World Academy of Science, Engineering and Technology
International Journal of Psychological and Behavioral Sciences
Vol:15, No:6, 2021

Early Depression Detection for Young Adults with a Psychiatric and AI Interdisciplinary Multimodal Framework

Raymond Xu, Ashley Hua, Andrew Wang, Yuru Lin

***Abstract*—**During COVID-19, the depression rate has increased dramatically. Young adults are most vulnerable to the mental health effects of the pandemic. Lower-income families have a higher ratio to be diagnosed with depression than the general population, but less access to clinics. This research aims to achieve early depression detection at low cost, large scale, and high accuracy with an interdisciplinary approach by incorporating clinical practices defined by American Psychiatric Association (APA) as well as multimodal AI framework. The proposed approach detected the nine depression symptoms with Natural Language Processing sentiment analysis and a symptom-based Lexicon uniquely designed for young adults. The experiments were conducted on the multimedia survey results from adolescents and young adults and unbiased Twitter communications. The result was further aggregated with the facial emotional cues analyzed by the Convolutional Neural Network on the multimedia survey videos. Five experiments each conducted on 10k data entries reached consistent results with an average accuracy of 88.31%, higher than the existing natural language analysis models. This approach can reach 300+ million daily active Twitter users and is highly accessible by low-income populations to promote early depression detection to raise awareness in adolescents and young adults and reveal complementary cues to assist clinical depression diagnosis.

***Keywords*—**Artificial intelligence, depression detection, facial emotion recognition, natural language processing, mental disorder.

I. Introduction

DEPRESSION (a mental health disorder) is characterized by persistently depressed mood or loss of interest in activities, causing significant impairment in daily life including disruption in education, work, relationships, and roles people play in society. According to the World Health Organization, depression is the leading cause of disability worldwide, affecting an estimated 350 million people [1].

Depression onset increases dramatically during the middle and high school years, with nationally representative prevalence rates for adolescents estimated at 7.5% [2]. “At its worst, depression can lead to suicide. Close to 800,000 people die due to suicide every year. Suicide is the second leading cause of death in 15-29-year-olds” [1].

During the COVID-19 pandemic, depression rates increased by over 40.9% comparing the national survey data in mid-July 2020 and the data in 2019 conducted by the U.S. Census Bureau [3]. The evidence shows that young adults are most vulnerable to the mental health impacts of the pandemic. The timely diagnosis of depression becomes an utmost necessity globally. The delayed detection and intervention process (2–6 years) leads to the drop of the recovery rates, approximately 60% at 2 years, 40% at 4 years, and 30% at 6 years with comorbid anxiety having a key role in limiting recovery [4].

This research’s objective is to develop an interdisciplinary multimodal AI framework that leverages Natural Language Processing (NLP), image processing as well as psychiatric clinical best practices to analyze the multi-channel depression indicative signals from adolescents and young adults. The benefits include the increased accuracy, large scale, being accessible at low or no cost to low-income families, and the early screening before further progression to neurovegetative and neurocognitive symptoms, according to the mental disorder symptoms taxonomy.

II. Related Work

A. Review of the Existing AI Research

Several machine learning approaches have been applied to social media data for depression detection.

Since 2000, machine learning algorithms have been popular for NLP sentiment analysis. More recently, deep learning has been applied for text classification. AI depression detection research results have been reported using Twitter, Facebook, and other social media posts.

Table I summarizes the research advancements in the reverse chronological order.

B. Review of Psychiatric Survey Approach

The number one impacting factor to AI solutions is the quality of the dataset. The data should contain the depression indicative signals. For this reason, we researched the depression symptoms definition and current clinical detection approaches in order to define the scope of our dataset.

Depression Symptoms

The APA Diagnostic and Statistical Manual defines the nine depression symptom groups as shown below [21]:

1. Sadness (Dysphoria)
2. Loss of Interest (Anhedonia)
3. Appetite
4. Sleep
5. Thinking/concentration

Raymond Xu is with The Harker School, United States (e-mail: readingclub25@gmail.com).

