ATHABASCA UNIVERSITY

STUDENT FACING LEARNING DASHBOARD DESIGN FOR ONLINE LEARNERS

BY

ANTHONY AZIMI

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN INFORMATION SYSTEMS

FACULTY OF SCIENCE AND TECHNOLOGY

ATHABASCA, ALBERTA

NOVEMBER, 2024

Approval Page

Athabasca University
Faculty of Graduate Studies

Approval of Thesis

The undersigned certify that they have read the thesis entitled

STUDENT FACING LEARNING DASHBOARD FOR ONLINE LEARNERS

Submitted by

Anthony Azimi

In partial fulfillment of the requirements for the degree of

Master of Science in Information Systems

The thesis examination committee certifies that the thesis
and the oral examination is approved

Supervisors:
Dr. Ali Dewan
Athabasca University

Dr. Oscar Lin
Athabasca University

External Examiner:
Dr. Eric Wang
Athabasca University

December 20, 2024

1 University Drive, Athabasca, AB, T9S 3A3 Canada
Toll-free (CAN/U.S.) 1.800.788.9041 ex. 6821
fgs@athabascau.ca | fgs.athabascau.ca | athabascau.ca

Dedication

I dedicate this thesis to everyone who supported me over the past few years. To my family, who loved me unconditionally, taught me by example to work hard, and encouraged me to pursue my dreams. And to my wonderful girlfriend, who supported and inspired me, standing by my side every step of the way.

Acknowledgement

I would like to express my deepest gratitude to everyone who has supported me throughout the completion of this thesis. First and foremost, I extend my sincere thanks to Dr. Ali Dewan, whose invaluable guidance, expertise, and encouragement were instrumental in shaping this research. His mentorship and insights challenged me to reach new heights, and I am immensely grateful for his support and patience throughout this journey. I am also profoundly grateful to Dr. Oscar Lin for his thoughtful advice and unwavering support. His constructive feedback and dedication to excellence have greatly enriched my work and inspired me to pursue my research with greater depth and precision. I am thankful for his time, encouragement, and the many discussions that helped refine my understanding and approach. Finally, I would like to acknowledge the support of my family and my wonderful girlfriend, who have been a constant source of strength and motivation. Their encouragement made this journey possible, and I am deeply appreciative of their love and understanding.

Abstract

The shift from traditional classrooms to online learning has highlighted the need for self-regulated learning. This study investigates how learning analytics dashboards (LADs) can enhance self-regulation by visualizing performance data. An intelligent student-facing learning analytics dashboard (SF-LAD) was developed and evaluated to support online learners by improving planning, organization, and engagement. Key features include tracking progress over time, comparing performance with peers, analyzing discussion board sentiment, and monitoring login patterns to boost self-awareness. Findings show that younger students (<30) find dashboards more beneficial for self-reflection and engagement, though some features, like peer comparison and sentiment analysis, received mixed feedback, suggesting a need for customization. While the dashboard demonstrated usability within the 50th percentile, the net promoter score (-6%) indicated room for improvement, such as adding "call-to-action" prompts. Overall, SF-LADs proved effective for managing time, motivating action, and creating visibility, offering valuable insights for future iterations.

Keywords: Intelligent Learning dashboard, online learning, learning management system, information visualization, data mining, self-regulated learning, machine learning, predictive modeling.

Preface

The design and development of a Student-Facing Learning Analytics Dashboard (SF-LAD) presented in this thesis was inspired by the growing need to support self-regulation among online learners. My background in educational technology and personal interest in improving the online learning experience led me to explore the potential of intelligent, AI-driven dashboards that could help students monitor and manage their learning. This work is driven by the belief that creating visibility into students' progress and performance can empower them to set goals, plan effectively, and take control of their educational journey.

In this research, I have contributed to each stage of the study— from conceptualizing the research questions to designing the SF-LAD, analyzing the data, and presenting the findings. The SF-LAD is designed to foster awareness, self-reflection, and self-regulated learning by providing students with insights into their cognitive, metacognitive, behavioral, and emotional competencies. Through machine learning, natural language processing, and data visualization, this tool equips students in self-paced online learning environments with the resources to manage their academic goals and engage meaningfully with the material.

Thesis Structure and Contents

This thesis is organized into seven chapters, each building on the foundation of how learning analytics can enhance self-regulated learning:

- **Chapter 1**: This chapter introduces the role of online education within higher education, emphasizing the challenges it faces, including student isolation, reduced interactivity, and the need for tools that promote engagement and self-regulation.

- **Chapter 2**: A review of the literature on learning analytics dashboards (LADs) is provided, exploring their development from instructor-focused risk detection tools to student-centered instruments for enhancing self-awareness and guiding learning behaviors.
- **Chapter 3**: Here, the methodology of the research is outlined, detailing the design and evaluation process for the SF-LAD, with a focus on its role in supporting self-regulated learning in online environments.
- **Chapter 4**: This chapter covers the setup and execution of the usability study for the SF-LAD, describing the sampling strategy, data collection methods, and analytical tools, including the System Usability Scale (SUS) for evaluating dashboard performance, reliability, and validity.
- **Chapter 5**: The results of the usability testing and hypotheses testing are presented in this chapter, followed by an analysis of the predictive modeling and sentiment analysis outcomes for the SF-LAD, demonstrating its effectiveness in fostering planning, organization, and self-monitoring.
- **Chapter 6**: The thesis concludes by discussing the impact of self-regulation on online learning success and how the SF-LAD can support students in developing this essential skill. The chapter underscores the value of integrating data visualization, predictive modeling, and sentiment analysis to help learners manage their academic responsibilities.

Importance of the Study

This research is significant as it advances the understanding of how student-facing learning dashboards can support self-regulated learning, an increasingly important skill in the digital learning landscape. By identifying effective features and highlighting the dashboard's impact on students' engagement and reflection, particularly among younger learners, this study provides valuable insights for future dashboard designs that prioritize usability and flexibility.

Acknowledgements

I would like to express my heartfelt gratitude to my mentors, Dr. Ali Dewan and Dr. Oscar Lin, from Athabasca University, for their unwavering support and guidance throughout this project. Their mentorship and encouragement led to several successful publications based on this research. I am deeply appreciative of their belief in my work and their commitment to fostering my academic and professional growth.

Table of Content

List of Tables

List of Figures

List of Symbols, Nomenclature, or Abbreviations

CMS: Content Management System

KPI: Key Performance Indicator

LAD: Learning Analytics Dashboard

LMS: Learning Management System

MOOC: Massive Open Online Courses

NLP: Natural Language Processing

SF-LAD: Student Facing Learning Analytics Dashboard

SPOL: Self-Paced Online Learning

SUS: System Usability Scale

UX: User Experience

Chapter 1. Introduction

In recent years online learning has grown in popularity. Many higher educational institutions are adopting online education and they are working towards improving the quality of their online courses while overcoming several challenges such as increasing teaching efficacy, removing the feeling of isolation, and improving technologies to support learning online (Farahmand, Dewan, & Lin, 2021). Given the rapid growth of online learning, it is important for institutions to provide high quality education for learners. Many institutions are now using learners' achievements and satisfaction as a means of assessing the quality of their programs. Although several research studies showed mixed reviews regarding these assessments, researchers pointed out that online learning could be as effective as traditional classroom-based learning (Kim & Bonk, 2006).

A predominate part of face-to-face learning is the social and communicative interaction between students and instructors and students and students (Ya Ni, 2013). A student's ability to ask questions, share opinions, or agree/disagree with point of view is fundamental learning activity. This is because conversation and discourse with instructor and other students is where new concepts are clarified, skills are practiced and original ideas are formed (Ya Ni, 2013). This goal could be achieved in online learning using great technological support with web-based tools (Ya Ni, 2013). The tools such as discussion boards, synchronous chats, electronic bulletin boards and emails can create an integrated learning environment, which promotes student centered learning. This approach should encourage wider participation by students and produce more in-depth discussions than what is possible in traditional face-to-face classrooms (Ya Ni, 2013).

Students enrolled in online learning have access to highly advanced educational material in their field of study from any geographical location (Durall & Gros, 2014). Furthermore, adult learners can combine their professional experience and higher education with their work responsibilities in a flexible manner (Whitehill, Serpell , Lin, Foster , & Movel, 2014; Cohen & Nachmias, 2006; Holder B. , 2007). Online learning benefits institutions by providing more learners access to educational services while reducing the overall cost to the institution (Cohen & Nachmias, 2006).

Online-courses heavily rely on "learning activities" to deliver course material and facilitate knowledge acquisition (e.g., videos, self-assessments, online articles, etc.). However, traditional online learning processes lack methods to incentivize and motivate students to complete these learning activities. This causes students to feel isolated as they are taught through tutorially designed learning material (Farahmand, Dewan, & Lin, 2021). This lack of interaction often affects students' progress and performance negatively (The University of British Columbia, 2014). A student facing learning dashboard (SF-LAD) can be adopted as an important tool to enforce students' learning activities in an online environment by providing immediate feedback on students' progress, which can be accomplished by compiling and visualizing data on individual students' work on learning activities. Data can be captured from log data in learning management systems (LMS), video platforms, and course websites with domain–specific learning activities (The University of British Columbia, 2014).

Self-regulated learning is a set of individual strategies that students undertake in order to learn (Aguilar, Karabenick, Teasley, & Baek, 2021). Based on the literature, there is a gap in adopting novel approaches to motivate self-regulated learning in an online learning environment (Li & Lajoie, 2022; Onah, Pang, Sinclair,, & Uhomoibhi,, 2021; Boekaerts, 1996). For example, a novel approach to motivate self-regulated learning is integrating it with cognitive engagement (i.e., psychological investment of the students in the learning process), which is marked by learners' effort to understand what they are studying and their attempt to reach the highest levels of comprehension on a particular area of study. It worth noting that integrating these theories can advance our understanding of what drives student efficiency to support more effective learning (Li & Lajoie, 2022). Adopting such novel approaches can bring more focus on the interacting effects of cognitive and motivational processes that account for student learning outcomes (Li & Lajoie, 2022). Accordingly, developing tools such as learning dashboards with novel features provides a path to support self-regulated learning skills to aid learners steer in the right direction to achieve their learning goals (Onah, Pang, Sinclair,, & Uhomoibhi,, 2021).

This research aims to minimize this gap by exploring ways to transition from traditional learning processes in online studies to innovative approaches such as designing and developing tools with novel features to help online learners improve their self-regulation skills by understanding their cognitive, metacognitive, behavioral, and emotional competencies through an intelligent student facing learning dashboard (SF-LAD). Research shows that learning dashboards can have cognitive and metacognitive benefits for students (Durall & Gros, 2014). Learning dashboards can support students by providing information about their progress in a course and flag if they are at risk of failing a course (Durall & Gros, 2014). Furthermore, learning dashboards can give students personalized recommendations to read material and suggest learning activities that could greatly improve students' engagement in online courses (Bodily, et al., 2018).

1.1. Statement of the Problem

Online learning is fast becoming a key part of how post-secondary institutions offer courses; however, there is a distinct gap between students' satisfaction in online courses, compared to traditional face-to-face courses (Kim & Bonk, 2006). Students often feel isolated as teaching in on-line courses is based on tutorially designed learning materials. Therefore, students lack enough interactions with the instructors and other students to engage and compare their progress and performance. Student facing learning dashboards (SF-LAD) can help to minimize this gap by providing online students a tool to accurately measure their performance to improve their planning, organization, and self–assessment. This can result in improved self–regulated learning by measuring how learners manage cognition, motivation, and behavior to meet their learning goals. Research has recognized that these can be important predictors of achievement in an online learning environment because self–regulated learners can adapt and monitor their cognition and motivation to better align with task demands, resources, and goal settings (Zhang & Qin, 2018).

1.2. Purpose of Research

According to Jivet et al. (2017) the most common goal of learning dashboards is increasing awareness and reflection (Jivet, Scheffel, Drachsler, & Specht, 2017). The purpose of this research is to leverage the potential of using an intelligent student facing learning analytics dashboard to increase individuals' self-regulation by creating visibility into their performance.

1.3. Research Objective

The goal of this research is to design and test an intelligent student facing learning dashboard (SF-LAD) that could assist self-paced online learning (SPOL) students. This study addresses research challenges in designing learning dashboards that focus on fostering awareness and reflection, students goal setting/managing and planning, and measuring self-regulated learning and engagement in the course. This is accomplished by creating visibility into cognitive, metacognitive, behavioral, and emotional competencies of online learners by using predictive modeling, natural language processing modeling, and data visualization.

1.4. Research Hypothesis

The research hypothesis is that there is a correlation between creating immediate visibility to students' progress (compared to both self and others in the class) and self-regulation, which leads to enhanced performance in the online courses.

1.5. Research Question

This research attempts to answer the following two major research questions:

1. How to design an effective SF-LAD to enhance students' organizational skills, self-regulation, and motivation in online learning?
2. How to evaluate the effectiveness of a SF-LAD in terms of enhancing students' organizational skills, self-regulation, and motivation in online learning?

1.6. Delimitations of Research

This research is focused on the four main issues with the existing learning dashboards. First, few learning dashboards have focused on fostering awareness and reflection, considering that reflection and awareness are difficult to measure objectively (Jivet, Scheffel, Drachsler, & Specht, 2017). Second, existing learning dashboards are not designed to support students' goal setting, planning, as well as helping students manage their self-set goals (Jivet, Scheffel, Drachsler, & Specht, 2017). Third, existing learning dashboards are focused on peer comparison by just looking at how learners are connected with and are supported by their peers to measure motivation and engagement (Jivet, Scheffel, Drachsler, & Specht, 2017). Fourth, existing dashboards are focused on short-term improvements by constantly exposing students to motivational triggers that rely on social comparison to peers rather than long-term improvement by fostering mastery of knowledge, acquiring skills, and developing competencies (Jivet, Scheffel, Drachsler, & Specht, 2017). To increase long term improvement through awareness and reflection, this research aimed to design and evaluate a learning dashboard that objectively fosters mastery of knowledge, acquiring skills, and developing competencies by allowing evaluation of cognitive, behavioral, emotional state and metacognitive abilities to enable goal setting, planning, and motivation through engagement.

Chapter 2. Review of Literature

2.1. Overview

Given the rapid growth of online education, it's important for academic institutions to provide high quality online programs (Kim & Bonk, 2006). Learning analytics dashboards (LADs) can be used in online education to improve the quality of online learning experience. However, historically, business intelligence (BI) and data visualization in academic institutions have been focused on transactional processes, for example, students' enrollment and admission, rather than students' learning behavior. While using learning analytics dashboards is rather a new adoption in education, recently, the focus of learning dashboards has shifted to support students by looking at the students' behavioral patterns and progress (Durall & Gros, 2014).

2.2. Methods for Learning Analytics

Learning analytics dashboards (LADs) can be defined as applications that can create visibility into learners' online behavior patterns in a virtual learning setting (Park & Jo, 2015). LADs are a specific class of "personal informatics" applications. These tools support users in collection of personal information about various aspects of their life, behavior, habits, and thoughts (Durall & Gros, 2014). LADs can be adopted in online learning as a supporting tool, which can be deployed to track students' self-awareness and performance in the course. This objective can be achieved through log-files and data mining to discover meaning, followed by visualization of the results to bring visibility into students' performance (Park & Jo, 2015).

Originally, dashboards were developed with educators as target users with focus on creating visibility to students with high risk of failing academically. However, in recent years, student facing learning dashboards are becoming more and more attractive as a research topic (Lim, Dawson, Joksimovic, & Gašević, 2019). One of the major benefits of LADs is their ability to track students' performance by allowing learners to discover hidden patterns, unknown correlations, and other useful information that allow them to make evidence based decisions (Durall & Gros, 2014). As a result, nowadays student facing learning dashboards are becoming more common as a tool to provide students with personalized feedback (Lim, Dawson, Joksimovic, & Gašević, 2019).

Evidence shows that learning dashboards can have a number of cognitive and metacognitive benefits for students (Durall & Gros, 2014). For example, learning dashboards can assist students by providing information about their progress in a course, and flag if they are at risk of failing a course (Durall & Gros, 2014). Other benefits of learning dashboards include providing customized recommendations to students in relation to reading material and learning activities and measuring if a particular student is improving (Durall & Gros, 2014). Finally, interactive learning dashboards can tailor the course offering based on students' needs (Durall & Gros, 2014).

It is worth noting that although many institutions have accepted that learning analytics dashboards (LADs) can be very helpful to students' success, few institutions have implemented LADs as part of their educational system (Lim, Dawson, Joksimovic, & Gašević, 2019). The reason is that research into learning analytics has found that most of the feedback generated by LADs is not developed from theoretically established instructional strategies, particularly when giving feedback to students. In other words, the feedback generated by LADs are purely fact–based. The issue is that the fact–based feedback with detailed instructive messages on how to address learning deficiencies can be difficult to process for students (Lim, Dawson, Joksimovic, & Gašević, 2019).

Current literature shows that a key component of students' success in online courses is engagement in the course, which can have a significant impact on students' performance (Mushtaq, Zhu, Zhang, & Abidi, 2018). In online courses we can define students' engagement as the amount of time students spend on the learning process for course specific content. Several studies indicated that the level of student engagement can have a direct effect on students' final grades, retention of material, and the course dropout rate (Mushtaq, Zhu, Zhang, & Abidi, 2018).

To improve students' engagement, including active participation and collaboration in an online learning environment, students need to have a satisfactory level of interaction with the course (Pellas, 2014). According to a study by McNamara et al. (2017), language has a fundamental importance in the field of education as it is a conduit for the communication and understanding of information (McNamara, Allen, Crossley, Dascalu, & Perret, 2017). Learning management systems (LMS) can generate a considerable amount of information about language generated by the students within discussion board forums and emails. This data can be mined using natural language processing (NLP) to better understand the depth of students' grasp of course topics. In addition, the analysis of discussion forum posts using NLP can provide insights into students' social interaction with colleagues and instructors, which is a good indicator of their course engagement (McNamara, Allen, Crossley, Dascalu, & Perret, 2017). As a result, NLP analysis can be summarized in the dashboard to present students with information about their level of understanding for various course topics, which topics they found challenging, and how they interact with the instructor and other students in the course (Alblawi & Alhamed, 2017). The theory is that by showing this information to students, we could potentially help students enrolled in self-paced online learning (SPOL) courses to feel less isolated, and increase their self-awareness, which could affect their motivation and engagement level in the course (Farahmand, Yan, Dewan, & Lin, 2022). So, to summarize, the information collected from learning management systems (LMS) can be used to create visibility and awareness into the dynamics of how a student communicates with his/her peers as part of learning dashboards' design features.

Also, literature suggests that academic self-concept or self-awareness can be formed through both *internal comparison* (comparing with one's own performance in various areas or over time) and *external comparison* (comparing with performance of the other students). This concept is referred to as *frame of reference* (Lim, Dawson, Joksimovic, & Gašević, 2019), and can be used in student facing learning dashboard to improve self-awareness.

Furthermore, literature indicates that the online learning environment is designed for autonomy. As a result, self-regulation is critical for success in online learning (Barnard, Lan, To, Paton, & Lai, 2009)**.** Studies have demonstrated that enhancing self-regulatory behavior can improve academic performance in traditional classrooms (Barnard, Lan, To, Paton, & Lai, 2009). If self-regulation is important to face-to-face learning, then it can be expected that self-regulation skills play a critical role in online learning (Barnard, Lan, To, Paton, & Lai, 2009).

Another important variable that is associated with learning achievement and can be considered in learning dashboards' design is learning time. Learning time can be described as the amount of time that students were engaged in learning (Hwang & Wang, 2004). This information can be used to assess students' performance. Learning time distribution consists of the distribution of individuals' web-based course inter-arrival time and the distribution of their on-line duration (the inter arrival time refers to the length of time between two successive logins to the learning management system) (Hwang & Wang, 2004).

Learning dashboards use data mining to draw insight and discover patterns from the data captured from the learning management system (LMS). This information can be used to predict students' behavior, study patterns, students' planning and organization ability, motivation, and engagement level in a course. The information can even be used to identify different learner types, as well as the characteristics of their interactions with the learning management system (LMS) (Bienkowski, Feng, & Means, 2012). By mining student data from both LMSs and instructor records, researchers can predict students' performance in relation to the course material. This information is both descriptive and predictive and can lead to prescriptive recommendations. From the descriptive to perspective, data mining may help answer the questions such as "what happened?", "where was the problem?" and "what actions should be taken?" (Dietz-Uhler & Hurn, 2013).

AI based (intelligent) student facing learning dashboards can help students' metacognition by helping them to monitor their learning behaviors and learning process and assess their own learning achievement using criteria to regulate their learning based on the results of the monitoring. To be effective, students have to monitor their activities in relation to learning goals or criteria set in the planning phase and then regulate their learning by revising their strategies or methods or making other adjustments (Chen, Min, Goda, Shimada, & Yamada, 2021). Supporting students' self-monitoring activities means that the dashboard should help students in self-observation or self-recording processes, such as recording a specific behavior that occurred and specific learning achievements that met the criteria or learning goal (Chen, Min, Goda, Shimada, & Yamada, 2021).

AI based (intelligent) student facing learning dashboard design is enabled by the current content management systems (CMS) that collect a vast array of information about students' activities including: discussion board posts, access to reading material, and completion of course quizzes and other assessments, which can generate thousands of transactions per student (every activity performed by a student will be considered a transaction) (Picciano, 2012). It is worth noting that modern learning management systems (LMSs) collect data about the number of times students access resources, date and time of access, the number of discussion board posts generated, the number of discussion boards read, and the type of resources accessed. Aggregated trace data can support inferences about (i) learner's preferred work schedule that mildly support inferences about procrastination, (ii) which resources are judged more relevant and appealing, (iii) motivation to calibrate judgement of learning and efficacy, and (iv) value attributed to contributing, acquiring, or clarifying by exchanging information with peers (Winne, 2017).

The students' digital footprint (logged data) can then be mined and analyzed to identify patterns of learning behavior that can provide insights into students' educational practices. This information can then be used to predict students' learning success and provide proactive feedback to students (Gašević, Dawson, & Siemens, 2014).

Analyzing such transactions can benefit both the student and the learning institution by providing both groups with accurate information based on learner's performance. It is important to note that intelligent students facing learning dashboards enable learners to have instant feedback about their performance in the course, which means that they don't have to wait for instructor's feedback to make critical decisions. This will result in better course experience for students as they will be able to measure their performance against other students in the class, thus will be able to immediately identify areas for improvement (Picciano, 2012).`

Our research is focused on creating a dynamic AI based (intelligent) student facing learning dashboard to enhance students' self-regulated learning (SRL) and ultimately boost their engagement in online courses, empowering them to reach their academic goals. This study acknowledges the pivotal role of engagement in students' success and underscores its importance in designing a student-friendly learning dashboard, in line with prior studies' findings.

2.3. Critical Discussion on Research Gaps and Proposed Solution

One of the main barriers in online learning is that teaching with technology is not a one size fits all approach (Gillett-Swan, 2017). Another barrier in online courses is that students may feel isolated as collaborative learning could be difficult in online learning environment (Molenaar, Horvers, Dijkstra, & Baker, 2019). For example, direct student-instructor interaction is not available, and students often learn in an individualized mode (Yan, Lin, & Kinshuk, 2020). It means that the social interaction of student-instructor and student-student is often missing (Yan, Lin, & Kinshuk, 2020). This causes students to feel isolated. In addition to the solitude, students lack information on how they are performing compared to their peers (Yan, Lin, & Kinshuk, 2020). Without social interaction and formative feedback, students usually lack self-awareness (e.g., how they are performing, what are their learning gaps, and what they can do to improve). The lack of learning awareness can cause several challenges including when to seek help, how to. motivate themselves, and how to regulate their learning (Yan, Lin, & Kinshuk, 2020). Finally, many online courses are designed in such a way that students can choose when to submit summative assessments (Yan, Lin, & Kinshuk, 2020). Thus, students who submit summative assessments towards the end of the course may receive minimal formative feedback during the course, which could affect their course performance (Yan, Lin, & Kinshuk, 2020).

Technology can play a significant role in removing barriers in a course (Molenaar, Horvers, Dijkstra, & Baker, 2019) by using data mining to analyze data collected from log files to discover meaning, followed by visualization of the results to bring visibility into students' performance (Park & Jo, 2015). For example, adaptive learning technologies and personalized visualizations in intelligent student facing learning dashboards can support learners.

Three critical applications are suggested by Verbert et al. (2013) to remove barriers in online learning while designing learning dashboard for students: *1) increase learning awareness*, *2) identify struggling students,* and *3) provide academic intervention* (Verbert, Duval, Klerkx, Govaerts, & Santos, 2013). In online courses, students' self-awareness of knowledge state and learning gaps can help them to reflect on their learning and decide when to ask for help. Learning dashboards can help with this through social comparison, goal achievement, and learning progress (Bodily, et al., 2018). Venant et al. (2016) found that social comparison with other students in online courses can be a motivating factor, which could cause students to work harder, increase their course engagement, and simulate a feeling of connectedness to peers (Venant, Vidal, & Broisin, 2016)**.**

Dashboards can assist students by providing information about students' progress in a course and flag if they are at risk of failing (Durall & Gros, 2014). The key benefits of learning dashboards from students' perspective are capturing information in real-time, allowing students to map their objectives to their learning path, and giving students an overview of their personal learning preferences and style. These would allow students to measure their progress more accurately, better understand their strengths, weaknesses, and learning preferences, make accurate predictions about their future progress and make better decisions for the next steps in their learning path (Howlin & Lynch, 2014). Furthermore, learning dashboards can provide customized recommendations concerning students' reading material and learning activities and measuring if a student is improving (Durall & Gros, 2014). To summarize, learning dashboards can tailor the course offering based on students' needs (Durall & Gros, 2014).

2.4. Review of Existing Learning Analytics Dashboards

Bodily, et al (2018) performed a comprehensive literature review on the existing student facing dashboards to examine methods and functionalities that enhance the rigour of dashboard design, as well as their impact on students' achievement, behaviour, and skills. The study found that while the most common objective of the designed dashboards was to improve student awareness and reflection, only 2% of the reviewed articles conducted an experiment to identify the effectiveness of the dashboard on students' skills (e.g., students' awareness and reflection). Another interesting finding by the study is that most of the student facing dashboards are not designed to improve students' engagement; instead, most of the student facing dashboards are focused on increasing students' awareness and reflection (Bodily, et al., 2018).

Course Signal is a dashboard designed by Purdue University (Arnold & Pistilli, 2012). Course Single is a student success system that allows faculty to give students meaningful feedback using a predicative model. The course signal was designed to use the data found in educational institutions including data collected from instructional tools and Learning Management Systems (LMS) to provide students with real time information if they are at risk of failing based partially on their effort on the course. Through application of analytics on large datasets, which were mined for insights, the dashboard also identifies if students might be falling behind. One of the goals of this dashboard is to produce action intelligence to guide students to improve their retention and performance outcomes. Course Signal uses summative evaluation to determine effectiveness of the dashboard. Variables that Course Signal uses to track students' performance and retention include *login trends*, *performance results*, *content usage* and *message analysis.*

Narcissus is a dashboard designed at the University of Sydney. The target users for Narcissus dashboard are students, and the dashboard is designed to enable students to collaborate more effectively and work in groups over a long period of time. Narcissus dashboard uses data from Trac (an open-source, web-based project management and bug tracking system), which can be used to generate a group model. This model is then visualized in an active interface. The dashboard uses a Wattle Tree to visualize the activities of the group members over time. The first version of the dashboard was then tested in Firefox 2.0.0.9 on Mac OSX. The Narcissus dashboard is written in Python using plugin from Trac and Quartz (a job scheduling library that can be integrated into different Java applications). This approach leverages Trac's ability to visualize information in real time. Summative assessment was used to determine the effectiveness of the dashboard in assisting individual students and groups to reflect on their performance. As part of the summative assessment, three usability studies were conducted: 1. individual level study, 2. group and individual level study and 3. facilitated study to determine the actual usefulness of the dashboard. Students who participated generally found that using the dashboard allowed them to work better in groups and gave them the opportunity to reflect on their group work, which was presented in their final report (Kay & Upton, 2006).

Student Inspector dashboard is designed by Scheuer and Zinn (2012). Student Inspector is designed for both teachers and students. Using Student Inspector, teachers are able to get to know each student at an individual level in the context of online learning. The dashboard is designed to capture each student's learning path by keeping track of learners' knowledge, higher level competencies or the lack thereof and provide instructors with the list of subjects that each student struggles with. This information can also be used in a student facing learning dashboard to provide the information regarding the course progress. Student Inspector was designed with three major components (the browser, performance measurement, and topic coverage). Student Inspector was evaluated through a summative approach, using a questionnaire to collect data on the students' interactions with the dashboard (usability testing). Student Inspector tracks student data automatically from the learning management system (LMS); this information is then visualized in the dashboard using pie and bar charts. Overall, Student Inspector received positive feedback from the students and instructors who interacted with the dashboard (Scheur & Zinn, 2007).

Glass is a dashboard designed for both teachers and students by Leony et al. (2012). Glass was designed as a visualization tool for instructors and student to show students' performance in comparison to the rest of the group. Glass dashboard tracked students' login trends, performance results, content usage, and message analysis. The dashboard collects data from the learning management system (LMS) and analyzes the data by using descriptive statistics, focusing on the representation of the raw data characteristics. This information is then visualized using graphs such as line charts and bar charts (Leony, Pardo, Valentín, Castro, & Kloos, 2012).

The Student Activity Meter (SAM) was designed by Govaerts, Duval and Verbert and Pardo in 2012 as a student and instructor facing dashboard, which aimed to visualize user actions to enhance learners and instructors' awareness to support self–reflection. SAM visualizes data tracked in learning environments, using metrics such as basic time spent, resources used, forum views, and post actions. SAM allows instructors to quickly identify which students are at risk, gain insights into how learners are spending their time on the course activities, and gain insight into which tools within the LMS are most used by students. SAM allows students to self reflect on their activities by allowing them to compare their performance against their peers (Park & Jo, 2015).

Stepup is a dashboard developed by Santos, Verbert and Duvel (2012) for both instructors and students. The goal of the dashboard is to promote student reflection and enhance awareness of their activities. The dashboard uses summative evaluation methodologies to prove the usefulness of the dashboard to students and instructors. Stepup tracks student activities by using descriptive statistics to analyze the data and visualize information for students and instructors (Santos,, Verbert, & Duval, 2012).

Loco–Analyst is a learning dashboard developed by Jovanović, Gašević, Brooks, Devedžić, & Hatala, (2007) to suggest meaningful feedback to educators and students based on analysis of the user tracking data. The goal of the dashboard is to track students' login trends, performance results, content usage, and message analysis. It uses bar graphs, pie charts and table matrix as well as a tag (word) cloud to display the feedback to instructors and students in an easy to understand manner. The dashboard tracks students' activities automatically from learning management systems (LMS) and analyzes the data, using descriptive statistics. Overall, 50% of the students and instructors found the feedback generated by the dashboard to be informative and helpful to improve learning (Park & Jo, 2015).

Student Success System (S3) is developed by Essa and Ayad (2012) and is an instructor facing dashboard designed to identify students at risk of failing a course. The dashboard uses student performance results, social network data, and students at risk scores to determine which students are at risk of failing. This information is visualized in the dashboard using risk quadrant, scatterplot chart, win-lost chart and sociogram. The dashboard was positively received, as instructors were able to successfully identify students who were at risk of failing and provided them with appropriate assistance (Essa & Ayad, 2012).

SNAPP is developed by Bakharia and Dawson (2011) and is an instructor facing dashboard designed to evaluate students' participation. The dashboard collects information on content usage and social network, visualizing this information using a sociogram. It uses proprietary tool to track students' content usage and social networking activities. SNAPP uses descriptive statistics to extract insight from the data. Overall, SNAPP is considered a powerful tool for visualizing the participants' relationships in discussion forums and identifying isolated discussers. It can also be used to provide real time feedback to students regarding their learning and performance (Bakharia & Dawson, 2011).

Learning Analytics for Prediction and Action Dashboard (LAPA) was designed by Park and Jo for students and teachers in online courses. This learning and teaching support tool is designed to stimulate students' learning (actions) from a narrow and short-term perspective. The goal of LAPA is to inform students' online learning behavior to students and instructors. The dashboard visualizes students' online activity summary, total login time, total login regularity, visits to the discussion boards, time spent on boards and visits to the repository. The information is visualized using a scatter plot, bar graph, and line graph. It tracks students' activities automatically from learning management system (LMS). The dashboard was evaluated using formative assessment through usability testing and summative assessment by conducting experiments using control groups in two classes and the results of a survey on the perceived usefulness and the degree of conformity (Park & Jo, 2015).

Goal-Oriented Visualization of Activity Tracking is a dashboard designed by Santos et al. (2012) to increase students' motivation by tracking their course activities to promote self-reflection and comparison with peers. It uses a bubble chart to display students' goals according to the importance of each goal. A line chart is used to visualize students' course activates over time. A bar chart displays the frequency of students' actions, course events, and the number of logins. The benefit of the dashboard is helping students reflect on their learning goals and support communication with the instructor. However, the dashboard's navigation is not intuitive, decreasing its usefulness. Moreover, this dashboard displays several redundant metrics in a series of tables, making the visualizations hard to understand (Santos, Govaerts, Verbert, & Duval, 2012).

Learning Analytics for Online Discussions is a dashboard designed by Friend Wise, Zhao & Hausknecht (2014) to increase awareness, using real-time information about students' interactions with other students in asynchronous online discussions. The dashboard uses starburst visualization to provide an outline of the flow of conversation, showing the connections between ideas. The visualization helps students to recognize which discussion threads are more active. However, starburst is a complex visualization, making it difficult for some students to understand the structure of the discussions. Furthermore, the dashboard only tracks the discussion structure, ignoring additional metrics such as students' agreement or disagreement with the concepts (Friend Wise, Zhao, & Hausknecht, 2014).

Notification System on Student Behaviors is a dashboard designed by Xu and Makos (2010) to increase students' awareness in a collaborative online learning environment through a personalized notification system. The dashboard uses correlation and ANOVA to predict students' performance. Correlation and ANOVA determine the relationship between students' self-efficacy, the use of notification, and students' collaborative behaviour (Xu & Makos, 2010). The benefit of this model is promoting students' awareness of their online behavior and its impact on group work. The limitation of this model is that it only focuses on the students' behavior and ignores the cognitive and metacognitive elements of the collaboration.

2.5. Review of Learning Analytics Machine Learning Models

Different predictive models have been used in recent studies, especially for educational data mining and students' progress predictions (Durairaj & Vijitha, 2014). Predictive models have also been used in some extent as a critical component in dashboard design. Wolff et al. (2013) uses a decision tree to predict students' failure by examining the changes in user activity by inspecting students past behaviours. Through this process, the model can identify consistent patterns in students' behaviours, which might be hard for instructors to identify, and uses this information to predict if students will pass or fail the course (Wolff, Zdrahal, Nikolov, & Pantucek , 2013). The benefit of this model is that it can identify at risk students at the very early stages of the course, giving the instructor more time to help at risk students. The limitation of the model is that it requires granular data to make predictions. Elakia et al. (2014) uses C4.5 and ID3 (Sakkaf, 2020) decision tree algorithms that iteratively divide features into two or more groups at each step (Elakia, Gayathri, Aarthi, & Naren J, 2014). Although the internal operation of the decision trees is easy to understand because of its simple rule-based approach, some researchers argue that the decision trees do not always provide reliable and repeatable results (Elakia, Gayathri, Aarthi, & Naren J, 2014).

Another data mining approach that was explored was clustering. The goal of clustering is to find data points that naturally fit together, splitting the full dataset into sets of clusters. Clustering is useful as it categorizes information based on the most common categories in the dataset that are not known in advance. One of the clustering methods that can be used is optimal clustering. This method works by looking at each data point within a category to see how similar they are to the other data points and categorizes them into other clusters. An advantage of clustering is that it can be used with or without a hypothesis. This allows clustering algorithms to postulate that each data point must belong to exactly one cluster or can postulate that some points belong to more than one cluster or no cluster. Furthermore, clustering allows analyzing the goodness of fit to determine the accuracy of the clusters. This is done by assessing how well the set of clusters fit the data, relative to how much fit might be expected solely by chance, given the number of clusters, using statistical metrics such as Bayesian information (Baker, 2010).

