|  |  |  |  |
| --- | --- | --- | --- |
| Title | Predicting Young Imposter Syndrome Using Ensemble Learning. | | |
| Author(s) Name | Khan, M. N. A.; Miah, M. S. U.; Shahjalal, M.; Sarwar, T. B.; and Rokon, M. S | | |
| Contact Email(s) | saef@aiub.edu | | |
| Published Journal Name | Complexity | | |
| Type of Publication | Journal | | |
| Volume |  | Issue |  |
| Publisher | Hindawi | | |
| Publication Date | Feb 2022 | | |
| ISSN |  | | |
| DOI | http://doi.org/10.1155/2022/8306473 | | |
| URL |  | | |
| Other Related Info. |  | | |
|  | | | |

|  |  |
| --- | --- |
| Abstract |  |
| *Background*. Imposter syndrome (IS), associated with self-doubt and fear despite clear accomplishments and competencies, is frequently detected in medical students and has a negative impact on their well-being. This study aimed to predict the students’ IS using the machine learning ensemble approach. *Methods*. This study was a cross-sectional design among medical students in Bangladesh. Data were collected from February to July 2020 through snowball sampling technique across medical colleges in Bangladesh. In this study, we employed three different machine learning techniques such as neural network, random forest, and ensemble learning to compare the accuracy of prediction of the IS. *Results*. In total, 500 students completed the questionnaire. We used the YIS scale to determine the presence of IS among medical students. The ensemble model has the highest accuracy of this predictive model, with 96.4%, while the individual accuracy of random forest and neural network is 93.5% and 96.3%, respectively. We used different performance matrices to compare the results of the models. Finally, we compared feature importance scores between neural network and random forest model. The top feature of the neural network model is Y7, and the top feature of the random forest model is Y2, which is second among the top features of the neural network model. *Conclusions*. Imposter syndrome is an emerging mental illness in Bangladesh and requires the immediate attention of researchers. For instance, in order to reduce the impact of IS, identifying key factors responsible for IS is an important step. Machine learning methods can be employed to identify the potential sources responsible for IS. Similarly, determining how each factor contributes to the IS condition among medical students could be a potential future direction. | |