6. Guilt (Worthlessness)
7. Tired (Fatigue)
8. Movement (Agitation)
9. Suicidal ideation

TABLE I
REVIEW OF EXISTING DEPRESSION DETECTION AI RESEARCH

Methodologies	Limitations	Year
Mixed approach: lexicon-based and classification based (Support Vector Regression and Multinomial Naïve Bayes) [5]	High precision at the expense of low recall. Subjective to lexicon DB quality. Lexicon is not depression symptom category based, so words can be ambiguous.	2019
Hybrid Neural Network combining both text and images [6]	No leverage of well-established Psychiatric clinical practices. Not examine depression symptoms in prolonged period (two weeks).	2018
Transfer learning method: Universal Language Model Finetuning (ULMFit) and the Google research project Bidirectional Encoder Representations for language training [7]	Default Stochastic Gradient Descent (SGD) is not able to converge as quickly, sub-optimal training loss curve, hence further experimentation is needed.	2018
Neural Network Multitask Learning (MTL) framework, use demographic attributes and mental states jointly to improve prediction accuracy [8]	The simplest way (feed-forward network with n-gram input layer) in this study does not support predicting any number of NLP tasks from a Convolutional Neural Network (CNN).	2017
Neural Network Image-based VGG-Net [9]	Visual and textual features are the two main features. Hence, neural network models that only use visual information are limited in their performance.	2015
Neural Network text-based Gated Recurrent Unit (GRU) [10]	Social media posts often come in inconsistent or incomplete syntax with very limited surrounding textual contexts. Therefore, it's challenging for text-only models to perform well.	2014

The National Institute of Mental Health (NIMH) presents a list of questions that an individual should consider answering if he/she wants to check for depression [22]. All these questions are aimed to surface the symptoms listed above.

Self-Report and Clinical Screens

Three major approaches are being widely used across populations: CES-D, PHQ-2 and PHQ-9.

CES-D and CESD-R (Revised)

The Center for Epidemiologic Studies Depression Scale (CES-D) was created in 1977 by Laurie Radloff and revised in 2004 by William Eaton and others. CES-D is a self-report screening test containing 20 survey questions based on the past week's experience [11], [12]. Each of the 20 questions was carefully designed towards one of the nine depression symptoms.

PHQ-2 and PHQ-9

The Patient Health Questionnaire is a clinical tool used to assist the clinician in the diagnosis of depression, to quantify depression symptoms and to monitor the severity.

When screening for depression the Patient Health Questionnaire (PHQ-2) can be used first. If this is positive, the PHQ-9 can then be used, which has 61% sensitivity and 94% specificity in adults [13]. The questionnaire was developed by Drs. R.L. Spitzer, J.B.W. Williams, K. Kroenke and colleagues, with an educational grant from Pfizer, Inc. [15]. Therefore, in the data collection process, this study aimed to collect the signals indicating the above common features.

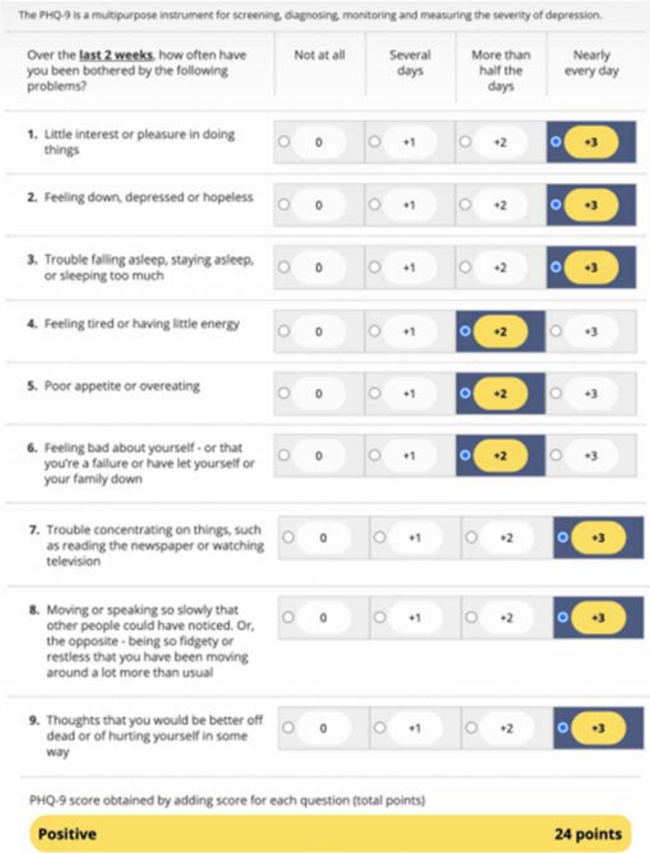

Fig. 1 Patient Health Questionnaire-9 [13]

III. PSYCHIATRIC AND AI INTERDISCIPLINARY MULTIMODAL FRAMEWORK

The proposal in this study leveraged the best practices from both psychiatric as well as AI fields. A multimodal AI framework was used on multi-channel information that was semantically correlated: videos/images non-verbal cues and natural language verbal cues. They complemented each other to reveal patterns that were not visible when working with individual channels alone. This proposed approach consolidated data from various sources to produce more accurate depression detection.