Classification is used to categorize students based on their personal learning data, demographic data, and interaction logs. This information can then be used to identify different types of learners, as well as the characteristics of their interactions with the learning system. The results of the analysis can then be used to group students based on their performance on the course. *Classification* techniques and trends or sequence analysis are often applied to create models for analyzing student needs and provide students with recommendations and feedback about their next actions and study habits in online learning environment (Bienkowski, Feng, & Means, 2012).

For this study, researcher also explored sentiment analysis (also known as opinion mining or emotion AI), which is a sub-field of natural language processing (NLP) that attempts to identify and extract opinions within a given text across blogs, reviews, social media, forums, and news (Genç, 2019). Sentiment analysis is a task that focuses on the polarity detection and recognition of emotions towards an entity, which can be an individual, topic, and/or event (Kumar, Srinivasan, Cheng, & Zomaya, 2020). In general, sentiment analysis aims to find users' opinions, identify the sentiment they express, and then classify their polarity into positive, negative, and neutral categories (Kastrati, Dalipi, Imran, Nuci, & Wani, 2021).

Several recent studies focused on NLP-based sentiment analysis of course discussion forums. Kastrati et al. (2021) conducted a systematic literature review on deep learning-based methods for sentiment analysis in educational settings (Kastrati, Dalipi, Imran, Nuci, & Wani, 2021). This study highlighted the need for dealing with students' opinions and recognized its challenges due to the nature of the language used by the students and the large volume of information. Wen et al. (2014) explored how the mining of forum posts can be used to monitor students' trending opinions toward the course. This study also identified a correlation between sentiment ratios, measured based on daily forum posts, and the number of students who dropped out each day (Wen, Yang, & Rose, 2014).

Li and Xing (2021) found discussion boards as a valuable platform for students learning as they promote knowledge exchange (Li & Xing, 2021). This study used recurrent neural networks (RNN) and generative pre-trained transformers (GPT) to provide students with emotional and community support using contextual replies. Chaplot et al. (2015) performed sentiment analysis using an artificial neural network (ANN) to predict student attrition and showed how students' sentiment would affect their decision to dropout (Chaplot, Rhim, & Kim, 2015).

Moreno-Marcos et al. (2018) analyzed students' sentiments in forum posts and detected patterns in students' behavior to identify complex emotions such as excitement, frustration, or boredom (Moreno-Marcos, Alario-Hoyos, & Munoz-Merino, 2018). This information can help students to better understand the emotions they exhibit in a course, which allows them to better engage with their peers through increased self-awareness.

Hew et al. (2020) developed a system using a supervised machine learning algorithm for sentiment analysis and hierarchical modeling to analyze course features that affected students' course satisfaction (Hew, Hu, Qiao, & Tang, 2020). The results showed that the course instructor, content, assessments, and course schedule had a positive impact on students' course satisfaction and completion. This study used both learner-level and course-level factors to predict student satisfaction. Estrada et al. (2020) created a system that uses multiple sentiment analyses to identify learner-centered emotions (e.g., engaged, excited, or bored) through students' sentiments (i.e., positive or negative) (Estrada, Cabada, & Bustillos, 2020).

Different machine learning models have been used in educational data mining and these models have great potential to be use in intelligent dashboard design (Farahmand, Dewan, & Lin, 2021). Decision trees, K-means clustering, artificial neural network, linear discriminant analysis, naive bayes classifier, support vector machines, and various deep learning models could be used in the dashboards for students' performance prediction, progress analysis, automated feedback generation and educational decision making. For real time analysis of data and visualizing the results, these methods need to be computationally efficient and optimized by tuning different parameters (Al-Jarrah, Yoo, Muhaidat, Karagiannidis, & Taha, 2015; IGI Global, n.d.; Wang, Xu, & He, 2016).

In this study, we designed a prototype intelligent student facing learning dashboard (SF-LAD) by applying clustering and sentiment analysis models, using Student Performance and Engagement Prediction in eLearning dataset with 485 students' records (Moubayed. et. Al, 2020); as well as Stanford MOOC posts dataset with 29,000 anonymized students' discussion board posts from eleven Stanford University public self-paced online learning (SPOL) courses (Agrawal & Paepcke, n.d.) and empirically evaluated the dashboard by collecting primary data.

2.6. Learning Analytics Dashboards (LADs)

Learning Analytics Dashboards are a relatively new addition to online learning research. As such, their results need to be interpreted within the context of their limitations (Roberts, Howell, Seaman, & Gibson, 2016):

- First, the primary subject of the research could affect the outcome of the learning analytics dashboard research. For example, students from health sciences may have different experience using a learning analytics dashboard compared to students studying computer science. One reason is that health science students may have less knowledge and understanding of learning analytics dashboards in comparison to students in computer science or other similar fields. Disciplinary differences and the type and frequency of assessments may also influence how students respond to the dashboard.

- The second limitation of the existing dashboards is that students may find it difficult to conceptualize approaches of learning analytics dashboards. This could lead to students having biases towards one type of dashboard over the others. Such biases could influence the students' perception of the learning analytics dashboard, making it more difficult to implement them for online learning institution wide.
- The third limitation of the existing dashboards is that students' attitude towards the learning analytics dashboards could impede the adoption of the dashboard in institutions offering online courses. Note that research has found that there is absence of a student voice in the decision–making about the features that should be included in a learning analytics dashboard. This is because until recently, the end user of learning analytics dashboards have been academics and not the students.

This research uses a requirement analysis based on learning dashboards reviewed from the previous studies (refer to section 2.1 Background) to clearly understand the strengths and limitations of learning analytics dashboards created by other researchers. The learning from the previously designed dashboards is then used in the design of the student facing learning dashboards (SF-LAD) for this research, which aims to address the three gaps identified in the literature and create a dashboard that can be used by students from different disciplines and fields of study, minimize student bias, and promote adoption by students, allowing post-secondary institution to implement the dashboard institution wide.

The first gap that the dashboard for this research tries to address is to *design a dashboard that can be used by students from multiple disciplines.* One can look at the strengths of the Course Signals dashboard, designed by Purdue university. Course Signals is a student facing learning dashboard focused on providing student with meaningful feedback, using a predictive model. It also utilizes easy to understand visualizations, which reduces the need for students to have statistical analysis skills. This means that students in the fields such as health care studies can easily process visualizations and insight derived from the data. This allows students to use the visualizations to make more effective decisions by removing visual-spatial uncertainty from the decision-making process. It allows student to make more rational decisions based on the facts and not the assumptions (visual-spatial reasoning is a process that enables students to use visualization and cognition to derive meaning from the visual information displayed in the dashboard) (Padilla, Creem-Regehr, Hegarty, & Stefanucci, 2018).

The next issue that the dashboard for this research tries to address is to *design a dashboard that minimizes students' bias.* This is an important limitation that is necessary to overcome, as it impedes adoption of the learning dashboards by post-secondary institutions. One approach to minimizing bias toward the dashboard designed for this research is to take advantage of the GLASS dashboard designed by Leony, Pardo and Valentin (2012). In Glass, the main focus of the dashboard is visualizing students' performance. By focusing the dashboard on student performance against oneself and peers, the dashboard can provide a unified experience to all the students in the class, while providing each student with the information about their individual performance. This approach allows to anonymize the data, meaning that students only see their performance against the class, rather than the individual students. Theoretically, with this approach students find the dashboard more useful, rather than feeling they are constantly being monitored. Furthermore, by showing anonymized and aggregated data, we can create a sense of harmony among students, which will also reduce bias (Jones, 2019)

The final issue that the dashboard for this research tries to address is to improve the *adaptation/implementation of the student facing learning dashboards by post-secondary institutions.* The previously stated reasons have made it difficult for institutions to implement learning dashboards as a tool for students to improve their organizational skills, self-awareness, and motivation. To address this issue, we are going to combine the solutions to the pervious issues. First, we are going to present students with aggregated performance data to put them at ease about being watched or tracked. The theory is that this will reduce their bias towards the learning dashboards and prompt students to see it as a beneficial tool to help them be more organized and become more self-aware of their performance in class. Second, the dashboard for this research will use easy to understand visualizations to create a tool that can be used by many students in multiple fields of study. The theory is that this solution will minimize the barriers for institution wide adaption, prompting post-secondary institutions to implement learning dashboards for their online courses.

2.7. Visualization Techniques for LADs

As demonstrated, learning dashboards enable students to make incremental changes during their learning process. This will allow students and instructors to make changes to learning and teaching strategies to mitigate different learning challenges (Park & Jo, 2015). In other words, insights derived from learning dashboards (i.e., data visualization) will provide a powerful tool for online students to make data driven decisions about their education and play a more active role in governance of their own learning (Saggi & Jain, 2018).

Learning dashboards provide visualization to communicate critical information that supports achievement of the learning goals. To be effective, learning dashboards must be designed in a way that the dashboard audience can perceive what is important in a glance. Some of the information that can be communicated in an intelligent student facing learning dashboard include frequency of resource use, interactions in discussion boards and time spent (Lim, Dawson, Joksimovic, & Gašević, 2019).

The visualization techniques used in the intelligent student facing learning dashboard display the learning traces based on learning analytics data to help learners monitor and reflect on their own learning processes and progress toward learning goals (Chen, Min, Goda, Shimada, & Yamada, 2021). Visualizations are important because they have the potential to influence students to alter their goal orientations (Aguilar, Karabenick, Teasley, & Baek, 2021). By integrating learning analytics approach with the concept of learning dashboard, it become possible to provide a visual display of data, revealing patterns in students' progress and behaviors that might otherwise go undetected. Learning dashboards can help students to recognize what they have been doing and what they should do by bringing subtle information to the foreground (Chen, Min, Goda, Shimada, & Yamada, 2021).

Chapter 3. Methodology

3.1. Theoretical Framework

The current shift from traditional classrooms to online learning in higher education calls for more attention to self-regulated learning due to the autonomy associated with online learning. Self-regulated learning and performance can be described as the processes of learners activating and sustaining their own cognitions, affects, and behaviors that are oriented in a systematic way toward the achievement of personal goals (Zimmerman & Schunk, 2011). Learners create self-oriented feedback loops by setting personal goals, which requires self-awareness to monitor their effectiveness to enable them to adapt their functioning (Zimmerman & Schunk, 2011). It means that self-regulated learners must be proactive in order to set goals and engage in a self-regulatory cycle; therefore, supportive motivational beliefs are also essential for self-regulation. It is worth noting that self-regulation is not defined as an individualized form of learning. The reason is that self-regulation also includes self-initiated forms of social learning, such as seeking help from peers, coaches, and teachers (Zimmerman & Schunk, 2011).

Self-regulated learning is a set of individual strategies that students undertake in order to learn. For example, students set personal learning goals to monitor their progress towards those goals, seek help when they don't meet those goals and reflect on the learning to understand if the strategies they used to reach a particular goal were useful (Aguilar, Karabenick, Teasley, & Baek, 2021). One way students can self-regulate their studying is to monitor their achievement and study tactics so that they are well calibrated (Aguilar, Karabenick, Teasley, & Baek, 2021).

Academic self-regulation helps learners transform their mental abilities into academic skills, which has led researchers to attribute individual differences in learning to some students' inability to self-regulate. Within this framework of self-regulation, students need to have the ability to recognize their own strengths and weaknesses in learning, set and revise their learning goals accordingly, and be aware of the work it takes to achieve those goals (Chen, Min, Goda, Shimada, & Yamada, 2021). To accomplish this, students require heightened metacognition and the motivation to orient their thoughts, emotions, and behaviors to attaining goals (Chen, Min, Goda, Shimada, & Yamada, 2021).

Self-regulated learning strategies are actions and processes needed for acquiring information or skills to produce a particular effect and perception of being the instrument of that quality and effect by learners. Self-regulated learning includes both metacognition and motivation dimensions. Students need to monitor their own learning behaviors in relation to their learning goals and reflect on their effectiveness, which will improve their motivation to continue their strategies and approach (Chen, Min, Goda, Shimada, & Yamada, 2021).

Metacognitions have two dimensions: knowledge about cognition and the self-regulation of cognition (Flavell, 1979). There are three types of knowledge about cognition: declarative, procedural, and conditional knowledge. Declarative knowledge means awareness and understanding of the basic strategies; procedural knowledge means knowing how to use these strategies (e.g. skimming, underlining, summarizing, using context, or finding the main idea while reading); and conditional knowledge means knowing when and why to use particular strategies in specific conditions that influence learning goals, tasks or context (Chen, Min, Goda, Shimada, & Yamada, 2021).

Self-regulation of cognition describes three general strategy phases: *planning*, *monitoring* and *regulating*.

- During the *planning phase*, students set learning goals related to their extrinsic motivation, that is the motivation that is driven by external rewards such as score and certification, or intrinsic motivation, that is driven by inherent satisfaction of the activity such as improvement of the skills and knowledge rather than the desire for a reward or specific outcome (Chen, Min, Goda, Shimada, & Yamada, 2021).

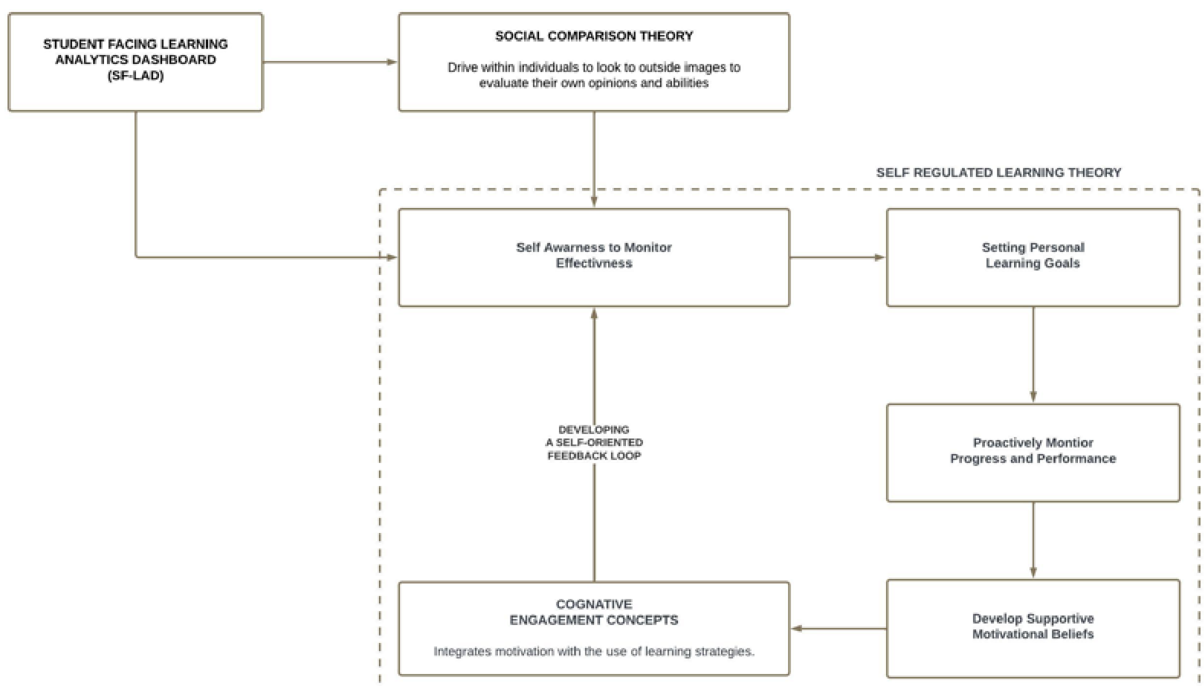
- In the *monitoring phase*, learners use different ways to track their progress; for example, assess their attention during individual lectures and learning activities, distinguish between what they knew or learned with what they don't know or haven't learned yet, and evaluate their understanding using questions of self-tests (Chen, Min, Goda, Shimada, & Yamada, 2021).
- In the *regulation phase*, students read or re-read the contents they are not familiar with, adjust their learning pace based on their learning abilities, review additional materials, or use test strategies such as skipping the questions they don't know. Regulation activities help learners determine appropriate learning or test strategies based on the self-evaluation and causal attribution of results, which can improve their learning behaviors (Chen, Min, Goda, Shimada, & Yamada, 2021).

Self-regulation doesn't happen consistently in learning and requires additional effort (da Moda, 2022). The additional effort can be motivated by visualizing student learning analytics that represent students' learning behaviour and achievement via dashboards. We can draw from social comparison theory that posits that there is a "drive within individuals to look to outside images in order to evaluate their own opinions and abilities" (Festinger, 1957, p. 16) to help online learners with self-regulation. Festinger (1957) hypothesized that the tendency to compare self with peers decreases as the difference between that person's abilities and one's own abilities become more divergent (Festinger, 1957). Therefore, the dashboard for this study can use data visualization to create awareness on how one is performing compared to class average in online environment to motivate additional effort and enhanced self-regulation.

Furthermore, this research draws from the concept of cognitive engagement to integrate motivation with the use of learning strategies (Blumenfeld, Kempler,, & Krajcik, 2006). Cognitive engagement is based on students' willingness to exert effort in their own learning by employing the required cognitive and metacognitive strategies, and making conscious effort to regulate attention, affect and effort in pursuit of their learning goals (Blumenfeld, Kempler,, & Krajcik, 2006). For example, metacognitive strategies such as setting goals, planning, monitoring, evaluating progress, and making the required adjustments. Figure 1 presents theoretical framework for this research.

Figure 1

Theoretical Framework



There is growing interest in the potential of using learning analytics dashboards (LADs) for increasing individuals' self-regulation by creating visibility into their performance in various applications. According to Jivet et al. (2017) self-regulated learning is the core theory that informs the design of learning analytics dashboards for students to make them aware of their learning process by visualizing their data. However, being aware doesn't imply taking action; therefore, self-awareness by itself is not sufficient. Dashboards should be designed with a broader objective, using awareness and reflection to enhance cognitive, behavioral or emotional competencies (Jivet, Scheffel, Drachsler, & Specht, 2017).

3.2. Application of Theoretical Framework

There are four main issues with the existing learning dashboards. First, few learning dashboards have focused on fostering awareness and reflection. Additionally, reflection and awareness are difficult to measure objectively. Second, existing learning dashboards are not designed to support students' goal setting, planning, as well as helping students manage their self-set goals. Third, learning dashboards are focused on peer comparison to measure motivation and engagement, which was accomplished by just looking at how learners connected with and were supported by their peers. Fourth, existing dashboards are focused on short-term improvements by constantly exposing students to motivational triggers that rely on social comparison to peers rather than long-term improvement by fostering mastery knowledge, acquiring skills, and developing competencies (Jivet, Scheffel, Drachsler, & Specht, 2017).

The contribution of this research is to create an intelligent student facing learning dashboard (SF-LAD), designed based on the theoretical framework for this research, to address the first three stated limitations of the existing dashboards by including the following three new features:

- *Supporting students' awareness and reflection.*
- *Supporting students with goal setting, planning and managing their self-set goals.*
- *Measuring self-regulated learning and engagement in the course.*

This will be accomplished by providing students with key performance indicators (KPIs) that support their cognitive, metacognitive, behavioral and emotional competencies while students are enrolled in the course (Jivet, Scheffel, Drachsler, & Specht, 2017). Key performance indicators (KPIs) are defined as quantifiable measures that can be used to determine to what extent the set of objectives are achieved. For example, in a student-facing dashboard, we can give students information about their *course progress*, *sentiment presented in the discussion board posts,* and *if they are active or passive learners.* This information can help students to overcome certain learning barriers by increasing self-awareness, better goal setting and increased course engagement, and improve their temporal progress in the course (Dipace, Fazlagic, & Minerva, 2019).

To foster awareness and self-reflection, the dashboard for this study provides students with insight into their self-performance and time management ability. Research has demonstrated that time management is a key component of students' self-regulation and students' overall performance in online courses. This is based on two reports, one by Piccoli, Ahmad and Ives (2001), and the other by Davies and Graff (2005), who found that login frequency is correlated to course satisfaction and participants' grades in the course ((Piccoli, Ahmad, & Ives, 2001; Davies & Graff, 2005). This means that the amount of time learners invest can be a correlative factor in learning performance and self-regulation. As a result, we can use *login frequency* as a predictor for learners' self-regulation, awarness and reflection (Kim, Jo, & Yoon, 2015).

To help students' with goal setting, planning, and managing their self-set goals, the dashboard provides students with information regarding their self-performance. This is because learning science research has identified that self-performance can be used to measure learners' *cognitive*, *metacognitive*, *affective,* and *motivational* constructs (Gašević, Dawson, & Siemens, 2015). For this reason, we can use self-performance to determine students' self-efficacy (an individual's belief in their ability to act in the ways required to reach specific goals), which in turn can be used to measure students' motivation. To achieve this, learning management system (LMS) data for the number of discussion board posts generated by students, number of discussion board posts read, communications with the course instructor, access to learning resources and how often student complete assignment on time or late will be used to automatically gauge students' self-performance. Students' self-efficacy and motivation can be used to identify successful learning strategies, and measure their overall effect on students' performance to achieve and manage their learning goals (Gašević, Dawson, & Siemens, 2015).

To measure students' engagement in a course, the student facing learning dashboard (SF-LAD) for this study focuses on students' communication between instructor/student and student/student. This is because research has found that language has a fundamental importance in the field of education as it is a conduit for the communication and understanding of information (McNamara, Allen, Crossley, Dascalu, & Perret, 2017). Learning management systems (LMS) can generate a considerable amount of information about students, including data via click-stream logs, assignments, and course performance, as well as language generated by students within discussion board forums and emails. This data can be mined using natural language processing (NLP) to better understand the depth of students' grasp of course topics. In addition, the analysis of discussion forum posts using NLP can provide insight into students' social interaction with colleagues and the instructor, which is a good indicator of their course engagement (McNamara, Allen, Crossley, Dascalu, & Perret, 2017). As a result, NLP analysis can be summarized in a dashboard to present students with information about how they are interacting with the course instructor and other students in the course (Alblawi & Alhamed, 2017). The theory is that by showing this information, we could potentially help students enrolled in online courses to feel less isolated and increase their self-awareness, which could affect their motivation and engagement level in the course.

The data captured from learning management system (LMS) about students performance, time management, and discussion forum posts are analyzed using two predictive models:

- *Natural Language Processing (NLP) model*.
- *K–mean clustering (unsupervised model)*

NLP sentiment analysis is used to predict students' engagement in the course. Visibility to their own sentiment allows students to understand their emotions in an objective way. This enable students to understand why they are struggling in the course (not engaged) and what they dislike about the courses (Altrabsheh, Gaber, & & Cocea, 2013). Analyzing the text can help students to better understand their emotions by categorizing their emotions into: anger, joy, frustration and neutral. This information is useful as it can show students patterns associated with their emotions, which could improve their engagement in the course (Altrabsheh, Gaber, & & Cocea, 2013).

K-Mean clustering model is used to identify meaningful clusters by partitioning student data into K clusters, where each observation is assigned to a cluster with its nearest mean. Based on how the clusters are formed, the students will be grouped into three clusters (Agnihotri, Aghababyan, Mojarad, Riedesel, & Essa, 2015):

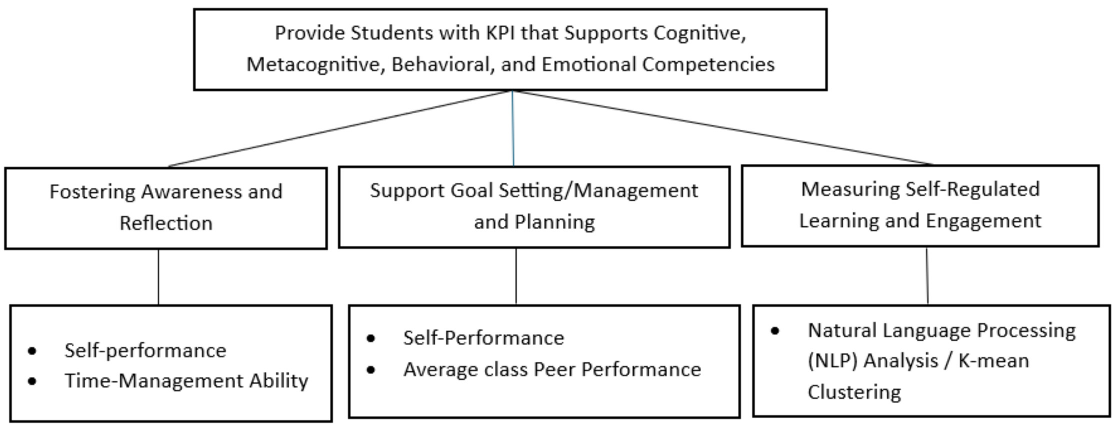
- o *Persistent Learner* (Medium scores)
- o *Regular Learner* (High score)

o *Irregular Learner* (Low scores)

Figure 2 presents application of the theoretical framework in designing the intelligent student facing dashboard (SF-LAD) for this study.

Figure 2

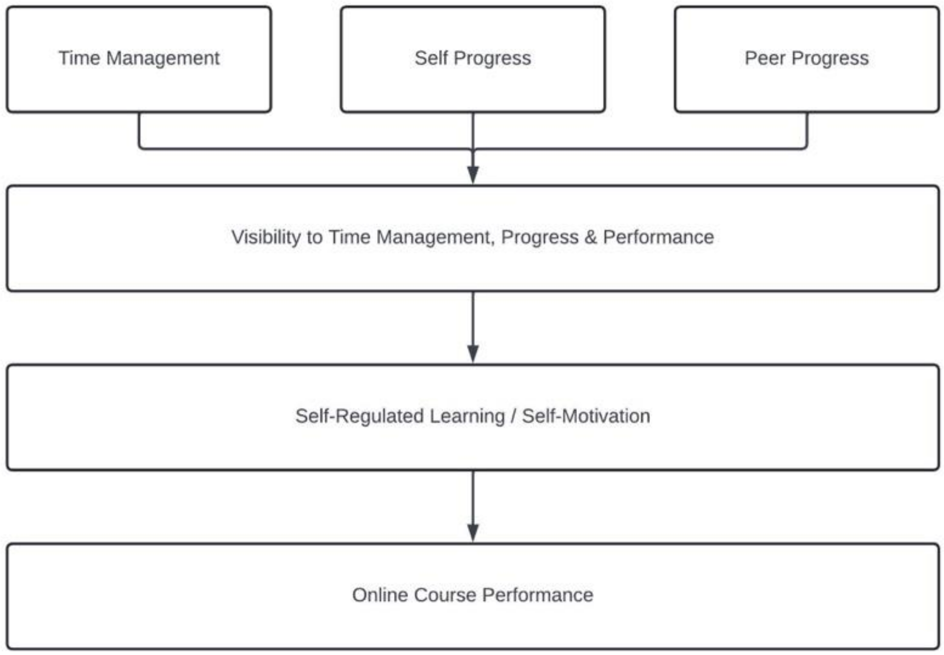
Application of Theoretical Framework for Intelligent Student Facing Dashboard for This Research



Based on the literature review, the research hypothesis is that there is a correlation between creating immediate visibility to students' time management and progress/performance (compared to both oneself and others in the class) (independent variables) and self-regulation and motivation (intermediary variables), which lead to enhanced performance (dependent variable) in the online course, as presented in Figure 3.

Figure 3

Research Hypothesis

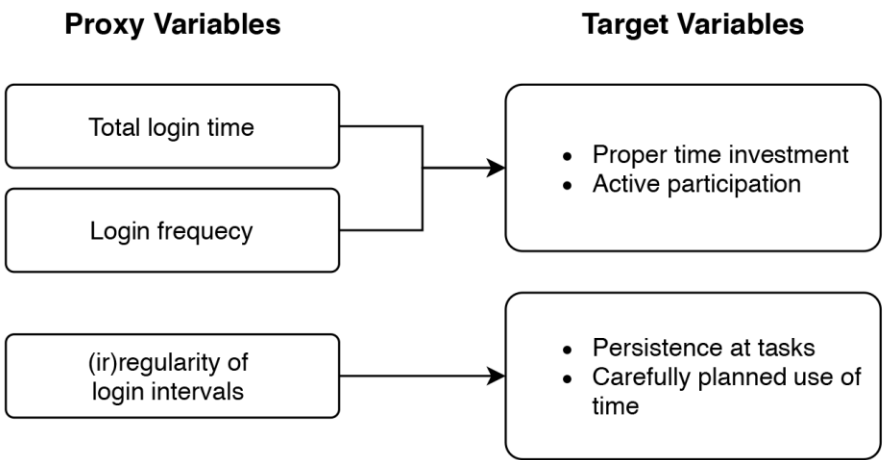


Research has shown that in online environment, learning occurs as a result of three factors: personal factors (i.e., students' beliefs), external factors (i.e., learning environment) and behavior (i.e., self–regulation) (Syahrial & Tanti, 2019). This research examines how an intelligent student facing learning dashboard (SF-LAD) could affect students' self–regulation and self-efficacy beliefs by measuring students' motivation, engagement, and academic achievements in an online learning environment (Koç, 2017).

To describe the theoretical background for selection of time management variables and time management strategy, Figure 4 depicts impact of the proxy variables (*i.e., variables that are not directly relevant, but serve in place of an unobservable or immeasurable variable*) on the targeted variables based on the work by Jo, Kim & Yoon (2014).

Figure 4

Relationship between Proxy and Targeted Time Management Variables



A key goal of designing an intelligent student-facing learning dashboard is to promote self-regulation and increase motivation in students enrolled in self-paced online learning (SPOL) courses (Farahmand , Dewan, & Lin, 2020). To accomplish this task, learning dashboards need to remove learning barriers identified for SPOL courses. Three critical learning dashboard applications that need to be considered when designing an intelligent learning dashboard are: (i) increase learning awareness; (ii) identify struggling students; and (iii) provide academic intervention (Verbert, Duval, Klerkx, Govaerts, & Santos, 2013). In SPOL, students' self-awareness of knowledge state and learning gaps can help them to reflect on learning and decide when to ask for help. Learning dashboards can achieve this objective through social comparison, goal achievement, and monitoring learning progress (Bodily, et al., 2018). Venant et al. (2016) found that social comparison with other students in online courses can be a motivating factor and could cause students to work harder, increase their course engagement and simulate a feeling of connectedness to peers (Venant, Vidal, & Broisin, 2016).

Learning dashboards need to identify struggling students in such a way that gives them enough time to seek extra support. Students get detailed information regarding the topics and concepts that are important when seeking help. This information also helps instructors to provide effective intervention (Yan, Lin, & Kinshuk, 2020). The dashboard needs to support students' sense-making from learning and teaching data. This allows students to monitor their course progress in real-time and help them to better understand their engagement and motivation level (Fazlagic, Dipace, & Minerva, 2019). To accomplish this, four main categories of indicators need to be considered: *key performance indicators, data hierarchy, dashboard design,* and *filters* (Fazlagic, Dipace, & Minerva, 2019):

- *Key performance indicators* can be defined as quantifiable measures that are used to determine to what extent students' set goals and objectives are achieved. For example, for students the goal could be to get over cognitive thresholds or the temporal progress in the course.
- *Data hierarchy* can be defined as both the structure (level) of the data and users' permission (privileges). For example, learning managers may need to have access to a dataset to enable them to explain the overall institution system performance. However, students should only be able to see their personal data such as progress in the course and the totality of their activities.
- *Dashboard design* can be defined as the appropriate way for an effective presentation and consolidation of the information. For example, is the relevant information displayed effectively on a student's mobile device?
- *Filters* can be defined as the methodology used to effectively display the relevant information. For example, showing an instructor the students' ID, or highlighting specific information in a quiz or assessment result (Fazlagic, Dipace, & Minerva, 2019).

To design an effective learning dashboard, reviewing several prior studies is essential. The review process allows developers to better understand the key performance indicators (KPIs) that students find useful and not useful. The dashboard design process starts with data ingestion from the learning management system (LMS). Learners enrolled in self-paced online learning (SPOL) courses naturally generate a considerable amount of data as they progress through different course activities. For example, clicking a hyperlink to open web resources is data about a learner's cognition and motivation based on the context of the link (Winne, 2017).

Once the data captured from the learning management system (LMS) is analyzed, using data mining, the next step is selecting the key performance indicators (KPIs) displayed to the students. KPIs can be defined as quantifiable measures that can be used to determine to what extent the set of objectives are achieved. For example, in an intelligent student-facing dashboard, we can give students information about their course progress, sentiment/level of understanding of the course topics presented in the discussion board posts, and if they are active or passive learners. This information can help students to overcome certain learning barriers or improve their temporal progress in the course (Dipace, Fazlagic, & Minerva, 2019).

Academic Institutions can exploit intelligent student facing learning dashboards to enhance self–regulated learning by (Moubayed, Injadat, Nassif, Lutfya, & Shami, 2018):

- *Enhancing course personalization* through identifying different people's learning styles and tweaking the instructional delivery of the content to meet individualized learning goals and preferences.
- *Improving recommendations provided to students* by leveraging self–assessment process.
- *Refining the delivery and transmission of content* via exploring the use of different methods for content delivery.

- *Monitoring and analyzing key performance indicators (KPIs)* such as productivity and time–management, as well as improving students' experience (i.e., adopting dashboards as a performance management system).

Students can exploit intelligent learning dashboards to better understand how they are positioned among their peers. Some of the other benefits of learning dashboards from students' perspective are (Howlin & Lynch, 2014):

- Capturing information in real time, which can give students the ability to measure their own progress more accurately. This allows students to better understand their strengths, weaknesses, and their learning preferences. They can then use these insights to make accurate predictions about their future progress.
- Allowing students to map their individual objectives to their learning path.

And giving students an overview of their personal learning preferences and style, which could help students make better decisions for the next steps in their learning path.

The goal of this research is twofold. First is to design an intelligent student facing learning dashboard (SF-LAD) that will assist students with self–regulated learning by using data visualization, predictive modeling, and natural language processing model. Second is to test the SF-LAD to assess usability of the designed SF-LAD. This study uses empirical investigation to answer the research questions, based on quantitative strategy, focusing on measuring the results through data collection from a sample group (rather than running controlled experiments) to evaluate the designed SF-LAD for this study. Data will be collected in an academic institution to evaluate the effectiveness of the dashboard. Students enrolled in online-learning courses make up the population for this study. Measurement instrument for data collection in this research is a questionnaire (System Usability Scale), using cross-sectional research design (i.e., collecting data from many different individuals at a single point in time).

In online education, data storytelling and visualizations can be used to communicate insights more effectively to students (Martinez-Maldonado, Echeverria, Nieto, & Shum, 2020). Data storytelling can be defined as an "information compression" technique to help an audience focus on what is most important to communicate via data visualizations (Echeverria, et al., 2018). The goal of data storytelling is to communicate insights through the combination of data, visuals, and narrative (Dykes, 2015). Modern businesses have been using static visualizations to support storytelling, usually in the form of diagrams and charts embedded in the body of text. In this format the text conveys the story, and the image typically provides supporting evidence (Segel & Heer, 2010). The following four principles should be used when communicating key insights to students using data storytelling and visualization in learning dashboards (Echeverria, et al., 2018): (i) have a title for visualizations that provides a succinct message intended to communicate the main message of the visualization explicitly; (ii) highlight a specific data series related to a specific key performance indicator (KPI); (iii) declutter visualizations by removing grids, eliminating indirect legend, and reducing the number of colors; and (iv) use narrative texts and labels in visualization to add context to interesting insights and minimize distraction (Duval, 2011).

Visualizations can also convert abstract and complex data into concrete and simple to understand information by amplifying human cognition. Thus, the proper visualization can lead to effective communication and correct decision-making. To effectively communicate the trends and important information, several design principals must be adopted. First, the visualization should highlight important data and trends in a way that stands out from the rest. Second, the information must be organized logically, support situated awareness, and help rapid perception by using diverse visualization techniques. Third, the information should be deployed so that the elements of the information support viewers' immediate goal for decision making (Park & Jo, 2015).

