

Using Data Mining Techniques for Investigating of Performance in ICT at G.C.E Advanced Level

P.N.W.A.L.K.Premarathne

149225D

Faculty of Information Technology
University of Moratuwa

May 2017

Using Data Mining Techniques for Investigating of Performance in ICT at G.C.E Advanced Level

P.N.W.A.L.K.Premarathne

149225D

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

May 2017

Declaration

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 Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given.

Name of Student

Signature of Student

P.N.W.A.L.K.Premarathne

.....

Date :

Supervised by

Name of supervisor

Signature of supervisor

S.C Premarathne

.....

Senior Lecturer

Date :

Faculty of Information Technology

University of Moratuwa

Acknowledgement

A final thesis project like this is never the work of anyone alone. The contribution of many different people, in their different ways, have made this possible. I would like to extend my sincere thanks to all of them.

I would like to express my sincere gratitude to my supervisor Mr Saminda Premaratne the Senior Lecture of University of Moratuwa Sri Lanka, especially, Prof. Asoka Karunanada for the knowledge and assistance they provided throughout the research project. And I would like to thank the rest of my respective lecturers of the project committee, for their encouragement and insightful comments during project evaluation, which helped me to make necessary improvements.

Next I would like to express my special thanks to Mr Anura Pushpakumara the Commissioner General of Examination, Mrs Gayathri Abeygunasekara Commissioner of Examination Research and Development, Mr D.A Chandrasiri Commissioners of Examination Technology and other staff members of the Research and Development and the ICT branch for their support in providing data on exam results.

And I would like to thank all my friends who helped me to distribute my questionnaire among selected students as a sample.

I thank my colleagues who were behind me since last two years, sharing all good and bad moments in this hard time period.

Last but not least, I would like to thank my family for their unconditional support throughout my postgraduate degree. Your constant inspiration and guidance kept me focused and motivated. Thank you very much.

Abstract

The statistical analysis of the examination results plays a vital role in the field of education as the quality of the education is revealed by the evaluation process. The conclusions which follow the analysis, can be a feedback of the quality of examination papers or the supplementary of the students' feedback of teachers teaching process. Therefore it provides a summary of the prevailing standard of education which lead to educational research and reform. These reforms contribute to the positive changes applied to the examination system and to the learning and teaching process continuously.

In year 2011 Information and Communication Technology was introduced to the G.C.E Advanced Level as a main subject. The main objective of this research is to identify the problems related to the ICT subject of G.C.E Advanced Level Examination and provide suggestions for the teaching process of the education system. In this issue Association Rule technique of data mining is used to identify certain patterns in the examination results.

Table of Contents

Declaration.....	iii
Acknowledgement	iv
Abstract.....	v
1 Introduction.....	1
1.1 Prolegomena.....	1
1.2 Background and Motivation	2
1.3 Problem statement.....	6
1.4 Hypothesis.....	6
1.5 Objectives.....	6
1.5.1 Sub Objectives.....	6
1.6 Data Mining Based Approach	7
1.7 Structure of the thesis	8
1.8 Summary.....	8
2 Literature Review - Development and Challenges in Data Mining.....	9
2.1 Introduction	9
2.2 Early developments	9
2.3 Modern trends in data mining.....	12
2.4 Major Issues and Future challenges of data mining.....	13
2.5 Problem Definition.....	14
2.6 Summary.....	15
3 Technology - Data Mining Tools and Techniques. Principles of opinion mining.....	16
3.1 Introduction	16
3.2 Data Mining In Education Sector	17
3.3 Data Mining Techniques	17
3.3.1 Association.....	17
3.4 Summary.....	18
4 Approach.....	19
4.1 Introduction	19
4.2 Hypothesis.....	19
4.2.1 Input to the system.....	19
4.2.2 Output of the system	19