In Fig. 2, the blue boxes represent AI components, the green boxes represent best practices from the psychiatric field, and the stars represent the unique design for adolescents and young adults.

Details are explained in the following subsections in A through E order.

A. Data Collection

This research used both twitter data as well as survey results from a newly designed multimedia-based survey.

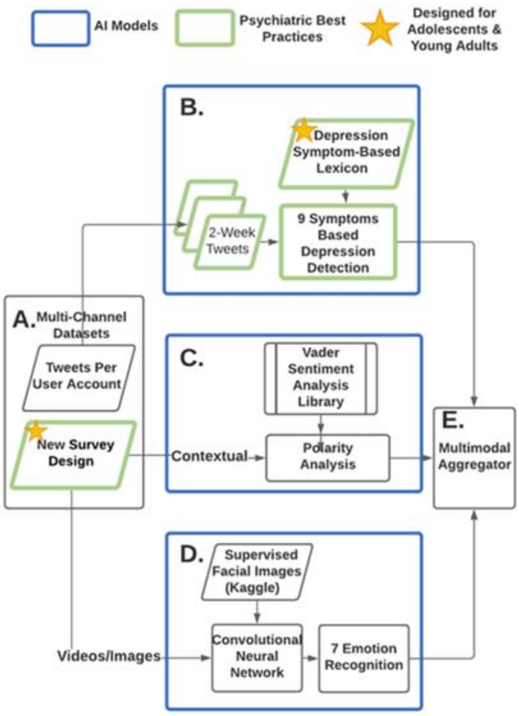

Fig. 2 Inter-Disciplinary Multimodal AI System Architecture: A. Data collection; B. 9 Symptoms Detection with NLP on Twitter Accounts; C. Depression Detection with Sentiment Analysis on Contextual Survey Results; D. Facial Emotion Recognition with CNN on Multimedia Survey Results; E.Multimodal Aggregator for the Final Prediction Result

1. Young Adult Age Relevant Survey Design for AI Analysis

The proposed solution transformed the traditional clinic questionnaire into a new survey design that enabled AI multimedia data analysis. Since the audience was adolescents and young adults who would take the survey outside of a clinical setting, the survey was made to be more age relevant and engaging. It included mood induction images intended to cause an emotional reaction. The image list used in the survey has been revised based on feedback from 51 US high school students.

Those surveyed were instructed to improvise a short video reflecting their emotional reaction to the chosen pictures.

Besides video data collection, the survey inherited some essential questions from the clinic questionnaire PHQ-9.

2. Twitter Data Collection and Pre-Processing

Binary logistic regression analysis revealed that youth who reported higher depression symptoms were two times more likely to disclose negative emotions online and were three times more likely to disclose various hassles online than their peers who reported lower symptoms [14]. Therefore, this study focused on Twitter data.

Twitter data per user account were retrieved using Twitter Search Engine Streaming API. The data were labeled into Depression and Non-Depression datasets based on the Twitter account's self-declared user profile (name, bio). Data cleansing and normalization was done before training.

Fig. 3 Proposed AI based depression survey form: video data collection

B. 9 Symptoms Detection with NLP on Twitter Accounts

To comply with psychiatric questionnaire-based diagnosis based on symptoms in the past two weeks, the Twitter data were grouped into 2-week batches. This depression analysis was based on a period of time instead of a snapshot of time per tweet to better reflect the status of a user. Also, to analyze each individual symptom, a symptom-based lexicon should be used. However, adolescents and young adults suffering from depression are not likely to directly describe their feelings. They tend to use their own vocabularies that are often overlooked in the general lexicons, like ulcers, bullying, down, irritable, and arguments, instead of the word “depressed”. In order not to miss the clues, this research defined young adults’ depression indicative lexicon after studying the following three areas:

- Adolescent and young adults’ online social media posts in the self-declared depression Twitter accounts
- The multi-media survey results analysis from 51 American high school students
- Existing research on depressed young people’s vocabulary choices [15]-[17]

How often have you been feeling the following ways in the past two weeks?

