Student-centric Model of Learning Management System Activity and Academic Performance: from Correlation to Causation

Varun Mandalapu
University of Maryland Baltimore County
varunm1@umbc.edu

Lujie Karen Chen
University of Maryland Baltimore County
lujiec@umbc.edu

Sushruta Shetty
University of Maryland Baltimore County
sshetty2@umbc.edu

Zhiyuan Chen
University of Maryland Baltimore County
zhchen@umbc.edu

Jiaqi Gong
Univerisy of Alabama
jiaqi.gong@ua.edu

In recent years, there is a lot of interest in modeling students' digital traces in Learning Management System (LMS) to understand students' learning behavior patterns including aspects of meta-cognition and self-regulation, with the ultimate goal to turn those insights into actionable information to support students to improve their learning outcomes. In achieving this goal, however, there are two main issues that need to be addressed given the existing literature. Firstly, most of the current work is course-centered (i.e. models are built from data for a specific course) rather than student-centered (i.e. models are built taking the perspective of students by analyzing data across courses); secondly, a vast majority of the models are correlational rather than causal. Those issues make it challenging to identify the most promising actionable factors for intervention at the student level where most of the campus-wide academic support is designed for. In this paper, we explored a student-centric analytical framework for LMS activity data that can provide not only correlational but causal insights mined from observational data. We demonstrated this approach using a dataset of 1651 computing major students at a public university in the US during one semester in the Fall of 2019. This dataset includes students' fine-grained LMS interaction logs and administrative data, e.g. demographics and academic performance. In addition, we expand the repository of LMS behavior indicators to include those that can characterize the time-of-the-day of login (e.g. chronotype). Our analysis showed that student login volume, compared with other login behavior indicators, is both strongly correlated and causally linked to student academic performance, especially among students with low academic performance. We envision that those insights will provide convincing evidence for college student support groups to launch student-centered and targeted interventions that are effective and scalable.

Keywords: Learning Management System, Student behavior modelling, Causal Analysis

1. INTRODUCTION

Understanding student learning behaviors is a complex process as it is affected by a multitude of factors like self-regulation, instructional design, and social environment (Ikink, 2019). However, it is not efficient to study all the influencing factors as some factors are not feasible to intervene compared to others. Some have a higher impact on the performance at an individual level than others. Earlier research in education identified self-regulation as a highly impactful and intervenable factor (Zimmerman and Kitsantas, 2014; Schapiro and Livingston, 2000). Self-regulation, in general terms, is defined as one's ability to control their own behaviors, thoughts, and emotions to achieve their goals to drive their learning experiences successfully.

Studying student self-regulation capabilities is an arduous task as it involves multidimensional and complex factors like planning, goal setting, time management, and self-monitoring. Many of these factors are hard to quantify without student self-reports or other psychometric evaluations, and these reports are prone to biases like social desirability and reference bias (Toering et al., 2012; Rosenman et al., 2011). To mitigate these biases, researchers in education focused on extracting data from computer-based systems like Learning Management Systems (LMS) as they are used to deliver course content and have the capability to capture student interaction and assessment behaviors non-invasively (Tempelaar et al., 2015; Salehian Kia et al., 2021). Student interaction data with LMS proved to be a valuable resource in identifying learning strategies and track patterns among students that strongly support their academic performance. The data logs collected by LMS systems are analyzed with scientific techniques published in the Educational Data Mining (EDM) domain. In their study, Romero and Ventura (Romero and Ventura, 2007) described that EDM methods rely on clustering and pattern recognition techniques to categorize students into various groups based on their interaction patterns. Categorization of students using clustering and pattern recognition supports instructors in making changes for a set of students. Teaching practices that impact the entire classroom can be evaluated using predictive analytics that tracks student learning and achievement from the vast amount of interaction data collected by LMS. Majority of these studies analyze student interaction data in a specific course to understand student learning behaviors and teaching methods.

While course level predictions are suitable for supporting instructor level decision making; however, if intervention is on student level behaviors such as study habits or self-regulation skills, it is beneficial to look at student-centered indicators so that interventions may be more targeted and cost-effective. Developing student-centric models that analyze student LMS interactions across courses in a college/university setting will help address the issues with course specific models. This study is the first step in developing feature extraction and modeling methods from LMS login data that are scalable across semesters and transferable across different undergraduate fields. Developing these early performance prediction models alone does not help improve student learning as identifying self-regulation components like time management and other student engagement factors play a crucial role in recommending methods to improve their current behaviors.

LMS systems also enabled researchers to study time management strategies, a compo-

nent of self-regulation adopted by students (Jo et al., 2015; Ahmad Uzir et al., 2020). In contradiction to traditional time management studies that utilize self-reported questionnaires, student time management strategies are analyzed based on login and submission patterns in LMS systems. However, only few studies focused on the relationship between time management strategies identified in LMS systems with the biological nature of human functionality. Therefore, these earlier modeling efforts are limited to predicting student performance as a function of their login patterns and fell short of identifying patterns in time management that support improving student academic improvement. These earlier studies also focused on student behaviors related to their time compliance with assignments, including procrastination behaviors and ways to support the earliness of student work (Ilves et al., 2018; Edwards et al., 2009; Martin et al., 2015). These studies reported that early starters have better learning outcomes, but there is little emphasis on what times these students work during a day and how these learning time patterns affect their performance. Inspired by these earlier studies and comparatively similar studies in biology about human activity/productivity during different times of the day, referred to as chronotype/chronological analysis, this study analyzes the relationship between student activity on LMS at different times in a day and their performance by employing clustering methods. In addition to this, the student chronotype clusters are added to LMS behaviors and demographics as features to predictive models to predict student performance.

Earlier research in the area of Learning Analytics and Educational Data Mining focused on the relationship between self-regulation and student performance through LMS systems used student interaction features as proxy variables for self-regulation (Dabbagh and Kitsantas, 2013; Landrum, 2020). These studies showed a significant correlation between student login patterns and their academic performance. However, it is still challenging to design interventions as correlations do not necessarily mean a meaningful cause and effect between student login behaviors and academic performance. In this work, we introduce a novel framework to perform causal inference and discovery of multimodal student data collected from LMS and student administrative system. In addition, we enrich the feature set by introducing new approach to extract login behaviors at a student level. Finally, this work also introduced concepts from human chronobiology to analyze student login behavior on an hourly basis (Chronotype analysis) to identify activity patterns at different times in a day and their relationship to academic performance. The proposed correlation/predictive modeling framework (Mandalapu et al., 2021), chronotype analysis and causal inference framework goes beyond traditional educational data mining approaches to answer the following research questions that are under-studied in current literature.

2. RESEARCH QUESTIONS & OUTLINE

RQ1. How student learning activity at different times in a day (chronotypes) relate to their academic performance?

RQ2. How student-centric characterization of LMS activity correlates with academic performance?

RQ3. What are the causal relationship between student-centric characterization of LMS activity and academic performance?

In the related work Section 3, we discuss earlier research works that use LMS system data to predict student academic achievement. Section 3.1, we details research work that focuses on the evolution of chronotype on the basis of human biology and in the following Section 3.2 explains how chronotypes are linked to individual cognitive abilities and academic achievement.The final related work sections 3.3 and 3.4 explains earlier research work in correlation and causation to identify relationship between student LMS interaction data and their academic performance.

Section 4 provides the details of data set used to perform analysis in this study. The methodology section 5 details the steps taken to extract student aggregate login features in 5.1. The correlation studies that focuses on machine learning modeling, and post hoc LIME based explanations are detailed in Section 5.2.1 and Section 5.2.2. The section 5.3 details the causal analysis and inference tools and methods adopted in this study. The results of the study related to student chronotype analysis are listed in section 6.1. Section 6.2.1 explains the results of predictive models by comparing their predictions based on multiple evaluation metrics, this section also details various correlation based explanation methods that details the relationship between student demographics, LMS interaction features and academic performance. The causal discovery and inference results are detailed in section 6.3. Finally in Section 7, we detail the key findings and contributions made by this study to the existing literature in subsections **??** and **??**.

3. RELATED WORK AND BACKGROUND

Universities and colleges around the world adopted LMS systems, such as Moodle and Blackboard, to provide onsite, hybrid, and online courses based on their capabilities to support communication, content creation, administration, and assessment(Alokluk et al., 2018; Berechet and Georgescu,) LMS systems build and swiftly distribute individualized learning materials and information in addition to automating and centralizing a variety of administrative processes including setting up and managing student accounts, establishing a syllabus, assignments, assessments, and grading, etc.(Shchedrina et al., 2021). These methods also encourage the reuse of instructional materials. Instructors can design content structures, distribute it in a sequential order, limit access, organize group activities, monitor student activity, load and replace learning materials, and give feedback on assessments using the systems. LMS systems use a variety of login roles based on user classification thanks to cutting-edge database software created by Oracle, IBM, and Microsoft that emphasizes interconnection, data independence, and security. These roles will permit instructors to create new content or privately address student issues and create discussion boards to capture student knowledge on specific topics.

The relationship between the use of LMS and student academic accomplishments has been examined in numerous papers in the fields of Educational Data Mining and Learning Analytics. According to Vengroff and Bourbeau's (Vengroff and Bourbeau, 2006) study, undergraduate students benefited from having more content available in LMS.

4

They also conclude that students who used LMS regularly did better in exams than their peers who have minimal interactions. In their study, Dutt and Ismail ([Dutt and Ismail, 2019](#)) found that keeping track of the LMS resources that students use encourages the creation of fresh approaches to learning that improve student achievement. Additionally, they examined the thresholds for student involvement elements like self-assessment quizzes, exercise time, discussion boards, and performance results. Another study by Lust et al. ([Lust et al., 2011](#)) investigated how differently students used various LMS capabilities, such as the amount of time spent on web-links, web lectures, quizzes, feedback, discussion board entries, and messages read. The findings of this research made a significant contribution to the creation of adaptive and creative recommendation systems. In their research, Hung and Zhang ([Hung and Zhang, 2012](#)) also discovered trends based on six indices that represented student effort: the frequency with which students access course materials, the number of LMS logins, the sum of interactions in discussion threads, the number of synchronous discussions, the number of posts read, and the course's final grades.

When investigating the relationship between students' online behavior on the LMS and their grades, Dawson et al. ([Dawson et al., 2010](#)) found a substantial difference between high and low performing students in the number of online sessions visited, overall time spent, and the number of posts in discussion boards. A multinomial logistic regression model was created by Damainov et al. ([Damianov et al., 2009](#)) based on the amount of time spent in the LMS. This study discovered a substantial correlation between the amount of time students spent studying and their grades, particularly for those who had grades between D and B. Other studies emphasized how frequently students accessed course materials in the LMS, as opposed to measuring time spent online. A study by Baugher et al. ([Baugher et al., 2003](#)) found that regularity in student hits is a reliable predictor of student performance compared to the total number of hits. In their study, Chancery and Haque studied the student interaction logs of 112 undergraduate students and discovered that those with lower LMS access rates had lower grades than their counterparts with greater access rates. Biktimirovan and Klassen ([Biktimirov and Klassen, 2008](#)), who observed a significant correlation between student hit consistency and success, provided additional support for this study. In their study, which measured access to numerous LMS features, it was discovered that the only significant predictor of student achievement is access to assignment solutions. However, these studies tend to be more descriptive than predictive.

3.1. CIRCADIAN RHYTHMS & CHRONOTYPES

Humans have a 24-hour internal clock running in the background that determines when to sleep and when to be productive. These 24-hour cycles are referred to as circadian rhythms. The behavioral manifestation of these circadian rhythms is termed chronotype ([Reppert and Weaver, 2002](#); [Dibner et al., 2010](#)). Understanding an individual's chronotype helps identify their routine and provides insights about their highly active and productive time. Earlier research argues that circadian rhythms differ between individuals as a group of clock genes conditions them ([Bell-Pedersen et al., 2005](#)). Nonetheless, these are not fixed and can vary during an individual's lifetime.

Students have to schedule their daily activities based on their predetermined schedules by taking social constraints like class schedules, outside work hours into consideration. General clock time preferences are more likely to match natural chronotypes when students are given more flexibility in organizing their schedules. Differences between a person's natural chronotypes and schedules influenced by social constraints cause a phenomenon referred to as social jetlag (Roenneberg et al., 2003; Wittmann et al., 2006). This social jetlag leads to an accumulation of sleep debt that causes tiredness and a decline in cognitive abilities throughout the workweek. Prior research suggested that individuals with a match in chronotypes and schedules during workweek don't show a change on weekends (Korczak et al., 2008; Vitale et al., 2015). This is in contrast with individuals whose weekday routine varies from natural routines. As activity and sleep are both indicators of natural chronotypes, an observation of individuals on weekends might give accurate insights into their natural habits defined by circadian rhythms.

Earlier studies primarily focused on two significant chronotypes in humans: Morning and Evening. These studies showed that these two chronotypes are based on an individual's age and gender and vary across an individual lifetime. The insights from earlier work showed that children are predisposed towards morningness, but a delay of phase preference can be observed when they reach adolescence (Carskadon et al., 1998; Crowley et al., 2007). This shift reaches a maximum of eveningness at the age of 20. Multiple studies also observed that most individuals return to morningness again at the age of 50 (Díaz-Morales and Sorroche, 2008; Baehr et al., 2000). Gender-based chronotype studies contradicted a lot in identifying variations in morning and eveningness as some studies showed women to have a greater tendency towards morningness compared to men. A meta-analysis showed that the significance between morningness and females is very weak (Randler, 2007). In addition to these two demographics, other researches focus on different variables like productivity, mood, temperament, caffeine consumption, avocational interest, and internal temperatures.

