RSGC
Royal St. George's College

The Young Researcher

2021 Volume 5 | Issue 1

Designing an Algorithm to Detect Depression in Users: A Quantitative Correlational Study

Taylor Hodan

Recommended Citation
Hodan, T. (2021). Designing an algorithm to detect depression in users: A quantitative correlational study. *The Young Researcher*, 5 (1), 38-53. http://www.theyoungresearcher.com/papers/hodan.pdf

ISSN: 2560-9815 (Print) 2560-9823 (Online) Journal homepage: http://www.theyoungresearcher.com

Designing an Algorithm to Detect Depression in Users: A Quantitative Correlational Study

Taylor Hodan

Although the diagnosis of mental disorders like depression has improved over the last decade, many cases continue to go undetected. The symptoms are often observable on social media platforms. This study seeks to address this issue by designing a program to predict the likelihood and severity of depression in users by analyzing their Twitter and iMessage histories. This study employs correlational research to develop and test the program. This research used two phases to measure the program's accuracy on different sources of data. In phase one, the program was tested on a sample of 2,741 Reddit posts and comments. The program achieved an accuracy rate of 94.80% across several subreddits. Suicidal behaviour/ideation and a depressed mood had the strongest correlations with depression (coefficients of 0.654997 and 0.58218, respectively). Phase two involved testing the program on 6 grade 12 students. During this phase, the program had a recall, precision, and F1 of 1.0, 0.5, and 0.67, respectively. The results suggest that ERDS should integrate data from platforms like Twitter, iMessage, and Reddit because they reflect users' mental states.

Keywords: depression, early risk detection system, algorithm, Reddit, iMessage

1. Introduction

According to the World Health Organization, "depression is a common mental disorder [that affects] more than 264 million people of all ages" (2020). It is the "leading cause of disability" and can differ in severity as it may only cause occasional mood fluctuations for some, while for others it can lead to persistent feelings of guilt and low self-worth (World Health Organization, 2020). However, Sharifa Williams, Instructor of Data Analysis and Applied Regression at Columbia University, and her associates argue that the most troubling statistic is that "around two-thirds of all cases of depression are undiagnosed" (Williams et al., 2017, p. 633). This is especially concerning because undiagnosed and thus untreated depression often significantly diminishes quality of life (Williams et al., 2017, p. 633). Furthermore, a cross-sectional study in the Medical Science Monitor concluded that the prevalence of anxiety and depression has doubled in areas that are directly affected by the COVID-19 lockdowns (Lei et al., 2020). This study aims to address this issue by using correlational research to examine the extent to which an algorithm can accurately detect signs of depression in users' data from their social media and iMessage. The program analyzes users' data and predicts the likelihood and severity of their depression. By alerting users to their symptoms, this program aims to mitigate the challenges and prevalence of people being unaware of their depression.

2. Literature Review

2.1. Overview of Depression

Depression is a common, yet severe, mood disorder that is classified as a major depressive disorder (MDD) or a depressive episode (World Health Organization, 2020). Alan Gelenberg, professor of psychiatry at the University of Wisconsin, and his colleagues explain that a depressive episode features "a

depressed mood or loss of interest or pleasure in usual activities that persists over a period of at least 2 weeks [but less than 2 years] and is accompanied by a constellation of depressive symptoms" (Gelenberg et al., 2010, p. 77-78). This differs from MDD which occurs when depressive symptoms like "indecision, thoughts of death or suicide, or feelings of worthlessness…are present for at least 2 years" (Gelenberg et al., 2010, p. 78). Furthermore, Lynn Boschloo, assistant professor of psychology at Vrije Universiteit Amsterdam, and her colleagues elaborate, classifying these symptoms as significant or insignificant factors in detecting depression (Boschloo et al., 2016). They conducted a prospective study analyzing the relationship that each symptom has with predicting MDD. They found that the most significant factors are fatigue, loss of interest or pleasure, difficulty concentrating, and a depressed mood. Additionally, hypersomnia and an increase or decrease in weight are the least significant symptoms (Boschloo et al., 2016, p. 184). Also, Anita Thapar, professor of adolescent psychiatry at Cardiff University, and her associates performed a narrative review examining depression in adolescents (Thapar et al., 2012). They argue that the significant and insignificant symptoms in predicting MDD are almost identical to those in adults except for difficulty concentrating, which is considered a minor factor. Thus, there are various symptoms of depression that can be used to predict it; however, they differ in significance between age groups.

2.2. Assessments to Clinically Diagnose Depression

The most common practices that are used to clinically diagnose depression are surveys and interviews. Brett Thombs, professor of psychology at McGill University, and his associates conducted a systematic review of these instruments. They found that a commonly used tool is the Patient Health Questionnaire (PHQ), which utilizes a cut-off score of 10 to identify depression. They claim that other popular assessments are interviews that involve trained psychiatrists. Although these consults are more expensive to administer, they often achieve a higher degree of accuracy as they rely on healthcare professionals (Thombs et al., 2019). Munmun De Choudhury, associate professor of computing at the Georgia Institute of Technology, and his colleagues note that the Center for Epidemiologic Studies of Depression (CESD) Scale is another "tool [that is used] to determine the depression levels" of its users through a self-report scale (De Choudhury et al., 2013, p. 3). Although this means that users can manipulate their responses, which would decrease the accuracy of this instrument, it is more accessible than alternative options as it does not require a psychiatrist. Thus, various tools can be used to diagnose depression; however, each has its own advantages and disadvantages.

2.3. Computer Programs that Detect Depression in Users

Over the past twenty years, many researchers have begun developing early risk detection systems (ERDS) that involve algorithms to predict whether users have depression by analyzing various types of data including linguistic and photographic for depressive symptoms. Sharath Chandra Guntuku, assistant professor of information science at the University of Pennsylvania, and his colleagues claim that the accuracy of these programs is often measured by whether the prediction matches the result of a clinical assessment (Guntuku et al., 2017). Yena Lee, Research Assistant at the University Health Network in Canada, and her associates argue that ERDS that rely on a single source of data for analysis cannot predict depression in users with a high degree of accuracy as they do not provide sufficient information about a user's mental state (Lee et al., 2018). David Losada, professor of computer science at the University of Santiago de Compostela, and Fabio Crestani, professor of information retrieval at the Universitá della Svizzera Italiana, acknowledge the limitations of utilizing only one data source. However, they refute Lee et al.'s argument, claiming that by using statistical processes like Logistic Regression classifiers[1] in conjunction with these programs, they can achieve a high degree of accuracy (Losada and Crestani, 2016). Moreover, many studies that involve creating ERDS with a single data source validate Losada

1 A statistical and data mining technique that is used to analyze and classify binary and proportional response data sets (Maalouf, 2011, p. 1).

and Crestani's claim as most have precision[2], recall[3] and F1[4] values between 0.6-0.8—higher than the 0.5-0.8 average for most ERDS (Ahmed et al., 2018; Paul et al., 2018). Therefore, ERDS that examine one data source can still effectively detect signs of depression by employing statistical processes.

Most ERDS involve semantic analyses of text data from social media platforms like Twitter (Tsugawa et al., 2015). De Choudhury et al. (2013, p. 2) argue that social media is a logical source of data to examine because "people [often] post about their depression and even their treatment on [it]". Furthermore, Guntuku et al. (2017, p. 46) elaborate, claiming that online forums like Reddit[5] are also valid data sources as they provide "public[ly] available text data related to mental health". These programs often measure the use of personal pronouns like "I" and sentiment to reach a conclusion (Vedula & Parthasarathy, 2018). Moreover, many inspect the user's metadata for features like retweets and social neighbourhood[6] (Losada & Crestani, 2016; Paul et al., 2018). Thus, social media is a valid source of data because it is publicly available and offers an array of features to analyze.

