Predictive Analysis Of Social Media Sentiments And Its Impact On Mental Health

By Kashmira Ghag

This dissertation is submitted for the degree of

Masters in Business Analytics (MSc. BA)

At

Dublin Business School (DBS)

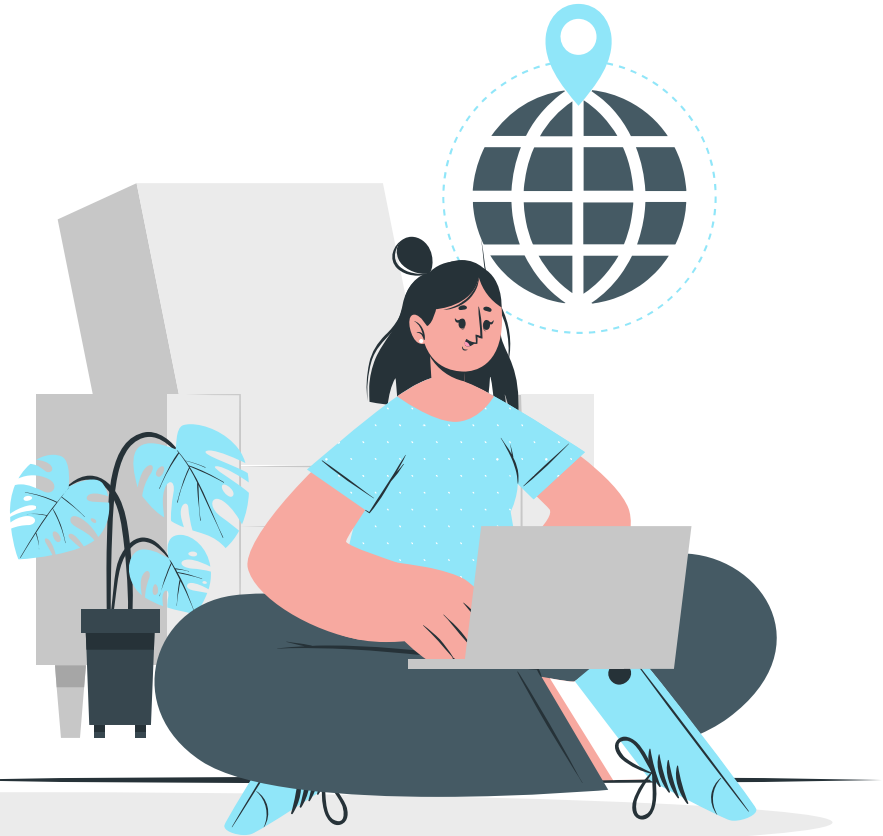
Supervisor - Paul McEvoy

Predictive Analysis of Social Media Sentiments And Its Impact on Mental Health

TABLE OF CONTENT

INTRODUCTION

Social media significantly impacts mental health, contributing to issues like anxiety and depression, especially among younger generations. This research uses machine learning to analyse sentiments from social media interactions, aiming to identify emotional patterns, predict mental health outcomes, and promote a healthier online environment.



PROBLEM STATEMENT

- In recent years social media usage has increased drastically and several works have focused on the psychological impact of social media on every age factor, very little empirical research has been undertaken to provide a scientific method of determining the impact using advanced statistical techniques.

- However, the precise mechanisms through which social media emotions influence mental health are not fully understood. To recommend actions on social media and their impact on mental health, there is a need to have robust and sound predictive models. This study aims to solve this problem by utilizing efficient techniques and tools.

