**Harnessing NLP to Detect Stress**

**in**

**Social Media**

**Early Intervention for Mental Wellbeing**

**Chong Chee Chen (Jimmy)**

**19 Oct 2024**

Table of Contents

[1. Problem Statement 4](#_heading=h.1fob9te)

[2. Industry & Domain 5](#_heading=h.3znysh7)

[2.1. Value Chain 5](#_heading=h.gqrr3vykc9ld)

[2.2. Key Concepts 5](#_heading=h.3h42pv57zmip)

[2.3. Potential Applications 5](#_heading=h.wzpw6v2cfowq)

[3. Stakeholders 7](#_heading=h.2et92p0)

[3.1. Stakeholder Concerns and Motivations 7](#_heading=h.eezbauvz6dx7)

[3.2. Stakeholders’ Expectations 7](#_heading=h.x9qqdtolvj3z)

[4. Business question 9](#_heading=h.tyjcwt)

[4.1. Core Business Question 9](#_heading=h.15w4qbd3m4fl)

[4.2. Quantifying Business Value 9](#_heading=h.7o69kjyv5z9q)

[4.3. Accuracy and Implications 9](#_heading=h.f1dlg3ewcxjt)

[5. Data question 11](#_heading=h.3dy6vkm)

[5.1. Core Data Question 11](#_heading=h.36w1imc9slvz)

[5.2. Essential Data Requirements 11](#_heading=h.m2sfxh8cl5g3)

[6. Data 12](#_heading=h.1t3h5sf)

[6.1. Data Origin 12](#_heading=h.jtfzpsr5jemr)

[6.2. Data Volume and Attributes 12](#_heading=h.bojgqced7jp)

[6.3. Data Reliability 12](#_heading=h.fr6p5zz2nm6j)

[6.4. Raw Data Quality 12](#_heading=h.kkgv1ktoxhc6)

[6.5. Data Generation 13](#_heading=h.rjw36hpnpzvo)

[6.6. Data Availability 13](#_heading=h.9uq2u4cbmrbf)

[7. Data science process 14](#_heading=h.4d34og8)

[7.1. Data analysis 14](#_heading=h.2s8eyo1)

[7.1.1. Data Wrangling Pipeline 14](#_heading=h.tltj3777buym)

[7.1.2. EDA Highlights 14](#_heading=h.rgqlwtol2ee0)

[7.1.3. Pipeline Reusability 14](#_heading=h.2r1rg5pq9xrb)

[7.1.4. Intermediary Data Structures 14](#_heading=h.qupfqensmp0q)

[7.2. Modelling 15](#_heading=h.17dp8vu)

[7.3. Outcomes 15](#_heading=h.3rdcrjn)

[7.4. Implementation 15](#_heading=h.26in1rg)

[8. Data answer 16](#_heading=h.lnxbz9)

[9. Business answer 17](#_heading=h.35nkun2)

[10. Response to stakeholders 18](#_heading=h.1ksv4uv)

[11. End-to-end solution 19](#_heading=h.44sinio)

[12. References 20](#_heading=h.2jxsxqh)

# 

# Problem Statement

This project investigates the feasibility of automatically detecting stress in social media text using Natural Language Processing (NLP) techniques. Early detection of stress can be crucial for improving mental health outcomes. Social media platforms offer a vast amount of user-generated text data that can be potentially analysed to identify individuals who might be experiencing stress.

Currently, identifying stress in individuals often relies on self-reporting or clinical assessments. These methods can be subjective and time-consuming. Analysing social media data offers a potential avenue for more objective and scalable stress detection.

The desired state is to develop a reliable and efficient NLP-based system that can automatically detect stress in social media text. This system could be used to:

* Flag individuals who might be experiencing stress for further evaluation or intervention.
* Provide targeted support resources to those in need.
* Gain insights into population-level stress trends.

There has been some research on stress detection using social media data. Existing studies have explored various NLP techniques, such as sentiment analysis. This project aims to build upon this existing work by exploring the use of more advanced techniques like TF-IDF vectorization and machine learning models for improved accuracy.

# Industry & Domain

The project primarily falls within the **Natural Language Processing (NLP)** and **Mental Health** domains. Specifically, it focuses on applying NLP techniques to social media data to detect stress.

The NLP industry is rapidly evolving, driven by advancements in machine learning and deep learning algorithms. While there have been significant breakthroughs, challenges remain in areas such as handling ambiguity, context understanding, and domain adaptation.

* **Mental Health:** The mental health landscape is facing increasing pressure due to rising rates of mental health disorders and the stigma associated with seeking help. There is a growing need for innovative solutions that can provide accessible and effective mental health support.