Most text-based ERDS also rely on corpora[7] to find trends in texts written by users with clinical depression. By comparing the users' texts against corpora, developers can recognize the similarity between the two documents and reach a prediction (Paul et al., 2018). Nikhita Vedula who has a Ph.D. in Computer Science from Ohio State University (OSU), and Srinivasan Parthasarathy, professor of computer science at OSU, argue that an alternative approach involves Word2Vec[8] programs. They claim that this technique is often used with Skip-gram[9] and binary classifiers[10] (Vedula and Parthasarathy, 2018). Furthermore, some studies have taken more unorthodox approaches that rely on classifiers rather than corpora or Word2Vec programs. For example, Ashir Ahmed, associate professor of information technology at Kyushu University, and his colleagues use k-Nearest Neighbour[11] to analyze text data. By analyzing standard linguistic dimensions, the algorithm was able to predict whether users had MDD 73% of the time (Ahmed et al., 2018). This is similar to Losada and Crestani's project which achieved an accuracy rate of 71%; however, the most precise program was from Vedula and Parthasarathy whose accuracy rate was 86% (Losada & Crestani, 2016; Vedula & Parthasarathy, 2018). Therefore, ERDS that rely on Word2Vec programs are particularly effective in recognizing signs of depression in text data.

Although most studies analyze linguistic data, some researchers have begun examining alternative sources of data. Andrew Reece who has a Ph.D. in Psychology from Harvard University, and Christopher Danforth, professor of mathematics at the University of Vermont, created an early risk detection system that analyzed the psychological indicators found in photographic data from Instagram. This included

2 A measure of accuracy that is calculated by dividing the number of true positives by the number of true positive and false positives (Losada & Crestani, 2016).

3 A measure of accuracy that is calculated by dividing the number of true positives by the number of true positive and false negatives (Losada & Crestani, 2016).

4 A measure of accuracy that is calculated by multiplying the precision and recall by two and then dividing the resulting number by the sum of the precision and recall. Precision, recall, and F1 are on a scale from 0 to 1. Numbers closer to 1 indicate that the program has a higher degree of accuracy as it has almost all true positives and true negatives, and nearly no false positives and false negatives (Losada & Crestani, 2016).

5 An online discussion website where members contribute content such as text and images (Reddit Inc., 2020).

6 Refers to the number of followers and followees (Vedula & Parthasarathy, 2018).

7 A language resource that consists of various bodies of texts that are related to a particular topic (Paul et al., 2018).

8 A program that uses a text corpus as an input to generate its word vectors as an output. It constructs a vocabulary from the training text data and then learns the vector representation of words. The resulting word vector file can be used in natural language processing and machine learning programs (Goldberg & Levy, 2014).

9 An unsupervised learning technique that infers the nearby contextual words (Inkpen et al., 2018, p.

10 Often used in algorithms to produce a prediction score, which reflect the program's certainty that the given observation belongs to the positive class (Sharp et al., 2018).

11 A non-parametric approach that is used to calculate the distances from the points of interest to the points in the training set (Ahmed et al., 2018, p. 6).

the presence of other people through facial recognition technology and the setting (Danforth, 2017). Thus, although text-based ERDS are the most common, studies have started to develop ERDS with non-text sources of data too.

While most ERDS rely solely on one data source, some researchers have begun using multiple sources in a single program (De Choudhury et al., 2013; Gutunku et al., 2017). H.T. Kung, professor of computer science at Harvard University, and his associates argue that this gives a more holistic image of the users' mental states, which allows better prediction of depression (Kung et al., 2016). Specifically, programs utilizing multiple sources have achieved accuracy rates between 84-88% (Kung et al., 2016; Vingerhoets et al., 2010). Therefore, although ERDS can identify signs of depression with only one data source, analyzing multiple sources of data increases the likelihood of the ERDS achieving higher degrees of accuracy.

2.4. Gap Analysis

Past literature has examined the symptoms and risk factors of depression in adolescents and adults (Boschloo et al., 2016; Thapar et al., 2012). Many researchers have also developed ERDS that use one or more sources and types of data to predict depression in users (Inkpen et al., 2018; Paul et al., 2018). This study most closely resembles that of Sho Tsugawa, assistant professor of engineering at the University of Tsukuba, and his colleagues. They created an early risk detection system—using machine learning algorithms and data from Twitter's application programming interface (API)12—to predict depression in users on a sliding scale based on the CESD Scale (Tsugawa et al., 2015). This study seeks to develop a similar system as it also analyzes text data from Twitter's API. Unlike past studies, however, the program is a part of an iOS application that alerts users of the likelihood and severity of their depression by analyzing their data from iMessage and Twitter. Moreover, this research differs from past studies because this program relies on the outputs of various coding files, which measure users' symptoms from different data sources, to reach a conclusion. Furthermore, no study has yet attempted to test the accuracy of such a program on Canadian adolescents. Therefore, by examining multiple sources and types of data, this research aims to create a unique early risk detection system with a higher precision, recall, and F1 than previous studies that have only used one data source.

3. Methodology

3.1. Study Design

This study uses quantitative correlational research to examine the extent to which a newly created algorithm can accurately detect signs of depression in users' data from Twitter and iMessage. By employing this type of research, this study establishes the relationship between symptoms of depression and their ability to predict it in users. John David Creswell, professor of psychology at Carnegie Mellon University, and John W. Creswell, Research Scientist at the University of Michigan, claim that this form of research uses "the correlational statistic to describe and measure the degree or association…between two or more variables" (Creswell and Creswell, 2018, p. 12). This lets researchers examine the "more complex relationships among variables [that are] found in techniques of structural equation modeling, hierarchical linear modeling, and logistic regression" (Creswell & Creswell, 2018, p. 12). Moreover, Oberiri Apuke, lecturer of communication at Taraba State University, adds that correlation coefficients can be placed on these relationships to describe them in more depth (Apuke, 2017). These coefficients "range from +1.00 to -1.00. Higher correlations (coefficients closer to

12 A computer interface that enables the app's developers to share an application's data and functionality to third-party developers (IBM Cloud Developers, 2020).

13 A programming language that allows coders to access Reddit's API (Python Software Foundation, 2021).

14 The number of users who wrote these posts and comments is unknown since the "subreddit.hot" feature on the Reddit API was only used to track the text in distinctive posts. These were the most upvoted posts on each subreddit at the time that they were collected.

15 This was calculated on Excel using the Z-score squared multiplied by the standard deviation and 1 minus the standard deviation. This was then divided by the confidence interval squared.

+1.00 or -1.00) indicate stronger relationships," with a positive correlation suggesting that the variables are directly related, whereas a negative correlation implies an inverse relationship (Apuke, 2017, p. 44). By quantitatively measuring each symptom's ability to predict depression and the relation that each model has with detecting depression, adjustments can be made to the weighting of each indicator and the model to produce a higher correlation coefficient and thus a more accurate algorithm.

This study employs a prediction design as it aims to predict an outcome—depression or the lack thereof—by analyzing symptoms of depression that act as the variables (Apuke, 2017). William Lammers, professor of psychology at the University of Central Arkansas, and Pietro Badia, professor of psychology at Bowling Green State University, explain the applications of correlational research. They argue that it is especially useful for measuring the accuracy of diagnostic tools—in this instance, the algorithm—as it provides insight on the different conditions associated with a disorder (Lammers and Badia, 2005). This is particularly beneficial in cases of difficulty in detection or diagnosis, which is the case for depression, and if the design "is successful and the correlation is strong, then both the speed and accuracy of identifying the disorder may be substantially increased" (Lammers & Badia, 2005, p. 7). Thus, correlational research suits this study because it measures the accuracy of the program and determines the strength of each indicator in predicting depression.

