

**iPonder: A Multimodal Approach to Teen Mental Health Diagnosis and Therapy via
Deep Transfer Learning**

Vivek Kogilathota
Reedy High School

Abstract

Suicide is the second leading cause of death for teenagers. Even with suicide rates in teenagers being horrifyingly high, over 50% of suicidal teens don't ask for help or reach out in any way. Whereas depressed adults have the facilities to independently seek therapy, depressed teens not only feel an obligation to deal with these serious issues by themselves to avoid burdening those around them but also keep quiet due to their own shame. Thus, I built iPonder, a cross-platform accessible mobile application for teenagers to privately discuss, confront, and overcome their depression. Using a four-level diagnosis approach based on user data, responses to psychological questionnaires, choice specification, and free-response answers, iPonder constructs a hyper-personalized user profile, which is used to pair teenagers anonymously online with similar profiles to discuss and vent about their issues. Questionnaires include the Patient Health Questionnaire- 9 and General Anxiety Disorder- 7. Using Reddit's Self-Reported Mental Health Diagnosis text-corpus with over 200,000 data points, six artificial neural networks were constructed to predict user risk of various mental illnesses based on their free-text-responses, using transfer learning from a token-based text embedding neural network with over 48 million parameters trained on the English Google News 7B corpus with over 4 billion data points. All neural networks were evaluated on their accuracy, precision, recall, and f1-score. Besting state-of-the-art metrics on similar tasks in the relevant literature, iPonder achieves over 95% accuracy, recall, precision, and f1-score per text-based neural network and over 80% for video-based neural networks, and successfully integrated these models into the mobile application via a Python Flask application programming interface. Based on the profile created by the application based on the neural network scores, iPonder anonymously pairs similar users online to discuss their issues via conversational therapy and provides a dedicated daily journaling section to track emotional progress. This provides the anonymous yet effective therapy deeply desired by teenagers, pointing to a new age of accessible, effective, and anonymous mental health therapy for teens.

Acknowledgment of Major Assistance

All novel research done as part of this study was accomplished solely through my supervision, modification, and creation. All steps of the procedure and all background research was done at my home.

I would like to pay my special regards to the Institute of Creative Technologies at the University of Southern California for permitting me to use the Distress Analysis Interview Corpus - Wizard of Oz to analyze visual, aural, and textual clues for depression, anxiety, and suicide watch. The novel research and meaningful contribution regarding multimodality and user privacy of this paper would not have been possible without this resource. The contents of this dataset were made available upon submitting a release form to the university and obtaining a confirmation email and login information back from the administrator.

I would also like to pay my special regard to the Reddit company for permitting me to use the Reddit Self-Reported Mental Health Diagnosis Corpus to analyze anonymous sources of labeled and classified text on depression, anxiety, suicide watch, and lack thereof. The novel multimodal dataset contributed by this paper would not have been possible without this source of rich and abundant data. The contents of this dataset were made available upon submitting a release form to the company and obtaining a confirmation email and file link back from the administrators.

iPonder: A Multimodal Approach to Teen Mental Health Diagnosis and Therapy via Deep Transfer Learning

Introduction

Problem

Suicide is the second leading cause of death for teenagers, and more teenagers and young adults die from suicide than from cancer, heart disease, AIDS, birth defects, stroke, pneumonia, influenza, and chronic lung disease, combined. Even with suicide rates in teenagers being extremely high, over 50 percent of suicidal teens do not request help nor reach out in any way [1, 2, 3]. Whereas depressed adults have the facilities to independently seek therapy, depressed teens not only feel an obligation to deal with these serious issues by themselves to avoid burdening those around them but also keep quiet due to their own shame. Furthermore, society is very poor at identifying those with mental health issues, evident in the fact that four out of five teenagers who have committed suicide have shown clear warning signs. To further stress the urgency of the issue, approximately 12 teen suicides happen every single day and 1.9 million teens in the United States have been diagnosed with depression [2].

