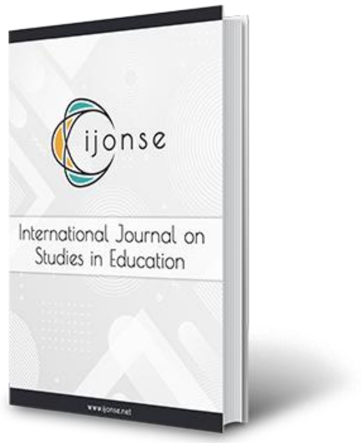

www.ijonse.net

Leveraging Machine Learning and AI to Enhance Educational Learning Analytics

Ikechukwu Ogbuchi
Atlantic Technological University, Ireland

To cite this article:

Ogbuchi, I. (2025). Leveraging machine learning and AI to enhance educational learning analytics. *International Journal on Studies in Education (IJonSE), 7*(3), 650-674. https://doi.org/10.46328/ijonse.1933

International Journal on Studies in Education

2025, Vol. 7, No. 3, 650-674 https://doi.org/10.46328/ijonse.1933

Leveraging Machine Learning and AI to Enhance Educational Learning Analytics

Ikechukwu Ogbuchi

Article Info

Article History

Received:
13 January 2025
Accepted:
4 June 2025

Keywords

Machine learning
Higher education
Artificial intelligence
Learning analytics

Abstract

Data-driven insights play a pivotal role in optimising learning analytics within higher education institutions. Despite their importance, much of the data in these institutions remains untapped, trapped in siloed data stores. This study addresses this challenge by applying machine learning and mathematical modelling using a learning analytics research framework (Khalil et al., 2022), encompassing phases of data insights, analytics, and intervention. The study aimed to identify features influencing learner progression and discover student groups in educational environments. Utilising data from 1,017 students over a 3-year period, the research employed unsupervised machine learning techniques and automated student feedback. Feature analysis identified attendance, interactions, time intervals, and activities related to quizzes and workshops as useful predictors of students requiring additional support. The study found that K-means model performs best with an average recall of 89% and an overall accuracy of 72% for identifying at-risk student groups using non-personally identifying data. The findings emphasise the utility of unsupervised machine learning for early identification of at-risk students, enabling timely interventions to prevent potential failure or dropout. Personalised and automated feedback forms summarising students' progression received high ratings, 93% rating for usefulness from students, highlighting their satisfaction with it as a learning analytics intervention.

Introduction

Learning analytics data is only valuable when it is transformed into a format which is interpretable and can provide actionable insights. This study investigated how some machine learning algorithms and mathematical modelling can enable digital and learning transformation in Higher Education by identifying factors affecting learners' success. Data for this study was from the School of Science and Computing within Atlantic Technological University (ATU) consisting of 3 academic sessions involving 1,017 first year students. Measuring student engagement has proven difficult, and data collected from interactions with content, grades in activities, and online attendance were often siloed and inaccessible.

Due to increased learning on online learning platforms, a vast amount of data was generated as students interacted with the learner content, and insights about the student's learning pattern were derived and potential behaviour predicted. Student engagement is a crucial factor in learning outcomes in higher education (Boulton et al. 2019). This study explored how machine learning can be leveraged to improve student engagement. More specifically, the research aims to establish patterns in the Moodle VLE data to identify student groups based on their engagement with their courses and interactions. The adoption of a data-science approach, enhanced by machine learning and AI, offers stakeholders greater insight, enhances student engagement, improves student outcomes, and facilitates operational improvements through real-time access to performance indicators.

This research was guided by the following key research questions:

1. Can historical data (three years) uncover patterns and insights between successful and unsuccessful students based on their interactions and learning behaviour?
2. Can machine learning models predict risk/struggling students with good accuracy? What are the early indicators that predict students' success or disengagement?
3. How effective is automated and personalised feedback forms as an intervention method within the learning analytics framework?

Literature Review

Learning Analytics, Student Engagement and Success

Learning analytics, an evolving research field, can help enhance the field of education and student success through more informed and focused interventions. The working definition of learning analytics according to Long & Siemens, (2011) is the measurement, collection, analysis, and reporting of data about learners and their contexts for purposes of understanding and optimising learning and the environments in which it occurs.Learning analytics can be a powerful tool to improve student success and engagement. It can be used to understand the relationships between students, educators, and their subject matter (O'Farrell, 2017). Learning analytics can be used to support at-risk students. Katerina et al., (2019) showed that learning analytics can support data-driven learning design when data are collected from various sources and come from a regular part of the students' learning process. Hongxin, (2020) highlighted that learning analytics can be used to support at-risk students in an institution. An institution can use learning analytics to identify students at risk of dropping out and provide them with additional support services. Learning analytics, although still a young field, is a powerful resource for informed decision making and better learning outcomes. (Flores-Vivar & García-Peñalvo, 2022) described how artificial intelligence (AI) in education can be useful to both students and educators. Lee O'Farrell, 2017 reiterates that "*Learning analytics can reveal a lot about the progress of learners and the suitability of the contexts in which learning takes place.*" Lee O'Farrell, 2017 reiterates that "*Learning analytics can reveal a lot about the progress of learners and the suitability of the contexts in which learning takes place.*"

It is critical that the science of learning and curriculum design is kept at the heart of analytics solution, recognising educators and students as key stakeholders and beneficiaries of learning analytics (Tsai et al., 2018). According to O'Brien, (2022) lower attendance rates and poor student engagement are among the current academic concerns as some education programs suffer high attrition rates, primarily due to the poor engagement of students with their classes. Farag , (2020) found high rate of attrition in the freshman year for an engineering programme. These challenges are primarily due to poor engagement of students in their classes. Research conducted by Hu et al. (2016) shows that there is little research on learning analytics focusing on student engagement online; emphasizing the importance and necessity of further research in this area.

Mohd et al., (2016) found that student engagement in learning correlates with good academic results and success. Hu et al. (2021) also found that student engagement is an important indicator of the effects of learning in higher education. Yoong, (2014) proposes that to improve student learning, there is a need to better understand students and analyse the data tracked from student learning activities online. Anthony, (2021) found that tracking activities and completion is useful to improve student engagement.