Recent studies observed a new type of chronotype referred to as intermediate (Valladares et al., 2018; Porcheret et al., 2018). A study performed by Putilov et al. in 2019 (Putilov et al., 2019) argued about four chronotypes based on time differences in alertness and sleepiness: morning, afternoon, napper, and evening. Most of these studies adopt different self-reported questionnaires that are developed to analyze an individual's daily preferences. Initially, most of these measures treated chronotypes as unidimensional: Diurnal Type Scale (DTS) and Circadian Composite scale (CCS) (Torsvall and Åkerstedt, 1980; Smith et al., 1989). But, psychometric studies questioned this unidimensional approach to morning-eveningness (Putilov and Putilov, 2005; Putilov and Onischenko, 2005). Recent studies came up with a multidimensional approach that treated chronotypes as independent dimensions. As most of the chronotype measurement methods are self-reported surveys, biases like social desirability, recall period, sampling approach, and selective recall are hard to eliminate. In order to mitigate this issue, our work explores the activity patterns of students based on login times captured by the LMS. We employ clustering techniques to identify and classify students based on their similar activity patterns.

Research in psychology documented the effects of time of day on complex simple and complex cognitive functions that are based on an individual's chronotype (Schmidt et al., 2007; Roberts and Kyllonen, 1999). An early study by Roberts and Kyllonen (Kyllonen and Christal, 1990) showed that individuals who were active during the evenings did well on the measures of memory, cognitive ability, and processing speeds even though these cognitive tasks are performed during the morning time. In addition to this, this study also reported a high correlation between individuals working memory which is a proxy of general intelligence, and morningness and eveningness (Killgore and Killgore, 2007). However, other researchers reported that the relationship between cognitive ability and chronotype is much more variegated. For example, a study by Killgore and Killgore(Killgore and Killgore, 2007) reported a significant correlation between eveningness and verbal cognitive ability but not between eveningness and math ability. Further analysis showed that the latter finding is only observed in female participants. A similar study by Song and Stough(Song and Stough, 2000) showed a significant eveningness effect on a spatial subtest of Multidimensional Aptitude Battery IQ and not on any other subtests. Most of the research in this domain is still inconsistent on which aspects of cognitive abilities do chronotypes impact.

Inspired by earlier studies that reported significant relations between cognitive ability and chronotypes, researchers in education focused on studying the relationship between academic performance and chronotypes, mainly student grade point averages and other measures extracted from their academic achievement indicators (Enright and Refinetti, 2017; Preckel et al., 2011a). Most of the studies in this area reported that student academic achievement is strongly and inversely related to their eveningness, whereas morningness is positively associated with their achievement. These patterns are similar in students at the high school level and students at the university level. A study by Precckel and Roberts (Preckel and Roberts, 2009) showed a significant negative correlation between academic achievement and eveningness. This study was conducted on 270 German secondary students where teachers assigned grades averaged over German, Math, English, and Physics. These results seem to be consistent even after introducing controls for gender and intelligence.

A recent study focused on university students that utilized three classification scales (Morning, Evening, and Neither) found that morningness students showed much better performance on both theoretical and practical examples compared to students belonging to other chronotypes (Montaruli et al., 2019). This study also showed that students belonging to neither evening nor morning chronotypes performed lower than their counterparts. Additionally, this study also observed that students who belong to eveningness did worse on practical examinations than morningness and students belonging to neither class. One interesting find in this study is related to the performance of eveningness students on theoretical exams. This study infers that the higher intelligence expressed by eveningness students compensated for their disadvantage on the theoretical exam but not on practical exams. However, it is still necessary to study the impact of an age-based shift in chronotypes with student's academic performance, as

earlier studies reported that individuals move from morningness to eveningness by the time they reach their 20's. In addition to this, it is also necessary to study the impact of chronotypes on student-centric models that analyze comprehensive student data in a semester or academic year instead of course-specific analysis. This understanding will support the development of interventions that effective and promise overall student success. Our work will utilize student clusters extracted from their hourly login patterns to develop and assess models that predict student performances.

3.3. CORRELATION STUDY OF THE RELATIONSHIP BETWEEN LMS ACTIVITY AND ACADEMIC PERFORMANCE

Online teaching tactics are mostly dependent on instruction design since each type of interaction—student/instructor, student/student, and student/content—has a favorable effect on learning. A study by Coldwell et al. (Coldwell et al., 2008) focused on the relationship between student participation in a fully online course and their final grades. They discovered a positive correlation between student participation and final grade. When Dawson et al. (Dawson et al., 2010) looked at the effects of several LMS technologies, they discovered a positive correlation between participation in discussion forums and student achievement. They found that the discussion forum, which is the main instrument for interaction in LMS, was where more than 80% of interactions took place. Asynchronous communication was discovered to be the main method of engagement in all online courses, according to a different study by Greenland (Greenland, 2011). Near the deadlines for assignments and exams, Nandi et al. (Nandi et al., 2011) discovered an increase in the volume of posts in discussion forums. Additionally, they discovered a strong relationship between exam performance and ongoing online class engagement, particularly among top achievers.

All of the studies mentioned used log files from LMS systems to gather objective information about activity and performance in order to establish a connection between independent interaction variables and student grades. The majority of the studies that have been addressed are based on univariate analysis and concentrate on one variable, or a group of variables that have a significant impact on the outcomes of students in one course or courses like it. To assess or comprehend student performance, especially across the range of on-campus courses available in a university context, is a very difficult task. Most of the authors discussed above noted the need for more in-depth works to investigate student performance across courses and based on multiple variables. These studies also lack an explanation of the variables they used to measure student performance, and it is clear that the authors chose the LMS variables based on their belief that these variables are highly correlated with student scores.

Even though there is a common agreement about the purpose of learning analytics, there are still several varying opinions on what data needs to be collected and analyzed to improve teaching and learning processes. It is extremely difficult to determine the net contribution of numerous interactions to the learning processes, according to a study by Agudo-Peregrina et al. (Agudo-Peregrina et al., 2014). Contrary to past studies that indicated a strong importance for student peer connections, their findings indicate that peer contact between students has less influence than peer interaction

between students and teachers. In order to forecast student grades at the end of the course, Dominquez et al. (Dominguez et al., 2016) used a variety of factors, including LMS logins, time stamps, and content access flags recorded in a biology course. According to the findings, the algorithm's forecast accuracy for succeeding semesters is 50%. In their recent study, Lerche and Keil (Lerche and Kiel, 2018) used Moodle log data from 369 students enrolled in three online courses over the course of three semesters to forecast their final grades for each course. For all three courses, their regression findings ranged from 0.17 to 0.6 in terms of predicting student scores in a course at the end of the semester. This wide variation in performance between courses is caused by the different variables used in each course according to its structure. The inconsistent results from past studies may be explained by examining the different instructional designs, factors in extracted data, statistical inferences, predictive modeling employed, interpreting model outcomes and pattern observations, etc.

Since LMS systems record student interactions in non-intrusive, ready-to-use settings, their data has gained prominence in LA and EDM circles. Several studies were discussed earlier in this research that utilized the LMS data to develop models that track student progress. However, it is currently difficult to develop extremely precise models that forecast student learning outcomes across courses and understand the impact of various variables gathered by LMS. Their failure to predict student success across courses in a particular semester is another important flaw in past research. One primary issue in predicting student performance across a semester is to find methods that aggregate student LMS variables across courses. This study demonstrates approaches for filling the knowledge gap identified by past investigations.

In this study, we approach the problem of tracking student achievement by developing student-centric models that build on aggregated LMS interaction variables collected across a semester irrespective of student year and course. One distinctive feature of our work is related to the study of model performance on longitudinal student data. We develop models that predict student end-of-term GPA based on four cumulative periods in a semester. The purpose of this paper is to explain how numerous aggregated LMS variables affect various student groups that have been grouped according to performance, race, gender, and student type. The importance of features is explained by adopting correlation statistics for univariate importance, a regression model for interaction effect, and LIME for model-based yet model agnostic explanations.

3.4. CAUSAL INFERENCE FROM STUDENT LMS INTERACTION DATA IN PERFORMANCE PREDICTIONS

The primary objective of educational research is to develop interventions that promote student academic achievement. Earlier studies in learning analytics and educational data mining focused on developing models that predict student achievements early on in their academic program to support the development of these interventions. However, most of these studies focused on course-level predictions and analysis, making it harder to develop interventions at the student level. Another gap observed in prior research is the lack of causal understanding related to features that causes a shift in student academic performance. Causal inferences promote the development of impactful

intervention techniques that positively impact student academic performance.

Traditional research utilizes randomized control trials to discover causal knowledge. However, these are time-consuming, expensive, and sometimes unethical and impractical to implement with real-world students. Studying causal relationships from student interaction with LMS systems allows us to explore a vast amount of noninvasive data collected by these systems. This method will provide a search for answers beyond common correlations. Yet, this type of analysis is not common in learning analytics and educational data mining domains. For example, a study by Fanscali (Fancsali, 2015) investigated the role of carelessness in analyzing the counter intuitive relationship between the learning outcomes of students and their affective states related to confusion and boredom. This study utilized observational data from the algebra cognitive tutor platform. An earlier study by Fanscali (Fancsali, 2014) on a similar dataset showed a relationship between student's affective states, gaming the system, prior knowledge, and their learning outcomes. This study utilized a causal framework termed causal discovery with model. In their research, Koedinger et al. (Koedinger et al., 2016) analyzed interaction data captured by the online learning environment. This study reported causal interaction between active student engagement and their learning outcomes. Inspired by this work, a recent study by Chen et al. (Chen et al., 2021) developed a causal discovery framework that utilized TETRAD (Ramsey et al., 2018), a causal discovery and inference toolkit. Our work adopts the framework proposed by Chen et al. (Chen et al., 2021) to study causal relationships between student login behaviors captured by LMS and learning outcomes.

4. DATASET

For this study, we chose undergraduate student data captured by LMS in Fall 2019 from a large public university in the United States. These students were part of either Information Systems (IS) or Computer Science (CS) departments. The Blackboard system is predominantly used as an LMS to deliver course material, assessment, and grading. The student demographic data captured by a standalone Student Information System (SIS) is used to categorize students based on different demographic variables. A total of 1651 students were enrolled in these two departments in the Fall 2019 semester. Based on student distribution in the undergraduate majors selected, we categorized students into three ethnicities based on their counts: White, Asian, and Minority(mostly African Americans and Hispanic population). This study also researches student performance based on their admit types, such as four-year regular student or transfer student. The transfer student here refers to students that completed at least one year of their undergraduate course work at a different college (mostly community college) before enrolling at current institution. The demographics of student data are provided in Table 1. The university's Institutional Review Board (IRB) approved this study, and all the student specific demographic and personal information are anonymized by following General Data Protection Regulation (GDPR) standards.

Table 1: Student demographics.

Demographic	Student Count
Total Students (N)	1651
No of unique courses	440
No of unique course instructor combinations	638
Male : Female	1302 (79%) : 369 (21%)
White : Asian : Minority	630 (38%) : 495 (30%) : 526 (32%)
4 – Year : Transfer	976 (59%) : 675 (41%)
Full Time : Part Time	1446 (88%) : 205 (12%)
IS : CS	934 (57%) : 717 (43%)
1st Yr : 2nd Yr : 3rd Yr : 4th Yr	115 (7%) : 329 (20%) : 515 (31%) : 692 (42%)
≤ 3 : 4-5 : $\gg 5$ (Courses Enrolled)	298 (18%) : 1035 (63%) : 318 (19%)

5. METHODOLOGY

The methodology section of this study is divided into three subsections. In Section 5.1, the focus is on the feature extraction methods that describes the methodology used to extract aggregated login related features from LMS data. Section 5.2 focuses on Correlation studies which comprises of machine learning approaches and correlation based model explanation methods deployed in the study. In section 5.3, the methodology to study the causal relationships between student login behaviors, chronotypes, and academic performance outcomes are detailed.

5.1. FEATURE EXTRACTION

From various LMS related features like student logins, content accesses, time spent, discussion posts, assignment submissions, and time intervals discussed in earlier literature, we identified that only three features could be commonly extracted from different courses to generate aggregate level student data: Student Login Counts, Time intervals and Prior performance.

