Journal of Computer and Communications, 2022, 10, 113-130
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Review on Psychology Research Based on Artificial Intelligence Methodologies

Rushit Dave, Kyle Sargeant, Monika Vanamala, Naeem Seliya

Computer Science Department, University of Wisconsin (Eau Claire), Eau Claire, USA
Email: rushitdave05@gmail.com

How to cite this paper: Dave, R., Sargeant, K., Vanamala, M. and Seliya, N. (2022) Review on Psychology Research Based on Artificial Intelligence Methodologies. *Journal of Computer and Communications*, **10**, 113-130.
https://doi.org/10.4236/jcc.2022.105008

Received: April 27, 2022
Accepted: May 28, 2022
Published: May 31, 2022

Abstract

The field of psychology and health care are often viewed as very different. One of the main differences between these two fields is the application of prediction. When it comes to physical health, the health care and the medical field can predict outcomes of an individual's health based on things like diet, exercise, and genetic makeup. Psychology is a little bit different; psychologists are great at predicting behavioral changes in patients, but the prediction of an individual's mental health is not up to par with the prediction of physical health in the medical field for several factors. The scope of this study is to highlight how artificial intelligence has been applied to the field of psychology. In particular, how machine learning and deep learning technologies have been used to predict developmental risks of mental health disorders and risk of suicidal/self-injurious behaviors, as well as how artificial intelligence can be used to detect levels of depression.

Keywords

Psychology, Mental Illness, Depression, Suicide, Self-Injury

1. Introduction

Artificial Intelligence (AI) is a popular topic among computer science studies and research. AI is constantly being implemented into everyday life including medical practices, transportation, economics, and many more. The core of AI is machine learning. Machine learning applies different algorithms to learn a specific task. These tasks include making predictions, making classifications, developing images, and much more. Deep learning is a subset of machine learning that achieves similar tasks but with a more complex structure. These aspects of AI have been put to many great uses such as diagnoses in the medical fields, simulations in mathematics and physics, image classifications in biology and chemistry, and countless more. However, one field that has not been around as long as the previously mentioned field has yet to fully implement the power of these AI techniques, psychology. Psychology as science seeks to study and identify how individuals' behaviors relate to their cognitive and emotional processing [1]. Some researchers believe that unfortunately, a majority of psychologists only focus on explaining behaviors. Yarkoni and Westfall believe that since this has become commonplace in the field, prediction of future behavior is uncommon or overlooked [2].

Recently psychologist researchers have started experimenting with AI to predict and classify in many areas of study: from quantifying levels of pain from brain scans [3], applying machine learning techniques to further understand personality [4], and detecting human needs in critical events [5], to predicting problematic social media use [6], and future alcohol abuse [7]. Researchers have even analyzed how to make these AI models better specifically for the psychology field [8] [9]. There are many ways psychologists have started applying AI and machine learning to important topics.

One of the most important topics psychologists deal with today is the topic of mental health and mental illnesses. The most common mental illnesses and disorders treated by psychologists are major depressive disorder (MDD), anxiety, post-traumatic stress disorder (PTSD), schizophrenia, and many others. Treatments for these ailments often take the form of different kinds of therapies or even medication when paired with a psychiatrist. As such, heterogeneity in these diseases is something psychologists have looked at by using machine learning techniques to further understand [10]. Mental illness is on the rise and something that millions of individuals have experienced and among these individuals, depression and anxiety are the most common.

This paper will review the recent studies that have applied AI methodologies to aid the field of psychology. The applications of AI in this review will include using AI and machine learning concepts to aid in the diagnosis and prognosis of mental illnesses and disorders, detecting levels of depression, and predicting the risk of suicidal behavior and self-injury.

2. Literature Review

2.1. Machine Learning Applications for PTSD

Papini *et al.* [11] and Karstoft *et al.* [12] both examine the efficacy of applying machine learning to predict PTSD development after admission to an emergency room or hospitalization. Papini and colleagues aimed to increase the accuracy of a previous work's attempt to predict PTSD by using an ensemble machine earning approach. Papini and researchers collected data from 271 patients who have been admitted to an emergency room. Multiple physical predictive variables were collected including, pulse, length of stay, level of consciousness, and injury severity among others. Psychological predictive variables were collected in addition such as the history of a mood or anxiety disorder, current mental health, any current PTSD symptoms, and others. PTSD screening was then completed at three, six, and 12 months after emergency room admittance. After preprocessing the collected data, 41 predictive features remained. The machine learning model the researchers used was extreme gradient boost (XGBoost) which consists of multiple decision trees. Decision trees are a common machine learning algorithm that runs testing data through a series of yes or no questions that are derived from the training data. The model was used to predict whether an individual would show positive PTSD symptoms (PC-PTSD score ≥ 3) or negative PTSD symptoms (PC-PTSD score < 3), represented as PTSD + and PTSD –, respectively. The accuracy of the model was determined by an area under the curve (AUC) accuracy score. **Table 1** shows the performance for the researcher's XGBoost model as well as two benchmark prediction comparisons. One benchmark, "Hospital features" used normal data collected at hospitals to make predictions. The second benchmark, "PTSD severity at hospital only" used logistic regression on only the most important predictive feature. Karstoft and her colleagues took a slightly different approach. In their study [12], data was collected from 957 trauma survivors. Among this data, 68 predictive features were filtered by how important they are in predicting PTSD development. Researchers then evaluated the accuracy of the prediction using a machine learning approach known as support vector machines (SVM). SVM is a supervised learning method, meaning training data is required to improve the model. SVM models are commonly used for classification and outlier detection using clustering. The AUC results show a promising 75% accuracy. The study mentions that this is a good starting point and that use of other data sets is important to determining other important predictive variables when it comes to predicting PTSD.