RESEARCH OBJECTIVE

Analyse the emotional impact of social media comments

Utilise machine learning for sentiment analysis

To test and compare machine learning models

To track changes in emotions over time

To develop actionable insights for mental health interventions

RESEARCH QUESTION

RQ 1

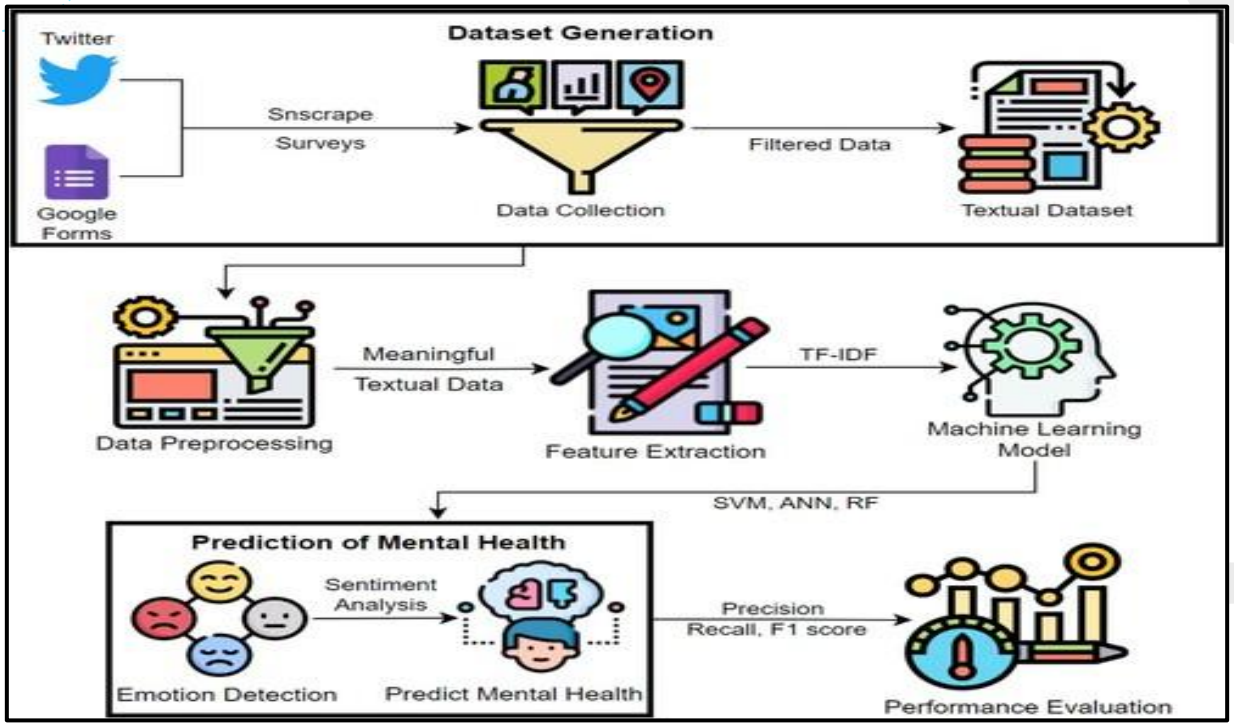
On which features can sentiment analysis be optimized to effectively assess its impact on mental health, focusing on specific issues such as depression, anxiety, or stress?

RQ 2

Can simpler machine learning models reliably predict mental health states from social media comments, and how do they compare to more complex approaches in terms of accuracy and efficiency?

DISSERTATION OUTLINE

It presents clear outline of the thesis; offering a short description of the chapters to give an overview of its content



BACKGROUND

- Sentiment Analysis in Social Media
- Impact of Social Media on Mental Health
- Machine Learning in Sentiment Prediction
- Text Classification

LITERATURE REVIEW

Research Highlights

- Some work shows that user activity, language patterns, and ego networks in these places, including Reddit and Twitter, can signal depression onset.
- Newer versions of such models, for instance, DistilBERT embeddings, have enhanced emotional tone prediction, but these need external verification.

Dataset Uniqueness

- Our dataset focuses on six major mental health categories—capturing key issues caused by social media.
- With comments from various platforms, it eliminates the need for cross-platform details, prioritizing focused categorization over complex analyses, making our study stand out.

Lack of Research & Focus

- Past works make their models more complex by adding unnecessary features into them.
- In our work, we focus on the comments as the primary aspect to examine its effect on the mental states directly, employing simple yet efficient models for outcome specificity.

METHODOLOGY

- Data Collection
- Data Cleaning and Normalization
- Feature Extraction
- Data Augmentation
- Machine Learning Models For Sentiment Analysis
- Data Preprocessing
- Text Cleaning
- Encoding
- Data Splitting
- Tools and Technologies

IMPLEMENTATION

Data Loading & Initial Analysis

Dataset is uploaded and all the required process is done on data

Handling Missing Values

Extracting potential flaws that pose a risk to dataset completion and correctness.

Data Visualization

Using visualizations to represent and illustrate patterns and trends in data

Data Mapping

Categorical data is converted into numerical formats to ensure machine learning compatibility.

Library Installation and Importation

Setting up the necessary Python libraries for data analysis and model creation.

Text Cleaning

Preprocessing text data involves removing noise, such as punctuation, stopwords, and unnecessary words.

Model Implementation

Using machine learning algorithms to evaluate sentiments and predict mental health outcomes

Evaluation Metrics

For evaluating the reliability and dependability of model predictions, we have used precision and recall

Model Evaluation

Analyzing model performance to determine the best effective strategy for sentiment analysis.