3.2. Data Sources

The first step was identifying the main symptoms of depression from past literature. These included a depressed mood, fatigue, insomnia, feelings of guilt/worthlessness, difficulty concentrating, suicidal ideation/behaviour, and the number of personal pronouns versus plural subject pronouns. Then, the program was tested and refined using Python[13] with data from Reddit. The program scanned for symptoms of depression by searching for related words such as "tired" and "depressed." The Reddit API was used to collect 2,741 posts and comments[14] from Reddit users on depression-related subreddits ("Depression," "Depression_help," "Depression Regimens," and "ForeverAlone") and four other randomly selected subreddits ("MovieSuggestions," "Technology," "Food," and "News") to test whether it can differentiate users discussing depression versus those who are not. Since the demographic information of Reddit users is not publicly available, this study seeks to examine a diverse sample of users by analyzing data from multiple subreddits and users. To ensure that the analysis had statistical significance, a sample size based on a 90% confidence level and standard deviation of 0.5 was calculated[15]. The calculated sample size was approximately 70; however, to ensure that the program was tested on a diverse data set, a larger sample of 2,741 was chosen. Participants were classified as depressed if they explicitly referenced a diagnosis of depression in the text collected, through phrases like "I was diagnosed with depression;" simply stating "I am depressed" did not qualify as a diagnosis. Further, if users referenced attempted suicide or phrases pertaining to suicide ideation, they were classified as depressed. There were four rounds of testing; after each round, the accuracy rate was calculated on Excel by finding the difference between the number of true positives (TP) and true negatives (TN) against false positives (FP) and false negatives (FN). The number of times each symptom was found was also recorded after each round. This information was used to make modifications including changing the weighting of symptoms—based on how frequently they were found—and expanding the number of phrases associated with each symptom—if they were not detected much. These adjustments were made at the end of each round to improve the accuracy rate—increase the number of TP and TN relative to FP and FN.

The second phase tested the algorithm (in the Swift[16], Python, and Jupyter[17] files) and the iOS app. Students aged 17-18 in grade 12 at Appleby College[18] were invited to participate via email. The demographic information of each participant was recorded to

16 A programming language that is compatible with all Apple platforms (Apple Inc., 2020).

17 A web-based interactive development environment for Jupyter code, data, and notebooks (Project Jupyter, 2021).

18 An independent school in Oakville, Ontario, Canada that teaches over 800 students in grades 7 to 12. Its student body consists of a roughly equal mix of genders and represents 51 cultures and nationalities (Appleby College, 2020).

ensure that the program was tested on a diverse party that included different genders and ages. To calculate the sample size, a 75% confidence level and standard deviation of 0.5 was used. A lower confidence level was used because the results would likely vary if the study was repeated since they are unique to users' iMessage and Twitter data. The calculations suggested a sample of 4, however, a larger sample of 6 was used to examine even more data.

Participants received two online surveys—the PHQ and Beck's Depression Inventory (BDI)[19]—and were asked to fill them out simultaneously. Both questionnaires used a Likert-type 4-point scale to respond to a series of statements like "little interest or pleasure in doing things…[and] feeling down, depressed, or hopeless" ("Patient Health Questionnaire," 2005). If the respondent did not feel comfortable answering a question, they were given the option to pass—at which point, the totals for each category were reduced accordingly. Once they submitted the surveys, they were sent the coding files to run on their MacBooks. These scanned their iMessages—both the text in each message and the timestamps—and their Tweets—if the user gave access to it.

The program returned a score that predicted depression on a scale from 0-100, based on the PHQ and BDI. A score between 0-18 suggested no to minimal depression, 19-37 was mild, 38-56 was moderate, 57-75 was moderately severe, and 76-100 was severe. Users were then told to enter their scores into the app which returned the corresponding category they were placed in. They then submitted the score they received on a Microsoft Form. To measure the program's accuracy, the score they were given was compared against the results of the PHQ and BDI to determine whether they matched. To be considered a match, the user had to have been placed in the same category in both questionnaires and the app. The prediction from the program was not a formal diagnosis of depression, rather it was a method of alerting the user that they were showing possible signs of depression.

To protect participants' privacy, the programs were run directly on their devices. Additionally, the results of the questionnaires and the app were kept confidential unless the app found signs of suicidal thoughts on the device, the user indicated suicidal behaviour on the questionnaires, or the user received a severe or moderately severe diagnosis from the app or questionnaires. If this was the case, a member of the school's guidance department was alerted and put in contact with the student. However, at no point was the data given to the guidance member. Further, all participants were given resources for help regardless of their score. By implementing these measures, this research was approved by Appleby College's Internal Ethics Review Board.

3.3. Hypotheses

The following are hypothesized:

1. The recall, precision, and F1 of the final program will be within 0.02 of 0.75. This is based on the results that similar studies like Inkpen et al. (2018), achieved.

2. The more sources and data—Tweets and iMessages—that the program is given access to, the more accurate the prediction will be. This is based on the findings of similar studies like Kung et al. (2016), who found that more data sources let ERDS detect additional signs of depression or a lack thereof.

3.4. Analysis

To measure the relationship that each symptom in phase one had with predicting depression, a trendline modelling the correlation coefficient was created. This was used to determine the type and strength of each relation. The correlation coefficient was calculated using the Pearson product-moment on Excel. For phase two, the accuracy of the program was measured through the number of TP[20], TN[21], FP[22], FN[23], precision, recall, and F1. These statistics were calculated on Excel.

19 A self-report questionnaire that consists of 21 items that the user assigns a score of 0-3. It diagnoses depression on a sliding scale from 0-63 (Morales et al., 2017).

20 Participants who were correctly placed in a category that suggested a moderate degree of depression or higher.

21 Students who were correctly put into the no to minimal depression category.

22 Participants who were incorrectly assigned to a category that implied a moderate degree of depression or higher.

23 Participants who were incorrectly assigned to the no to minimal depression category.

4. Results

4.1. Phase One Results

The program was tested on 2,741 posts from 8 subreddits: "MovieSuggestions," "Technology," "Food," "News," "Depression," "Depression_help," "Depression Regimens," and "ForeverAlone." The breakdown of posts that were collected from each subreddit is illustrated in Figure 1.

The program achieved a higher accuracy rate on subreddits that were not related to depression: "MovieSuggestions," "Technology," "Food," and "News." However, the algorithm still achieved an accuracy rate of 81% on those related to depression, which is shown in Figure 2.

To illustrate the correlation between depression and each symptom, the number of references to the symptom was measured on the y-axis and the percentage score that predicted depression was put on the x-axis. The first relationship, in Figure 3, has a correlation coefficient of 0.655 which indicates a moderate positive correlation.

