Recognizing Suicidal Intent in Depressed Population using NLP: A Pilot Study

Samiha Binte Hassan[1], Sumaiya Binte Hassan[1], Umme Zakia[2], Student Member, IEEE

Abstract—**Depression is a prevalent form of mental disorder that can affect productivity in daily activities and might lead to suicidal thoughts or attempts. Conventional diagnostic techniques performed by mental health professionals can help identify the level of depression present in a person. To facilitate such a diagnostic approach, in this paper, we present an automated conversational platform that was used as a preliminary method of identifying depression associated risks. The platform was developed to understand conversations using Natural Language Processing (NLP) via machine learning technique. In the proposed two-phased platform, the initial *intent recognition phase* would analyze conversation and identify associated sentiments into four categories of 'happy', 'neutral', 'depressive' and 'suicidal' states. In the final *emotion nurturing phase*, the platform continued with supportive conversations for the first three states while triggering a local call to a *suicide prevention helpline* for 'suicidal' state as a preventive measure. This multi-layer platform integrated Google Home mini, Google Dialogflow Machine Learning (ML) algorithm and Twilio API. Dialogflow ML obtained classification accuracy of 76% in recognizing user's mental state via NLP and was found efficient over the classic SVM classifier. As a pilot study, current focus of this paper was solely based on the usage of words and intent of the user and was found effective.**

Keywords—speech recognition, user intents, machine learning, depression, suicidal tendencies, natural language processing, Google Dialogflow

I. INTRODUCTION

Clinical depression or major depressive disorder classified by the Diagnostic and Statistical Manual of mental disorders (DSM), is a severe psychiatric illness that impacts an individual's cognitive and neuropsychological behaviors [1]. The most prevalent symptoms of depression consist of persistent negative emotions, pessimism, low energy, and suicidal thoughts [2]. Often this psychological mood disorder stems from an individual's difficulty in coping with stressful life events, persistent negative thoughts and emotional trauma [3]. It has been estimated by the World Health Organization (WHO) that 264 million people worldwide are affected by this psychological illness, making depression the fourth most leading cause of disability worldwide [4]. It is expected to become the second most dominant cause of disability in the world by 2030 [3]. The most severe symptoms of depression are suicidal thoughts and tendencies. According to WHO, each year approximately 800,000 cases of death is due to suicide itself [5]. It is estimated that suicide is the third most prevalent cause of death in 15-19-year age group, making younger generation more susceptible to depression and suicide [6]. Suicidal behavior is often combined with certain psychological disorders and is fatally evident in people with prior diagnosis of clinical depression. According to [2], 50% of individuals with clinically depressive symptoms commit suicide. Thus, there is a strong correlation between depressive traits and suicidal intentions.

Early detection of depression and suicidal thoughts can be a preventative measure for suicidal attempts [7]. Study shows that between 50-70% of people dealing with depressive traits and suicidal thoughts contacted their health care provider but only 50% of these cases were accurately detected and provided with an effective treatment plan [3][8]. This suggests that the proper diagnosis and immediate treatment plan for those individuals exhibiting severe depressive traits could prevent up to 70% casualties of suicide. Thereby, to prevent suicidal tendencies, a two-process system is required: (1) accurate diagnosis and (2) immediate treatment plan for the individual. Although there are many studies focusing on finding a diagnostic tool, not all are aware of providing immediate treatment after diagnosis. Hence, it is significant to emphasize on proper diagnostic tools for depression. Both early detection and effective aftercare are necessary to reduce the mortalities risks associated with suicide and depression.

This study focuses on developing an effective diagnostic tool for depression in people who demonstrate the need for clinical treatment. An interactive platform is proposed to carry out conversations to recognize users' sentiments or intents using natural language processing (NLP). The two-phased platform was designed to recognize suicidal intents from oral speech of a person by integrating machine learning approach with Google Home mini, a virtual personal assistant (VPA). It is imperative to identify words associated with strong negative emotions to correlate severe depression and suicidal tendencies. During the first stage i.e., *intent recognition phase*, Google Dialogflow machine learning algorithm was used for training a model to recognize user's mental or emotional states from his/her oral speech and/or conversations with the VPA. The model was trained with a dataset that had several sentimental statements along with actual suicidal notes categorized into four emotional states such as '*Happy*', '*Neutral*', '*Depressive*', and '*Suicidal*' states. In the second stage or the *emotion nurture phase,* a distress signal was triggered to a *suicide prevention helpline* in case severe depression or '*Suicidal*' intent was recognized; otherwise, the platform continued with supportive conversations. Therefore, this detection system focuses both on (1) the diagnosis and (2) providing immediate resources and treatment plan following the diagnosis. The use of such a novel tool can be used to initiate proper resource and treatment for depression in conjunction with traditional approach.

The rest of the paper is organized as follow: Section II describes related works present in literature while proposed model is described in Section III. Implementation, results with limitations and future works of the proposed system are discussed in Section IV and Section V, respectively. The paper

[1]Authors are with University of British Columbia, BC, Canada,
Email: {Samiha.bintehassan, sumaiya.bintehassan}@alumnai.ubc.ca.
[2]Author is with Simon Fraser University, BC, Canada,
Email: uzakia@sfu.ca.
Corresponding author: Samiha Binte Hassan,
Email: Samiha.bintehassan@alumnai.ubc.ca, Mobile: +1 (604)-7271161.

978-1-7281-8416-6/20/$31.00 ©2020 IEEE

concludes in Section VI.