4.3	Process	19
4.4	Users of system	22
4.5	Features	22
4.6	Summary	22
5	Design	23
5.1	Introduction	23
5.2	System Design	23
5.3	Top level Design	27
5.4	User interface Design	27
5.5	Summary	27
6	Implementation	28
6.1	Introduction	28
6.2	Collection of Data.....	28
6.3	RapidMiner.....	28
6.4	Data Pre-Processing	28
6.5	Analysis by Questions on Exam paper	32
6.5.1	Grade obtained for Mathematics	34
6.5.2	Whether followed ICT or Not followed.....	38
6.5.3	Analysis by Gender.....	41
6.6	Analysis by the Stream sat for at the Advanced Level Exam	48
7	Evaluation	52
7.1	Introduction	52
7.2	Analysis by Questions on Exam paper	52
7.2.1	Evaluation on Grade obtained for Ordinary Level Mathematics.....	52
7.2.2	Evaluation on whether followed ICT for Ordinary Level or Not followed.	56
7.2.3	Evaluation on Gender.	60
7.3	Analysis by the Stream sat for the G.C.E Advanced Level Examination.....	63
7.3.1	Commerce stream.....	63
7.3.2	Arts stream.....	64
7.3.3	Other stream.....	66
7.3.4	Engineering Technology stream.....	67
7.3.5	Bio-System Technology stream.....	68
7.4	Analysis by the Grades obtained by each stream	72

7.4.1	Grade A	72
7.4.2	Grade B.....	72
7.4.3	Grade C.....	73
7.4.4	Grade S.....	73
7.4.5	Grade F.....	74
8	Conclusion and Future work.....	75
8.1	Introduction	75
8.2	Conclusion.....	75
8.3	Limitations.....	76
8.4	Future Developments	77
8.5	Summary	77
9	References.....	78
10	Appendices.....	81

Table of Figures

FIGURE 1.1 MODERN EDUCATIONAL DATA MINING	3
FIGURE 3.1 THE DATA MINING PROCESS	16
FIGURE 4.1 THE MODEL	20
FIGURE 5.1 TOP LEVEL DESIGN	27
FIGURE 6.2 SELECT THE CELLS TO IMPORT	29
FIGURE 6.1 FORMAT YOUR COLUMNS.....	29
FIGURE 6.3 REPLACE MISSING VALUE OPERATOR.....	30
FIGURE 6.4 VALUES REPLENISHMENT	30
FIGURE 6.5 EDIT PARAMETER LIST: COLUMNS	31
FIGURE 6.7 MISSING VALUES REPLACED DATA SET	32
FIGURE 6.6 MISSING VALUES REPLACED DATA SET	32
FIGURE 6.8 LEVEL OF DIFFICULTY IN PYTHON CHAPTER	34
FIGURE 6.9 PROCESS.....	35
FIGURE 6.10 INSERT COLUMNS	36
FIGURE 6.11 SELECTED DATA SET	36
FIGURE 6.12 PROCESS.....	37
FIGURE 6.13 SELECTED DATA SET	37
FIGURE 6.14 PROCESS.....	38
FIGURE 6.15 SELECT COLUMNS	38
FIGURE 6.16 PROCESS.....	39
FIGURE 6.17 PROCESS.....	39
FIGURE 6.18 SELECT COLUMNS WEB DESIGN	39
FIGURE 6.19 PROCESS WEB DESIGN	40
FIGURE 6.20 SELECT COLUMNS	40
FIGURE 6.21 PROCESS DFD	41
FIGURE 6.22 SELECT COLUMNS.....	41
FIGURE 6.23 FORMAT COLUMN.....	41
FIGURE 6.24 FORMAT COLUMNS.....	42
FIGURE 6.25 SELECTED DATA SET	42
FIGURE 6.26 SELECT ATTRIBUTE	43
FIGURE 6.27 PROCESS WITH OPERATORS	43
FIGURE 6.28 SELECT CELLS TO IMPORT	44
FIGURE 6.29 CONVERTED NUMERICAL TO BINOMINAL	44
FIGURE 6.30 SELECT ATTRIBUTES	45
FIGURE 6.31 PROCESS WITH OPERATORS	45
FIGURE 6.32 SELECT DATA SET	46
FIGURE 6.33 SELECT THE CELLS TO IMPORT.....	46
FIGURE 6.34 SELECT ATTRIBUTES	47
FIGURE 6.35 PROCESS WITH OPERATORS	47
FIGURE 6.36 SELECT CELLS TO IMPORT	49
FIGURE 6.37 FORMAT COLUMNS.....	49
FIGURE 6.38 FORMAT COLOMNS	49
FIGURE 6.39 SELECT DATA SET.....	51
FIGURE 6.40 SELECTED DATA SET	50
FIGURE 6.41 PROCESS WITH OPERATORS	50
FIGURE 6.42 NUMERICAL TO BINOMINAL CONVERTED DATA SET.....	51