One way in which learning analytics can be used to improve student engagement is by tracking student activity and completion. Developing environments and tools to track these activities is useful in learning analytics. Pérez-Berenguer et al.,(2016) highlighted the benefit of developing environments to track these learning activities to improve learning analytics. Therefore, developing learning management systems and technology tools which track student activities and provide instructors with reports on student engagement is a valuable way of implementing learning analytics.

Research findings from Nurfadhlina et al.(2021) measured student engagement using learning analytics and concluded that course participation and achievement have a strong positive correlation. Although these studies have shown the usefulness of learning analytics and its value, there have been challenges in implementing it across educational institutions and deriving the greatest value.Tsai & Gasevic, (2017) noted that there are technical, organisational, and pedagogical challenges with learning analytics. One such is limited availability of learning analytics policies to address issues such as privacy. There are also limited number of studies empirically validating the impact of learning analytics intervention. Challenges have also been identified in the use and application of learning analytics (Banihashem et al., 2018).

Artificial Intelligence, Blended Learning and Learning Analytics

AI and machine learning are critical technologies used to enhance learning. Kuleto et al., (2021) showed that the use of analytics and AI can help universities mitigate emerging challenges with teaching students and design a curriculum that meets their needs. Waheed et al. (2020) highlighted that machine learning algorithms can be used to make predictions for at-risk university students from VLEs. According to Kew et al., (2017), the identification of learning indicators is of great importance, as it enhances prediction and identification of at-risk students. Mireilla (2017) highlights the importance of tracking digital footprint data on students' interaction with their feedback to identify students at risk of failing a module. These research findings demonstrate the relevance of technological tools in learning analytics.

Student engagement as a vital construct in understanding student learning behaviour could be used in evaluating technology-enhanced learning systems on their ability to properly impact students' learning. Fidelia et al. (2022) demonstrated that student engagement data are good indicators of academic performance and learning. By integrating learning analytics into blended learning process, students' learning can be better understood and promoted (Baolin, 2017). Mark (2021) revealed that all students and not just at-risk students could benefit from interventions informed by learning analytics, and research findings explored by Tuti (2020) show that predictive models leveraging machine learning methods are effective in exploratory data analytics of student data.

Blended learning refers to a learning method which combines traditional face-to-face learning style with online or e-activities (Attard & Holmes, 2022). This mode of learning is essential for providing institutions with a data source which can help harness the power of learning analytics. Adiguzel et al. (2020) found benefits of using digital products in student courses and saw the potential of blended learning products in the overall teaching and learning experience of students.

Learning Analytics and Machine Learning

The adoption of learning analytics enhanced by machine learning may offer stakeholders greater insight and enhance student engagement, thereby improving student outcomes and facilitating operational improvements through real-time access to key performance indicators. Machine Learning (ML) is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate how humans learn, gradually improving its accuracy (IBM, 2022).

Educational institutions store massive amounts of student data that are often left unanalysed for insights. Machine learning holds significant value for big data generated by educational institutions. Machine learning uses programming and mathematical tools to learn patterns in data, which are often difficult to discover. Types of Machine Learning include supervised, unsupervised, reinforcement learning, semi-supervised learning, and learning-to-learn (Oladipupo, 2010). Supervised learning techniques focus on teaching computers how to learn from data by showing what they should look out for. In unsupervised learning, data is fed into the model through programming, and the model automatically identifies patterns in the data without showing what to identify. In reinforcement learning, the model is trained using a reward system, whereby it acts in an environment and is penalised or rewarded according to the set of actions it takes.

Studies have shown that student dropout rates are increasing across various institutions, and concerns have been raised over low student engagement (O'Brien, 2022) as education receives large amount of government funding.Since 2016, there has been significant public reinvestment in higher education of €1.1 billion by the government, an increase of over 40% (Department of Furthur and Higher Education, 2022). Increased spending on education aims to create a high-quality learning environment for students and equip them for national development. However, student dropout rates are increasing (Dass et al., 2021).

Dass et al. (2021) considered the following features in predicting student performance: time and topics, topics mastered, topics practiced, and time spent. Daud et al. (2017) noted that the ability to predict the success or otherwise of a student is an interesting area because it could provide educational institutions with useful knowledge from their databases. This may aid institutions to provide additional support or adapt teaching styles to students proned to quitting. Daud et al. (2017) used information on family expenditure (electricity, gas, medical, and accommodation bills), family income (father, mother , personal, and miscellaneous income etc.) student personal information (gender, marital status, home ownership, previous institutions attended) and family assets (land , stock , house and vehicular values) to predict student performance to support analytics. However, the study acknowledged that in most cases, all these data were unavailable for a dynamic construction of student identity. Access to this information is also dependent on student's willingness to share them. Omar et al. (2020) applied a clustering approach for analysing student efficiency and performance using grade and GPA values in their dataset.

Nafuri et al. (2022) used clustering analysis to classify student academic performance in higher education using gender, registration age, marital status, place of birth, income status, entry qualification, sponsorship, residence, Curriculum Grade Point Average (CGPA) and employment status as predictors. The well-cited Gray (2015) did a thorough work exploring learner motivation, age relationship with performance, and self-efficacy. However, personal information of students, such as age, gender, and course of study, were used as features. The study found that these factors were not indicative or significant in creating classification models of students at risk of failing. Embarak (2020) also applied machine learning to predict students at risk using only high school math and English SAT scores, and grade. Although the study found good correlation between the features used and their eventual performance, it acknowledged discrepancies found between learners in different programs.

All these studies, have shown the use of personal information for model development. This means existing attempts to predict student performance and inform learning analytics have somehow been influenced by student's past, gender, and their past family circumstances.

Methodology

Methodology Framework

To decide on the appropriate methods to answer the research questions, it is important to choose an effective framework to guide the study. A critical analysis of learning analytics frameworks shows existence of valuable theoretical and pedagogical guidelines for educators (Kaliisa et al., 2022). A systematic review found a broad consensus on core elements of learning analytics frameworks, encompassing development, application, privacy, representation, data source, data type, focus/purpose, and educational context (Khalil et al., 2022).