One of the significant challenges while building a student-centric model on LMS data is to extract aggregated features that are least biased. As Blackboard's content is dependent on instructor and course, it is crucial to mitigate the variations caused by these
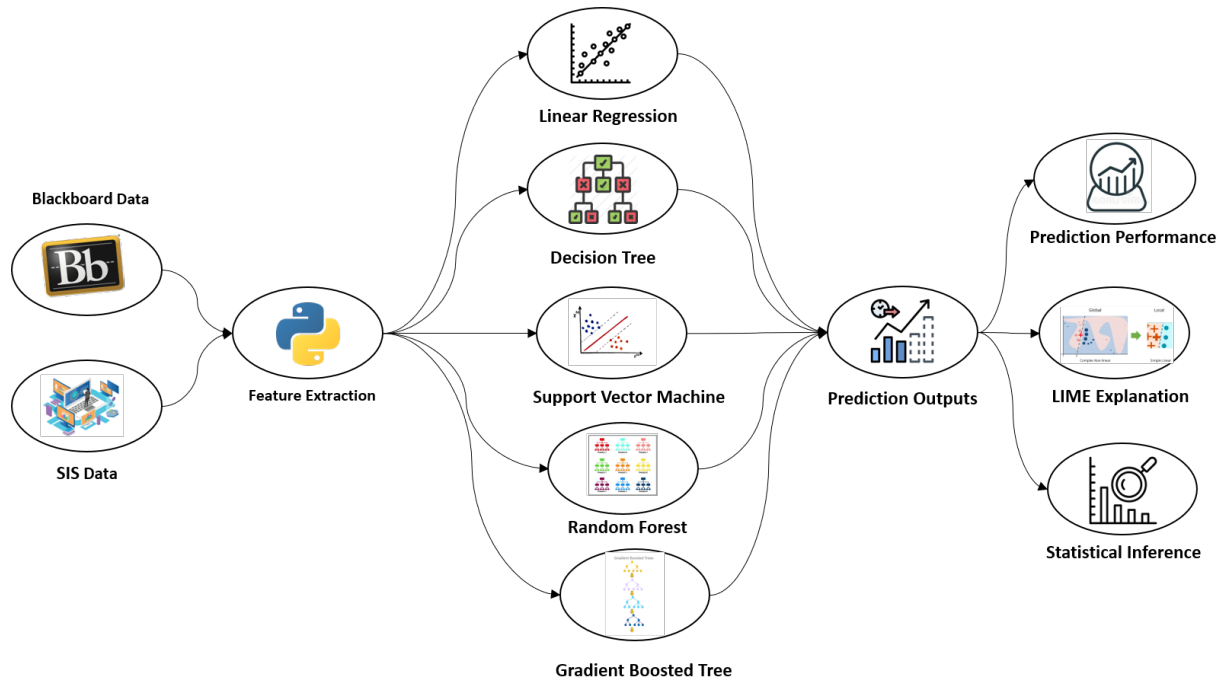
Figure 1: Student-centric Model Workflow

factors on aggregate student variables. This work employs multiple statistical measures to mitigate these issues. The details are explained in the below Section 5.1.1.

5.1.1. Normalized Login Volume

Earlier studies showed that student performance is strongly correlated with the volume of student LMS logins. Also, calculating the total login count introduces a hidden bias as courses with more content on Blackboard prompt students to login more often than other courses with less content and flexible deadlines. To mitigate this issue, our work followed the below steps to extract student login features.

1. Extract all courses enrolled by the students in our study sample which includes students majored in both IS and CS;

2. Count the total number of logins for all students enrolled in those courses, regardless if students are part of our study sample;

3. Calculate normalized login volume for each student in our study sample in each course they enrolled. For a given student S_i and a given course C_j the student is enrolled in, the normalized login volume N_{ij} is a Z-score calculated as $\frac{L_{ij}-\mu_j}{SD_j}$ where L_{ij} is the absolute login counts for student S_i in course C_j, and μ_j and SD_j is the mean and standard deviation aggregated from all absolute login counts for all students enrolled in course C_j.

12

This step is crucial as it allows us to mitigate the bias from variations in courses design which is reflected in the between-course variations of overall login volume. Z-scores provide a value that helps understand if student logins are higher or less than average logins in a specific course.

4. Once the z-scores are calculated for all courses, we extract a vector of login z-scores for each student based on their enrolled courses.

5. As most predictive models do not take vectors of variable length as input, we then extracts seven significant statistics from the login vector: mean, median, minimum, maximum, standard deviation, skewness, and kurtosis.

5.1.2. Login Regularity

Apart from student login volumes, the regularity of logins also provides valuable insights into student achievement as regularity is related to students' habitual study patterns which reflects their self-regulation capabilities. In this work, we utilize an *entropy-based method* to extract features that define student login regularity in each course. In information theory, entropy is used to define uncertainty or randomness (Gray, 2011). Entropy measure will explain if student's logins are regular (less random) or irregular (more random). Based on this concept, if the entropy value is high, then a student has an irregular login pattern, and if the entropy value is low, the student has a regular login pattern. The steps to calculate student regularity features are given below.

1. Extract all course accesses with timestamps for every student in our study sample.

2. Calculate the interval between login timestamps which will result to a vector of time intervals for each course enrolled by a student.

3. Calculate login interval entropy H_{ij} given a student S_i and course C_j. Entropy is calculated using the KL estimator with the k nearest neighbor method proposed by Kozachenko and Leonenko (**?**). KL estimator uses k-nearest neighbor distances to compute the entropy of distributions. The reason for adopting this method instead of the classicial Shannon entropy is due to the continuous nature of time intervals (Kraskov et al., 2004).

4. Once the entropies are calculated, we get a vector of entropies for each student based on the number of enrolled courses. We then calculate the seven statistics similar to student logins: mean, median, minimum, maximum, standard deviation, skewness, and kurtosis.

5.1.3. Login Chronotypes

Studies in chronobiology and chronopsychology showed variation in different individual active periods at different times of the day (Putilov et al., 2021; Romo-Nava et al., 2020). These studies classify an individual into either morning type or evening type based on their timing of high activity period. For example, if an individual is highly

active in the morning compared to the evening, they are considered morning type and vice versa. Inspired by this work in human psychology (Putilov et al., 2021), this work divides a day into 24 hourly time bands and in addition aggregates 24 hours in a day into four-time bands T1 (12 AM to 6 AM), T2 (6 AM to 12 PM), T3 (12 PM to 6 PM), and T4 (6 PM to 12 AM) and extract student logins. In addition to this, this work also extracts the logins on weekdays and weekends to study their influence on student performance.

1. Count the number of logins during each hour in a day, count number of logins based on 4 time bands T1-T4 and on weekdays and weekends for each course through out the semester.

2. Calculate the mean of login count vector across all days in a given semester for each of these time bands and weekday/weekend.

3. Normalize the login count with the number of courses enrolled by an individual student. This normalization will mitigate the bias introduced by the number of courses enrolled across the student cohort.

5.1.4. Student Demographics and GPA from prior semesters

This work also utilizes the demographic and prior performance measured by GPA features captured by the SIS system. These student-level features were listed in Table 2

Table 2: Student demographics and GPA at beginning of semester.

Features	Values
Start GPA (Prior Performance)	Cumulative GPA available till the start of semester
Gender	Male & Female
Ethnicity	White, Asian & Minority
Student Year	Freshman, Sophomore, Junior & Senior
Admit Type	Regular & Transfer
Enrollment Type	Full time & Part time
Student Age	Continuous variable

5.2. CORRELATION STUDIES

5.2.1. Machine Learning approach

This work studied five of the most common regression models for comparison purposes. The selected models include Regularized Linear Regression (LR), Decision Tree (DT), Support Vector Regressor (SVR), Random Forest (RF), and Gradient Boosted Regressor (GBR). As model hyperparameter influences their predictive performance, we utilized a grid search mechanism to select multiple parameters to predict with high accuracy. We also adopted a feature selection method based on a multi-objective evolutionary algorithm in addition to hyperparameter search. This feature selection algorithm evaluates each feature set based on pareto-optimal that balances model complexity and accuracy. The details of models and hyperparameter search criteria are discussed below.

Regularized Linear Regression: This model is extension from conventional linear regression. The regularization parameter is set so that the hyperparameter search space looks for an alpha value that fits between ridge and lasso regression. An alpha value of 1 represents lasso regression, and an alpha value of 0 represents ridge regression. This study searched for the best alpha value using a grid search between 0 and 1 in increments of 0.1.

Decision Tree: The decision tree algorithm is a collection of linked nodes intended to estimate the numerical target variable. Each node in the tree represents a rule used to split on an attribute value. The node uses a least-squares criterion to minimize the squared distance between the average value in a node when compared to the actual value. The hyperparameter search space for this algorithm evaluates both maximal depth and pruning. The maximal depth value varies between 1 and 100 in increments of 10. Pruning will make the DT algorithm use multiple criteria like minimal gain, minimal leaf size, and pruning alternatives to decide the stopping criterion.

Support Vector Machines: The SVM used in this study is built based on Stefan Reupping's mySVM (Ruping, 2000). This algorithm will construct a set of hyperplanes in a high dimensional space for regression tasks. A good hyperplane is decided based on the functional margin. The hyperparameter search space focused on both dot and radial kernel functions with a C (SVM complexity) value range between 10 and 200. The kernel gamma function is set for a radial kernel with a range of 0.005 and 5 with three logarithmic increments.

Random Forest: A RF model builds an ensemble of decision trees on bootstrapped datasets. The splitting criteria are similar to a decision tree. The regression outcome is the average of the observed train data GPA present at that end node. We only tuned the number of trees hyperparameter to reduce the time complexity of the execution. The number of tree searches varied between 10 and 1000 trees in 10 linear steps

Gradient Boosted Tree: The GBT model builds multiple regression trees in a sequence by employing boosting method. By sequentially applying weak learners on incrementally changed data, the algorithm builds a series of decision trees that produce and an ensemble of weak regression models. As GBT is a nonlinear model, we search hyperpa-

rameters related to the number of trees, learning rate, and maximal depth. The number of tree values varies between 1 and 1000 in five quadratic increments, the learning rate varies between 0.001 and 0.01 in five logarithmic increments, and the maximal depth parameter varies between 3 and 15 in three logarithmic increments

5.2.2. Model Explanation using LIME

The concept of Locally Interpretable Model Explanations (LIME) was introduced to explain the predictions made by black-box models that deal with classification problems. LIME explains each prediction made by a complex model by training a surrogate model locally (Ribeiro et al., 2016). However, this earlier methodology is not scalable to deal with categorical variables, tabular data, and regression problems. In this work, we adopt the correlation-based LIME method available in RapidMiner (Mierswa and Klinkenberg, 2021) to explain machine learning models' predictions(Hofmann and Klinkenberg, 2016; Mandalapu and Gong, 2019; Mandalapu et al., 2021; Mandalapu, 2021). The steps invovled in LIME is as following:

1. Perturb data in the neighborhood of each sample in the dataset. The number of simulated samples can be defined by user. A higher number of simulated samples will provide higher accuracy of explanations but at the cost of more computing time.

2. Make predictions using the ML model for all the simulated samples around each original sample in the dataset.

3. Calculate the correlation between each feature in the dataset and the target variable.

4. The features that have a positive correlation are considered supporting features, and features with negative correlation with predicted outputs are referred to as contradicting features.

As LIME provides feature importance value for each feature at each sample, we aggregate the importance value for all samples to build global importance for each variable. The significant advantage of this method compared to traditional global importance methods is its flexibility. As model global importance's are calculated across all samples in the data, the LIME based feature importance's can be calculated for subsets of data. This flexibility provides users with a deeper understanding of each feature's role for different sets of populations present in a dataset.

In addition to applying the LIME methodology, this work also studies univariate and multivariate feature importance on student performances by applying correlation and linear regression methods. The student dataset used in this study is divided into multiple subsets containing different student groups based on various demographics. A correlation value is calculated between input features and student end-of-term GPA.

This value provides us with an intuition about the impact of various features on student performances related to different demographics. As correlation only provides independent variable importance on student performance, we also adopt a linear regression model to explore the variation of feature importance based on coefficient values. Applying a linear regression model will also consider the interaction effect between input features to fit the outcome variable.

5.3. CAUSAL INFERENCE

One of the primary objectives of this study is to understand any causal relationships between student login behaviors and their performance outcomes. This understanding will support the development of efficient and effective intervention strategies that may positively impact student academic performance, effectively turning data into insights and into action. To achieve this goal, we adopt the processing pipeline shown in Figure 2.

The process shown in Figure 2 takes three inputs: student learning behaviors as characterized by login variables as described in Section 5.1 , demographics, and chronotype clusters. For login variables, we specifically focus on normalized student login volume, regularity of student logins, and weekday/weekend logins described in Table 3. As these features are a group of independent statistics (mean, median, standard deviation, minimum, maximum, skewness, and kurtosis) grouped into sets based on their relevance, we input them to Sparse Multiple Canonical Correlation Analysis (CCA) algorithm. This algorithm will help learn sparse representation for each feature group by maximizing the correlation among these feature sets with multiple features (Hardoon and Shawe-Taylor, 2011). This CCA approach results in a composite variable for each feature group. These are represented by a linear combination of a small set of features and are used as inputs for causal structural discovery and inference.

Student demographics are used in two ways. In the first method, we feed these demographics directly to the causal discovery model to study their relationships with student performances. In the second method, we use this demographic information to group students based on their demographics and explore causal relationships for students belonging to a specific group. Finally, we use chronotype clusters to study the causal relationship between chronotypes and student performance.