	Never	Rarely	Half the time	Most times	Always
Hopeless	○	○	○	○	○
Disinterest in doing things	○	○	○	○	○
Feel bad about yourself(ex. feeling like you have failed)	○	○	○	○	○
Feel like your family/friends are better off without you	○	○	○	○	○
Had thoughts about causing harm to yourself	○	○	○	○	○
Lost interest in things you used to enjoy	○	○	○	○	○
Pressured	○	○	○	○	○
Anxious	○	○	○	○	○
Loss of sleep or appetite	○	○	○	○	○
Seriously considered or have caused harm to yourself	○	○	○	○	○

Fig. 4 AI based depression survey form: contextual data collection

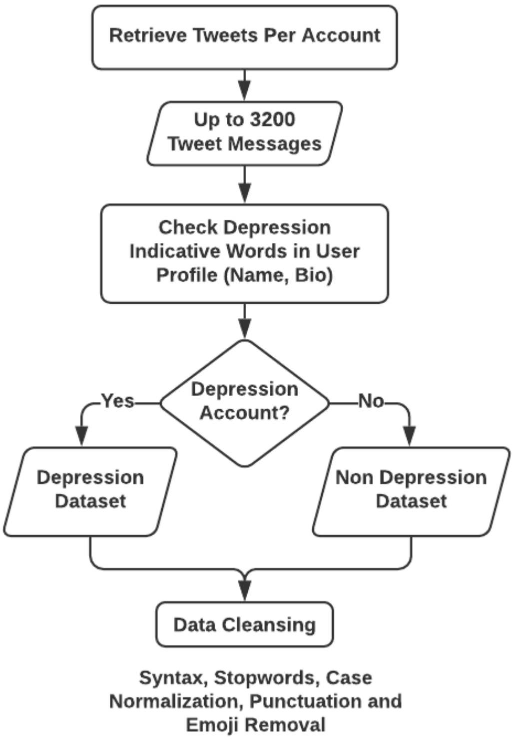

Fig. 5 Twitter data collection and pre-processing

For each batch, the system checked the symptom indicative words. For example, for symptom “Loss of Interest”, the system checked occurrences of words “absence of pleasure”, “moody”, “bored”, “no job”, etc. based on the Depression Symptom-Based Lexicon that was augmented with adolescents and young adults’ vocabularies.

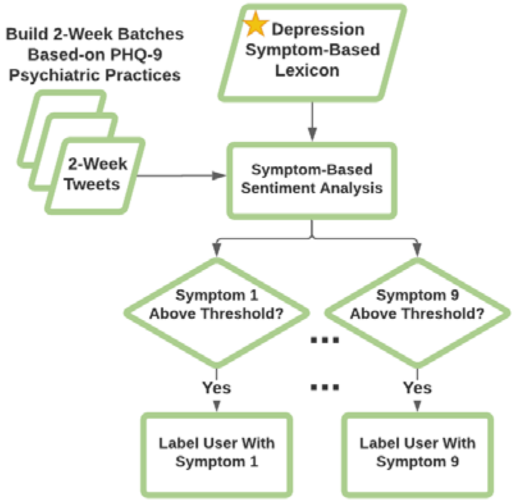

Fig. 6 Nine depression symptoms detection based on 2-week tweets

TABLE II
SYMPTOM-BASED LEXICON DEFINED FOR ADOLESCENTS AND YOUNG ADULTS

appetite disorder	sleep disorder	tired	concentration issue
can’t eat lost weight eating problems	can’t sleep sleep problems	stressed pressure ulcers tired all the time	distraction not focused disturbed
movement	**loss of interest**	**guilt**	**Suicidal ideation**
fatigue anxiety bullying	going through the motions moody reject love	blame deserve unhappy nobody ugly worthless unsuccessful	escape hurt wrong life no purpose no meaning
sadness			
break down irritable arguments	Alone Feel pain Lonely suck	worried uncomfortable not going to work whatever i do	nothing will change isolate hopeless uneasy

To assist age cohort analysis, this study used the AgePredictor library.