Figure 1

The Number of Posts/Comments from Each Subreddit

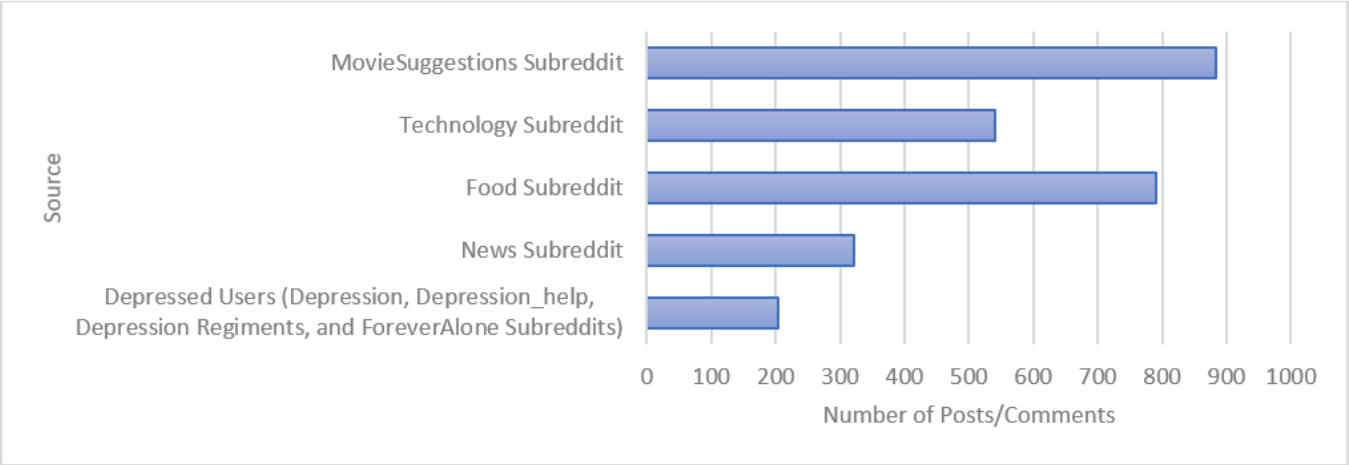


Figure 2

The Accuracy Rate for Each Source

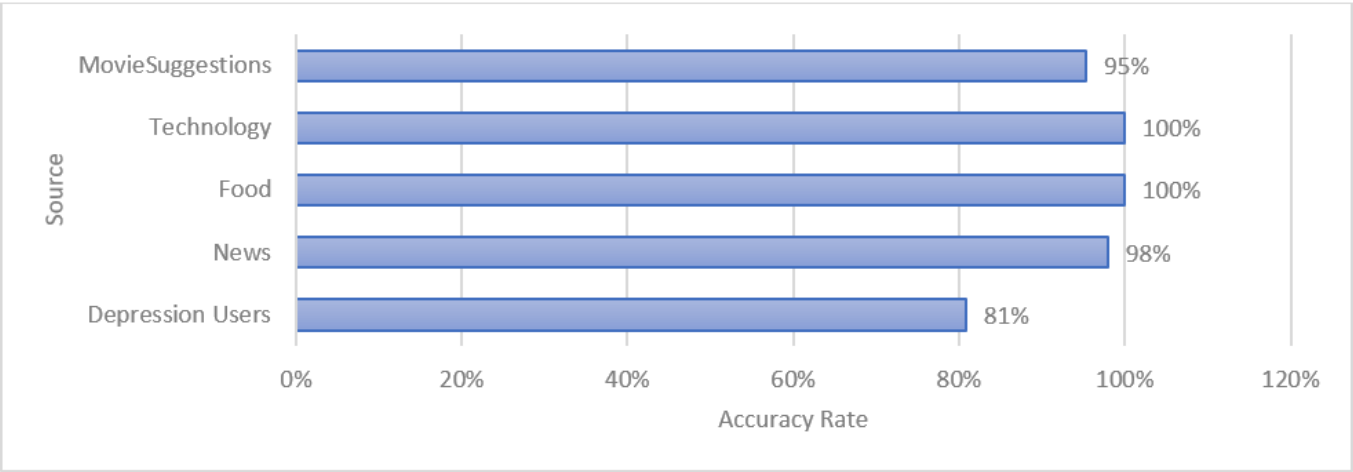


Figure 3

The Correlation Between Suicide and Depression

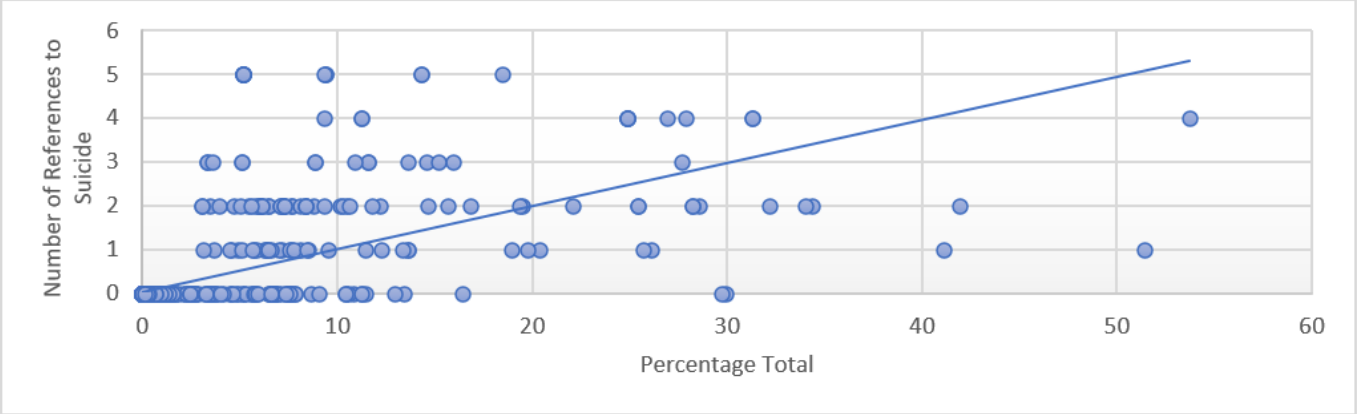
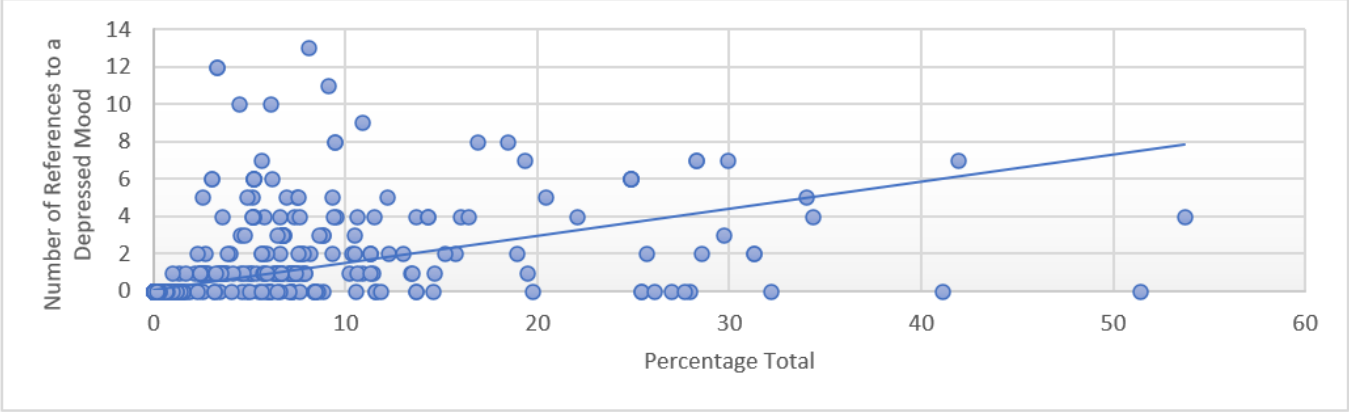


Figure 4

The Correlation Between a Depressed Mood and Depression



The trendline in Figure 4 showcasing the correlation coefficient of 0.5822 suggests that there is a moderate positive relationship between a depressed mood and depression.

The relationship between a loss of interest/pleasure and depression in Figure 5 has a correlation coefficient of 0.0337, which implies that there is no relation between them.