5.3.1. Causal Inference and Discovery

This work's causal discovery and inference are performed using a software suite named TETRAD (Ramsey et al., 2018). TETRAD was developed 20 years ago as a drag and drop suite of procedures to explore causal relationships in the input dataset. This suite can take continuous, categorical, mixed, covariance, and correlations as input for causal discovery. This tool consists of multiple proven algorithms that are selected based on the type of data that is inputted for causal discovery. This work adopts two algorithms: Greedy Fast Causal Inference (GFCI) and PC algorithm (Ogarrio et al., 2016; Spirtes and Glymour, 1991). These algorithms were selected based on their ability to accept prior knowledge that is useful to control the directionality of causal relations.
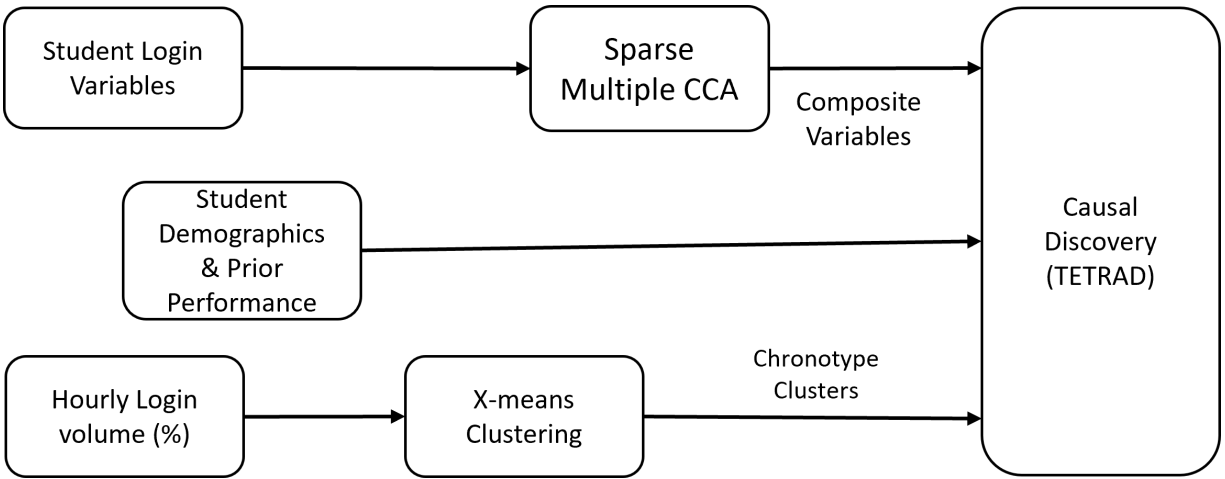
Figure 2: Analysis pipeline to study causal relationships between student learning behaviors, chronotypes and performance

The GFCI algorithm only works with continuous variables and can output Partial Ancestral Graphs (PAG). These PAG's are causal networks that include hidden confounder variables. The working of this algorithm is based on the combination of two algorithms named FGES and FCI (Spirtes et al., 1993). The FGES algorithm takes input sample data and background knowledge to score using a greedy search mechanism and applies it to a larger sample, and selects the Causal Bayesian Network (CBN) with a higher score. On the other hand, FCI is a constraint-based algorithm with similar functionality to FGES. In addition to this, it entails the set of conditional independence relations that will be satisfied at population-level data. FCI algorithm has two phases named as adjacency phase and orientation phase. The adjacency phase starts with an undirected graph and performs a sequence of conditional independence tests to filter edges between two adjacent variables that are independent. As the output of the adjacency phase is an undirected graph, the directionality between adjacent sets is provided by the orientation phase. In the orientation phase, the stored conditional settings used to remove the adjacencies to reduce edges are used to determine directionality. GFCI algorithm uses FGES to improve both adjacency and orientation phases of FCI by providing a more accurate initial graph for causal network development.

Though GFCI is a robust causal inference algorithm for continuous datasets, it is unable to handle categorical variables like student demographics. As such, we only used the GFCI algorithm to study the causal relationship between login variables and student performance.

We adopt a PC algorithm to explore causal relationships with both categorical and continuous variables to mitigate this issue. PC algorithms utilize a pattern search method that assumes the underlying structures in data are acyclic. This algorithm also assumes that no two variables have the same latent variable. The only drawback of the PC search algorithm is its inability to show confounding relations in its data, whereas GFCI algorithm mitigates this issue. This is also one of the primary reasons to adopt two algorithms in this study instead of one for causal structure discovery and inference.

5.3.2. Sparse Multiple CCA

The high dimensional dataset in this study is reorganized into multiple sets based on their roles in student LMS interaction process. The primary reason for reorganizing features variables into different groups is to make the causal inference part of this study more interpretable. The variables are groups with meaningful roles is shown in Table 3. But reorganizing variables into different groups alone is not sufficient for interpretability as causal algorithm needs a single variable to represent each feature concept related to LMS interactions. We generates a composite variable from the group of features for every feature subsets using Sparse Multiple Canonical Correlation Analysis (mCCA) (Hardoon and Shawe-Taylor, 2011). The traditional CCA method takes two matrices as inputs to generate a linear combination of variables in each feature set with a high correlation between the two feature sets. This is similar to the way Principal Component Analysis (PCA) works, but on multiple datasets at a time. In addition to this, CCA also maximizes the correlation between datasets while generating composite variables. Sparse CCA is an extension to traditional CCA, where sparsity constraints are imposed. This makes it more compact and yields a more interpretable representation of the data. This study adopts the mCCA function from the PMA package in R. The standardized student login variables were inputted to the mCCA algorithm. A grid search is performed to select the model with the best total correlation and compactness. This model is used to output weights for each feature and then generate composite variables for each feature set. These composite variables are then fed to GFCI or PC variant algorithm to perform causal structural discovery and inference.

6. RESULTS

In this section 6, we discuss the outcomes of experiments performed in this study. Section 6.1 reports the results of clustering student activity at different times in a day, referred to as chronotype analysis. In the section 6.2, we report the performance of five machine learning algorithms to predict student performance at the end of semester by learning from student login, interactions, and demographic data till the middle semester. This subsection also reports the importance of different features based on the correlation between predictors and target variable and also using LIME post hoc model explanation. The final section 6.3 reports the outcomes of causal inference and discovery experiments to study the cause and effect between different predictor variables and also between predictor variables and outcome variable.

6.1. CHRONOTYPES ANALYSIS

We cluster the student hourly login percentages using X-means with Dynamic Time Warping (DTW) (Müller, 2007) distance calculation to study different chronotypes. The results of the clustering algorithm are shown in Figure 3. As shown, the X-means algorithm found three clusters to be optimal for the input data. The x-axis in Figure 3 shows the hours in a day based on the 24-hour clock. It starts with H1 representing 12 AM to 1 AM and ends with H24 describing 11 PM to 12 AM with 1-hour incremental

Table 3: Feature set for causal analysis relationships.

Feature	Description
Demographics	Age, Gender, Student Year, Type of Admit, Full time/Part time, Student Ethnicity.
Enrolled Courses	The total number of courses enrolled by student in Fall 2019
Student Start GPA	The cumulative GPA of student till the start of Fall 2019 Semester
Normalized Login Volume	Statistics related to login z scores of students per each enrolled course: Mean, Median, Standard Deviation, Minimum Score, Maximum Score, Skewness, and Kurtosis.
Student Regularity in Logins	Statistics from KL Entropy of student login intervals per each enrolled course: Mean, Median, Standard Deviation, Minimum Score, Maximum Score, Skewness, and Kurtosis.
Weekday and Weekend volumes	Login counts on weekdays and weekends.
Hourly Login Volume Percentage	Count cumulative login volumes per hour and normalize them by calculating the percentage of logins per hour.
Hourly Time spent	Count the aggregate time spent per 60 minutes every hour till the middle of the semester and then calculate the percentages by normalizing across 24-hour bands in a day.

time bands. The y-axis in Figure 3 represents cluster centroids for each corresponding time band and cluster.

From cluster 0, we can observe that students' login percentages slowly start to increase from the morning at 8 AM (H9) and reaches a peak between 12 AM and 1 AM (H1). We can classify these students as most active in the evening with peak activity at night. There are a total of 239 students clustered into this group. Cluster 1 in Figure 3 starts to see an activity increase from 8 AM and reaches a peak activity at 2 PM and 4 PM and suddenly drops till 7 PM in the evening and gradually stabilizes there till 11 PM and then dips during the nighttime. There is a total of 560 students in cluster 1 that are assigned to this pattern. In final cluster 2, consisting of 889 students, the pattern is similar to cluster 1, but it doesn't stabilize at 7 PM. The dip in activity of these students continues with very few of them active at night. Based on the patterns found in these three clusters, we can classify cluster 0 as "Afternoon to Night," cluster 1 as "Afternoon to Evening," and cluster 2 as "Active Afternoon." Even though these three clusters have some similarities in the afternoon login patterns, where most are active, we can also see a peak in the evening between 7 PM and 8 PM. This specific time slot is interesting as the variations in cluster patterns started in the evening except for this particular time band. To understand any relationship between student demographics and chronotype
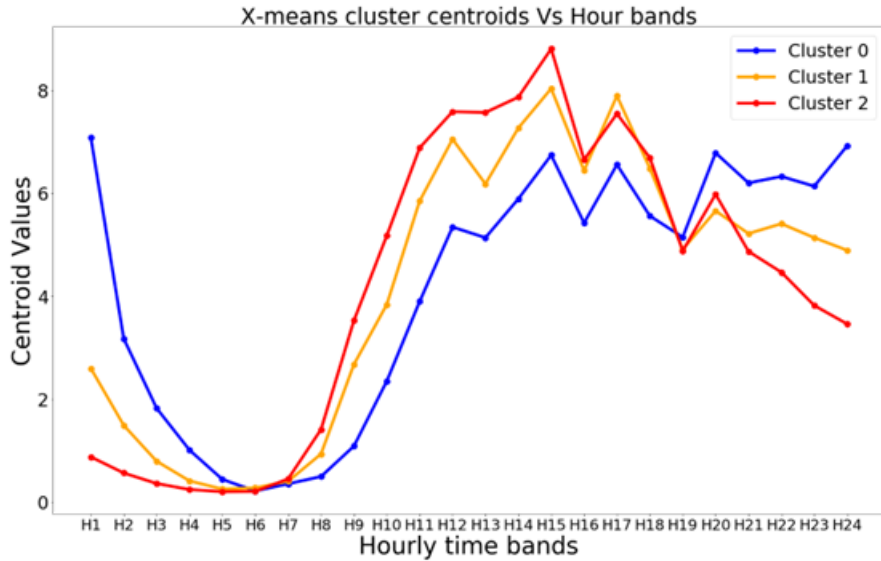
Figure 3: Clustering outcome of normalized hourly login volume data to study student chronotype

clusters, we adopt the chi-square significant test and record the p-values that are listed in Table 4. From Table 4, we can observe a meaningful relationship between student ethnicity and chronotypes. Another significant relation is found between student enrollment type (full-time or part-time) and their chronotypes. This can be related to the varying course loads and a difference in the nonacademic workload of part-time students compared to full-time students. Finally, as a student progresses through their undergraduate years, we can see a shift in chronotypes.

Table 4: Student Demographics and their corresponding p-values based on chi-square test with Chronotype Clusters.

Demographic	P-value
Gender	0.0571
Ethnicity	0.0001
Full Time or Part Time	0.007
Regular or Transfer	0.384
Student Year	0.018
GPA Bands	0.855

6.2. CORRELATION ANALYSIS

6.2.1. Predictive Model Performance

The five machine learning models adopted in this study were evaluated using a five-fold cross-validation method. In this method, the student data is divided into five

Table 5: Performance evaluation of predictive models based on R squared and RMSE values.

Model	R Squared	RMSE
LR	0.26	0.73
DT	0.24	0.74
SVM	0.20	0.77
RF	0.36	0.68
GBT	0.36	0.68

Table 6: Validating the generalizability of models developed on Fall 2019 student data to different terms.

Term	R-squared	RMSE	Student Count (IS & CS)
Fall 2019 (Main Model)	0.362	0.622	1559
Spring 2019	0.370	0.683	1545
Fall 2018	0.360	0.722	1486

equal folds at a student level. Four of the five folds are used for model training in every iteration, and one fold is used for model testing. The machine learning models are evaluated based on two performance metrics: R squared (R^2) and Root Mean Squared Error (RMSE). The output performance metrics are the average of five test fold performances. From the table 5 we can see that both Random Forest and Gradient Boosting Tree performed similarly with a R-squared value of 0.36 and RMSE value of 0.68.

We also validated the best performing model (GBT) resulted from current and earlier study on Fall 2019 middle of semester data by testing them on Spring 2019 and Fall 2018 student data related among IS and CS students. The results in Table 6 shows that student-centric models can be generalized across different semesters. Additionally, this work tested further the the model's external validity by validating them on student data from other departments in the Fall 2019 semester. From Table 7, we can observe that the test performance of the model developed on IS and CS student data showed similar performance metrics on students enrolled in four other departments.