Goldberg *et al.* [13] conducted a study on the automatic detection of student attentiveness in the classroom. Although it is not mentioned explicitly in the paper, this research could be used as a stepping-stone to diagnose attention deficit-hyperactivity disorder (AD-HD). The researchers had three main research questions, are visual indicators of engagement and disengagement correlated to a student's learning of the material, is it possible to predict this level of comprehension of the student using machine learning and visible engagement/ disengagement and does student attentiveness affect surrounding students' attention. A total of 52 students from a university in Germany volunteered to participate in the study. The students were informed on the topic of the material that would be discussed in a 90-minute lecture. They were then asked to complete questionnaires before the lectures collecting data about the students' background as well as the individuals' learning prerequisites. Students were then recorded during the lecture using three cameras at differing angles in the classroom. After the lecture, participants completed a knowledge test of the topic they were lectured on. The machine learning model used variables like gaze, head-pose, and facial expressions to predict the class's overall attention. The researchers used a deep learning approach to analyze individual students' attentiveness and related the measures to other students to get an overall attentiveness score of a class using the OpenFace library. However, since the cameras were unable to capture variables from all the students, measurements were taken from a subsample size of 30 students. Although there is some support in the idea that machine learning can be used to predict engagement levels, researchers suggested that this is a good starting point and that a larger sample size should be used in the future. Table 2 shows the prediction of the post-test variables created by the machine learning model and the manual ratings from the teacher. The model consisted of two sets of ratings, one with analysis of head pose and gaze, as well as one with analysis of head pose, gaze and surrounding students' attentiveness ("sync").

Table 1. Performance metrics with bootstrapped 95% confidence interval.

Performance Metric	Full model (N features = 41	PTSD Severity at hospital only	Hospital features (N features = 22)
Area under curve	0.85 [0.83, 0.86]	0.78 [0.76, 0.80]	0.75 [0.73, 0.76]
Sensitivity	0.69 [0.66, 0.72]	0.69 [0.66, 0.72]	0.57 [0.53, 0.61]
Specificity	0.83 [0.80, 0.85]	0.87 [0.86, 0.88]	0.76 [0.73, 0.79]
Positive predictive value	0.65 [0.62, 0.69]	0.63 [0.61, 0.66]	0.53 [0.50, 0.56]
Negative predictive value	0.86 [0.84, 0.87]	0.80 [0.79, 0.81]	0.80 [0.79, 0.81]
Overall accuracy	0.78 [0.77, 0.80]	0.78 [0.77, 0.80]	0.70 [0.68, 0.72]

Note: The model's metrics reported above reflect performance when "PTSD +" was greater than or equal to 50%.

Table 2. Prediciton of post-test variables.

Rating System	Estimated rating (head pose + gaze)				
	b	SE	P	R^2	F
Knowledge Test	1.37	8.09	0.867	0.001	0.03
Cognitive Engagement	7.74	3.82	0.053	0.136	4.1
Involvement	13.94	6.05	0.03	0.17	5.31*
Situational interest	5.64	5.17	0.286	0.044	1.19
Rating System	Estimated rating (head pose + gaze) + sync				
	b	SE	P	R^2	F
Knowledge Test	1.14	2.98	0.704	0.006	0.15
Cognitive Engagement	3.03	1.4	0.04	0.152	4.67*
Involvement	5.37	2.42	0.023	0.184	5.87*
Situational interest	2.63	7.88	0.175	0.07	1.95
Rating System	Manual rating				
	b	SE	P	R^2	F
Knowledge Test	0.63	0.86	0.468	0.02	0.54
Cognitive Engagement	1.38	0.35	0	0.38	15.91***
Involvement	2.33	0.54	0	0.414	18.34***
Situational interest	1.54	0.48	0.003	0.286	10.42**

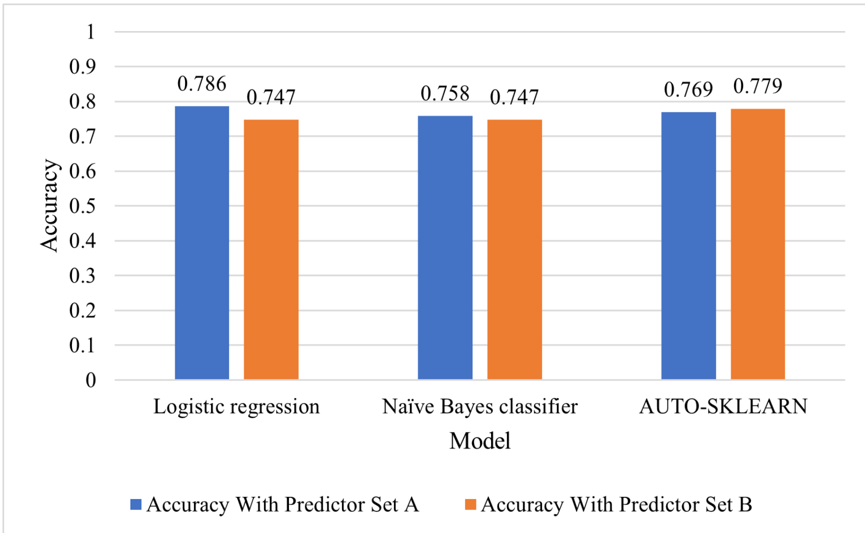
Note: *p < 0.05; **p < 0.01; ***p < 0.001.