A pictorial representation of the chosen framework is shown in Figure 1 which visually shows its constituents and considerations. The chosen framework below revolves around learners, processes, data interpretation, optimisation and some components involved in each stage of the framework. The goal of this framework is to enhance learning and lead to well-informed decision making in education institutions.

Based on the design goals formulated from reviewing existing literature, the framework used for this research is based on the learning analytics framework proposed by Khalil & Ebner, (2015) and Clow (2012) which consist of four key parts:

1. Learning Environment: This refers to the environment that produces data and everything constituting

this environment. Clow (2012) referred to this as “Learners” in his framework.

2. Big Data: This is a diverse collection of datasets captured from the learning environment. This data entails interactions, grades, learning activities, and academic information; referred to as the “Data” phase in the Clow framework.
3. Analytics: Different data analysis methods are employed to identify trends, patterns, and derive insights. Qualitative and quantitative data are analysed using statistics and are visualised. Predictive models are employed in this phase such as K-means, Agglomerative and Spectral clustering techniques. Clow called this the “Metric” phase.
4. Action: Here, the results from the analysis are used to improve the learning environment. It includes steps taken to intervene; personalisation of learning and change of learning policies to improve learner experience. The Clow framework calls it “intervention”.

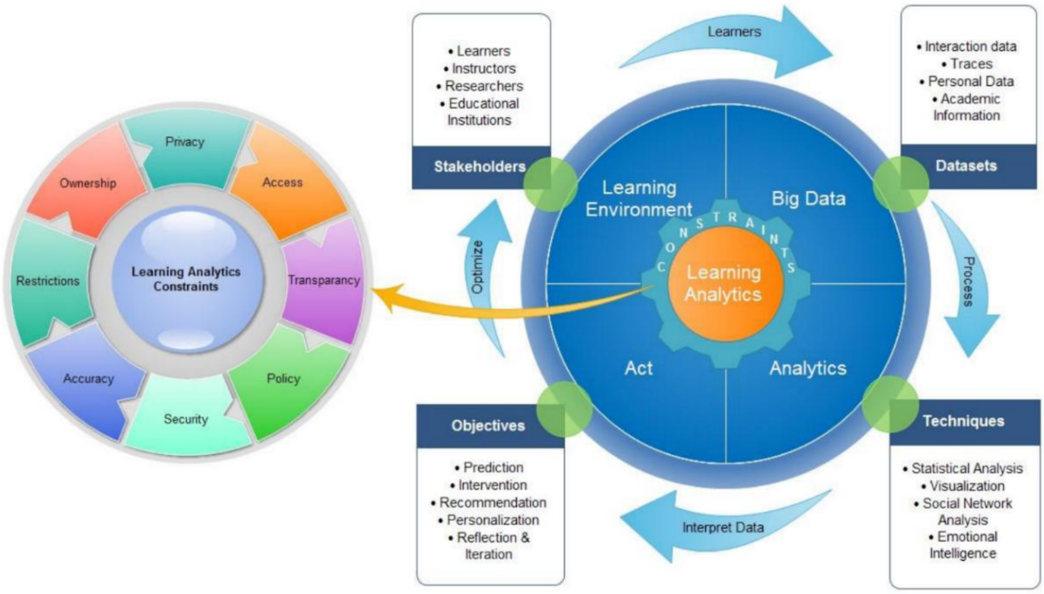

Figure 1. The Learning Analytics Framework Proposed by Khalil and Ebner (2022)

Ethical Guidelines Followed

1. The research ensured strict compliance with the data governance structure at the institution managed by the data protection office.
2. Clear guidelines on data collection, storage and usage were established. Ensuring that data was retrieved and conveniently stored on a system approved by the institution for analysis.
3. Transparency in the usage of the data across the learning analytics team was followed.
4. Informed consent was obtained from the student groups who participated in the pilot study both digitally and in classrooms, students were also aware of the purpose behind the research.
5. Measures were taken to prevent algorithmic bias by only including data generated from student learning for model development and not retrieving or including data related to demographics.
6. Latest guidelines and policies regarding data protection at the university were continuously reviewed and followed in alignment with the continuously evolving data protection norms and GDPR regulations.

Datasets Used for this Research

Datasets from three academic sessions with first-year students participating in a math module were used in this research study. The participating students were from the science and computing departments and gave their consent for the data to be used for this purpose.

Table 1 shows a summary of the datasets from the 3-year period with the number of participants in each year and aliases which will be used to refer to the datasets and the science and computing departments involved. Log file data for all datasets were also extracted. The log files data contained time series information on the interactions the students had with different components of their courses. It was used to get information on how engaged the students were on the course from a time perspective. Data was extracted twice for the MC2, MS2 at week 5 of the course and at week 12. Data was also extracted twice- at week 7 and week 12 of the semester for MC3 and MS3. It was only extracted once for the MC1, MC2 datasets as the course had already been completed at the beginning of the analysis.

Table 1. Summary of Datasets from the Courses Used for This Study

Year	Module	Alias	Number of Participants (n)
2020 - 2021	Essential Maths for Computing	MC1	106
2020 - 2021	Mathematics 1.1 for science	MS1	202
2021-2022	Essential Maths for Computing	MC2	132
2021-2022	Mathematics 1.1 for science	MS2	241
2022 – 2023	Essential Maths for Computing	MC3	122
2022-2023	Mathematics 1.1 for science	MS3	214
	Total		1017

Datasets Description

The datasets used in this research were extracted from the Moodle VLE in a structured form. These ranged between 71 to 103 features each. The rows represented feature observations for individual students. In all the datasets, there are five preceding features with only basic student details: name (first and last name), ID number, email address, and department. Those features were only used for identification purposes that facilitated the automated intervention and were not included in the analytics or modelling. The remaining features are described in Table 2:

Table 2. Summary Description of Features in Datasets

Feature type	**Feature Description**
Quiz	This represent quiz performance of the student across different quizzes and stored as a numeric value.
Attendance	Captures the students availability and participation during either their general lecture or journal classes.

