2020 19th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)

Predicting the level of generalized anxiety disorder of the coronavirus pandemic among college age students using artificial intelligence technology

Hana Alharthi, PhD.
Dept. of Health Information Management & Technology. College of Public Health
Imam Abdulrahman Bin Faisal University (IAU)
Dammam, Saudi Arabia
e-mail: Halharthi@iau.edu.sa

Abstract— Introduction: **Emerging reports indicate heightened anxiety among university students during the Corona pandemic. Implications of which can impact their academic performance. Artificial intelligence (AI) through machine learning can be used to predict which students are more susceptible to anxiety which can inform closer monitoring and early intervention. To date, there are no studies that have explored the efficacy of AI to predict anxiety among college students.** ***Objective:*** **to develop the best fit model to predict anxiety and to rank the most important factors affecting anxiety.** ***Method:*** **Data was collected using an online survey that included general information; Covid-19 stressors and (GAD-7). This scale categorizes level of anxiety to none, mild, moderate, and severe. We received 917 survey answers. Several machine learning classifiers were used to develop the best fit model to predict student level of anxiety.** ***Results:*** **the best performance based on AUC is AdaBoost (0.943) followed by neural network (0.936). Highest accuracy and F1 were for neural network (0.754) and (0.749) respectively, then neural network selected to be the best fit model. The three scoring methods revealed that the top three features that predicted anxiety to be gender; sufficient support from family and friends; and fixed family income.** ***Conclusion:*** **Neural network model can assist college counselors to predict which students are going through anxiety and revealed the top three features for heightened student anxiety to be gender, a support system, and family fixed income. This information can alter college councilors for early mental intervention.**

Keywordst; Machine learning; Artificial intelligence; Classifiers; anxiety ; AI ; ML.

I. INTRODUCTION

Anxiety profoundly affect quality of life of those afflicted with these conditions. College students are particularly vulnerable for anxiety as challenges of college life may be a trigger for these conditions. This is of especial concern during Covid-19 pandemic as emerging reports points to heightened anxiety among college age students. Implications of which can impact their academic performance. Artificial intelligence (AI) through machine learning can be used to predict which students are more susceptible to anxiety which can inform closer monitoring and early intervention. To date, there are no studies that have explored the efficacy of AI to predict anxiety among university students. We will assess this relationship among college students using an online survey. Data collected will include general information; Covid-19 stressors and the 7-item Generalized Anxiety Disorder Scale (GAD-7). This tool categorizes level of anxiety to none, mild, moderate, and severe. The orange data mining software will be employed to develop the best fit model to predict student level of anxiety.

II. LITERATURE REVIEW

The coronavirus pandemic, to date, has sickened millions of people and killed thousands world-wide [1]. It is a respiratory borne virus that causes severe acute respiratory distress with multi-organ involvement. Also, it has a higher human-to-human transmission rate and as such it can be widely spread among communities. Most of the world countries have ordered their citizens to stay home and business and schools were ordered to shut down.

Growing reports indicate that this public health response has significant mental health repercussions, including emotional imbalance, insecurity, sadness, vulnerability, and depression [2]. The causes of which can be underlying mental health disease vulnerability and/or fear or anxiety about contracting the disease or dealing with its financial consequences [3]. Currently, research efforts are underway to understand the effects of this disease on mental health.

College students reported high level of psychological discomfort [4]. As a result, this might affect their concertation and have adverse effect on their academic performance. A pre-Covid-19 study found that 35% of students reported symptoms of anxiety, mood swings and drug addiction [5]. This population seems to be particularly vulnerable to mental health and it is unclear how specifically Covid-19 has impacted this population. A worldwide study showed that 33.8% of medical students were showing symptoms of anxiety and the majority of which were in the Middle East and Asia [6]. Students from Switzerland have shown deterioration of their mental health since Covid-19 spread as they were afraid of physical isolation and lack of emotional support,[7]. In addition, 24% of 7000 medical students reported anxiety [4]. Although Universities often have student counselling center to support mental health needs of students and provide clinical help if need it, given the stigma associated with mental health, not all students seek help and consequently their mental health may deteriorate. Artificial intelligence (AI) technology can provide a valuable resource to identify most susceptible to mental health concerns including anxiety. Once identified,

