Predictive Analysis of Dropouts in Information Technology Higher Education

U. G. N. Kumari
169317F

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the partial fulfillment of the requirements of Degree of Master of Science in Information Technology

April 2019

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Name of Student

U.G.N.Kumari

Signature of Student

…….

Date:

Supervised by

Name of Supervisor:

Mr. S.Premaratne

Signature of Supervisor:

…….

Date:

Acknowledgment

I special gratitude pass to Mr. S. Premaratne, Senior Lecturer, Faculty of Information Technology, University of Moratuwa for his help, direction, and supervision was given to me all through my undertaking, making it a triumph.

Very special gratitude goes out to all down at Advanced Technological Institute, Labudawa for helping and providing the enormous support for data collection.

My genuine appreciation to Mrs. S.C.P. De Silva, Department of IT, Faculty of IT, [1]University Of Moratuwa for the animating talks, occupied hours we were cooperating before due dates.

I‘m grateful to all the staff members of the Faculty of Information Technology, University of Moratuwa.

To wrap things up I might want to thank my family for supporting me profoundly throughout writing this thesis and my life in general.

Abstract

At this point attention on educational data mining methods have impact highly on predicting academic performance as the increased higher education dropout rates especially in information technology education has received huge attention in recent years due to the quality of higher education has been a topic of debate for many years. There is a huge necessity of mining educational data system and exact hidden knowledge to understand the factors affecting student dropouts and to understand the patterns that can lead to predict student performance at the entrance and improve and monitor performance of students enrolling in Information technology higher education by building early warning indicators based on factors affecting to dropouts and manage students drop out from the higher education. Data mining strategies have been utilized to effectively extract new, conceivably imperative Knowledge and diverse data mining techniques, for example, association, classification, clustering, prediction, sequential patterns, and decision trees are being utilized by numerous sorts of research. Identification of relevant attributes which affect to dropouts in ICT higher education is a leading concern in the field of education data mining as there are no significant studies that can be applied to understand the complex-inter correlated and distinct factors affecting to dropouts in ICT higher education Hence, the research has been conducted to identify the complex-inter correlated and distinct factors affecting to dropouts in ICT higher education. It is hypothesized that an experimental methodology can be adapted to generate a database that includes relevant information for extracting knowledge. The raw data will be preprocessed in terms of filling up missing values, transforming values in one form into another and relevant attribute/ variable selection. Thereby select student records, which can be used for classification prediction model construction. In constructing a classifier model different classifying algorithms can be applied and in this study evaluation of different classification algorithms is done to identify the most accurate algorithm. Finally, a predictive analyzing model will be building for student profile analyzing using the identified algorithm. Then this classification model will be used in developing an application to predict the students' dropouts. The overall research will be designed using the WEKA data mining tool and using java WEKA library for developing the application.

Table of Contents

List of Figures

List of Tables