



INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Early Prediction of Dyslexia Using Machine Learning

A dissertation by

Ms. Udari Kaushalya Jayasinghe

Supervised by

Mr. Torin Wirasingha

Submitted in partial fulfilment of the requirements for the
BSc (Hons) Computer Science degree

Department of Computing

April 2019

©The copyright for this project and all its associated products resides with
Informatics Institute of Technology

Abstract

According to the statistical analysis of the International Dyslexia Association and The American Dyslexia Association, 1 in 10 people are dyslexic, as well as over 40 million American Adults are dyslexic, among them, only 2 million know they have dyslexia. Lack of awareness about the Dyslexia disability makes the situation more critical. Since it should be identified in its' early stages of the development before the situation gets critical.

The purpose of this study is to, predict dyslexia learning disability in its early developmental stage of preschool children to reduce series of impairments due to the dyslexia learning disability and improve school performances as well as the quality of their lives.

“Early Prediction of Dyslexia Using Machine Learning” is the idea of the research project. This is based on a questionnaire to predict dyslexia and make parents aware of the child's disability. The questionnaire, specially designed for dyslexia prediction in primary school children, followed by International standards of DSM-5 criteria (Diagnostic and Statistical Manual of Mental Disorders (DSM–5) dyslexia, 2013). Prediction will be done using KNN and Logistic Regression Algorithms to analyse answers and predict the Dyslexia disability. There are two prediction models were included in this application as it facilitates to identify dyslexia in two age categories which are age 6 to 7 and 8 to 9 categories. Two separate questionnaires were included in the proposed solution. A Mobile Application will be delivered which enables children to play with their parents to identification of Dyslexia.

the accuracy of the 6-7 model and the 8-9 model are in the range of 80% to 95%. The accuracy of the prediction model for both the age categories with KNN Algorithm was performed well according to the test results. The future enhancements and the limitation of this study also included in the conclusion chapter of this research study.

Subject Descriptors:

Machine Learning

Supervised Learning

Supervised Learning- Classification

Key words: Learning Disability, Dyslexia Prediction, Machine Learning, KNN Algorithm, DSM-5, Mobile Application