Still, these findings do not change the fact that depression and mental illness are indeed best addressed through conversational and intimate mental health therapy. As opposed to the popular means of antidepressant medication, cognitive and conversational therapy not only engages in similar neural mechanisms but also better reduces the risk of relapse [4]. The central dilemma here between the social stigma of therapy for teens yet still obtaining this effective conversational treatment poses a severe threat, a threat exacerbated by the COVID-19 pandemic.

Background Research

To maximize accessibility and effectiveness, but also incorporate privacy, an artificially intelligent solution interfaced with a mobile application was chosen. With a rapidly advancing body of research and investment, artificial intelligence has demonstrated its capacity to perceive and classify human characteristics, including speech and behavioral cues, serving as a step-in for a diagnostician.

Facial Action Coding System (FACS). The FACS is a way to classify human expression based on distinct movements and the position of key features on the face. Used by psychologists, it is a common standard to analyze facial expressions. Different movements in the face such as a raised eyebrow, closed eyes, or puckered lips can correspond to different action units. When put

together, a combination of action units can explain the current user emotion. The FACS is highly computerized through the use of clinical machines to determine action units in a three-dimensional space; however, for this study, the units must not only be determined through a two-dimensional space but also accessible from a remote mobile application.

Neural Networks and Deep Transfer Learning. Neural Networks are a popular tool in the subset of artificial intelligence known as deep learning. Using layers of nodes connected via linear equations with weights and biases, these neural networks receive a vector on inputs and return an output. These networks are so popular because they are designed in a way to mimic the brain's function to make predictions and learn. Neural networks are trained on a dataset of input vectors and corresponding outputs. However, some models are trained with much larger sources of data and can be more accurate due to this than others. Thus, a popular and relatively new process known as transfer learning is used to incorporate the accuracy and mathematics behind models trained with large sources of data into new models with similar sources of data.

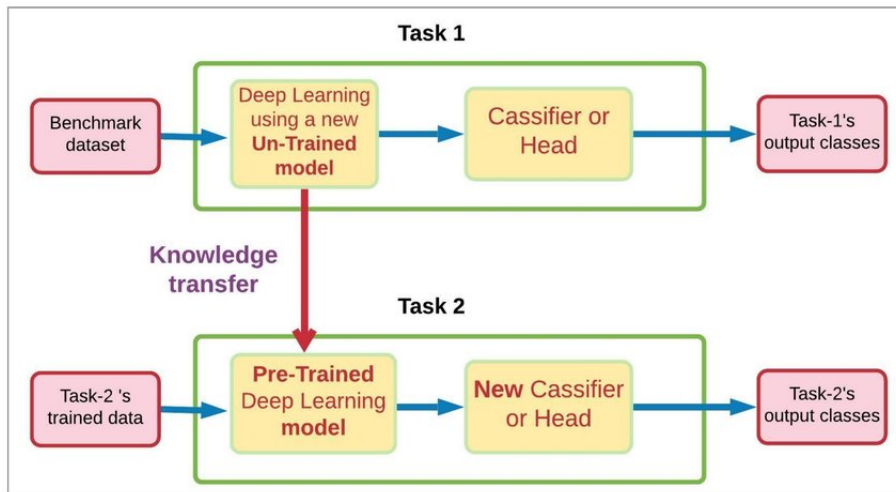


Figure A. Diagram of Transfer Learning through Neural Networks [6]

Long Short-Term Model (LSTM). The Long-Short Term Model is a type of neural network that is capable of handling a time-series of data. Intending to solve the outlined problem, analysis of a user video with sequential text, video, and audio is critical for an accurate diagnosis. Used extensively in computer science as a recurrent neural network, given a series of data, such as encodings for facial expression or audio characteristics over time and a target value, the LSTM can train on these values and be utilized in the future to classify other videos.