Figure 5

The Correlation Between a Loss of Interest/Pleasure and Depression

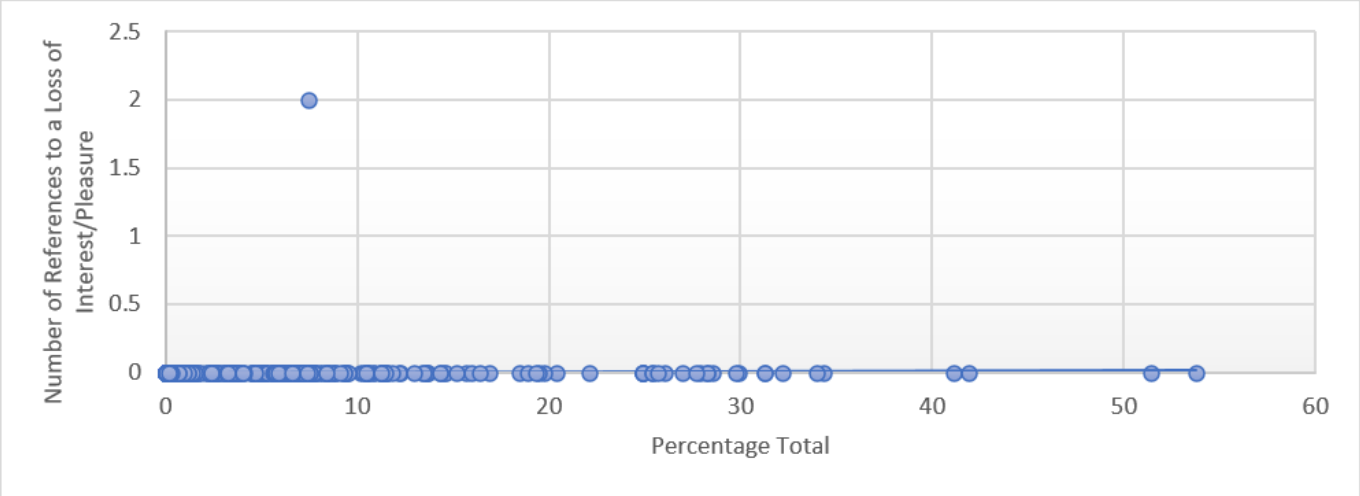
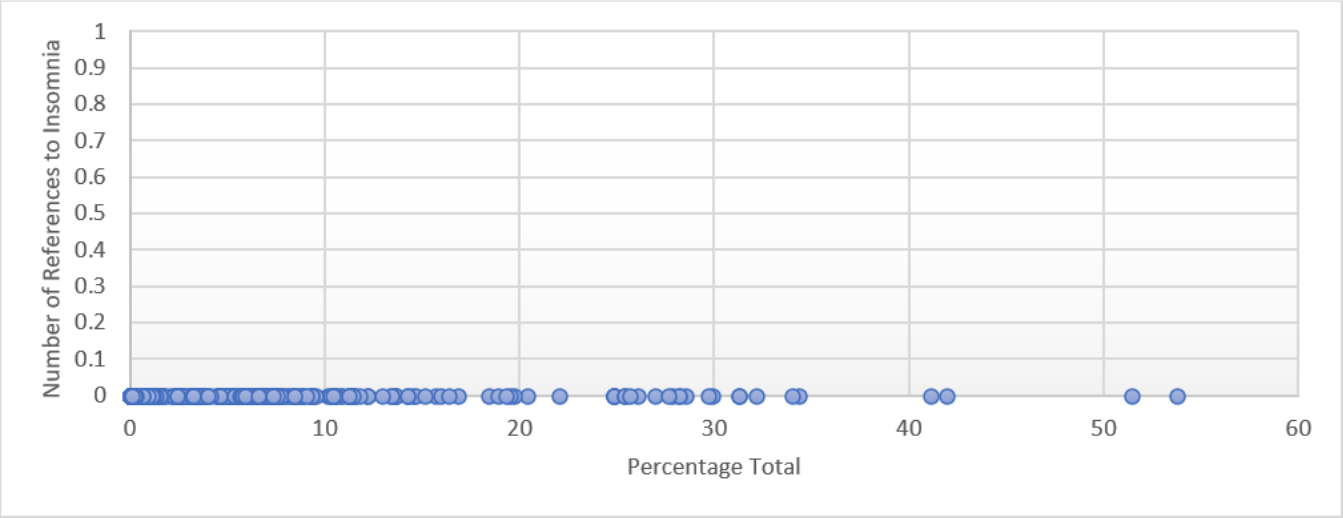


Figure 6

The Correlation Between Insomnia and Depression



In Figure 6, the correlation coefficient for insomnia and depression is 0, which indicates that there is no relationship between the two.

In Figure 7, the correlation coefficient for fatigue and depression is also 0, which indicates that there is no relationship.

The correlation coefficient for feelings of guilt/worthlessness and depression in Figure 8 is 0.1806, which suggests that there is a weak correlation.

The correlation coefficient for concentration problems versus depression in Figure 9 is 0, which indicates that there is no relation.