3.3. SF-LAD Design

This section describes SF-LAD design process for this research. The dashboard design draws from several prior studies. For example, research by Jo, kim and Yoon (2014) identified that students' log-in frequency, total log-in time, and log-in regularity, as well as visits on the board and repository are variables that can predict students' learning performance; and Hodges (2008) identified the correlation between students' motivation and self-awareness with their performance ((Jo, Kim, & Yoon, 2014; Hodges, 2008). Therefore, the dashboard functionalities are designed to visualize variables, which based on literature, can influence online learners' organizational skills, self-awareness, and motivation.

A rapid prototyping methodology is used for User Interface (UI) design to develop a SF-LAD that is aligned with students' needs. Rapid prototyping can be described as an iterative approach for development of the UI of software applications and involves a quick development of system mock-ups prior to building for production. The benefit of rapid prototyping is to enable validation of the UI by target users and other stakeholders prior to full-scale deployment. Rapid prototyping allows feedback to be received earlier in the development process to improve efficiency of the long-term development, as fewer updates are required after full deployment (Smith, 2021). The first version of the dashboard prototype for this research is designed for online learners (only). Figure 5 shows the dashboard prototype that was developed for this research.

Figure 5

Intelligent Student Facing Learning Dashboard (SF-LAD) Prototype

An important SF-LAD feature is time management. One of the variables used in SF-LAD time management feature is students' *Login frequency*, because we can use the data generated from how often students' login into the learning management system (LMS) to predict students' performance. Two studies found strong evidence for correlation between *login frequency* and students' satisfaction / performance. A study by Piccolo, Ahmad and Ives (2001) found a strong correlation between login frequency and course satisfaction, and a second study by Davis and Graff (2005) found evidence of a significant association between online activities and learners' grades (Jo, Kim, & Yoon, 2014).

Another variable used for SF-LAD time management is students' *total login time*, which is the total time that a learner dedicates to learning and course materials in the learning management system (LMS). The degree to which learners invest their time has been recognized as a powerful factor that is correlated with students' performance. In this research, the concept of

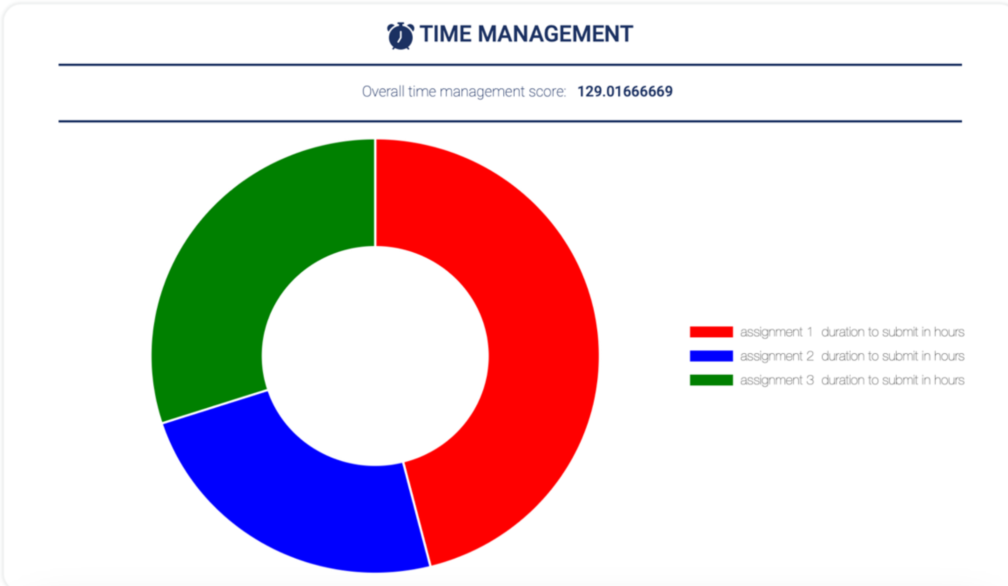
"Total Study Time" (Cotton & Savard, 1981) is adopted to support the importance of measuring the *total login time* (the total login time is used to represent the "Total Study Time"). Although the *total login time* in the LMS can hardly be regarded as a meaningful metrics to measure learning, we can hypothesize that the total login time is mainly the *time on task* by taking into account that learning activities for online courses occur in the LMS; for example, interacting with peers, watching lectures, and submitting assessments (Jo, Kim, & Yoon, 2014).

The last time management variable in SF-LAD is the *amount of time it took students to complete assessments* in the learning management system (LMS), which allows to gauge students' *planning and organization* skills (self-regulation). Self-regulation is a key concept in online learning. In recent years, research has demonstrated a positive correlation between learning success and students' self-regulation capabilities, which is important for initiating and sustaining the learning process (Schumacher & & Ifenthaler, 2018). In self–regulated learning, learners metacognitively manage cognition, motivation, and behavior to increase their chances to achieve their learning goals (Marzouk, et al., 2016).

Students will be able to see their overall time management score and how long it took them to complete each assignment in Figure 6.

Figure 6

Students' Time Management and How Long It Took Students to Complete Assignments 1, 2, 3



The next key metric that is going to be tracked for SF-LAD is students' performance. The *performance evaluation* section of the SF-LAD presents *students' performance* against *self* (over time), and against *class average* in the same course, including discussion boards, quizzes, assignments, and projects. The hypothesis is that visibility into one's performance over time can lead to motivation by developing self-efficacy beliefs. Self-efficacy can be defined as one's belief in his or her capability to adopt behaviors required to achieve their performance goals (Bandura, 1982). Rigorous research has found that self-efficacy beliefs have a strong positive correlation with learners' performance (Hodges, 2008).

For performance evaluation, data will be aggregated based on the assessments. For example, for an assignment, the aggregation time interval could be a week, a day, or a specific time, depending on the module/unit. The data can also be aggregated by dividing the course into two periods before or after mid-term. A simple example of this could be how long a student spent in course website during the week. A more complex example could be to look at the correlation between how long before an assignment deadline learner starts to work on an assignment and their grade. This will involve measuring the average time for the student between the submission date and the due date. For this complex calculation, two main factors are the *timestamp for the submission time* and the *timestamp for the student's actual start time* (Veeramachaneni, O'Reilly, & Taylor, 2014).

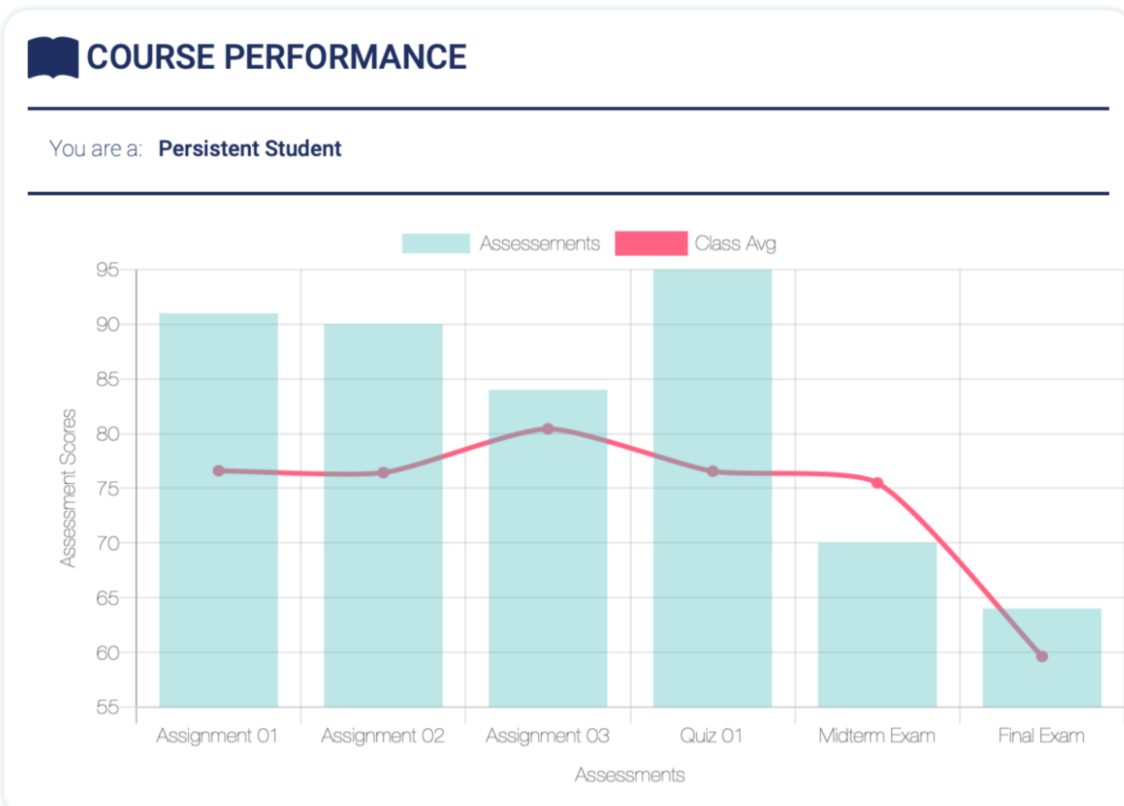
A challenge that could be encountered when implementing the metrics for student performance is how to calculate student start time of working on an assessment and present this information in the SF-LAD. One solution is to look at the timestamp of when the students first opened the assessment and the timestamp for when they first closed the assessment. Once this information is captured, we can calculate the number of minutes that have elapsed between opening and closing the assessment. The results of the calculation can then be used in a series of conditional statements, where if N number of minutes has passed between opening and closing the assessment for the first time, then the SF-LAD program begins tracking the student's progress.

In the performance section of SF-LAD, we also use k-mean clustering to identify if the student is considered a regular, ir-regular, or persistent student. The *average score*, *average number of logins*, *average number of content reads*, *average assessment lateness indicator* and *average hours to submit assignments* are used to calculate the K-Mean clusters. K-mean clustering allows the dashboard to partition the data into K clusters where each observation will be assigned to its nearest mean. The clustering process start by choosing k random observations as the initial cluster centroids. Then, each observation will be assigned to its nearest centroids by calculating Euclidean distance, which is the distance between the data points in each of the clusters. After this, a new centroid will be calculated using the data points in each cluster. We repeat this process until we have out clusters. By using K-mean clustering, the data can be represented as a set of N observations $\{x_1, x_2, x_n\}$, where each observation is a D-dimensional vector of D attributes. K-mean cluster partitions the N observations into k clusters $\{c_1, c_2, c_n\}$, where it can be minimized as $argmin \sum_{k=1}^{k} \sum_{X \epsilon c_k} \| X - \mu_k \|^2$, where μ_k is the mean of the points in c_k. This is to accommodate the scale of the dataset as it allows k-mean clustering to process the information with speed and efficiency (Agnihotri, Aghababyan, Mojarad, Riedesel, & Essa, 2015).

You can see the students' performance compared to the class average for each assessment in Figure 7.

Figure 7

Display Students' Performance in the Intelligent Student Facing Learning Dashboard (SF-LAD)



The next component of SF-LAD is capturing data from LMS, such as Moodle or D2L (Desire to Learn by Brigthspace). For example, to capture the learning data that we need from

Moodle, we can use its Experience API (XAPI), which is an open data specification for granular, in-the-moment data collection across the learning tool. This approach by Moodle allows to specify exactly the data that we need from LMS, which will increase the efficiency and accuracy of the SF-LAD. The XAPI enables us to capture information about the course content, course components, login information and other resources used by learners when they are logged into the LMS (eThink, 2017). You can see the students' sentiment throughout the course in Figure 8.

Figure 8

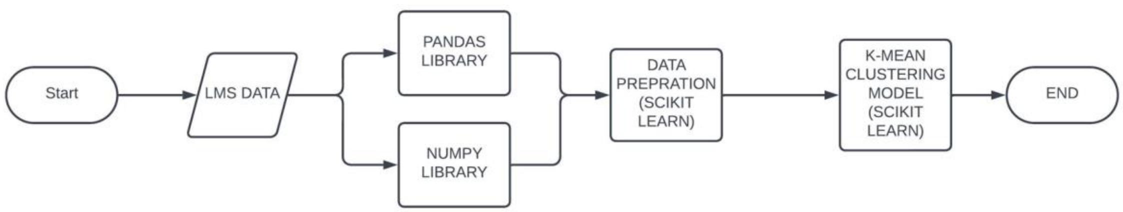
Students' Sentiment for the Duration of the Course

3.4. Predictive Modeling

This section describes methodology for predictive modeling for SF-LAD design for this study. We explored different predictive models for the dashboard and adopted K–mean clustering because it is efficient and easy to implement compared to other models. K–mean clustering results are also easy to analyze and interpret. We used the Python's *pandas* and *NumPy* libraries to process the data and the *SciKit* library for data analysis, using K-means clustering (Macarini, Cechinel, Machado, Ramos, & Munoz, 2019). Figure 9 shows the overview of the data processing and data analysis.

Figure 9

K-mean Clustering Model Methodology Overview (Macarini, Cechinel, Machado, Ramos, & Munoz, 2019)



The purpose of using K-mean clustering is to group the students into one of the three categories: *Persistent, Regular, or Irregular* based on *the Average Score, Average Number of Logins, Average Number of Contents Read*, *Average Assessment Lateness Indicator*, and *Average Hours Took to Submit Assessments* (refer to Table 1).

Table 1

Students Grouping Based on Cluster Statistics

	Average Score	Average Number of Logins	Average Number of Content Reads	Average Assessment Lateness Indicator	Average Hours to Submit Assignments
Persistent Student	Medium	Medium	Medium	Low	Medium
Regular Student	High	High	High	Low	Low
Irregular Student	Low	Low	Low	High	High

We used *Average Score* for each cluster because a study by Conijn, Beemt, & Cuijpers (2017) found evidence that students with a high final grade generally have higher performance and are more likely to complete the course (Conijn, Van den Beemt, & Cuijpers, 2017). We used *Average Number of Logins* because a study by Piccolo, Ahmad, and Ives (2001) found a strong correlation between login frequency and course satisfaction (Piccoli, Ahmad, & Ives, 2001), and a second study by Davis and Graff (2005) found evidence of a significant association between online activities and learners' grades (Davies & Graff, 2005; Jo, Kim, & Yoon, 2014). We used *Average Number of Contents Read* because according to a study by Huang and Hew (2016), there is a correlation between motivation and self-regulation of students and the number of instructional materials they read (Huang & Hew, 2016). We used *the Time it Takes Students (in Hours) to Complete Assessments*, and *the Delay in the Completion of the Assessments* because a study by Moubayed et al. (2018) found an association between the time it takes students to complete assessments and the delay in the submission of the assessments with students' level of engagement (Moubayed, Injadat, Shami, & Lutfiyya, 2018).

The clustering results allows us to identify and examine a particular set of study behaviors. By analyzing these behaviors, student facing learning dashboard (SF-LAD) provides students with the ability to self-monitor and make micro-adjustments as required to ensure the achievement of their academic goals.

3.5. Natural Language Processing Model

This section describes methodology for Sentiment Analysis in SF-LAD designed for this study. We explored different sentiment analysis models and decided to use XLNet model. The XLNet architecture for sentiment classification starts with the language model training. The first component of the language model is a word-embedding matrix where a fixed-length vector is assigned for each token (word) in the vocabulary and the sequence is converted to a set of vectors (McGoldrick, Cao, & Prince, 2019). Next, researchers need to relate the embedded tokens in a sequence. In XLNet this is achieved with transformers (McGoldrick, Cao, & Prince, 2019). A transformer is a deep learning model that adopts the mechanism of self-attention, differentially weighting the significance of each part of the input data (Wikipedia, n.d.). In XLNet training, the objective of the model is to learn the conditional distribution of all permutations of the tokens in a sequence. XLNet accomplishes this by using samples from all possible permutations without needing to see every single relation (McGoldrick, Cao, & Prince, 2019). A premutation is a language model that is trained to predict one token, given the preceding context; unlike the traditional model where tokens are predicted in sequential order, premutation predicts tokens in some random order (Karan , 2019). The way this works is that given sequence X, an AR model calculates the probability $\Pr(X_i|X_{<i})$. Here researchers calculate the probability of a token X_i in

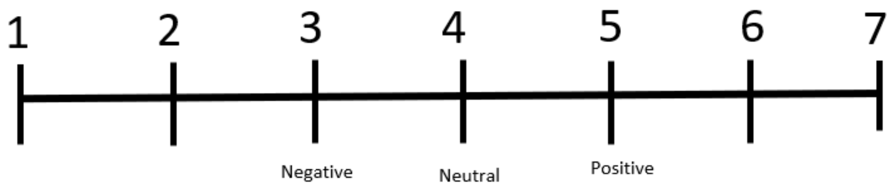
the sentence, conditioned on the tokens $X_{<i}$ preceding it. These conditioning words provide the context for the model. From this, the AR model learns the relationship between tokens (McGoldrick, Cao, & Prince, 2019). Next, researchers use masking as a method to perform word prediction by intentionally hiding certain words in a sentence (Verma, 2022). This process works by replacing certain tokens (words) with a generic mask token, then asking the model to recover the originals (McGoldrick, Cao, & Prince, 2019). This allows the model to incorporate the context of the tokens both to the left and right of the word being predicted to get the best prediction (McGoldrick, Cao, & Prince, 2019).

The purpose of the XLNet-based sentiment analysis model for this study is to categorize students' discussion board posts as *negative, neutral, and positive* to discover insights about how students' sentiment affects their course engagement over time. Although the application of sentiment analysis in educational research is limited, its use is important to identify students' opinion over time and allows educators and students to reflect on teaching and studying strategies and make changes where required (Nkomo & Daniel, 2021).

Sentiment classifiers are often used for binary classification (just positive or negative sentiment). In this study, the sentiment analysis focuses on the polarity of a text (*positive, negative, and neutral*) and uses three discreet classes - between 1 to 7 – to categorize sentiment (refer to Figure 10). As demonstrated, we categorize students' sentiments as *negative*, *neutral*, or *positive*. The sentiment intensities often vary because of the subtleties of human language. In the model, if the polarity score is less than or equal to 3.5 then the sentiment is negative. If the polarity score is greater than 3.5 but less than 4.5 then the sentiment is neutral. If the polarity score is greater than or equal to 4.5 then the sentiment is positive (Rao , 2019).

Figure 10

The Class Labels for Sentiment Classification



Now that we have introduced the research methodology, we proceed with presenting experiment design, data collection, research instrument, and data analysis approach for this study in the next chapter.

Chapter 4. Experiment

This chapter starts by describing experiment set up to evaluate the usability of the intelligent student facing learning dashboard (SF-LAD) designed for this research and data collection approach (including sampling for SF-LAD evaluation), followed by introducing research instrument (System Usability Scale – SUS). Then data analysis approach will be described, followed by reliability and validity. Finally, ethics consideration is delineated.

4.1. Experiment Design

Following development of the rapid prototype as the first version of the intelligent student facing learning dashboard (SF-LAD), the dashboard was evaluated by administering a survey to online students to test and measure the effectiveness of the dashboard in terms of usability (quality of the user experience). A sample of online students were asked to interact with the dashboard and fill in a survey and answer a ten item questionnaire to determine the intention to adopt the tool by assessing the usability of the dashboard. Dashboard usability was assessed by calculating usability score for each participant based on their responses to ten item questionnaire and calculating the average usability score for the dashboard. The average usability score was compared with all the systems in the database to interpret the usability, using five different methods to interpret the score (i.e., percentile method, letter grade method, adjective method, acceptability method, and promoter/detractor method). In addition to comparing the usability score with all the systems in the database for interpreting the test results for usability, this study consulted literature to gather insight about usability of the other similar educational technologies and used the average usability score for 77 surveys of internet platform educational technologies, using the same ten item usability questionnaire, as a benchmark for verifying the interpretation of the dashboard usability score. Furthermore, the collected demographical data was used to visualize usability score for each participant to identify potential patterns and trends in data to potentially identify specific groups of online learners that can mostly benefit from the designed intelligent dashboard.

Moreover, two additional questions were included in the survey to verify research hypothesis and one additional question to gather survey participants' feedback and comments on how to improve the dashboard. The hypothesis verification questions were analyzed, using descriptive statistics and the patterns were visualized to draw insight. Also, detail feedback and comments by the survey participants were categorized and analyzed to identify areas for improvement for the dashboard in the next iteration.

4.2. Data collection

For empirical investigation to test the usability of the intelligent student facing learning dashboard (SF-LAD) designed for this research, data is collected from a sample group in an academic institution. The goal is to assess the effectiveness of the dashboard and evaluate online learners' intention to adopt this tool.

Students enrolled in online courses make up the population for this study.

The participants for this study are students enrolled in Athabasca university; specifically, students enrolled in COMP200 and COMP 214 courses. Approximately 460 students (380 students enrolled in COMP 200 course and 80 students enrolled in COMP 214 course) received an invitation to participate in the research after their first assessment. This is because during the initial evaluation of the study, it was determined that students enrolled in COMP200 and COMP 214 are well suited to evaluate the effectiveness of the dashboard designed for this research. The invitations were posted in the course discussion board by an administrator in the School of Computing and Information Systems (SCIS), who helped to distribute recruitment letter to the potential participants. The course teachers were not involved in the recruitment process.

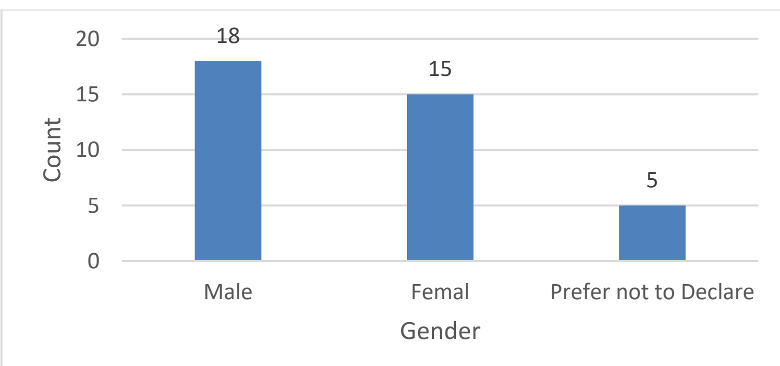
The more students participate in the survey, the better; however, with too few students it will be hard to validate any results. The sample size of 30 participants was selected as the minimum number of survey participant as it is large enough to provide meaningful results and is achievable. After administering the survey from the 460 students enrolled in COMP 200 and COMP 214, 38 participants completed the survey that is the sample size for this study. The first 30 participants who completed the survey received a $15 Amazon gift certificate as a token of appreciation for their participation.

During the survey, participants were instructed to login to the student facing learning dashboard using the account information provided in the survey invitation. Once participants were logged in, they had the ability to interact with the dashboard.

The survey participants consisted of 18 male students, and 15 female students and 5 students who preferred not to declare (refer to Figure 11).

Figure 11

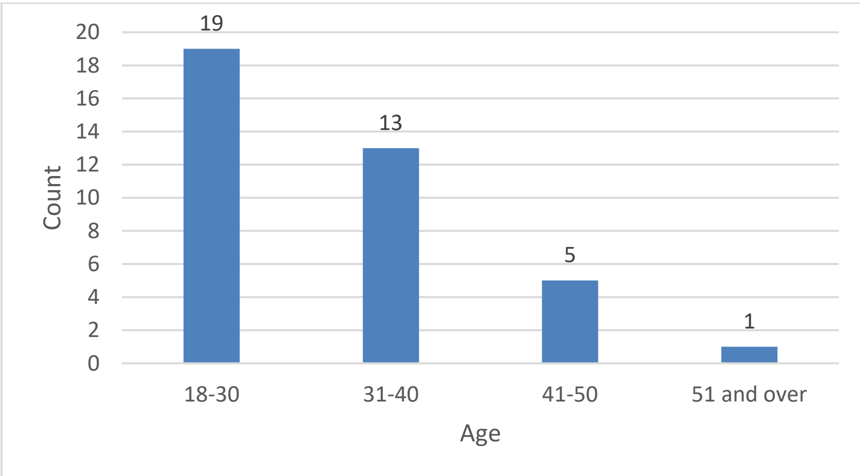
Research Sample



The age range for participants is classified in categories 18-30, 31-40, 41-50, and 51 and above. Figure 12 presents the distribution of age ranges for survey participants.

Figure 12

Research Sample Age Distribution



Demographic information is important because it allows the researchers to better understand how different genders and age ranges understand and use the dashboard. The demographic information allows us to ensure that survey was fair and representative of the target population. It also provides context as to how different groups will use and adopt the student facing dashboard designed for this research (Tankala, 2022).

The measurement instrument for data collection in this research is a questionnaire (System Usability Scale - SUS), using cross-sectional research design (i.e., analyzing the collected data from the sample at a single point in time). System Usability Scale (SUS), its benefits, challenges, considerations when using SUS, delimitations, and the questionnaire are described in this section.

The System Usability Scale (SUS) is a questionnaire based method to assess the usability of a system or product through a series of standardized questions, and in today's industry, it's considered an indispensable component of user experience (UX) (Soegaard, 2024). System usability scale uses a simple ten item scale to provide subjective assessment of usability. There are five responses for each question in the scale, ranging from Strongly agree to strongly disagree. The scores range between *1 – 100* (Klug, 2017). Scores are affected by the complexity of both the system and tasks that survey participants may perform before taking the SUS survey (Klug, 2017).

System Usability Scale (SUS) was designed by John Brook in 1986 and provides an efficient method for measuring the usability of a variety of products (Klug, 2017). It is a broadly used tool to measure system usability and can be quickly applied and interpreted (Blattgerste, Behrends, & Pfeiffer, 2022). It has been used extensively by different industries to test a range of applications including hardware, software, mobile devices, and web applications (Klug, 2017). Originally, SUS questionnaire was considered as a one-dimensional "quick and dirty" approach for measuring usability, however, the research since 1986 has revealed that SUS provides helpful insights and dimensions to contextualize and compare each SUS score (Blattgerste, Behrends, & Pfeiffer, 2022) and it is now considered the industry accepted standard for measuring the overall user-friendliness of a system (Soegaard, 2024). With the validation efforts of SUS for various languages and recent application of SUS in new contexts, SUS is considered a fairly quick but apparently not dirty approach to measure usability of a system (Blattgerste, Behrends, & Pfeiffer, 2022).

Here're the list of questions for System Usability Scale Questionnaire for the intelligent student facing learning dashboard (SF-LAD) designed for this research:

1. I think I would like to use the student facing learning dashboard frequently.
2. I found the dashboard unnecessarily complex.
3. I thought the dashboard was easy to use.
4. I think that I would need the support of a technical person to be able to use the student facing learning dashboard.

5. I found the visualizations in this dashboard appropriate for monitoring students' self-regulation.

6. I found that there was too much inconsistency in the dashboard for monitoring students' self-regulation.

7. I would imagine that most people would learn to use this dashboard very quickly.

8. I found the dashboard cumbersome to use.

9. I felt very confident using the dashboard.

10. I need to learn a lot of things before I could get going with this dashboard.

System Usability Scale (SUS) provides several potential benefits to researchers and system usability testers. SUS has been in use for over 30 years and is a reliable, tested tool for evaluating a wide range of systems and applications (Klug, 2017).

SUS provides developers and product managers real-time feedback along with a number of other benefits and is pivotal in ensuring that software and interactive systems are aligned with users' needs and expectations. The emphasis is on the user's perspective, making it an integral part of user-centered design and evaluation (Soegaard, 2024). SUS employs a standardized questionnaire and numerical rating system that facilitates quantitative analysis. The data allows for direct comparisons of the usability of different systems, tracks improvement over time, and sets measurable usability benchmarks (Soegaard, 2024).

SUS is famous for its simplicity and its exact use of words. It consists of just ten questions, enabling quick administration, rendering it an efficient approach for collecting usability data (Soegaard, 2024). The speed at which SUS can be administered enable iterative design process to run smoothly. It allows for frequent usability assessments and rapid adjustments to users' needs and expectations (Soegaard, 2024).

SUS uses a common language for communication between designers, developers and users. It’s straightforward and numerical results make it easier for all stakeholders to discuss and address usability issues, fostering a collaborative atmosphere for ongoing system refinement (Soegaard, 2024).

The standardized nature of SUS ensures consistency in usability evaluation across different systems over time. This consistency is important for tracking progress and making objectives comparisons, which ultimately leads to more reliable insights and decision-making (Soegaard, 2024).

The concise and user-friendly format of SUS encourages users to provide feedback on their experiences and thus helping the organization collect valuable insights directly from their target audience (Soegaard, 2024). The data collected using SUS can assist in allocation of resources by pinpointing specific areas of software system or functionality that requires attention. This ensures that limited resources are used for maximum impact (Soegaard, 2024).

There are implementation challenges with using System Usability Scale (SUS). One of the challenges is SUS grading, which is a complex process and includes normalization for calculated grades to produce a percentage (Klug, 2017). Another possible issue is that questions in the SUS scale alternates between statements that are positive and negative statements, which can lead to confusion for some participants when they complete the scale. “Odd” number questions on the scale (i.e., questions 1, 3, 5, 7, and 9) are worded positively, therefore number “5” is the highest score for “Odd” questions, and “Even” number questions on the scale (i.e., questions 2, 4, 6, 8, and 10) are worded negatively, therefore number “1” is the highest score for “Even” questions. So, it is important that SUS scoring be executed correctly to prevent errors (Klug, 2017).

As noted, SUS is comprised of ten question items. One way for enhancing SUS is to add more question items to gauge additional aspects of usability, such as speed, task efficiency, accessibility, mobile responsiveness, and learnability (Brown, 2023). For this study, the original ten questions in SUS were customized for the intelligent student facing learning dashboard (SF-LAD) to assess the usability of the current design for online learners, and two additional questions were added to verify the research hypothesis. The first additional question is added to verify if the SF-LAD can improve self-reflection (i.e., the activity of thinking about your own feelings and behaviour, and the reason that may lie behind them). The second additional question is if SF-LAD improves course engagement and performance. Note that for the purpose of SUS data analysis, to maintain the integrity of the analysis and scoring for SF-LAD usability, only the original ten customized questions were used for calculating the SUS score. The results from the additional two questions were analyzed separately to verify the research hypothesis.

Another way to enhance SUS is to add open-ended questions to the questionnaire to gather qualitative insights. By prompting the users to provide specific feedback and suggestions for enhancement of the system usability, the system under evaluation can be improved by receiving deeper insights about the system users' perceptions (Brown, 2023). For this study, participants were prompted to provide comments on how the intelligent student facing learning dashboard (SF-LAD) can be improved to gather qualitative insight from the participants. The answers to this qualitative question are analyzed to identify areas for improvement for the SF-LAD designed for this study.

Considering the popularity of System Usability Scale (SUS) and its fitness for analyzing the perceived usability of systems, products, or services, it is crucial to understand that it has a number of drawbacks. When using SUS as part of usability tests, the following limitation should be considered (Question Pro, n.d.):

- SUS can not provide accurate information on a product's weakness.
- It is not possible to compare two systems and their functionality using SUS systematically.
- SUS results depend on each participant's subjective response.
- Scores and calculations for SUS can confuse some users.

4.3. Data Analysis

Before we describe how to interpret System Usability Scale (SUS) scores, we clarify how to calculate SUS score for each participant, which is the first step in SUS data analysis (UIUX Trend, n.d.):

Step 1: The first step in data analysis is to convert the scale into numbers for each one of the 10 questions (Strongly Disagree=1, Disagree=2, Neutral=3, Agree=4, Strongly Agree=5)

Step 2: The second step is to calculate the SUS score for each participant:

SUS Score = (X + Y) * 2.5

X = "Sum of the points for all odd-number questions" – 5

Y = 25 – "Sum of the points for all even-numbered questions"

The rationale behind the calculation is that the total score is 100 and each question has a weight of 10 points. Since all the Odd-numbered questions are in positive tone, if the response is Strongly Agree the maximum points will be given, which is 10 for each question (after normalization). If the response is Strongly Disagree the minimum points will be given, which is 0 (after normalization). To achieve that you need to shape your data for Odd numbered questions to be scored 0 – 4 (instead of 1 – 5) by subtracting 1 from each Odd numbered question to ensure that the minimum point is 0. In other words, by subtracting 1 from your numbered scale, you change the points for each response as follows (UIUX Trend, n.d.):

Strongly Disagree:	1 – 1 = 0
Disagree:	2 – 1 = 1
Neutral:	3 – 1 = 2
Agree:	4 – 1 = 3
Strongly Agree:	5 – 1 = 4

In contrast, since the Even-numbered questions are in negative tone, if the response is Strongly Agree the minimum points will be given, which is 0. If the response is Strongly Disagree the maximum points will be given, which is 10 (after normalization). So, the responses will be shaped as follows (UIUX Trend, n.d.):

Strongly Disagree:	5 = 4
Disagree:	4 = 3
Neutral:	3 = 2
Agree:	2 = 1
Strongly Agree:	1 = 0

Step 3: is normalizing SUS Score. It is important to recognize that SUS scores are not percentages, and it is required to be normalized to produce a percentile ranking (Klug, 2017). We normalize the shaped data (score between 0 – 4) by multiplying the score by 2.5 to generate a total SUS score (*for all 10 items*) between 0 -100. This will ensure that the maximum is 10 for each question (Soegaard, 2024).

A SUS score ranges from 0 – 100 with higher scores indicating better usability for the system (Soegaard, 2024). SUS can be difficult to interpret by itself; for example, what does SUS score of 50 average is considered (good or not)? Fortunately, over 30 years of experience with SUS has provided researchers with significant volume of data from more than 10,000 responses for hundreds of systems that can be used for interpretation of SUS scores (Sauro, 5 Ways to Interpret a SUS Score, 2018). Currently, there are at least five different ways to interpret SUS scores, trying to contextualize what a specific SUS score actually means (Blattgerste, Behrends, & Pfeiffer, 2022). These methods are Percentiles, Grades, Adjectives, Acceptability, and Promoters and Detractors (Sauro, 5 Ways to Interpret a SUS Score, 2018).

4.3.1. Percentile Method

The first method is using percentile ranking. SUS scores range from 0 to 100; however, it is important to note that SUS scores are not percentages. A SUS score of 60 is not necessarily 10% better than a SUS score of 50. Instead, the SUS score should be considered as a percentile rank. The percentile rank is not linear; it follows a ‘S’ curve. Therefore, increasing SUS score from 5 to 20 or from 90 to 95 will hardly raises the percentile rank (Lawson, 2022).

Percentiles help to understand how a certain SUS score compares to the larger dataset. For example, SUS score of 80 means that it falls at the 80th percentile, which implies that this system outperforms 80% of the scores in the database (Soegaard, 2024).

The average score (at the 50th percentile) is 68. That means a raw SUS above 68 is above average and below 68 is below average. For example, SUS score of 75 is at the 73rd percentile (scoring better than 73% of the scores in the database); and a raw SUS score of 52 falls at the 15th percentile (scoring worse than 85% of the scores in the database) (Sauro, 5 Ways to Interpret a SUS Score, 2018).

SUS score is presented as percentage, ranging from 0 to 100. In this context a perfect score of 100 indicates flawless usability and exceptional user experience. Based on the previous studies that have used SUS, scores between 68 and 70 indicate a decent level of usability and any score below 50 indicated potential for serious deficiency in usability and suggests a need for substantial improvement to enhance user satisfaction (Sauro, 2018). For accessing S curve for SUS score (curve with percentile ranking), refer to Sauro (2018).

Analysis of more than 5,000 user scores from almost 500 studies across various types of applications attests that the average SUS score is 68 with a standard deviation of 12.5 (Sauro, 2011a). A SUS score of 68 is in the 50th percentile (i.e., the score is higher than 50 percent of all tested systems and applications in the database) (Sauro, 2011a). Table 2 shows SUS Score vs Percentile range (Sauro, 2018).