## Value Chain

* **Data Collection:** Gathering social media data from platforms like Reddit and Twitter.
* **Data Preprocessing:** Cleaning, filtering, and preparing the data for analysis.
* **Feature Engineering:** Creating relevant features from the text data (e.g., TF-IDF vectors, word embeddings).
* **Model Development:** Training and evaluating NLP models (e.g., LinearSVC, LSTM) to classify text as indicative of stress or not.
* **Deployment:** Integrating the trained model into a practical application, such as a chatbot or a web-based tool.
* **Evaluation and Refinement:** Continuously monitoring the model's performance and making improvements based on feedback.

## Key Concepts

* **NLP:** Tokenization, stemming, lemmatization, stop word removal, feature extraction (TF-IDF, word embeddings), machine learning algorithms (e.g., SVM, Naive Bayes, neural networks).
* **Mental Health:** Stress, anxiety, depression, emotional well-being, mental health resources.

## Potential Applications

* **Customer Service:** Analysing customer feedback to identify areas of frustration or dissatisfaction.
* **Human Resources:** Monitoring employee sentiment and identifying potential signs of burnout.
* **Healthcare:** Analysing patient records to detect early signs of mental health conditions.
* **Marketing:** Understanding consumer sentiment towards products or brands.

# Stakeholders

* **Individuals:**
  + Social media users who may be experiencing stress.
  + Mental health professionals and organisations.
  + Family members and friends of individuals struggling with mental health.
* **Organisations:**
  + Social media platforms (e.g., Reddit, Twitter).
  + Mental health research institutions.
  + Technology companies developing AI-powered mental health solutions.
  + Healthcare providers and insurers.
  + Government agencies responsible for mental health initiatives.

## Stakeholder Concerns and Motivations

* **Individuals:**
  + Early detection of stress can lead to timely intervention and improved mental health outcomes.
  + Access to personalised support resources can help individuals cope with stress and prevent it from escalating into more severe conditions.
* **Mental health professionals and organisations:**
  + This technology can supplement traditional mental health assessments and provide a more objective and scalable approach to identifying individuals at risk.
  + Early intervention can reduce the burden on mental health services and improve overall outcomes.
* **Organisations:**
  + Social media platforms have a responsibility to promote the well-being of their users.
  + Technology companies can leverage their expertise to develop innovative solutions for mental health challenges.
  + Healthcare providers and insurers can benefit from tools that can help identify individuals who may require mental health services.
  + Government agencies can use this technology to inform public health policies and allocate resources effectively.

## Stakeholders’ Expectations

* **Individuals:**
  + Accuracy and reliability of the stress detection system.
  + Privacy and confidentiality of personal data.
  + Accessibility and ease of use of the technology.
* **Mental health professionals and organisations:**
  + Integration with existing mental health services.
  + Evidence-based approach to the development and validation of the technology.
  + Ethical considerations regarding the use of AI in mental health.
* **Organisations:**
  + Scalability and cost-effectiveness of the solution.
  + Positive impact on user engagement and well-being for social media platforms.
  + Contribution to advancements in AI and mental health research.
  + Alignment with ethical guidelines and regulations related to data privacy and mental health.

# Business question

## Core Business Question

Can we accurately detect stress in social media text using NLP techniques, and if so, how can this information be used to improve mental health outcomes?

## Quantifying Business Value

* **Social media platforms:** By identifying users who may be experiencing stress, platforms can provide targeted support resources and potentially reduce negative outcomes such as self-harm or online harassment. Assuming a 1% reduction in user churn due to improved mental health support, and considering a hypothetical platform with 100 million active users, this could translate to a significant increase in user retention and revenue.
* **Mental health organisations:** Accurate stress detection can help identify individuals who may benefit from mental health services earlier, leading to improved outcomes and reduced costs associated with treating severe mental health conditions. Assuming a 10% reduction in the cost of treating severe mental health conditions for individuals identified early through stress detection, and considering a hypothetical population of 1 million individuals at risk, this could result in substantial cost savings for mental health organisations.
* **Technology companies:** Developing a successful stress detection system can position technology companies as leaders in the emerging field of AI-powered mental health solutions, opening up new market opportunities and partnerships.

## Accuracy and Implications

The required accuracy for a stress detection system depends on the specific application and the consequences of false positives and false negatives. However, in the context of mental health, a high degree of accuracy is crucial to minimise the risk of misidentifying individuals who may be experiencing stress.

* **False positives:** False positives can lead to unnecessary distress and stigma for individuals who are not actually experiencing stress. It is important to balance the need for sensitivity in detecting stress with the potential for false alarms.
* **False negatives:** False negatives can result in missed opportunities for intervention and support for individuals who are struggling with stress. This can have serious consequences for mental health, including increased risk of self-harm and suicide.