II. Related Work

The diagnostic process of depressive and suicidal tendencies often follows a traditional approach in a controlled interview setting to elicit verbal behavior and non-verbal cues [6]. Analysis of verbal speech content is a dominating indicator of an individual's mental health status. If implemented correctly, this diagnostic tool can accurately reveal depressive traits; however, it has its limitations [9]. The main drawback of this tool is that the diagnostic process is predominated by subjective inferences of the interviewer; introducing biases in the diagnosis which can lead to either false negative or false positive diagnostic results [9]. Although false positive is generally harmless, however, false negatives can have fatal consequences as this type of diagnosis can fail to predict and prevent suicidal tendencies, therefore leading to mortalities. Another constraint of this traditional approach is that tremendous amount of resources is required to train health care individuals to utilize the tool for proper diagnosis and treatment. This can often be time consuming and cost ineffective; and most of the time this diagnostic tool lacks accessibility for the general population. This demands the notion of an objective diagnostic tool that provides higher level of accuracy while eliminating any form of subjective biases. Therefore, the use of different biomarkers such as voice recognition, specifically analysis of natural language processing in suicidal populations can be an effective method for objective inferences [6][10]-[14].

Recent research on the screening process of depression and suicidal intention focuses on finding distinctive biomarkers which do not require a clinical setting for screening. Tracking different biomarkers such as geospatial activity (using GPS and Wi-Fi), kinesthetic activity (using multi-axial accelerometers), sleep duration and verbal speech content analysis using speech detection algorithms provide unobtrusive monitoring [10]-[14]. These can lead to effective screening of depression remotely. Studies focusing on semantic content reveal that lexical word choices of an individual can provide information about their current emotional state, therefore providing pure and unbiased inference of their mental condition [15][16]. Study indicates that vocabulary of people dealing with severe depression often contain certain lexical word categories [16]. Recent adoption of automated processes such as smartphones coupled with virtual assistance opens a new door of possibility for screening various mental illnesses [16][17]. Research have shown that verbal and non-verbal cues such as facial expression, semantic content and syntax are effective indicators of major depression and suicidal intention [3][15][17][18][19]. Unfortunately, the implementation of these diagnostic tools in the general population remains obscure and to the best of the authors knowledge none of the research work emphasizes on the aftercare following the diagnosis.

To understand and analyze the natural dialogue, these diagnostic tools require Artificial Intelligence (AI). To establish communication between human and machine, the interactive conversational systems, or Virtual Personal Assistants (VPAs) based on AI are getting popular. A variety of applications where VPAs can be deployed are in home automation, customer service, education or medical assistance, automated vehicles, security control and many more. Google's Home mini, Amazon's Alexa, Apple's Siri, or Microsoft's Cortana are some well-known VPAs [20]. For these applications, understanding client's perception or emotion is important to provide appropriate feedback. Thus, sentiment analysis through natural language processing is the key features of the VPAs. Natural language process (NLP) is a subfield of cognitive systems that overlaps in linguistics, computer science, and artificial intelligence. This processing system specifically focuses on human-machine interactions using natural language. Recognizing emotion through text mining requires identifying polarity of the text and classifying it into positive, negative, or neutral states [21]. Research show that NLP can be used in classifying suicidal intention in a population from verbal cues in an effective manner [22]-[24], also effective in analyzing written language [25].

The purpose of this study was to develop an objective diagnostic tool for recognizing depression that would be affordable and easily available. It could categorize emotional states of a person and performed sentiment analysis to recognize depressive symptoms. The proposed novel platform utilized a popular VPA for interactive conversations and provided preventive and supportive measures in cases of extreme depression. NLP was implemented for processing the conversations between the person and the VPA, termed as '*VPA-DR (virtual personal assistant for depression recognition)*' in the rest of the paper. This cloud-based personal mental health assistance could be used to complement the traditional diagnosis process in recognizing user's current mental states at an early stage.

III. Proposed Suicidal Intent Recognition System

A. Modules Integrated

1) Dialogflow API

The Dialogflow ES is a Google's cloud-based platform that can recognize and process natural language from a conversational user interface. Few popular applications are such as to build interactive voice response system mobile/web application, or to create chat-bot by integrating multi-layer platforms/applications to Diaglogflow ES. It uses NLP to identify the *intent* of the user by speech recognition and sentiment analysis. The Dialogflow ES trains an *Agent* with its built-in Machine Learning (ML) algorithm to perform Natural Language Understanding (NLU). It utilizes two concepts for NLU operations: *Intent* and *Entity* [26].

2) Google Home Mini

Google Home Mini is a virtual personal assistive device (in our case, the VPA-DR) used for performing assistive tasks based on conversational language, initiated by a user. It comes with Quad-core 64-bit ARM CPU 1.4 GHz, Wi-Fi, Bluetooth support and high-performance ML hardware engine. It can be integrated to communicate with Google Dialogflow ES platform.

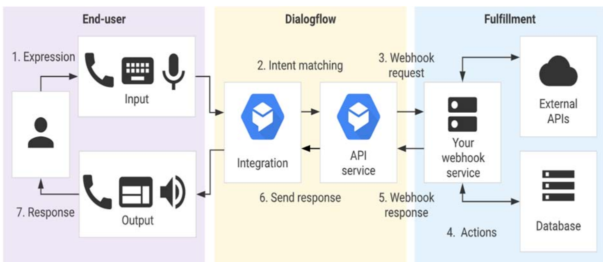


Fig. 1. Multi-layer platform [25].