Table 2

SUS Scores Interpretation Using Percentile

SUS Score	Percentile Range
84.1 - 100	96-100
80.8 – 84	90-95
78.9 – 80.7	85-89
77.2 – 78.8	80-84
74.1 – 77.1	70-79
72.6 - 74	65-69
71.1 – 72.5	60-64
65 - 71	41-59
62.7 – 64.9	35-40
51.7 – 62.6	15-34
0 – 51.6	0-14

4.3.2. Letter Grades Method

Another way to interpret SUS scores is by using letter grades, which can be easy to understand for stakeholders. Since the average SUS score is 68, it is preferred to grade on a curve in which a SUS score of 68 is at the center of the range for a ‘C’. Therefore, as shown in Table 3, the full “C” range is from 65.0–71.0. Also, a SUS score of 78.9 or above would constitute an “A-”, and a SUS score of 51.6 or below would constitute an “F” (Klug, 2017).

Interpretation of SUS scores using letter grade closely relates to percentile interpretation of SUS scores. For example, Grade A indicates “superior” performance, grade C indicates “average” performance, and grade F indicates “failing” performance. Table 3 presents SUS scores interpretation based on letter grade and percentile methods. (Sauro, 2018).

Table 3

SUS Scores Interpretation Using Letter Grade and Percentile Methods

Grade	SUS Score	Percentile Range
A+	84.1 - 100	96-100
A	80.8 – 84	90-95
A-	78.9 – 80.7	85-89
B+	77.2 – 78.8	80-84
B	74.1 – 77.1	70-79
B-	72.6 - 74	65-69
C+	71.1 – 72.5	60-64
C	65 - 71	41-59
C-	62.7 – 64.9	35-40
D	51.7 – 62.6	15-34
F	0 – 51.6	0-14

4.3.3. Adjectives Method

Another way to interpret SUS scores is by using adjectives (words) instead of numbers to describe users’ experience with the system (Sauro, 2018). A benefit of using an adjective scale is that it makes it easy to convey interpretation of the mean SUS score to others involved in the study (Bangor, Kortum, & Miller, 2009).

Bangor, Kortum and Miller (2009) conducted an analysis to determine, using the adjective rating scale, how well the ratings matched the corresponding SUS score via individual ratings by the participants (*i.e. they added a seven-point adjective-anchored Likert scale questions to SUS*

Survey of 964 participants as an eleventh question with adjectives: Worst Imaginable, Awful, Poor, OK, Good, Excellent, Best Imaginable). They found highly significant correlation ($\propto < 0.01$ *level of significance that presents the probability of obtaining the results by chance and r = 0.822 Pearson correlation coefficient that measures the strength of the relationship between two variables*) (Bangor, Kortum, & Miller, 2009).

The finding that adjective rating scale very closely matches the SUS scale attest that it helps to interpret mean SUS score for individual studies to effectively communicate perceived usability of the system as a subjective label to the target audience (Bangor, Kortum, & Miller, 2009). Table 4 presents SUS scores interpretation based on adjective label, letter grade and percentile methods (Sauro, 2018).

Table 4

SUS Scores Interpretation Using Adjective Label, Letter Grade and Percentile Methods

Grade	SUS Score	Percentile Range	Adjective
A+	84.1 - 100	96-100	Best Imaginable
A	80.8 – 84	90-95	Excellent
A-	78.9 – 80.7	85-89	Good
B+	77.2 – 78.8	80-84	Good
B	74.1 – 77.1	70-79	Good
B-	72.6 - 74	65-69	Good
C+	71.1 – 72.5	60-64	Good
C	65 - 71	41-59	OK
C-	62.7 – 64.9	35-40	OK
D	51.7 – 62.6	15-34	OK
F	25.1 – 51.6	2-14	Poor
F	0 - 25	0-1.9	Worst Imaginable

4.3.4. Acceptability Method

Another method for using words to describe SUS is in terms of what’s acceptable and not acceptable. These terms can be used when SUS score is well above average or well below average.

Bangor, Kortum and Miller (2008) assigned "Acceptable" to SUS scores roughly above the average score of 68 (with letter grades A and B) and "Unacceptable" to SUS scores lower than 50 (with a letter grade F). They assigned "Marginally Acceptable" to SUS scores between 50 – 68 (with letter grades C and D). Table 5 presents SUS scores interpretation based on acceptability, adjective label, letter grade and percentile methods (Sauro, 2018).

Table 5

SUS Scores Interpretation Using Acceptability, Adjective Label, Letter Grade and Percentile Methods

Grade	SUS Score	Percentile Range	Adjective	Acceptability
A+	84.1 - 100	96-100	Best Imaginable	Acceptable
A	80.8 – 84	90-95	Excellent	Acceptable
A-	78.9 – 80.7	85-89	Good	Acceptable
B+	77.2 – 78.8	80-84	Good	Acceptable
B	74.1 – 77.1	70-79	Good	Acceptable
B-	72.6 - 74	65-69	Good	Acceptable
C+	71.1 – 72.5	60-64	Good	Acceptable
C	65 - 71	41-59	OK	Marginally Acceptable
C-	62.7 – 64.9	35-40	OK	Marginally Acceptable
D	51.7 – 62.6	15-34	OK	Marginally Acceptable
F	25.1 – 51.6	2-14	Poor	Not Acceptable
F	0 - 25	0-1.9	Worst Imaginable	Not Acceptable

4.3.5. Promoters and Detractors Method

Another way to interpret SUS scores is using Net Present Score (NPS). In 2003 Fred Reichheld introduced the Net Promoter Score (NPS) that uses a single question ("*How likely is it that you would recommend our company to a friend or colleague?"*) with 11 scale steps from 0 (*not at all likely*) to 10 (*Extremely likely*), which has become a popular metrics for customer loyalty because it identifies the likelihood if someone would recommend a product (Sauro, 2012).

"Promoters" are respondents who select a 9 or 10, "Detractors" are those selecting 0 through 6, and "Passives" are all the others. The "percentage of Promoters" minus the "percentage of Detractors" is the NPS from a survey. For example, if you've collected 100 "*LTR" (stands for likely to recommend)* ratings for a company for which 25 ratings fall between 0 and 6 (25% Detractors), 25 fall between 7 and 8 (25% Passives), and 50 fall between 9 and 10 (50% Promoters). The resulting NPS is the percentage of Promoters minus the percentage of Detractors, which in this case is: 50% - 25% = 25% (Sauro, 2012).

Net Promoter Score (NPS) is a metric that is simple to understand by managers, and they can use NPS to track improvements over time because NPS improvements have a strong relationship to company growth. NPS can also help managers when it is compared to industry benchmarks (Sauro, 2012).

When the SUS score is divided by 10, it roughly matches the LTR (*likely to recommend*) rating. Analysis indicates that the SUS scores can explain about 39% of the variation in Likelihood to Recommend (LTR) ratings, which corresponds to a statistically significant correlation (about .623) between SUS and LTR (Sauro, 2012).

Table 6 presents SUS scores interpretation based on promoters and detractors, acceptability, adjective label, letter grade and percentile methods (Sauro, 2018).

Table 6

SUS Scores Interpretation Using Promoters and Detractors, Acceptability, Adjective Label, Letter Grade and Percentile Methods

Grade	SUS Score	Percentile Range	Adjective	Acceptability	NPS
A+	84.1 - 100	96-100	Best Imaginable	Acceptable	Promoter
A	80.8 – 84	90-95	Excellent	Acceptable	Promoter
A-	78.9 – 80.7	85-89	Good	Acceptable	Promoter
B+	77.2 – 78.8	80-84	Good	Acceptable	Passive
B	74.1 – 77.1	70-79	Good	Acceptable	Passive
B-	72.6 - 74	65-69	Good	Acceptable	Passive
C+	71.1 – 72.5	60-64	Good	Acceptable	Passive
C	65 - 71	41-59	OK	Marginally Acceptable	Passive
C-	62.7 – 64.9	35-40	OK	Marginally Acceptable	Passive
D	51.7 – 62.6	15-34	OK	Marginally Acceptable	Detractor
F	25.1 – 51.6	2-14	Poor	Not Acceptable	Detractor
F	0 - 25	0-1.9	Worst Imaginable	Not Acceptable	Detractor

4.4. Reliability and validity

The accuracy and adequacy of our measurement procedures are evaluated in scientific research by “reliability” and “validity” (jointly called the “psychometric properties” of measurement scales) (Wellington & Szczerbinski, 2007). In this section, reliability and validity of measurement scales are described and applied to the measurement instrument for this study (System Usability Scale - SUS) to determine accuracy and adequacy of the measurement procedure for this research.

4.4.1. Reliability

Reliability helps to determine if we use this scale to measure the same construct multiple times, assuming the underlying phenomenon is not changing, do we get almost the same result every time? (Wellington & Szczerbinski, 2007). Therefore, the reliability of a measurement instrument refers to the extent to which it consistently measures concepts.

Cronbach's Alpha is one way to measure the strength of the consistency (University of Virgina, n.d.). In this study Cronbach Alpha is used to assess the reliability of the scale. Cronbach's alpha coefficient measures the reliability (internal consistency) of a scale (i.e., set of survey items). Using Cronbach's alpha coefficient helps determine whether a scale (i.e., collection of items) consistently measures the same characteristic. Cronbach's alpha quantifies the level of agreement between collection of items on a standardized 0 to 1 scale. Cronbach's alpha coefficient 0 indicates that there is absolutely no correlation between the items, which means the items are completely independent. Therefore, knowing the response to a question doesn't provide any information about the responses to the other questions. In contrast, Cronbach's alpha coefficient 1 indicates that the items are perfectly correlated, which means that knowing the value of one response provides complete information about the other items. Analysts often use 0.7 as a benchmark value for Cronbach's alpha. Cronbach's alpha coefficient of 0.7 and higher indicate that the items are sufficiently consistent to indicate that the measure is reliable (Frost, n.d.).

This study uses System Usability Scale (SUS) as research instrument. With significant application of SUS in different fields and in new contexts SUS has shown good reliability (Blattgerste, Behrends, & Pfeiffer, 2022).

To verify reliability of the scale for this study, Cronbach's alpha coefficient was calculated with the primary data collected for this research by using the following formula (Griffin, 2023):

$$\alpha = \frac{k}{k-1} * (1 - \frac{\sum_{i=1}^{k} \sigma_y^2}{\sigma_x^2})$$

Equation 1: Cronbach's alpha coefficient

where

$\alpha = Chronback\ Alpha$

K = *Number of Items in the scale (10 items for SUS)*

$\sum_{i=1}^{k} \sigma_y^2$ = *Summed variance of each item*

$\sigma_x^2 = Variance\ of\ total\ scores$

Cronbach's alpha coefficient for the SUS survey for this study is 0.8, which is considered good reliability (internal consistency).

4.4.2. Validity

The extent to which a measure adequately represents the underlying construct that it is supposed to measure is referred to as validity (often called construct validity). Validity of a measure can be evaluated by using "theoretical" or "empirical" approaches. Note that the types of validity described here refer to the validity of the measurement procedures (i.e., scales), which is different from the validity of hypotheses testing procedures, for example "internal validity (causality)", "external validity (generalizability)", or "statistical conclusion validity" (Wellington & Szczerbinski, 2007).

Theoretical evaluation of validity determines how well the theoretical construct is represented in the scale. This type of validity is called representational validity (translational validity) and consists of two subtypes (Wellington & Szczerbinski, 2007):

- "Face validity": evaluating whether an indicator seems to be a reasonable measure of its underlying construct

- "Content validity": evaluation of how well a set of scale items matches with the relevant content domain of the construct that it is attempting to measure.

Representational validity (both "face validity" and "content validity") is evaluated by using a panel of expert judges, who rate each item (indicator) on how well they fit the conceptual definition of that construct, using a qualitative technique called Q-sort, which is the systematic study of participant viewpoints (Wellington & Szczerbinski, 2007).

Empirical evaluation of validity uses empirical observations to determine how well a given measure relates to the external criterion (also called criterion-related validity), and includes four sub-types (Wellington & Szczerbinski, 2007):

- *"Convergent" validity*: closeness with which a measure relates to the construct that it is supposed to measure
- *"Discriminant" validity*: degree to which a measure does not measure (or discriminates from) other constructs that it is not supposed to measure
- "*Concurrent" validity*: how well one measure relates to other concrete criterion that is presumed to take place simultaneously. For example, can the students' scores in a calculus class presumed to correlate well with students' scores in a linear algebra class (since both calculus and algebra are mathematics)?
- "*Predictive" validity*: how well a measure can successfully predicts a future outcome that it is theoretically expected to predict. For example, can standardized test scores such as Scholastic Aptitude Test predict the academic success in college?

Typically, convergent validity and discriminant validity are evaluated together for related constructs. For example, if you expect that an organization's knowledge and organisation's performance are related, how for convergent validity you can make sure that your measure of organizational knowledge is actually measuring organizational knowledge, and how for discriminant validity you make sure that your measure is not measuring organizational performance? You can establish convergent validity by demonstrating similarity (i.e., high correlation) between values of indicators of the construct by comparing the observed values of one indicator of one construct with that of other indicators of the same construct. You can establish discriminant validity by demonstrating that indicators of one construct have low correlation with other constructs (i.e. are dissimilar from) (Wellington & Szczerbinski, 2007).

One approach to ensure validity of the research instrument is to use established and validated scales for measuring the constructs of interest (Blattgerste, Behrends, & Pfeiffer, 2022). The research instrument for this study is System Usability Scale (SUS), which has excellent psychometric properties. Research has consistently demonstrated that SUS is reliable and has acceptable levels of concurrent validity (Lewis & Sauro, 2009). SUS highly correlates with other questionnaire-based measurements of usability (called concurrent validity) (Sauro, 2011b). Validity refers to how well something can measure what it is intended to measure, which in our study is perceived usability. SUS has been shown to effectively discriminate between unusable and usable systems (Sauro, 2011b).

4.5. Ethical considerations

In the literature review and development of learning dashboard the distinction between ethical and legal principles is sometime blurred. A range of ethical issues that were considered and planned for in this research were *privacy, data ownership, control, transparency, consent, anonymity, non-maleficence (corruption) and beneficence (benefaction), data management, and security and access* (Corrin, et al., 2019).

In learning analytics literature, ethics and privacy are often conflated. However, there are some critical distinctions between the two terms that need to be defined and considered. While ethics is part of philosophy that seeks to resolve moral questions around right and wrong, privacy relates to the right to freedom from surveillance or unauthorized disclosure of one's personal information. Ethics can be defined as moral code of norms and conventions that exists in society externally to a person, whereas privacy is the intrinsic part of a person's identity and integrity (Corrin, et al., 2019). For this reason, privacy involves a level of self-determination in that individuals are imbued with the capacity to determine their level of privacy or disclosure of personal information. However, it is worth noting that technological development has increasingly impacted privacy and the way it is understood, as users are not always aware of the data that is being collected about them or the way data is being used (Corrin, et al., 2019).

Data ownership refers to the possession of, control of, and responsibility of information. Questions regarding the ownership of data include considerations of who determines what data is collected, who has the right to claim possession over the data, who decides how any analytics applied to the data are created, used and shared, and who is responsible for the effective use of data. Ownership of data relates to the outsourcing and transfer of data to third parties (Corrin, et al., 2019).

Transparency is the provision of information about data, intentions and processes used by organizations that a data subject can employ to inform decision making. In the context of learning analytics, this could relate to transparency of the data collected and techniques used for analysis and the generated outcomes. The lack of transparency can cause unease and distrust for data subjects. To alleviate these issues, it is important that higher education institutions are clear and transparent about what data is collected and the purpose of the data collection and how the data will be used, processed, stored, and shared (Corrin, et al., 2019)

Consent involves making a contract with the data subject by asking their consent for the data to be collected and analyzed. For consent to be valid it must be informed; therefore, the subject should be provided with clear and transparent information on the purpose for the data collection so that they are in a position to give informed consent. As part of informed consent, participants should be provided with an option to opt-out of having data collected at any stage. This is a challenging ethical consideration in the context of learning analytics, and one that has received increased attention in recent studies. A solution to this is to provide clear and transparent information to participants to obtain their consent (Corrin, et al., 2019)

Anonymity is the process of offering individuals the choice to conceal or reveal their identity and any identifying information about themselves. In the area of learning analytics anonymity may involve de-identification of individuals prior to data sharing or analysis. However, although it is generally agreed that institutions should make every effort to anonymize data, experts have also argued that anonymity cannot always be one hundred percent guaranteed (Corrin, et al., 2019).

Another important ethical consideration is protecting and managing data collected for the purpose of learning analytics. Data management is related to the storage of data collected and includes considerations of where data should be stored, how long it should be stored, and who has access rights to the data. Employing measures to protect and safeguard access to the data is also important (Corrin, et al., 2019).

Access should also be a consideration, this relates to the extent to which student data should be controlled, protected and restricted. It raises questions surrounding who should have access to the student data, including teaching staff and potential employers as well as the extent to which students should be able to have access to the data held about them and analytics performed on their data (Corrin, et al., 2019).

Now that we have described the research design, data collection and data analysis approach, we will present the results from predictive modeling, sentiment analysis modeling, and student facing learning dashboard evaluation, using system usability scale (SUS) in the next chapter.

Chapter 5. Results and Discussion

This chapter starts by the results and analysis of the intelligent student facing dashboard (SF-LAD) usability test and hypotheses verification, followed by presenting and analyzing the results of the predictive modeling and sentiment analysis modeling for the SF-LAD design.

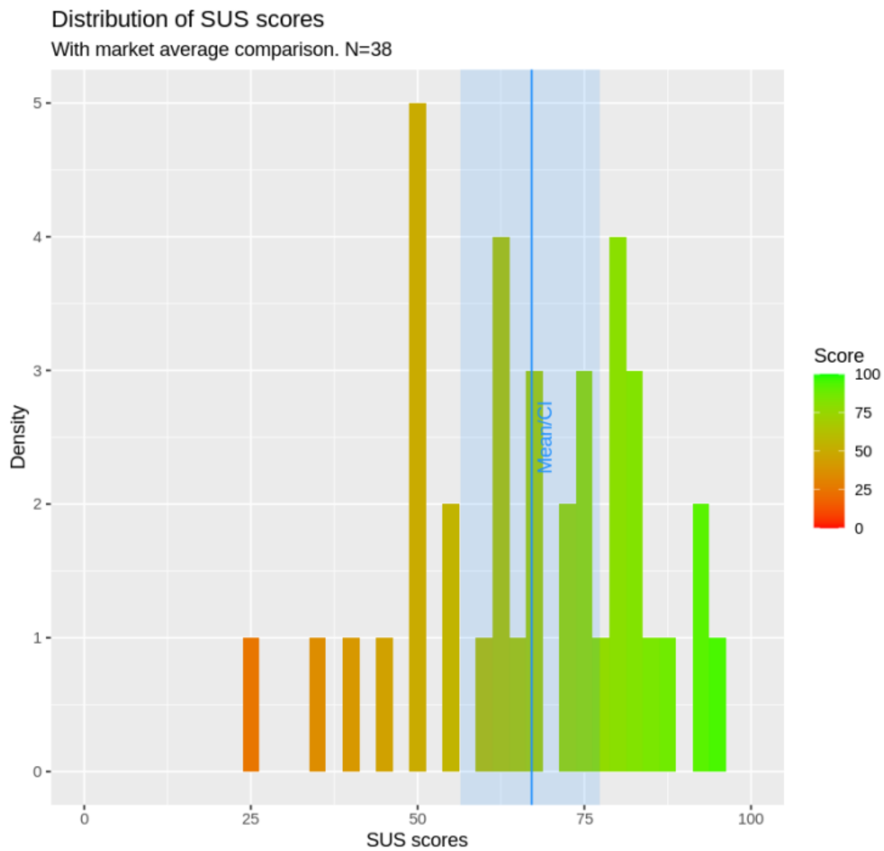
5.1. SF-LAD Usability Evaluation Results

In this section, the result of usability evaluation of the SF-LAD created for this study, using primary data from survey of 38 online learners from Athabasca University is described (*refer to "section 4.4.3 Data Collection" for detailed information about the sample data*). System Usability Scale (SUS) was used as research instrument (*refer to "section 4.4.3 Data Collection > Research Instrument" for detail information about the System Usability Scale)*.

To calculate the SUS score and interpret the results we use R programming language. Figure 13 presents SUS score distribution for SF-LAD.

Figure 13

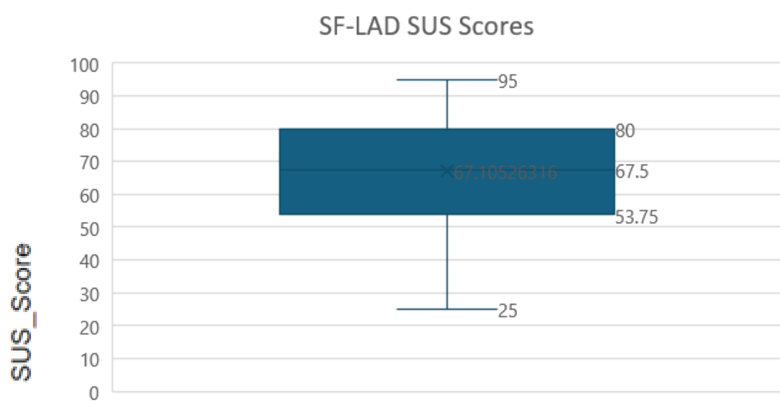
Distribution of SUS Scores for SF-LAD Created for This Study



To graphically summarize data values for SUS scores for SF-LAD, a box and whisker plot is created. A box and whisker plot can provide additional detail compared to a histogram and is considered an effective and easy way to summarize data to display the results in a single graph. The box presents the range of SUS scores that around 50 percent of the data points fall. There are five marks in the graph. The mark with the greatest value is the maximum, which is far outside the box. The mark with the lowest value is the minimum, which is also far outside the box on the opposite side of the maximum. The box itself includes the lower quartile, the upper quartile, and the median in the center. The median is presented in the graph as the value separating the higher half of the box from the lower half of the box. The lower quartile is the 25th percentile, and the upper quartile is the 75th percentile. The median is the middle, and it helps give a better sense of what to expect from these measurements. The whiskers (*the lines extending from the box on both sides to present upper and lower fence*) typically extend to 1.5 times the Interquartile Range (*IQR),* which is the box *(i.e., IQR = Q3 – Q1),* to set a boundary beyond which would be considered outliers (Box and Whisker Plots, n.d.). Figure 14 presents box and whisker plot for SF-LAD SUS scores. SF-LAD data does not include any outliers.

Figure 14

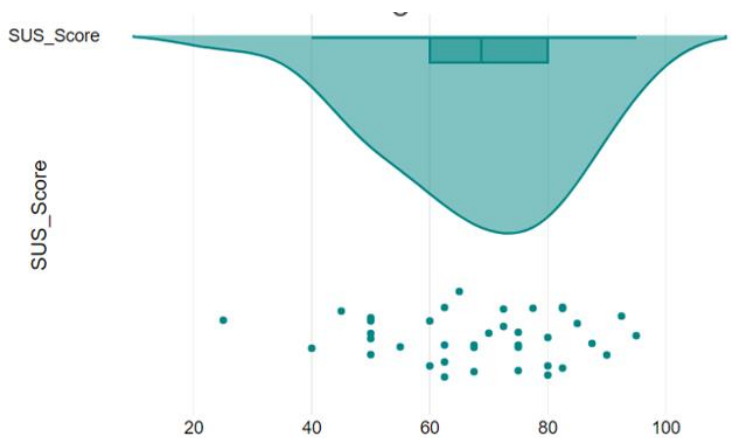
Box and Whisker Plot to Summarize SF-LAD SUS Scores



Effective data visualization is important to data analysis, data interpretation, and data communication (Allen, et al., 2021). To provide further graphical illustration of SF-LAD's SUS scores, a raincloud graph is plotted. The raincloud plot combines data distribution (*the 'cloud' in the graph*), with jittered raw data (*the 'rain' in the graph*). It supplements boxplot presentation of data. Figure 15 presents raincloud graph for SF-LAD SUS scores to illustrate distribution of data. It shows that SUS scores have single mode (one peak).

Figure 15

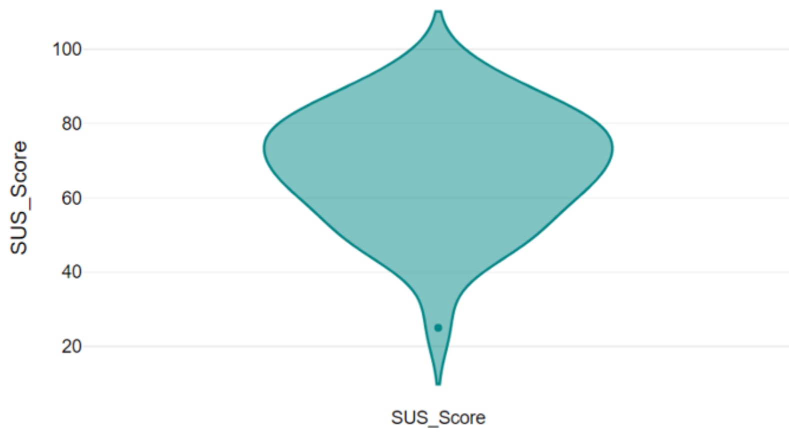
Raincloud Graph to Illustrate SF-LAD SUS Scores



To understand where values are clustered around, a violin plot is created for SF-LAD SUS scores. Violin plot provides additional insight about the distribution of the data. For example, is data clustered around the mean or is data clustered around the median? Or is the data mostly clustered around the minimum or maximum? This insight complements summary statistics presented in the box and whisker plot and data distribution and jittered raw data presented in the raincloud graph. The box plot only provides for visualizing basic distributions. Boxplot provides a convenient way for comparing summary statistics (such as range and quartiles), but it doesn't allow visualization of variations in the data (Carron, 2021). Figure 16 presents violin plot for SF-LAD SUS scores.

Figure 16

Violin Plot for SF_LAD SUS Scores



The following section presents data analysis for SF-LAD results, using five different methods (*Percentile method, Letter Grade Method, Acceptability Method, Adjective Method, and Promoters and Detractors Method - refer to section 4.5 "Data Analysis" for details about each method*), based on the calculated SUS score for each participant.

5.2. SF-LAD SUS Survey Results' Interpretation

After completing SUS survey and data normalization, SUS score was calculated for each participant, and the mean SUS score was calculated to assess usability for SF-LAD. Summery statistics for this research's survey are shown in Table 7.

Table 7

Summary Statistic for SF-LAD SUS Scores for The Survey Participants

Mean	67.1
Std Dev.	16.8
Max.	95
Min.	25
Median	67.5

5.2.1. Percentile Method

The mean SUS score for SF-LAD usability test survey for all the 38 participants is 67.1, which is in the 50th percentile as shown in Table 8 (Sauro, 2018). It means that this system (SF-LAD) outperforms 50% of all the systems in the database (Sauro, 2018).

Table 8

Percentile Interpretation of SUS Score

SUS Score	Percentile Range
84.1 - 100	96-100
80.8 – 84	90-95
78.9 – 80.7	85-89
77.2 – 78.8	80-84
74.1 – 77.1	70-79
72.6 - 74	65-69
71.1 – 72.5	60-64
65 - 71	41-59
62.7 – 64.9	35-40
51.7 – 62.6	15-34
25.1 – 51.6	2-14
0 - 25	0-1.9

Literature was consulted to interpret mean SUS score of 67.1 in the context of educational internet platform technology. A systematic evaluation of usability for educational technology (using SUS as research instrument) by Vlachogianni and Tselios (2021) found that mean usability levels change depending on the type of educational technology (Vlachogianni & Tselios, 2021). Vlachogianni and Tselios (2021) examined a total of 156 studies from 95 research papers that were composed of five major categories of educational technologies: 1) internet platforms such as LMS, MOOC, Wiki, etc., 2) mobile applications, 3) university websites, 4) multimedia and 5) affective tutoring systems (*ATS - the next-generation learning approach that can detect the affective status of learning to increase performance*) (Vlachogianni & Tselios, 2021). Table 9 presents the result of the study in terms of mean SUS score and standard deviation for each type of educational technology (Vlachogianni & Tselios, 2021).

Table 9

Mean SUS Score for Various Educational Technologies

Educational Technology	Number of Surveys Examined for the Educational Technology	Mean SUS Score for the Educational Technology	Standard Deviation
The Internet Platforms (e.g., LMS, MOOC, Wiki, etc.)	77	66.25	12.42
Mobile Applications	33	73.62	13.49
University Websites	12	63.82	16.52
Multimedia	21	76.43	9.45
Affective Tutoring Systems			

The category of interest to our study is internet platforms (LMS, MOOC, Wiki, etc.), which we can use as a benchmark to compare with SF-LAD in terms of usability to identify where SF-LAD stands in comparison with other internet platform educational technologies. Comparing the mean SUS score for 77 surveys of internet platform educational technologies (66.25) with SF-LAD mean SUS score (67.1) provides insight about the usability of SF-LAD compared to similar category of educational technology.

Therefore, SUS score of 67.1 is interpreted to have a decent level of usability. This is in-line with Vlachogianni and Tselios (2021) study's findings that internet platforms generally have good usability levels but not without flaws (Vlachogianni & Tselios, 2021).

5.2.2. Letter Grade Method

Using letter grade interpretation streamlines the communication of SUS score interpretation with some stakeholders who may not be familiar with statistical percentile. Letter grade interpretation of SUS scores is closely related to percentile. Using curved grading scale, the grades range from A+ to F as shown in Table 10 (Sauro, 2018).

Table 10

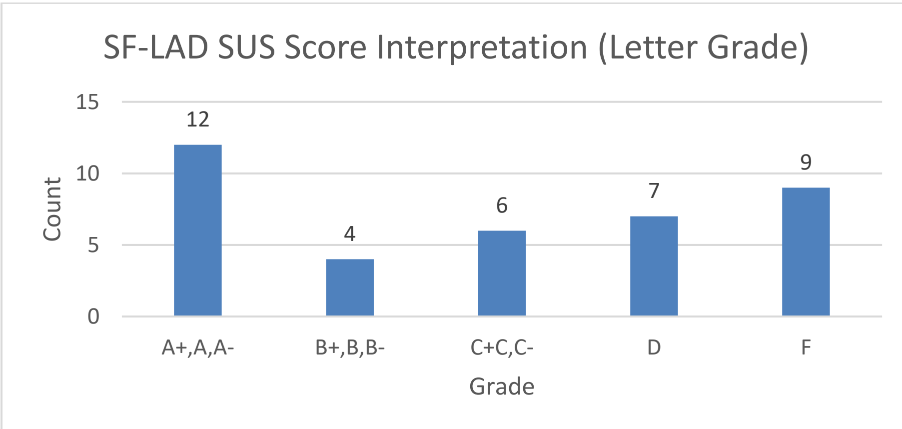
Curved Grading Scale and Percentile interpretation of SUS Scores

SUS Score	Percentile Range	Grade
84.1 - 100	96-100	A+
80.8 – 84	90-95	A
78.9 – 80.7	85-89	A-
77.2 – 78.8	80-84	B+
74.1 – 77.1	70-79	B
72.6 - 74	65-69	B-
71.1 – 72.5	60-64	C+
65 - 71	41-59	C
62.7 – 64.9	35-40	C-
51.7 – 62.6	15-34	D
25.1 – 51.6	2-14	F
0 - 25	0-1.9	F

Figure 17 presents count of letter grades for individual participant's SUS score letter grade for all 38 respondents.

Figure 17

SF-LAD SUS Letter Grade Count



The SUS score for 32% of the respondents correspond to letter grades A+, A, or A-; for 11% of the respondents corresponds to letter grades B+, B, or B-, for 15% of the respondents correspond to letter grades C+, C, or C-, for 18% of the respondents corresponds to letter grade D; and for 24% of the respondents corresponds to letter grade F.

The letter grade for mean SUS score for all 38 participants is 67.1 that corresponds to letter grade C in the curved grading scale interpretation of SUS Scores (Sauro, 2018). Letter grade C is consistent with our benchmark (mean SUS score of 77 surveys of internet platform educational technologies used in the systematic review of educational technologies' usability: 66.25 is interpreted as letter grade C– refer to Table 13) (Vlachogianni & Tselios, 2021).

5.2.3. Adjective Method

It is beneficial to also use relative judgement of the system usability to comparable applications, using SUS large database as a benchmark (Bangor, Kortum, & Miller, 2009). Relative Adjective interpretation of SUS scores is presented in Table 11.

Table 11

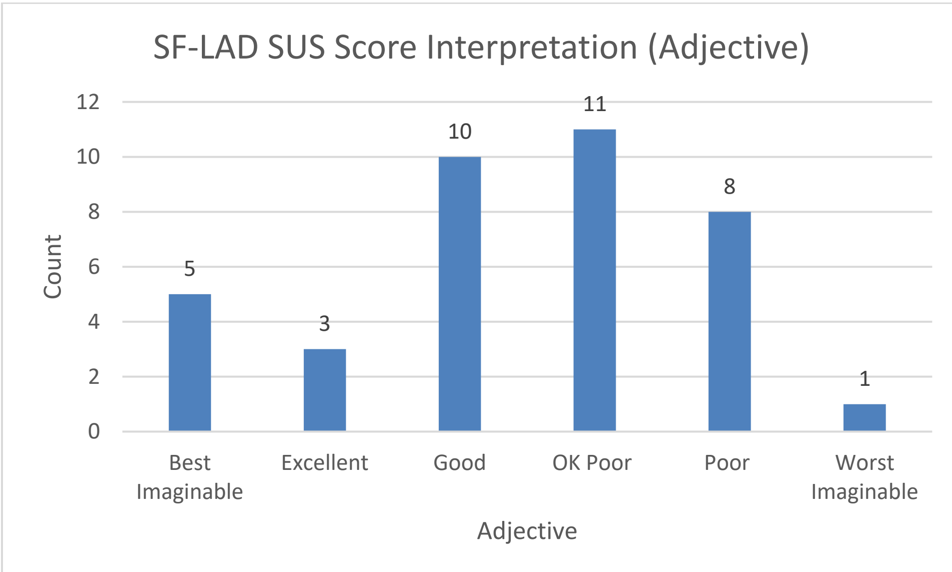
Relative Adjective, Curved Grading Scale, and Percentile Interpretation of SUS Scores

SUS Score	Percentile Range	Grade	Adjective
84.1 - 100	96-100	A+	Best Imaginable
80.8 – 84	90-95	A	Excellent
78.9 – 80.7	85-89	A-	Good
77.2 – 78.8	80-84	B+	Good
74.1 – 77.1	70-79	B	Good
72.6 - 74	65-69	B-	Good
71.1 – 72.5	60-64	C+	Good
65 - 71	41-59	C	OK
62.7 – 64.9	35-40	C-	OK
51.7 – 62.6	15-34	D	OK
25.1 – 51.6	2-14	F	Poor
0 - 25	0-1.9	F	Worst Imaginable

Figure 18 presents the count of various adjectives for interpretation of SUS score for all 38 respondents.

Figure 18

SF-LAD SUS score Adjective Count



The SUS score for 13% of the respondents correspond to Best Imaginable adjective, 8% corresponds to Excellent, 26% corresponds to Good, 29% corresponds to OK, 21% corresponds to Poor, and 3% corresponds to Wort Imaginable.

With SUS score of 67.1, usability of SF-LAD corresponds to relative adjective "OK" (Sauro, 2018). Adjective OK is consistent with our benchmark (mean SUS score of 77 surveys of internet platform educational technologies used in the systematic review of educational technologies' usability: 66.25 is interpreted as adjective OK– refer to Table 13) (Vlachogianni & Tselios, 2021).

5.2.4. Acceptability Method

Whether the dashboard will be accepted or not accepted is also an important factor (Vlachogianni & Tselios, 2021). This is another variation on using words to describe SUS score (Sauro, 2018). Table 12 presents acceptability of SUS scores.

