Mental Health and Depression: A Comparison of Logistic Regression and Naive Bayes Models

Course: IN316: Data Mining
Instructor: Miao Qiang
Yuan Ze University (元智大学)

Group Members:

- Mikollito Ong (林米克) - 1103558
- Mlandvo - 1113523
- Rey - 1111549
- Zithile - 1113552

Abstract

This project explores how machine learning can be utilized to gain insights in mental health related **binary classification**. Using a dataset scraped from Reddit posts labeled as either related or unrelated to depression, we compare two models: **Logistic Regression** and **Multinomial Naive Bayes.** By applying Natural Language Techniques (NLP), our goal is to detect patterns that could indicate if the content is related to any mental health topic, or specifically depression related topics.

We trained two models based on a pre-labeled dataset containing two classes (**1 = Related to depression** , **0 = Not related to depression)**. The dataset was cleaned and pre-processed with tokenization, parts-speech-tagger, and lemmatization to improve model generalization.

Evaluation results demonstrate that our first model, **Logistic Regression model**, achieves **an F1-Score of 0.95**. Despite the high accuracy, we identified a good amount of false negatives. The other model, **Multinomial Naive Bayes model,** reached an **F1-Score of 0.90**. Though not as impressive as the former, its **high recall rate of 0.99** for **class 1** indicates that it performs better at accurately detecting True Positives.

Further improvements can be achieved by integrating sentiment analysis to improve the accuracy even better. Another idea for future improvement is to improve the quality of our dataset.

Though our project ***does not intend to diagnose individuals***, they offer a promising potential for **Content Tagging** , **Mental Health Monitoring (Non-Diagnostic)** and **Targeted Well Being Content Recommendations**.

Introduction

Mental health plays a critical role in human functioning and well-being. Depression is often invisible yet deeply impactful. Through this project, we do not intend to provide diagnoses but rather develop models that recognize patterns in language possibly associated with depression. Using data mining and NLP techniques, we explored how specific word choices or text structures might serve as early indicators of mental health-related topics.

Methodology

Dataset Overview
We used a dataset sourced from Kaggle, originally scraped from Reddit mental health subreddits. The dataset consisted of 7,731 posts, which we refined down to 7,289 after removing duplicates and short entries. Each post was pre-labeled as depression-related (1) or not (0), making this a supervised binary classification problem.

Data Cleaning, Preprocessing and Feature Engineering
To prepare the data for modeling, we performed the following steps:

1. **Text Cleaning**: Standardized the text by converting to lowercase,removing punctuation, numbers, and irrelevant characters.
2. **Filtering**: Removed posts with fewer than five words and duplicates.
3. **Tokenization & Stopword Removal**: Split posts into individual words, removing common words like “the” , “is” , “and” , “at” , etc.
4. **Lemmatization**: Converted words to their base forms using POS tagging.
5. **TF-IDF Vectorization**: Represented the textual data numerically using Term Frequency-Inverse Document Frequency.

These steps ensured the data was cleaned and suitable for input into machine learning models.

Model Selection

1. **Logistic Regression**
    - **Configuration**: max_iter = 1000
    - **Strength**: High interpretability, revealing which words most influence the classification outcome.
    - **Purpose**: Ideal for general screening due to its precision.
2. **Multinomial Naive Bayes**
    - **Strength**: Effective at handling word frequency, with a high recall rate.
    - **Purpose**: Useful when the goal is to capture every potential case, even at the risk of false positives.

Both model’s output is binary (**1** for depression-related, **0** for not), with associated confidence scores between 0 and 1.

Training & Evaluation

- **Train-Test Split**: 80 / 20
- **Evaluation Metrics**:
    - Accuracy
    - Precision
    - Recall
    - F1 Score