Feature type	Feature Description
Journal	Represent students journal performance. Journals are hands on classes where students are guided to complete work for each week related to topic taught during the week.
Grades	Represents the overall grade of the student as reflected by an aggregation of performance across different topics.
Assignments	Represents dataset features that shows assignment completion rates for students in group.
Workshops	Represents scores from workshop (group) activities assigned to students.
Exams	Represent exam scores for the various exams taken by the students at different times in the semester.
Checklists	Represents the task completion checklist progress for course.
Interactions	Captured from the log files, these give an aggregated count of the number of interactions the student made on the learner platform in the form of clicks.
Average interactions per session	Represent the aggregated number of interactions per session when students logged on.It considers only the days when the student logged into to the course.
Average interactions per day	The average number of interactions per day for each student which considers all the days in the semester including those when there was no log in activity.
Days between interactions	The average time in days between clicks aggregated from the time stamps in the log files.

Descriptive Statistics

To gain insight into the relevant features in the cleaned data, descriptive statistics about the features in the datasets were performed. Table 3 shows some descriptive statistics for individual features selected from the activity types for the MC1 dataset.

Table 3. Descriptive Statistics of Selected Features in Dataset MC1

	Attendance	Checklist	Quiz Total	Workshop	Journal Total	Average Grade	Total Interactions
Count	106	106	106	106	106	106	106
Mean	69.6	84.1	70.0	72.7	69.5	87.8	1217.3
Standard deviation	32.6	30.2	32.7	42.6	36.6	30.5	526.7
Min	0	0	0	0	0	0	15

Similar values were obtained for the MC2 and MC3 datasets within computing groups. For the science group, Table 4 shows a sample of descriptive statistics for MS2. Statistical descriptions of the selected features for MS2 were also similar in the MS1 and MS3 computing groups. Descriptive statistics for all datasets were computed using the Python Describe function which computes the mean, median, standard deviation, and the maximum and minimum values across all selected features.

Table 4. Descriptive Statistics of Selected Features in Dataset MS2 (n=241)

	Attendance	Checklist	Journal Total	Quiz Total	Maths Grade	Days Between Interactions	Total Interactions
Count	241	241	241	241	241	241	241
Mean	54.2	58.1	9.9	6.7	43.7	4.0	1059.0
Standard deviation	26.1	35.9	6.8	5.4	28.1	4.2	684.6
Min	0	0	0	0	0	1.4	3

Data Extraction and Analysis

Data for model development were manually extracted from Moodle and application programming interfaces (APIs). APIs act as a bridge for various programs and applications to communicate and exchange information. The manual method of extraction involved logging into the Moodle platform with admin privileges for educators and downloading the gradebook and log file datasets. APIs were developed for an easier transfer of information from Moodle to Jupiter Notebook environment without requiring manual extraction. An API was shared with the open-source community to assist fellow researchers with extracting data from the Moodle database. Subsequently, the data was processed and cleaned. Log-file data was merged with course data to explore interaction relationships. This was achieved using software development tools, such as Power BI, the Python Merge function, and Excel. Figure 2 shows the distribution of log interactions across days of the week. The results showed the least interactions during the weekends and maximum interactions on Monday for MC1, MC2, and Wednesday for the rest of the datasets.

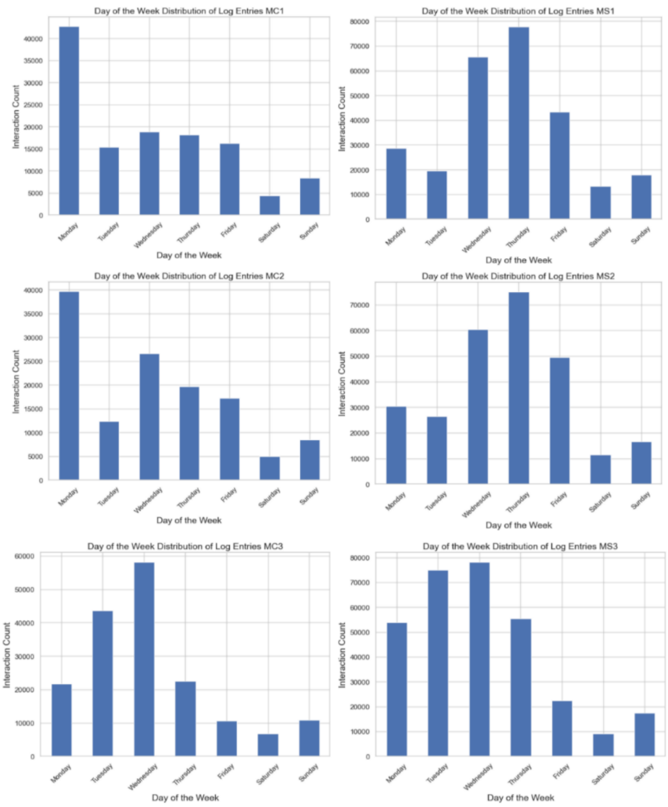

Figure 2. Histogram of Log Interaction by Day of the Week

Figure 3 shows scatterplots of attendance against grades across all year groups. The correlation coefficient “r” across all groups ranged from 0.64 and 0.84.

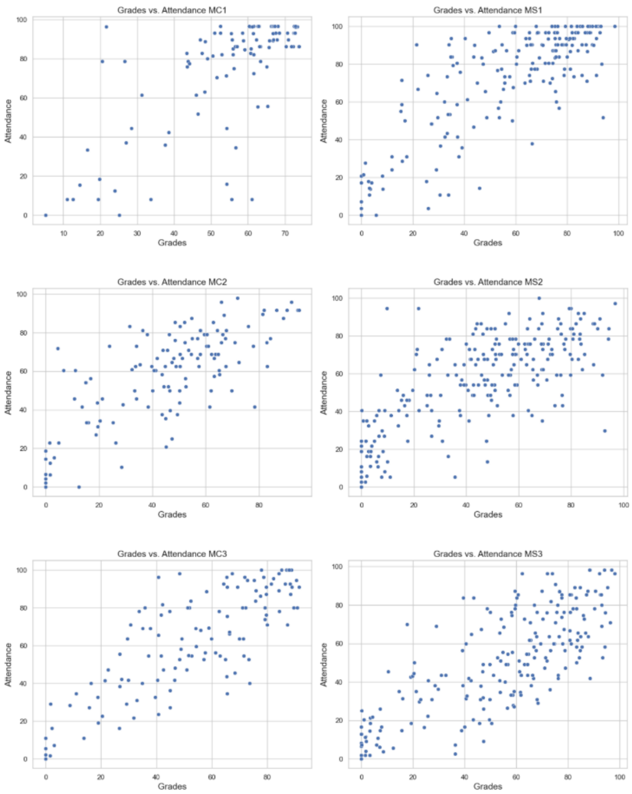

Figure 3. Scatterplot of Attendance against Final Grade across Groups

Correlation maps were also used to evaluate the relationships between the selected features, as shown in Figure 4. The correlations are shown in fractions up to 1. Strong positive correlation coefficients (>0.5) can be observed between interactions and all other features. As well as between Math grade and other features.