Figure 7

The Correlation Between Fatigue and Depression

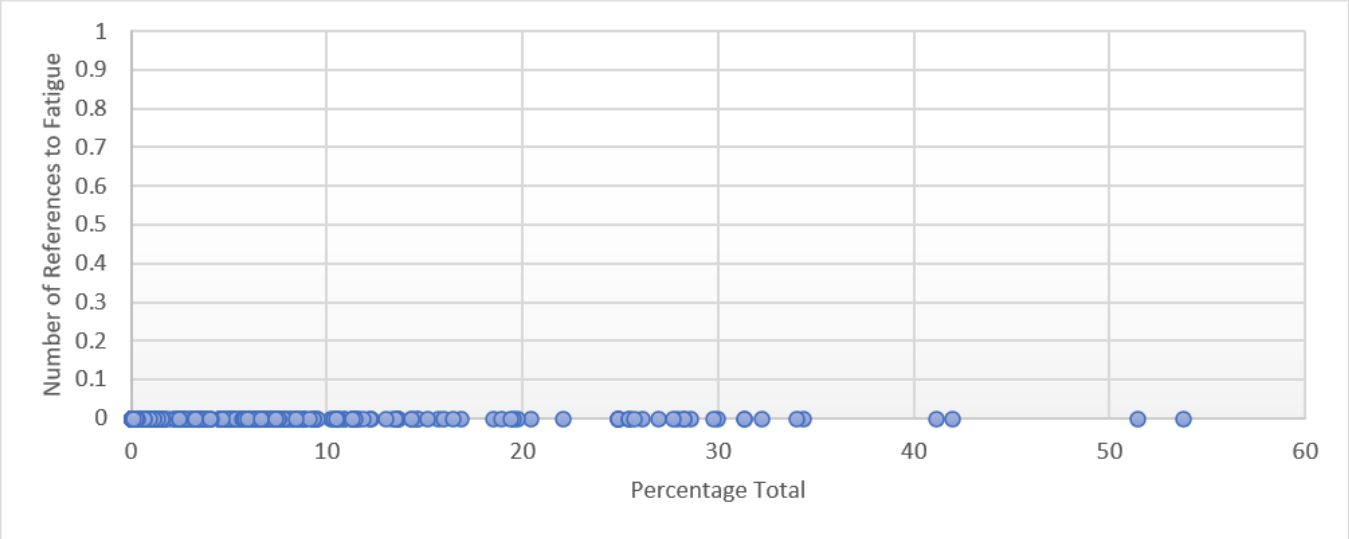


Figure 8

The Correlation Between Feelings of Guilt/Worthlessness and Depression

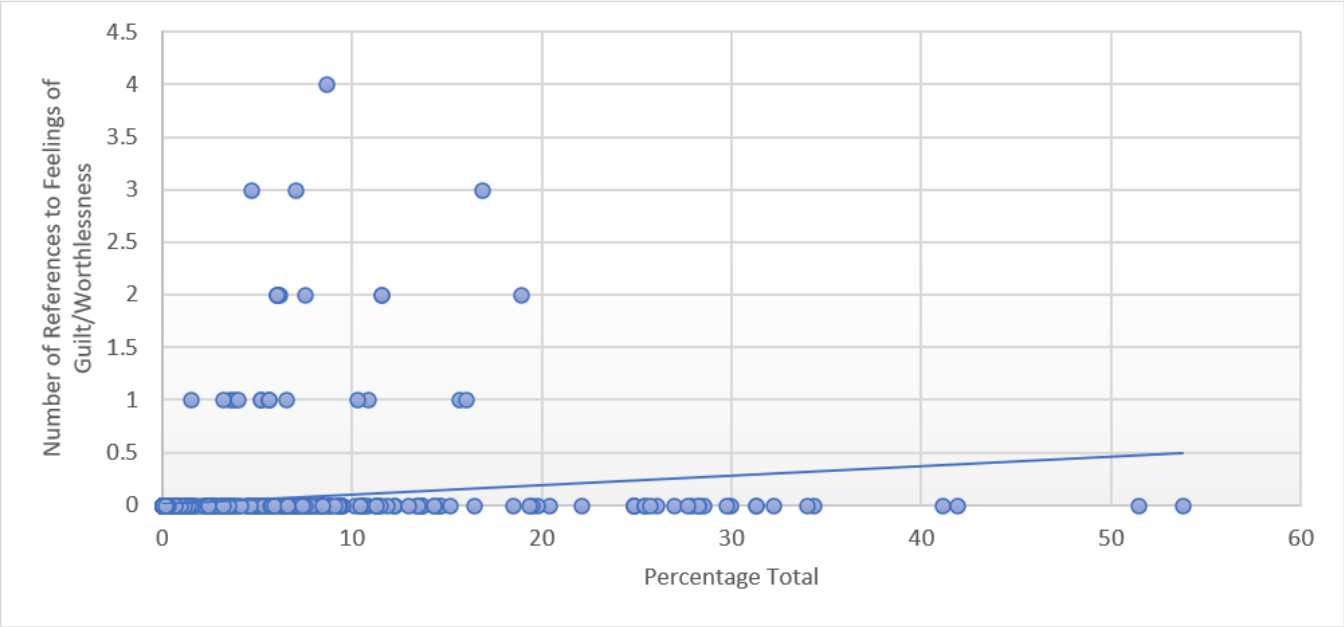


Figure 10

The Correlation Between Personal Pronouns and Depression

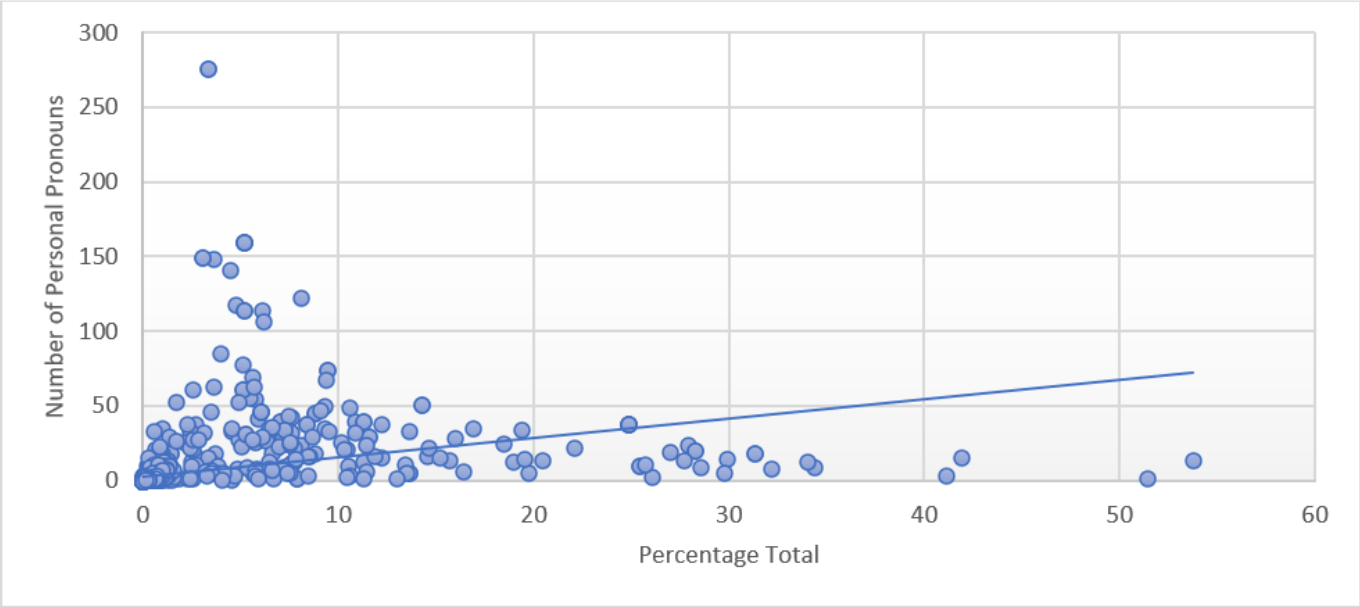
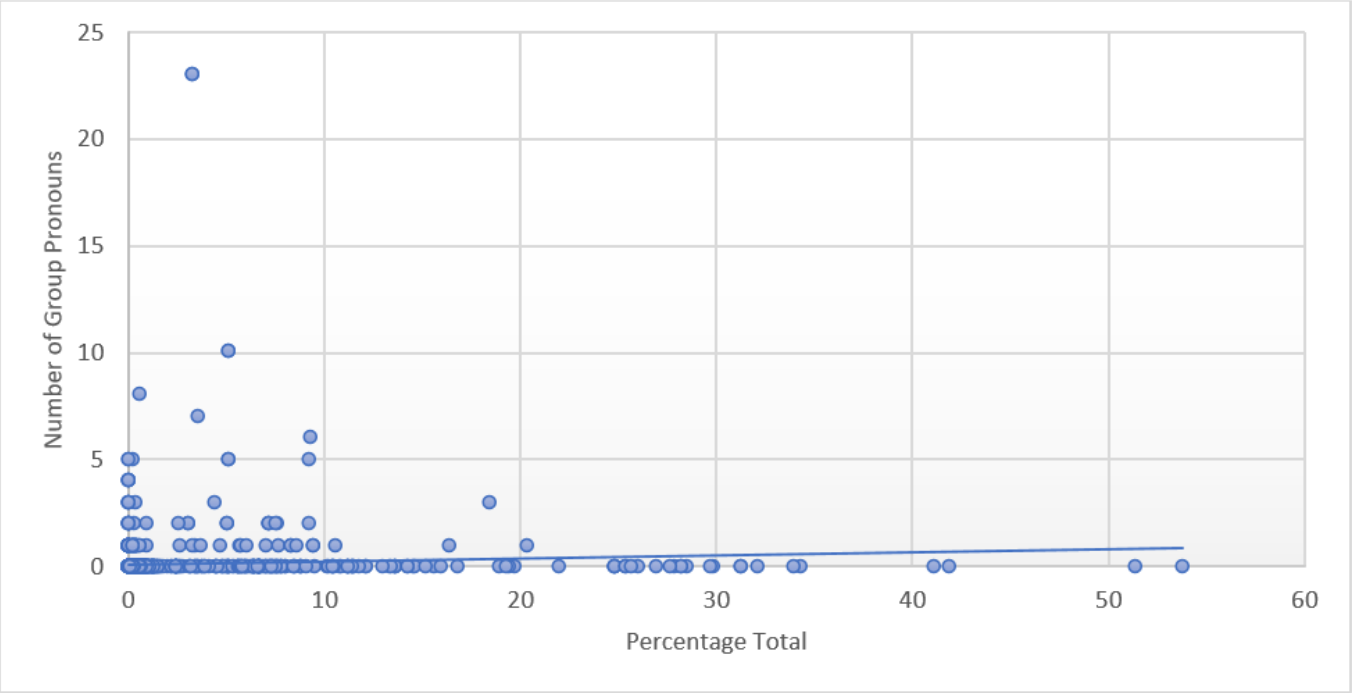


Figure 11

The Correlation Between Groups Pronouns and Depression



In Figure 10, personal pronouns and depression have a correlation coefficient of 0.3319, which indicates that there is a weak positive correlation.

Finally, in Figure 11, the correlation coefficient for group pronouns versus depression is 0.0642, which indicates that there is no relation.

4.2. Phase Two Results

There were 6 grade 12 students in this phase, which involved testing the program on users' iMessage and Twitter histories. There were 3 female and 3 male participants (3 aged 17 and 3 aged 18). The results are shown in Table 1. The precision reflects the difference between the TP and TP plus FP. Recall indicates the number of TP against TP plus FN. F1 shows the weighted average of precision and recall.

5. Discussion

5.1. Phase One

This research reveals new information about the factors that can be used in ERDS to predict depression. The program achieved an average accuracy rate of 94.80% across all subreddits. It was particularly successful in predicting the lack of symptoms on subreddits that did not explicitly relate to depression: "MovieSuggestions," "Technology," "Food," and "News" as it achieved rates between 95-100% (Figure 2). However, the program was still able to detect depression in users classified as depressed 81% of the time.