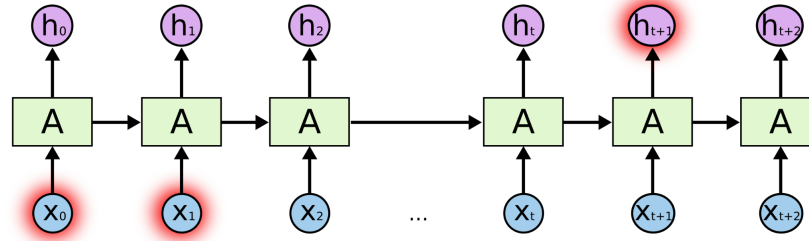


Figure B. Diagram of LSTM. Each node uses its output as input for the next node [6]

Hypothesis

The researcher hypothesized that if text-based deep learning neural networks were trained via transfer learning on substantial data, then these neural networks would be able to predict mental health disorder tendencies in teens with accuracy, precision, and recall of over 95%. Secondly, if LSTMs were trained on formulated data from the text-based neural networks combined with visual and aural characteristics per frame, the neural networks would be able to predict mental health disorder tendencies in teens with accuracy, precision, and recall of over 80%. Text, image, and audio inputs would be received from the user through a brief video from the researcher-made mobile application outlining their struggles. The networks would return a percentage likelihood that the user has a specific mental health disorder/issue based on the video, and the resulting percentages combined with scores on clinical behavioral questionnaires would be used to anonymously pair users on the application to allow them to discuss their issues. The key mental health issues targeted are suicidal tendency, depression, and anxiety. For clarification, accuracy is the number of correct predictions the model makes out of the total number of predictions. Recall is the number of correct positive predictions out of the total number of positive predictions. Precision is number of correct positive predictions out of the total number of actual positive labels.

Materials, Methods, and Procedures

Materials

Due to the software nature of the research, the necessary physical materials include a computer with a graphics processing unit capable of training deep learning models, Python 3.8 installed, and the modules Keras, NumPy, Pandas, and Tensorflow. However, tools and data from other sources were incorporated and listed below.

Distress-Analysis Interview Corpus (DAIC). The Distress Analysis Interview Corpus-Wizard of Oz Corpus was created by the Institute of Creative Technology at the University of Southern California [1, 5]. This database is a culmination of a series of interviews carried out at the university on patients with mental health issues. There are 189 40-minute recorded interviews in the corpus with 189 distinct patients, which were scanned for features such as tone, eye movement, posture, and facial expression, and asked to take the Patient Health Questionnaire 9. This data consists of a primarily important feature that uses a 68-point facial landmarking system that plots points on the face to determine key changes in action units and encode the expression of participants over time. Additionally, it includes a transcript of the interview with frames during which a certain sentence was said. Lastly, it includes key aural characteristics per frame (taken directly from the DAIC user guide): F0, VUV, NAQ, QOQ, H1H2, PSP, MDQ, peakSlope, Rd, Rd_conf, MCEP_0-24, HMPDM_0-24, HMPDD_0-12 [7, 8]. This time-series data was essential to the development of a computer vision algorithm that predicted depression tendency based on a face-camera video.

frame	timestamp	confidence	success	AU01_r	AU02_r	...
1	0	0.939744	0	0	0	...
2	0.0333333	0.944596	1	0	0	...
3,	0.0666667	0.947792	1	0	0.612177	...
4	0.1	0.948344	1	0	0.767003	...
5	0.133333	0.984772	1	0	0.515646	...

Table A. Sample of DAIC Action Unit data, AU## represents the presence of an action unit, ## continues until 28.

Reddit Mental Health Corpora (SMHD). To accurately predict depressive tendencies in teens via text-based methods, the most important tool to use is conversational data. The most accurate text-based classified mental health data on the internet is available via the Reddit

Self-Reported Depression Diagnosis Corpus (RSDD) and the Self-Reported Mental Health Diagnosis Corpus (SMHD) [9, 10]. These two corpora contain over 1.5 million text components. This corpus contains an abundance of negative text data to predict tendencies of the following mental health issues: Addiction, Alcoholism, Anxiety, ADHD, Autism, Bipolar Disorder, Borderline Personality Disorder, Depression, Loneliness, PTSD, Schizophrenia, and Suicide Watch. Additionally, it also contains an abundance of positive or neutral text data of the following categories: Fitness, Finance, Teaching, Parenting, and Law [12].