2473-3636/20/$31.00 ©2020 IEEE
DOI 10.1109/DCABES50732.2020.00064

early intervention may avoid more severe cases of mental health decline. AI through machine learning techniques have been utilized by researchers to predict mental health disorders among different populations [8] where several machine learning classifiers have been applied. These classifiers are to predict which elderly inpatients are suffering from anxiety and random forest classifier showed the top AC of 91%. A study measured level of anxiety among internet seafarers using different machine learning classifiers showed that Catboost classifier was best fit model for prediction with accuracy score of 82.6% and 84% for precision [9]. Predicting anxiety among internet bloggers were also examined using their texts where several machine learning classifiers were used. CNN has provided the best prediction model with accuracy scoring 78% and recall scores 72%, [10]. Suicide prediction among one online community reported decision tree classifier to have the highest accuracy scoring 92% of suicide prediction [11]. Suicide prediction among twitter users showed that CNN scored a higher accuracy rate, 78%, outperforming support vector machine (SVM),[12]. Machine learning classifiers can predict which people are struggling with anxiety based on their reading habits and the model reported with best prediction was Naive Bayes, [13].

Despite this demonstrated usefulness of AI in mental health area, factors that affect prediction of severity of anxiety among college students is not clear. Here, we will survey college students attending the university of Imam Abdulrahman Bin Faisal and applied artificial intelligence tool based on the (GAD-7) using machine learning classifiers to predict level of anxiety among the student population. Our findings will inform best practices in AI to predict mental health among college students. In addition, we set to find the most important factors predicting generalized anxiety disorder.

III. METHODOLGY

Students answered the online survey voluntarily after reading the reason behind the study in google form. The survey was anonymous.

A. Data collection

Data of 917 instances were collected through the link distributed to the students in a local university through their coordinators by emails in addition to a student WhatsApp group between august 21 through august 27. The survey includes three sections. 1) basic information, 2) stressors factors that has been established through the literature and 3) the (GAD-7) a well-established and widely used screening tool for anxiety, (GAD-7).

TABLE 1. : STRESSORS RELATED TO COVID-19

Has the pandemic caused you to worry about the economic impacts?
Has the pandemic caused you to worry about academic delay?
I feel the pandemic affected my daily life
There was sufficient support from family and friends

General questions included gender, marital status, whether the student has children, whether the student lives at home, whether family income is fixed, household income (below or above 15000 Saudi Riyal (SR), and if someone they know got the disease. Table (1) list the stressors relayed to Covid-19. Several machine learning classifiers were used to develop the best fit model to predict anxiety level.

B. Data preprocessing

The collected data was imported into an excel datasheet to organize the data into a format that can be imported to the ML software. For the stressor's questions, each question asked for a response using psychometric methods extending from "strongly disagree (1) to "strongly agree" (5).

Each stressor scale was collapsed, such that responses of "1", "2", or "3" were grouped as "0" and responses of "4" or "5" were grouped as "1". For the (GAD-7), individual stressor answers were totaled to obtain anxiety category. Between zero and four has none, between five and nine mild, between ten and fourteen moderate, between fifteen and twenty one is severe. Dataset has thirteen features, one target variable with four values and 917 instances.

C. Classification models

To develop the machine learning models, several supervised machine learning classifiers in the Orange data mining software were used. Each classifier was trained and tested on the dataset provided. The instances were separated to 70:30 for to train and test respectively and K-Fold cross validation to avoid the models from overfitting and underfitting. Fig.1 demonstrates the workflow for anxiety dataset.