RESULTS

Evaluation Metrics/ Models	Precision	Recall
Support Vector Classifier	0.93	0.93
Naive Bayes Classifier	0.88	0.89
Logistic Regression	0.93	0.93

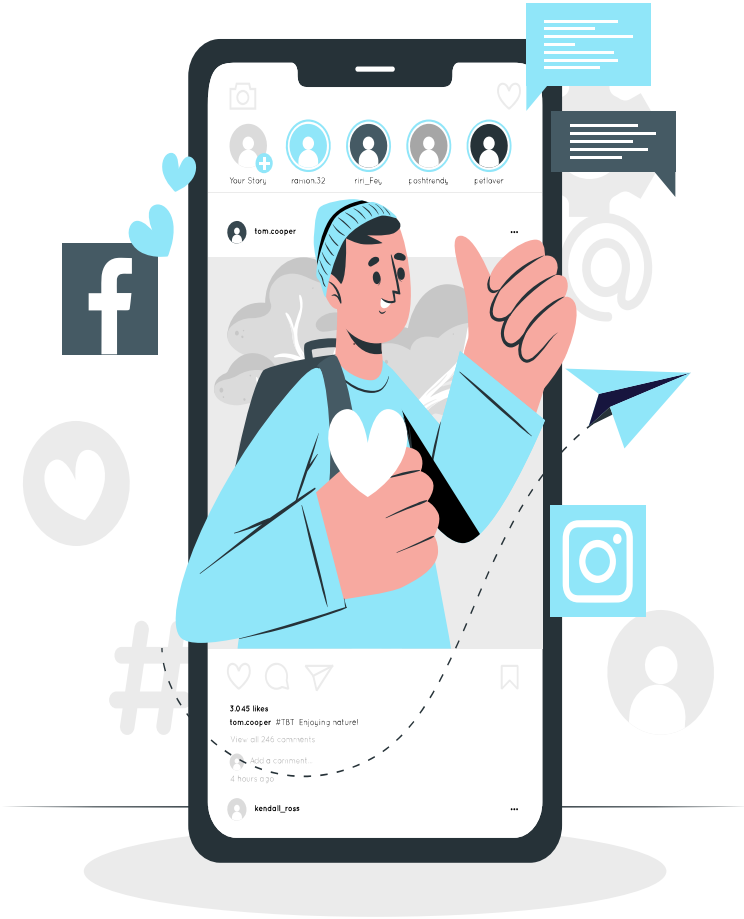
- All the models showed good performance level, SVM had the highest precision and higher recall compared to the other models.
- For sentiment analysis, the accuracy of the models was confirmed by an ability to accurately predict mental conditions.
- These results indicate that there is value in simple, yet robust models for garnering practical data for early treatment and continual psychiatric care.

CONCLUSION

- The study was able to use machine learning models when analyzing the social media sentiments while providing accurate predictions of effects on mental health like stress, anxiety, and depression levels.
- The models showed high levels of precision and recall, which prove their feasibility in a variety of sentiment analysis issues.
- The proposed models were agreed because of their simplicity and effectiveness, which is as evaluated by both the dependent variables of SM interactions and mental wellbeing.

FUTURE SCOPE

- Future research can expand on the current work by including real-time sentiment analysis to improve the responsiveness and relevance of interactions.
- Exploring different methodologies, such as deep learning and hybrid methods, can lead to even greater increases in accuracy and scalability.
- Furthermore, advancements in user-friendly devices and software for mental health screening can result to new therapy and early intervention options, benefiting society.

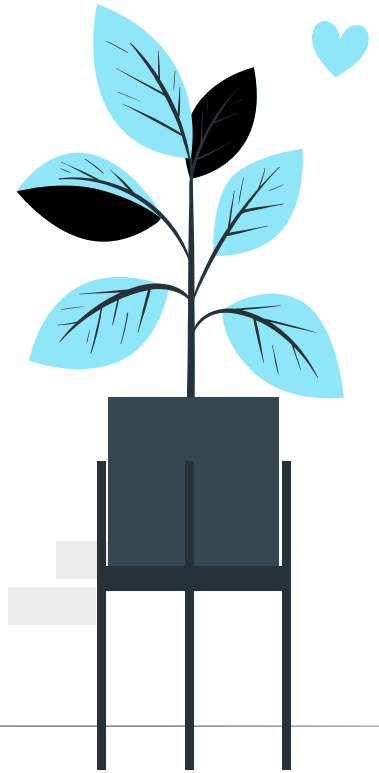


"Your mental health is a priority. Your happiness is an essential. Your self-care is a necessity."

—Melody Beattie

REFERENCES

- Kaggle, n.d. **Sentiment analysis for mental health dataset**. [online] Available at: https://www.kaggle.com/datasets/suchintikasarkar/sentiment-analysis-for-mental-health/data
- Beyari, H., 2023. The relationship between social media and the increase in mental health problems. *International Journal of Environmental Research and Public Health*, 20(3), p.2383.
- Valdez, D., Thij, M., Bathina, K., Rutter, L., & Bollen, J. (2020). Social media insights into us mental health during the covid-19 pandemic: longitudinal analysis of twitter data. Journal of Medical Internet Research, 22(12), e21418. https://doi.org/10.2196/21418
- IntechOpen, n.d. **Chapter on text classification and machine learning**. [online] Available at: https://www.intechopen.com/chapters/88300
- AAAI, 2023. **Exploring advanced social media sentiment analysis techniques**. [online] Available at: https://ojs.aaai.org/index.php/ICWSM/article/view/14432



Thank You!