Table 7: Validating the applicability of models developed on IS & CS student data to students from other departments

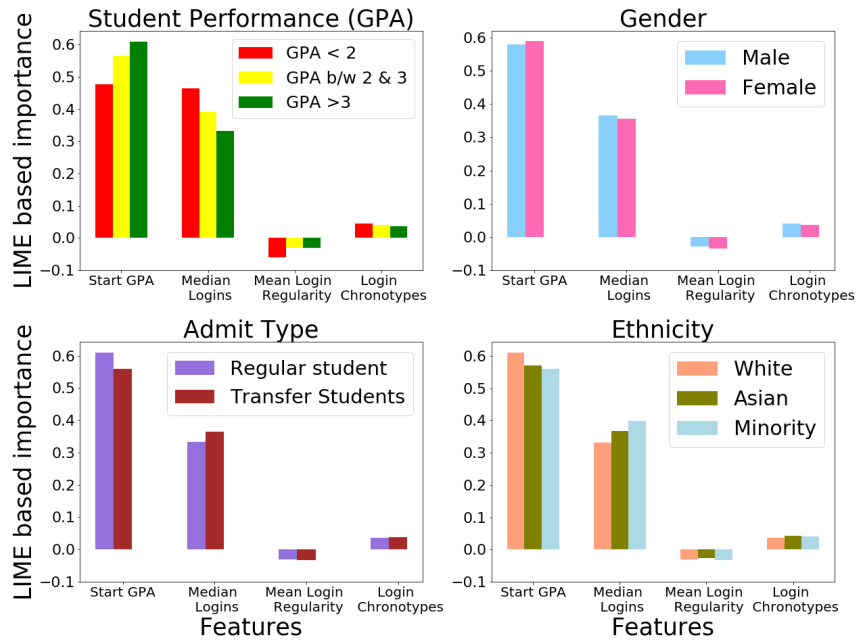| Department	Degree	R-squared	RMSE	Student Count
IS & CS (Main Model)	BS	0.362	0.622	1559
Biological Science	BS	0.381	0.675	491
Bio Chemical & Molecular Biology	BS	0.381	0.737	226
Mechanical Engineering	BS	0.385	0.657	391
Psychology	BS	0.317	0.849	215

Figure 4: LIME importance's for different student groups divided based on GPA, ethnicity, admit type and gender.

6.2.2. Model Explanation

Given the large set of features we are exploring, we are interested in gaining insights of the relative importance of features in predicting students' performance. This step is crucial when we move the model from the lab into the real-world to inform decision-making. We explored three different approaches that can provide complementary views into the model insights. While LIME method attempts to explain blackbox models, simple and linear model based approach interpretable by design. In explaining the model, we pay special attention to the role of students' demographics attributes may play in revealing the heterogeneity of feature importance in predicting student outcomes.

APPROACH #1: LIME BASED FEATURE IMPORTANCE LIME-based approach extract feature importance at the local level also called local fidelity. By applying the LIME method explained in the methodology section, we extract feature importance for different student groups categorized based on their demographics.

From Figure 4, we can observe that cumulative student GPA at the start of the semester is an important feature to predict student end-of-term GPA. Student login volumes are the second important feature set for model predictions on different student demographics. This study's focus is also on student self-regulation capability measured by the regularity of logins (entropy). This entropy measure is labeled as "mean login regularity" in 4 and in general higher entropy means less regular. We observe that

23

for students with GPA values less than 2, the regularity of the logins feature played a key role compared to a student with a higher GPA. In case of minority students, we observe that login volumes have higher importance on their performance when compared to other student ethnic groups. We also observe that the importance of login regularity to predict academic performance does not have significant variation when compared between different ethnic groups. One implication from these observations is that introducing teaching practices that guide LMS use and time management will significantly impact students with low GPA and a minority race. Start GPA played a slightly less significant role in transfer students than regular students as transfer students join in different years. Their cumulative GPA might not be available at the start of the semester, similar to freshman.

Even though there is a huge imbalance in the number of male and female students present in the dataset, we do not observe any significant difference in feature importance between these two genders. One limitation of the LIME method is related to global importance. The importance showed by LIME at the local level do not necessarily correspond to global importance.

APPROACH #2: CORRELATION-BASED FEATURE IMPORTANCE As earlier LIME based feature importance methods showed a significant impact of login volumes to predict student performance, we adopt Pearson correlation statistic to infer this relationship for different student groups. To do this, we create subsets of student data based on different groups: student GPA, gender, ethnicity, and admit type.

From Figure 5, we observe that the student login count and regularity in logins are highly significant for a student with a GPA lower than 2. We can also observe that as the entropy increases (i.e. students following a less regular schedule), the GPA reduces, as inferred from the negative correlation between the two variables. This observation seems to corroborate the understanding that regularity in student logins reflects their higher self-regulation capabilities which correlate with better learning outcomes. Earlier research showed that students with good self-regulation capabilities perform better in class (Kim et al., 2014; Park and Jo, 2015). For other student groups divided based on gender and admit type, there is no significant variation in the importance of login volume and login regularity on student performances.

APPROACH #3: REGRESSION MODEL BASED FEATURE IMPORTANCE One significant limitation of earlier methods is their inability to capture interaction effects as feature importance might change in the presence of other features. To study the interaction effects, we apply a linear regression model on different categories of student login data collected till the middle of the semester. These student categories were divided based on GPA, gender, admit type, and ethnicity of students. Even though linear regression models are applied on all features discussed in earlier sections, we only report the coefficients of median login volume and mean login regularity in Table 8, as these variables are the focus of this study. From Table 8, we observe that login volume and login regularity features follow similar directions for students with lower GPAs and students from minority ethnic backgrounds as observed in the LIME and correlation-based anal-
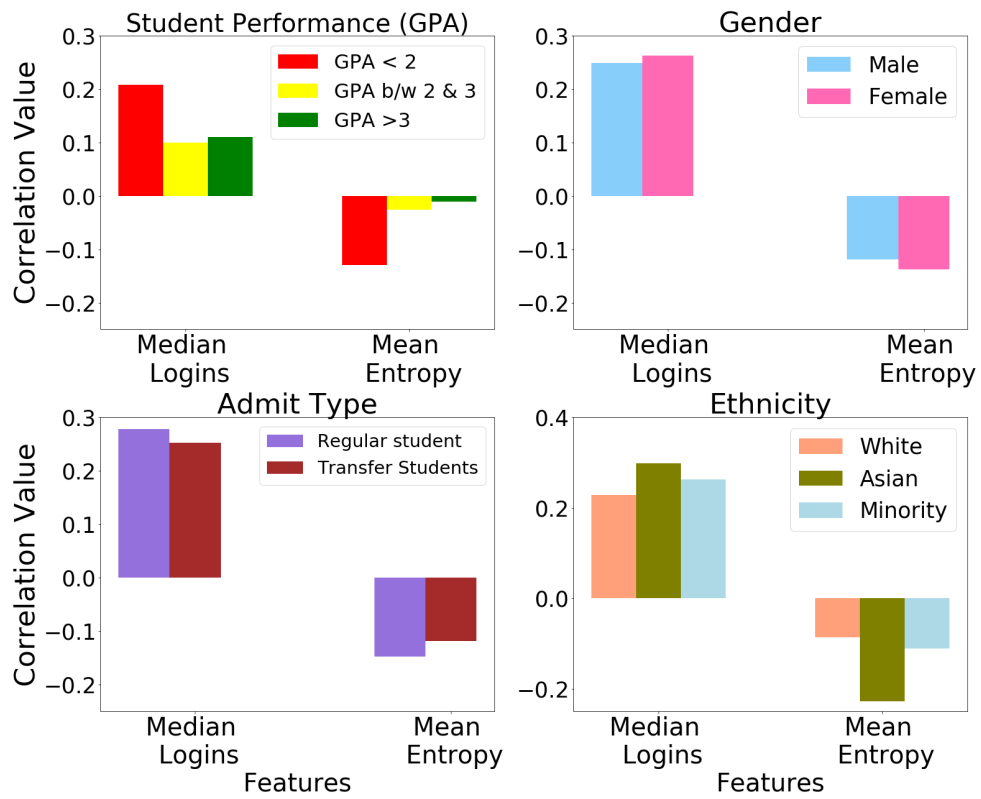
Figure 5: Correlation values for different student groups divided based on GPA, ethnicity, admit type and gender.

Table 8: Regression coefficients (Significance marked with *)

Student Demographic	Student Groups	Median Logins Coefficient	Mean Login Regularity Coefficient
GPA	GPA ≤ 2	**0.171***	**-0.398***
	GPA >2 & ≤ 3	-0.013	0.200
	GPA >3	0.065	-0.002
Gender	Male	0.135	0.130
	Female	-0.021	-0.157
Admit Type	Regular	0.399	-0.004
	Transfer	0.611	0.191
Ethnicity	White	0.204	-0.029
	Asian	-0.085	0.201
	Minority Race	**0.201***	**-0.153***

ysis. Another reason for focusing on a student from these two groups is their higher attrition rates found in earlier studies (Kupczynski et al., 2011; Liu and Cavanaugh, 2011). At least at surface level these findings suggest a potential benefit for targeting these two student groups as these groups of students happens to have higher attrition rates.

6.3. CAUSAL ANALYSIS

6.3.1. Aggregated Level Causal Analysis

This part of this study explores the causal relationships between student login variables, chronotypes, and performance outcomes treating students from a homogeneous group. We adopt PC variant and GFCI causal discovery algorithms. As discussed in 5.3.1 the PC variant algorithm gives high level understanding of variables and is able to incorporate categorical variables whereas GFCI algorithm can indicate confounding relationships and can be used to quantify relationships based on Structural Equation Modeling. Our focus is to study how login variables are interacting with performance and the impact of demographics incase of PC variant algorithm and studying indepth quantifiable interaction between login behaviors and student performance with GFCI algorithm. To do this, we apply a PC Variant algorithm in TETRAD that can handle both categorical and continuous variables. The composite variable input for the PC algorithm is calculated based on mCCA. In addition, we also incorporate common-sense knowledge to constraint the directionality of cause and effect. For example, we stipulate that demographic variables can only be a cause not an effect as they cannot

Table 9: Graph connector type descriptions

Connector Type	Description
A → B	A is a cause of B. It may be direct or indirect cause
A o→ B	Either A is a cause of B or there is an unmeasured confounder of A and B
A ←→ B	There is definitely an unmeasured confounder of A and B
A → B	There is a direct relationship between A and B (no confounders)
A —o B	This can be any of the above cases or a combination of them. A relation exists but not clear on directionality of directness.

be changed. The login variables are both cause and effect as they are malleable with interventions or behavioral changes. These variables act as a causal explanation for end-of-term GPA but can be affected by student demographics.

The feature weights related to mCCA are listed in Appendix Table 10. The causal relationships between variables are shown in Figure 6. From Figure 6, we observe that the student academic performance has a direct causal relationship with Normalized login volume and student prior performance. In addition, this analysis also shows that chronotype is causally linked to student enrollment type (part-time or full-time), and student year. Other demographics like gender and admit type (regular/transfer students) does not causally linked to login variables nor performance. Even though there are other causal relationships between different student variables, they are not directly impacting the performance. All the relationships displayed in Figure 6 are significant as we set a p-value cutoff at 0.05.

One of the limitation of the PC variant algorithm is the lack of confounder identification. To identify if the causal relations are more robust or if there is a presence of any confounder, we employ the GFCI algorithm in TETRAD. Additionally, we also explore the causal relationships by dividing student data into subsets based on their different demographic features. The interpretation of causal connectors related to GFCI is given in Table 9.

Figure 7 shows the output of the GFCI algorithm related to continuous student login variables. It shows a direct causal relationship between student normalized login volume and their academic performance from this Figure 7. The relationship between student prior performance (Cumulative GPA till the start of the semester) and the current end of term GPA observed in PC variant analysis shown in Figure 6 also exists in GFCI output. But it is not clear if this is a direct relationship or if there might be an unmeasured confounder between these two. Consistent relations identified by PC Variant and GFCI methods are found between student login regularity & student login volume, student hourly login volume & regularity, and Weekend/WeekDay login volumes & login volumes as well as regularity. Even though multiple other relationships
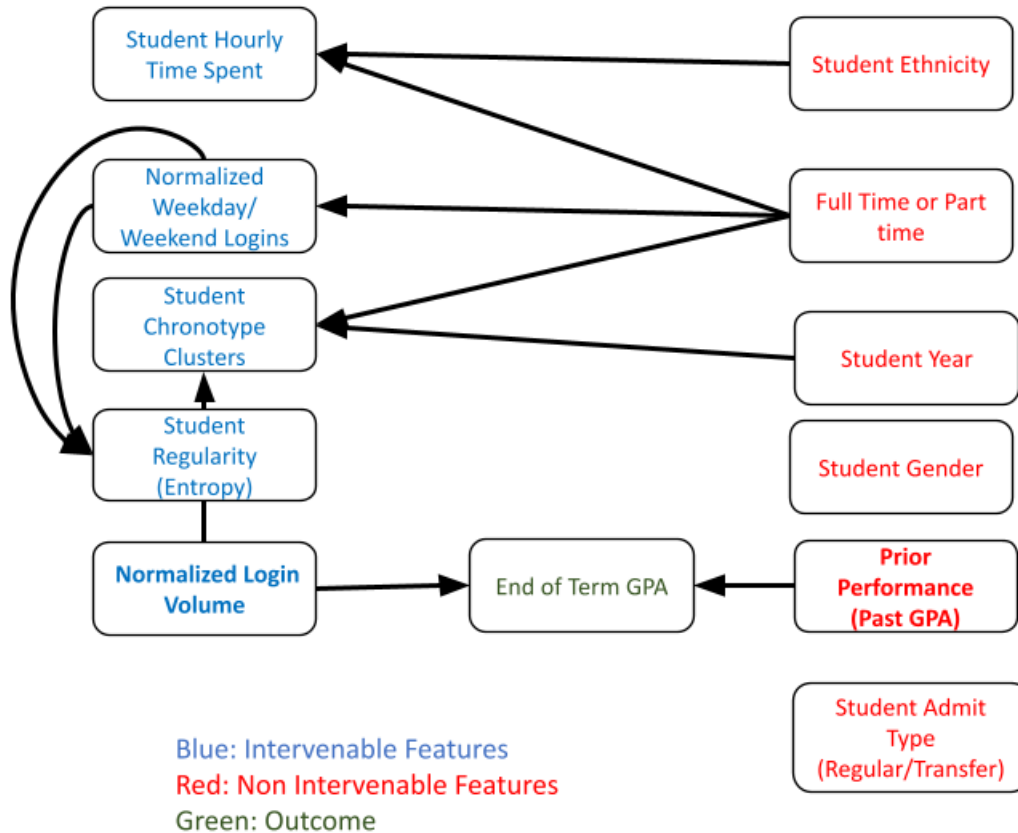
Figure 6: Causal relationship exploration between student demographics, login behaviors and chronotypes using PC Variant Algorithm.

are common in both PC Variant analysis and GFCI for different variables, we can see no direct relationships that impact student performance except login volume and prior performance.