The prediction result for one of the experiment Twitter account is shown in Fig. 7. Twitter Account “Al...n” is obfuscated for privacy consideration.

```
Ranked Symptom Indicative Words ----------->
{'born': 0, 'die': 1, 'forgiveness': 2, ...}

Iterate Through Symptoms -------->
<Symptom 0>
Suicidal_Ideation
<Symptom 1>
Sadness
```

Fig. 7 Prediction result for one of the Twitter accounts

After verifying against the labeled test data (label was extracted from the Twitter account self-declared user profile), the Precision yielded at 77.77 %, recall at 98.34% and accuracy at 86.30% for Twitter Account Depression Detection (label -1 for depression).

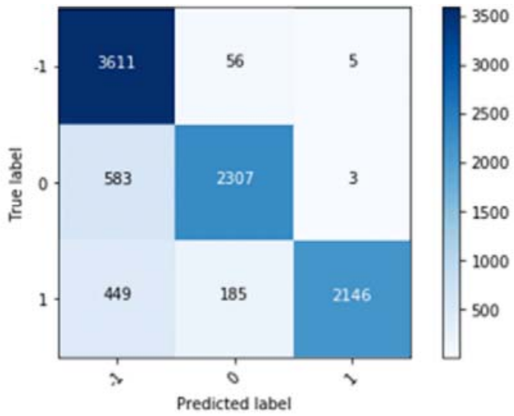

Fig. 8 ConfusionMatrix for experiment on 10K dataset

To visualize the keywords in the collected message, WordCloud library was used. The word cloud may show partial words because this study obfuscated some inappropriate words for the audience.

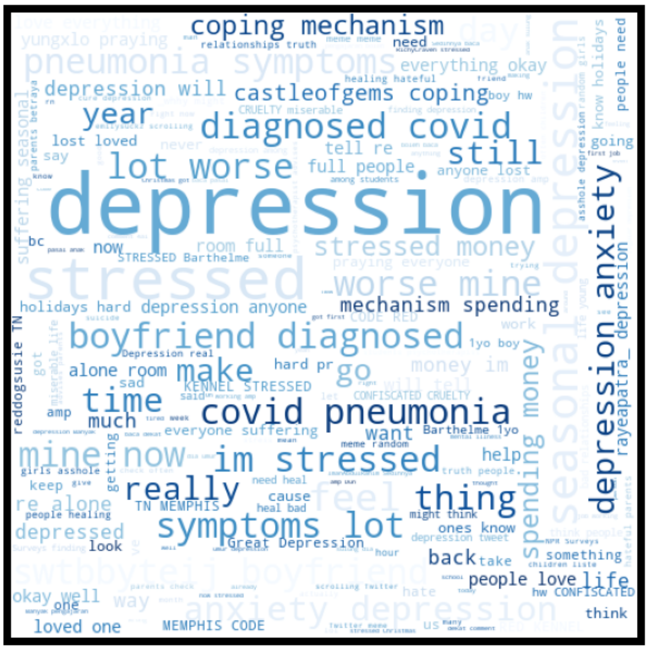

Fig. 9 Depression WordCloud

C. Depression Detection with Sentiment Analysis on Contextual Survey Results

The new survey was sent to 300+ US high school and college students covering both depressed and non-depressed users. A total of 223 participated. After filtering out the partial results, 89 were used as quality survey results. Among the quality ones, 18 were labeled with Depression and 71 with Non-Depression based on users' self-declaration.

The contextual content from the young adults' survey results (user typed in content as well as audio-to-text conversion content) was sent to AI component C for sentiment polarity analysis.

First, each survey result went through the sentiment polarity check using Vader library. Vader is a rule-based sentiment analysis library. Comparing with TextBlob, Vader works slightly better with user generated short content. The sentiment score was assigned to each survey result. Then, TF-IDF was used for text vectorization and feature extraction.

Term Frequency-Inverse Document Frequency (TF-IDF) measures how relevant a word is to a document in respective to a collection of documents, a.k.a. corpus. The more frequently a word occurs in a document and the less documents in the corpus contain this word, the more relevant this word to this document. E.g., if a document has high occurrence of word "medicine", and this word is not very popular in the whole corpus, this means that "medicine" is one of the keywords that represent the meaning of this document. However, though words such as "this", "and" occur very frequently in a document, they also occur in many other documents in the corpus, so they are not good representations of the meaning of the document and hence are ranked low.

In this AI component C, Vader's sentiment scores and TF-IDF extracted features were sent to the logistic regression model for classification. The accuracy peaked at 87.76%.