3) Twilio API

The Twilio API lets Dialogflow *Agent* program to call or send texts to a Twilio supported number over Public Switch Telephone Network (PSTN) or session initiation protocol (SIP) layer. Twilio Media Streams enables integration between Twilio and Dialogflow by accessing audio in real-time from an ongoing phone call [27]. Twilio API either calls or sends messages on behalf of the user to phone numbers around the world via Twilio's REST API, as shown in Fig. 1.

B. Process Model: Suicidal Intent Recognition Platform

In the proposed multilayer platform, Dialogflow ES played vital role to serve an end user. It utilized cloud resources and acted as a linchpin to manage communication between the applications, as shown in Fig. 2. The hybrid Machine learning (ML) algorithm embedded in Dialogflow was trained with a preprocessed dataset to identify depressive and suicidal speech patterns. Dialogflow *Agents* were trained for NLU by analyzing conversations between the user and the VPA-DR via *intent* and *entity*. By default, the Dialogflow *Agent* responded to the user by identifying an *intent* behind a statement by matching against a static set of predefined response. On the other hand, *entity* was instrumental in extracting any important data from the user's speech [26].

Once the Dialogflow Agent received the input stream from the VPA-DR, it handled the request for *intent matching*, and responded with audio for the VPA-DR. For a more dynamic approach, Dialogflow was enabled for *fulfillment*, i.e. responding to a certain intent by sending a request to the webhook service via Twillo API with information about the matched intent. A unique String Identifier (*SID)* key to identify specific resources and an *Access Token* were generated by the Twillo API. Once a Twilio number was assigned with the specific token, it was used to send outgoing SMS and calls to a dedicated *suicide prevention helpline*.

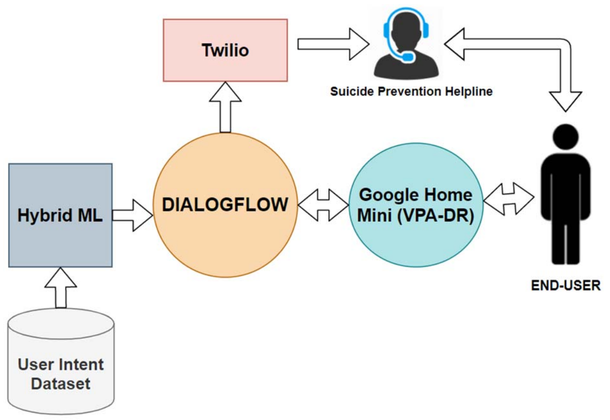


Fig. 2. Proposed suicidal intent recognition platform.

C. Process Phases

1) Intent Recognition Phase

In the proposed system, four categories of user intents were implemented for speech analysis using the machine learning model. These emotional states were: *'Happy', 'Neutral', 'Depressive',* and *'Suicidal'*. Dialogflow-ML returned a predicted value indicating the emotional/ sentimental state through the built-in natural language processing. Dialogflow *Agent* identified depression in a user by matching against the predefined intents when the ML threshold for '*Depressive*'/ '*Suicidal*' intent was detected.

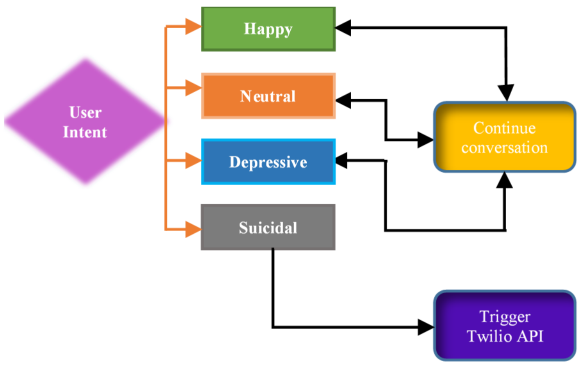


Fig. 3. Activating different phases based on user intent.

2) Emotion Nurture Phase

For the supportive role, VPA-DR, the Google Home mini would continue conversations for personal mental assistance in accordance with the emotional state being identified. If the machine learning model identified or mapped '*Suicidal*' state in a user, Dialogflow triggered a webhook service to Twilio API. A client with the SID and access token was created to connect phone numbers to call or send text messages. Twilio API would sent a text message to a local suicide prevention center, sending a distress signal that requested professional mental health assistance for the user. Fig. 3 shows the emotion states and the appropriate process phases that were designed to address associated issues.

IV. Implementation

A. Dataset for Suicidal Intent Recognition

The dataset used in this study was segregated into 4 different categories: 1. '*Happy*', 2. '*Neutral*', 3. '*Depressive*' and 4. '*Suicidal*'. Each categories consisted of sentences and statements that depicted certain emotions expressed by real people. The '*Happy*' and '*Neutral*' intents were curated from Kaggle dataset [28] available online. Statements for '*Neutral*' intent did not convey a strong emotion, but rather stated facts or asked questions to complete daily activities. The '*Depressive*' intents consisted of depressive speech tendencies and words that were observed in a population suffering from mild to medium level of depression. The fourth dataset consists of statements depicting '*Suicidal*' intent; extracted from the suicide notes of severely depressed population who have committed or attempted suicide [29]. For each category of user intent, 30 statements/ expressions were included in the suicidal intent dataset. Each category was extracted and compiled into a singular comma separated values (CSV) file which was then fed to the Dialogflow hybrid-ML model. The suicidal intent dataset was split into 75:25 ratio as training and test data for each category of user intent, as shown in Table I.