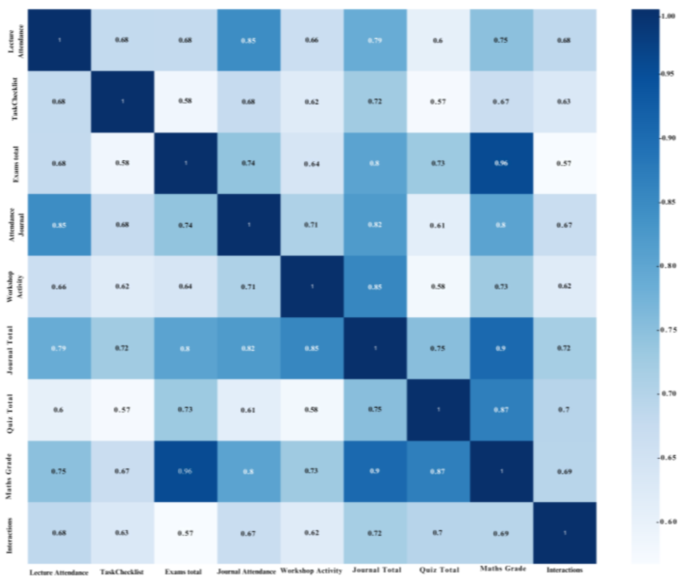

Figure 4. Correlation Heatmap of Selected Features for MS2.

Machine Learning Model Development and Approach

Once relevant features were identified and filtered based on their respective MI Scores (machine learning feature selection technique described later), machine learning models were evaluated. To determine the predictive model, data from the 2020/2021 session were analysed, comprising 205 science students in MS1 and 106 computing students MC1. The results of the analysis informed the machine learning model selection which was subsequently validated on the 2021/2022 data, consisting of 241 math students and 132 science students. The 2021/2022 dataset was used to develop an automated feedback system for students. The 2022/2023 dataset, which included data for 214 math students and 122 science students, was employed to analyse student feedback to better understand student motivations and enhance student support services.

Predictive models using unsupervised learning were used to forecast student outcomes. The model was not provided with a target feature or any explicit guidance on what should be predicted. Instead, it analysed the dataset and identified inherent patterns and structures. K-means clustering was the chosen model for this study after evaluation with two other models._ K-means is an iterative algorithm that partitions a dataset into group of clusters denoted as "k". The goal is to minimise the cluster variance within the identified cluster groups, which refers to the sum of the squared distances between data points and centroids for each cluster. A centroid is the geometric centre of a cluster which represents the mean location of all the data points within that cluster. The mathematics of the k-means model are denoted below as:

- k is the number of clusters.
- where n denotes the number of data points.
- x_i refers to the ith data point.
- c_j is the centroid of cluster j.
-

The goal was to minimise the objective function denoted as J which is the minimum sum of the squared Euclidean distances between data points and their corresponding cluster centroids.

$$J = \sum_{i=1}^{n} min_j \left|x_i - c_j\right|^2$$

$\left|x_i - c_j\right|^2$ represents the Euclidean distance between a given data point x_i and c_j which is the centroid of a given cluster j.

The clustering algorithm randomly assigns and initialises a specified number of clusters "k" based on input, then assigns each data point from the dataset to the nearest centroid of each of the clusters. It continues this assignment process and updates the position of the initial clusters until it reaches a point of convergence at which J does not improve with further centroid position update.

The clustering models were trained on the features in each dataset that met the MI score threshold. Subsequently, models were evaluated over different cluster numbers using silhouette scores which were used to calculate the goodness of fit of the clustering model (Shutaywi & Kachouie, 2021). Its values ranged from -1 to 1, where 1 represents perfect clustering model and -1 represents poor model.

Mathematically, the silhouette score is calculated as follows:

$$s(i) = \frac{b(i) - a(i)}{\max\big(a(i), b(i)\big)}$$

Where:

- $s(i)$ represents the Silhouette Score for data point i
- $a(i)$: The average distance of data point i to all other data points within the same cluster representing the cohesion within the cluster.
- $b(i)$: The average distance of data point i to all data points in a different cluster, where i does not belong representing separation from other data points.

The k-means model was selected based on its superior silhouette scores to generate cluster predictions on MC1 and MS1 for prediction evaluation.

Although the k-means model is an unsupervised learning technique with no labels assigned to a response feature, it can be evaluated using known labels which were excluded when training (process of teaching a computer to make predictions on given data) the model in a cluster validation process. This was done to evaluate the effectiveness of k-means in identifying groups of students in its predicted clusters and was treated as a supervised classification problem. This technique of "cluster validation" has been successfully used in other works (Aziz et al., 2021).

The actual labels which were the "Pass/Fail" feature created in the data phase were now treated as the ground truth and compared against the predicted clusters from the k-means model. The recall, precision and overall accuracy evaluation metrics were used.

Explanation of Metrics

The metrics were calculated using the model classifications of true positives (TP), true negatives (TN), False Positives (FP), and False Negatives (FN).

The task of assessing the ability of a model to classify whether a patient has a disease is used to explain each classification type.

A *True Positive (TP)* indicates that the predictive model predicted the patient as sick, and that the patient was sick. In this case, the illness was classified as positive. Therefore, true positives are correctly classified positive predictions.

A *True Negative (TN)* represents a correct negative prediction. If a patient is not sick, and the model predicts this correctly, it is a true negative. The negative prediction was true.