It should be noted that Figure 7 does not fully specify causal relationships as there are some relationships with confounding factors. In addition to this, the GFCI causal graph in Figure 7 only shows which variable acted as cause and which variable had an effect but it does not detail the strength and directionality of cause and effect relationship. The causal graph shown in Figure 7 is used to generate a Directed Acyclic Graph (DAG) based on domain knowledge. This DAG is used as an input to estimate linear Gaussian Structural Equation Models (SEM). We then calculate the model's goodness of fit and coefficients to study if the causal relationship between the features is as expected based on domain knowledge. For example, for the pair of variables prior performance and End of Term GPA, we believe that there is a direct relationship between these two as prior performance can only be a cause but not an effect. For other variables with confounding measures, we can expect any directionality of cause and effect relationship as they all are relevant and drawn from the similar statistic. This is the reason we assign a double-headed arrow to other relationships with confounding factors.
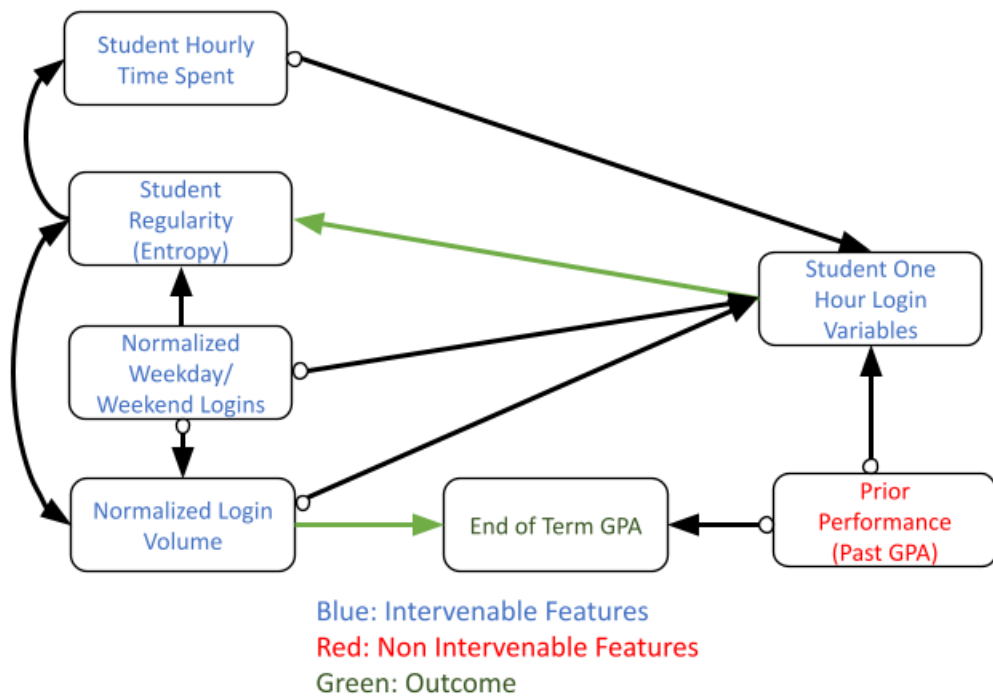
Figure 7: Causal relationship between Student login variables and their academic performance using GFCI algorithm
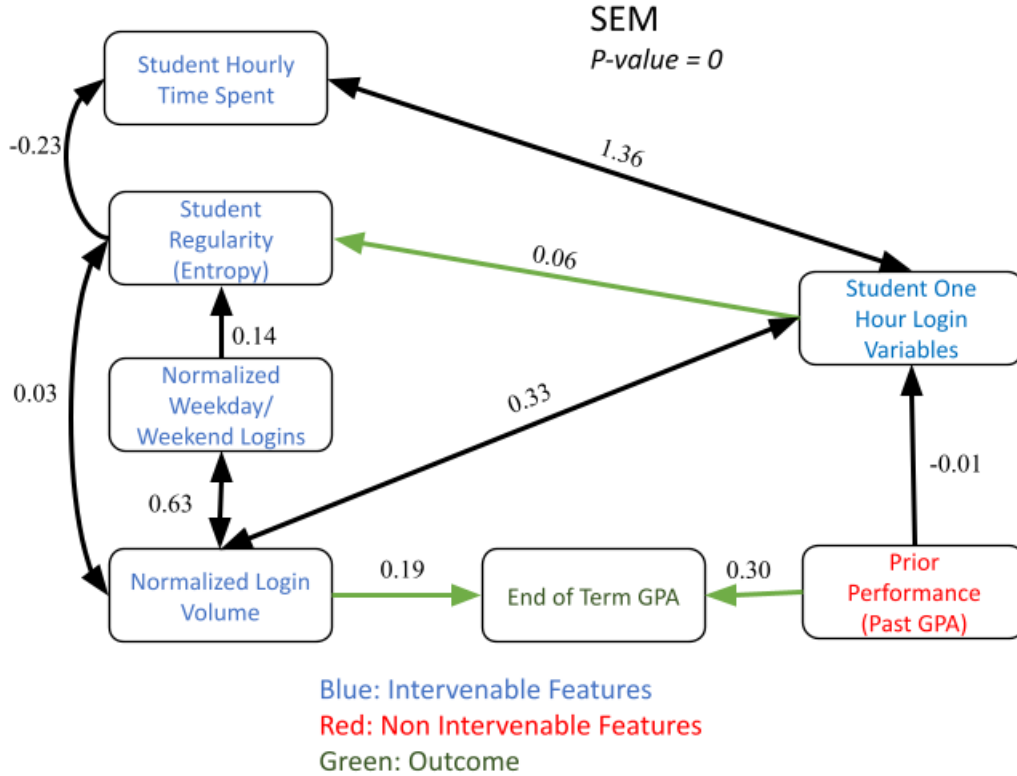
Figure 8: DAG derived from GFCI causal graph. Numerical values on directed arrows represent coefficients, and the values bidirectional arrows represent covariance reported by SEM

Figure 8 represents DAG derived from the causal graph in Figure 7. The values indicated on arrows represent coefficients if it is a unidirectional arrow and covariance if it is a bidirectional arrow. From this DAG, we can observe that Login volume causes End of term GPA and if the login volume increases, student GPA increases as the coefficient value (0.19) is positive. A similar relationship can be observed between student prior performance and current term end GPA. The model's goodness of fit reveals a significant relationship between all the edges as all of their p values are close to 0 and less than 0.05. From Figure 8, we can confirm that the causal relationships observed in the GFCI algorithm are sensible.

6.3.2. Group specific causal analysis

In the next step, we divide the data into subsets based on gender, race, and GPA to study causal relationships based on these demographics. In line with our demographic based correlation studies in 6.2.2, We only focus on "login volume", "Student Regularity (Entropy)" and "Prior performance". The gender based feature weights related to mCCA are mentioned in Appendix tables 11 and 12. From 9 (a), we can observe that there is no direct relationship between login variables and performance for different
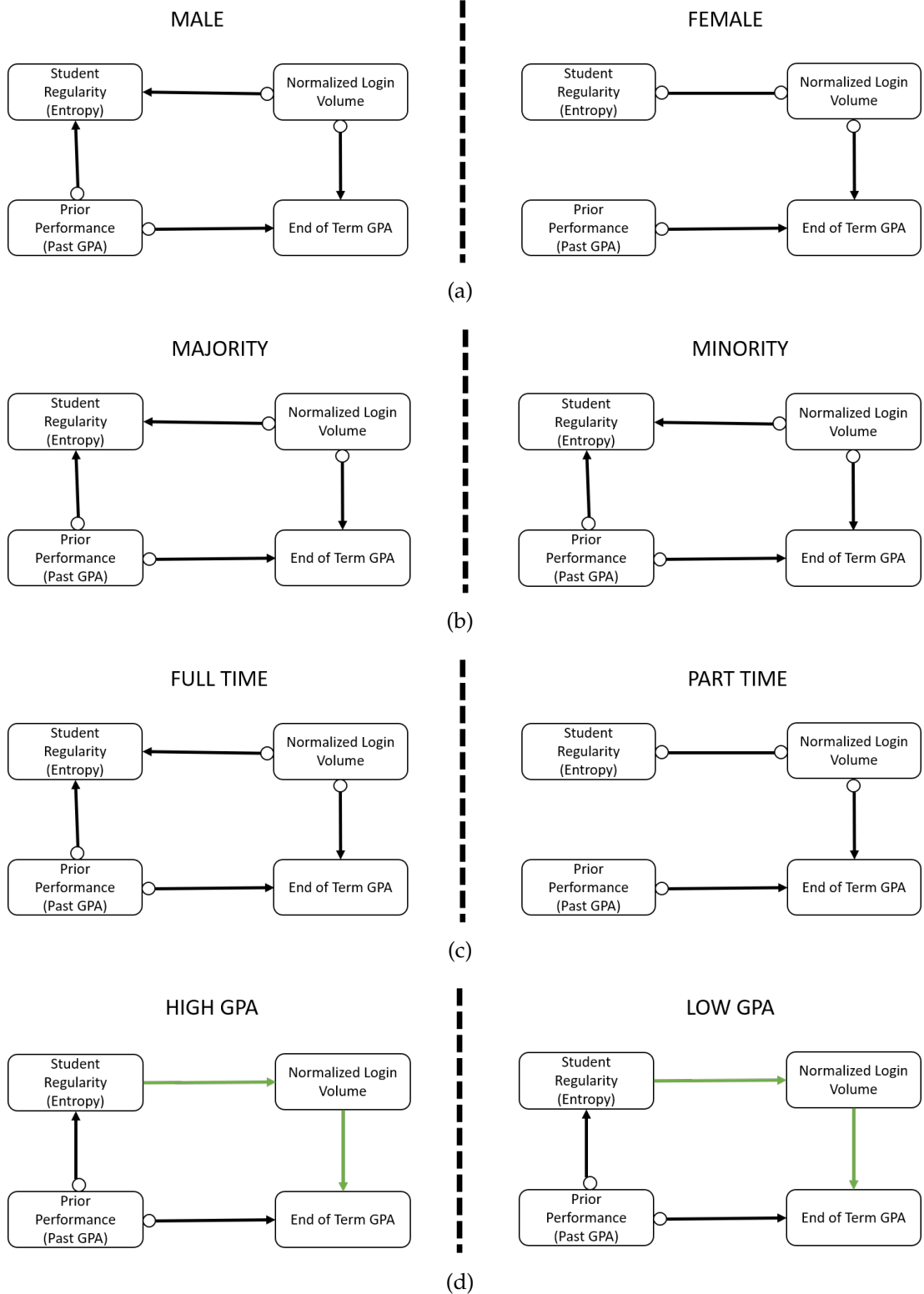
Figure 9: (a) Causal relationships based on gender-specific dataset (b) Causal relationships based on student ethnicity (c) Causal relationships based on student enrollment type (Full-time/Part-time) (d) Causal relationships based on student GPA

gender and race classifications. One difference between Male and Female gender is the impact of student prior performance on student login regularity.

There is a relationship between student regularity and their performance for Males, but for females, this does not exist. From 9 (b), we do not see any significant difference in relationships between students from Majority races and students from Minority races. The mCCA feature weights related to ethnicity based datasets are shown in Appendix tables 13 and 14. One interesting observation is the causal relationship between login volume and performance for minority race students, which we observed in our earlier work [210] with predictive analytics as well. The relationship between prior performance and regularity exists in full-time students but not part-time students, as shown in 9 (c). The corresponding mCCA weights for Full time and Part time students are shown in Appendix tables 15 and 16. There is a significant and direct relationship between login variables and student performance for students with varying bands of GPA (< 2 Low GPA & > 3 High GPA), as shown in 9 (d). The mCCA weights based on GPA are shown in Appendix Table 17 and Table 18. Even though there are no clear relationships without confounders based on gender and race, it is still valuable to learn that a relationship exists between student logins and their performance.