TABLE I. Dataset Used for Recognizing User Intent

User Intent	DATASET	
	Training data	***Test data***
HappyDB	25 entries	5 entries
NeutralDB	25 entries	5 entries
Depressive DB	25 entries	5 entries
SuicidalDB	30 entries	10 entries

These data were collected from a population consisted of both male and female with ages between 15-65 years and with a diverse race and culture. An HP laptop [Intel core i7 processor with 16 GB RAM] was used to implement the proposed system.

Fig. 4. Training dataset used for learning user intents of 4 categories.

B. Data Preprocessing

By default, Dialogflow performed stemming on words which had a common root by stripping each word of its derivational and inflectional suffixes for intent mapping [30], [31]. The automatic spelling correction enabled in Dialogflow could detect and correct spelling and grammar mistakes, so the user intent was mapped as correctly written statements as the test dataset. Also, words that were considered redundant/ unnecessary (such as “a”, “the” etc) were filtered out to identify the user intent.

C. Training DialogFlow Machine Learning

For training the model, *term reinforcement* processing was implemented to put more weight of certain words and phrases that ranked higher. Similarly, words and phrases that were not repeated as much during the training phrase were given less weight for mapping purpose. In order to utilize *term reinforcement*, each *intent* was trained with ≥ 25 training phrases and statements. There were four *intents* used in the proposed system: ‘*Happy*’, ‘*Neutral*’, ‘*Depressive*’ and ‘*Suicidal*’, with corresponding phrases/statements from the following datasets: *HappyDB, NeutraDBl*, *DepressiveDB* and *SuicidaDBl*. Few glimpses of the training dataset and test dataset are listed in Fig. 4, Table II and Fig. 6.

D. Triggreing States based on User Intent

To refine the response by the VPA-DR, Dialogflow Machine Learning (ML) classification threshold was tuned to a certain value. This filtered out false positive results while providing a variety of matched natural language inputs for the designed *Agent*. The Dialogflow ML scored each potential match with an intent detection confidence value, also known as the confidence score that ranges from 0 (completely uncertain) to 1 (completely certain) [32]. Once the ML scored an intent, there were three possible outcomes:

- If the highest scoring intent had a confidence score greater than or equal to the ML Classification Threshold setting, it was returned as a match.
- If no intents met the threshold, a fallback intent was matched.
- If no intents met the threshold and no fallback intent was defined, no intent was matched [26].

In this study, the machine learning (ML) classification threshold was set to 0.3. This confidence score indicated how closely words were matched. The dialogflow Agent had no *actions* or *fulfillments* when *Happy* and *Neutral intents* were identified. In contrast, if the Agent mapped a conversation to the *Depressive intent*, it would trigger the *fulfillment* by continuing the conversation with encouraging words. However, if at any point in the conversation *Suicidal intent* was recognized, then the Agent would *fulfill* that *intent* with a webhook service call (via Twilio) to a local *suicide prevention helpline* service center, as shown in Fig. 5.

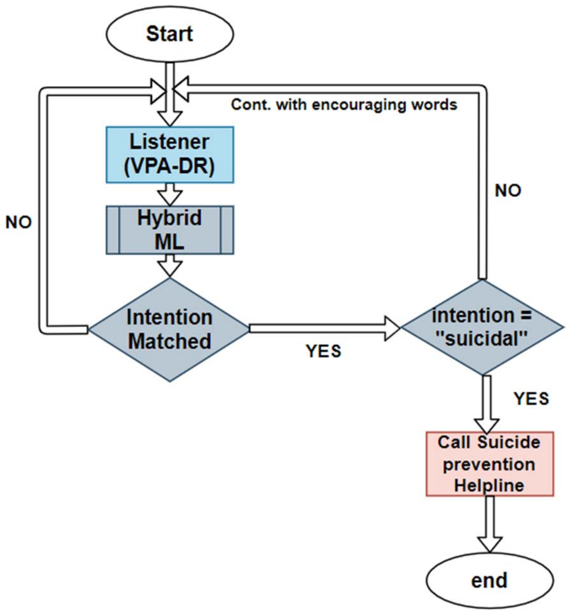


Fig. 5. Decision flow in recognizing user’s ‘Suicidal’ intent.

Fig. 6. Test dataset used for suicidal intent recognition.

E. SVM classifier in recognizing suicidal intent using NLP

To compare the performance of the built-in ML technique, we implemented an SVM classifier for natural language

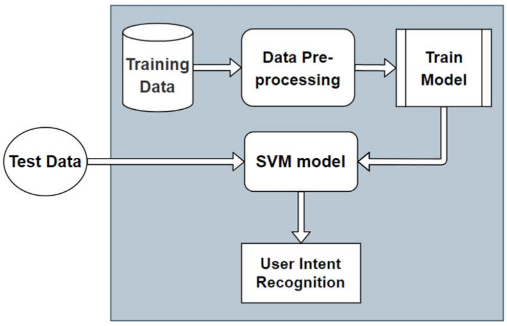


Fig. 7. Multi-class classification using SVM model.

processing in Matlab 2020 [MathWorks, MA, USA]. The error-correcting output codes (ECOC) function was implemented to train SVM models [One vs One] and returned a multiclass fully trained model using predictors and class labels in training data. Here, a simple classification model was created using word frequency counts as predictors, as shown in Fig. 7. The same training and test datasets as described in Section IV.A were used to train and evaluate the model. Before training, raw data was preprocessed for word analysis. This involved tokenization of the data for case conversion, removing list of stop words (such as "and", "of", and "the"), erasing punctuations and generating a cleaned normalized dataset. Text data was then converted to numeric data and was used to train the multiclass linear classification model using the *fitcecoc* with a linear learner that could support sparse data input. A bag of words was generated for predicting. This supervised SVM classifier used word frequency counts from the bag-of-words model and the labels for prediction as depicted in Fig. 8. Test data was also preprocessed similarly for evaluation purpose.