Therefore, the ideal stress detection system would achieve a high level of both sensitivity and specificity, minimising both false positives and false negatives.

# 

# Data question

## Core Data Question

The primary data question is: Can we effectively extract features from social media text that accurately predict stress levels?

## Essential Data Requirements

To answer this question, the following types of data are essential:

* **Social media text:** A large dataset of text posts from platforms like Reddit and Twitter.
* **Stress labels:** Corresponding labels indicating whether the text post is associated with high stress levels (1) or not (0).
* **Metadata:** Additional information about the posts, such as user demographics, time of posting, and engagement metrics.
* **Annotated data:** A subset of the data that has been manually annotated by human experts to ensure accuracy and reliability of the stress labels.
* **Benchmark datasets:** Established datasets for stress detection in text to compare the performance of the developed model.

## 

# Data

## Data Origin

The data was sourced from Reddit and Twitter, and subsequently made available on Kaggle.

## Data Volume and Attributes

The dataset consists of four main components:

* **Reddit Title:** Contains titles of articles collected from stress-related and non-stress-related subreddits on Reddit.
* **Reddit Combi:** Combines titles and body text of articles from both stress-related and non-stress-related subreddits on Reddit.
* **Twitter Full:** Includes stress-related and non-stress-related tweets collected from Twitter.
* **Twitter Non-Advert:** A denoised version of the Twitter Full dataset, excluding promotional content.

## Data Reliability

The reliability of the data depends on several factors:

* **Data quality:** The quality of the text data can vary depending on the source (e.g., user-generated content on social media) and may contain noise, inconsistencies, or biases.
* **Sample size:** A larger dataset can generally improve the reliability of the results, but even with a large dataset, there may be limitations due to the inherent variability of human language and behaviour.

## Raw Data Quality

The raw data contain a variety of issues, including:

* **Noise:** Spelling errors, typos, and grammatical mistakes.
* **Inconsistencies:** Variations in language style, tone, and vocabulary.
* **Biases:** Personal biases of the users who generated the content.
* **Spam and promotional content:** Irrelevant or misleading information.

## Data Generation

The data was generated by users posting content on Reddit and Twitter. The specific methods used to collect and preprocess the data may vary depending on the source and the researchers' procedures.

## Data Availability

The availability of ongoing data depends on the policies of Reddit and Twitter. However, given the nature of social media platforms, it is likely that new data will continue to be generated and made available over time.

# Data science process

## Data analysis

### Data Wrangling Pipeline

* **Loading:** Loading the datasets from Excel files into Pandas DataFrames.
* **Exploration:** Inspecting the data to understand its structure, dimensions, and content.
* **Preprocessing:** Cleaning the data by handling missing values, removing unnecessary columns, and normalising text.
* **Feature engineering:** Creating new features from the text data, such as TF-IDF vectors.
* **Splitting:** Dividing the data into training and testing sets.

### EDA Highlights

* **Data distribution:** Analysing the distribution of stress labels and text lengths.
* **Word clouds:** Visualising the most common words used in stressed and non-stressed text.

### Pipeline Reusability

The pipeline is generally reusable for processing future data. The code can be adapted to load new data, apply the same preprocessing steps, and train and evaluate the models. However, the specific steps and parameters may need to be adjusted based on the characteristics of the new data.

### Intermediary Data Structures

The intermediary data structures used in the pipeline include:

* **Pandas DataFrames:** To store and manipulate the data.
* **NumPy arrays:** For numerical operations and machine learning algorithms.
* **TF-IDF vectors:** To represent text data as numerical features.
* **Lists and dictionaries:** For storing intermediate results and metadata.

### 

## Modelling

* What are the main features used?
* Did you find any interesting interactions between features?
* Is there a subset of features that would get a significant portion of your final performance? Which features?
* How did you select features?
* What feature engineering techniques are used?
* What are the models used?
* How long does it take to train your model?
* What are the tools used? (cloud platform, for example)
* What are the model performance metrics?
* Which model was selected?

## Outcomes

* What are the main findings and conclusions of the data science process?

## Implementation

* What are the considerations for implementing the model in production?

# Data answer

* Was the data question answered satisfactorily?
* What is the confidence level in the data answer?

# Business answer

* Was the business question answered satisfactorily?
* What is the confidence level in the business answer?

# Response to stakeholders

* What are the overall messages and recommendations to the stakeholders?

# End-to-end solution

* What is the overall end-to-end solution to use the model developed in the project?

# References

* <https://www.kaggle.com/datasets/mexwell/stress-detection-from-social-media-articles>