Table 12

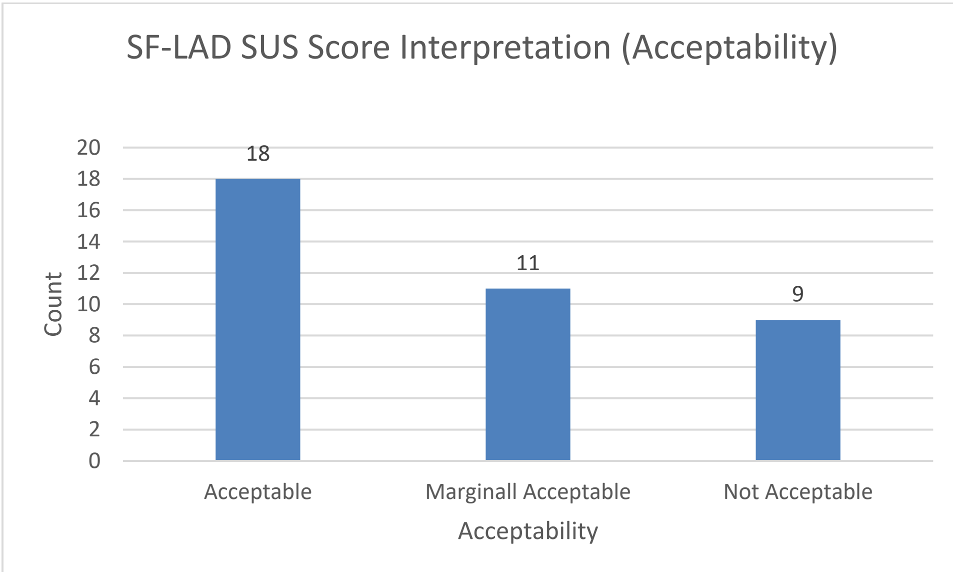
Acceptability, Relative Adjective, Curved Grading Scale, and Percentile Interpretation of SUS Scores

SUS Score	Percentile Range	Grade	Adjective	Acceptability
84.1 - 100	96-100	A+	Best Imaginable	Acceptable
80.8 – 84	90-95	A	Excellent	Acceptable
78.9 – 80.7	85-89	A-	Good	Acceptable
77.2 – 78.8	80-84	B+	Good	Acceptable
74.1 – 77.1	70-79	B	Good	Acceptable
72.6 - 74	65-69	B-	Good	Acceptable
71.1 – 72.5	60-64	C+	Good	Acceptable
65 - 71	41-59	C	OK	Marginally Acceptable
62.7 – 64.9	35-40	C-	OK	Marginally Acceptable
51.7 – 62.6	15-34	D	OK	Marginally Acceptable
25.1 – 51.6	2-14	F	Poor	Not Acceptable
0 - 25	0-1.9	F	Worst Imaginable	Not Acceptable

Figure 19 presents the count of different acceptability levels for interpretation of SF-LAD for all 38 participants.

Figure 19

SF-LAD SUS score Acceptability Count



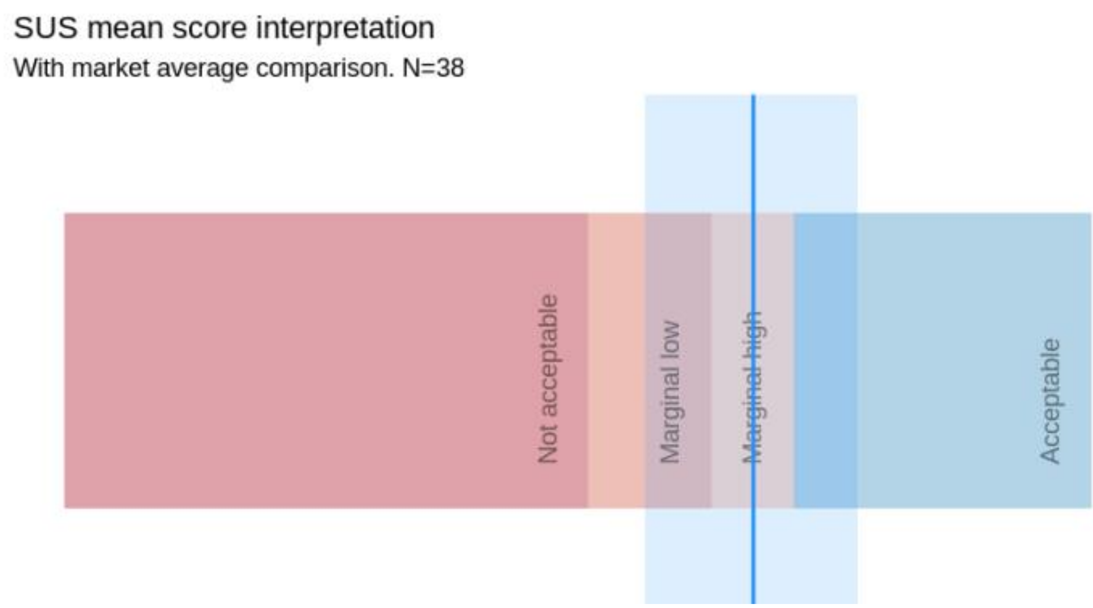
SUS score for 47% of the participants corresponds to Acceptable, for 29% corresponds to Marginally Acceptable, and for 24% corresponds to Not Acceptable.

With mean SUS score of 67.1, SF-LAD is interpreted as "Marginally Acceptable" (Bangor, Kortum, & Miller, 2009). Marginally Acceptable is consistent with our benchmark (mean SUS score of 77 surveys of internet platform educational technologies used in the systematic review of educational technologies' usability: 66.25 is interpreted as Marginally Acceptable– refer to Table 13) (Vlachogianni & Tselios, 2021).

Figure 20 is a graphical presentation of SF-LAD acceptability based on mean SUS score (67.1).

Figure 20

SF-LAD Mean SUS Score Interpretation



5.2.5. Promoters and Detractors Method

There is strong correlation between the SUS score and Net Promoter score (NPS); with SUS score on average explaining between 30% and 50% of variations in users Likelihood to Recommend (LTR). Consequently, we also use NPS as a method for interpretation of SUS results (Sauro, 2018).

As a rule of thumb, LTR can be calculated by dividing SUS score by 10 (Sauro, 2012). Therefore, we can use SUS score to determine LTR classification (*Promoters, Passive, and Detractors*) (Sauro, 2018). *Promoters* are loyal and enthusiastic audience for the system or product. They are likely to share positive feedback or get engaged on social media and recommending the system or product with others; *Passives* are typically satisfied with the system or product but not enough happy to share their experience with others; *Detractors* are unhappy audience who are not likely to use the system or purchase the product again, and may even discourage others from using the system or purchasing the product (What is a good Net Promoter Score?, n.d.).

Sauro (2018) analyzed data from 4,664 respondents to estimate the average SUS score for each LTR classification. Based on his analysis, to achieve a *Promoter* classification, a SUS score needs to be around 81 or higher. On the other hand, SUS scores around 53 and lower are associated with *Detractor* classification. SUS scores between these two ranges are associated with *Passive* classification (Sauro, 2018).

Once we identify LTR classification for each respondent, we can calculate the percentage of each classification. By subtracting the percentage of detractors from the percentage of promoters we can identify the Net Promoter Score (NPS) for the system (*Net Promoter Score (NPS) = total % of promoters – total % of detractors*). Note that the percentage of passives is not used in the Net Promoter Score formula (How to calculate & measure Net Promoter Score (NPS), n.d.).

Calculating NPS for a system helps to determine how much our audience like SF-LAD. Using NPS allows us to understand our audience better and make continuous improvements (10 benefits of Net Promoter Score and how to maximize it, n.d.).

Table 13 presents classification of participants based on their Likelihood to Recommend (LTR) (Sauro, 2012).

Table 13

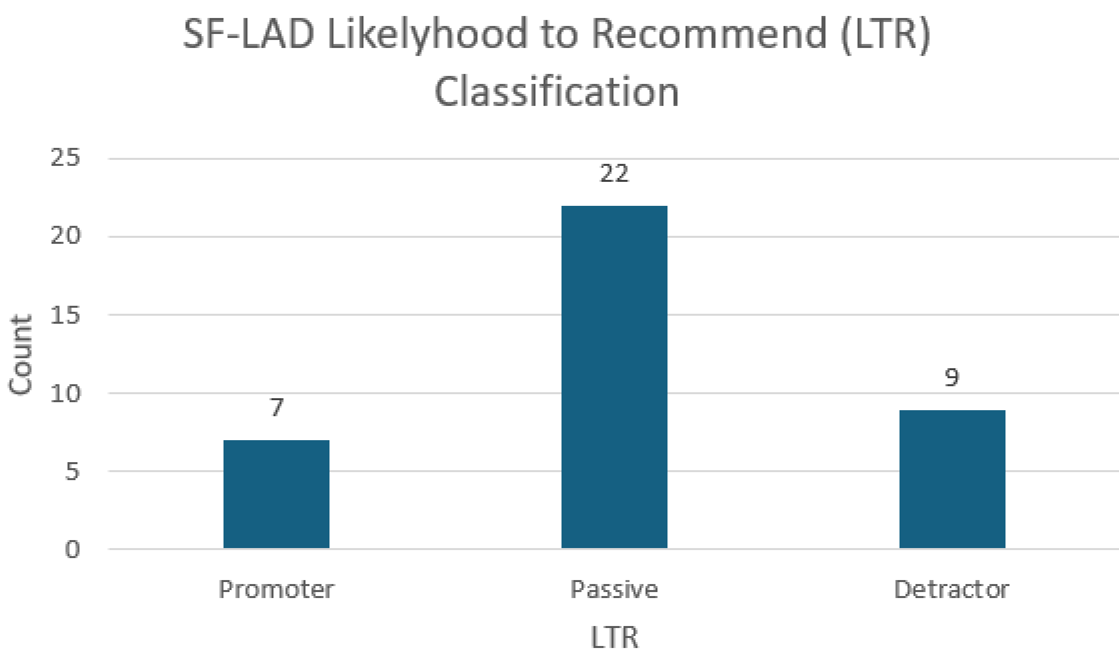
Likelihood to Recommend (LTR), Acceptability, Relative Adjective, Curved Grading Scale, and Percentile Interpretation of SUS Scores (Shaded Row Corresponds to SF-LAD SUS Score)

SUS Score	Percentile Range	Grade	Adjective	Acceptability	NPS
84.1 - 100	96-100	A+	Best Imaginable	Acceptable	Promoter
80.8 – 84	90-95	A	Excellent	Acceptable	Promoter
78.9 – 80.7	85-89	A-	Good	Acceptable	Promoter
77.2 – 78.8	80-84	B+	Good	Acceptable	Passive
74.1 – 77.1	70-79	B	Good	Acceptable	Passive
72.6 - 74	65-69	B-	Good	Acceptable	Passive
71.1 – 72.5	60-64	C+	Good	Acceptable	Passive
65 - 71	41-59	C	OK	Marginally Acceptable	Passive
62.7 – 64.9	35-40	C-	OK	Marginally Acceptable	Passive
51.7 – 62.6	15-34	D	OK	Marginally Acceptable	Detractor
25.1 – 51.6	2-14	F	Poor	Not Acceptable	Detractor
0 - 25	0-1.9	F	Worst Imaginable	Not Acceptable	Detractor

Figure 21 presents count of promoter/detractor interpretation of SUS score for SF-LAD for all 38 participants.

Figure 21

SF-LAD LTR vs Count



In the context of educational technology, deploying the Net Promoter Score for learning products can be used as a powerful metrics because it allows to gauge whether learners are happy. This information can then be used to continually improve the learning system (Theodotou, 2022).

The percentage of promoters for SF-LAD is 18% (7/38) and the percentage of detractors is 24% (9/38). Therefore, NPS (*Net Promoter Score*) for SF-LAD is -6 (*i.e., the percentage of Promoters minus the percentage of Detractors: 18% - 24% = -6%*).

The following criteria can be used for interpretation of Net Promoter Score (NPS) (Baquero, 2022):

- NPS of 50 to 100 is considered excellent.
- NPS of 0 to 49 is considered sufficient.
- NPS of −49 to 0 is considered insufficient.
- NPS of −100 to −50 is considered deficient.

Based on this criteria NPS for SF-LAD (-6%) is interpreted as insufficient. It means that the quality of the system is perceived insufficient by the users and requires improvement (Baquero, 2022). Therefore, the feedback gathered from the participants was classified and analyzed to identify the suggested areas for improvement to incorporate in the next iteration of SF-LAD (refer to section 5.4 Survey Participants' Comments Analysis).

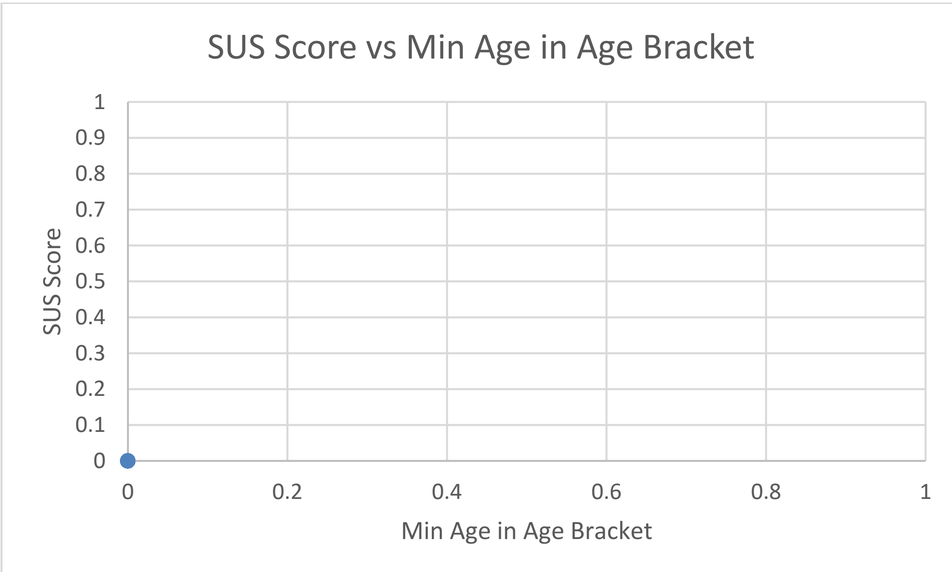
Furthermore, to gather deeper insight, the LTR classifications were analyzed by using the collected demographic data to observe if any pattern emerges. Promoters consist of 4 females and 3 males; and Detractors consist of 4 females, 4 males, and 1 participant that preferred not to declare. In terms of age range, Promoters consist of 4 participants 18-30 and 3 participants 31-40; and Detractors consist of 5 participants 18-30, 2 participants 31-40, and 2 participants 41-50. Data does not indicate any patterns related to age or gender.

5.3. Correlation Between Participants' Age and SUS score

In a sample of 38 participants, we investigated the possible correlation between participants' age and the SUS score for each participant. In the survey, participant's age was classified under four ranges (18-30, 31-40, 41-50, 51 and above). 50% (19) of the participants are 18-30, 34% (13) of the participants are 31-40, 13% (5) of the participants are 41-50, and 3% (1) participant is 51 or over. To create a scatter plot for individual participant's SUS score and age, Minimum age in each range was used. Figure 22 presents the scatter plot to examine correlation between SUS score and Age (i.e., min age in each age bracket).

Figure 22

SUS Score vs Participants Age



Figures 23 and 24 present mean SUS score for each age bracket (column chart), and distribution of data for each age bracket (box and whisker plot) for SUS scores respectively.

Figure 23

Mean SUS Score by Age

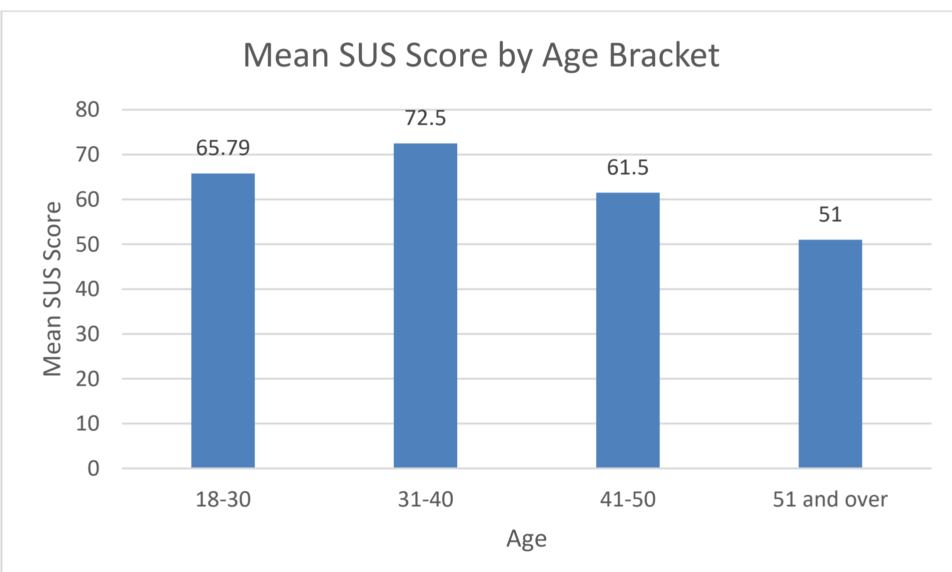
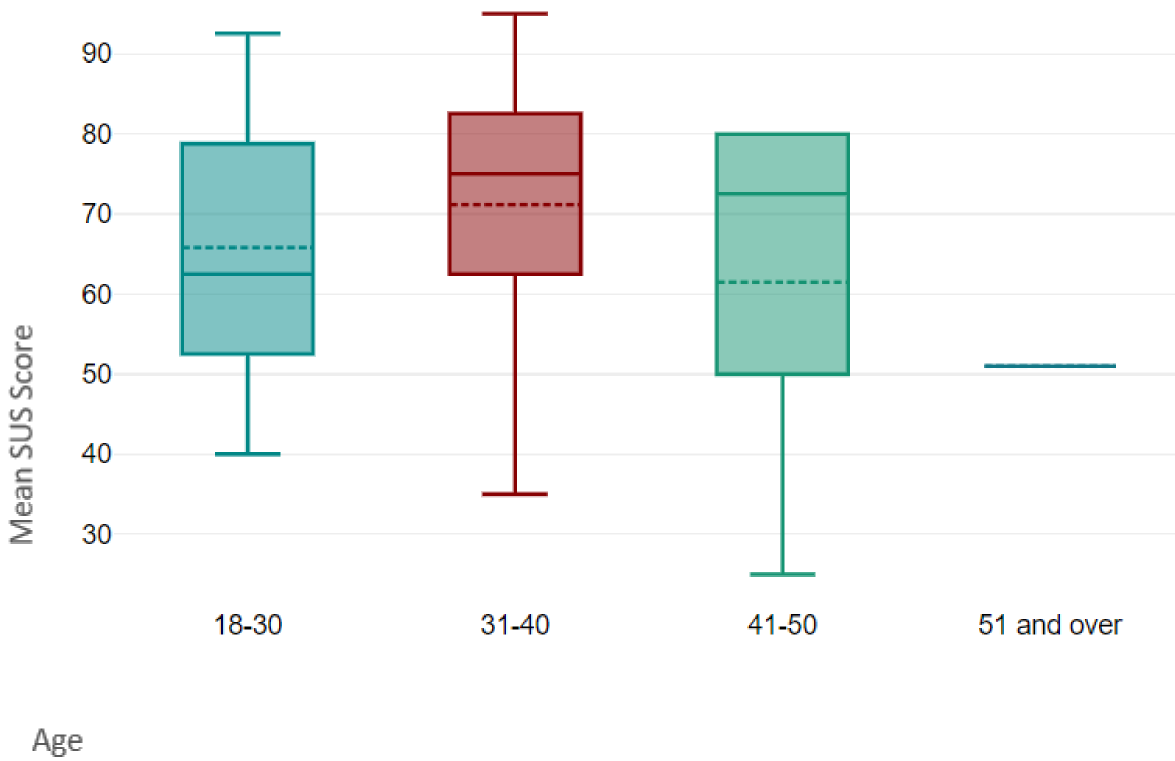


Figure 24

Mean SUS Score Distribution by Age



To assess if there is a correlation between SUS score and participants age, we used Pearson's correlation measure (r). Pearson is one of the most popular correlation methods. It produces a score from − 1 to + 1. If the Pearson coefficient for the two variables is high (close to + 1), it is interpreted that the two variables are directly correlated. If the Pearson coefficient is high but negative (close to -1), it is interpreted that the two variables are inversely correlated. If the Pearson coefficient is small (close to 0), it is interpreted that the two variables are not correlated (Pearson Correlation, n.d.).

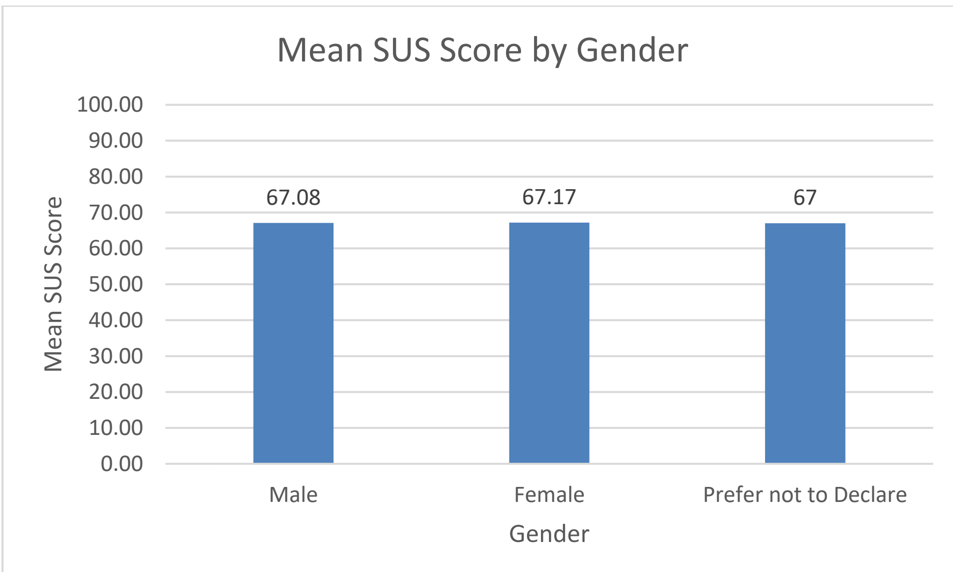
Pearson coefficient for each participant's SUS score vs the Minimum age in the age bracket that the participant belongs to is 0.009 (r = 0.009), which indicates no correlation between participants age and SUS score.

5.4. Correlation Between Participants' Gender and SUS Score

In a sample of 38 participants, 47% (18) of the participants are male, 39% (15) of the participants are female, and 13% (5) of the participants prefer not to declare. The mean SUS score for different genders is visualized to provide insight, as presented in Figure 25.

Figure 25

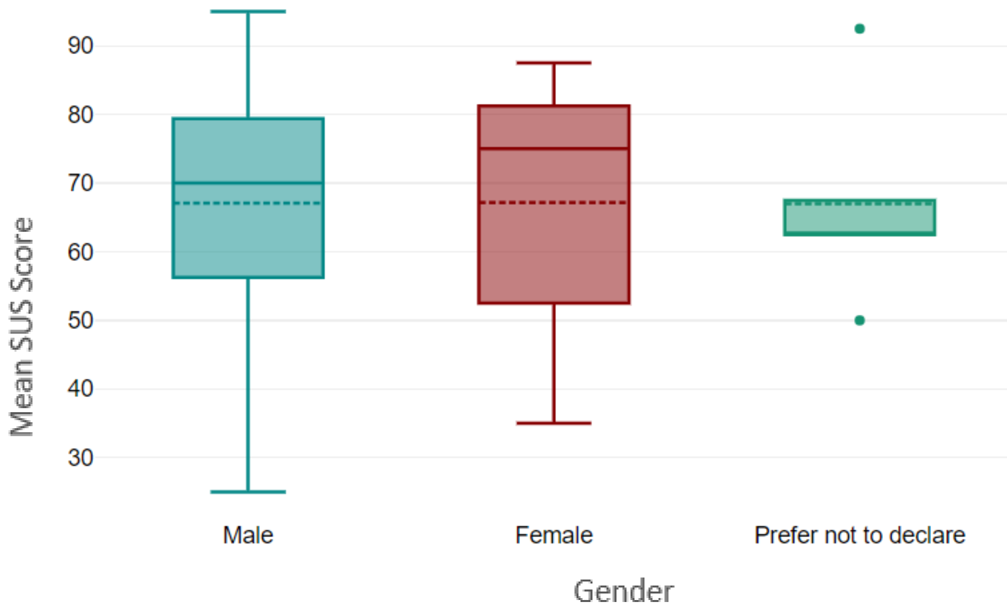
Mean SUS Score by Genders



The mean SUS score for females is 71.25, for males is 66.25, and for participants that prefer not to declare is 63.75. To provide additional insight about the distribution of the SUS scores by gender, box and whisker plot for SUS score is created by gender, as presented in Figure 26.

Figure 26

Mean SUS Score Distribution by Genders



5.5. Hypothesis Verification Results

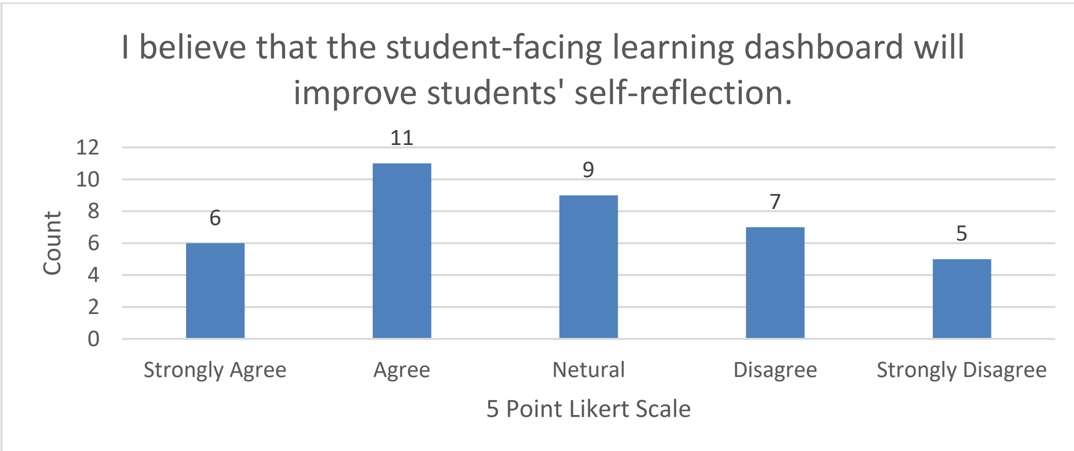
In the survey of 38 participants from Athabasca University, two additional questions were included to assess the two hypotheses for this research. In this section, the participants' responses to these questions are graphically presented and analyzed.

5.5.1. First Hypothesis Verification Question

In response to the first question (*I believe that the student-facing learning dashboard will improve students' self-reflection. Self-reflection is defined as the activity of thinking about one's own feelings and behaviour, and the reasons that may lie behind them*), 44.73% of the participants Strongly Agree or Agree that SF-LAD will help them increase their self-reflection, 31.57% of the participants Strongly Disagree or Disagree that the dashboard will help them increase their self-reflection, and 23.68% of the participants believed that the dashboard would Neither Help Nor Not Help increase their self-reflection. The responses to the first question are captured in Figure 27.

Figure 27

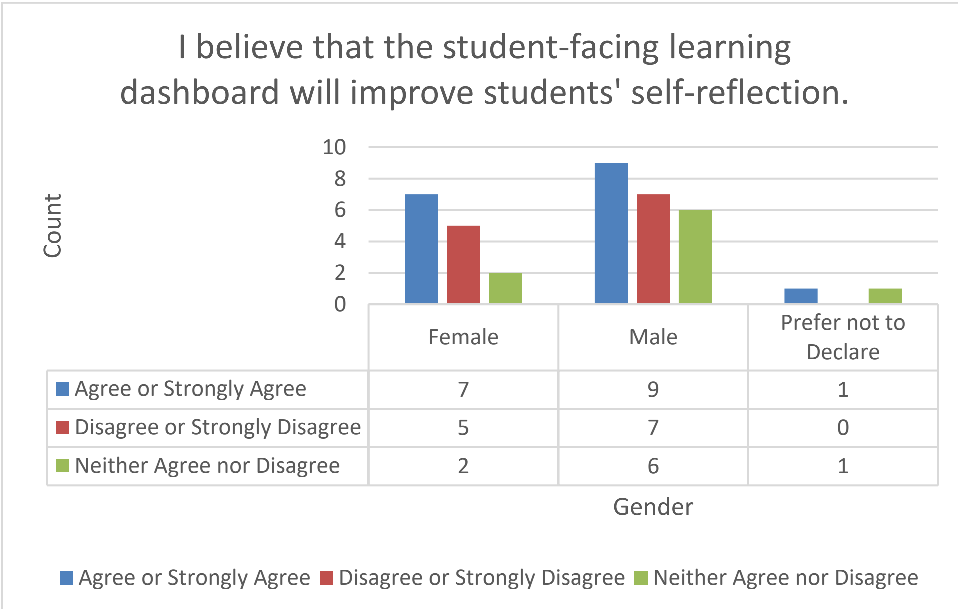
SUS Bar Chart Showing the Count of Student Responses to the First Research Question



Based on this result, we can see that the opinion on whether the dashboard would be helpful in increasing students' self-reflection is divided. However, a significant portion of the students (44.73%) found that the dashboard would provide them enough information about the measured metrics to provide context to increase their self-reflection. To gain further insight, two more graphs are created to breakdown responses by gender and age, as shown in Figures 28 and 29.

Figure 28

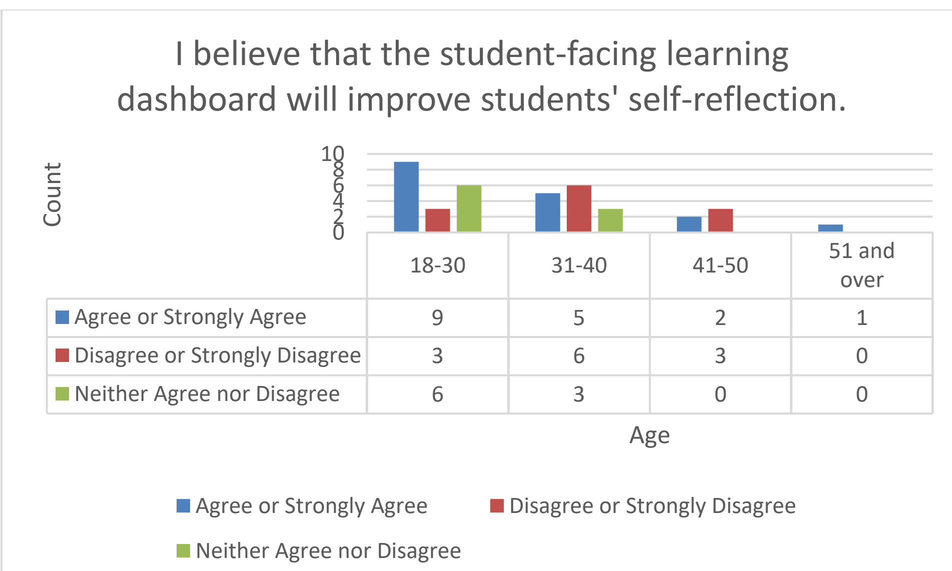
Responses to The First Research Question (by Gender)



	Female	Male	Prefer not to Declare
Agree or Strongly Agree	7	9	1
Disagree or Strongly Disagree	5	7	0
Neither Agree nor Disagree	2	6	1

Figure 29

Responses to The First Research Question (by Age)



	18-30	31-40	41-50	51 and over
Agree or Strongly Agree	9	5	2	1
Disagree or Strongly Disagree	3	6	3	0
Neither Agree nor Disagree	6	3	0	0

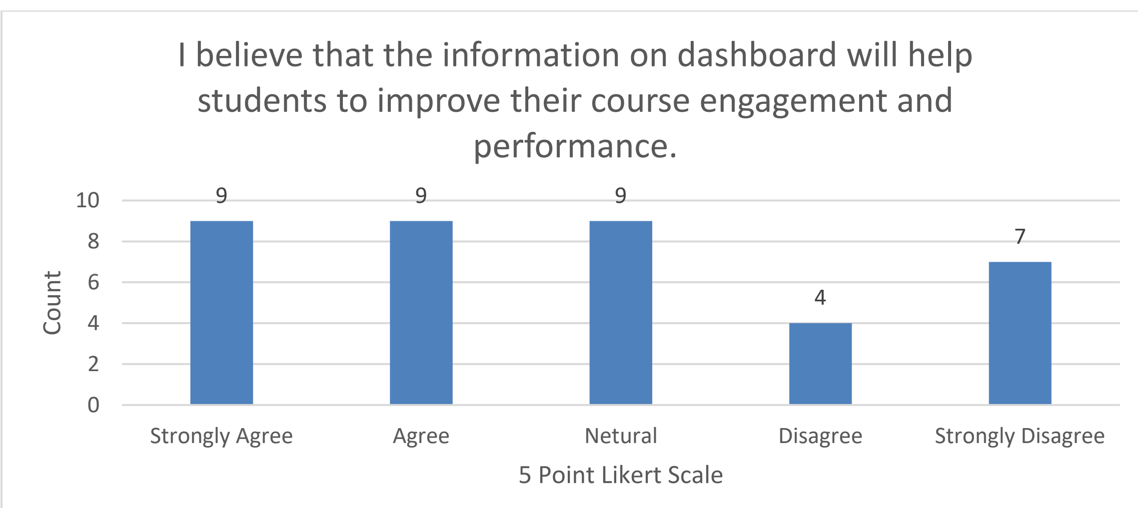
A pattern emerging from Figure 29 is that only 17% of age bracket 18-30 Disagree or Strongly Disagree, in contrast to age bracket 31-40 with 43% Disagree or Strongly Disagree and age bracket 41-50 with 60% Disagree or Strongly Disagree. The pattern shows that younger respondents are more likely to find the student-facing dashboard helpful for self-reflection, compared to more mature respondents. This is an interesting observation and can be further investigated in future surveys.

5.5.2. Second Hypothesis Verification Question

In response to the second question (*I believe that the information on dashboard will help students to improve their course engagement and performance*), 47.36% of the participants Strongly Agree or Agree that the information presented in the dashboard will help them to improve their course engagement and performance, 28.94% of the participants Strongly Disagree or Disagree that the dashboard will provide enough context to increase their engagement and performance, and 23.68% of the participants believed that dashboard would Neither Help Nor Not Help them improve course engagement and performance. The responses to the second question is captured in Figure 30.

Figure 30

Count of Responses to the Second Research Question



Based on this result, we can see that opinion on whether the dashboard would be helpful in increasing students' course engagement and performance is divided. However, a significant portion of the participants (47.36%) found the dashboard helpful to provide enough information to provide context to help increase their course engagement and performance. To gain further insight more graphs are created to breakdown responses by gender and age, as shown in Figures 31 and 32.