Fig. 8. Bag-of-words generated during NLP training with SVM classifier.

V. Results & Discussion

A. Performance matrice

For performance analysis, classification accuracy was calculated to evaluate the trained model in predicting user intents. The classification accuracy was the proportion of the intent labels that the model predicted correctly and was calculated based on the following equation:

$$Pred\_Sentiment\_Intent_{Acc} = \sum \frac{(YPred==YTest)}{num(YTest)} \quad (1),$$

where Y_{Pred} was the predicted intent and Y_{Test} was the true label/ true intent in the test dataset. It is a value between 0 and 1. Higher value indicates the accuracy/quality of prediction of the user intents by the models.

To further compare the classifications performed by the Dialogflow ML and the SVM models, Receiver Operating Characteristic (ROC) curve was used. It was used to plot ROC curve for each output class or user intent predicted by a classification model [33]. *True positive rate* (TPR) or *sensitivity* [TP/(TP+FN)] and *false positive rate* (FPR) or *false sensitivity* [FP/(FP+TN)] were used by determining the appropriate thresholds across the interval of [0, 1] to outputs. As each curve tilted towards the top left edge of the plot, the classification prediction was improved. Area Under the ROC curve (AUC) scores were used to compare the two classification models through *sensitivity* versus *1-specificity* [1- TN/(TN+FP)].

B. Evaluating Dialogflow ML & SVM in recognizing suicidal intent

To understand the underlying ML performance, we trained both the Dialogflow hybrid-ML model and the SVM model with the training dataset that consisted of 4 categories of user intents. The test data was a smaller dataset that also consisted of the similar categories, as stated in Table I. User intents were predicted by the proposed Dialogflow ML and the SVM

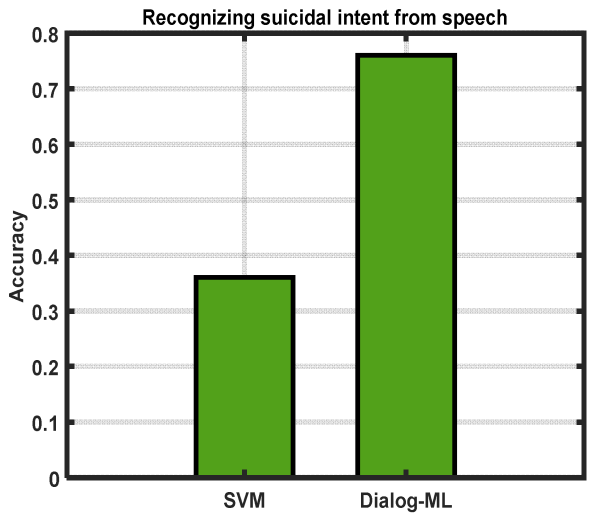
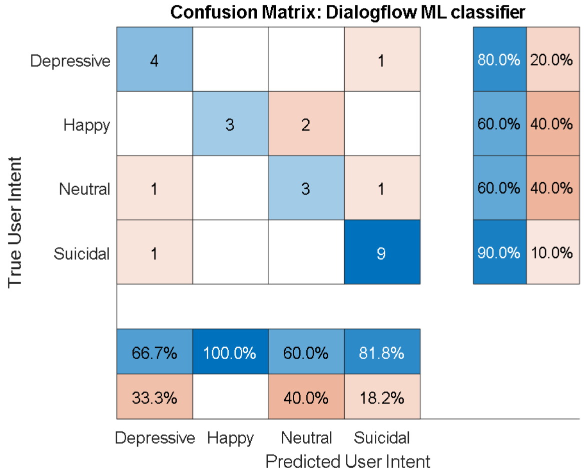


Fig. 9. Recognizing suicidal intent using Dialogflow ML and SVM model.

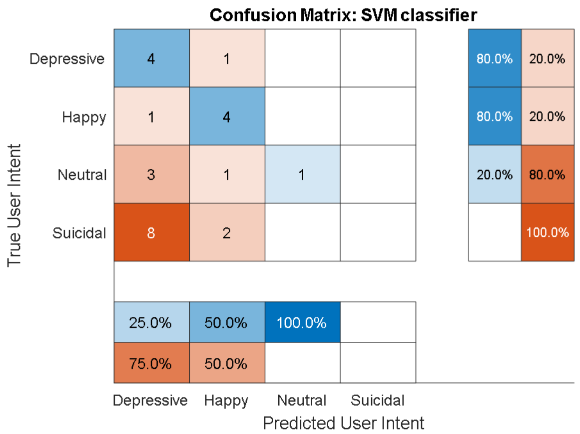
model and the classification accuracies were calculated according to Equ. (1). Dialogflow ML obtained overall 76% accuracy in classifying user intents in the four emotional states with a threshold value of 0.3, as shown in Fig. 9.

A key aspect to notice that the dialogflow ML algorithm is not open sourced, so the machine learning and speech recognition analysis happens as a black box. Therefore, to evaluate the proposed model performance, NLP of user intent was carried out with the traditional SVM model. The SVM classifier obtained 36% accuracy in predicting user's emotional state which was lower than the Dialogflow ML.