D. Model evaluaiton

Models were evaluated based on several matrices, which included area under the curve (AUC), accuracy, and F1-measure. These outcomes ensured that the best model for predicting level of anxiety among the university students can be identified

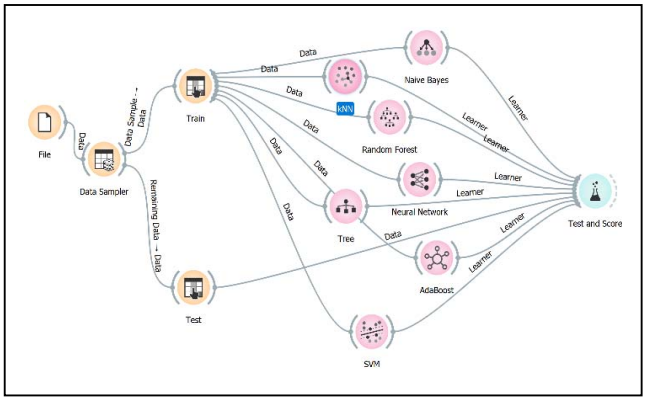


Figure 1. Orange workflow for anxiety dataset

IV. Results and Discussions

Seven ML classifiers were applied to classify the generalized anxiety disorder among college students during Covid-19 pandemic into different categories (none, mild, moderate, severe). Comparison of performance for the different models was analyzed based on the following metrics: Accuracy of prediction, area under the curve (AUC) curve and F1. Table 2 shows the best performance based on AUC is AdaBoost (0.943) followed by neural network (0.936). Highest accuracy and F1 were for neural network (0.754) and (0.749) respectively, then neural network selected to be the best fit model.

Network was selected as the suitable model to predict anxiety level. Different studies using machine learning to predict psychological disorders reported different results. Two studies one to assess anxiety, depression and stress and the other one to evaluate suicide intentions in social media reported that neural network outperformed other classifiers, which is consistent with our findings [14], [12]. One study that looked at predicting anxiety and depression among elderly patients showed random forest classifier produced the best model, [8]. In a study to predict anxiety, depression and stress in modern life naïve Bayes classifier outperformed other classifiers [15].

To examine which features were the most important to predict anxiety, three scoring methods were selected. Info. Gain, Gain ration and ReliefF. Gender, support from family and friends, and fixed family income to be the top features predicating generalized anxiety disorder, as shown in Fig 2. Therefore, psychological disorder can be identified by these attributes.

V. Conclusion

This research focused on applying several machine learning classifiers to predict the level of students' anxiety in Covid- 19 pandemics. Level of anxiety score was calculated based on the (GAD-7) which categorizes level of anxiety to none, mild, moderate, and severe.

TABLE 2: Classifiers Evaluation Matrix

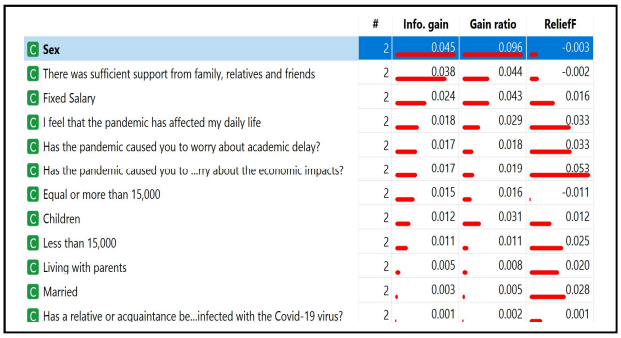
Evaluation Results

Model	AUC	CA	F1	Precision	Recall
kNN	0.758	0.448	0.426	0.493	0.448
Tree	0.859	0.596	0.587	0.600	0.596
SVM	0.834	0.765	0.760	0.776	0.765
Random Forest	0.839	0.590	0.574	0.595	0.590
Neural Network	0.936	0.754	0.749	0.769	0.754
Naive Bayes	0.660	0.437	0.420	0.433	0.437
AdaBoost	0.943	0.743	0.739	0.744	0.743

The neural network classifier produced a suitable model with AUC (0.936), accuracy (0.75) and F1(0.749). The result of using three scoring methods revealed that the top three predictors of anxiety were gender ; family/friendship support, and family income. The neural network model can assist college counselors to predict which students are going through anxiety which will facilitate early mental intervention and/or additional academic support.

Acknowledgment

I thank Dr. Irfan Ullah Khan (IAU) for informative discussions, Ms. Fatmah Alshakh (IAU) for assistance with data collection.