Figure 31

Count of Responses to The Second Research Question (by Gender)

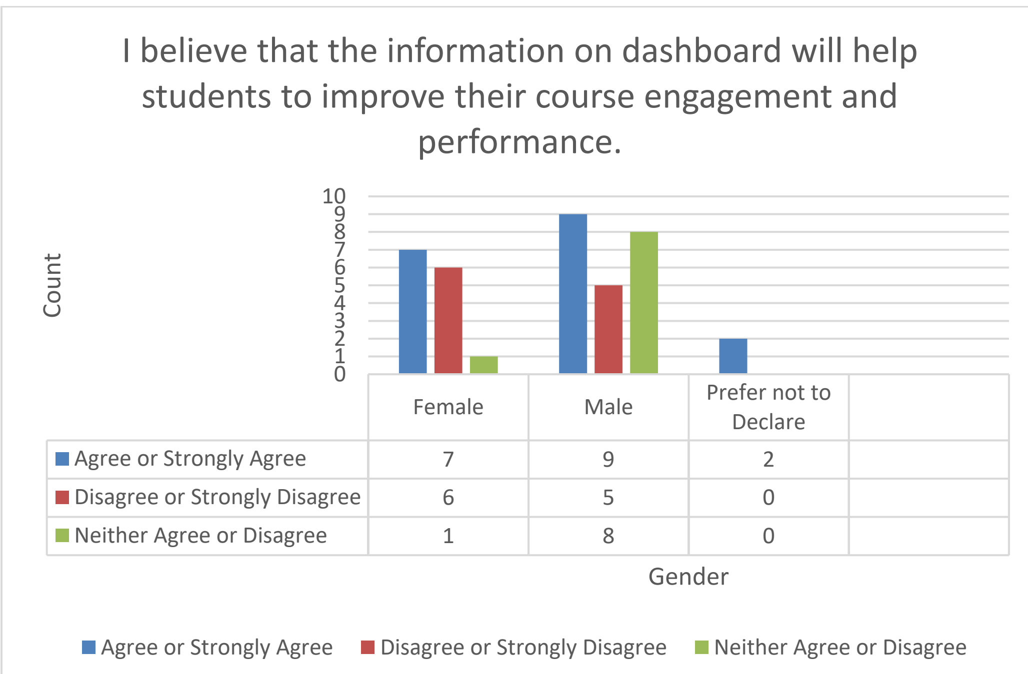
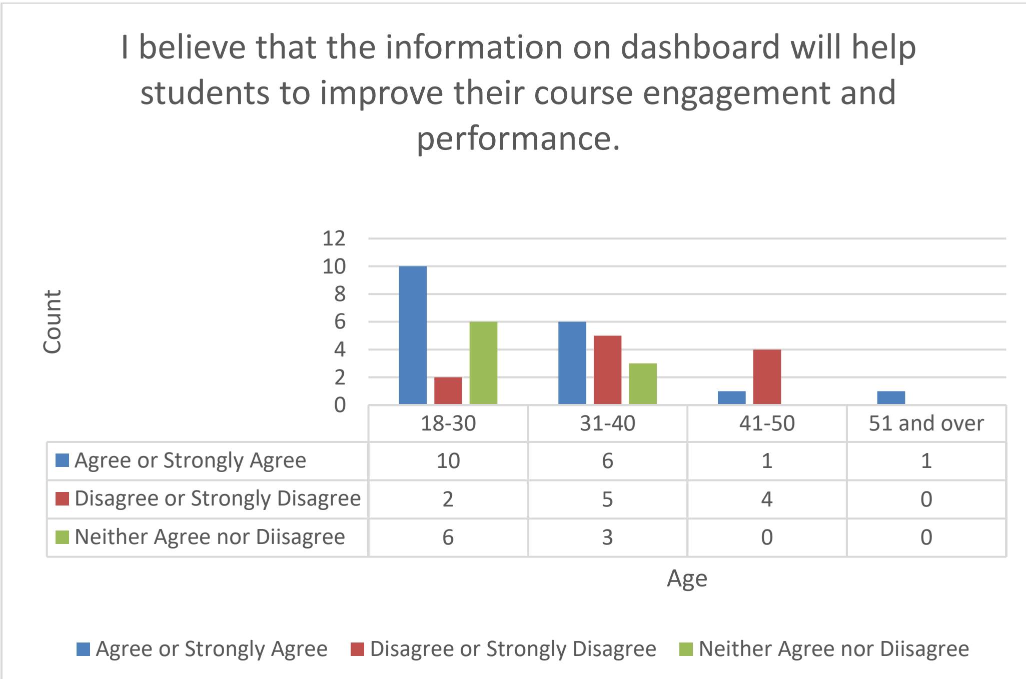


Figure 32

Count of Responses to The Second Research Question (by Age Range)



The pattern emerging from Figure 32 is consistent with Figure 29. Only 11% of age bracket 18-30 Disagree or Strongly Disagree, in contrast to age bracket 31-40 with 36% Disagree or Strongly Disagree and age bracket 41-50 with 80% Disagree or Strongly Disagree. The pattern indicates that younger respondents are more likely to find the student-facing dashboard helpful to improve course engagement and performance compared to more mature respondents. This is an interesting observation and can be further investigated in future surveys.

5.6. Survey Participants' Comments Analysis

The goal of data collection via survey for SF-LAD is to measure the perceived usability of the new student facing dashboard designed for this study. In addition to using SUS scale and hypothesis verification questions, participants were asked to comment on how the student-facing learning dashboard can be improved.

Overall, from 38 participants in the survey, 28 participants provided comment. Participants commented on various aspects of the dashboard. These comments are classified and recorded by category. The categorized comments are presented in Table 18 to derive insight. The sum of the number of participants who commented on each category is presented in Table 14 and visualized in Figure 33. In Table 14, the sum of the count for various categories is more than the number of respondents that left comment for improvement of SF-LAD (*28 participants*). The reason is that some participants provided comments that touches various categories; therefore, these comments are split and recorded in their relevant category. Another observation is that some of the comments/recommendations received from the participants contradict each other. For example, some participants have discouraged peer comparison, and some participants have encouraged peer comparison. That is the reason that the exact wording of the participants' comments is included in Table 14 to provide accurate insight.

Table 14

Participants Comments (Classified by Category)

	Categories of Comment/ Recommendation	**Participants' Comments**	**Number of Participants That Provided This Category of Comment**
		SF-LADs Strengths	
1.	Finds SF-LAD effective.	1) I would love to have access to this type of dashboard to visualize how I feel during my semester and having an overview of it all. 2) I love this dashboard and wish it was already implemented." 3) So far so good 4) Overall, I think this is a good dashboard that will help students manage their time and help them understand how well they are doing in the course.	10

		5) I believe the dashboard could be positive to show where you stand in each course as well as your average in the classroom. 6) Lastly, the grades versus class average seems like a VERY helpful feature. It's the sort of thing that I find to be lacking from Athabasca in general: visualising my grades is difficult, and knowing the class average would give me a lot of perspective on how I'm performing that is sorely lacking from distance ed. These widgets look nice overall, and fine to use, but I find the colours on the pie chart to be really hard on the eyes. The UI is nice aside from that. I like that I can see more detailed information by mousing over elements. 7) I like that you can see where you are compared to the class average. 8) I like that you can see where you are compared to the class average. 9) It is very user-friendly and self-explanatory. 10) Good as is.	
Areas for SF-LAD Improvement			
2.	Add description for measurements in the SF-LAD to provide context.	1) Include a brief description of what is being measured for each section. This will help inform users of what exactly they are looking at. 2) The dashboard provides data with no context and no calls to action. I do not know what I need to do to improve, or what actions will drive desired outcomes. 3) It would be nice to know where the information is coming from. Such as what goes into the 'sentiment' metric. For the course performance dashboard, I believe that the following could improve it: Add the ability to drill down in the data, ie redirect each point to the where the marked assignment can be accessed. Add small data labels so that you don't need to hover your mouse over the data point to see the value. For the course performance dashboard, I believe that the	4

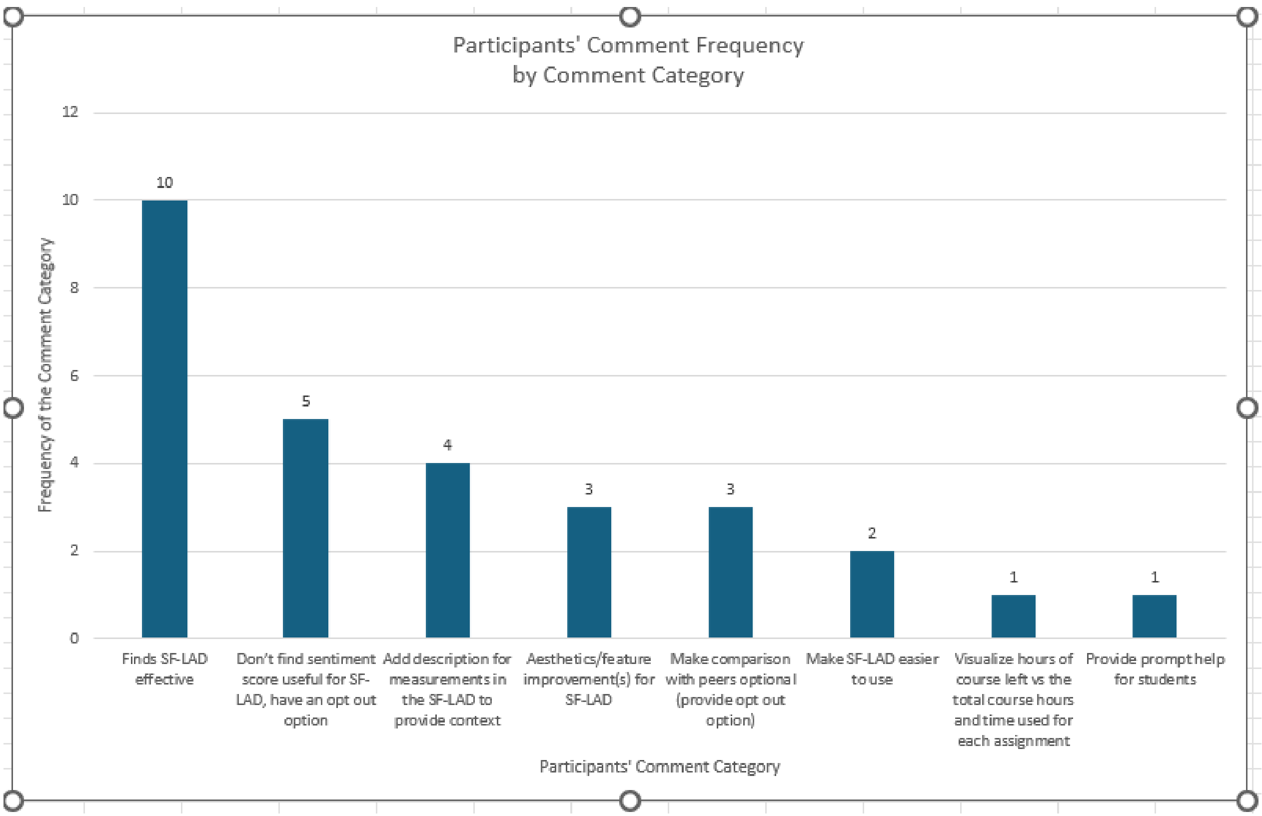
		following could improve it: Add the function where when you select an assignment, it will redirect to the assignment specific page (ie where you can see what the assignment is). Add small data labels so that you don't need to hover your mouse over the graph to see the value. Add the ability for students to customize the dashboard, so that they can add or hide assignments being presented. Overall, I think this is a good dashboard that will help students manage their time and help them understand how well they are doing in the course. I think that it could be improved with additional functionality though, and I think that it could also use some text that explains what these graphs present and how to use the dashboard. For example, what does the ""Overall time management score"" of 129.0, mean?" 4) add a section for feedback and suggestions for improvements so the student can know how to do better if needed	
3.	Aesthetics/feature improvement(s) for SF-LAD.	1) I think overall, it is a pretty good job done on the ideas, I would improve the look of the student sentiment chart, other than that, I was naturally attracted to look at them and explore them. 2) With all of the free tools available for school and time management I do not believe that this would meet the standards. The dashboard would need a lot more features and better esthetics. It looks dated and the tools and information is lacking. 3) More visually pleasing	3
4.	Make SF-LAD easier to use.	1) I think the dashboard ought to be made easier to use. 2) I believe the layout could be worked on, so it is easier to catch all of the information at a glance without scrolling	2

5.	Make comparison with peers optional (provide opt out option).	1) "Don't show how they're doing compared to other students. Evidence is emerging that social media causes depression. The primary theory is that people feel worse comparing themselves to others who are doing better. So, for students doing below average, seeing they're doing below average may contribute to depression." 2) I think showing these metrics to students will, in the best case, barely help the students that are already doing well and, in the worst case, heavily discourage students that are already doing poorly. I don't think these kinds of metrics should be shown to students at all. The students should be in the classes to learn, not to watch a number go up. 3) I felt anxiety almost instantly looking at it. It would just put more pressure on me.	3
6.	Don't find sentiment score useful for SF-LAD, have an opt out option.	1) I don't see a use in recording my sentiment score, unless it is viewed by a professor to improve the course. 2) Everything seemed straightforward besides the student sentiment part, needed to assume more on what that was about than the rest. 3) I would get rid of the Student Sentiment chart, it seems the least useful of the charts or at the very least change the order of the charts presented so that it is the last item shown. 4) I am unsure about the sentiment grading. It seems to get into tricky territory that could lead to a negative learning environment (I think discussion board interactions are fine as-is.) The time taken to complete assignments would, to me, be EXTREMELY anxiety inducing, as a student who performs well but does distance learning to allow myself more control over my own schedule. For some people, it might be positive, but there should be an option to opt out of any of these given features.	5

		5) The student sentiment chart feels inappropriate and a bit of a privacy violation.	
7.	Suggestion to visualize hours of course left vs the total course hours and time used for each assignment.	1) Perhaps some way to calculate the total amount of course hours a student will have in a term, then show the hours left and the time used for each assignment.	1
8.	Provide prompt help for students.	1) More question forums so students can promptly be helped.	1

Figure 33

Frequency of Each Comment Category



Areas for improvement for SF-LAD captured from the participants' comments include:

1) Make sentiment analysis graphs in SF-LAD optional (i.e., students can opt-out)

2) Make peer comparison graph in SF-LAD optional (i.e., students can include or exclude the additional line chart for "class average" in the Course Performance graph).

3) Add more description to SF-LAD to provide additional context.

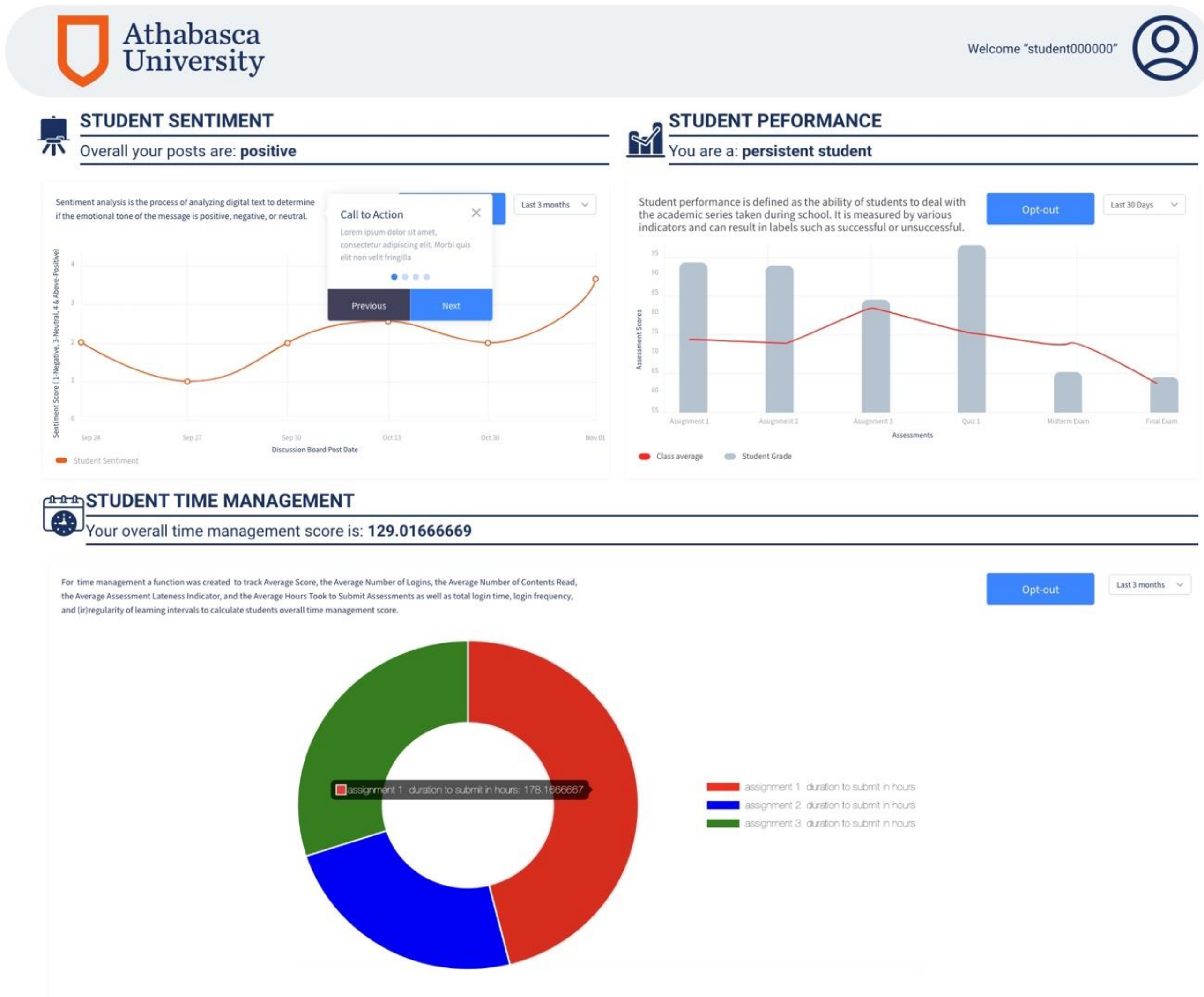
4) Explore valuable suggestions provided to enhance dashboard aesthetics and functionality (refer to table 18).

5.7. Revised SF-LAD for Next Iteration Based on Survey Feedback

Figure 34 shows the revised version of the student facing learning dashboard for the next iteration of the intelligent student facing learning dashboard (SF-LAD) to incorporate the feedback received from the survey participants. As shown in Figure 34, the next iteration of SF-LAD includes a short explanation of each key performance indicator (KPI), call to action to provide students with feedback, and the ability to opt-out of the features that a student might not find beneficial for them.

Figure 34

The Revised Version of the Student Facing Learning Dashboard for Next Iteration

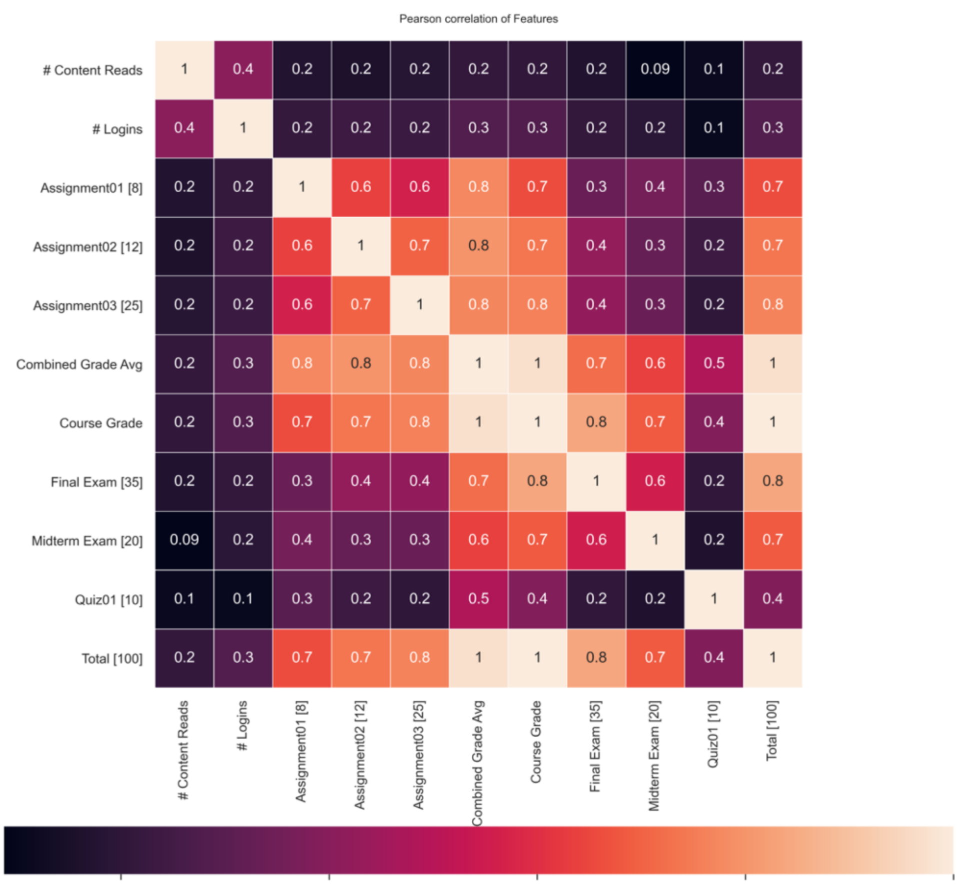


5.8. K-mean Model Results

The K-mean predictive model for this study was developed using the Student Performance and Engagement Prediction in eLearning dataset (Moubayed, Injadat, Shami, BouNassif, & Lutfiyya, 2020). The dataset has 485 records for students. After importing the dataset, the Pearson correlation coefficient for the data was calculated and visualized using a heatmap (refer to Figure 11). The correlation plot visualizes the associated metrices. The magnitude of the correlation coefficient is visualized by a combination of numbers with two colors, printed with varying intensity. The colors indicate the sign of the coefficient and the intensity of the color increases proportionally with the magnitude of the correlation coefficient (Haarman, et al., 2015). Figure 35 shows the correlation plot for the positively correlated features that have the most effect on the formation of clusters in this study. In this heatmap, the color black represents features that are weakly correlated, and light orange represents features that are strongly correlated.

Figure 35

Pearson Correlation Coefficient Displaying The Positively Correlated Features



We examined the correlation for all the 20 available features using the Pearson correlation coefficient and identified the following features as being positively correlated with each other: *Quizzes*, *Assignments (1 to 3)*, *Midterm Exam*, *Final Exam*, *Number of Logins*, *the Number of Times*

Students Accessed the Course Resources, and *the Course Overall Grade.* For example, Assignments 1 and 2 have a correlation of 0.7 and Assignment 3 has a correlation of 0.8 with students' Overall Course Grade, which demonstrates a strong relationship. Additionally, Final Exam has a correlation of 0.8 with the Overall Course Grade, which demonstrates a strong relationship. Midterm Exam has a correlation of 0.6 and Quiz 1 has a correlation of 0.4 with students' Overall Course Grade, which demonstrates a moderate relationship. Furthermore, Number of Logins has a correlation of 0.3 and the Number of Contents Read has a correlation of 0.2 with students' Overall Course Grade, which demonstrates a weak relationship. Therefore, Assignments (1 to 3), Midterm Exam, and Final Exam (highly correlated features) contribute more to forming the clusters.

Refer to Table 15 for K-mean clustering summary statistics. The clusters (represented by numbers 0, 1 and 2) represent the grouping of the students. Students in cluster 0 are categorized as Persistent Students based on the *Average Course Grade (78.60)*, *the Average Number of Logins (76.69)*, *Average Content Read (276.13)*, *Low Lateness Indicator for Assessments (ranging between 0.01 - 0.02),* and *the Average Hours to Submit the Assignments (254.82)*. Students in cluster 1 are categorized as Regular Students based on the *Average Course Grade (83.26)*, *the Average Number of Logins (84.42)*, *Average Content Read (272.09)*, *Low Lateness Indicator for Assessments (0.0),* and *the Average Hours to Submit Assignments (103.72).* Note that students who are considered Regular were able to complete course assessments at a faster pace compared to Persistent students. Students in cluster 2 are categorized as Irregular Students based on the *Average Course Grade (31.00)*, *the Average Number of Logins (28.14)*, *Average Content Read (126.86)*, *High Lateness Indicator for Assessment (ranging between 0.71 – 1.00),* and *the Average Hours to Submit Assignments (448.13)*.

Table 15

K-mean Clustering Summary Statistics

	Cluster 0 (Persistent Students)	Cluster 1 (Regular Students)	Cluster 2 (Irregular Students)
Quiz01	77.03	79.73	38.57
Assignment01	75.59	77.44	0.00
Midterm Exam	74.16	80.13	59.57
Assignment02	76.30	79.67	10.71
Assignment03	81.05	82.39	19.14
Final Exam	59.17	65.77	26.00
Course Grade	78.60	83.26	31.00
Total	78.69	83.38	31.14
Number of Logins	76.29	84.42	28.14
Number of Content Reads	276.13	272.09	126.86
Number of Forum Reads	1.07	3.15	0.14
Number of Forum Posts	0.07	0.21	0.14
Number of Quiz Reviews before submission	2.06	2.06	1.00
Assignment 1 lateness indicator	0.02	0.00	1.00
Assignment 2 lateness indicator	0.03	0.00	0.71
Assignment 3 lateness indicator	0.01	0.00	0.71

Assignment 1 duration to submit (in hours)	305.49	151.43	558.00
Assignment 2 duration to submit (in hours)	215.43	65.27	277.39
Assignment 3 duration to submit (in hours)	243.53	94.45	509.00
Average time to submit assignment (in hours)	254.82	103.72	448.13

Further analyzing the clustering results, we can see that from the 485 students in the original dataset, *52.9% of students are Regular*, *45.7% of students are Persistent,* and *1.4% are Irregular*. The student categorization is visualized in Figure 36.

Figure 36

Pie Chart Showing the Percentage of Students Categorized as Regular, Persistent and Irregular

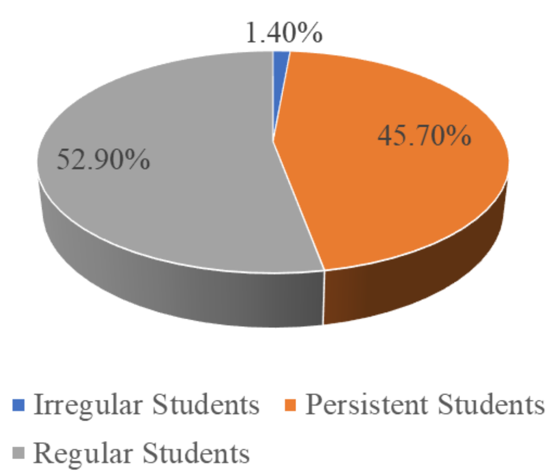
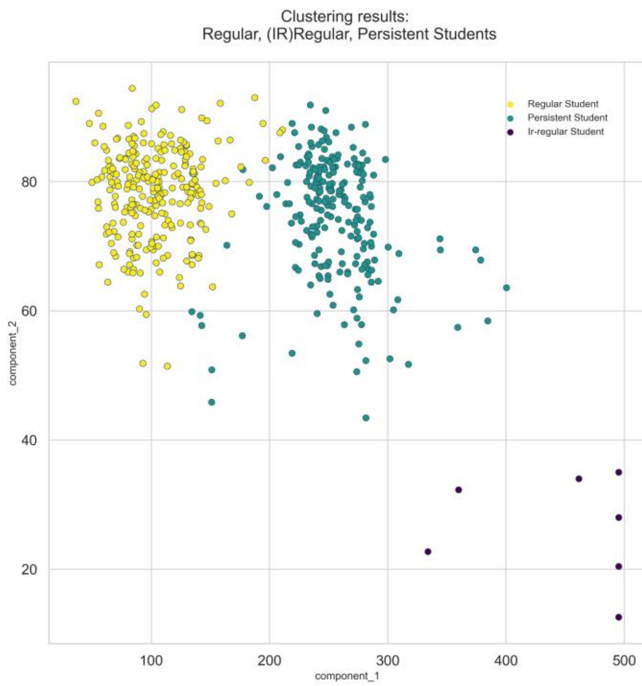


Figure 37 presents K-mean clustering model results, representing three distinct categories of students (Persistent, Regular, Irregular). Using this model, any new learner can be compared with the other learners based on his/her performance (according to the criteria described above) and identified in one of the groups in the SF-LAD.

Figure 37

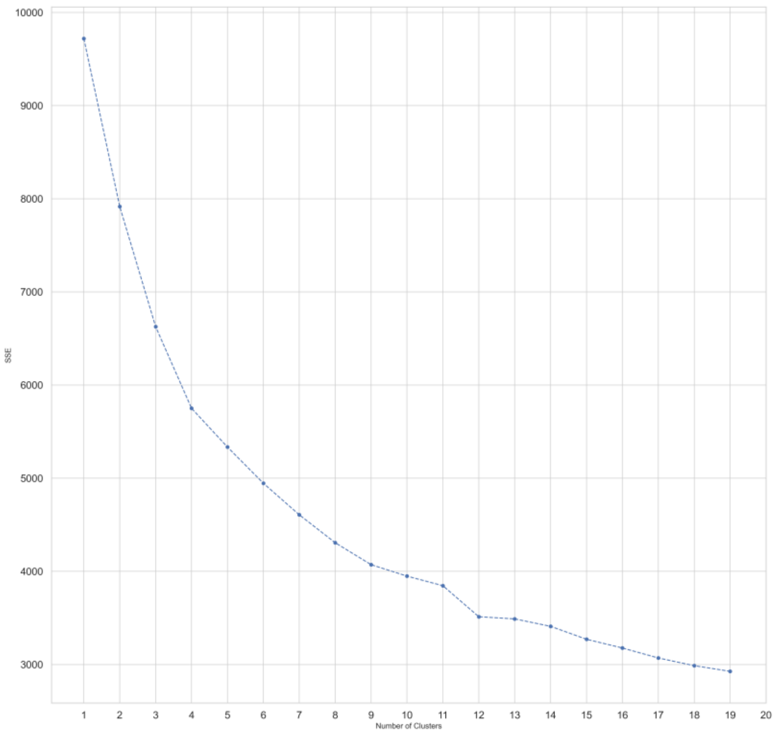
K-mean Clustering Model Results



To evaluate whether the clusters are meaningful, we used the elbow method. This method evaluates the goodness of the clusters by looking at the percentage of variance explained as a function of the number of clusters (Bholowalia & Kumar, 2014). The elbow method works by plotting the average of the intra-cluster sum of squares of the distance between cluster centroids and the cluster points and the number of clusters (K) (Khandelwal, 2020). The results of the elbow method are plotted in Figure 38. From the plot, we can see that the $K = 3$ separates the data points with a minimal error.

Figure 38

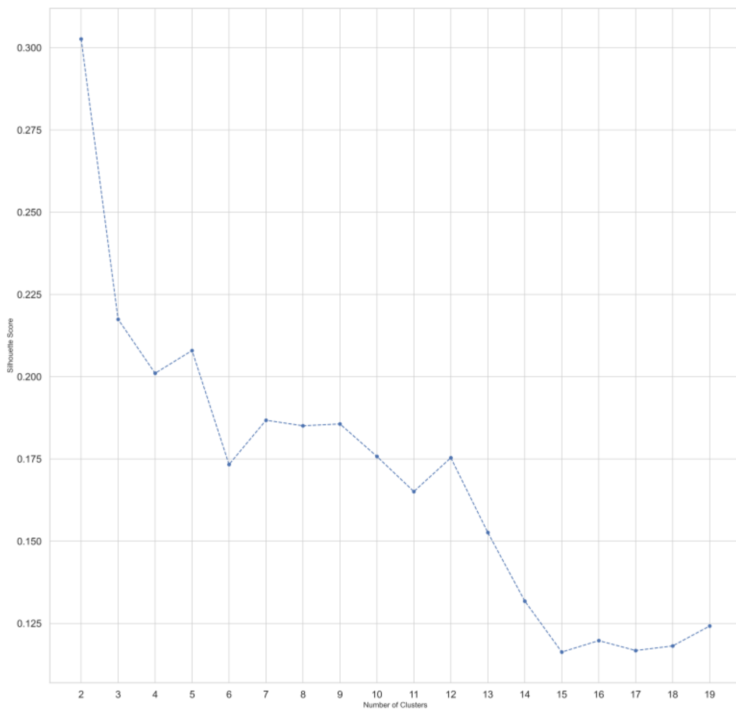
Elbow Plot (Sum of Square Distances)



We have also determined the cohesion and the separation of the clusters. We used the Silhouette coefficient to measure the cluster cohesion and separation and plotted the results in Figure 39.

Figure 39

Plot of the Silhouette Scores for The K-mean Model



The Silhouette score shows how well the clustering algorithm performs. Equation 2 represents the Silhouette score formula, where $b(i)$ represents the average distance of point *i* to all the points in the nearest cluster and $a(i)$ represents the average distance of point *i* to all the other points in its own cluster (Khandelwal, 2020).

$$(i) = \frac{b(i) - a(i)}{\max(a(i), b(i)} \qquad \text{(number??)}$$

Equation 2 Silhouette score formula

The Silhouette score varies between $+1$ and -1, with $+1$ being the best score and -1 being the worst overlapping clusters (Khandelwal, 2020).

In this study we looked at cognitive factors such as students' ability to complete assessments on time (students' ability to complete assessment on time is an indicator of students' planning and organizational ability, which is a cognitive process). We examined behavioral factors such as students' login frequency and how long it takes students to complete assessments. Both of these behaviors indicate students' engagement in the course. Metacognitive processes can be thought of as students' ability to self-regulate for the purpose of evaluating their behaviors to select effective learning behaviors. Metacognition occurs when students demonstrate awareness of their cognitive processes and then monitor and analyze those processes to reach learning goals. The insight gathered about each student's cognitive, behavioral and metacognitive characteristics can be used to predict if a particular student is going to achieve his/her learning goals. For example, Student000040 is considered a *Regular Student* based on his *Course Grade (92%)*, *Number of Logins (77)*, *Number of Contents Read (492), Low Lateness Indicator (0.0)* and *the Average Hours to Submit Assignments (110.05).* Using this information, we can predict that this student will pass the course with letter grade A, as shown in Table 16. The datapoint related to student000040 is labeled in Figure 40.

Figure 40

K-mean Clustering Model Results with Points for Student00040 Labeled

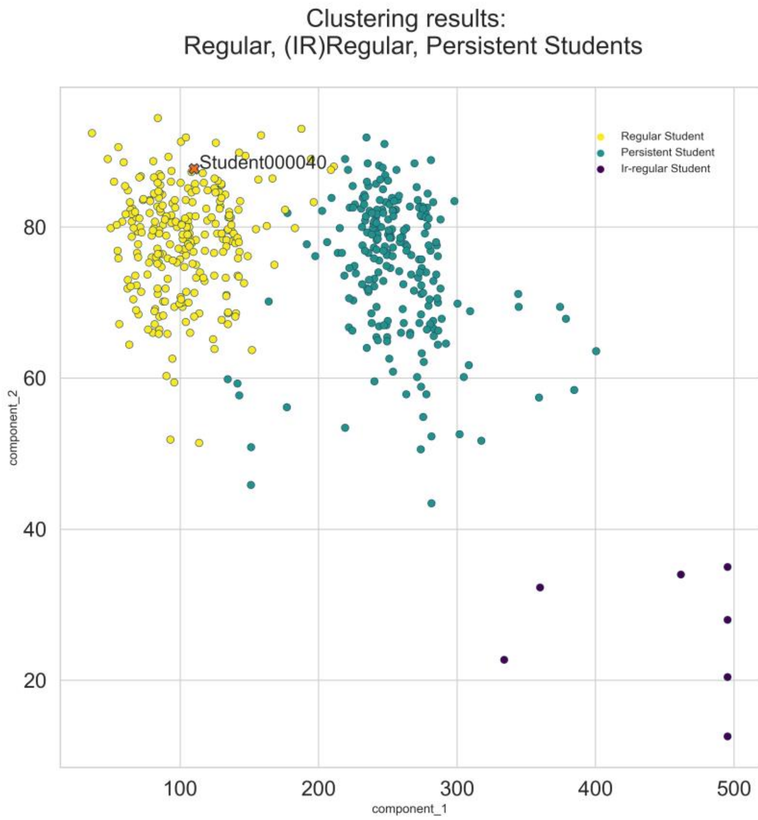


Table 16

Descriptive Statistics for Student 00040

	student000040
Quiz01	90
Assignment01	93
Midterm Exam	85
Assignment02	91
Assignment03	92
Final Exam	71
Course Grade	92
Total	92
Number of Logins	77
Number of Content Reads	492
Number of Forum Reads	0
Number of Forum Posts	0
Number of Quiz Reviews before submission	4
Assignment 1 lateness indicator	0
Assignment 2 lateness indicator	0
Assignment 3 lateness indicator	0
Assignment 1 duration to submit (in hours)	165.2
Assignment 2 duration to submit (in hours)	69.21666667
Assignment 3 duration to submit (in hours)	95.73333333
Student_Category	Regular Student
Letter_Grade	A
Course_Outcome	P

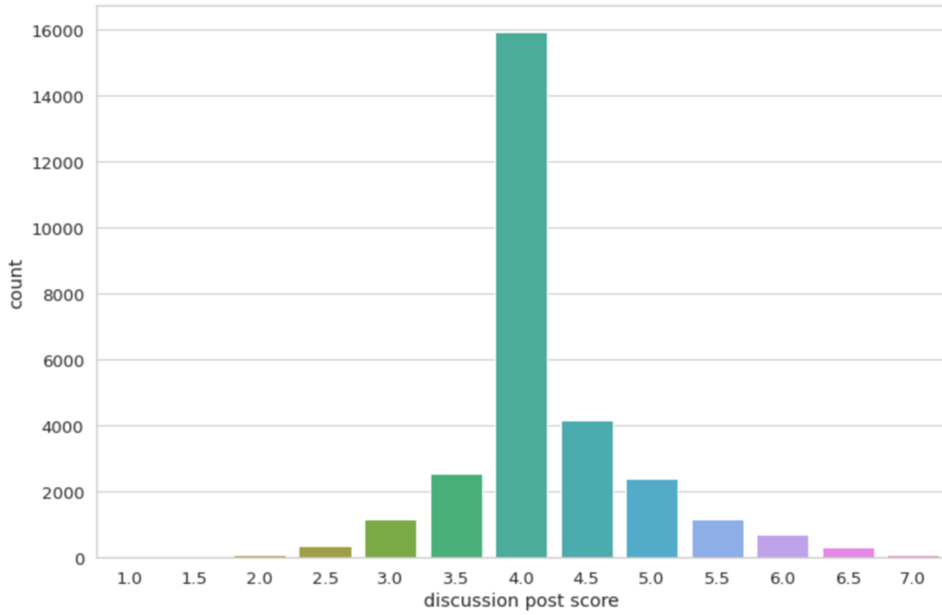
Once a learner is identified as *Persistent, Regular*, or *Irregular*, according to the predictive model, the information will be reflected on the SF-LAD.