FIGURE 6.43 PROCESS WITH OPERATORS	51
FIGURE 7.1 ASSOCIATION RULE RESULT MATHEMATICS AND PYTHON	52
FIGURE 7.2 ASSOCIATION RULE FOR PYTHON CHAPTER BY MATHEMATICS	53
FIGURE 7.3 ASSOCIATION RULE RESULT WEB DESIGN AND MATHEMATICS.....	54
FIGURE 7.4 ASSOCIATION RULE FOR WEB DESIGN CHAPTER BY MATHEMATICS.....	54
FIGURE 7.5 ASSOCIATION RULE RESULT DFD AND MATHEMATICS.....	55
FIGURE 7.6 ASSOCIATION RULE RESULT FOR DFD CHAPTER	56
FIGURE 7.7 ASSOCIATION RULE RESULT IT SUBJECT	57
FIGURE 7.8 ASSOCIATION RULE PYTHON IT SUBJECT	57
FIGURE 7.9 ASSOCIATION RULE RESULT WEB DESIGN IT SUBJECT.....	58
FIGURE 7.10 ASSOCIATION RULE WEB DESIGN FOR IT SUBJECT.....	58
FIGURE 7.11 ASSOCIATION RULE WEB DESIGN TABLE FOR IT SUBJECT.....	59
FIGURE 7.12 ASSOCIATION RULE RESULT DFD IT SUBJECT	59
FIGURE 7.13 ASSOCIATION RULE DFD QUESTION BY IT SUBJECT	59
FIGURE 7.14 ASSOCIATION RULE RESULT PYTHON BY GENDER.....	60
FIGURE 7.15 ASSOCIATION RULE PYTHON BY GENDER.....	60
FIGURE 7.16 ASSOCIATION RULE RESULT WEB DESIGN BY GENDER	61
FIGURE 7.17 ASSOCIATION RULE WEB DESIGN BY GENDER	61
FIGURE 7.18 ASSOCIATION RULE RESULT DFD BY GENDER	62
FIGURE 7.19 ASSOCIATION RULE DFD BY GENDER	62
FIGURE 7.20 ASSOCIATION RULE BY COMMERCE.....	63
FIGURE 7.21 ASSOCIATION RULE BY ARTS STREAM	65
FIGURE 7.22 ASSOCIATION RULE BY OTHER STREAM	66
FIGURE 7.23 ASSOCIATION RULE BY ENG_Tech STREAM	67
FIGURE 7.24 ASSOCIATION RULE BY GRADE	69
FIGURE 7.25 ASSOCIATION RULE GENERATED USING CONDITIONS.....	71
FIGURE 7.26 ASSOCIATION RULE BY A GRADES.....	72
FIGURE 7.27 ASSOCIATION RULE BY B GARDES.....	72
FIGURE 7.28 ASSOCIATION RULE BY C GRADES.....	73
FIGURE 7.29 ASSOCIATION RULE BY S GARDES.....	73
FIGURE 7.30 ASSOCIATION RULE BY F GRADES.....	74

List of Tables

TABLE 2-1 ACHIEVEMENTS AND THE LIMITATIONS OF THE KEY RESEARCH DISCUSSION IN THIS CHAPTER.	14
TABLE 5-1 DETAILS AN INITIAL QUESTIONNAIRE	24