A *False Positive (FP***)** represents a prediction of a positive case (sick), which is false. Therefore, if someone were not sick but predicted to be sick, then it is false positive.

Finally, a *False Negative (FN)* represents a prediction of a negative case (not sick), which is false. Therefore, if someone is sick and the model predicts that the person is healthy, this prediction is false, resulting in a false negative prediction.

Metrics derived from the classification types are as follows:

Recall: Recall is a performance metric that evaluates a model's ability to correctly identify all relevant instances of a given class in an output category. In disease detection, recall measures the accuracy of a model in identifying all sick patients in each dataset. Recall does not consider false positives, as its primary objective is to correctly identify all sick patients. This prioritisation comes from the notion that accurately diagnosing as many sick patients as possible is crucial, even if it erroneously leads to some false positives where healthy patients are classified as sick. Recall is better metric in such prioritisation as the risk of missing a sick patient is far greater than that of misclassifying a healthy patient as sick. This metric is particularly important when the goal is not to miss any instance of the case. Recall was the main objective of this research project. It was more important to identify as many students needing support as possible. The risk of missing any students outweighed offering more support to a few misclassified students who may not have needed it.

The formula for recall is as follows:

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{All\ actual\ Positive}$$

Precision: Precision evaluates a classification model's positive predictions. It evaluates the correctness of all positive predicted cases by the model. This metric is concerned with correctness of positive predicted cases and does not prioritise all positive cases .In the context of disease detection, it measures the accuracy of correctly classified sick patients compared with overall number of patients classified as sick. If there is a higher risk of misclassifying patients as sick which leads to unnecessary treatment or procedures, then precision is a more relevant metric.

If a model predicts a patient as positive for a disease that would require surgery, it is expected to be as precise as possible in classifying cases as positive. It is important to identify all truly sick patients, the risk of falsely classifying a healthy patient as sick and subjecting them to unnecessary treatment, anxiety or surgery is significantly higher than the risk of missing a sick patient. Therefore, prioritising precision ensures that the model is used responsibly and minimizes the potential for harm to patients especially the cost implications.

The formula for the precision is as follows:

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{All\ predicted\ Positive}$$

Overall Accuracy: This evaluates the overall correctness of a classification model in identifying correct classes as compared to actual cases, so it considers all true predictions (positives and negatives) and compares it with the overall predictions. It measures how well a model predicts the different classes for a given dataset. If greater priority is not given to a predicted class, the accuracy is a better metric. So based on the disease detection context, if there was no emphasis on either predicted class or no real danger misclassifying for any class, accuracy would be a preferred metric.

The formula for calculating the accuracy is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{Overall\ correct\ predictions}{All\ Predictions}$$

Based on the 2020/2021 and 2021/2022 results metrics three clusters for the predictive model were chosen:

Machine Learning Model Results

The k-means, agglomerative and spectral clustering models were evaluated to identify different student groups in the data. The silhouette scores were used to evaluate effective cluster identifiers. The unsupervised learning models were evaluated across cluster values, k, from 2 and 10 for both the MC1 and MS1 datasets at week 12 (semester end) as shown in figure 5.

Figure 5, the k-means model performs best in 9 out of 10 tests, followed by the Agglomerative clustering model which was better in 1 case for both MC1 and MS1.

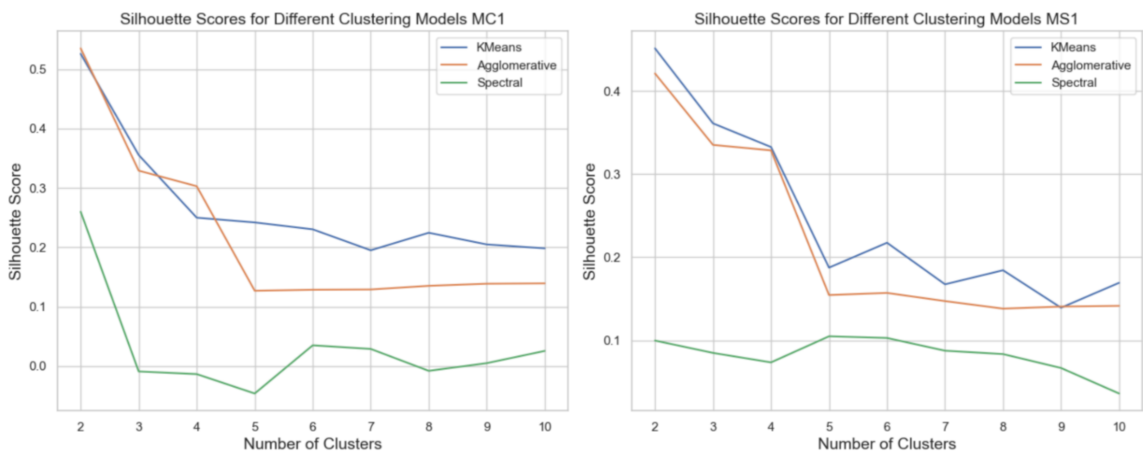

Figure 5. Silhouette Score Graphs Across All Models Tested

Results on Optimum Number of Student Clusters

An Elbow Plot was first used to investigate the computing cluster numbers using values from 1 to 10 to determine the optimum choice. Using this method, the point at which the rate of decrease for the sum of the squared distances changes significantly was selected as elbow. The sum of the squared distances against the number of clusters was computed. Figure 6 was obtained using the elbow method. It showed an unclear choice of elbow point for cluster values 2 and 3. Therefore, additional tests were conducted to determine the suitability of each cluster.

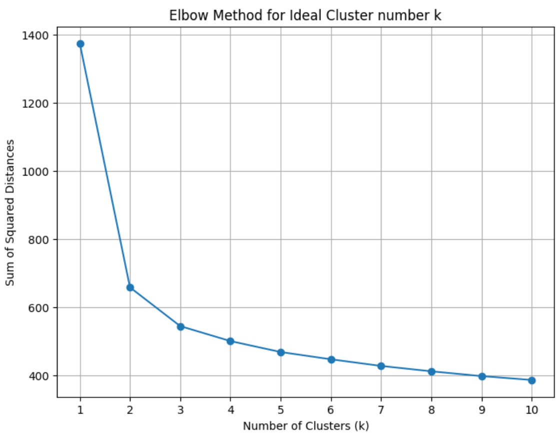

Figure 6. Elbow Method used to Determine Optimum Number of Clusters.