2.2. Machine Learning Applications for Anxiety Disorders

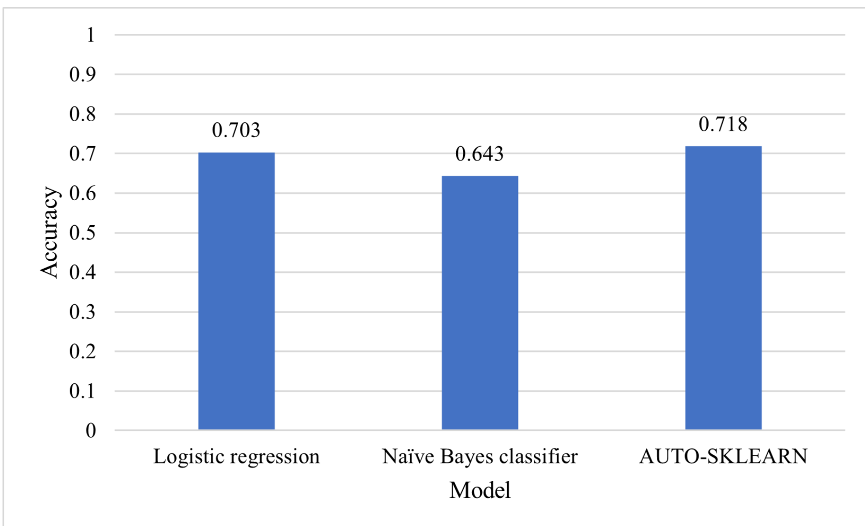
van Eden *et al.* [14] and Bokma *et al.* [15] explored the application of machine learning on anxiety disorders. One of the main goals of van Eden and colleagues' research was to compare performance between three machine learning models: logistic regression, naïve Bayesian classifier, and Auto-sklearn. A naïve Bayesian classifier is a machine learning algorithm that assumes no relation between other present features. Auto-sklearn is a tool kit that will automatically manage the hyperparameters of the model. These methods were used to predict DSM-IV-TR psychiatric diagnoses at year 2, 4, 6, and 9 follow up. The variables used as predictors were classic demographic, clinical diagnoses, and any participant's self-reported depression and/or anxiety. The models were trained on 50% of the sample data and tested on the other 50%. The results of the study show that Auto-sklearn outperformed the other two models at every follow-up period. Researchers also found that for the most part, the three models average accuracy increased as the follow-ups progressed: year 2 (0.668), year 4 (0.714), year 6 (0.744), year 9 (0.742). **Figure 1(a)** shows the three models' predictions at year 2. The models predicted both binary outcomes (presence of psychiatric illness or not) as well as categorical outcomes (healthy, presence of a mood disorder, presence of an anxiety disorder and presence of comorbid disorder) **Figure 1(b)** Bokma *et al.* [15] applied machine learning to predict the recovery of anxiety disorders. Researchers gathered patients with a diagnosed anxiety disorder (panic disorder, generalized anxiety disorder, agoraphobia, or social phobia). Many variables were collected from the participants including clinical variables, psychological variables, and demographic variables. In total, 569 variables were collected from the patients. In the study, random forest classification was chosen to make the prediction of recovery and performed relatively well amongst the different predictor variables. The accuracy was 61.7% using clinical predictive variables, 61.0% for psychological predictive variables, 53.1% for socio-demographic variables, 52.7% for biological variables, 50.2% for lifestyle variables, and 62.4% for the combination of all predictive variables. Researchers reported their findings were only moderately successful as too many false positives and negatives occurred meaning the implementation of the current model was unlikely in clinical practice.

2.3. Machine Learning Applications for OCD Disorders

Hoexter *et al.* [16] explored how machine learning and neuroimaging can be used to predict the severity of obsessive-compulsive disorder (OCD) in an individual. The main goal of the study was to look to see if the grey matter in the brain can be identified as a predictor for the severity of OCD in an individual. Researchers collected data from 37 adult patients with diagnosed but untreated OCD. The participants were then put through a clinical assessment to quantify the severity of their OCD symptoms. This was completed using three differing assessments (SCID-I, Y-BOCS, and DY-BOCS). The patients then took an MRI scan. The MRIs were processed using a software package called Freesurfer. This software took the MRIs and automatically segmented the cortical and subcortical regions of the participants' MRIs. Once the cortical and subcortical were labeled on the MRIs, they were passed into a support vector regression model (SVR). SVR is slightly different than SVM in that it is a regression algorithm. 16 brain regions in the segmented MRIs were used as predictive variables for the severity of OCD. The model's accuracy when predicting severity scores related to the DY-BOCS severity measure was 0.49. For the model's accuracy when predicting severity related to Y-BOCS the result was 0.44.



(a)



(b)

Figure 1. The accuracy of the Logistic regression, Naive Bayes classifier, and AUTO-SKLEARN using the data from the two-week follow-up. (a) The accuracies of the binary outcomes using two different predictor sets. (b) The overall accuracies of the three models predicting categorical outcomes: Healthy, Mood disorder, Anxiety disorder, and Co-morbidity.

Bracher-Smith *et al.* [17] conducted a review of machine learning models used in the genetic prediction of psychiatric disorders. Bracher-Smith and colleges compiled over 63 papers. Of these 63 papers, 77 models were selected for further investigation. Researchers looked at seven studies involving schizophrenia, five involving bipolar disorder, 3 on autism, and 1 on anorexia. In terms of the most common model used in these studies, SVM and neural networks were among the most commonly used. Neural networks are also used for a variety of the same reasons. The power of neural networks comes from the number of nodes and the weights associated with them. One primary difference is that neural networks can be both supervised and unsupervised. After the composition of the models, researchers expressed that there were a few studies that had a high risk of bias in their models. Bias refers to the type of input data that is used to train a model in supervised learning. Models that are given input data that is biased to one outcome will likely reduce accuracy for your testing data. Researchers reported a wide range of variance in each model's accuracy depending on what psychiatric illness was involved. The best performing models were: XGBoost with schizophrenia (0.86), neural network with bipolar disorder (0.77) and autism (0.74), and for anorexia LASSO and SVM tied (≈0.69) as showed in **Figure 2**.

Tate *et al.* [18] used machine learning to predict the future development of mental health problems in mid-adolescence. In addition to predicting mental illness development researchers in this study were also curious to see if modern machine learning approaches would outperform traditional logistic regression. Researchers collected 474 predictors from parents of a total of 7638 children. Five different models were used to make these predictions, random forest, XGBoost, logistic regression, and SVM. The study found that SVM and random forest performed the best with about 74% accuracy. In the study, it was mentioned that 74% percent accuracy is not high enough to be used as a clinical practice but is a good start. Kaur and Sharma [19] explore previous studies of psychological disorders that used supervised machine learning techniques and/or nature-inspired computing techniques. Nature-inspired computing is a class of algorithms that are based around natural occurrences in the world including neural networks. The research goes over studies involving stress, depression, autism, anxiety, AD-HD, Alzheimer's, Parkinson's, insomnia, schizophrenia, and dementia. The most important part of this review of studies is the comparison of accuracies with and without the implementation of feature selection machine learning approaches. The compilation of accuracies shows that in most studies, models that use feature extraction outperform the models that do not.