7. DISCUSSION

With the increasing adoption of LMS systems in colleges and universities, there is growing interest in understanding student learning behaviors based on their interaction with these systems. Studies in Learning Analytics and Educational Data Mining have been focusing on developing models that predict student performance based on their interaction data captured by these systems. While those models provides certain insights on the correlational relationship between *what students do*(as characterized by their LMS interactions) and *what will happen in the future*(e.g. their academic performance), they may not be directly translated into actionable insights due to the inherent limitation of correlation analysis. Taking a **student-centric** view in analyzing LMS activities, we explored an analytical framework that can triangulate multiple aspect of students login behaviors including login volume, regularity and chronotype and quantify their relationships with academic performance. In order to elucidate the contributing causal factors to students' learning outcome, we goes beyond the correlation analysis by employing a framework of causal discovery and inference. Here, we summarize a few key findings:

7.1. INSIGHTS FROM CAUSAL ANALYSIS

From the analysis in this study, we noted that student login volume is not only correlated, but also causally linked to students' academic performance. Moreover, it also revealed that login volume is influenced by student other self-regulation indicators such as logins regularity. This is an important finding as this shows that targeting an intervention that can increases student logins or their login regularity might improve their overall academic performance. The group-based study also demonstrated interesting

and important heterogeneity of the impact of logins on students from different backgrounds. For example, we note the relationships between self-regulation variables and performance are stronger in students with lower GPA and also for students from minority groups. However, we found no significant gender difference in terms of the impact of LMS features on outcomes, nor noteable difference between regular and transfer students. These findings confirm that studies at the student level reveal valuable insights for various groups of students and can act as strong inputs for effective intervention development.

7.2. INSIGHTS FROM MODEL EXPLANATION ON RICH SET OF LMS BEHAVIORS

In this study, we focused on understanding the impact of a rich set of LMS behavior indicators on student performances. In addition to login volume, which was the primary focus of earlier studies, we also studied the impact of login regularity measured by entropy statistics on student performance by implementing LIME explanation, correlation, and linear regression methods. From our model explanation studies, we observed that students who login regularly into the LMS system have a positive relationship with performance improvement. The model explanation outcomes on this data also showed a positive relationship between increase in student login volumes and GPA. Even though this observation is accurate across all students, it has a slightly higher importance in students from minority races and student with GPA lower than 2 based on the data analyzed. One major limitation with LIME and other correlation-based methods is their lack of explanation of cause-and-effect relationship that is crucial for interventions. To overcome this, we are performed a causal analysis and discovery to study the relationship between LMS features and student performance.

7.3. INSIGHTS FROM CHRONOTYPE ANALYSIS

One of the primary contributions of this study is the identification of student chronotypes based on their interaction variables. Earlier studies discussed the chronotype patterns in individuals and their impact on productivity (Díaz-Morales and Sorroche, 2008; Preckel et al., 2011a; Preckel et al., 2011b). These studies showed that individuals at younger ages are highly active in the morning, but this shifts to evening as they reach adolescence. The clustering methodology employed in this study showed a similar pattern where students in an undergraduate university showed high activity during the afternoon to evening hours. Prior studies also showed that these chronotypes have statistically significant relationships with academic achievement. However, these studies are primarily performed on a single domain or course (Zavgorodniaia et al., 2021). Our analysis explicitly targets students as a single entity and aggregates their activity across all enrolled courses.The statistical findings in our work showed significant relationships between student demographics and chronotypes but not between chronotypes and their performance. In addition to this, we also performed a predictive modeling methodology to study if these chronotypes contribute to student performance prediction. The findings from this analysis revealed that student performance prediction stayed similar with and without the chronotype variables. These observations

33

align with some earlier studies ([Preckel et al., 2011a](); [Zavgorodniaia et al., 2021]()) that contradicted the finds of relationships between chronotypes and performance. Overall, research on the relationship between student performance and chronotypes needs more investigation as the findings between different studies are inconsistent.

8. LIMITATION

There are also some significant limitations in this study. The data captured by LMS is only a snapshot of activity in students' day-to-day lives. Student performance factors can be influenced by many other external factors like study environment, family background, and student perception towards a course. The login-based chronotypes discussed in this study are only a part of time management. Students' time management can be observed from multiple other factors like assignment submissions, time spent on course contents, and exploring time lag between lecture delivery and student reading content access. Another limitation is related to the dataset. This aggregate-level dataset only captures the student login information but not content-level information. Content level information is much more fine-grained and provides much more insights into what a student might be working on when they are logging into the system. One significant challenge with content-level information is their diversity based on course and instructor style. This will make it hard to extract student-level aggregated features.

9. CONCLUSION

To conclude, the causal analysis in this study strengthened our earlier findings that showed significant relationships between student login variables and their performance. The findings of this study are valuable to Educational Data Mining and Learning Analytics community as they support the design and development of interventions techniques based on LMS variables to improve student performance. As the data in this study is collected till the middle of the semester, the development and deployment of interventions at this stage provides valuable time to students for improvement and contribute to their academic achievement. In addition, This study explored a multi-prong methodology frameworks that involves model explanation and causal analysis, working together, they provide convincing evidence for interventions that targets students' LMS activity. The group-specific analysis further provide guidance on what specific student group to target. We envision that this type of methodological frameworks may stimulate many follow up work in this area to turn data into useful insights and further into action and intervention that has the promise to improve students' outcome.

10. APPENDIX

Table 10: Composite variables with associated original features and weights estimated by mCCA for all students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Max Volume	0.36
	Median Volume	0.93
Login Regularity (Entropy)	Mean Entropy	-0.36
	Median Volume	-0.93
Hourly Login Volumes	H1 to H9	<0.18
	H10 to H24	0.24
Hourly Time Spent	H1 to H8	0.14 to 0.17
	H9 to H21	0.21 to 0.26
	H22 to H24	0.19
Weekday/Weekend Login Volumes	WeekDay	0.99
	WeekEnd	0.1

Table 11: Composite variables with associated original features and weights estimated by mCCA for MALE students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Min Volume	-0.3
	Max Volume	-0.52
	Mean Volume	-0.54
	Median Volume	-0.55
	Standard Deviation Volume	-0.2
Login Regularity (Entropy)	Min Entropy	0.41
	Max Entropy	0.43
	Mean Entropy	0.57
	Median Entropy	0.57
	Standard Deviation Entropy	0.06
	Skewness Entropy	-0.02
	Kurtosis Entropy	0.002

Table 12: Composite variables with associated original features and weights estimated by mCCA for FEMALE students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Mean	0.11
	Median	0.99
Login Regularity (Entropy)	Mean	-0.99
	Median	-0.11

Table 13: Composite variables with associated original features and weights estimated by mCCA for Majority Race (White & Asian) students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Mean	-0.11
	Median	-0.99
Login Regularity (Entropy)	Mean	0.99
	Median	0.11

Table 14: Composite variables with associated original features and weights estimated by mCCA for Minority Race students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Mean	0.11
	Median	0.99
Login Regularity (Entropy)	Mean	-0.11
	Median	-0.99

Table 15: Composite variables with associated original features and weights estimated by mCCA for Full-time students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Mean	-0.56
	Median	-0.59
	Min	-0.13
	Max	-0.53
	Std	-0.07
Login Regularity (Entropy)	Mean	0.62
	Median	0.62
	Min	0.30
	Max	0.37

Table 16: Composite variables with associated original features and weights estimated by mCCA for Part-time students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Max	-0.99
	Median	-0.11
Login Regularity (Entropy)	Mean	0.99
	Median	0.11

Table 17: Composite variables with associated original features and weights estimated by mCCA for Low GPA ($\ll 2$) students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Mean	0.11
	Median	0.99
Login Regularity (Entropy)	Mean	-0.11
	Median	-0.99

Table 18: Composite variables with associated original features and weights estimated by mCCA for High GPA (≥ 3) students

Composite Variable	Feature Name	Weights
Normalized Login Volume	Mean	-0.36
	Median	-0.93
Login Regularity (Entropy)	Mean	0.93
	Median	0.36

References

AGUDO-PEREGRINA, Á. F., IGLESIAS-PRADAS, S., CONDE-GONZÁLEZ, M. Á., AND HERNÁNDEZ-GARCÍA, Á. 2014. Can we predict success from log data in vles? classification of interactions for learning analytics and their relation with performance in vle-supported f2f and online learning. *Computers in human behavior 31*, 542–550.

AHMAD UZIR, N., GAŠEVIĆ, D., MATCHA, W., JOVANOVIĆ, J., AND PARDO, A. 2020. Analytics of time management strategies in a flipped classroom. *Journal of Computer Assisted Learning 36,* 1, 70–88.

ALOKLUK, J. A. ET AL. 2018. The effectiveness of blackboard system, uses and limitations in information management. *Intelligent Information Management 10,* 06, 133.

BAEHR, E. K., REVELLE, W., AND EASTMAN, C. I. 2000. Individual differences in the phase and amplitude of the human circadian temperature rhythm: with an emphasis on morningness–eveningness. *Journal of sleep research 9,* 2, 117–127.

BAUGHER, D., VARANELLI, A., AND WEISBORD, E. 2003. Student hits in an internet-supported course: How can instructors use them and what do they mean? *Decision Sciences Journal of Innovative Education 1,* 2, 159–179.

BELL-PEDERSEN, D., CASSONE, V. M., EARNEST, D. J., GOLDEN, S. S., HARDIN, P. E., THOMAS, T. L., AND ZORAN, M. J. 2005. Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics 6,* 7, 544–556.

BERECHET, L. D. AND GEORGESCU, M. The road from blackboard learning management system to moodle learning management system in modern universities.

BIKTIMIROV, E. N. AND KLASSEN, K. J. 2008. Relationship between use of online support materials and student performance in an introductory finance course. *journal of education for business 83,* 3, 153–158.

CARSKADON, M. A., WOLFSON, A. R., ACEBO, C., TZISCHINSKY, O., AND SEIFER, R. 1998. Adolescent sleep patterns, circadian timing, and sleepiness at a transition to early school days. *Sleep 21,* 8, 871–881.

CHEN, L. K., RAMSEY, J., AND DUBRAWSKI, A. 2021. Affect, support, and personal factors: Multimodal causal models of one-on-one coaching. *Journal of Educational Data Mining 13,* 3, 36–68.

COLDWELL, J., CRAIG, A., PATERSON, T., AND MUSTARD, J. 2008. Online students: Relationships between participation, demographics and academic performance. *Electronic journal of e-learning 6,* 1, 19–30.

CROWLEY, S. J., ACEBO, C., AND CARSKADON, M. A. 2007. Sleep, circadian rhythms, and delayed phase in adolescence. *Sleep medicine 8,* 6, 602–612.

DABBAGH, N. AND KITSANTAS, A. 2013. Using learning management systems as metacognitive tools to support self-regulation in higher education contexts. In *International handbook of metacognition and learning technologies*. Springer, 197–211.

DAMIANOV, D. S., KUPCZYNSKI, L., CALAFIORE, P., DAMIANOVA, E. P., SOYDEMIR, G., AND GONZALEZ, E. 2009. Time spent online and student performance in online business courses: A multinomial logit analysis. *Journal of Economics and Finance Education 8,* 2, 11–22.

DAWSON, S., HEATHCOTE, L., AND POOLE, G. 2010. Harnessing ict potential: The adoption and analysis of ict systems for enhancing the student learning experience. *International Journal of Educational Management*.

Díaz-Morales, J. F. and Sorroche, M. G. 2008. Morningness-eveningness in adolescents. *The Spanish journal of psychology 11,* 1, 201–206.

Dibner, C., Schibler, U., and Albrecht, U. 2010. The mammalian circadian timing system: organization and coordination of central and peripheral clocks. *Annual review of physiology 72,* 517–549.

Dominguez, M., Bernacki, M. L., and Uesbeck, P. M. 2016. Predicting stem achievement with learning management system data: Prediction modeling and a test of an early warning system. In *EDM*. ERIC, 589–590.

Dutt, A. and Ismail, M. A. 2019. Can we predict student learning performance from lms data? a classification approach. In *3rd International Conference on Current Issues in Education (ICCIE 2018)*. Atlantis Press, 24–29.

Edwards, S. H., Snyder, J., Pérez-Quiñones, M. A., Allevato, A., Kim, D., and Tretola, B. 2009. Comparing effective and ineffective behaviors of student programmers. In *Proceedings of the fifth international workshop on Computing education research workshop*. 3–14.

Enright, T. and Refinetti, R. 2017. Chronotype, class times, and academic achievement of university students. *Chronobiology international 34,* 4, 445–450.

Fancsali, S. 2014. Causal discovery with models: behavior, affect, and learning in cognitive tutor algebra. In *Educational Data Mining 2014*. Citeseer.

Fancsali, S. 2015. Confounding carelessness? exploring causal relationships between carelessness, affect, behavior, and learning in cognitive tutor algebra using graphical causal models. In *EDM*. 508–511.

Gray, R. M. 2011. *Entropy and information theory*. Springer Science & Business Media.

Greenland, S. 2011. Using log data to investigate the impact of (a) synchronous learning tools on lms interaction. In *Proceedings of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE) Conference*.