OpenFace. Used by the creators of the DAIC, OpenFace is a tool intended for computer vision and machine learning researchers, the computing community, and people interested in building interactive applications based on facial behavior analysis [5]. OpenFace is the first toolkit capable of facial landmark detection, head pose estimation, facial action unit recognition, and eye-gaze estimation with available source code for both running and training the models. The computer vision algorithms which represent the core of OpenFace demonstrate state-of-the-art results in all of the aforementioned tasks. Furthermore, our tool is capable of real-time performance and can run from a simple webcam without any specialist hardware. This tool was used to landmark faces and obtain facial action encodings in real-time to maintain consistency with the datasets.

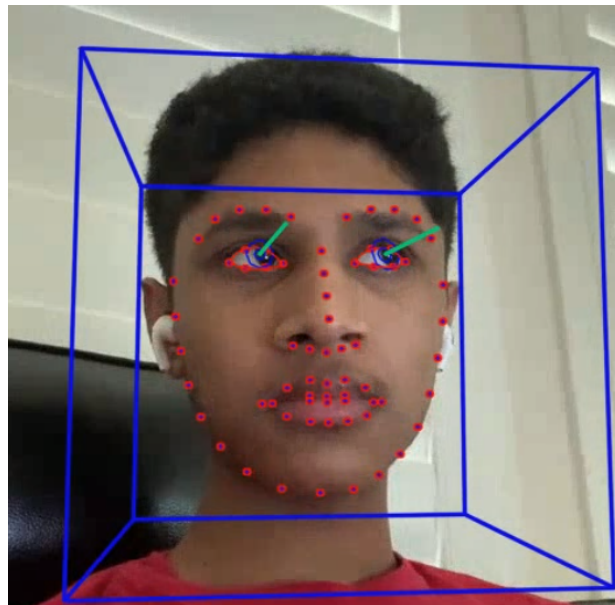


Figure C. OpenFace demonstration including 68-point landmarks and gaze.

Methods and Procedures

The development of the solution/experimental design has been split into several parts as it involves several diverse machine learning aspects due to the facet of multimodality. All deep learning models were trained via Keras, a Python deep learning framework used to build neural networks. Keras is also used to carry out the process of transfer learning.

Preprocessing of Datasets. Firstly, the SMHD was required to be reformatted to obtain equal quantities of training, test, and validation data. The current format of the corpus is a list of text content in a csv file per category. However, the data had to be transformed to be one file per explored mental health issue including several labeled text contents rather than several files of the same text content. To accomplish this, all data points from the depression, anxiety, and suicide watch folders were initially put into nine CSV files (a train, test, and validation file per depression, anxiety, and suicide watch) and labeled with a value of 1 to indicate the presence of the mental health issue. Next, a balanced mix of data was taken from the fitness, finance, teaching, parenting, and law folders and distributed throughout the same three CSV files, but labeled with a value of 0 to indicate a lack of any mental health issue.

Text	Value
I feel more and more sad and angry each and every day. I find it hard to keep carrying out the basic tasks in my life.	1
While teaching students, it's important to remember to keep things simple. Pretend like they don't know a single thing... because they probably don't!	0
Trading stocks regularly may be addicting, but it's important to maintain a balanced mind. Take a rest day every now and then. Don't let the profit dictate how you feel.	0
I'm sad beyond my expectations. Truly, I find no happiness in things. My once rose-colored glasses are a haunting black color. I know I need help, but I simply can't ask.	1
I'm beginning to feel numb from how numb I've been feeling for the past year. I	1

would pay good money to see myself in the mirror with a truly meaningful smile, but money can't always do miracles.	
---	--

Table B. Sample of preprocessed depression dataset.

Next, the DAIC data included distinct files for textual, visual, and aural characteristics (previously discussed in the description of the dataset) However, to properly use both modes of data in the same LSTM, the data tables were joined. The new data included one data vector per frame with sentence uttered, action unit encodings, and aural characteristics.