Like the previously mentioned studies, Dwyer *et al.* [9] applied machine learning to a wide variety of mental health illnesses. However, this study included one of the most common mental illnesses, MDD. The study explains the importance of using machine learning as a tool of diagnosis. The study explains how a researcher was able to apply machine learning to decrease the misdiagnosis of bipolar disorder from 75% to only 31%. Dwyer and colleagues also touch

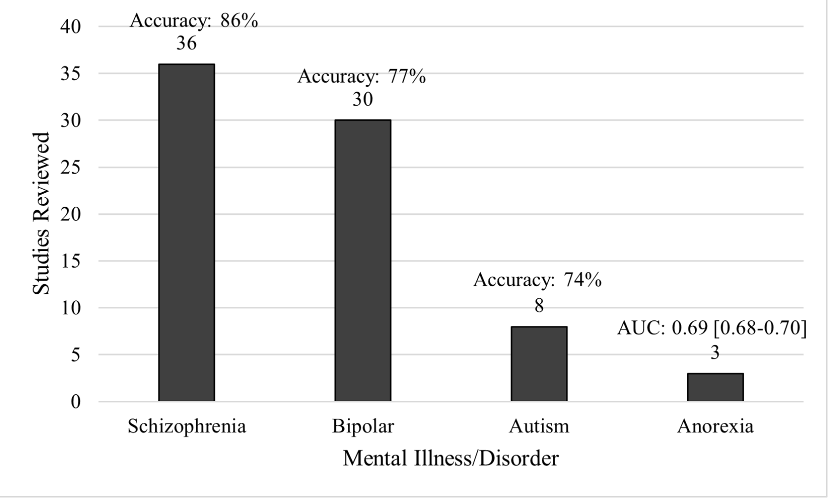


Figure 2. Shows the number of studies reviewed in the study conducted by Bracher-Smith et al. The highest reported accuracy for models involving Schizophrenia, Bipolar, and Autism were 86%, 77%, and 74% respectively. The highest reported AUC for Anorexia was: 0.69 [0.68 - 0.70].

upon the use of machine learning in the prognosis and treatments of mental illnesses. They mention that a big issue with the current approach of prognosis and treatments is that they are much generalized. Currently, treatments are based on generalized symptoms and often result in many instances of changing medication. Machine learning techniques have been tested to recommend different treatment approaches. However, these did not use biological data and were restricted by computational function. Today, machine learning models have been paired with large sample sizes to predict responses to antidepressants.

2.4. Machine Learning Applications for Depression

Na *et al.* [20] used machine learning to predict the development of depression amongst differing data. Na and colleagues used a nationwide dataset to predict the future development of depression among individuals living in a community. A total of 6,588 individuals were selected, 521 of which were labeled with having depression. Researchers used a random forest model to produce the predictions. The prediction accuracy was reported as a promising 87% with "satisfaction for leisure", "familial relationship", and "social relationship" variables having the most impact on the model's predictions. Although the study only used demographic predictive variables, the inclusion of biological or psychological variables may increase the accuracy of the model. On an alternative note, relating to prediction, Nelson *et al.* [21] used depression as a predictive variable in their research. The main goal of the study was to predict the development of psychosis in patients with high-risk or recent-onset depression. 668 patients and controls were selected from multiple countries in Europe. Researchers implemented a learning model called NeuroMiner that was trained on differing predictors of psychosis transition from three sources. The prediction of psychosis development among the datasets was reported with 75.7% accuracy.

Symptoms of depression vary across individuals, making diagnosis very important. Many studies have approached the diagnosis of depression using machine learning. Guo *et al.* [22] explore the detection of depression using a neural network approach. Researchers developed both a network that classifies static two-dimensional face images 2D-SADN as well as a network that classifies three-dimensional geometric patterns of faces 3D-DGDN. The main difference in the 3D-DGDN model is it detects motion and depth using a Kinect camera. Researchers collected the highest accuracy when combining both the two-dimensional network and the three-dimensional network (between 76% and 77%). **Figure 3** shows the structure of this hybrid approach.

Mumtaz *et al.* [23] also use visual data to diagnose depression using EEG scans. Researchers in this study tested to see if models could discriminate between healthy patients and depressed patients based on EEG data. EEGs are used to monitor electrical activity in the brain. 34 patients with MDD and 30 healthy patients were passed into three classification models: SVM, naïve Bayes, and logistic regression. Researchers reported that the SVM model had the highest accuracy in discrimination of MDD patients and healthy patients (98% accuracy). Similarly, Priya *et al.* [24] applied traditional machine learning models to detect depression as well as anxiety and stress. Collected data was passed into five models, decision tree, random forest tree (RFT), naïve Bayes, SVM, and K-nearest neighbor (KNN). KNN is a common classifier that clusters like data based on Euclidian distance. Researchers of this study found that naïve Bayes performed the best in detecting depression, resulting in an accuracy of 85.5%. Lastly, Morales *et al.* [25] surveyed interesting approaches to detect depression using facial recognition. Researchers mentioned one indicator of depression is through the visual cues in facial features. In one surveyed study, this indicator was measured and detected using Facial Action Coding System (FACS). Using FACS they found depression could be predicted by a downward angle of gaze, less intense smile, shorter duration of smiles, longer self-touches, and fidgeting

Kumar *et al.* [26] apply machine learning to depression in a slightly different fashion than diagnosis. Kumar and colleagues' research includes using eight machine learning models to predict the severity of depression as well as anxiety and stress. Data points were collected from participants who had answered questions on the Depression Anxiety Stress Scale 43 (DASS42). This questionnaire scored participants on depression, anxiety, and stress using a 4-point scale. Once completed, total scores represented severity. Depression severity was categorized as the following: 0 - 9 (Normal), 10 - 13 (Mild), 14 - 20 (Moderate), 21 - 27 (Severe), and 28+ (Extremely Severe). The eight models used fall into four types: Bayes classification, KNN, neural network, and Tree-based classification. An additional hybrid category was created by the researchers using a combination of the random forest model and the k-star model. As depicted in **Figure 4**, the hybrid approach did increase the accuracy of severity classification for both models, but the radial basis function network (RBFN) outperformed every model approach with a depression severity accuracy of 96.03%.