The elbow visualiser was applied using a distortion score and silhouette score. The distortion score calculates the sum of squared distances between each data point and its closest centroid within a cluster, while the silhouette score evaluates how similar a data point is to its own cluster compared to other clusters as shown in Figure 7. According to the "distortion score" metric, 3 clusters was the optimal number of clusters for the k-means model to identify useful groups. This was explored using the technique of computing silhouette scores for different cluster numbers from 2 to 10. The silhouette scores below showed that value 2 gave the highest score of 0.524.

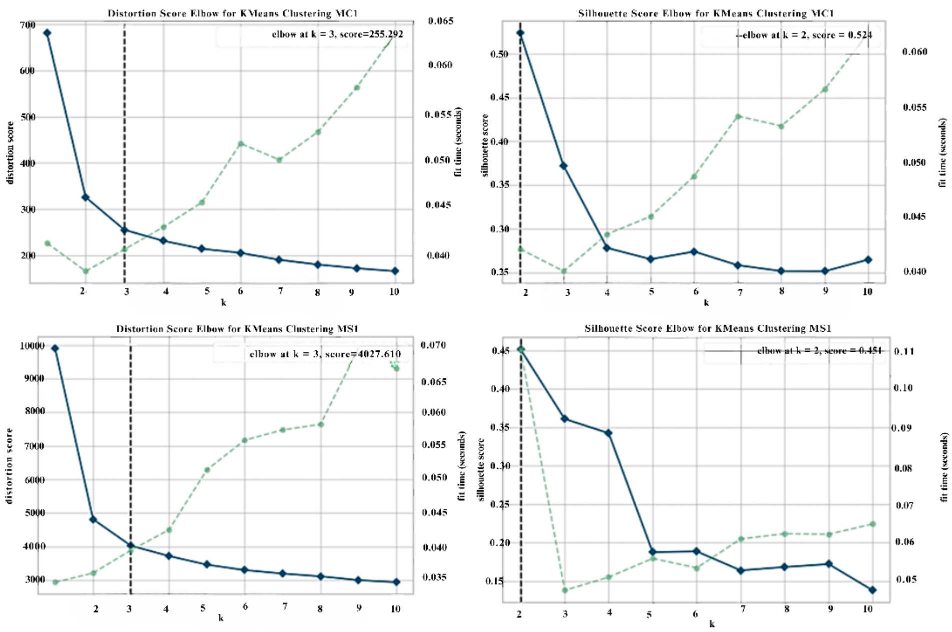

Figure 7. Cluster Analysis

Because k=3 is the optimum value according to the distortion method and k=2 is shown as ideal using the silhouette score metric, both values were used for cluster predictions, and groups were compared against the actual final performance.

Results from Predictive Models to Identify Student Groups

The k-means model and cluster numbers chosen were implemented on the 2020/2021 datasets MC1 (computing students) and MS1 (science) students to gauge the effectiveness of the model in identifying students who were not successful in the course by failing the course or dropping out. To determine this, cluster groups identified by k-means using 2 clusters and 3 clusters were compared to the performance of the students during training. The model was also evaluated for week 5 and week 12 dataset for MC2 and MS2. Finally, it was evaluated for week 7 and week 12 data for MC3 and MS3.

Predictions from the 2-cluster groups were compared with final student pass/fail category. The 2-cluster k-means identified two groups that closely matched with the pass/fail group, whereas the 3-cluster k-means identified 3 groups; one closely matching the pass group, one group that was either way, and one group closely matching the fail group. Additional observations of the group which could go either way, showed average math grade values indicating that this group was likely to either pass or fail the course.

To calculate these metrics for the 3-cluster k-means, the likely at risk and at-risk clusters were combined for evaluation against the actual pass/fail feature which is ground truth. The following predictions for true positives, true negatives, false positives, and false negatives for 3 clusters across all datasets are given in Table 5:

Table 5. K-Means Predictions Using 3 Clusters

Dataset/Period	**True Positive (TP)**	**True Negative (TN)**	**False Positive (FP)**	**False Negative (FN)**
MC1	23	70	12	1
MS1	61	118	23	0
MC2(Week 5)	40	64	27	1
MC2(Week 12)	41	69	22	0
MS2(Week 5)	95	110	33	6
MS2(Week 12)	95	116	27	3
MC3(Week 7)	34	26	60	2
MC3(Week 12)	36	59	27	0
MS3(Week 7)	41	118	38	17
MS3(Week 12)	58	122	34	0

Table 6 shows values for TP, TN, FP, and FN for MC1, MS1, MC2 and MS2 for weeks 5 and 12 using cluster value of 2. K-means predictions for 2-clusters for MS3 and MC3 were not computed as the previous datasets showed cluster value of 3 was better at recall.

Table 6. K-means Predictions Using 2 Clusters

Dataset/Period	True Positive (TP)	True Negative (TN)	False Positive (FP)	False Negative (FN)
MC1	18	81	1	6
MS1	60	129	12	1
MC2(Week 5)	30	88	3	11
MC2(Week 12)	39	89	2	2
MS2(Week 5)	80	134	9	18
MS2(Week 12)	83	134	9	15

Figure 7 visualises the three clusters identified by k-means across three axes related to grade (course), Journal and Quiz totals for different MC1 and MS1, MC2, MS2, MC3, and MS3. In the clusters shown, blue represents the group (not at risk) most likely to progress through the course, the green cluster represents the group (likely at risk) which could go either way as regards the curriculum with indicators around average level. The red cluster group (at risk) are those most likely to not progress through the course based on indicators.