Identification of Existing Model for Transfer Learning. The first mode to analyze for mental health cues are text or speech content. Because text can be interpreted for emotion via its vocabulary/content and its sentiment, the existing model that needed to be incorporated was required to have the same characteristics. After analysis of several models on TensorFlow Hub, an open-source resource for machine learning models and papers, the model chosen was Version 2 of the nnlm-en-dim50, which is a token-based text embedding. This means that the model is meant to be embedded as a layer into a larger neural network and takes text inputs, separates them into tokens or words, and maps them to a unique vector in space with number coordinates. These numbers can be fed through the rest of the neural network similar to how a standard artificial neural network functions and outputs the percentage probability of demonstrating mental health issues. This model is incredibly reliable as it has been initialized on the Google News English 7B corpus with over 4 billion tokens in the token bank.

Training the Three Neural Networks. Once the model was chosen, it was trained on the three datasets created in the first portion of the procedure. The layers of the neural networks are as follows: a Sequential Layer, the Embedding Layer, 2 Dense Layers, 1 Dropout Layer, and 2 more Dense Layers. The chosen activation function is “relu”, the optimizer is “adam”, and the loss function is “Binary_Crossentropy”. These specific algorithms were chosen after fine-tuning the network to determine the optimal hyperparameters. Three models were trained (one per suicide watch, depression, and anxiety datasets). The models were then tuned based on test accuracy. Lastly, As mentioned in the hypothesis, these models were evaluated via the final validation data on accuracy, precision, and recall, each calculated from a confusion matrix.

Formulating the New Multimodal Datasets. Once all three text-based neural networks were trained, tested, and evaluated, they were ready to be used in the creation of the primary multimodal dataset. For each sentence captured in the transcript, it was run through each text-based neural network to return a score. Each score was then used in place of the sentence in each data vector under the input name “text_score” per each mental health issue. This text value was stitched to the preprocessed DAIC dataset to obtain the final multimodal data set. The target value for each data vector was the associated questionnaire score. For example, the PHQ-9 score greater than 10 was the positive target value for the depression model, the GAD-7 greater than 8 was the positive target value for the anxiety, model, and the PHQ-9 score with the last question being answered yes was the positive target value for the suicide watch model [11]. With text, video, and audio data included in a single data vector per each mental health issue dataset, the multimodal dataset was accomplished [13].

frame	text_score	AU01_r	AU02_r	F0	VUV	...
1	0.98	0	0	-9.3103	2.0175	...
2	0.98	0	0	-9.2498	2.0553	...
3,	0.98	0	0.612177	-8.8736	2.1421	...
4	0.77	0	0.767003,	-8.8355	2.2538	...
5	0.77	0	0.515646	-8.9991	2.0058	...

Table C. Sample of multimodal depression data, AU## represents the presence of an action unit, ## continues until 28.

Training the Three Long-Short Term Models. Once the new datasets were complete, an LSTM was trained for each of the three datasets. The layers of the neural networks are as follows: a Sequential Layer, an LSTM layer, and a Dense layer. The chosen activation function is “sigmoid”, the optimizer is “adam”, and the loss function is “Binary_Crossentropy”. These specific algorithms were chosen after fine-tuning the network to determine the optimal

hyperparameters [6]. The models were then tuned based on test accuracy. Lastly, As mentioned in the hypothesis, these models were evaluated via the final validation data on accuracy, precision, and recall, each calculated from a confusion matrix.

Hyperparameters, layers, train, and evaluate.

Incorporating Models into Mobile Application. While the bulk of the research was determining a private yet accessible and effective means for teens to receive an accurate diagnosis, the facet of conversational therapy and interfacing the technology through a mobile application required an extra portion of the procedure to make the results applicable to real-world use [8]. The structure discussed here is a four-level diagnostic process.

Firstly, the user inputs high-level data such as name, age, and sex. Secondly, the user responds to the PHQ-9 and GAD-7 to provide a more specific but still high-level evaluation of the mental profile. Thirdly, the user responds to a multiple-choice form to further narrow down the source of their issues. Several common stimuli including bullying, academics, relationships, and more are included as stimuli. These choices are used to provide the algorithm a lens to analyze the user's video. Moreover, the user video constitutes the fourth and the most specific level of this diagnostic process and collects all necessary data used to diagnose and makes an educated pairing of users.