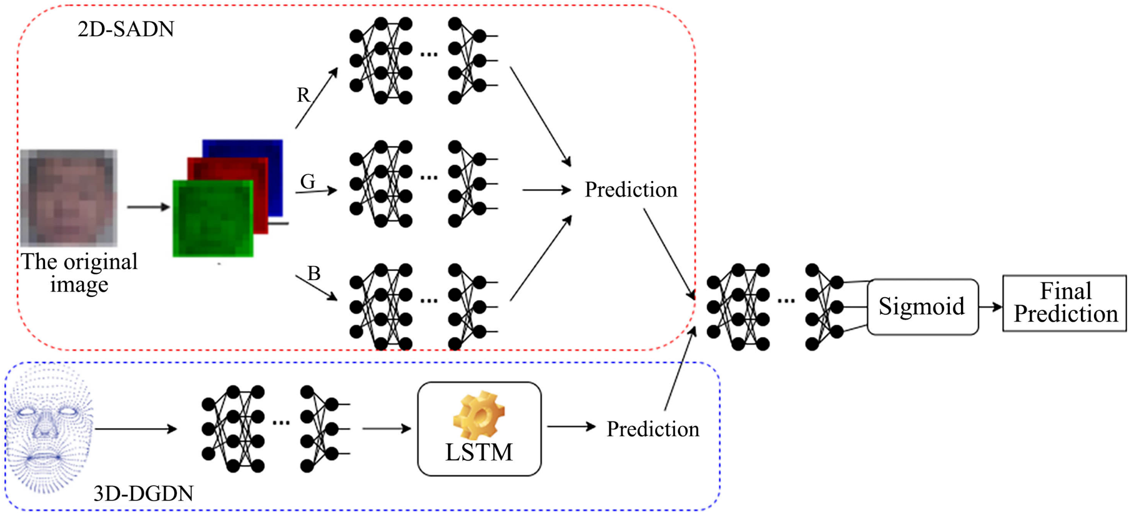


Figure 3. The general structure of the combination of the 2D-static appearance deep network (2D-SADN) and the 3D-dynamic geometry deep network (3D-DGDN). The 2D image is separated into three RGB channels and passed into a neural network to make a prediction. The 3D face points are also passed into a neural network followed by long short-term memory (LSTM) architecture, outputting a prediction. The two predictions then pass through concatenation and fully connected layers before going through a final sigmoid function.

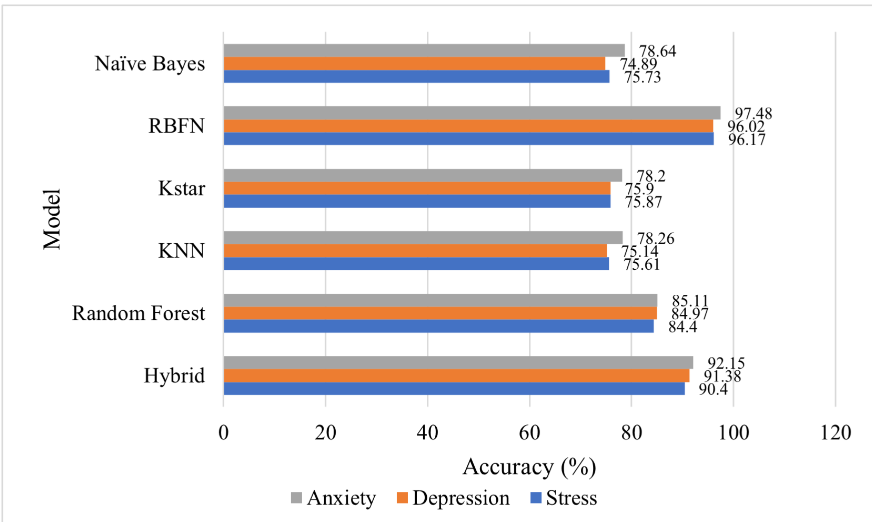


Figure 4. The accuracy of six different models each making classifications on severity of anxiety, depression, and stress.

Machine learning is also a powerful approach for predicting suicidal behavior, and many of these approaches are analyzed in Cox's *et al.* [27] survey paper. One of the most interesting ways machine learning is applied to this prediction is highlighted in the research paper by Cohen *et al.* [28]. Researchers in the study applied natural language processing (NLP) and machine learning in adolescent therapy sessions to predict suicide risk. NLP is a branch of AI that focuses on allowing computers to analyze/understand human voice and vocabulary. Researchers worked with mental health professionals to collect data from 60 students over 267 interviews/sessions. The mental health professionals provided an initial measurement of suicide risk based on an initial meeting where several emotional questions were asked. Patients with no risk were labeled as control patients and those who did show risks were labeled as case-patients. During therapy sessions, the adolescent's responses were recorded using an app called MHSAFE. Researchers pre-processed the long strings of texts paired with the diagnosis from the mental health professional before passing them into 3 machine learning models: SVM, logistic regression, and XGBoost. The study mentioned that previous works showed promising results using SVM. However, in the current study, they found that XGBoost performed the best (AUC: 0.78) resulting in feasible predictions of suicide risk using NLP. Similarly, Walsh *et al.* [29] predict suicide risk using a larger data set. Data was collected from a repository at Vanderbilt University Medical Center. Training data in this study included 5543 patients that had some indication of suicide attempt or self-injury on their record. Eight predictors were made using a random forest model. Predictions ranged from 7 days prior to a suicide attempt to 720 days prior to a suicide attempt. Results are very positive as all eight predictions on the testing data had an AUC accuracy equal to or above 80% (7 days: 84%, 14 days: 83%, 30 days: 82%, 60 days: 82%, 90 days: 81%, 180 days: 81%, 365 days: 83%, and 720 days: 80%). Edgcomb *et al.* [30] take an alternate approach with the prediction of readmission to psychiatric hospitals for suicidal and self-injurious behaviors. Longitudinal electronic health records were collected from the UCLA and Research Data Repository. Patients' records were only collected if they were 18+ years old and were diagnosed with a depressive disorder, bipolar disorder, or schizophrenia/schizoaffective disorder as well as having two or more hospitalizations. Each patient's data contained at least records of one year prior and one year after hospitalizations. Medical diagnoses were categorized by the International Classification of Diseases (ICD) ICD-9 and ICD-10. The ML model used was the Classification and Regression Tree algorithm (CART). The model was implemented using equal weights and all analyses were created using 10-fold cross-validation. The model was able to predict readmission to hospitals for suicide-related behavior and self-injury with an accuracy of 86%. The most important predictors in the dataset were: history of suicide attempt or self-harm, medical comorbidity, in-hospital mortality score, age, number of medical hospitalizations in the previous year, alcohol use disorder, and bipolar disorder.