Fig. 10 shows the confusion matrices for the Dialogflow-ML and the SVM classifier in predicting user intents of the four emotional states. The classification scores such as: *Accuracy*, *Error, Sensitivity, Specificity, Precision, False Positive Rate (FPR), F1 Score, Matthews Correlation Coefficient (MCC),* and *Kappa* are reported in Table III, while these exhibited their well-known expressions [34]. For the Dialogflow-ML, scores for sensitivity: 0.7250, specificity:



(a)



(b)

Fig. 10. Confusion matrix: a) Dialogflow-ML classifier, b) SVM classifier.

0.9167, precision: 0.7712, FPR: 0.0833, F1 score: 0.7336, MCC: 0.6627, kappa:0.3600 were obtained.

For the SVM classifier, some scores could not be calculated due to fatal failure in recognizing a certain intent.

It is noteworthy to mention that the Dialogflow ML could achieve 80% and 90% accuracies in predicting '*Depressive*' and '*Suicidal*' states, respectively. Misclassification occurred mainly in recognizing '*Neutral*' and '*Happy*' states, as shown in Fig. 10 (a). In case of misclassification for '*Suicidal*', the model predicted '*Depressive*' state. This revealed that the proposed system could identify depression in a person fairly well. It could become an issue when the model misclassified the non-suicidal intents as '*Suicidal*', because it would wrongfully trigger the *suicide prevention helpline*.

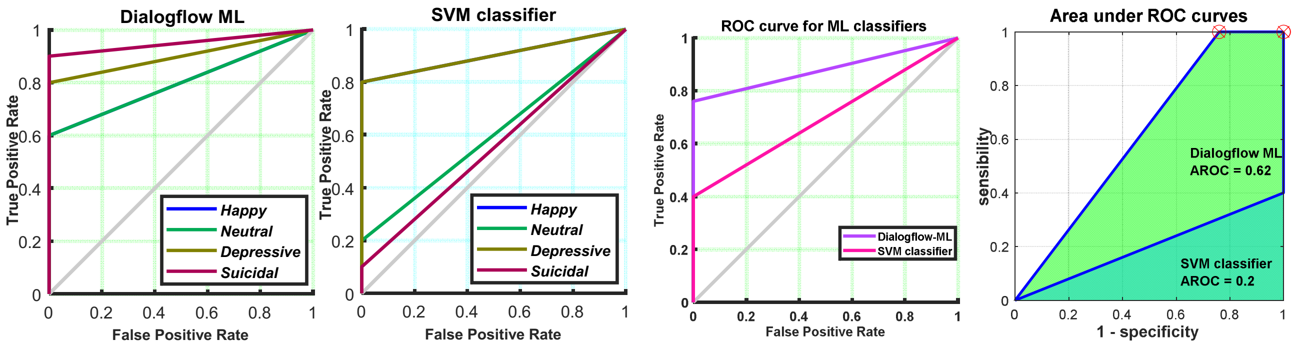
The SVM classifier was able to recognize '*Happy*' and '*Depressive*' intents well. Though the model misclassified '*Neutral*' as '*Depressive*' state. Unfortunately, it failed to recognize '*Suicidal'* intent entirely, as shown in Fig. 10 (b). This led to conclude the unsuitability of the model in predicting severe risks associated to depression in the population.

The classification of the four user intent classes were compared using the ROC curve. These curves showed that the Dialogflow ML classifier could effectively recognize the *'Depressive'* and '*Suicidal*' state. Also, performances of the Dialogflow-ML model and the SVM classifier were compared using ROC and area under ROC curve, as shown in Fig. 11. These results further support the superiority of the proposed platform in predicting the user emotional states in depressed population.

C. Limitations & Future Works

This paper aimed to identify the suicidal tendencies in a depressive population, however, there were a few limitations that we need to acknowledge. One of the key limitations was the nature of VPA-DR conversation initiation technique. It was observed in a depressive population with potential suicidal inclinations that they rarely initiated conversations about their mental states [2]. Since the VPA-DR was depended on the end user's willingness to activate the system, it could not initiate diagnosis for a person lacking in self-motivation to trigger this feature. The other thing to note that, in sentiment analysis, it is imperative to include non-verbal behavioral cues such as, pitch, response time, sleep pattern, anxiety levels. In a traditional clinical setting, the most prevalent and effective form of diagnostic tool for the screening process of depression (leading to suicidal tendencies) is based on these non-verbal behavioral patterns. But the proposed system did not include these non-verbal cues as a diagnostic criterion due to the lack of resources.

Though it is not known how the Dialogflow ML worked, but it could predict user intents or sentiments better than the traditional SVM classifier. Even if the proposed model based on Dialogflow ML model with multi-layer platform design provided better performance, still it requires further improvements for proper diagnosis in practical applications. It was observed that the model performed quite well in '*Happy'*



(a) (b)

Fig. 11. Receiver Operating Characteristic curve: a) comparing each classes of user intents, b) comparing classifiers with ROC curve and area under ROC curve.

and '*Suicidal*' intent. In few test trials, '*Depressive'* was recognized with '*Suicidal*' which would trigger the *suicide prevention helpline.* It would rather be acceptable to have a healthcare worker contacting the individual to acknowledge his/her mental health with probable depressive state. The system posed false positive and false negative assertions in a very few cases, which could lead to incorrect diagnosis of the mental state of a user and would impact negatively in real-life application.