The user transcript is generated via Google's Speech to Text API, which is a state-of-the-art model to transform speech into text, which the previously described models dissect [3]. Secondly, the Amazon Comprehend Keyword Extraction API is used to pull keywords from user-text and the Hugging Face Zero-Shot Classification model is used to compare these keywords to other keywords from other users to evaluate a pairing [9].

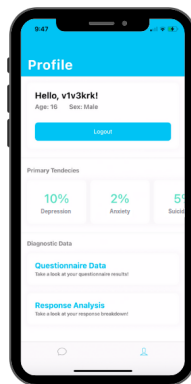


Figure D. Profile page of the iPonder mobile application.

Results

Text-Based Models

	Precision	Accuracy	Recall
Depression	99.28	98.99	99.3
Anxiety	98.33	98.12	98.9
Suicide Watch	99.1	98.9	99.29

Table D. Numerical results for evaluation metrics across models. All values are percentages.

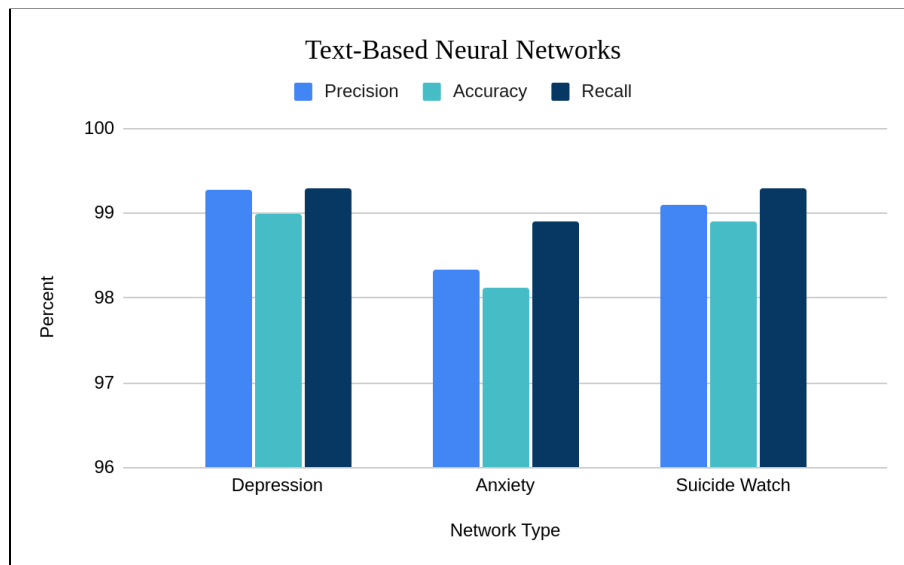


Figure E. Graphical results and comparisons for evaluation metrics across models.

Long Short-Term Models

	Precision	Accuracy	Recall
Depression	85.67	84.35	86.98
Anxiety	83.22	80.27	85.90
Suicide Watch	87.57	86.89	88.99

Table E. Numerical results for evaluation metrics across models. All values are percentages.