As noted above, self-harm can be a relatively good predictor of suicide risk. Xu *et al.* [31] leveraged deep learning to predict self-harm risks instead of suicide risks. Researchers collected data from 2323 patients with IDC-9-CM codes: E950-E959, meaning they have been admitted to a hospital due to self-injurious behavior. The sample also contained 46,460 control samples. The model the researchers used was a patient embedding method called Diagnosis to Vector (Dx2vec). This method consists of calculation of comorbidity of two diseases, max pooling, and finally feeding data into a Long Short-Term Memory (LSTM) deep learning model. Max pooling is a common practice in deep learning to prevent the over-fitting of a model. Over-fitting is an issue where a model's performance is 100% accurate for its training set. This is an issue because the model will have learned a bias for the training set and will not be able to make accurate predictions on the test data. Researchers split up the sample into 80% training and 20% testing. Using the Dx2vec based deep neural network (DNN) there was a 72% accuracy in identifying patients who are at risk of self-injury in the next year. Researchers in the paper mention that compared to other regression-based approaches, DNN approaches tend to be superior. Similarly, Fox *et al.* [32] applied multiple models with varying complexity to predict non-suicidal self-injury (NSSI). Fox and colleagues used a sample of 1021 high-risk self-injurious and/or suicidal individuals who responded to questions assessing a wide range of variables related to NSSI and self-injurious thoughts and behaviors. They used this sample to test 3 different model types, each being more complex than the previous. The first "simple" model was tested for accuracy in prediction using traditional risk factors of NSSI and logistic regression. The second "more complex" model took the same data and added the implementation of multiple logistic regression analysis. In the third model, they implemented random forest classification identified as the "most complex" model. As an additional test, they decided to remove the strongest predictor for NSSI from the data to see how prediction accuracy changed then. As expected, added complexity has increased prediction accuracy and decreased false positives and false negatives. **Table 3** shows the accuracy of each model at three different baselines. It was identified that self-cutting episodes in the month prior to baseline were the best predictor. Among these predictors, other good predictors are suicidal thoughts and behaviors, psychopathology, self-disgust, agitation, and other relevant clinical measures.

Table 3. Random forest and multiple logistic regression model performance measures.

Model	AUC [95% CI]	Precision	Recall
Logistic Regression			
T2	0.56 [0.52, 0.59]	0.26	0.56
T3	0.56 [0.53. 0.59]	0.42	0.61
T4	0.56 [0.53, 0.59]	0.49	0.57
Multiple Logistic Regression			
T2	0.72 [0.69, 0.76]	0.43	0.71
T3	0.70 [0.67, 0.73]	0.57	0.72
T4	0.70 [0.68, 0.73]	0.63	0.71
Random Forest			
T2	0.87 [0.84, 0.90]	0.94	0.76
T3	0.89 [0.87, 0.92]	.83	0.09
T4	0.90 [0.88, 0.92]	0.86	0.09

Note: T2 = 3 days after baseline; T3 = 14 days after baseline; T4 = 28 days after baseline.

3. Discussion and Analysis

This current study reviews and analyzes other studies that lend themselves to the advancement of the field of psychology using AI approaches. Among these approaches, machine learning and deep learning models are the most common tools used in the collection of studies above. The researchers applied these tools to answer questions as well as predict diagnoses and prognoses related to a variety of mental disorders, depression being the main focus, as well as self-injurious and suicidal behavior.

Although machine learning and deep learning are a part of the hierarchy of AI, there are some subtle differences between the two. Machine learning is a method that applies mathematical algorithms to input data to understand relationships in hopes of performing tasks without human intervention. Many of the studies reviewed above apply a variety of machine learning algorithms to conduct their research. SVM and RF were the most used models in the collection of studies above. The most notable studies using SVM were the studies produced by Mumtaz *et al.* [23] and Tate *et al.* [18]. The work of Mumtaz *et al.* is an important study because it gives proof to the idea that EEG images can be used to detect MDD in patients. The results found in this study could eventually be built upon to diagnose MDD and potentially other mental disorders using EEG images. Mumtaz and colleagues mentioned that adaptations of their work could be done by using fMRI images, better isolation of confounding variables like lingering effects of medication, and use of a larger sample size. The work of Tate *et al.* is also notable as a good foundation for future research. The idea of predicting a child's future development of mental health problems in mid-adolescence is undoubtedly only a job machine learning has the power to do. Although it was mentioned in the paper that 74% accuracy is not enough to be put into clinical practice, the study has plenty of potential to be improved especially in the selection of the predictors of a child. A significant study conducted by Kumar *et al.* [26] showed the importance of random forests in the classification of depression stress and anxiety. The random forest model in this study performed very well on its own with an accuracy of 84%-85% however the most notable part of the study is when researchers combined random forest with the k-star resulting in a near 7% increased accuracy. The combination of machine learning models could be used to increase the accuracy/results of many studies mentioned in this review and something that should be leveraged more.

Although deep learning is a subsection of machine learning the difference comes in the overall structure of the two. Deep learning models are often structured similarly to human neurons in the brain. The structure consists of multiple layers of perceptions that are like neurons in the brain. These perceptions apply equations and biases to input data from layer to layer. The significance of deep learning in the field of psychology is highlighted in the work done by Guo *et al.* [22]. The detection of depression among two-dimensional and three-dimensional images/data is a quite unique approach to the diagnosis of MDD. The use of

2D-SADN and 3D-DGDN were interesting approaches to the detection of MDD when convolutional neural networks are more likely a better tool for facial and image detection.

4. Limitations and Future Strategies

Studies reviewed in this literature review each stated limitations to their attempts in applying AI techniques to the field of psychology. Some of the most common limitations of these studies are the absence of larger sample sizes [4] [8] [16] [21] [23], structural flaws in models, and data-related issues. Having a large sample size for models to train/validate on is very important to the performance of the models used. Large sample sizes limit the possibility and the effects of unwanted biases in the model while also improving precision and accuracy. More challenges arose for some researchers as they were creating the structure for the models they used. This is especially important with neural networks. Savci *et al.* [6] expressed their concerns about only implementing one hidden layer in the neural network they were using. Implementation of additional hidden layers would allow the neural network to understand more complex relationships between training/testing data and their respective labels. Finally, data collection and manipulation seemed to be listed in many studies' limitations. Training a model on data requires a close to even representation of classifiers to decrease biases. This is a problem Kumar *et al.* [26] and Priya *et al.* [24] ran into when using data from the DASS21 and DASS42 databases. Both datasets were reported as imbalanced therefore interfering with a more honest accuracy score of the models tested.