When examining the correlation that each factor has with predicting depression, the strongest correlations involved suicidal behaviour/ideation and a depressed mood. The correlation coefficient for suicidal behaviour/ideation versus depression was 0.64997 (Figure 3). According to Patrick Schober, Medical Specialist at the Universitair Medische Centra, and his colleagues, this is a moderate positive correlation (Schober et al., 2018). This means that as the number of references to suicide increases so does the percentage, which suggests that the user has more severe depression. Moreover, Swenda Moreh, Vice President of BridgePoint Healthcare, and Henry O'Lawrence, professor of healthcare at California State University, argue that suicide is a common sign of depression in adolescents. They claim that although "the rates of completed suicide are low…[at] 8.2 per 100,000 among 15- to 19-year-olds, there are many more attempted and suicidal ideations" (Moreh and O'Lawrence, 2016, p. 285). Many adolescents document these attempts on online platforms like Reddit that provide anonymity, which likely contributed to the high correlation between suicidal ideation and depression (Moreh & O'Lawrence, 2016). Thus, suicidal ideation/behaviour can be a significant factor in predicting depression on anonymous servers like Reddit.

The second factor with a moderate correlation was a depressed mood at 0.5822 (Figure 4). This coincides with past literature that found a depressed mood to be a statistically significant factor in detecting depression. Daantje Derks, professor of psychology, and her associates argue that it is "easier to express negative emotions [in computer-mediated communication (CMC)]…because one does not know the other…[and] one is less aware of the social effects of one's own expressions" (Derks et al., 2008, p. 4). Moreover, the "reduced visibility of emotions strengthens emotional style and content and makes it easier to express emotions, especially when individuals find it difficult to express them in real life," which applies to depression that often goes undiagnosed for these reasons (Derks et al., 2008, p. 4). This supports the conclusion of a depressed mood being a significant factor in predicting depression on CMC channels.

The two factors with low correlations are feelings of guilt/worthlessness and personal pronouns at 0.1806 and 0.3319, respectively (Figures 8 and 10). The low correlation for feelings of guilt/worthlessness was expected since previous studies, like that of Tsugawa et al., had similar findings, and it is not a factor that is used to clinically diagnose depression in instruments like the BDI (Tsugawa et al., 2015).

Findings that differed from past studies included the absence of correlation between depression and insomnia, fatigue, and a loss of interest/pleasure (Figures 5-7). The lack of correlation with loss of interest/pleasure is likely a consequence of the program's inability to identify activities that the user no longer enjoys as it only searched for generalizable terms like "disinterested." The lack of correlation with insomnia and fatigue is surprising given the findings of past research and new studies relating to "Zoom fatigue."24 For example, Manyu Jiang, recipient of the Facebook Journalism Project Scholarship, notes that CMC involving video is harder than face-to-face communication as it "requires more focus." Furthermore, although people's minds are together, their bodies are not; "that dissonance, which causes people to have conflicting feelings, is exhausting" (Jiang, 2020). Since Reddit users do not see the people that they are con-

versing with and can contribute whenever they want to, the fatigue that is often felt on other CMC channels or in-person does not apply. Therefore, the findings suggest that the most significant factors in predicting depression are suicidal behaviour/ideation and a depressed mood whereas the least significant are insomnia, fatigue, and a loss of interest/pleasure.

5.2. Phase Two

The results indicate that there is a positive correlation between the factors used in the programs and a users' degree of depression. This is likely attributed to many of the symptoms being used in clinical assessments like the PHQ. Further, the sliding scale that was used to predict depression was based on clinical assessments. Additionally, this implies that users' messages and Tweets contain data that does reflect their mental state which coincides with findings from past studies like that of Kung et al. (2016). Moreover, Prakash Date, professor of engineering at Cummins College, and his associates agree with these conclusions. They argue that social networking sites (SNS) display an "individual's [way of] thinking, mood, activities, and socialization." Further, they claim that it can show whether someone has depression since users often share "their thoughts, feelings, [and] emotions… [related to] guilt, worthlessness, [and] helplessness" (Date et al., 2018, p. 6016). Thus, the data from phase two shows similar trends to phase one and demonstrates the real-world applicability of these findings.

5.3. Limitations

Since there remains a stigma around discussing mental health illnesses, participants in phase two may have altered their responses on the questionnaires to avoid a diagnosis. This may have increased the number of FP as the programs may have detected signs of depression in cases where questionnaires would have concluded that they were not depressed. Roberta Heale, professor of nursing at Laurentian University, and Alison Twycross, professor of healthcare at London South Bank University, propose a solution of using two surveys that are similar to one another to achieve parallel-form reliability. If the difference between the final scores of the surveys was greater than ten, it suggested that the data was unreliable, and thus, the participant was excluded from the results (Heale & Twycross, 2015).

Another limitation involved the semantic analyses of users' texts. The algorithm could only detect symptoms if the personal pronouns that the participants used, and the description of the symptoms were on the same line in the text file. For example, if participants displayed signs of self-hatred by using negative adjectives to describe themselves, the algorithm only recognized that as a symptom if the pronouns were on the same line. Moreover, if users expressed their symptoms in a manner that was not explicitly coded for, which included using a short form, slang, or incorrectly spelled words, the program was unable to detect them. Both limitations resulted in the program returning a higher quantity of FN as it was unable to recognize the symptoms in the way that the user presented them. Finally, since there were many requirements for participants in phase two and the program was still in beta testing, there were only 6 volunteers. The limited sample size affected the recall, precision, and F1, leading to either extremely high or low values since a maximum of 6 numbers were used to calculate each. Therefore, to better generalize these results and the program's accuracy, a larger sample is needed.

5.4. Implications

This research indicates that ERDS that analyze data from CMC channels—without videos—should examine suicidal behaviour/ideation because it has the strongest correlation with predicting depression. However, since Reddit does not publish data regarding users' ages, no definite conclusions can be drawn that apply to specific age groups or that relate to non-text-based CMC channels. Furthermore, these findings suggest that the key symptoms of depression—those with the highest correlation coefficients—may differ based on the type of CMC. Thus, future research should measure the variation in the correlation between symptoms on different CMCs. Furthermore, this research indicates that as many SNS as possible

24 Anxiety, tiredness, or worry that results from overusing virtual videoconferencing platforms (Wiederhold, 2020).

should be integrated into ERDS as data sources because they reflect a user's mental state. However, to generalize these findings, further research should examine the accuracy of similar ERDS on different ages and in regions beyond Canada.

6. Conclusion

This program was tested on data from Reddit in phase one and grade 12 students in phase two. The study's hypotheses were partially correct. The recall of the final program was above the expected 0.75 at 1.0 since the program had no false negatives. However, due to the single false positive, the precision and F1 were lower than expected at 0.5 and 0.67, respectively (Table 1). Finally, having more sources of data to examine did result in more accurate predictions as the participants who shared their Twitter and iMessage data received TN or TP and their scores were closer to their PHQ and BDI counterparts than those who only gave access to iMessage.

The findings of this research are important to this field because they present new trends that differ from past literature, specifically regarding the correlation that each symptom has with predicting depression on CMC channels. Additionally, they suggest that future research is needed to further explore the correlation of these factors on different CMC mediums. Moreover, it is unclear whether the trends identified in phase two can be extrapolated due to the small sample. Thus, additional research should test the accuracy of ERDS on different ages from various regions.

Author Note: Please email taylorhodan@gmail.com to see the original code used in the study.

References

Ahmed, A., Islam, Md. R., Kabir, M. A., Kamal, A. R. M., Wang, H., & Ulhaq, A. (2018). Depression Detection from Social Network Data using Machine Learning Techniques. *Health Information Science and Systems*, 6(1). https://doi.org/10.1007/s13755-018-0046-0

Apple Inc. (2020). *About Swift*. Swift.Org. https://swift.org

Appleby College. (2020). Appleby College. https://www.appleby.on.ca/

Apuke, O. (2017). Quantitative Research Methods: A Synopsis Approach. *Arabian Journal of Business and Management Review* (Kuwait Chapter)., 6, 40–47. https://doi.org/10.12816/ 0040336

Boschloo, L., van Borkulo, C. D., Borsboom, D., & Schoevers, R. A. (2016). A Prospective Study on How Symptoms in a Network Predict the Onset of Depression. Psychotherapy and Psychosomatics, 85(3), 183–184. JSTOR. https://doi.org/10.2307/48540676

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (Fifth Edition).* SAGE Publications Inc.