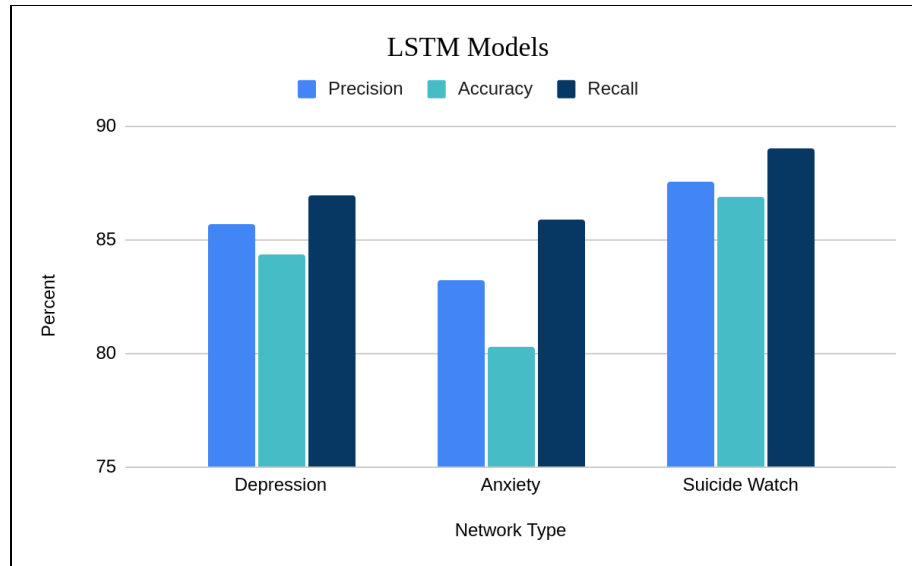


Figure F. Graphical results and comparisons for evaluation metrics across models.

Discussion

The researcher hypothesized that if text-based deep learning neural networks were trained via transfer learning on substantial data, then these neural networks would be able to predict mental health disorder tendencies in teens with accuracy, precision, and recall of over 95%. Secondly, if LSTMs were trained on formulated data from the text-based neural networks combined with visual and aural characteristics per frame, the neural networks would be able to predict mental health disorder tendencies in teens with accuracy, precision, and recall of over 80%.

Analysis

Based on the percentages in Table D and Table E, the hypothesis was correct. All text-based model percentages were above 95%, while all LSTM evaluation metrics were above 80%. These values are considered very reputable in the field of artificial intelligence, especially when the data involves video. Suicide watch and depression models generally functioned better than anxiety, but only by a relatively small margin. This may be due to the nature of behavior regarding anxiety or simply a dataset that allowed for less association to be drawn.

Moreover, for each model, the recall values were higher. In binary classification situations, this indicates that the models are strict. It means that if a person seems as if they are on the fence of having or not having a disorder, the model weighs them towards having it. This is

MULTIMODAL MENTAL HEALTH DIAGNOSIS

important as it ensures that no one with mental health issues is left out even if some users without mental issues are classified. For each model, the user ensures a high chance of being diagnosed correctly, which is a testament to the effectiveness of artificial intelligence.

Impact

The impact of the findings of the iPonder project is far-reaching. With modern technology in transfer learning and LSTMs, it reached metrics with models that have never been accomplished, which allows the integration of artificial intelligence into the field of behavioral science. Additionally, artificial intelligence seems to be the only viable and effective solution to the problems discussed initially. Teens with mental health disorders have a way to deal with their problems early on while avoiding the social stigma of visiting a therapist or talking through it with anyone else. This contributes to the conversational therapy portion of what already is researched in the field of behavioral science. Additionally, the cost of this project was under \$10, which means there is a substantial market for those who cannot afford or spend on mental health therapy. Ultimately, with iPonder, the community strives to rectify the ongoing mental health crisis in the youth to ensure that they have the right mental state to change the world on their own.

Limitations

The major limitation to this study is the lack of real participants aside from those studied in the existing datasets. The evaluation of the mobile application and its accompanying technologies researched would be a key result in concluding whether neural networks and LSTMs could accomplish the task of teen mental health diagnosis and therapy. Secondly, this study aimed to reproduce the circumstances in which the DAIC data was collected to the best of its ability; however, this may not always be the case for the user. In certain scenarios, OpenFace malfunctioned and didn't produce an accurate stream of aural, visual, and textual data, which may have skewed metrics slightly. Lastly, any differences in metrics between the models may have been a product of different real-world predictive capabilities of speech and expression or an imbalance of the data. This is still ambiguous.