TABLE III
TF-IDF EXPERIMENTS PERFORMANCE METRICS, TUNING PARAMETER: SENTIMENT SCORE THRESHOLD

Sentiment score threshold	precision	recall	f-score	accuracy
0.6	72.01%	94.01%	81.55%	87.10%
0.4	74.51%	92.69%	82.61%	91.19 %
0.2	81.47%	91.02%	85.98%	92.32 %
0	91.61%	79.66%	85.22%	92.53%
-0.2	98.73%	73.95%	84.56%	89.07%
-0.4	99.45%	68.94%	81.43%	85.51%
-0.6	99.76%	69.91%	82.21%	85.03%

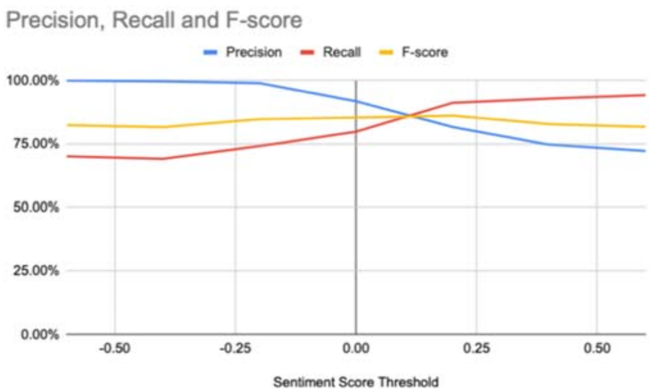

Fig. 10 TF-IDF metrics line chart

The average accuracy for TF-IDF was 88.96%. Considering TF-IDF yielded a more consistent trend, TF-IDF was adopted in the final AI framework in this study.

D. Facial Emotion Recognition with CNN on Multimedia Survey Results

This study took a multimodal approach to analyze both non-verbal and verbal cues for increased accuracy.

Kinesics of the face, or facial expressions are part of the essential nonverbal cues. Facial emotion recognition models have three main elements: face and facial component detection, feature extraction, and expression classification. The key variations in different recognition techniques lie within feature extraction and the classifier. Conventional approaches to this problem include several different feature extraction techniques: 1) geometric feature extraction, 2) appearance feature extraction based on recording various statistics of the pixels' values within the face image, and 3) deep learning CNN.

Among all above approaches, this study chose approach 3) for supervised learning. CNN is a better fit for recognizing and accurately categorizing image and video input. In the convolution layer, the computer extracts high level features by moving a kernel, also known as filter, over the entire image and reducing its dimensions. The max pooling layer also reduces the size of the features by extracting the dominant convolved features. Both of these are used to reduce computational power. It is also resistant against shifts or distortions in the input.

The labeled dataset contains 35 thousand labeled images of people with different facial emotions: anger, fear, disgust, happiness, sadness, surprise, and neutral.

In the data preprocessing phase, the features' values were normalized to a common scale between 0 and 1 to facilitate efficient gradients process and avoid oscillation during training.

In the CNN architecture, three hidden layers were built with kernel size of 3x3 and activation function of rectified linear unit (ReLU). Max Pooling was applied to reduce the dimension sizes and prevent model overfitting. Additionally, Dropout function was added to avoid overfitting. Towards the end of the architecture, neurons were linearized with the Flatten function to enter the fully connected Dense layer. For the output layer, the Softmax activation function was used to output the multiclass classification results: categorizing the images to seven emotion groups.

Adam optimizer was applied to optimize the weights by minimizing the cost. Batch and shuffle were implemented to prevent the potential interference from the local optima during training.

During the prediction phase, Haarcascade was used to recognize facial areas to drive higher accuracy. The videos from the survey result were processed to frames of images using Cv2 library.

After tuning the hyper-parameters and adjusting the epochs, the model peaked at 83% accuracy when epochs reached 80.

E. Multimodal Aggregator for the Final Result

Towards the end of AI prediction, the system aggregated the AI scores from both facial non-verbal cue analysis and natural language-based depression symptoms analysis to determine the final prediction result for the given user.

