

MACHINE LEARNING IN DIAGNOSIS OF MENTAL ILLNESS; A CASE STUDY OF UNIVERSITY OF ILORIN MEDICAL STUDENTS

Dere Abdulhameed Abiola



Introduction

Despite having similar or healthier profiles than peers at the start of medical school, medical students report higher levels of psychological discomfort than their same-age peers, according to a study [1].

Mental health illness refer to a wide range of mental health conditions, such as depression, stress, anxiety, and eating disorders. Although many studies have focused extensively on diagnosis and treatment, little attention has been given to preventing or alleviating mental illness among medical students using state of the art AI techniques [2].

High prevalence of mental illness and the need for proactive as well as effective mental health care, combined with recent advances in AI, has led to an increase in explorations [3,4] of how the field of machine learning (ML) can assist in the detection, diagnosis and treatment of mental health illnesses. Predicting which student will go on to develop mental health symptoms during the course of training is critical for early intervention and preventing future, severe negative outcomes. It is no news that medical undergraduates have higher levels of mental distress than the general population and their fellow peers [5] and seeing that these are the future doctors, it is important that more research be done to achieve a long lasting solution.

Therefore, with ML techniques, I aimed to I.) Develop a model that can predict mental health problems among medical students in the University of Ilorin II.) Investigate what factors could be more alarming in diagnosing these problems as they ensue.

Data Collection

This study was conducted at the Faculty of Clinical and Basic Medical Sciences, University of Ilorin for medical students ranging from 100 to 600 level between March to June 2021. We utilized Questionnaire for Assessment of Mental Health Literacy (QuaALiSMental) [6] and a self-constructed questionnaire which was carefully drafted by Consultant Physicians in the department of Psychiatry of the University of Ilorin Teaching Hospital. The questionnaires contained 49 attributes and over 600 participants were recorded. Questionnaires were administered to students physically.

Data Processing

- Handling of missing, duplicate and incorrect data was done with the use of Pandas library.
- Model was trained and built using the SciKit learn library.
- Outliers and noise in the model were detected using Pandas library and subsequently dropped.
- A consultant clinician certified the labels as mentally ill which includes but not limited to depression, severe anxiety, eating disorder, Attention Deficit Hyperactivity Disorder(ADHD).

Methodology

- The algorithms were trained on 80% (480) of the data and tested on the remaining 20% (120) through random shuffling.
- Model tuning was done with grid search and the best performing model was selected.
- Hyperparameter optimization with cross validation.
- An algorithm with the best mean f1 score was retained and trained on the most predictive variables.
- This model was used to identify medical students within the University with a higher likelihood of having a mental illness during the course of the training.