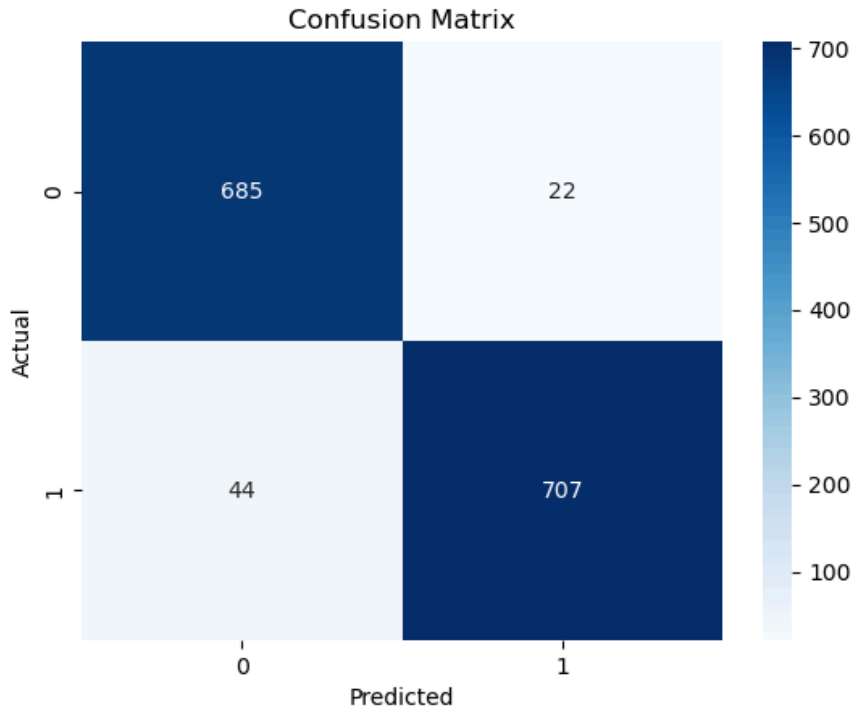
Implementation Tools

- **Python** - Main programming language
- **Jupyter** Notebook - Coding Environment
- **Anaconda** - Virtual Environment & Isolated Package
- **NLTK** - Text Processing Tool
- **Pandas** - Data Manipulation and cleaning
- **NumPy** - Numerical Operation and structure
- **Scikit-Learn** - Deploying Machine Learning models
- **Matplotlib** - Data Visualization
- **Seaborn** - Data Visualization

Results

Logistic Regression

Accuracy: 95.5%			
Class	Precision	Recall	F1-Score
0	0.94	0.97	0.95
1	0.97	0.94	0.96



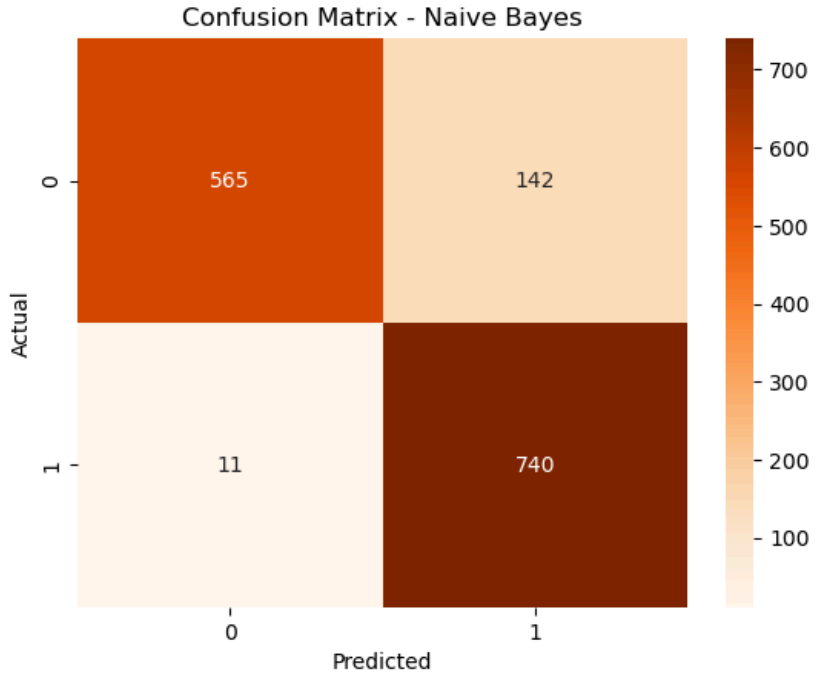
Confusion Matrix for Logistic Regression

Primary Observation:

- Strong performance in minimizing false positives, making it reliable for general use.
- High false negatives.

Multinomial Naive Bayes

Accuracy: 89.5%			
Class	Precision	Recall	F1-Score
0	0.98	0.80	0.88
1	0.84	0.99	0.91



Confusion Matrix for Naive Bayes

Primary Observation:

- Exceptional at catching true positives, thus preferred where missing a case is too risky.
- High false positives.

Conclusion: Depending on the application, a hybrid approach may offer the best balance between recall and precision.

Conclusion

This study demonstrates how text classification models can contribute to understanding mental health-related content in online communities. Logistic Regression performs well in accuracy and interpretability, while Multinomial Naive Bayes prioritizes recall, making it perform well at capturing true positives. A balanced approach that takes both model's strengths could improve accurate labelling across both classes. Most importantly, this work highlights the significance of human stories behind every data point. While not intended for professional diagnosis, Mental health online content classification can open new opportunities in terms of Content Tagging, Mental Health Monitoring (Non-Diagnostic) and Targeted Well Being Content Recommendations.

Appendix

Project Code:
The complete implementation is provided in Jupyter Notebooks (**Depression_Content_Detection.ipynb**), with all dependencies documented in a **README.txt**.

Teamwork and Job Breakdown:
Mikollito Ong :
- Data Cleaning & Preprocessing
- Naive Bayes implementation
- PowerPoint Presentation
- README preparation
- Project Report

Mlandvo :
- Feature extraction (TF-IDF)
- Project Report

Rey :
- Logistic Regression implementation and evaluation
- PowerPoint Presentation
- Project Report

Zithile :
- Data cleaning
- Lemmatization
- Model testing
- PowerPoint Presentation
- Project Report

All members contributed equally to discussions, design choices, and final conclusions.
Dataset source: https://www.kaggle.com/datasets/infamouscoder/depression-reddit-cleaned