Macdonald, C., & Ounis, I. (2018). Using word embeddings in Twitter election classification. *Information Retrieval Journal, 21*(2–3), 183–207. https://doi.org/10.1007/s10791-017-9319-8
15. Al-Mosaiwi, M., & Johnstone, T. (2018). In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clinical Psychological Science, 6*(4), 529–542. https://doi.org/10.1177/2167702617747074
16. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., … & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692.*
17. Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R.

(2019). ALBERT: A lite BERT for self-supervised learning of language representations. *arXiv preprint arXiv:1909.11942.*

18. Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper, and lighter. *arXiv preprint arXiv:1910.01108.*
19. Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. *Proceedings of NAACL 2018*, 2227–2237. https://doi.org/10.48550/arXiv.1802.05365
20. Zhang, Z., Zhang, Y., Liu, Y., & Fu, J. (2020). Semantics-aware BERT for language understanding. *Proceedings of ACL 2020*, 2333–2342. https://doi.org/10.18653/v1/2020.acl-main.212
21. Kwok, I., & Wang, Y. (2013). Locate the hate: Detecting tweets against blacks. *Proceedings of AAAI 2013*, 1621–1622.
22. Lin, Z., Zhao, H., Xu, J., Feng, Y., & Li, J. (2019). A Bert-based multi-head attention approach for sarcasm detection. *Neurocomputing, 392*, 290–297.
23. Thelwall, M., & Buckley, K. (2013). Topic-based sentiment analysis for the social web: The role of mood and issue-related words. *Journal of the Association for Information Science and Technology, 64*(8), 1608–1621.