CHAPTER 1

INTRODUCTION

This chapter includes the following subtopics, namely: (1) Rationale; (2) Theoretical Framework; (3) Conceptual Framework/Paradigm; (4) Statement of the problem; (5) Hypothesis (Optional); (6) Assumption (Optional); (7) Scope and Delimitation; and (8) Importance of the study.

1.1 Rationale

Depression is an issue worldwide that affects millions with its serious impact on health and well being. Depression detection is the task of determining whether a human being is depressed or non-depressed by using some features or indicators. Ideally, machine learning for detecting depression is trained with the same indicators (data) that doctors use in the process of forming a diagnosis. The features used for the classification include face and gesture, voice, and language [10]. One of the features used in this study is a linguistic indicator or detecting depression from writing.

Social media platforms such as Twitter have become an outlet for users to share their emotions in brief and casual posts. These messages offer a chance, for automated systems to identify depression indicators. The informal style of tweets presents a challenge for accurate interpretation using machine learning models. The use of slang terms, shortened words, and spelling errors in tweets may impede the accuracy of the interpretation process.

The text was taken from the X data. X (Twitter) is one of the largest micro-blogging services that allows users to post a tweet. With X being a service where users can write anything, X has become a place for users to express their feelings, emotions, and complaints about a problem, including health problems such as depression. Therefore, this study focuses on detecting depression from the writings of X users. Detecting mental illness / depression in social media presents multiple challenges:

1. Unstructured data : Social media posts are frequently unstructured and feature informal language, spelling errors, and grammar mistakes. As a result, extracting valuable information and identifying potential indicators of mental illness can be difficult[5].

2. Detecting mental illness requires an understanding of individuals' experiences and their contextual backgrounds. It can be challenging to interpret social media posts with accuracy and account for individual differences[5].

3. Feature engineering is also necessary. Extracting meaningful features from social media posts to develop accurate predictive models is a complex task. The identification

of the most valuable features and their consequential contribution to the improved accuracy are challenging[2].

Overall, considering tweets are frequent use of abbreviations and slang, and lack of grammatical correctness, tweets often present a particular challenge. The utilization of particular slang phrases in the context of social media posts, specifically on the platform Twitter, has the potential to result in vocabulary mismatches [12]. Based on the vocabulary mismatch problem, several strategies have been implemented, including 'enriching' features by combining User-based, Content-based, and LIWC Features. Then, feature expansion is also implemented along with Feature Enrichment. This study uses Tweet data from user X who has agreed to participate in the DASS-42 survey. Then, the label of depression or not is obtained from scaling the value generated automatically by the DASS-42 questionnaire which has several levels of severity. This research tries to detect whether a user is depressed or not from the tweet data and user profile.

1.2 Theoretical Framework

Depression detection is the task of determining whether a user is depressed or non-depressed by using some features or indicators. There are several indicators to detect depression, including visual, speech, linguistic, and multimodal [10]. Visual indicators have been used widely for depression analysis, including body movements, gestures, subtle expressions, and periodic muscular movements. On the other hand, speech indicators are very complex because the system must consider changes in intonation and acoustics of the voice.

The features or indicators used in this study is a linguistic indicator or detecting depression from writing (text data taken from tweet) and user behaviour. This research focuses on implementing feature enrichment for the task of identifying depression. The system that is implemented can reduce the vocabulary mismatch that is usually found in tweet data. Feature enrichment that is implemented includes User-based Features, Content-based Features, and LIWC-Features.

1.3 Conceptual Framework/Paradigm

User-based and content-based features are types of features used in Depression Detection Systems to identify signs of depression in individuals. When it comes to user-based features, they center on how individuals interact with technological tools and online platforms. User-based features, also known as user profiles in studies[13], or user activity features in [1], are characteristics that can be derived from user behavior, such as the number of followers, following, and tweets.

In contrast, content-based features revolve around the characteristics of the language they use, such as words, sentence structure, and grammar[1]. These features indicate a user's level of activity on social media and whether they are active or passive. In the other hand, LIWC Features are features extracted from a text using LIWC analysis tools. Linguistic Inquiry and Word Count (LIWC) is a text analysis software used to detect depression by analyzing individual language usage and identifying potential depression-indicative patterns[4].

1.4 Statement of the Problem

Detecting depression through media platforms such as Twitter presents notable obstacles because of the casual and unorganized nature of content posted by users (like X). Tweets commonly feature shortened words and phrases and informal language usage, like slang or typos that differ from grammar conventions. This issue of vocabulary complicates the functionality of conventional machine learning algorithms that depend on precise word matches or basic lexical attributes.

An advanced system for detecting depression could use feature expansion methods like word embedding to grasp the connections between words meanings. For example "sad", "unhappy, "Depressed" could be interchangeable in users language patterns. Without recognizing these similarities, between terms a model may struggle to flag tweets that hint at depression. Expanding features through the utilization of word embedding like FastText or Word2Vec aids in linking words, within similar vector spaces that minimize vocabulary discrepancies and enhance classification precision.

Due to their sparse layout, frequent use of abbreviations and slang, and lack of grammatical correctness, tweets often present a particular challenge. Machine learning frequently encounters vocabulary mismatches in capturing the meaning of the user's writing (in this case, it is used to detect depression) and produces less than-optimal accuracy. The implementation of feature enrichment, utilizing user-based, content-based, and LIWC features, along with feature extraction and expansion, will result in improved accuracy in depression detection systems using tweet data on X. The basic classification methods to be used are SVM, Naïve Bayes, KNN, Decision Tree, and Random Forest.

1.5 Objective and Hypotheses

The objective of this research is to enhance the accuracy of depression detection systems on Twitter by addressing vocabulary mismatches through feature enrichment and feature expansion techniques. Specifically, the study aims to evaluate the effectiveness of combining user-based, content-based, and LIWC features with advanced word embedding models like FastText in improving the system's ability to accurately classify depressive and non-depressive tweets.

The hypothesis of this study is that the combination of user-based, content-based, LIWC and feature expansion using word embedding will have significant impact to the accuracy of systems for detecting depression by overcoming the limitations of vocabulary mismatch in tweets.

1.6 Assumption

Depression is a common health condition that affects people globally. Its defining features are persistent feelings of sadness, hopelessness, and disinterest in previously enjoyable activities. The task of depression detection involves determining whether an individual is depressed or not based on certain features or indicators. The depression detection system must be able to detect people who are depressed by using linguistic indicators taken from text and user profiles. This research assumes the propose method can handle the mismatch problem that can improve the performance of the classification system.

1.7 Scope and Delimitation

Depression detection can be classified into three different prediction targets depressed or not depressed, severity (normal, mild, moderate, severe, and very severe), and score level prediction [10]. In this study, it is only limited to predicting whether the user is depressed or not. The scope of this study focuses on dealing with vocabulary mismatch. These issues should be rectified as they impact the results of depression classification. Vocabulary mismatch is caused by the fact that tweet data tends to have many variations of words, both slang and abbreviations, then many tweets are short, and grammatically inappropriate. The dimension of the tweet data will become sparse due to these problems. Therefore, Feature Enrichment aims to enrich these features so that the classifier does not misunderstand whether the person is depressed or not.

1.8 Significance of the Study

This research, which implements Feature Enrichment, aims to improve accuracy in the scope of detecting a person's depression with handwriting indicators. This research focuses on handling vocabulary mismatches that often occur in text datasets, especially Tweet data.