Training the Dialogflow-ML with more diverse data would make the model versatile and useful in real-life scenarios. Although the proposed system predicted the user intents quite impressively, it is not known how the model would behave in uncertain situations. Also, it is hard to find real and suitable datasets to train the model. Therefore, we could only implement the four basic emotion states. Implementing deep learning with big data would be preferable in recognizing suicidal intent. Further improvements of this virtual assistant can be achieved by incorporating voice and sentiment recognition through the user's tone, pitch, cadence to yield a more accurate level of diagnosis of such suicidal intentions. Also, it can be implemented in smart phones or in smart homes for early diagnosis of depression with low cost, with ease of implementation and operation. This cost-effective multi-layer platform can be easily integrated and might be useful in future identifying level of assistance required to treat depression and avoid suicidal attempts.

VI. Conclusion

In this study, a novel multi-platform model was implemented for recognizing user intents in depressed population. It utilized speech content analysis as an objective method for detecting depression and suicidal tendencies. The proposed diagnostic tool focused both on diagnosing depression and immediately providing resources and follow up treatment plans. Thus, the methodology described in this paper yielded application-based results and performed reasonably good in detecting users' various emotional intents with an overall accuracy of 76%, and 80-90% accuracies in detecting *Depressive* and *Suicidal* intents.

With the advent of Internet of Things (IoT), the proposed affordable diagnostic tool based on personal assistance can easily detect depression at earlier stage and obtain necessary support. This diagnostic tool can also be prescribed for a person with depression and provide an informal, continuous support at home. Therefore, the application of the proposed model can be manifold either working standalone or in congestion with mental health services. The cost-effective and user-friendly platform can be an ideal candidate in smart home services for early screening of depression in mass population.

References

[1] Diagnostic and Statistical Manual of Mental Disorders, 5th ed.; DSM–5; American Psychiatric Association [APA], 2013.

[2] K. Hawton, C.C. Comabella, C. Haw, K. Saunders, 'Risk factors for suicide in individuals with depression: A systematic review,' *Journal of Affective Disorders*, vol. 147, May 2013, pp. 17–28.

[3] N. Cummins, S. Scherer, J. Krajewski, and et al., 'A review of depression and suicide risk assessment using speech analysis,' *Speech Communication*, vol. 71, Mar. 2015, pp. 10–49.

[4] World Health Organization. Depressions. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression

[5] World Health Organization. Mental Disorders. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-disorders.

[6] World Health Organization. Suicide. [Online]. Available https://www.who.int/news-room/fact-sheets/detail/suicide.

[7] J.J. Mann, A. Apter, J. Bertolote and et al. 'Suicide Prevention Strategies: A Systematic Review,' *JAMA*, vol. 294, no.16, Oct. 2005, pp. 2064–2074.

[8] C. Sikorski, M. Luppa, H. König, and et al., 'Does GP training in depression care affect patient outcome? - A systematic review and meta-analysis,' *BMC Health Serv Res*, vol. 10, Jan. 2012, pp. 12.

[9] Regier DA, Kaelber CT, Rae DS, and et al., 'Limitations of Diagnostic Criteria and Assessment Instruments for Mental Disorders: Implications for Research and Policy,' *Arch Gen Psychiatry*, vol. 55, no. 2, 1998, pp. 109–115, doi:10.1001/archpsyc.55.2.109.

[10] R.W. Lam, R. Milev, S. Rotzinger, and et al., 'Discovering biomarkers for antidepressant response: protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort,' *BMC Psychiatry*, vol. 16, no. 105, 2016, https://doi.org/10.1186/s12888-016-0785-x.

[11] J. K Payne, and et.al., 'Effect of Exercise on Biomarkers, Fatigue, Sleep Disturbances, and Depressive Symptoms in Older Women With Breast Cancer Receiving Hormonal Therapy,' *Oncology Nursing Forum,* vol. 35, no. 4, Pittsburgh, Jul 2008, pp. 635– 642.

[12] G.Orrùa1, and et.al., 'Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review,' *Neuroscience & Biobehavioral Reviews*, vol. 36, no. 4, Apr 2012, pp. 1140–1152.

[13] P. Peter, C-G. Mardi, 'Meditative Movement for Depression and Anxiety,' *Frontiers in Psychiatry*, vol. 4, 2013, pp. 71, https://www.frontiersin.org/article/10.3389/fpsyt.2013.00071 , doi:10.3389/fpsyt.2013.00071.

[14] N.Cumminsab, and et al., 'A review of depression and suicide risk assessment using speech analysis,' *Speech Communication*, vol. 71, Jul. 2015, pp. 10–49, https://doi.org/10.1016/j.specom.2015.03.004.

[15] M. Cannizzaro, N. Reilly and P. J. Snyder. 'Speech Content Analysis in Feigned Depression,' *Journal of Psycholinguistic Research*, vol. 33, no. 4, Jul. 2004.

[16] T.E. Oxman, S.D. Rosenberg, P.P. Schnurr, and G.J. Tucker. 'Diagnostc classification through content analysis of patients' speech,' *The American Journal of Psychiatry*, vol.145(4), pp. 464–468. 1988.

[17] Z. Huang, J. Epps and D. Joachim, "Speech Landmark Bigrams for Depression Detection from Naturalistic Smartphone Speech," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, United Kingdom, 2019, pp. 5856–5860, doi: 10.1109/ICASSP.2019.8682916.

[18] D. Ben-Zeev, E.A. Scherer, R. Wang, H. Xie, and A.T. Campbell, 'Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health,' *Psychiatric Rehabilitation Journal*. vol. 38, pp. 218–226, Sept, 2015.