Five experiments each conducted on 10k data entries reached consistent results with an average accuracy of 88.31%, higher than the existing natural language analysis models.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 46, 46, 64)        640
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 44, 44, 64)        36928
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 22, 22, 64)        0
_________________________________________________________________
dropout_1 (Dropout)          (None, 22, 22, 64)        0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 20, 20, 64)        36928
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 18, 18, 64)        36928
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 9, 9, 64)          0
_________________________________________________________________
dropout_2 (Dropout)          (None, 9, 9, 64)          0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 7, 7, 128)         73856
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 5, 5, 128)         147584
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 2, 2, 128)         0
_________________________________________________________________
flatten_1 (Flatten)          (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 1024)              525312
_________________________________________________________________
dropout_3 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_2 (Dense)              (None, 1024)              1049600
_________________________________________________________________
dropout_4 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_3 (Dense)              (None, 7)                 7175
=================================================================
Total params: 1,914,951
Trainable params: 1,914,951
Non-trainable params: 0
_________________________________________________________________
```

Fig. 11 Overview of the CNN architecture

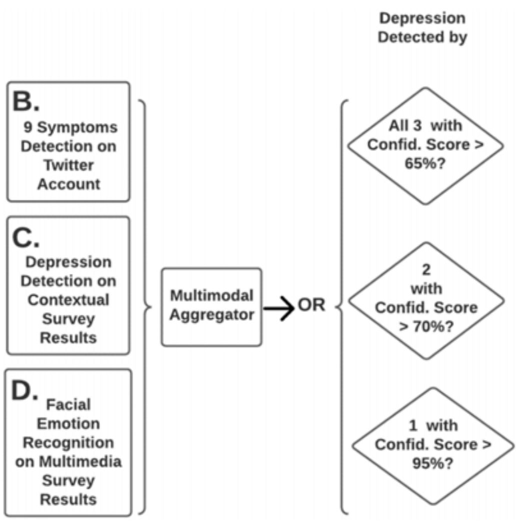

Fig. 12 Multimodal Aggregator

IV. Next Steps

A. Dataset Expansion

This study was based on five sets of 10K Tweets and 89 quality survey results from US adolescents and young adults. In the next phase, the survey should be expanded to a broader population covering varieties of genders, regions, and ethnicities.

B. Longitudinal Data Analysis

In a longitudinal study, researchers repeatedly examine the same individuals to detect any changes that might occur over a period of time, whereas cross-sectional study examines different samples (or a “cross-section”) of the population at one point in time [18]. Longitudinal study is an observational one and it is often applied in human psychology to understand societal and human behaviors and to come up with predictors of some human nature problems such as diseases or mental disorders. There have been longitudinal studies focusing on specific aspects of depression, such as the gender differences in clinical depression [19] and the association between depression and mortality [20]. As the next step of this research, a longitudinal study is highly recommended.

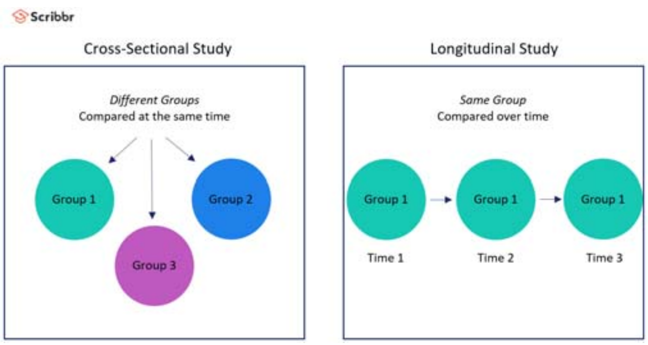

Fig. 13 Longitudinal vs. Cross-sectional [18]

V. Conclusions

Based on the studies on the current major self-report and clinical screening tools (CESD, PHQ-2 and PHQ-9), and the AI experiments on the NLP depression symptoms, sentiment analysis, and the facial emotional recognition on the 50k+ Twitter data and the 89 multimedia survey results from the US adolescents and young adults, we concluded that the interdisciplinary multimodal AI framework can produce 88.31% accuracy in detecting depression signals. Therefore, AI tools built on top of the psychiatric best practices can play a big role in depression detection before the symptoms progress to neurovegetative and neurocognitive degrees. AI can be used as an early screen tool to raise awareness in adolescents and young adults and can reveal complementary cues to assist clinical depression diagnosis.

