



A BI approach for Student Engagement and Retention along with Cognitive Load analysis for Educator

M. N. Algewatta

Reg. No.: MS23004436

A THESIS
SUBMITTED TO
SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

January 2025

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Dr. Kalpani Manathunga

Approved for MSc. Research Project:

MSc. Program Co-ordinator, SLIIT

Approved for MSc:

Head of Graduate Studies, FoC, SLIIT

DECLARATION

This is to certify that the work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). The work has not previously been submitted in any form to the Sri Lanka Institute of Information Technology or to any other institution for assessment for any other purpose.



Sign:

Malithi Nethmini Algewatta

Date: 16th January 2025

ABSTRACT

A BI approach for Student Engagement and Retention along with Cognitive Load Analysis for Educator

Malithi Nethmini Algewatta

MSc. in Information Technology

Supervisor: Dr. Kalpani Manathunga

January 2025

This research introduces a systematic approach to monitoring student engagement, retention, and cognitive load in higher education, combining Business Intelligence (BI) tools with cognitive load analysis to address the complex challenges educators face in understanding and supporting student learning. By integrating diverse data sources, such as attendance, academic performance, mental health indicators, demographic factors, and feedback, the study demonstrates how a data-driven framework can provide educators with real-time, actionable insights into student behavior patterns. The BI system revealed valuable trends, including attendance irregularities, declining grades, and demographic influences, which helped educators identify at-risk students early and respond proactively. Additionally, the research employed cognitive load analysis to evaluate the mental demands of course content, categorizing learning objectives according to Bloom's Taxonomy. This analysis helped educators identify content that may overwhelm students and adjust complexity levels to support effective learning. The combined use of BI insights and cognitive load data presented a holistic approach that allowed educators to monitor student engagement more effectively and tailor instructional content to meet students' needs without cognitive overload. Findings from this study suggest that integrating BI tools with cognitive load metrics provides a structured, insightful framework that not only facilitates proactive intervention for student retention but also supports educators in creating a balanced, engaging, and supportive learning environment. This research offers a practical model that empowers institutions to foster student success through data-driven instructional adjustments and comprehensive engagement monitoring, meeting the evolving demands of modern higher education.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Dr. Kalpani Manathunga, my esteemed supervisor, for her invaluable guidance, unwavering support, and insightful feedback throughout this research endeavor. Her expertise and encouragement have been instrumental in shaping the direction of this work and enhancing its quality. Additionally, I extend my heartfelt thanks to all those who have contributed to this project in various capacities. Your assistance and collaboration have been deeply appreciated.

TABLE OF CONTENTS

DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENT.....	iv
TABLE OF CONTENTS	v
LIST OF FIGURES.....	ix
LIST OF TABLES	xii
Chapter 1 Introduction	1
1.1 Background of the Study.....	2
1.2 The Significance of the Student Engagement and Retention.....	4
1.3 The Significance of the Cognitive Load Analysis for Educator	6
1.4 Research Problem	7
1.5 Scope of the Study	8
1.6 Chapter Profiles.....	9
Chapter 2 Literature Review	11
2.1 The factors influence for the student engagement and retention in Higher Education	11
2.2 Business Intelligence and Higher Education.....	13
2.3 Available Products	15
2.3.1 Google Classroom.....	15
2.3.2 Edmodo	16
2.3.3 Sakai	16
2.3.4 Research Gap	16
2.4 Cognitive Load Analysis.....	17
2.4.1 Available Products	20
2.4.2 Research Gap	21
2.5 Aim of the Research.....	22

2.6 Research Questions	23
2.7 Research Hypotheses	23
2.8 Research Objectives	23
Chapter 3 Methodology	24
3.1 Research Strategy	24
3.2 Research Process	25
3.2.1 Problem Identification and Motivation	26
3.2.2 Definition of the Objectives for a Solution.....	27
3.2.3 Design and Develop the Artifact.....	27
3.2.4 Demonstrate the Artifact.....	28
3.2.5 Evaluate the Artifact	28
3.2.6 Communicate the Results	29
3.3 Sampling Procedure	29
3.3.1 Population	29
3.3.2 Sample	30
3.3.3 Sampling Techniques.....	30
3.3.4 Sample Size.....	33
3.4 Data Collection Instruments.....	33
3.4.1 Questionnaire of the research.....	36
3.5 Validity and Reliability	37
3.5.1 Validity	37
3.5.2 Reliability.....	37
3.6 Ethical Considerations	38
3.7 Role of the Researcher	40
3.8 Constraints	41
Chapter 4 Detailed Architecture of the Prototype	43
4.1 Architecture of the Prototype	44