Future Steps

As an ultimate priority, the first step will be to test the final product on several students and request feedback on how technology could be improved. As for statistical improvement, a future study may be completed in conditions similar to that of the DAIC study to increase the

reliability of the results. Secondly, further tests can be done with the existing models to determine the direct association between various parameters such as action units, text content, and audio characteristics to see which has the highest predictive power. This can be highly contributive to behavioral scientists trying to make a real diagnosis or helpful to determine which parameters are merely noise without association. Additionally, further increasing the accuracy of the models by integrating ensemble learning may be a fruitful endeavor. In this case, for storage optimization, rather than training 6 separate models, 3 were trained with ensemble learning to differentiate between tendencies of distinct disorders with higher predictive power [6].

Conclusion

In conclusion, the hypothesis initially formulated was proven correct through the results obtained from building 9 machine learning models: 3 text-based and 3 LSTMs. The study targeted depression, anxiety, and suicide watch, producing fruitful results, highly applicable to the field of behavioral science. The diagnostic process formulated and tested in this study is not clinically approved and is still an informal prediction with speculated conclusions drawn from it. With further research springboarded from this study, anonymous yet effective and approved conversational therapy for teens will become a reality.

Literature Cited

- [1] S. Alghowinem, R. Goecke, M. Wagner, J. Epps, M. Hyett, G. Parker, and M. Breakspear, "Multimodal depression detection: Fusion analysis of paralinguistic, head pose and eye gaze behaviors," *IEEE Transactions on Affective Computing*, vol. 9, pp. 478–490, 10 2018.
- [2] M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," *IEEE Xplore*, p. 858–862, 12 2017. [Online]. <https://ieeexplore.ieee.org/document/8389299>
- [3] S. Alghowinem, R. Goecke, M. Wagner, G. Parkerx, and M. Breakspear, "Head pose and movement analysis as an indicator of depression," *IEEE Xplore*, p. 283–288, 09 2013. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/6681444>
- [4] S. Alghowinem, R. Goecke, M. Wagner, G. Parker, and M. Breakspear, "Eye movement analysis for depression detection," *IEEE Xplore*, p. 4220–4224, 09 2013. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/6738869>
- [5] A. P. Shah, V. Vaibhav, V. Sharma, M. Al Ismail, J. Girard, and L.-P. Morency, "Multimodal behavioral markers exploring suicidal intent in social media videos," *2019 International Conference on Multimodal Interaction*, 10 2019.
- [6] J. F. Cohn, T. S. Kruez, I. Matthews, Y. Yang, M. H. Nguyen, M. T. Padilla, F. Zhou, and F. D. I. Torre, "Detecting depression from facial actions and vocal prosody," *IEEE Xplore*, p. 1–7, 09 2009. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/5349358>
- [7] Y. Yang, C. Fairbairn, and J. F. Cohn, "Detecting depression severity from vocal prosody," *IEEE Transactions on Affective Computing*, vol. 4, pp. 142–150, 04 2013.
- [8] J. R. Williamson, E. Godoy, M. Cha, A. Schwarzentruer, P. Khorrami, Y. Gwon, H.-T. Kung, C. Dagli, and T. F. Quatieri, "Detecting depression using vocal, facial and semantic communication cues," *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge - AVEC '16*, 2016.
- [9] M. Trotzek, S. Koitka, and C. M. Friedrich, "Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, p. 588–601, 03 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/8580405>

- [10] S. G. Burdisso, M. Errecalde, and M. Montes-y Gómez, “A text classification framework for simple and effective early depression detection over social media streams,” *Expert Systems with Applications*, vol. 133, pp. 182–197, 11 2019.
- [11] C. Karmen, R. C. Hsiung, and T. Wetter, “Screening internet forum participants for depression symptoms by assembling and enhancing multiple nlp methods,” *Computer Methods and Programs in Biomedicine*, vol. 120, p. 27–36, 06 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0169260715000620>
- [12] S. C. Guntuku, D. B. Yaden, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt, “Detecting depression and mental illness on social media: an integrative review,” *Current Opinion in Behavioral Sciences*, vol. 18, pp. 43–49, 12 2017.
- [13] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, “Depression detection from social network data using machine learning techniques,” *Health Information Science and Systems*, vol. 6, 08 2018.