Date, P., Sonawane, N., Padmane, M., Suralkar, V., & Wable, S. (2018). *Predicting Depression Level Using Social Media Posts.* 7(5), 4.

De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). *Predicting Depression via Social Media.* 10.

Derks, D., Fischer, A. H., & Bos, A. E. R. (2008). The role of emotion in computer-mediated communication: A review. Computers in Human Behavior, 24(3), 766–785. https://doi.org/10.1016/j.chb.2007.04.004

Gelenberg, A. J., Freeman, M. P., Markowitz, J. C., Rosenbaum, J. F., Thase, M. E., Trivedi, M. H., Rhoads, R. S. V., Reus, V. I., DePaulo, J. R., Fawcett, J. A., Schneck, C. D., & Silbersweig, D. A. (2010). *Treatment of Patients with Major Depressive Disorder.* 152.

Goldberg, Y., & Levy, O. (2014). Word2Vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method. *ArXiv:1402.3722 [Cs, Stat]*. http://arxiv.org/abs/1402.3722

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

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. Evidence Based Nursing, 18(3), 66–67. https://doi.org/10.1136/eb-2015-102129

IBM Cloud Developers. (2020, August 31). *What is an Application Programming Interface (API)?* https://www.

ibm.com/cloud/learn/api

Inkpen, D., Buddhitha, P., Husseini Orabi, M., & Husseini Orabi, A. (2018). Deep Learning for Depression Detection of Twitter Users. *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic,* 88–97. https://doi.org/10.18653 /v1/W18-0609

Jiang, M. (2020). The Reason Zoom Calls Drain Your Energy. https://www.counsellingresources. co.nz/uploads/3/9/8/5/3985535/the_reason_zoom_calls_drain_your_energy.pdf

Kung, H.-T., Williamson, J. R., Godoy, E., Cha, M., Schwarzentruber, A., Khorrami, P., Gwon, Y., Dagli, C., & Quatieri, T. F. (2016). Detecting Depression using Vocal, Facial and Semantic Communication Cues. Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge - AVEC '16, 11–18. https://doi.org/10.1145/2988257. 2988263

Lammers, W., & Badia, P. (2005). *Fundamentals of Behavioural Research.* https://uca.edu/psychology/files/2013/08/Ch15-Nonexperimental-Designs_Correlational_Ex-Post-Facto_Naturalistic-Observation_Qualitative.pdf

Lee, Y., Ragguett, R.-M., Mansur, R. B., Boutilier, J. J., Rosenblat, J. D., Trevizol, A., Brietzke, E., Lin, K., Pan, Z., Subramaniapillai, M., Chan, T. C. Y., Fus, D., Park, C., Musial, N., Zuckerman, H., Chen, V. C.-H., Ho, R., Rong, C., & McIntyre, R. S. (2018). Applications of Machine Learning Algorithms to Predict Therapeutic Outcomes in Depression. *Journal of Affective Disorders,* 241, 519–532. https://doi.org/10.1016/j.jad.2018.08.073

Lei, L., Huang, X., Zhang, S., Yang, J., Yang, L., & Xu, M. (2020). Comparison of Prevalence and Associated Factors of Anxiety and Depression Among People Affected by versus People Unaffected by Quarantine During the COVID-19 Epidemic in Southwestern China. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research,* 26, e924609-1-e924609-12. https://doi.org/10.12659/MSM.924609

Levenberg, A (2020) GitHub source code (Version 2.0) [Source code]. https://github.com/yortos/ imessage-analysis/blob/master/imessages-data-extract-and-prep.ipynb

Losada, D. E., & Crestani, F. (2016). A Test Collection for Research on Depression and Language Use. *Experimental IR Meets Multilinguality, Multimodality, and Interaction,* 28–39. https://doi.org/10.1007/978-3-319-44564-9_3

Maalouf, M. (2011). Logistic regression in data analysis: An overview. International Journal of Data Analysis Techniques and Strategies, 3, 281–299. https://doi.org/10.1504/IJDATS.2011.041335

Morales, M., Scherer, S., & Levitan, R. (2017). A Cross-modal Review of Indicators for Depression Detection Systems. Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality, 1–12. https://doi.org/10.18653/v1/W17-3101

Moreh, S., & O'Lawrence, H. (2016). Common Risk Factors Associated with Adolescent and Young Adult Depression. Journal of Health and Human Services Administration, 39(2), 283–310.

Patient Health Questionnaire (PHQ-9). (2005). 2. https://www2.gov.bc.ca/assets/gov/health/practiti oner-pro/bc-guidelines/depression_patient_health_questionnaire.pdf

Paul, S., Kalyani, J., & Basu, T. (2018, August). Early Detection of Signs of Anorexia and Depression Over Social Media using Effective Machine Learning Frameworks. https:// www.researchgate .net/publication/326998357_Early_Detection_of_Signs_of_Anorexia_and_Depression_Over_Social_Media_using_Effective_Machine_Learning_Frameworks

Project Jupyter. (2021, May 7). Overview of Jupyter. https://www.jupyter.org

Python Software Foundation. (2021). Welcome to Python. Python.Org. Retrieved January 21, 2021, from https://www.python.org/about/

Reddit Inc. (2020). *About Reddit.* https://www.redditinc.com/

Reece, A. G., & Danforth, C. M. (2017). Instagram Photos Reveal Predictive Markers of Depression. EPJ Data Science, 6(1), 1–12. https://doi.org/10.1140/epjds/s13688-017-0110-z

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesthesia & Analgesia, 126(5), 1763–1768. https://doi.org/10.1213/ANE. 0000000000002864

Sharp, T. A., Thomas, S. L., Cubuk, E. D., Schoenholz, S. S., Srolovitz, D. J., & Liu, A. J. (2018). Machine Learning Determination of Atomic Dynamics at Grain Boundaries. *Proceedings of the National Academy of Sciences of the United States of America,* 115(43), 10943–10947.

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. Lancet, 379(9820), 1056–1067. https://doi.org/10.1016/S0140-6736(11)60871-4

Thombs, B. D., Benedetti, A., Kloda, L. A., Levis, B., Nicolau, I., Cuijpers, P., Gilbody, S., Ioannidis, J. P. A., McMillan, D., Patten, S. B., Shrier, I., Steele, R. J., & Ziegelstein, R. C. (2014). The Diagnostic Accuracy of the Patient Health Questionnaire: A Systematic Review. *Systematic Reviews,* 3, 124. https://doi.org/10.1186/2046-4053-3-124

Tsugawa, S., Kikuchi, Y., Kishino, F., Nakajima, K., Itoh,

Y., & Ohsaki, H. (2015). Recognizing Depression from Twitter Activity. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15, 3187–3196. https://doi.org/10.1145/2702123.2702280

Vedula, N., & Parthasarathy, S. (2017). Emotional and Linguistic Cues of Depression from Social Media. *Proceedings of the 2017 International Conference on Digital Health* - DH '17, 127–136. https://doi.org/10.1145/3079452.3079465

Vingerhoets, A., Balsters, M., Krahmer, E., & Swerts, M. (2010, January). *Measuring Potential Cues for Depression in Adolescents.* ResearchGate. https://doi.org/10.1145/1931344.1931386

Wiederhold, B. K. (2020). Connecting Through Technology During the Coronavirus Disease 2019 Pandemic: Avoiding "Zoom Fatigue." *Cyberpsychology, Behavior, and Social Networking,* 23(7), 437–438. https://doi.org/10.1089/cyber.2020.29188.bkw

Williams, S. Z., Chung, G. S., & Muennig, P. A. (2017). Undiagnosed Depression: A Community Diagnosis. SSM - *Population Health,* 3, 633–638. https://doi.org/10.1016/j.ssmph.2017.07.012

World Health Organization. (2020, January 30). *Depression.* https://www.who.int/news-room/fact-sheets/detail/depression