In their current state, a majority of the studies reviewed need a significant improvement in the performance metrics in order for these approaches to be practical in professional practices. However, challenges to applying these practices in the field of psychology will not be easily solved with increased performance metrics. The acceptability in the medical-social community is also a challenge that needs to be addressed before the implementation of these models. Society is highly concerned with how their personal data is being used and secured, proving a significant challenge to provide models with accurate or enough data to be effective [33]. In reality, the social acceptance of AI being used in a field like psychology would be much higher if clients had insight into how their data was being collected and how it would help them in the long run. Another aspect to consider is the practical acceptance of AI in psychology. As it stands there are some minor applications of AI to psychology that most seem fine with, such as a bot that can guide a user through the process of cognitive-behavioral therapy. However, if patients solely relied on AI to make a diagnosis, there may be some speculation. As a result, the situational acceptability would greatly benefit if the AI methodologies were used as tools by psychologist. As such, the reviewed models above would likely perform the best if implemented as a tool alongside a trained psychologist [34].

The studies reviewed above all have limitations that could be improved upon. Some are as simple as collecting more data to train the models, but some are more complicated. AI models need a large supply of data to train and test from. The more data that is collected the more opportunities the models must understand complex relationships in order to accurately predict or classify aspects. In most cases, machine learning models will perform the best-using 75% of the data set as training data and the rest as testing data. Selection of what type of model you use is also an important task for researchers. Some models/structures perform the best when they are selected for specific reasons, therefore research on when and why models are better for these reasons is an important step. Structure implementation in deep learning models specifically is also a factor that can influence the accuracy of models. Neural networks have many aspects that are customizable such as the number of hidden layers or the number of epochs a model should run. These aspects are important when trying to avoid over-fitting of a model and false positives or false negatives.

5. Conclusion

AI is a very powerful tool that has the capability to understand complex relationships among variables that humans cannot. Machine learning and deep learning have been used to solve many research problems in the field of science in many ways. The scope of this study is the review the ways machine learning and deep learning have been used in the field of psychology. Psychology is a field of science that works closely with psychiatry to tackle mental health disorders that individuals all around the world suffer with. Reviewing these studies shows a promising future with applying AI techniques to diagnose and predict outcomes of these issues of mental disorders, self-injury, and suicide. The research reviewed in this study promotes a core foundation for the implementation of more advanced applications of machine learning and deep learning approaches in psychology. Many of the studies in this review do have some limitations but leave room for adaptation and future improvement.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] American Psychological Association (2019) What Do Practicing Psychologists Do? https://www.apa.org/topics/psychotherapy/about-psychologists

[2] Yarkoni, T. and Westfall, J. (2017) Choosing Prediction over Explanation in Psychology: Lessons from Machine Learning. *Perspectives on Psychological Science*, **12**, 1100-1122. https://doi.org/10.1177/1745691617693393

[3] Lee, J., Mawla, I., Kim, J., Loggia, M.L., Ortiz, A., Jung, C., Chan, S.-T., Gerber, J., Schmithorst, V.J., Edwards, R.R., Wasan, A.D., Berna, C., Kong, J., Kaptchuk, T.J., Gollub, R.L., Rosen, B.R. and Napadow, V. (2018) Machine Learning-Based Predic-

tion of Clinical Pain Using Multimodal Neuroimaging and Autonomic Metrics. *Pain*, **160**, 550-560. https://doi.org/10.1097/j.pain.0000000000001417

[4] Bleidorn, W. and Hopwood, C.J. (2018) Using Machine Learning to Advance Personality Assessment and Theory. *Personality and Social Psychology Review*, **23**, 190-203. https://doi.org/10.1177/1088868318772990

[5] Alharthi, R., Guthier, B. and El Saddik, A. (2018) Recognizing Human Needs during Critical Events Using Machine Learning Powered Psychology-Based Framework. *IEEE Access*, **6**, 58737-58753. https://doi.org/10.1109/ACCESS.2018.2874032

[6] Savci, M., Tekin, A. and Elhai, J.D. (2020) Prediction of Problematic Social Media Use (PSU) Using Machine Learning Approaches. *Current Psychology*, **41**, 2755-2764. https://doi.org/10.1007/s12144-020-00794-1

[7] Afzali, M.H., Sunderland, M., Stewart, S., Masse, B., Seguin, J., Newton, N., Teesson, M. and Conrod, P. (2018) Machine-Learning Prediction of Adolescent Alcohol Use: A Cross-Study, Cross-Cultural Validation. *Addiction*, **114**, 662-671. https://doi.org/10.1111/add.14504

[8] Jacobucci, R., Littlefield, A.K., Millner, A.J., Kleiman, E. and Steinley, D. (2020) Pairing Machine Learning and Clinical Psychology: How You Evaluate Predictive Performance Matters. https://doi.org/10.31234/osf.io/2yber

[9] Dwyer, D.B., Falkai, P. and Koutsouleris, N. (2018) Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology*, **14**, 91-118. https://doi.org/10.1146/annurev-clinpsy-032816-045037

[10] Schnack, H.G. (2019) Improving Individual Predictions: Machine Learning Approaches for Detecting and Attacking Heterogeneity in Schizophrenia (and Other Psychiatric Diseases). *Schizophrenia Research*, **214**, 34-42. https://doi.org/10.1016/j.schres.2017.10.023

[11] Papini, S., Pisner, D., Shumake, J., Powers, M.B., Beevers, C.G., Rainey, E.E., Smits, J.A.J. and Warren, A.M. (2018) Ensemble Machine Learning Prediction of Posttraumatic Stress Disorder Screening Status after Emergency Room Hospitalization. *Journal of Anxiety Disorders*, **60**, 35-42. https://doi.org/10.1016/j.janxdis.2018.10.004

[12] Karstoft, K.-I., Galatzer-Levy, I.R., Statnikov, A., Li, Z. and Shalev, A.Y. (2015) Bridging a Translational Gap: Using Machine Learning to Improve the Prediction of PTSD. *BMC Psychiatry*, **15**, Article No. 30. https://doi.org/10.1186/s12888-015-0399-8

[13] Goldberg, P., Sümer, Ö., Stürmer, K., Wagner, W., Göllner, R., Gerjets, P., Kasneci, E. and Trautwein, U. (2019) Attentive or Not? Toward a Machine Learning Approach to Assessing Students' Visible Engagement in Classroom Instruction. *Educational Psychology Review*, **33**, 27-49. https://doi.org/10.1007/s10648-019-09514-z