Results

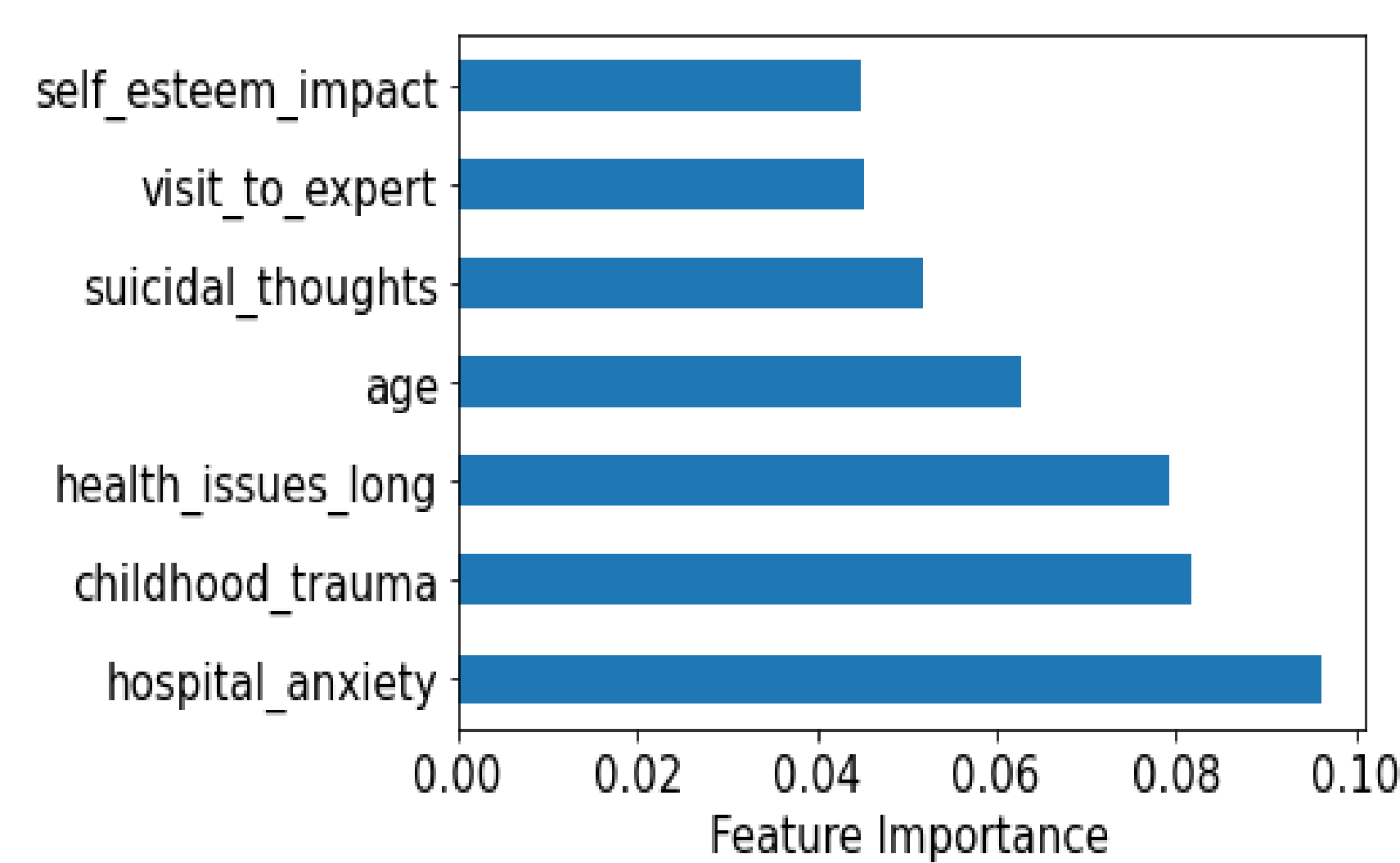


Chart 1. Feature Importance of the Random Forest Model

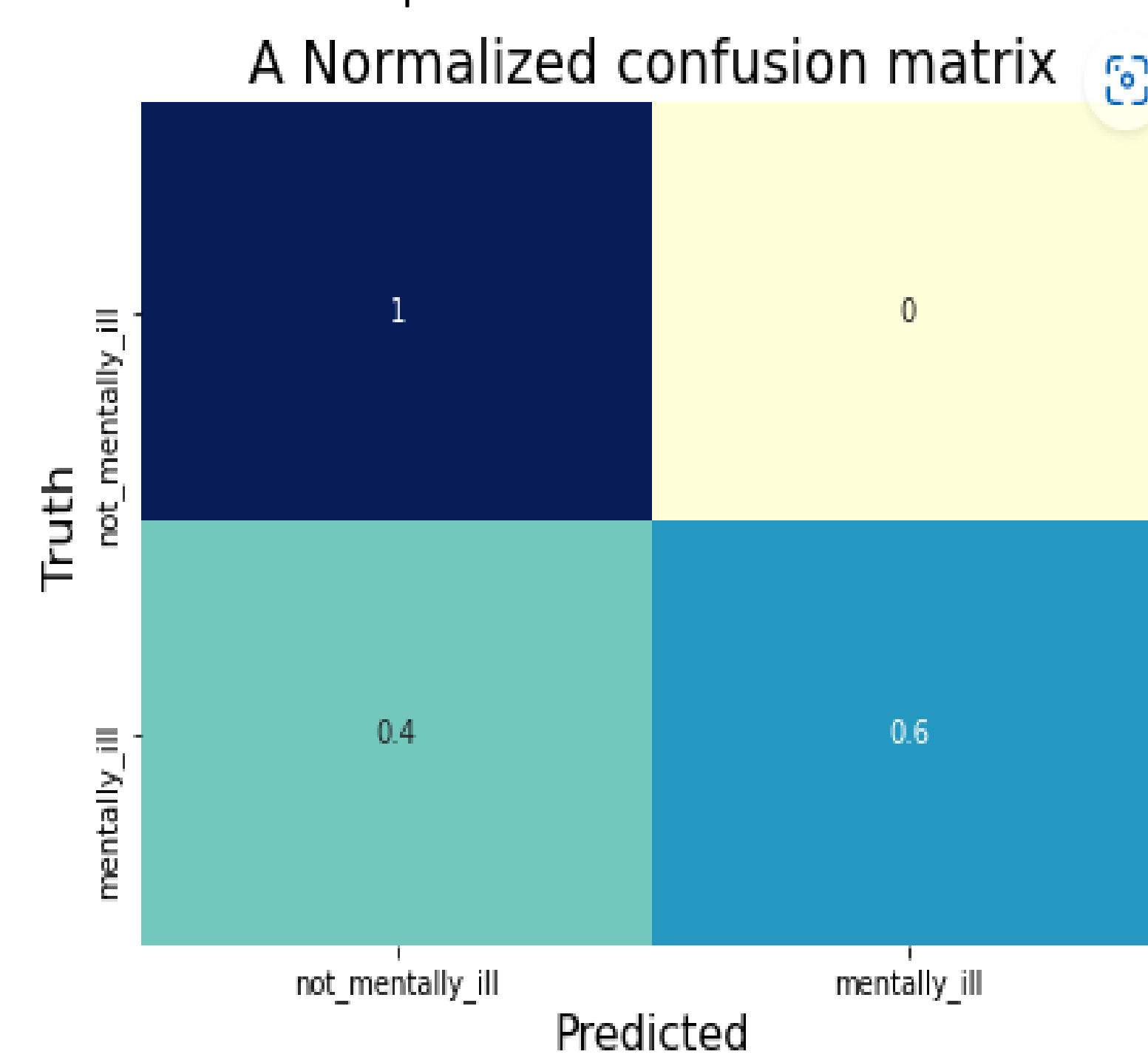


Chart 2. Confusion matrix of the Random Forest Model

Challenges

- Insufficient sample dataset.
- Lack of previous research studies on this topic within this locality.
- Data Labeling and feature imputing – it was hard to encode because of the fear of unnecessary weights.

Next Steps

- Gathering robust dataset for the entire university to get a more reliable model.
- More hyper parameter tuning to achieve further variance.
- Deployment on better frameworks.

Table 1. Performance Evaluation of the models

	Accuracy(%)	F1 score
Random Forest	89.0	0.77
XGBoost	82.0	0.75
SVM	79.1	0.67
Decision Trees	66.3	0.54

Discussion

Our analysis revealed no significant differences in the prevalence of depression and suicidal ideation between genders and no significant difference in the prevalence of depression between individuals of different ages (20 years and older or younger than 20 years). This is similar to a study conducted in China amongst medical students [7].

Hospital anxiety, childhood trauma, self esteem, suicidal ideation and repeat in medical school were all found to be significant as causative factors of mental illness amongst medical students. This is also taking into consideration other familial and environmental factors.

In another similar study conducted in Sweden [8] using ML techniques to predict mental disorder, random forest model achieved highest accuracy. This makes it evident that Machine Learning techniques can indeed be of great value when considering mental illnesses [8]. For this model to be highly optimal, students from their third year in college should be made to make use of the model and consequently attend counseling session.

Conclusion

The major goal of this study was to identify the significant academic and environmental factors associated with mental illness among medical students and using AI models as a means of getting preventive and predictive solution. One could argue that if a medical student has low self esteem, suicidal thoughts, had gone through childhood trauma, then such individual might be going through some form of mental illness ranging from depression to some others.