5.9. XLNet (NLP) model results

The sentiment analysis model for this study was developed using the Stanford MOOC posts dataset (Agrawal & Paepcke, n.d.). The dataset has 29,000 anonymized students' discussion board posts from eleven Stanford University public SPOL courses (Agrawal & Paepcke, n.d.). A bar plot was created to gain a better understanding of the distribution of sentiment of posts in the dataset as shown in Figure 41.

Figure 41

Discussion Board Posts' Sentiment Score vs Count for Stanford MOOC Dataset



Of the 29,000 discussion posts, 8,830 discussion posts have a sentiment score of 4.5 or higher, which represents a positive sentiment, and 15,937 discussion posts have a sentiment score between 3.5 and 4.5, which represents a neutral sentiment. The remaining 4,233 discussion posts have a sentiment score of 3.5 or less, which represents a negative sentiment.

Research has shown that student-to-student engagement and interactions through discussion boards are an important learning activity because they creates a venue for students to support each other (Li & Xing, 2021). As a result, providing students with information about their discussion post sentiment can create self-awareness, helping students to make micro adjustments to their communication strategy with their peers. A study by Wen, Yang, and Rosé (2014) posits that making students aware of their course sentiment could give them a chance to engage in social learning in self paced online learning (SPOL). It can be achieved by providing students with important information about their attitudes before and during the course (Wen, Yang, & Rose, 2014).

To motivate the students who generate stand-alone discussion board posts (i.e., no engagement or response from their peers), this study hypothesizes that the students' sentiment scores generated by the sentiment analysis model for their discussion board posts can act as a signal to create self-awareness about their sentiment and how it could affect their engagement (or lack of) with their peers. Such information can be meaningful to students because previous studies have shown a positive relationship between learning outcomes and social support. The studies found that information support in an online setting could help students achieve the intended course outcomes (Li & Xing, 2021).

To present how the model works to create transparency in students' sentiments, the sentiment scores are shown for one student in the Stanford MOOC posts dataset (Agrawal & Paepcke, n.d.). Table 17 presents the discussion posts with the positive and negative scores for one student enrolled in a SPOL course during June and July 2013. The results show that the student's sentiment in the discussion posts is neutral and positive. Figure 42 visualizes variation in this student's discussion board posts' sentiment score during the course.

Table 17

Posts, Predicted Sentiment and Scores for June and July 2013 Discussion Board Posts for a Student Enrolled in a Stanford University Online Education Course

Discussion Post	Date	Sentiment	Positive Score	Negative Score
Hello! In my opinion, I prefer to use Graph pad prism above SPSS and other statistical software. This software combines scientific graphing, comprehensive curve fitting, understandable statistics, and data organization. Prism is now used much more broadly by all kinds of biologists, as well as social and physical scientists.	2013-06-11 6:24:00 PM	Neutral	0.9956	0.0010
I can't see any reason to exclude Excel for small datasets that wouldn't also apply to most other software let's say you send me an Excel spreadsheet 1) you and I have different versions 2) I may not have that software even installed 3) it costs too much for me to buy 4) it doesn't run on Linux (or possibly Mac) but like I said, those reasons apply to almost any software R of course if free, runs on windows, Linux and Mac but the learning curve may be more difficult than Excel	2013-06-11 7:15:00 PM	Neutral	0.9983	0.0012
I am not looking for an alternative to R. I am fine with R. I was wondering how statistics were going to be taught without any open-source software for commonality of use and	2013-06-11 8:34:00 PM	Neutral	0.9962	0.0028

explanation. Also, discussion of model answers to homework.				
nonsense	2013-06-11 9:07:00 PM	Neutral	0.9974	0.0011
I seem to remember reading somewhere ages ago that Excel can introduce significant errors and shouldn't be used at all for what we could call serious analyses (i.e., for publications). Since there are so many much better free packages available, it makes sense *not* to use it in my opinion. Doing these simple calculations by hand also reinforces the concept - the mere fact of physically using your hands helps memory. I just finished a course on Special Relativity, lots of equations, and I can say it certainly did help using paper and pencil.	2013-06-12 9:05:00 AM	Positive	0.0389	0.0007
Excel is great but it does not have built-in functions for nonparametric tests as well as anova, but I could be wrong. examples include Wilcoxon rank-sum, spearman correlation, and testing for homogeneity to name a few.	2013-07-24 4:47:00 AM	Neutral	0.9983	0.0010

Figure 42

Discussion Board Posts' Sentiment for a Student Enrolled in Stanford University Online Education Course

Note that to accurately determine sentiment, two scores are calculated for each discussion post, a score for the negative sentiment and a score for the positive sentiment. For example, post 1 has positive sentiment score of 0.9956 and negative sentiment score of 0.0010 and is labeled as positive.

To verify the validity and accuracy of the predicted sentiment for each post, the Precision, Recall, and F1 scores were calculated. Precision is defined as the ratio of true positives to all items that the model has marked as positive (i.e., the number of true positives plus the number of false positives) (Stratis, 2020). The Recall is defined as the ratio of true positives to all reviews that are actually positive; it can be calculated by dividing the number of true positives by the total number of true positives and false negatives (Stratis, 2020). F1-score is defined as the harmonic mean of Precision and Recall values of a model (Shung , 2018). Support is the number of actual occurrences of the class in the specified dataset (Shung , 2018). Using these measurements, we can determine the overall accuracy of the prediction. The final Precision, Recall, F1-score, and Support for the model for the test dataset are presented in Table 18.

Table 18

Precision, Recall, and F1-Score for Sentiment Predictions

	Precision	Recall	F1 Score	Support
Negative	0.64	0.62	0.63	1084
Neutral	0.80	0.80	0.80	3981
Positive	0.78	0.80	0.79	2185
Model Accuracy (overall accuracy of prediction for test data): 0.77				**7250**

Chapter 6: Conclusion

This research is based on recognition that self-regulation doesn't happen consistently in learning because it usually requires additional effort (da Moda, 2022). Self-regulation can be defined as the degree of students' active participation in their own learning process meta-cognitively, motivationally and behaviourally (da Moda, 2022) The additional effort can potentially be motivated by visualization, using students' generated learning analytics (representation of students' learning behaviour and achievements via dashboards). This study explores how data visualization can be integrated with online learning to improve learners' performance through enhancing their skills in planning and organization by improving their self-regulation. To enhance self-regulatory behavior and address some of the challenges faced by students in SPOL courses, a SF-LAD is designed and evaluated for this study aiming to increase teaching efficacy, remove the feeling of isolation, and improve technologies to support learning online.

The research hypothesis is that there is a correlation between creating immediate visibility into students' progress (compared to both oneself, and others in the class) and self-regulation, which leads to enhanced performance in the online courses. It worth noting that students' engagement can have a direct impact on students' final grades, retention of material, and the course dropout rate (Mushtaq, Zhu, Zhang, & Abidi, 2018) and could lead to self-reflection, which influences students' motivation (Barnard, Lan, To, Paton, & Lai, 2009). Online learning environment is autonomous; therefore, self-regulation is vital for success in online learning and improving self-regulatory behavior can lead to improved academic performance (Barnard, Lan, To, Paton, & Lai, 2009).

The main research questions that this study attempted to answer are how to design an effective SF-LAD to enhance students' organizational skills, self-regulation, and motivation in online learning; and how to evaluate the effectiveness of this dashboard in terms of enhancing students' organizational skills, self-regulation, and motivation in online learning. To answer the research questions, a prototype SF-LAD was designed based on theory and analysis of the existing learning dashboards. Also, SF-LAD effectiveness (usability) was tested using empirical investigation based on primary data collection from a sample group in a post-secondary institution to examine the effectiveness of the dashboard and evaluate if online learners will adopt this tool.

The goal of this study is to use theory, focusing on the micro-level of learning analytics to support metacognitive regulation (self-regulation), and adopting the strategies that online students can use to plan, monitor, and regulate their cognition (Fleur, Bos, & Bredeweg, 2023). Some of the strategies related to self-regulation that were examined to design the SF-LAD are self-evaluation, organizing and transforming, goal setting and planning, seeking information, and keeping records and monitoring (da Moda, 2022) to enhance engagement, motivation and metacognition to improve academic achievement.

SF-LAD designed for this research includes (1) a function to track and visualize students' progress over time (compared to self and class average) and time to submit each assignment (giving a time management score to the student) to improve organizational skills and motivation; (2) a function to track Average Score, the Average Number of Logins, the Average Number of Contents Read, the Average Assessment Lateness Indicator, and the Average Hours Took to Submit Assessments. Predictive analytics clustering model (K–mean clustering) is used to sort data from LMS into separate groups based on similar attributes (Persistent Students, Regular Students, and Irregular Students), using secondary data: the Student Performance and Engagement

Prediction in eLearning dataset (Moubayed, Injadat, Shami, BouNassif, & Lutfiyya, 2020), the goal of separating students into groups is to discover insight into students' engagement and self-regulation patterns of behaviour; and (3) a function to track student posts' sentiment in the online course discussion board based on Natural Language Processing (NLP) to extract and identify sentiments in text data and classify the sentiments expressed as positive, neutral, and negative, using an auto-regressive language model (XLNet) and secondary data: the Stanford MOOC posts dataset (Agrawal & Paepcke, n.d.). Sentiment analysis was used to allow discovering how students' sentiment affects their course engagement over time.

Subsequently, the designed SF-LAD was evaluated in a survey to assess how users perceived the usability of the dashboard. The mean SUS score from survey of 38 online students is 67.1, which falls in the 50th percentile. It is interpreted that SF-LAD outperforms 50% of the systems in the database and has a decent usability. To verify this finding, we examined the systematic evaluation of education technology by Vlachogianni & Tselios (2021). They found that mean usability levels change depending on the type of educational technology (Vlachogianni & Tselios, 2021). Therefore, we used the mean SUS score for the 77 internet platform educational technologies (66.25) by Vlachogianni & Tselios (2021) as benchmark. Comparing the benchmark (66.25) with SF-LAD mean SUS score (67.1) provides insight about the usability of SF-LAD compared to similar category of educational technology; and confirms a decent usability for SF-LAD. This is in-line with Vlachogianni and Tselios (2021) study's findings that internet platforms generally have good usability levels but not without flaws (Vlachogianni & Tselios, 2021).

For deeper insight and effective communication of mean SUS score for SF-LAD, additional methods of interpretation were used, including letter grade (C), adjective (OK), and acceptability (marginally acceptable), all of which are the same as our benchmark for internet platform educational technology. We also used NPS as a powerful metrics in the educational technology for continual improvement. NPS for SF-LAD (-6%) is interpreted as insufficient. It means that the quality of the system is perceived insufficient by the users and requires improvement (Baquero, 2022). The valuable insight gathered through the participants' feedback and suggestions can be used to improve users' satisfaction in the next iteration of the design.

The survey gathered demographic information about the participants (gender and age range) to verify representation of diversified online students' population and examine if any interesting pattern in observed. The analysis of SUS scores didn't reveal any specific pattern based on the demographic data.

To substantiate the research hypothesis, two additional questions were included in the survey. The important finding is that 45% percent of the respondents Strongly Agree or Agree that SF-LAD improves self-reflection, and 47% Strongly Agree or Agree that SF-LAD helps students to improve their course engagement and performance; which is a higher percentage compared to 24% being Neutral; and 32% Strongly Disagree or Disagree that SF-LAD improves self-reflection and 29% Strongly Disagree or Disagree that SF-LAD helps students improve their course engagement and performance. The observation is that the opinion about the hypothesis is divided with a higher percentage of the survey respondents either Strongly Agree or Agree with the hypotheses. Another interesting pattern emerged from the analysis of the hypothesis verification questions. There is a significant difference between how younger respondents viewed the hypotheses, compared to more mature respondents. In age bracket 18-30 only 17% Strongly Disagree or Disagree with the first hypothesis and 11% to the second hypothesis; compared to age bracket 31-40 (43% Strongly Disagree or Disagree with the first hypothesis and 36% to the second hypothesis) and 41-50 age bracket (60% Strongly Disagree or Disagree with the first hypothesis and 80% to the second hypothesis). The finding is that the younger online students are more likely to find student facing learning dashboards helpful for self-reflection, leading to enhanced engagement in the course and improved performance.

To gather deeper insight, participants were asked to provide their comments, feedback and suggestions for improvement of SF-LAD. 28 respondents complied and provided comments, with contradictory feedback about some of the SD-LADs features that speaks to different type of needs and/or concerns by various users. This indicates an opportunity for introducing customization to student facing dashboards design by adding opt-out option for some features that evoke conflicting response in users to better serve the diverse needs of a diverse audience.

It is worth noting that the survey respondents' view on creating transparency to students' performance compared to class average and discussion board posts' sentiments was divided. While some participants found these features helpful, some of the respondents pointed out that the information about students' performance compared to the class average would contribute to anxiety and depression. Also, some of the participants expressed privacy concerns about revealing the projected sentiment in the discussion board posts. Therefore, these two features are the candidates for customization in the next iteration of SF-LAD by adding opt-out option to create flexibility and effectively serve online students' diverse needs.

To better understand the divided perception about the peer comparison, we looked at the Social Comparison Theory, which posits that there is a "drive within individuals to look to outside images in order to evaluate their own opinions and abilities" (Festinger, 1957, p. 16). The hypothesis associated with Social Comparison Theory is that humans have a drive to evaluate themselves by comparing their abilities with others; however, the tendency to compare oneself with others decreases as the difference between their ability or opinion diverge" (Festinger, 1957).

The practical use of social comparison with similar people is the idea that it enables individuals to generate an accurate evaluation about their aptitudes and opinions (Festinger, 1957). Multiple models have been developed to specify to what extent similarity with peers and the (preferred) direction of the comparison play a role in an individual's self-assessment and self-worth (Festinger, 1957). A study by Fleur et al. (2023) explored application of social comparison in learning analytics dashboards for supporting motivation and academic achievement, and provided evidence that carefully designed social comparison, rooted in theory and empirical evidence, can be used to boost motivation and performance (Fleur, Bos, & Bredeweg, 2023). The meta-analysis reveals that the impact of social comparison on motivation changes based on its orientation. It means that upward comparison, which is defined as comparison with better-off peers, seems to be generally favored over downward comparison, which is defined as comparison with worse-off peers, because it creates a pressure for conformity with the better off group. It is important to note that upward comparison is effective only if threat to self-esteem is not present. In other words, matching the upward performance of peers appears more achievable when they are only slightly ahead than when they seem to be far and potentially out of reach performance (Fleur, Bos, & Bredeweg, 2023).

These findings are consistent with the divided feedback about the impact of peer comparison for SF-LAD, which supports adding customization and flexibility to opt-out for some features in the dashboard. Taking into account that motivation plays a key role in whether learners will engage in self-regulated learning, social comparison can be a significant driver in increasing motivation performance (Fleur, Bos, & Bredeweg, 2023). Therefore, based on theory and the collected data, the finding of this study confirms that social comparison features are important for learning dashboard effectiveness; however, their application should integrate Social Comparison Theory in the dashboard design to ensure the feature is implemented effectively.

To summarize the data analysis findings from survey comments, the SF-LAD features are most effective in 1) helping students with managing their time, 2) helping students to understand how well they are doing in the course, 3) allowing to have a sense of how one felt during the entire semester, 4) creating visibility into where one stands in the course (by comparing to self and peers) to motivate action, and 5) being user-friendly, self-explanatory, and visually appealing.

The study found that visualizations by themselves do not provide the full context. Based on the feedback received, the visualizations should include a brief description for each KPI (e.g., "what is measured" or "how is it measured") to help users to better understand and interact with the dashboard. In addition, the participants felt that dashboard would benefit from "call to actions". They felt that this additional information would improve their performance by letting them know what they need to change to reach their desired outcomes. This is supported by emerging research, which has found that to be successful in online learning environment, it is of great importance to develop self-directed learning and self-regulation as learners have to make self-decisions about their learning process and effectively manage this process. Students who still haven't developed self-directed learning and self-regulation skills require external support and guidance to enable them to make effective decisions about their learning process and improve their skills. A key part of this process is feedback (i.e., call to action) to support student's decision making (Yilmaz & Yilmaz, 2020).

This study identified that it is more likely that younger students (=<30 years old) to find SF-LAD helpful for self-reflection, leading to increased engagement and improved performance. The reason might be due to a wider gap in self-regulatory skills and the need for additional support to develop these skills in younger students (=<30), compared to more mature learners with more experience with autonomous tasks and self-regulation. Moreover, increased exposure of younger students to internet platform learning technologies while growing up compared to more mature learners might also play a role in their openness to using educational technology for increased virtual engagement in an online environment.

Another finding is the need to consider variation in motivating factors for student population and the need for student facing learning dashboard designers to be mindful that a motivating approach for a group of students might have a diverse impact on another group of students (e.g., peer comparison). Therefore, the design should proceed with flexibility in mind. This can be achieved by enabling customization in dashboards that allows online students to opt-in or opt-out of some of the dashboard features (e.g., sentiment analysis and peer comparison for SF-LAD).

This research contributes to knowledge and practice by designing and developing a SF-LAD based on theory to create visibility into time management behaviour, performance (compared to self and peers), and sentiment in the course to improve self-regulatory behavior; and testing it to determine its usability and insight about factors that can impact its adoption by online students. The theoretical contribution of this study is integrating the existing literature findings related to SF-LAD and developing a model to show how the SF-LAD designed for this research can hypothetically contribute to improved performance in online environment and collecting primary data to test the hypothesis. The practical contribution of this research is providing a tool for academic institutions to potentially enhance online learners' self-regulation and motivation.

This research demonstrated that a significant percentage of students would find a theory based SF-LAD dashboard useful, and the information presented in the dashboard could help them to improve their self-reflection and increase their course engagement and performance in self-paced online environment.

The findings of this study extend our understanding of student facing learning dashboards technology in online courses. Also, the study shed light on where this tool can potentially be most helpful (i.e., younger online learners). The findings also identified the areas that require customization to add more flexibility, depending on individual learners' preference (i.e., sentiment analysis and comparison with peers). The findings will be useful for educators and researchers to further enhance the student facing dashboards to better support online learning and more effective application in online learning institutions.

The study findings are limited due to the narrow time frame to test the SF-LAD; as learners' adoption and engagement with the tool can increase over time and can change their learning outcome. This presents an opportunity for further research by measuring the correlation between the SF-LAD developed for this research and learners' performance over an extended period of time. The study findings are limited by the absence of a longitudinal study to observe how student interactions and the use of SF-LAD, designed specifically for this research, impact and alter their online learning behavior over an extended time.

It worth noting that the significance of this study is taking into account the affective domain in development of the SF-LAD by creating visibility into students' sentiment in their discussion board posts. This information can help students to better understand the emotions they exhibit in a course, which based on literature allows them to better engage with their peers through increased self-awareness (Moreno-Marcos, Alario-Hoyos, & Munoz-Merino, 2018).

Future research opportunities include using the insights gained from testing SF-LAD to provide students with pedagogically sound feedback and recommendations (call to action), which can lead to students having better engagement and performance in online courses. For generation of the feedback for students, we can use state of the art large language models (LLMs) such as ChatGPT to provide students with contextualized feedback (Yan, Maldonado, & Gašević, 2024). Integration of GenAI tools and LLM models allow researchers to ensure that feedback provided to students align with educational goals and uphold the principals of inclusive education while navigating the ethical complexities introduced by advanced technologies. This is because these tools have demonstrated an ability to generate learning products indicative of high metacognitive, cognitive and social skills that often surpasses the quality produced by average students (Yan, Maldonado, & Gašević, 2024)

Another opportunity is to visualize more detailed information in the future iterations of the dashboard, such as students' time-management efficiency, and using the data obtained from the LMS and clickstream as predictor of academic performance. It can reveal students who are struggling with effective time-management and create visibility into the impact of their time management skills with their course performance. This can be a motivating factor in improving the students' self-regulation and course performance. Other opportunities for future research are further development of the sentiment chart to provide students with more contextual information, adding additional time management features to improve the usability of the dashboard, updating the dashboard interface so that it feels more modern and in line with current trends, and enabling customization (e.g., opt out for some features) in the dashboard.

An important aspect of online learning is students learning behavior. A study by Qiu, et. Al 2016 found that tracking of students learning behavior in online courses is key for students to be able to evaluate their own performance in an objective and quantitative way (Qiu, et al., 2016). As such, another opportunity is to perform a longitudinal study for duration of an entire semester in a course to observe how the SF-LAD designed for this study changes students' learning behavior and if any noteworthy patterns emerge.

References

10 benefits of Net Promoter Score and how to maximize it. (n.d.). Retrieved from QuestionPro: https://www.questionpro.com/blog/benefits-of-net-promoter-score/#:~:text=The%20benefits%20of%20Net%20Promoter%20Score%20can%20significantly%20boost%20your,and%20make%20your%20business%20grow.

Agnihotri, L., Aghababyan, A., Mojarad, S., Riedesel, M., & Essa, A. (2015). Mining Login Data For Actionable Student Insight. *Conference: Educational Data Mining*, 472-475.

Agrawal , A., & Paepcke, A. (n.d.). *The Stanford MOOCPosts Data Set*. (datastage.stanford.edu) Retrieved January 31, 2022, from https://datastage.stanford.edu/StanfordMoocPosts/

Aguilar, S. J., Karabenick, S. A., Teasley, S. D., & Baek, C. (2021). Associations between learning analytics dashboard exposure and motivation and self-regulated learning. *Computers & Education*, 162.

Alblawi, A. S., & Alhamed, A. A. (2017). Big Data and Learning Analytics in Higher Education Demistifying Variety, Acquisition, Storage, NLP, and Analytics. *2017 IEEE Conference on Big Data and Analytics*, (pp. 124-129).

Al-Jarrah, O., Yoo, P., Muhaidat, S., Karagiannidis, G., & Taha, K. (2015). Efficient Machine Learning for Big Data: A Review. *Big Data Research, 2*(3), 87-93.

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., van Langen, J., & & Kievit, R. A. (2021). *Raincloud plots: a multi-platform tool for robust data visualization.* Retrieved from NIH: National Library of Medicine: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6480976/

Altrabsheh, N., Gaber, M. M., & & Cocea, M. (2013). SA-E: sentiment analysis for education. In *Intelligent Decision Technologies* (pp. 353-362). IOS Press.

Arafat, S., Aljohani, N., Abbasi, R., Hussain, A., & Lytras, M. (2017). Connections between e-learning, web science, cognitive computation and social sensing, and their relevance to learning analytics: A preliminary study. *Computers in Human Behavior*(92), 478-486.

Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using Learning Analytics to Increase Student Success. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 267-270.

Azar, S., & Zafer, D. (2013). Confirmatory Factor Analysis of Time Management Behaviour Scale: Evidence from Pakistan. *Interdisciplinary Journal of Contemporary Research in Business*, 946-959.

Baker, R. S. (2010). Data Mining for Education. *International Encyclopedia of Education (3rd edition)*, 1-14.

Bakharia, A., & Dawson, S. (2011). SNAPP: A Bird's-Eye View of Temporal Participant Interaction. *Proceedings of the 1st international conference on learning analytics and knowledge*, 168-173.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 122-147.

Bangor, A., Kortum, P., & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of User Experience*, 114-123.

Baquero, A. (2022). Net Promoter Score (NPS) and customer satisfaction: relationship and efficient management. *Sustainability*, https://www.mdpi.com/2071-1050/14/4/2011.

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S. L. (2009). Measuring self-regulation in online and blended learning environments. *Internet and Higher Education*, 1-6.

Bholowalia, P., & Kumar, A. (2014). EBK-Means: A Clustering Technique based on Elbow Method and K-Means in WSN. *International Journal of Computer Applications, 105*(9), 0975-8887.

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics. *U.S. Department of Education*, 1-26.

Blattgerste, J., Behrends, J., & Pfeiffer, T. (2022). A web-based analysis toolkit for the system usability scale. *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments*, (pp. 237-246).

Blumenfeld, P. C., K. T., & Krajcik, J. S. (2006). Motivation and cognitive engagement in learning environments. In *Motivation and cognitive engagement in learning environments* (pp. 475-488).

Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). Open learner models and learning analytics dashboards: A systematic review. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, (pp. 41-50).

Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist*, 100-112.

Box and Whisker Plots. (n.d.). Retrieved from Tableau from Salesforce: Understanding and using Box and Whisker Plots (Box Plots): https://www.tableau.com/data-insights/reference-library/visual-analytics/charts/box-whisker

Brown, V. (2023, August 08). *How to use the System Usability Scale in modern UX*. Retrieved April 24, 2024, from LogRocket: https://blog.logrocket.com/ux-design/use-system-usability-scale-modern-ux/#case-study

Burrus, A. (2019). *What does time mamangement mean to you? Exploring measures of time management and group differences.* University of Missouri-Saint Louis.

Carron, J. (2021). *Violin plots 101: visualizing distribution and probability density* . Retrieved from Mode: https://mode.com/blog/violin-plot-examples

Chaplot, D., Rhim, E., & Kim, J. (2015). Predicting Student Attrition in MOOCs using Sentiment Analysis and Neural Networks. *CEUR Workshop Proceedings.*

Chen, L., Min, L., Goda, Y., Shimada, A., & Yamada, M. (2021). Learning Analytics Dashboard Supporting Metacognition. *Learning Analytics Dashboard Supporting Metacognition*, 129-149.

Cohen, A., & Nachmias, R. (2006). A quantitative cost effectiveness model for Web-supported academic instruction. *The Internet and Higher Education, 9*(2), 81-90.

Cohen, A., & Nachmias, R. (2006). A quantitative cost effectiveness model for Web-supported academic instruction. *The Internet and Higher Education, 9*(2), 81-90.

Conijn, R., Van den Beemt, A., & Cuijpers, P. (2017). Predicting student performance in a blended MOOC. *Journal of Computer Assisted Learning*, 615-628.

Çöpgeven, N., & Firat, M. (2024). Effects of dashboard usage on learning interactions and academic achievement of distance education students. *Journal of Educators Online*, 16-20.

Corrin, L., Kennedy, G., French, S., Shum, S. B., Kitto, K., Pardo, A., . . . Colvin, C. (2019). The Ethics of Learning Analytics in Australian Higher Education. *Accessed online*, 1-32.

Cotton, K., & Savard, W. (1981). *Time factors in learning. Research on school effectiveness project: Topic summary report.* Retrieved 07 14, 2020, from Proquest: http://search.proquest.com.access.ewha.ac.kr/docview/6354 4228? accountid=10785.

da Moda, R. N. (2022). *Visual dashboard for self-regulation and co-regulation of online learning.* Retrieved from https://repositorio-aberto.up.pt/bitstream/10216/145145/2/590865.pdf

Davies , J., & Graff, M. (2005). Performance in e-learning: online participation and student grades. *British Journal of Educational Technology, 36*(4), 657–663.

Dey, R. (2023). *Understanding Principle Component Analysis (PCA)* . Retrieved from Medium: https://medium.com/@roshmitadey/understanding-principal-component-analysis-pca-d4bb40e12d33

Dietz-Uhler, B., & Hurn, J. E. (2013). Using Learning Analytics to Predict (and Improve) Student Success: A Faculty Perspective. *Journal of Interactive Online Learning, 12*(1), 17-26.

Dipace, A., Fazlagic, B., & Minerva, T. (2019). The Design of A Learning Analytics Dashboard: Eduopen MOOC Platform Redefinition Procedures. *Journal of e-Learning and Knowledge Society, 15*(3), 30-47.

Duan, X., Wang, C., & Rouamba, G. (2022). Designing a Learning Analytics Dashboard to Provide Students with Actionable Feedback and Evaluating Its Impacts. *Proceedings of International Conference on Computer Supported Education*, 1-11.

Durairaj, M., & Vijitha, C. (2014). Educational Data mining for Prediction of Student Performance Using Clustering Algorithms. *International Journal of Computer Science and Information Technologies, 5*(4), 5987-5991.

Durall, E., & Gros, B. (2014). Learning Analytics as a Metacognitive Tool. *6th International Conference on Computer Supported Education*, 380-384.

Duval, E. (2011). Attention please!: learning analytics for visualization and recommendation. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, (pp. 9-17).

Dykes, B. (2015). Data storytelling: What it is and how it can be used to effectively communicate analysis results. *Applied Marketing Analytics, 1*(4), 299-315.

Echeverria, V., Martinez-Maldonado, R., Shum, S. B., Chiluiza, K., Granda, R., & Conati, C. (2018). Exploratory versus Explanatory Visual Learning Analytics: Driving Teachers' Attention through Educational Data Storytelling. *Journal of Learning Analytics, 5*(3), 72-97.

Elakia, Gayathri, Aarthi, & Naren J. (2014). Application of Data Mining in Educational Database for Predicting Behavioural Patterns of the Students. *International Journal of Computer Science and Information Technologies, 5*(3), 4649-4652.

Essa, A., & Ayad, H. (2012). Student Success System: Risk Analytics and Data Visualization using Ensembles of Predictive Models. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 158 - 161.

Estrada, M., Cabada, R., & Bustillos, R. (2020). Opinion mining and emotion recognition applied to learning environments. *Expert Systems With Applications, 150*, 1-12.

eThink. (2017, May 29). *How to Collect and Measure Moodle Learning Data with an xAPI LRS*. Retrieved June 14, 2020, from eThink: https://ethinkeducation.com/blog/measure-moodle-data-xapi-lrs/

Farahmand , A., Dewan, M., & Lin, F. (2020). Student-facing educational dashboard design for online learners. *IEEE Cyber Science and Technology Congress.* Calgary, Canada.

Farahmand, A., Dewan, M., & Lin, F. (2021). Constructing Intelligent Learning Dashboard for Online Learners. *IEEE Cyberscience and Technology conference 2021.*

Farahmand, A., Yan, H., Dewan, M., & Lin, F. (2022). Removing Barriers in Self-Paced Online Learning through Designing Intelligent Learning Dashboards. In A. Farahmand, H. Yan, M. A. Dewan, & F. Lin, *Artificial Intelligence in STEM Education* (p. 14). CRC Press.

Fazlagic, B., Dipace, A., & Minerva, T. (2019). The Design of a Learning Analytics Dashboard. *Journal of E-Learning and Knowledge Society, 15*(3), 30-47.

Festinger, L. (1957). *Social comparison theory.* Retrieved from Selective Exposure Theory: http://www.bahaistudies.net/asma/selective_exposure-wiki.pdf#page=18

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 906-911.

Fleur, D. S., Bos, W. v., & Bredeweg, B. (2023). Social comparison in learning analytics dashboard supporting motivation and academic achievement. *Computers and Education Open*, 1-10.

Friend Wise, A., Zhao, Y., & Hausknecht, S. (2014). Learning Analytics for Online Discussions: Embedded and Extracted Approaches. *Journal of Learning Analytics*, 48-71.

Frost, J. (n.d.). *Cronbach's Alpha: Definition, Calculations & Example*. Retrieved from Statistics by Jim: https://statisticsbyjim.com/basics/cronbachs-alpha/#:~:text=Analysts%20frequently%20use%200.7%20as,industries%20have%20different%20minimum%20values.

Gašević, D., Dawson, S., & Siemens, G. (2014, December 21). Let's not forget: Learning Analytics are about Learning. *TechTrends, 59*, 64-71.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 64-71.

Genç, Ö. (2019). *The basics of NLP and real time sentiment analysis with open source tools*. (towards data science) Retrieved 12 12, 2021, from https://towardsdatascience.com/real-time-sentiment-analysis-on-social-media-with-open-source-tools-f864ca239afe

Gillett-Swan, J. (2017). The Challenges of Online Learning Supporting and Engaging the Isolated Learner. *Journal of Learning Design, 11*(1), 13-23.

Griffin, B. (2023). *Calculating Cronbach Alpha Using Excel*. Retrieved from Psychology of Business: https://psychologyofbusiness.beehiiv.com/p/calculating-cronbachs-alpha-excel

Haarman, B. C., Riemersma-Van der Lek, R. F., Nolen, W. A., Mendes, R., Drexhage, H. A., & Burger, H. (2015). Feature-expression heat maps – A new visual method to explore complex associations between two variable sets. *Journal of Biomedical Informatics, 53*, 156-161.

He , H., Dong, B., Zheng, Q., Di, D., & Lin, Y. (2019). Visual Analysis of the Time Management of Learning Multiple Courses in Online Learning Environment. *2019 IEEE Visualization Conference*, 1-5.

Hew, K., Hu, X., Qiao, C., & Tang, Y. (2020). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. *Computer and Education, 145*.

Hodges, C. B. (2008). Self-Efficacy in the context of online learning environments. *Performance Improvement Quarterly*, 7-25.

Holder, B. (2007). An investigation of hope, academics, environment, and motivation as predictors of persistence in higher education online programs. *The Internet and Higher Education, 10*(4), 245-260.

Holder, B. (2007). An investigation of hope, academics, environment, and motivation as predictors of persistence in higher education online programs. *The Internet and Higher Education, 10*(4), 245-260.

How to calculate & measure Net Promoter Score (NPS). (n.d.). Retrieved from qualtrics.XM: https://www.qualtrics.com/experience-management/customer/measure-nps/#:~:text=Subtract%20the%20percentage%20of%20detractors%20from%20the%20percentage%20of%20promoters.&text=(The%20percentage%20of%20passives%20is,be%2070%2D10%20%3D%2060.

Howlin, C., & Lynch, D. (2014, October 1). Learning and Academic Analytics in the Realizeit System. *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, 1-11.

Huang, B., & Hew, K. (2016). Measuring Learners' Motivation Level in Massive Open Online Courses. *International Journal of Information and Education Technology, 6*(10), 759-764.

Hwang, W.-Y., & Wang, C.-Y. (2004). A study of learning time patterns in asynchronous learning environments. *Journal of Computer Assisted Learning, 20*(4), 292-304.

IGI Global. (n.d.). *What is Computational Cost*. Retrieved October 03, 2021, from https://www.igi-global.com/dictionary/computational-cost/82637

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. *In Data Driven Approaches in Digital Education" 12the European Conference on Technology Enhanced Learning* (pp. 82-96). Tallinn, Estonia: Springer International Publishing.

Jo, I.-H., Kim, D., & Yoon, M. (2014, March 28). Analyzing the Log Patterns of Adult Learners in LMS Using Learning Analytics. *LAK '14 Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, 183-187.