[14] van Eeden, W.A., Luo, C., van Hemert, A.M., Carlier, I.V.E., Penninx, B.W., Wardenaar, K.J., Hoos, H. and Giltay, E.J. (2021) Predicting the 9-Year Course of Mood and Anxiety Disorders with Automated Machine Learning: A Comparison between Auto-Sklearn, Naïve Bayes Classifier, and Traditional Logistic Regression. *Psychiatry Research*, **299**, Article ID: 113823. https://doi.org/10.1016/j.psychres.2021.113823

[15] Bokma, W.A., Zhutovsky, P., Giltay, E.J., Schoevers, R.A., Penninx, B.W.J.H., van Balkom, A.L.J.M., Batelaan, N.M. and van Wingen, G.A. (2020) Predicting the Naturalistic Course in Anxiety Disorders Using Clinical and Biological Markers: A Machine Learning Approach. *Psychological Medicine*, **52**, 57-67. https://doi.org/10.1017/S0033291720001658

[16] Hoexter, M.Q., Miguel, E.C., Diniz, J.B., Shavitt, R.G., Busatto, G.F. and Sato, J.R. (2013) Predicting Obsessive-Compulsive Disorder Severity Combining Neuroimaging and Machine Learning Methods. *Journal of Affective Disorders*, **150**, 1213-1216. https://doi.org/10.1016/j.jad.2013.05.041

[17] Bracher-Smith, M., Crawford, K. and Escott-Price, V. (2020) Machine Learning for Genetic Prediction of Psychiatric Disorders: A Systematic Review. *Molecular Psychiatry*, **26**, 70-79. https://doi.org/10.1038/s41380-020-0825-2

[18] Tate, A.E., McCabe, R.C., Larsson, H., Lundström, S., Lichtenstein, P. and Kuja-Halkola, R. (2020) Predicting Mental Health Problems in Adolescence Using Machine Learning Techniques. *PLoS ONE*, **15**, e0230389. https://doi.org/10.1371/journal.pone.0230389

[19] Kaur, P. and Sharma, M. (2019) Diagnosis of Human Psychological Disorders Using Supervised Learning and Nature-Inspired Computing Techniques: A Meta-Analysis. *Journal of Medical Systems*, **43**, Article No. 204. https://doi.org/10.1007/s10916-019-1341-2

[20] Na, K.-S., Cho, S.-E., Geem, Z.W. and Kim, Y.-K. (2020) Predicting Future Onset of Depression among Community Dwelling Adults in the Republic of Korea Using a Machine Learning Algorithm. *Neuroscience Letters*, **721**, Article ID: 134804. https://doi.org/10.1016/j.neulet.2020.134804

[21] Nelson, B., Yung, A.R. and McGorry, P.D. (2019) Importance of Variable Selection in Multimodal Prediction Models in Patients at Clinical High Risk for Psychosis and Recent-Onset Depression. *JAMA Psychiatry*, ***76***, 339-340. https://doi.org/10.1001/jamapsychiatry.2018.4234

[22] Guo, W., Yang, H., Liu, Z., Xu, Y. and Hu, B. (2021) Deep Neural Networks for Depression Recognition Based on 2D and 3D Facial Expressions under Emotional Stimulus Tasks. *Frontiers in Neuroscience*, **15**, Article ID: 609760. https://doi.org/10.3389/fnins.2021.609760

[23] Mumtaz, W., Ali, S.S., Yasin, M.A. and Malik, A.S. (2017) A Machine Learning Framework Involving EEG-Based Functional Connectivity to Diagnose Major Depressive Disorder (MDD). *Medical & Biological Engineering & Computing*, **56**, 233-246. https://doi.org/10.1007/s11517-017-1685-z

[24] Priya, A., Garg, S. and Tigga, N.P. (2020) Predicting Anxiety, Depression and Stress in Modern Life Using Machine Learning Algorithms. *Procedia Computer Science*, **167**, 1258-1267. https://doi.org/10.1016/j.procs.2020.03.442

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

[26] Kumar, P., Garg, S. and Garg, A. (2020) Assessment of Anxiety, Depression and Stress Using Machine Learning Models. *Procedia Computer Science*, **171**, 1989-1998. https://doi.org/10.1016/j.procs.2020.04.213

[27] Cox, C.R., Moscardini, E.H., Cohen, A.S. and Tucker, R.P. (2020) Machine Learning for Suicidology: A Practical Review of Exploratory and Hypothesis-Driven Approaches. *Clinical Psychology Review*, **82**, Article ID: 101940. https://doi.org/10.1016/j.cpr.2020.101940

[28] Cohen, J., Wright-Berryman, J., Rohlfs, L., Wright, D., Campbell, M., Gingrich, D., Santel, D. and Pestian, J. (2020) A Feasibility Study Using a Machine Learning Suicide Risk Prediction Model Based on Open-Ended Interview Language in Adolescent Therapy Sessions. *International Journal of Environmental Research and Public*

Health, **17**, 8187. https://doi.org/10.3390/ijerph17218187

[29] Walsh, C.G., Ribeiro, J.D. and Franklin, J.C. (2017) Predicting Risk of Suicide Attempts over Time through Machine Learning. *Clinical Psychological Science*, **5**, 457-469. https://doi.org/10.1177/2167702617691560

[30] Edgcomb, J.B., Shaddox, T., Hellemann, G. and Brooks, J.O. (2021) Predicting Suicidal Behavior and Self-Harm after General Hospitalization of Adults with Serious Mental Illness. *Journal of Psychiatric Research*, **136**, 515-521. https://doi.org/10.1016/j.jpsychires.2020.10.024

[31] Xu, Z., Zhang, Q. and Yip, P.S. (2020) Predicting Post-Discharge Self-Harm Incidents Using Disease Comorbidity Networks: A Retrospective Machine Learning Study. *Journal of Affective Disorders*, **277**, 402-409. https://doi.org/10.1016/j.jad.2020.08.044

[32] Fox, K.R., Huang, X., Linthicum, K.P., Wang, S.B., Franklin, J.C. and Ribeiro, J.D. (2019) Model Complexity Improves the Prediction of Nonsuicidal Self-Injury. *Journal of Consulting and Clinical Psychology*, **87**, 684-692. https://doi.org/10.1037/ccp0000421

[33] Blashki (2019, December 15) Would You Trust AI with Your Mental Health? Pursuit. https://pursuit.unimelb.edu.au/articles/would-you-trust-ai-with-your-mental-health

[34] Abrams (2021, November 1) The Promise and Challenges of AI. American Psychological Association, Washington DC. https://www.apa.org/monitor/2021/11/cover-artificial-intelligence