	#	Info. gain	Gain ratio	ReliefF
Sex	2	0.045	0.096	-0.003
There was sufficient support from family, relatives and friends	2	0.038	0.044	-0.002
Fixed Salary	2	0.024	0.043	0.016
I feel that the pandemic has affected my daily life	2	0.018	0.029	0.033
Has the pandemic caused you to worry about academic delay?	2	0.017	0.018	0.033
Has the pandemic caused you to ...ry about the economic impacts?	2	0.017	0.019	0.053
Equal or more than 15,000	2	0.015	0.016	-0.011
Children	2	0.012	0.031	0.012
Less than 15,000	2	0.011	0.011	0.025
Living with parents	2	0.005	0.008	0.020
Married	2	0.003	0.005	0.028
Has a relative or acquaintance be...infected with the Covid-19 virus?	2	0.001	0.002	0.001

Figure 2 . Top three features predicting anxiety

References

[1] "Worldometers," *Coronavirus statistics*, 2020. [Online]. Available: https://www.worldometers.info/coronavirus. [Accessed: 03-Jun-2020].

[2] J. Wang *et al.*, "Psychological status of Wuhan medical staff in fighting against COVID-19," *Med. J. Wuhan Univ.*, vol. 41, no. 4, pp. 547–550, 2020.

[3] J. Cui, "Origin and evolution of pathogenic coronaviruses," *Nat. Rev. Microbiol.*, vol. 17, no. March, pp. 181–192, 2019, doi: 10.1038/s41579-018-0118-9.

[4] W. Cao *et al.*, "The psychological impact of the COVID-19 epidemic on college students in China," *Psychiatry Res.*, vol. 287, no. March, p. 112934, 2020, doi: 10.1016/j.psychres.2020.112934.

[5] R. P. Auerbach *et al.*, "WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders," *J. Abnorm. Psychol.*, vol. 127, no. 7, pp. 623–68, 2018, doi: doi.org/10.1037/abn0000362.

[6] T. T. Quek *et al.*, "The Global Prevalence of Anxiety Among Medical Students : A Meta-Analysis," 2019.

[7] E. I. Fried, F. Papanikolaou, and S. Epskamp, "Mental Health and Social Contact During the COVID-19 Pandemic : An Ecological Momentary Assessment Study," 2020.

[8] A. Sau and I. Bhakta, "Predicting anxiety and depression in elderly patients using machine learning technology," vol. 4, pp. 238–243, 2017, doi: 10.1049/htl.2016.0096.

[9] A. Sau and I. Bhakta, "Informatics in Medicine Unlocked

Screening of anxiety and depression among seafarers using machine learning technology," *Informatics Med. Unlocked*, vol. 16, no. August, p. 100228, 2019, doi: 10.1016/j.imu.2019.100228.

[10] Y. Tyshchenko, "Depression and anxiety detection from blog posts data," *Inst. Comput. Sci. Univ. Tartu, Est.*, 2018.

[11] S. R. Braithwaite, C. Giraud-carrier, J. West, M. D. Barnes, C. Lee, and C. L. Hanson, "Validating Machine Learning Algorithms for Twitter Data Against Established Measures of Suicidality Corresponding Author :," vol. 3, pp. 1–10, 2016, doi: 10.2196/mental.4822.

[12] J. Du *et al.*, "Extracting psychiatric stressors for suicide from social media using deep learning," vol. 18, no. Suppl 2, 2018.

[13] Y. Hou, J. Xu, Y. Huang, and X. Ma, "A Big Data Application to Predict Depression in the University Based on the Reading Habits," no. Icsai, pp. 1085–1089, 2016.

[14] P. Kumar, S. Garg, and A. Garg, "Assessment of Anxiety, Depression and Stress using Machine Learning Models," *Procedia Comput. Sci.*, vol. 171, no. 2019, pp. 1989–1998, 2020, doi: 10.1016/j.procs.2020.04.213.

[15] A. Priya, S. Garg, and N. P. Tigga, "Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 1258–1267, 2020, doi: 10.1016/j.procs.2020.03.442.