References

[1] “Depression.” https://www.who.int/news-room/fact-sheets/detail/depression (accessed Apr. 29, 2021).

[2] S. Avenevoli, J. Swendsen, J.-P. He, M. Burstein, and K. Merikangas, “Major Depression in the National Comorbidity Survey- Adolescent Supplement: Prevalence, Correlates, and Treatment,” *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 54, Oct. 2014, doi: 10.1016/j.jaac.2014.10.010.

[3] R. Feintzeig, “Is It OK to Reveal Your Anxiety or Depression to Your Boss?,” *Wall Street Journal*, Sep. 13, 2020.

[4] “Depression,” *Cmu.edu*, 2021. https://www-sciencedirect-com.proxy.library.cmu.edu/science/article/pii/S0140673618319482?via%3Dihub (accessed Apr. 29, 2021).

[5] P. Arora and P. Arora, “Mining Twitter Data for Depression Detection,” *IEEE Xplore*, 2019. https://ieeexplore.ieee.org/document/8938353/authors#authors (accessed Apr. 29, 2021).

[6] S. Moon, L. Neves, and V. Carvalho, “Multimodal Named Entity Recognition for Short Social Media Posts,” *arXiv:1802.07862 (cs)*, Feb. 2018, Accessed: Apr. 29, 2021. (Online). Available: https://arxiv.org/abs/1802.07862.

[7] J. Howard and S. Ruder, “Universal Language Model Fine-tuning for Text Classification,” *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2018, doi: 10.18653/v1/p18-1031.

[8] A. Benton, M. Mitchell, and D. Hovy, “Multitask Learning for Mental Health Conditions with Limited Social Media Data,” *ACLWeb*, Apr. 01, 2017. https://www.aclweb.org/anthology/E17-1015/ (accessed Apr. 29, 2021).

[9] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *arXiv.org*, 2014. https://arxiv.org/abs/1409.1556.

[10] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling,” *arXiv.org*, 2014. https://arxiv.org/abs/1412.3555.

[11] CESD-R, “Center for Epidemiologic Studies Depression Scale Revised Online Depression Assessment” (Online) Available: https://cesd-r.com

[12] Center for Epidemiologic Studies - Depression Scale (CES-D) | Measurement Instrument Database for the Social Sciences. (Online) Available: https://www.midss.org/content/center-epidemiologic-studies-depression-scale-ces-d

[13] PHQ-9 Depression Test Questionnaire (Online) Available: https://patient.info/doctor/patient-health-questionnaire-phq-9 and https://www.hiv.uw.edu/page/mental-health-screening/phq-9

[14] Ariel Shensa, Jaime E. Sidani, Cesar G. Escobar-Viera, et al. “Emotional support from social media and face-to-face relationships: Associations with depression risk among young adults” (Online) Available: https://www.tandfonline.com/doi/full/10.1080/09638237.2019.1581357?scroll=top&needAccess=true

[15] Y. Ophir, C. S. C. Asterhan, and B. B. Schwarz, “Unfolding the notes from the walls: Adolescents’ depression manifestations on Facebook,” *Computers in Human Behavior*, vol. 72, pp. 96–107, Jul. 2017, doi: 10.1016/j.chb.2017.02.013.

[16] K. Harvey, “Disclosures of depression,” *International Journal of Corpus Linguistics*, vol. 17, no. 3, pp. 349–379, Dec. 2012, doi: 10.1075/ijcl.17.3.03har.

[17] R. Dryden-Edwards, “Teen Depression Facts, Treatment, Symptoms, Statistics & Tests,” *MedicineNet*, 2019. https://www.medicinenet.com/teen_depression/article.htm.

[18] Thomas, Lauren. *What is a longitudinal study?*. (Online) Available: https://www.scribbr.com/methodology/longitudinal-study

[19] B. L. Hankin, L. Y. Abramson, T. E. Moffitt, P. A. Silva, R. Mcgee, and K. E. Angell, “Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study.,” *Journal of Abnormal Psychology*, vol. 107, no. 1, pp. 128–140, 1998.

[20] S. E. Gilman, E. Sucha, M. Kingsbury, N. J. Horton, J. M. Murphy, and I. Colman, “Depression and mortality in a longitudinal study: 1952–2011,” *CMAJ*, 23-Oct-2017. (Online). Available: https://www.cmaj.ca/content/189/42/E1304.short. (Accessed: 17-May-2021).

[21] “APA PsycNet,” *psycnet.apa.org*. https://psycnet.apa.org/record/2013-14907-000.

[22] “NIMH» Depression Basics,” *www.nimh.nih.gov*. https://www.nimh.nih.gov/health/publications/depression/.