

A Comparative Machine Learning Approach for Identification of the Pervasiveness of Obsessive Compulsive Disorder (OCD) among Sri Lankan Undergraduates

K.G.J Nimeshika* and Akila Brahmana

Department of Information Communication Technology, Faculty of Technology, University of Ruhuna, Kamburupitiya, Matara, Sri Lanka.

*jananimeshi@gmail.com

Abstract

Obsessive-compulsive disorder (OCD) is a common mental health problem described by obsessive thoughts and compulsive behaviors. Obsessions are intrusive, unpleasant thoughts, ideas, or desires that cause excruciating feelings. Compulsions are behaviors that an individual participates in to try to eliminate obsessions and reduce suffering. Everyone experiences obsessions and compulsions at some point. The majority of the Sri Lankan undergraduates have OCD but they are unaware of it. According to psychologists, contamination & washing, doubt and checking ordering, and arranging are the most common characteristics shown by students. If the student is aware that they have OCD, they can act to avoid it. These unpleasant thoughts and acts become more irritating as they interfere with students' learning. It also increases students' anxiety and depression. The study aims to develop a computerized model based on a data mining strategy to identify the pervasiveness of Obsessive-Compulsive Disorder (OCD) among Sri Lankan Undergraduates. This model identifies whether the students suffer from OCD as Low Moderate or High Level. Before the development of the model, a formal questionnaire was developed to examine the behaviors of undergraduates, and data was collected from Sri Lankan Undergraduates. This study used three unsupervised machine learning techniques, namely Kmeans, Agglomerative clustering, and DBSCAN clustering for the detection of patterns or relationships among datasets. To find the accuracy of the algorithms the Silhouette value, Davies Boulding value, and the Calinski-Harabasz value were used. Among those algorithms, the best accuracy is shown in the K-means algorithm. Hence, this algorithm has been selected to predict the pervasiveness of Obsessive-Compulsive Disorder (OCD) as low, moderate, or high among Sri Lankan Undergraduates.

Keywords: *Obsessive Compulsive Disorder, Sri Lankan Undergraduates, data mining strategy, unsupervised machine learning techniques.*