[19] J.R. Williamson, E. Godoy, M. Cha and et al., 'Detecting Depression using Vocal, Facial and Semantic Communication Cues,' *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge*, Oct. 2016, pp. 11–18, https://doi.org/10.1145/2988257.2988263.

[20] V. Këpuska and G. Bohouta, 'Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home),' *IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, 2018, pp. 99–103, doi: 10.1109/CCWC.2018.8301638.

[21] H. Kaur, V. Mangat and Nidhi, 'A survey of sentiment analysis techniques,' *International Conference on IoT in Social, Mobile, Analytics and Cloud (I-SMAC)*, Palladam, 2017, pp. 921–925, doi: 10.1109/I-SMAC.2017.8058315.

[22] G. Coppersmith, R. Leary, P. Crutchley, and A. Fine, 'Natural language processing of social media as screening for suicide risk,' *Biomedical Informatics Insights*, Aug. 2018, https://doi.org/10.1177/1178222618792860.

[23] M. Guidère, 'NLP Applied to Online Suicide Intention Detection,' *HealTAC* 2020, Mar. 2020.

[24] Q. Zhong, E.W. Karlson, B. Gelaye, and et al., 'Screening pregnant women for suicidal behavior in electronic medical records: diagnostic codes vs. clinical notes processed by natural language processing,' *BMC Med. Info. Decis. Mak.*, vol.18, no. 30, May 2018.

[25] K. Haerian, H. Salmasian and C. Friedman, 'Methods for identifying suicide or suicidal ideation in EHRs,' *Annual Symposium proceedings, AMIA Symposium*, vol. 2012, pp. 1244–1253, Nov. 2012.

[26] Intent matching - google clouds. Online. Available https://cloud.google.com/dialogflow/es/docs/intents-matching.

[27] Twilio Conversations API. Online. Available: https://www.twilio.com/conversations.

[28] Traning Dataset. [Online].
Available: https://www.kaggle.com/ritresearch/happydb.

[29] Traning Dataset. [Online].
Cognitive Systems,' *AI Magazine*, vol. 38, no. 4, pp. 43–56, 2017. https://doi.org/10.1609/aimag.v38i4.2745.

[32] Low Code Dialogflow bots. Mining Business Data. [Online]. Available: https://miningbusinessdata.com.

[33] A. Bradley, and P. Andrew, 'The use of the ares under the ROC curve in the evaluation of machine learning algorithms,' *Pattern Recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.

Available: https://github.com/CodeWritingCow/suicide-notes.

[30] M. Canonico, L. De Russis, 'A Comparison and Critique of Natural Language Understanding Tools,' Ninth International Conference on Cloud Computing, GRIDs, and Virtualization. ((Intervento presentato al convegno CLOUD COMPUTING, pp. 110-115, Barcelona, Spain, Feb. 18–22, 2018.

[31] M. McShane, 'Natural Language Understanding (NLU, not NLP) in

[34] R. Silipo, and Maari. Widmann, 'Confusion Matrix and Class Statistics, towards data science,'. [Online]. Available: https://towardsdatascience.com/confusion-matrix-and-class-statistics-68b79f4f510b.

TABLE II. TEST DATASET USED IN EVALUATING ML TECHNIQUES

Category	*Test Speech*
Happy	This is the best day of my life!
Happy	What an exciting news
Happy	I am getting a promotion today
Happy	I have found the love of my life
Happy	Amazing art
Neutral	seven days in a week.
Neutral	math is science
Neutral	trees live longer.
Neutral	Walk straight until you reach Main street
Neutral	Day and night are opposite.
Depressive	Isn't it so tiring to be alive?
Depressive	I don't want to wake up anymore
Depressive	I am feeling depressed
Depressive	darkness around us makes us invisible
Depressive	It's so meaningless to do anything
Suicidal	I really want help now i just cant so this anymore alone i feel hopeless
Suicidal	I cannot keep doing this anymore. Even after 3 years of self-healing, and therapy, I cannot find peace in anything I do.
Suicidal	I've lost all my ambitions and I'm emotionless. I still care about my loved ones and friends but I'm not strong enough to keep going. I'm trying to get the courage to end my life. I want to end my life next month. How? I still don't know.
Suicidal	I'm finally accepting that the best thing for me to do is just die. I won't have to deal with anything anymore and everyone can just forget I ever excited and it's a win win for everybody.
Suicidal	I just want to kill myself. I am so fucking tired. That's all. Too many people in my fucking house. I want this to be done.
Suicidal	i wish I wasn't born. I just can't do this anymore. I wanna disappear off of the face this planet, someone please end me
Suicidal	Why am I even bothering to write this?? Maybe I feel less lonely when I do so.. I'm sorry. God! I hate myself so much... I just want to leave.
Suicidal	I just wanna kill myself. But I cannot find the courage to. I just need the courage to.
Suicidal	I just want to kill myself. I am so fucking tired. That's all. Too many people in my fucking house. I want this to be done.
Suicidal	I'm going to kill myself

TABLE III. CLASSIFICATION SCORES OF ML MODELS

ML model	*Accuracy*	*Error*	*Sensitivity*	*Specificity*	*Precision*	*False Positive Rate*	*F1 Score*	*Matthews Correlation Coefficient*	*Kappa*
Dialogflow-ML	0.7600	0.2400	0.7250	0.9167	0.7712	0.0833	0.7336	0.6627	0.3600
SVM classifier	0.3600	0.6400	0.4500	0.8000	NaN	0.2000	NaN	NaN	0.4141