I created a model as a result that can take into account all of these variables and predict outcomes correctly. Overall, the top performing model might not be suitable for clinical use yet because of the little dataset used and the need for more tuning of the model to achieve more precision.

However it lays important groundwork for future models seeking to predict general mental health outcome especially when considering a larger and more inclusive data.

Contact

Name: Dere Abdulhameed Abiola
Affiliation: University of Ilorin
Email: daabiola3@gmail.com
Phone: +2348062688490

References

1. Hill MR, Goicochea S, Merlo LJ (2018). In their own words: stressors facing medical students in the millennial generation. Med Educ Online. 2018 Dec;23(1):1530558. doi: 10.1080/10872981.2018.1530558. PMID: 30286698; PMCID: 60286698
2. Anja Thieme, Danielle Belgrave, and Gavin Doherty. 2020. Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support the Development of Effective and Implementable ML Systems. ACM Trans. Comput.-Hum. Interact. 27, 5, Article 34 (October 2020), 53 pages. <https://doi.org/10.1145/3398069>
3. Thieme, Anja & Belgrave, Danielle & Doherty, Gavin. (2020). Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support the Development of Effective and Implementable ML Systems. ACM Transactions on Computer-Human Interaction. 27. 1-53. 10.1145/3398069.
4. Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., & Jeste, D. V. (2019). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. Current psychiatry reports, 21(11), 116. <https://doi.org/10.1007/s11920-019-1094-0>
5. Saravanan C, Wilks R. Medical students' experience of and reaction to stress: The role of depression and anxiety. The Scientific World Journal. 2014;2014
6. Zeng, Wen; Chen, Ruiqi; Wang, Xingyue; Zhang, Qin; Deng, Wei (2019). Prevalence of mental health problems among medical students in China: A meta-analysis. Medicine: May 2019 - Volume 98 - Issue 18 - p e15337 doi: 10.1097/MD.0000000000001533
7. Tate AE, McCabe RC, Larsson H, Lundstrom S, Lichtenstein P, Kuja-Halkola R (2020) Predicting mental health problems in adolescence using machine learning techniques. PLoS ONE 15(4): e0230389. <https://doi.org/>