Jones, K. M. (2019). Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. *International Journal of Educational Technology in Higher Education, 16*(24), 1-22.

Karan , P. (2019, July 10). *Understand how the XLNet outperforms BERT in Language Modelling*. (medium) Retrieved December 28, 2022, from https://medium.com/saarthi-ai/xlnet-the-permutation-language-model-b30f5b4e3c1e

Kastrati, Z., Dalipi, F., Imran, A., Nuci, K., & Wani, M. (2021). Sentiment Analysis of Students' Feedback with NLP and Deep Learning: A Systematic Mapping Study. *applied sciences*, 1-23.

Kay, J., & Upton, K. (2006). Narcissus: Group and Individual Models to Support Small Group Work. *Conference: Proceedings of the 17th International Conference on User Modeling, Adaptation, and Personalization*, 3-13.

Khandelwal, R. (2020, December 5). *Evaluating goodness of clustering for unsupervised learning case*. (Towards Data Science) Retrieved Feburary 7, 2021, from https://towardsdatascience.com/evaluating-goodness-of-clustering-for-unsupervised-learning-case-ccebcfd1d4f1

Kim, D., Jo, I.-H., & Yoon, M. (2015). Constructing Proxy Variables to Measure Adult Learners' Time Management Strategies in LMS. *Educational Technology & Society, 18*(3), 214-225.

Kim, K.-J., & Bonk, C. J. (2006). The Future of Online Teaching and Learning in Higher Education: The Survey Says…. *EDUCAUSE QUARTERLY, 4*, 22-30.

Klug, B. (2017). An Overview of the System Usability Scale in Library Website and System Usability Testing. *Journal of Library User Experience, 1*(6).

Koç, M. (2017, November 10). Learning Analytics of Student Participation and Achievement in Online Distance Education: A Structural Equation Modeling. *EDUCATIONAL SCIENCES: THEORY & PRACTICE, 17*(6), 1893-1910.

Kumar, A., Srinivasan, K., Cheng, W.-H., & Zomaya, A. Y. (2020). Hybrid Context Enriched Deep Learning Model for Fine-grained Sentiment Analysis in Textual and Visual Semiotic Modality Social Data. *Information Processing & Management, 57*(1), 2-35.

Lawson, P. (2022, 09 14). *How to SUS out usability scores*. Retrieved from Think Company: https://www.thinkcompany.com/blog/how-to-sus-out-usability-scores/

Leony, D., Pardo, A., Valentín, L., Castro, D., & Kloos, C. (2012). GLASS: A Learning Analytics Visualization. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 162-163.

Lewis, J. R., & Sauro, J. (2009). The Factor Structure of the System Usability Scale. *Human Centered Design: First International Conference*, (pp. 1-11).

Li , C., & Xing, W. (2021). Natural Language Generation Using Deep Learning to Support MOOC Learners. *International Journal of Artificial Intelligence in Education , 31*, 186-214.

Li, S., & Lajoie, S. P. (2022). Cognitive engagement in self-regulated learning: an integrative model. *European Journal of Psychology of Education*, 833-852.

Lim, L., Dawson, S., Joksimovic, S., & Gašević, D. (2019, March 04). Exploring Student Sensemaking of learning analytics dashboards: does frame of reference make a difference. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 250-259.

Macan, T., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students' time management: Correlations with academic performance and stress. *Journal of educational psychology*, 760-768.

Macarini, L., Cechinel, C., Machado, M., Ramos, V., & Munoz, R. (2019). Predicting students success in blended learning-evaluating different interactions inside learning management systems. *Applied Sciences, 9*(24), 1-23.

Martinez-Maldonado, R., Echeverria, V., Nieto, G., & Shum, S. B. (2020). From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics. *Proceedings of the 2020 chi conference on human factors in computing systems*, (pp. 1-15).

Marzouk, Z., Rakovic, M., Liaqat, A., Vytasek, J., Samadi, D., Stewart-Alonso, J., & ... & Nesbit, J. C. (2016). What if learning analytics were based on learning science? *Australasian Journal of Educational Technology*.

McGoldrick, G., Cao, Y., & Prince, S. (2019, July 10). *Understanding XLNet*. (borealisai) Retrieved December 21, 2022, from https://www.borealisai.com/research-blogs/understanding-xlnet/

McNamara, D. S., Allen, L. K., Crossley, S. A., Dascalu, M., & Perret, C. A. (2017). Chapter 8: Natural Language Processing and Learning Analytics. *Handbook of Learning Analytics*, 93-100.

Molenaar, I., Horvers, A., Dijkstra, R., & Baker, R. (2019). Designing Dashboards to support learners' Self-Regulated Learning. *In companion proceedings of the 9th International learning analytics & knowledge conference.*

Moreno-Marcos, P., Alario-Hoyos, C., & Munoz-Merino, P. (2018). Sentiment Analysis in MOOCs: A case study. *IEEE Global Engineering Education Conference (EDUCON).*

Moubayed, A., Injadat, M., Shami, A., & Lutfiyya, H. (2018). Relationship between Student Engagement and Performance in e-Learning Environment Using Association Rules. *IEEE world engineering education conference (EDUNINE).*

Moubayed, A., Injadat, M., Shami, A., BouNassif, A., & Lutfiyya, H. (2020). Student performance and engagement prediction in e-Learning datasets. *IEEE DataPort*.

Moubayed, A., Injadat, M., Shami, A., BouNassif, A., & Lutfiyya, H. (2020). Student Performance and Engagement Prediction in E-Learning Datasets. *IEEE DataPort*.

Mushtaq, H., Zhu, W., Zhang, W., & Abidi, S. (2018). Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores. *Computational Intelligence and Neuroscience*.

Nkomo, L. M., & Daniel, B. K. (2021). Sentiment Analysis of Student Engagement with Lecture Recording. *TechTrends, 65*, 213-224.

Onah, D. F., Pang, E. L., S. J., & U. J. (2021). An innovative MOOC platform: the implications of self-directed learning abilities to imporve motivation in learning and to support self-regulation. *The International Journal of Information and Learning Technology*, 283-298.

Padilla, L. M., Creem-Regehr, S. H., Hegarty, M., & Stefanucci, J. K. (2018, July 11). *Decision making with visualizations: a cognitive framework across disciplines*. Retrieved from SpringerOpen: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6091269/

Park, Y., & Jo, I.-H. (2015, 01 01). Development of the Learning Analytics Dashboard to Support Students' Learning Performance. *Journal of Universal Computer Science, 21*(1), 111-133.

Pearson Correlation. (n.d.). Retrieved from SienceDirect: https://www.sciencedirect.com/topics/computer-science/pearson-correlation

Pearson's Correlation Coefficient: a comprehensive overview. (n.d.). Retrieved from Complete Dissertation: by statistics solutions: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/pearsons-correlation-coefficient/#:~:text=Perfect%3A%20Values%20near%20%C2%B11,0.49%20indicate%20a%20moderate%20correlation.

Pellas, N. (2014). The influence of computer self-efficacy, metacognitive self-regulation and self-esteem on student engagement in online learning programs: Evidence from the virtual world of Second Life. *Computers in Human Behavior, 35*, 157-170.

Picciano, A. G. (2012). The Evolution of Big Data and Learning Analytics in American Higher Education. *Journal of Asynchronous Learning Networks, 16*(3), 9-20.

Piccoli, G., Ahmad, R., & Ives, B. (2001). Web-based virtual learning environments: A research framework and a preliminary assessment of effectiveness in basic IT skills training. *MIS quarterly*, 401-426.

Qiu, J., Tang, J., Liu, T., Gong, J., Zhang, C., Zhang, Q., & Xue, Y. (2016, February). Modeling and Predicting Learning Behavior in MOOCs. *Proceedings of the ninth ACM international conference on web search and data mining,*, 93-102.

Question Pro. (n.d.). *System Usability Scale: What it is, Calculation + Usage*. Retrieved April 24, 2024, from QuestionPro: https://www.questionpro.com/blog/system-usability-scale/

Rao , P. (2019, September 4). *Fine-grained Sentiment Analysis in Python (Part 1)*. (Towards Data Science) Retrieved January 05, 2023, from https://towardsdatascience.com/fine-grained-sentiment-analysis-in-python-part-1-2697bb111ed4

Rebetez, M., Rochat, L., & Linden, M. (2015). Cognitive, emotional, and motivational factors related to procrastination: A cluster analytic approach. *Personality and Individual Differences*, 1-6.

Richardson, M. (2009). *Principal Component Analysis*. Retrieved from https://people.duke.edu/~hpgavin/SystemID/References/Richardson-PCA-2009.pdf

Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016, December 19). *Student Attitudes toward Learning Analytics in Higher Education: "The Fitbit Version of the Learning World"*. Retrieved February 03, 2020 , from Frontiers in Psychology: https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01959/full

Saggi, M., & Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing and Management, 54*(5), 758-790.

Sakkaf, Y. (2020, 03 31). *Decision Trees: ID3 Algorithm Explained*. Retrieved 05 23, 2020, from Towardsdatascience: https://towardsdatascience.com/decision-trees-for-classification-id3-algorithm-explained-89df76e72df1

Santos, ,. J., Verbert, K., & Duval, E. (2012). Empowering students to re ect on their activity with StepUp!: Two case studies with engineering students. *Proceedings of ARETL'12 2nd workshop on Awareness and Reflection*, 2-14.

Santos, J., Govaerts, S., Verbert, K., & Duval, E. (2012). Goal-oriented visualizations of activity tracking: a case study with engineering students. *Proceedings of the 2nd international conference on learning analytics and knowledge*, (pp. 143-152).

Sauro, J. (2011a). *SUStified? Little known System Usability Scale facts.* Retrieved from User Experience: The Magazine of the UXPA: http://uxpamagazine.org/sustified/

Sauro, J. (2011b). *Measuring Usability with the System Usability Scale (SUS)*. Retrieved from MeasuringU: https://measuringu.com/sus/#:~:text=SUS%20is%20Reliable&text=SUS%20has%20been%20shown%20to,and%20still%20generate%20reliable%20results.

Sauro, J. (2012). *Predicting net promoter scores from system usability scale scores*. Retrieved from MeasuringU: https://measuringu.com/nps-sus/

Sauro, J. (2018). *5 Ways to Interpret a SUS Score*. Retrieved from Measuring U: https://measuringu.com/interpret-sus-score/

Scheur, O., & Zinn, C. (2007). How did the e-learning session go The Student Inspector. *Frontiers in Artificial Intelligence and Applications*, 158-487.

Schumacher, C., & & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in human behavior*, 397-407.

Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, 15*(6), 1139-1148.

Shung , K. (2018). *Accuracy, Precision, Recall or F1?* (towards data science) Retrieved January 29, 2022, from https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9

Sluijs, M., & Matzat, U. (2023). Predicting time-management skills from learning analytics. *Journal of Computer Assisted Learning*, 525-537.

Smart, P., Simperl, E., & Shadbolt, N. (2014, August 23). A Taxonomic Framework for Social Machines. *Computational Social Sciences*, 51-85.

Smith, J. (2021). *What is Rapid Prototyping for UX*. Retrieved from American Graphics Institute: https://www.agitraining.com/ux/classes/rapid-prototyping-ux

Soegaard, M. (2024). *System Usability Scale for Data-Driven*. Retrieved from Interaction Design Foundation: https://www.interaction-design.org/literature/article/system-usability-scale#what_is_a_system_usability_scale?-0

Stratis, K. (2020). *Use Sentiment Analysis With Python to Classify Movie Reviews*. (Real python) Retrieved January 1st, 2022, from https://realpython.com/sentiment-analysis-python/

Syahrial, M., & Tanti, S. (2019, June 01). research has shown that in online learning, learning occurs as a result of the following three factor personal factors (students' beliefs), external factors (learning environment). *Journal of Baltic Science Education, 18*(3), 389-403.

Tankala, S. (2022, November 27). *Why and How to Use Demographics in UX*. Retrieved June 2024 , from Nielson Norman Group: https://www.nngroup.com/articles/demographics-in-ux/

The University of British Columbia. (2014, October 1). *Learning analytics examples*. Retrieved January 12, 2021, from The University of British Columbia: https://isit.arts.ubc.ca/learning-analytics-examples/

The University of British Columbia. (2014, October 1). *Learning Analytics Examples*. Retrieved December 27, 2019, from The University of British Columbia: https://isit.arts.ubc.ca/learning-analytics-examples/

Theodotou, M. (2022). *Net Promoter Score: are your learners happy?* Retrieved from eLearning Industry: https://elearningindustry.com/net-promoter-score-are-your-learners-happy

UIUX Trend. (n.d.). *Measuring and Interpreting System Usability Scale (SUS)*. Retrieved January 2024, from UIUX Trend: https://uiuxtrend.com/measuring-system-usability-scale-sus/

University of Virgina. (n.d.). *Using and Interpreting Cronbach's Alpha*. Retrieved April 11, 2024, from University of Virgina: https://library.virginia.edu/data/articles/using-and-interpreting-cronbachs-alpha#:~:text=Cronbach's%20alpha%20is%20a%20measure,of%20scale%20or%20test%20items.

Uzir, N., Gašević, D., Jovanović, J., Matcha, W., Lim, L.-A., & Fudge, A. (n.d.). Analytics of time management and learning strategies for effective online learning in blended environments. *Tenth International Conference on Learning Analytics & Knowledge*, (pp. 392-401).

Veeramachaneni, K., O'Reilly, U.-M., & Taylor, C. (2014, July 20). Towards Feature Engineering at Scale for Data from Massive Open Online Courses. *Computer Science*, 1-16.

Venant, R., Vidal, P., & Broisin, J. (2016). Evaluation of learner performance during practical activities: an experimentation in computer education. *International Conference on Advanced Learning Technologies.* Austin, USA.

Venant, R., Vidal, P., & Broisin, J. (2016). Evaluation of Learner Performance During Practical Activities: An Experimentation in Computer Education. *IEEE 16th international conference on advanced learning technologies*, (pp. 237-241).

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. (2013). Learning Analytics Dashboard Applications. *American Behavioral Scientist, 57*(10), 1500-1509.

Verma, Y. (2022, March 4). *A complete tutorial on masked language modelling using BERT*. (analyticsindiamag) Retrieved December 28, 2022, from https://analyticsindiamag.com/a-complete-tutorial-on-masked-language-modelling-using-bert/

Vlachogianni, P., & Tselios, N. (2021). Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review. *Journal of Research on Technology in Education*, 1-19.

Wang, G., Xu, J., & He, B. (2016). A novel method for tuning configuration parameters of spark based on machine learning. *IEEE 18th International Conference on High Performance Computing and Communications.*

Wellington, J., & Szczerbinski, M. (2007). *Research method for the social sciences.* A&C Black.

Wen, M., Yang, D., & Rose, C. P. (2014). Sentiment Analysis in MOOC Discussion Forums: What does it tell us? *Educational data mining.*

What is a good Net Promoter Score? (n.d.). Retrieved from qualtrics.XM: https://www.qualtrics.com/experience-management/customer/good-net-promoter-score/

Whitehill, J., Serpell , Z., Lin, Y.-C., Foster , A., & Movel, J. (2014). The faces of engagement: automatic recognition of student engagement from facial expressions. *IEEE Transactions on Affective Computing, 5*(1), 86-98.

Wikipedia. (n.d.). *Transformer (machine learning model)*. Retrieved February 9, 2023, from https://en.wikipedia.org/wiki/Transformer_(machine_learning_model)

Winne, P. H. (2017). Chapter 21: Learning Analytics for Self-Regulated Learning. *Handbook of Learning Analytics – First edition*, 241-249.

Wolff, A., Zdrahal, Z., Nikolov, A., & Pantucek , M. (2013). Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment. *proceedings of the third international conference on learning analytics and knowledge.*

Xu, Z., & Makos, A. (2010). Investigating the Impact of a Notification System on Student Behaviours in a Discourse-Intensive Online Course.

Ya Ni, A. (2013). Comparing the Effectiveness of Classroom and Online Learning: Teaching Research Methods. *Journal of Public Affairs Education, 19*(2), 199-215.

Yan, H., Lin, F., & Kinshuk. (2020). Including Learning Analytics in the Loop of Self-Paced Online Course Learning Design. *International Journal of Artificial Intelligence in Education*, 1-18.

Yan, L., Maldonado, R., & Gašević, D. (2024, March). Generative Artificial Intelligence in Learning Analytics: Contextualising Opportunities and Challenges through the Learning Analytics. *Proceedings of the 14th Learning Analytics and Knowledge Conference*, (pp. 101-111).

Yilmaz, F. G., & Yilmaz, R. (2020). Student Opinions About Personalized Recommendation and Feedback Based on Learning Analytics. *Technology, Knowledge and Learning*, 753-768.

Zhang, W., & Qin, S. (2018). A Brief Analysis of the Key Technologies and Applications of Educational Data Mining on Online Learning Platform. *IEEE 3rd International Conference on Big Data Analysis*, 83-86.

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance. An introduction and an overview. In *Handbook of self-regulation of learning and performance* (pp. 15-26). Tylor & Francis.

Appendix A: Full System Usability Scale (SUS) Survey Result

SUS 1	SUS 2	SUS 3	SUS 4	SUS 5	SUS 6	SUS 7	SUS 8	SUS 9	SUS 10	SUS Score	Percentile	Grade	Adjective	Acceptability	Promoters and Detractors
2	1	2	3	3	2	1	3	2	1	50	2-14	F	Poor	Not Acceptable	Detractor
1	2	3	3	1	2	2	3	2	3	55	15-34	D	OK	Marginally Acceptable	Passive
4	3	4	4	3	3	4	1	3	4	82.5	90-95	A	Excellent	Acceptable	Promoter
0	3	3	4	2	3	4	3	3	4	72.5	60-64	C+	Good	Acceptable	Passive
4	3	4	4	3	2	4	2	4	3	82.5	90-95	A	Excellent	Acceptable	Promoter
0	0	0	2	0	0	2	2	2	2	25	0-1.9	F	Worst Imaginable	Not Acceptable	Detractor
2	2	2	3	3	2	2	3	3	3	62.5	15-34	D	OK	Marginally Acceptable	Passive
4	3	3	3	4	3	3	3	3	3	80	85-89	A-	Good	Acceptable	Passive
3	3	3	4	4	3	3	3	3	3	80	85-89	A-	Good	Acceptable	Passive
3	4	3	3	2	3	3	3	3	3	75	70-79	B	Good	Acceptable	Passive
3	0	3	3	3	3	3	3	3	3	67.5	41-59	C	OK	Marginally Acceptable	Passive
2	4	3	4	3	2	4	4	4	4	85	96-100	A+	Best Imaginable	Acceptable	Promoter
1	4	4	4	3	3	4	3	3	4	82.5	90-95	A	Excellent	Acceptable	Promoter
2	2	2	2	2	2	2	2	2	2	50	2-14	F	Poor	Not Acceptable	Detractor
3	4	3	2	3	3	3	3	2	4	75	70-79	B	Good	Acceptable	Passive
3	1	3	1	3	1	3	1	3	1	50	2-14	F	Poor	Not Acceptable	Detractor
3	4	3	4	3	1	3	3	4	4	80	85-89	A-	Good	Acceptable	Passive
2	3	3	3	3	2	4	4	3	4	77.5	80-84	B+	Good	Acceptable	Passive
2	2	3	0	1	1	2	2	3	2	45	2-14	F	Poor	Not Acceptable	Detractor
4	3	3	4	4	4	4	4	3	4	92.5	96-100	A+	Best Imaginable	Acceptable	Promoter
2	0	0	4	2	2	2	0	1	1	35	2-14	F	Poor	Not Acceptable	Detractor
4	3	3	4	3	3	4	4	4	3	87.5	96-100	A+	Best Imaginable	Acceptable	Promoter
2	2	3	1	3	1	3	1	3	1	50	2-14	F	Poor	Not Acceptable	Detractor
3	3	3	3	3	2	3	3	3	3	72.5	60-64	C+	Good	Acceptable	Passive
2	3	3	4	3	0	3	3	3	3	67.5	41-59	C	OK	Marginally Acceptable	Passive
1	1	2	3	2	0	1	0	3	3	40	2-14	F	Poor	Not Acceptable	Detractor
3	3	3	4	2	3	2	3	3	4	75	70-79	B	Good	Acceptable	Passive
1	3	3	3	1	3	3	3	3	3	65	41-59	C	OK	Marginally Acceptable	Passive
2	2	3	2	3	3	3	3	2	2	62.5	15-34	D	OK	Marginally Acceptable	Passive
4	4	3	4	3	3	4	4	4	4	92.5	96-100	A+	Best Imaginable	Acceptable	Passive
4	4	4	4	3	3	4	4	4	4	95	96-100	A+	Best Imaginable	Acceptable	Promoter
1	3	3	3	3	3	3	1	1	3	60	15-34	D	OK	Marginally Acceptable	Passive
0	2	2	3	3	2	1	3	1	3	50	2-14	F	Poor	Not Acceptable	Detractor
1	3	1	3	2	2	2	2	3	3	55	15-34	D	OK	Marginally Acceptable	Passive
1	3	3	4	2	1	2	3	3	3	62.5	15-34	D	OK	Marginally Acceptable	Passive
1	2	3	3	1	3	3	3	3	3	62.5	15-34	D	OK	Marginally Acceptable	Passive
3	2	3	2	3	2	3	3	3	3	67.5	41-59	C	OK	Marginally Acceptable	Passive
3	3	3	3	4	3	4	4	4	1	80	85-89	A-	Good	Acceptable	Passive

Appendix B: Participant Consent Form

Student Facing Learning Dashboard Usability Survey

Principal Researcher:

Amthony Azimi aazimi1@learn.athabascau.ca

Supervisors:

Dr. Ali Dewan adewan@athabascau.ca

Dr. Oscar Lin oscarl@athabascau.ca

You are invited to participate in a research study about student-facing learning dashboard to frequently inform each student about how they perform in a course. I am conducting this study as a part of my thesis of Master of Science in Information Systems.

As a participant, you are asked to participate in this study by completing a short anonymous online questionnaire to assess the impact of Student-Facing Learning Dashboard on self-paced online learners' self-regulation behaviour. As an appreciation of your participation, the first 30 participants who submit completed responses to the survey questionnaire will receive a $15.00 amazon gift card. Please provide your email address in part 4 of the questionnaire to receive your gift card. The email address will only be used to send the gift card and will be deleted from the collected data after the gift cards are sent to keep the survey completely anonymous.

Note that the student facing dashboard for this research is presented with simulated data from the following data sources:

- Stanford MOOC posts dataset
- Student Performance and Engagement Prediction in eLearning dataset

You are supposed to access the dashboard using the following information before answering the survey questions:

- Dashboard link: https://sf-ildresearchmvp.github.io/SF_iLD/
- Username: student000000
- Password: password

Involvement in this study is entirely voluntary and you may refuse to answer any questions or to share information that you are not comfortable with. You may withdraw from the study at any time by simply closing out of your browser. Once you submit your completed survey, however, data cannot be withdrawn as the survey is completely anonymous. Results of this study may be presented in academic conferences and/or published in academic journals. As a research participant, you can request to receive a copy of the published material by sending an email to the principal researcher at aazimi1@learn.athabascau.ca

This study has been reviewed by the Athabasca University Research Ethics Board. Should you have any comments or concerns regarding your treatment as a participant in this study, please contact the Research Ethics Officer at 780.213.2033 or by e-mail to rebsec@athabascau.ca

Thank you for your assistance in this project.

CONSENT:

The completion of the survey and its submission is viewed as your consent to participate.

BEGIN THE SURVEY

Appendix C: Survey

Part 1: Demographic Data

The demographic information will be used to assess if different genders or age groups react to the learning dashboard differently. Demographic information will help us to get a sense of how age categories/genders interact with the dashboard, giving us valuable

insights into various demographic groups' engagement with the dashboard and its perceived effect on their course performance.

Please answer the following questions:

1) What is your age

   Please choose only one of the following:

   - ○ 18-30
   - ○ 31-40
   - ○ 41-50
   - ○ 51 or above

2) What is your gender:

   Please choose only one of the following:

   - ○ Male
   - ○ Female
   - ○ Other
   - ○ Prefer not to declare

Part 2: Student Facing Learning Dashboard Usability Evaluation

Please use the link below and the provided username and password to access the student-facing learning dashboard designed for this research. You are supposed to access the dashboard using the following information before answering the survey questions.

After interacting with the dashboard answer questions 4 to 15:

- Dashboard link: https://sf-ildresearchmvp.github.io/SF_iLD/
- Username: student000000
- Password: password

1) I think I would like to use the student facing learning dashboard frequently

Please choose **only one** of the following:

- ○ Strongly Disagree
- ○ Disagree
- ○ Neither Agree nor Disagree
- ○ Agree
- ○ Strongly Agree

2) I found the dashboard unnecessarily complex

Please choose **only one** of the following:

- ○ Strongly Disagree
- ○ Disagree
- ○ Neither Agree nor Disagree
- ○ Agree
- ○ Strongly Agree

3) I thought the dashboard was easy to use.

Please choose **only one** of the following:

- Strongly Disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

4) I think that I would need the support of a technical person to be able to use the student facing learning dashboard.

Please choose **only one** of the following:

- Strongly Disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

5) I found the visualizations for this dashboard appropriate for monitoring students' self-regulation.

Please choose **only one** of the following:

- Strongly Disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

6) I found that there was too much inconsistency in the dashboard for monitoring students' self-regulation.

Please choose **only one** of the following:

- ◯ Strongly Disagree
- ◯ Disagree
- ◯ Neither Agree nor Disagree
- ◯ Agree
- ◯ Strongly Agree

7) I would imagine that most people would learn to use of this dashboard very quickly.

Please choose **only one** of the following:

- ◯ Strongly Disagree
- ◯ Disagree
- ◯ Neither Agree nor Disagree
- ◯ Agree
- ◯ Strongly Agree

8) I found the dashboard cumbersome to use.

Please choose **only one** of the following:

- ◯ Strongly Disagree
- ◯ Disagree
- ◯ Neither Agree nor Disagree
- ◯ Agree
- ◯ Strongly Agree

9) I felt very confident using the dashboard.

Please choose **only one** of the following:

- ◯ Strongly Disagree
- ◯ Disagree
- ◯ Neither Agree nor Disagree
- ◯ Agree
- ◯ Strongly Agree

10) I need to learn a lot of things before I could get going with this dashboard.

Please choose **only one** of the following:

- Strongly Disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

Part 3: Hypothesis Verification

Please use the link below and the provided username and password to access the student-facing learning dashboard designed for this research. You are supposed to access the dashboard using the following information before answering the survey questions.

- Dashboard link: https://sf-ildresearchmvp.github.io/SF_iLD/
- Username: student000000
- Password: password

1) I believe that the student-facing learning dashboard will improve students' self-reflection.

   Self-reflection can be defined as the activity of thinking about your own feelings and behaviour, and the reasons that may lie behind them.

Please choose **only one** of the following:

- ◯ Strongly Disagree
- ◯ Disagree
- ◯ Neither Agree nor Disagree
- ◯ Agree
- ◯ Strongly Agree

2) I believe that the information on dashboard will help students to improve their course engagement and performance.

Please choose **only one** of the following:

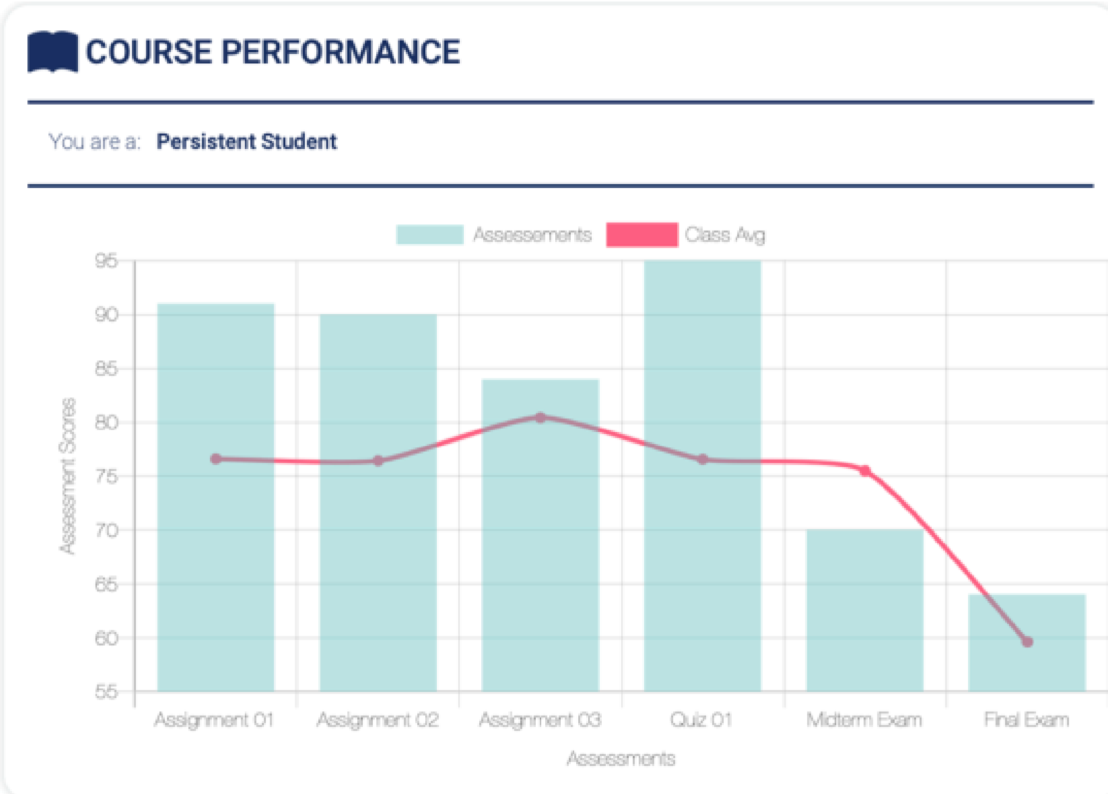
- ◯ Strongly Disagree
- ◯ Disagree
- ◯ Neither Agree nor Disagree
- ◯ Agree
- ◯ Strongly Agree

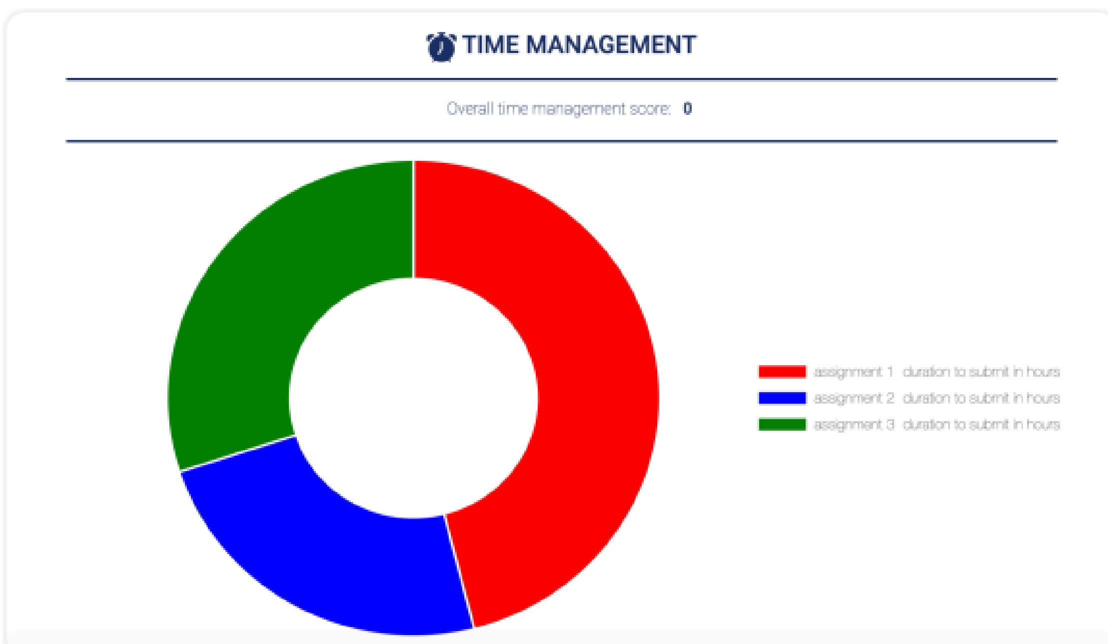
Part 4: Comments for Dashboard Improvement

1) Please provide your comments on how the student facing learning dashboard can be improved.

Please write your answer here:

Appendix D: Prototype SF-LAD Designed for This Study


COURSE PERFORMANCE
You are a: Persistent Student
Assessements
Class Avg
95
90
85
80
75
70
65
60
55
Assessment Scores
Assignment 01
Assignment 02
Assignment 03
Quiz 01
Midterm Exam
Final Exam
Assessments


TIME MANAGEMENT
Overall time management score: 0
assignment 1 duration to submit in hours
assignment 2 duration to submit in hours
assignment 3 duration to submit in hours

Appendix E: Certification of Ethics Approval

Athabasca University
RESEARCH CENTRE

CERTIFICATION OF ETHICAL APPROVAL

The Athabasca University Research Ethics Board (REB) has reviewed and approved the research project noted below. The REB is constituted and operates in accordance with the current version of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2) and Athabasca University Policy and Procedures.

Ethics File No.: **25084**

Principal Investigator:
Mr. Arta Farahmand, Graduate Student
Faculty of Science & Technology\Master of Science in Information Systems (MScIS)

Supervisor/Project Team:
Dr. Fuhua (Oscar) Lin (Co-Supervisor)
Dr. Ali Dewan (Co-Supervisor)

Project Title:
Student Facing Intelligent Learning Dashboard for Online Learners

Effective Date: January 18, 2023 **Expiry Date**: January 17, 2024

Restrictions:

Any modification/amendment to the approved research must be submitted to the AUREB for approval prior to proceeding.

Any adverse event or incidental findings must be reported to the AUREB as soon as possible, for review.

Ethical approval is valid *for a period of one year*. An annual request for renewal must be submitted and approved by the above expiry date if a project is ongoing beyond one year.

An Ethics Final Report must be submitted when the research is complete *(i.e. all participant contact and data collection is concluded, no follow-up with participants is anticipated and findings have been made available/provided to participants (if applicable))* or the research is terminated.

Approved by: **Date: January 18, 2023**

Dunwei Wen, Chair
School of Computing & Information Systems, Departmental Ethics Review Committee

Athabasca University Research Ethics Board
University Research Services Office
1 University Drive, Athabasca AB Canada T9S 3A3

Certification of Ethics Approval (RENEWAL)

CERTIFICATION OF ETHICAL APPROVAL - RENEWAL

The Athabasca University Research Ethics Board (REB) has reviewed and approved the research project noted below. The REB is constituted and operates in accordance with the current version of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2) and Athabasca University Policy and Procedures.

Ethics File No.: **25084**

Principal Investigator:
Mr. Arta Farahmand, Graduate Student
Faculty of Science & Technology\Master of Science in Information Systems (MScIS)

Supervisor/Project Team:
Dr. Fuhua (Oscar) Lin (Co-Supervisor)
Dr. Ali Dewan (Co-Supervisor)

Project Title:
Student Facing Intelligent Learning Dashboard for Online Learners

Effective Date: January 18, 2024 **Expiry Date**: January 17, 2025

Restrictions:

Any modification/amendment to the approved research must be submitted to the AUREB for approval prior to proceeding.

Any adverse event or incidental findings must be reported to the AUREB as soon as possible, for review.

Ethical approval is valid *for a period of one year*. An annual request for renewal must be submitted and approved by the above expiry date if a project is ongoing beyond one year.

An Ethics Final Report must be submitted when the research is complete *(i.e. all participant contact and data collection is concluded, no follow-up with participants is anticipated and findings have been made available/provided to participants (if applicable))* or the research is terminated.

Approved by: **Date: January 18, 2024**

Paul Jerry, Chair
Athabasca University Research Ethics Board

Athabasca University Research Ethics Board
University Research Services Office
1 University Drive, Athabasca AB Canada T9S 3A3
E-mail rebsec@athabascau.ca
Telephone: 780.213.2033