4.2 The Components of the BI Approach	45
4.2.1 Attendance Prediction	45
4.2.2 Grade Analysis and Prediction.....	50
4.2.3 Demographic Analysis.....	56
4.2.4 Lecturer Feedback Analysis.....	59
4.2.5 Mental Health Analysis.....	62
4.2.6 Report Portal	70
4.3 Cognitive Load Analysis Approach	71
Chapter 5 Results and Discussion	88
5.1 Results of the Survey	88
5.1.1 Reliability of the Key Factors	88
5.1.2 Demographic Analysis.....	89
5.1.3 Analysis of Key Factors.....	92
5.2 Analysis of Relationships using Correlation.....	110
5.3 Analysis of Relationships using Regression	114
5.4 Testing Hypotheses	116
Chapter 6 Conclusion and Recommendations.....	119
6.1 Summary of the Findings	119
6.2 Addressing the Research Questions	120
6.2.1 Research Question 01	120
6.2.2 Research Question 02	121
6.3 Strengths and Achievements	123
6.4 Limitations of the Research	124
6.5 Future Recommendations	125
6.6 Personal Reflection	126
References	128
Appendix.....	132

Appendix 1: Questionnaire – Lecturer	132
Appendix 2: Dataset	139
Appendix 3: Database of the implemented system	142
Appendix 4: Interfaces of the implemented system	142

LIST OF FIGURES

Figure 2.1 B-Alert Cognitive State Software	20
Figure 3.1 Design Science Research general steps (Vaishnavi & Kuechler, 2004).....	25
Figure 3.2 Probability Sampling Techniques	31
Figure 3.3 Non- probability Sampling Techniques	32
Figure 4.1 Architecture of the Prototype.....	44
Figure 4.2 Attendance Prediction Flowchart.....	46
Figure 4.3 Preprocessing data - Attendance Prediction.....	47
Figure 4.4 Balancing the data - Attendance Prediction.....	47
Figure 4.5 Training the model - Attendance Prediction	48
Figure 4.6 Predicting Attendance	48
Figure 4.7 Attendance Prediction Results	49
Figure 4.8 Flowchart of the Grades Prediction	51
Figure 4.9 Calculate Improvement Ratios - Grades prediction.....	52
Figure 4.10 Forecasting Future Grades Using Linear Regression	53
Figure 4.11 Generate Visual Forecast Charts - Grades Prediction.....	54
Figure 4.12 Generate Overall Assessment – Grades Prediction	54
Figure 4.13 Grade Analysis and Prediction Results.....	55
Figure 4.14 Demographic Analysis Flowchart.....	56
Figure 4.15 Gender Distribution	57
Figure 4.16 Location Distribution	57
Figure 4.17 Age Distribution.....	57
Figure 4.18 Enrollments trends by year	58
Figure 4.19 Average age by gender	58
Figure 4.20 Average age by locations	58
Figure 4.21 Sentiment Analysis - Lecturer Feedback Analysis	59
Figure 4.22 Feedback Categorization - Lecturer Feedback Analysis	60
Figure 4.23 Lecturer Feedback Analysis Results 1.....	61
Figure 4.24 Lecturer Feedback Analysis Results 2.....	61
Figure 4.25 Mental Health Analysis Flowchart	63
Figure 4.26 Fetch data - Mental Health Analysis.....	64
Figure 4.27 Calculate average levels - Mental Health Analysis	64