Hardoon, D. R. and Shawe-Taylor, J. 2011. Sparse canonical correlation analysis. *Machine Learning 83,* 3, 331–353.

Hofmann, M. and Klinkenberg, R. 2016. *RapidMiner: Data mining use cases and business analytics applications*. CRC Press.

Hung, J.-L. and Zhang, K. 2012. Examining mobile learning trends 2003–2008: A categorical meta-trend analysis using text mining techniques. *Journal of Computing in Higher education 24,* 1, 1–17.

Ikink, G. 2019. Vu research portal.

Ilves, K., Leinonen, J., and Hellas, A. 2018. Supporting self-regulated learning with visualizations in online learning environments. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. 257–262.

Jo, I.-H., Kim, D., and Yoon, M. 2015. Constructing proxy variables to measure adult learners' time management strategies in lms. *Journal of Educational Technology & Society 18,* 3, 214–225.

Killgore, W. D. and Killgore, D. B. 2007. Morningness-eveningness correlates with verbal ability in women but not men. *Perceptual and motor skills 104,* 1, 335–338.

Kim, J. H., Seodaemun-gu, S., Park, Y., Song, J., and Jo, I.-H. 2014. Predicting students' learning performance by using online behavior patterns in blended learning environments: comparison of two cases on linear and non-linear model. *Korea 120,* 750.

KOEDINGER, K. R., MCLAUGHLIN, E. A., JIA, J. Z., AND BIER, N. L. 2016. Is the doer effect a causal relationship? how can we tell and why it's important. In *Proceedings of the sixth international conference on learning analytics & knowledge*. 388–397.

KORCZAK, A., MARTYNHAK, B., PEDRAZZOLI, M., BRITO, A., AND LOUZADA, F. 2008. Influence of chronotype and social zeitgebers on sleep/wake patterns. *Brazilian Journal of Medical and Biological Research 41*, 914–919.

KRASKOV, A., STÖGBAUER, H., AND GRASSBERGER, P. 2004. Estimating mutual information. *Physical review E 69*, 6, 066138.

KUPCZYNSKI, L., GIBSON, A. M., ICE, P., RICHARDSON, J., AND CHALLOO, L. 2011. The impact of frequency on achievement in online courses: A study from a south texas university. *Journal of Interactive Online Learning 10*, 3.

KYLLONEN, P. C. AND CHRISTAL, R. E. 1990. Reasoning ability is (little more than) working-memory capacity?! *Intelligence 14*, 4, 389–433.

LANDRUM, B. 2020. Examining students' confidence to learn online, self-regulation skills and perceptions of satisfaction and usefulness of online classes. *Online Learning 24*, 3, 128–146.

LERCHE, T. AND KIEL, E. 2018. Predicting student achievement in learning management systems by log data analysis. *Computers in Human Behavior 89*, 367–372.

LIU, F. AND CAVANAUGH, C. 2011. High enrollment course success factors in virtual school: Factors influencing student academic achievement. *International Journal on E-learning 10*, 4, 393–418.

LUST, G., VANDEWAETERE, M., CEULEMANS, E., ELEN, J., AND CLAREBOUT, G. 2011. Tool-use in a blended undergraduate course: In search of user profiles. *Computers & Education 57*, 3, 2135–2144.

MANDALAPU, V. 2021. Profiling and modeling student learning behaviors & outcomes from digital learning environments. Ph.D. thesis, University of Maryland, Baltimore County.

MANDALAPU, V., CHEN, L. K., CHEN, Z., AND GONG, J. 2021. Student-centric model of login patterns: A case study with learning management systems. *International Educational Data Mining Society*.

MANDALAPU, V. AND GONG, J. 2019. Studying factors influencing the prediction of student stem and non-stem career choice. In *EDM*.

MANDALAPU, V., GONG, J., AND CHEN, L. 2021. Do we need to go deep? knowledge tracing with big data. *arXiv preprint arXiv:2101.08349*.

MARTIN, J., EDWARDS, S. H., AND SHAFFER, C. A. 2015. The effects of procrastination interventions on programming project success. In *Proceedings of the eleventh annual International Conference on International Computing Education Research*. 3–11.

MIERSWA, I. AND KLINKENBERG, R. 2021. Rapid miner.

MONTARULI, A., CASTELLI, L., GALASSO, L., MULÈ, A., BRUNO, E., ESPOSITO, F., CAUMO, A., AND ROVEDA, E. 2019. Effect of chronotype on academic achievement in a sample of italian university students. *Chronobiology international 36*, 11, 1482–1495.

MÜLLER, M. 2007. Dynamic time warping. *Information retrieval for music and motion*, 69–84.

NANDI, D., HAMILTON, M., HARLAND, J., AND WARBURTON, G. 2011. How active are students in online discussion forums? In *Proceedings of the Thirteenth Australasian Computing Education Conference-Volume 114*. 125–134.

OGARRIO, J. M., SPIRTES, P., AND RAMSEY, J. 2016. A hybrid causal search algorithm for latent variable models. In *Conference on probabilistic graphical models.* PMLR, 368–379.

PARK, Y. AND JO, I.-H. 2015. Development of the learning analytics dashboard to support students' learning performance. *Journal of Universal Computer Science 21,* 1, 110.

PORCHERET, K., WALD, L., FRITSCHI, L., GERKEMA, M., GORDIJN, M., MERRROW, M., RAJARATNAM, S. M., ROCK, D., SLETTEN, T. L., WARMAN, G., ET AL. 2018. Chronotype and environmental light exposure in a student population. *Chronobiology international 35,* 10, 1365–1374.

PRECKEL, F., LIPNEVICH, A. A., SCHNEIDER, S., AND ROBERTS, R. D. 2011a. Chronotype, cognitive abilities, and academic achievement: A meta-analytic investigation. *Learning and Individual Differences 21,* 5, 483–492.

PRECKEL, F., LIPNEVICH, A. A., SCHNEIDER, S., AND ROBERTS, R. D. 2011b. Chronotype, cognitive abilities, and academic achievement: A meta-analytic investigation. *Learning and Individual Differences 21,* 5, 483–492.

PRECKEL, F. AND ROBERTS, R. 2009. Schulleistung und chronotyp [school achievement and chronotype]. In *Paper presented at the 10th meeting of the section of Differential Psychology of the German Psychological Society, University of Koblenz-Landau, Landau.*

PUTILOV, A. A., MARCOEN, N., NEU, D., PATTYN, N., AND MAIRESSE, O. 2019. There is more to chronotypes than evening and morning types: results of a large-scale community survey provide evidence for high prevalence of two further types. *Personality and Individual Differences 148,* 77–84.

PUTILOV, A. A. AND ONISCHENKO, M. A. 2005. The 52-item sleep-wake pattern assessment questionnaire: Evaluation of its psychometric features. *Ergonomia 27,* 2, 131–144.

PUTILOV, A. A. AND PUTILOV, D. A. 2005. Sleepless in siberia and alaska: Cross-validation of factor structure of the individual adaptability of the sleep-wake cycle. *Ergonomia 27,* 3, 207–226.

PUTILOV, A. A., SVESHNIKOV, D. S., PUCHKOVA, A. N., DOROKHOV, V. B., BAKAEVA, Z. B., YAKUNINA, E. B., STARSHINOV, Y. P., TORSHIN, V. I., ALIPOV, N. N., SERGEEVA, O. V., ET AL. 2021. Single-item chronotyping (sic), a method to self-assess diurnal types by using 6 simple charts. *Personality and Individual Differences 168,* 110353.

RAMSEY, J. D., ZHANG, K., GLYMOUR, M., ROMERO, R. S., HUANG, B., EBERT-UPHOFF, I., SAMARASINGHE, S., BARNES, E. A., AND GLYMOUR, C. 2018. Tetrad—a toolbox for causal discovery. In *8th International Workshop on Climate Informatics.*

RANDLER, C. 2007. Gender differences in morningness–eveningness assessed by self-report questionnaires: A meta-analysis. *Personality and Individual Differences 43,* 7, 1667–1675.

REPPERT, S. M. AND WEAVER, D. R. 2002. Coordination of circadian timing in mammals. *Nature 418,* 6901, 935–941.

RIBEIRO, M. T., SINGH, S., AND GUESTRIN, C. 2016. "why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.* 1135–1144.

ROBERTS, R. D. AND KYLLONEN, P. C. 1999. Morningness–eveningness and intelligence: early to bed, early to rise will likely make you anything but wise! *Personality and Individual differences 27,* 6, 1123–1133.

ROENNEBERG, T., WIRZ-JUSTICE, A., AND MERROW, M. 2003. Life between clocks: daily temporal patterns of human chronotypes. *Journal of biological rhythms 18,* 1, 80–90.

ROMERO, C. AND VENTURA, S. 2007. Educational data mining: A survey from 1995 to 2005. *Expert systems with applications 33,* 1, 135–146.

ROMO-NAVA, F., BLOM, T., GUERDJIKOVA, A., WINHAM, S., CUELLAR-BARBOZA, A., NUNEZ, N., SINGH, B., BIERNACKA, J., FRYE, M., AND MCELROY, S. 2020. Evening chronotype, disordered eating behavior, and poor dietary habits in bipolar disorder. *Acta Psychiatrica Scandinavica 142,* 1, 58–65.

ROSENMAN, R., TENNEKOON, V., AND HILL, L. G. 2011. Measuring bias in self-reported data. *International journal of behavioural & healthcare research 2,* 4, 320.

RUPING, S. 2000. mysvm-manual. *http://www-ai. cs. uni-dortmund. de/SOFTWARE/MYSVM/*.

SALEHIAN KIA, F., HATALA, M., BAKER, R. S., AND TEASLEY, S. D. 2021. Measuring students' self-regulatory phases in lms with behavior and real-time self report. In *LAK21: 11th International Learning Analytics and Knowledge Conference*. 259–268.

SCHAPIRO, S. R. AND LIVINGSTON, J. A. 2000. Dynamic self-regulation: The driving force behind academic achievement. *Innovative Higher Education 25,* 1, 23–35.

SCHMIDT, C., COLLETTE, F., CAJOCHEN, C., AND PEIGNEUX, P. 2007. A time to think: circadian rhythms in human cognition. *Cognitive neuropsychology 24,* 7, 755–789.

SHCHEDRINA, E., VALIEV, I., SABIROVA, F., AND BABASKIN, D. 2021. Providing adaptivity in moodle lms courses. *International Journal of Emerging Technologies in Learning (iJET) 16,* 2, 95–107.

SMITH, C. S., REILLY, C., AND MIDKIFF, K. 1989. Evaluation of three circadian rhythm questionnaires with suggestions for an improved measure of morningness. *Journal of Applied psychology 74,* 5, 728.

SONG, J. AND STOUGH, C. 2000. The relationship between morningness–eveningness, time-of-day, speed of information processing, and intelligence. *Personality and Individual differences 29,* 6, 1179–1190.

SPIRTES, P. AND GLYMOUR, C. 1991. An algorithm for fast recovery of sparse causal graphs. *Social science computer review 9,* 1, 62–72.

SPIRTES, P., GLYMOUR, C., AND SCHEINES, R. 1993. Discovery algorithms for causally sufficient structures. In *Causation, prediction, and search*. Springer, 103–162.

TEMPELAAR, D. T., RIENTIES, B., AND GIESBERS, B. 2015. In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior 47,* 157–167.

TOERING, T., ELFERINK-GEMSER, M. T., JONKER, L., VAN HEUVELEN, M. J., AND VISSCHER, C. 2012. Measuring self-regulation in a learning context: Reliability and validity of the self-regulation of learning self-report scale (srl-srs). *International Journal of Sport and Exercise Psychology 10,* 1, 24–38.

TORSVALL, L. AND ÅKERSTEDT, T. 1980. A diurnal type scale: construction, consistency and validation in shift work. *Scandinavian journal of work, environment & health*, 283–290.

VALLADARES, M., RAMÍREZ-TAGLE, R., MUÑOZ, M. A., AND OBREGÓN, A. M. 2018. Individual differences in chronotypes associated with academic performance among chilean university students. *Chronobiology international 35,* 4, 578–583.

VENGROFF, R. AND BOURBEAU, J. 2006. In-class vs. on-line and hybrid class participation and outcomes: Teaching the introduction to comparative politics class. In *annual meeting of the APSA Teaching and Learning Conference*.

VITALE, J. A., ROVEDA, E., MONTARULI, A., GALASSO, L., WEYDAHL, A., CAUMO, A., AND CARANDENTE, F. 2015. Chronotype influences activity circadian rhythm and sleep: differences in sleep quality between weekdays and weekend. *Chronobiology international 32,* 3, 405–415.

WITTMANN, M., DINICH, J., MERROW, M., AND ROENNEBERG, T. 2006. Social jetlag: misalignment of biological and social time. *Chronobiology international 23,* 1-2, 497–509.

ZAVGORODNIAIA, A., SHRESTHA, R., LEINONEN, J., HELLAS, A., AND EDWARDS, J. 2021. Morning or evening? an examination of circadian rhythms of cs1 students. In *2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)*. IEEE, 261–272.

ZIMMERMAN, B. J. AND KITSANTAS, A. 2014. Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary educational psychology 39,* 2, 145–155.