Figure 4.28 Make conclusions - Mental Health Analysis	65
Figure 4.29 Risk Level and Video Recommendation Logic - Mental Health Analysis.....	66
Figure 4.30 Mental Health Analysis Results 1.....	67
Figure 4.31 Mental Health Analysis Results 2.....	68
Figure 4.32 Mental Health Analysis Results 3.....	69
Figure 4.33 Handle filters - Report Portal.....	70
Figure 4.34 DataFrame creation - Report portal	71
Figure 4.35 Report Generation in Excel and PDF Format - Report Portal	71
Figure 4.36 Bloom's taxonomy levels (Bautista, 2022)	72
Figure 4.37 Cognitive Load Analysis Flowchart	74
Figure 4.38 Data Preparation and Dataset Understanding	76
Figure 4.39 Label Encoding	76
Figure 4.40 Data Cleaning and Splitting.....	77
Figure 4.41 Tokenization	77
Figure 4.42 Model Initialization.....	78
Figure 4.43 Training Configuration	78
Figure 4.44 Training the Model	78
Figure 4.45 Model Evaluation.....	79
Figure 4.46 Calculate Cognitive Load Percentage.....	79
Figure 4.47 Explain Cognitive Load Results	80
Figure 4.48 Calculate Percentage Distribution.....	80
Figure 4.49 Sample Module Outline 1.....	82
Figure 4.50 Cognitive Load Analysis Approach Results 1.....	83
Figure 4.51 Sample Module Outline 2.....	85
Figure 4.52 Cognitive Load Analysis Approach Results 2.....	86
Figure 5.1 Demographic question 1 - Gender	90
Figure 5.2 Demographic question 2 - Department	90
Figure 5.3 Demographic question 3 - Experience	91
Figure 5.4 Dashboard and Reporting Q1	92
Figure 5.5 Dashboard and Reporting Q2	93
Figure 5.6 Dashboard and Reporting Q3	94
Figure 5.7 Dashboard and Reporting Q4	95
Figure 5.8 Cognitive Load Analysis Q1	96
Figure 5.9 Cognitive Load Analysis Q2	97

Figure 5.10 Cognitive Load Analysis Q3	97
Figure 5.11 Cognitive Load Analysis Q4	98
Figure 5.12 Data visualization Q1	99
Figure 5.13 Data visualization Q2	100
Figure 5.14 Data visualization Q3	101
Figure 5.15 Data visualization Q4	101
Figure 5.16 Predictive Analytics Q1	103
Figure 5.17 Predictive Analytics Q2	103
Figure 5.18 Predictive Analytics Q3	104
Figure 5.19 Overall System Effectiveness Q1	105
Figure 5.20 Overall System Effectiveness Q2	106
Figure 5.21 Overall System Effectiveness Q3	107
Figure 5.22 Overall System Effectiveness Q4	108
Figure 5.23 Overall System Effectiveness Q5	108
Figure 5.24 Overall System Effectiveness Q6	109
Figure 5.25 Overall System Effectiveness Q7	110
Figure 5.26 Correlation Between Dashboard and Reporting Usability (DAR) and Overall System Effectiveness (SYSN)	111
Figure 5.27 Correlation Between Cognitive Load Analysis (CLA) and Overall System Effectiveness (SYSN)	112
Figure 5.28 Correlation Between Predictive Analytics (PAN) and Overall System Effectiveness (SYSN).....	113
Figure 5.29 Correlation Between Data Visualization (DVN) and Overall System Effectiveness (SYSN).....	113
Figure 5.30 Regression between Dashboard and Reporting Usability (DAR) and Overall System Effectiveness (SYSN)	114
Figure 5.31 Regression between Cognitive Load Analysis (CLA) and Overall System Effectiveness (SYSN)	115
Figure 5.32 Regression between Predictive Analytics (PAN) and Overall System Effectiveness (SYSN).....	115
Figure 5.33 Regression between Data Visualization (DVN) and Overall System Effectiveness (SYSN).....	116

LIST OF TABLES

Table 2.1 Research Gap	17
Table 3.1 Format of the 5-point Likert scale.....	36
Table 3.2 Structure of the questionnaire.....	36
Table 5.1 Cronbach's alpha values	88
Table 5.2 Results of Reliability Analysis.....	88
Table 5.3 Dashboard and Reporting Questions.....	92
Table 5.4 Cognitive Load Analysis Questions.....	95
Table 5.5 Data Visualization Questions	99
Table 5.6 Predictive Analytics Questions	102
Table 5.7 overall system effectiveness Questions.....	105
Table 5.8 Correlation Values	111