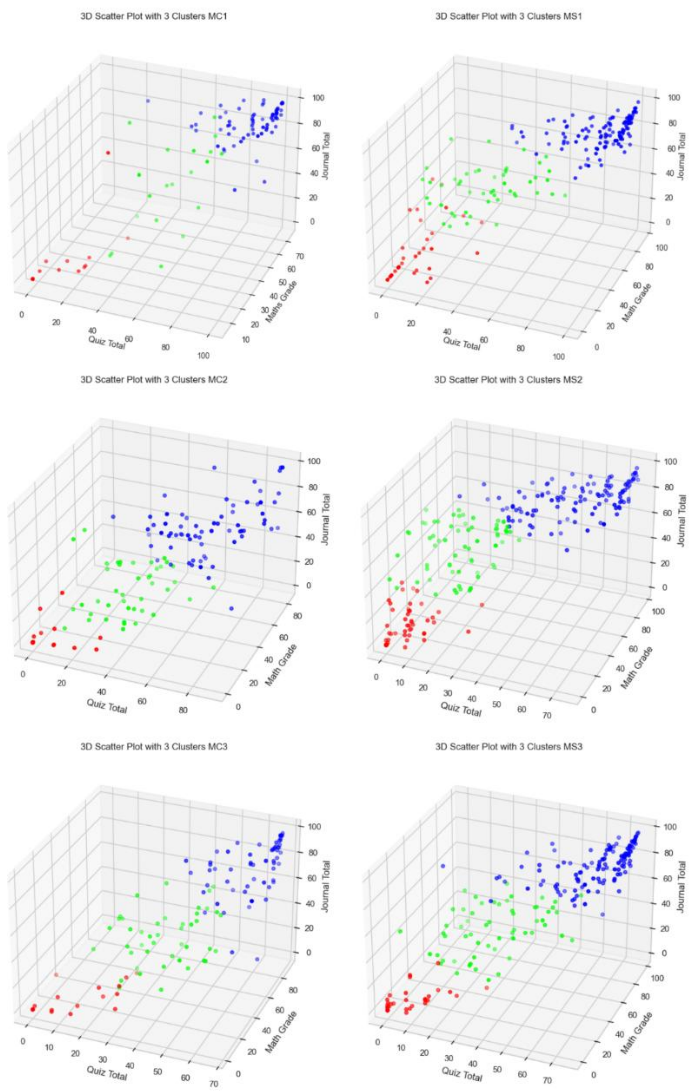

Figure 7. Three Clusters Analysis of Patterns Across Datasets

From the analysis of cluster predictions and comparisons shown in the following tables with cluster results, those formed at the end of the semester were more accurate in identifying students who were at risk of failing the course than those formed early on implying that with more data coming through over the weeks of the semester, the model gets better.

The results from Table 7 validated on the MS2 and MC2 datasets showed that a cluster value of 3 was better at identifying more students who were at risk of not being successful with the course as early as week 5. The same process of cluster analysis for MS1 and MC1 applied to MS2 and MC2 datasets for validation, showed that 3 clusters were better at identifying students who were at risk of not passing the course as early as week 5. This was also applied to MS3 and MC3 with 3 clusters still better.

A table showing the results of Recall, Precision and Accuracy scores when the predicted clusters were compared with the actual student categories from k-means cluster validations for MS1 and MC1 is presented in Table 7:

Table 7. Cluster Validation Evaluation Results for MS1, MC1 for 2 and 3 Clusters.

Dataset (Week 12)	Recall %	Precision %	Accuracy %	Number of Clusters
Science 20/21 MS1	98.36%	83.33%	93.56%	2
Science 20/21 MS1	100%	72.62%	88.61%	3
Computing 20/21 MC1	75.00%	94.74%	93.40%	2
Computing 20/21 MC1	95.83%	65.71%	87.74%	3

Exploring Evaluation Metrics for K-means

Based on the results in Table 7 from the 2020/2021 datasets, it is evident that the precision and overall accuracy decrease when the number of clusters is increased from two to three, but the recall increases for both science and computing datasets. Because of the greater importance of identifying at-risk students, a cluster value of 3 is preferable. This choice of cluster number and model is validated on the 2021/2022 datasets (MS2 and MC2), and data were extracted at two separate times, at the fifth week when new students were mostly settled and familiar with their learning environment, and data were also collected at the end of the semester (Week 12). The k-means model, using two and three clusters, was implemented for data extracted at week 5 and the final week before exams. Predictions were made for the datasets, and when the overall results of the students were obtained, these predictions were compared with the actual results of the students. Table 8 shows the results from the MC2 and MS2 datasets captured at week 5 for two and three clusters.

Table 8. Week 5 Model Prediction Evaluated across Different Metrics

Dataset (Week 5)	Recall	Precision	Accuracy	Number of Clusters
Computing 21/22 (MC2)	73.17%	90.91%	89.39%	2
Computing 21/22 (MC2)	97.56%	59.70%	78.79%	3
Science 21/22 (MS2)	81.63%	89.89%	88.80%	2
Science 21/22 (MS2)	93.88%	73.60%	83.82%	3

As seen from Table 8, the pattern of an improvement in recall with an increase from two to three clusters is consistent with previous patterns while precision and accuracy drop for both groups.

Table 9 shows metric results obtained for MC2 and MS2 when week 12 predictions were compared with final student outcome of pass/fail:

Table 9. Model Evaluation for MC2 & MS2 at Term End

Dataset (week 12)	Recall	Precision	Accuracy	Number of Clusters
Computing 21/22 (MC2)	95.12%	95.12%	96.97%	2
Computing 21/22 (MC2)	100.00%	65.08%	83.33%	3
Science 21/22 (MS2)	84.69%	90.22%	90.04%	2
Science 21/22 (MS2)	96.94%	77.87%	87.56%	3

The results show better recall values using the 3 clusters approach for predictions made using k-means at week 5 and the end of the semester. The results also show that the performance metrics improve across recall, precision, and accuracy when predictions are made using datasets at the end of the semester. A final model evaluation test was also carried out on MS2 and MC2 using only a cluster value of 3 and time periods at week 7 and week 12. Following this approach, the results shown in Table 10 were obtained. Minimum recall percentage of 70.69% was achieved in week 7.

Table 10. Model Evaluation on MC3 and MS3 for Three Clusters at Week 7

Dataset	Recall	Precision	Accuracy	Number of Clusters
Week 7 (MS3)	70.69%	51.90%	74.30%	3
Week 12 (MS3)	100.00%	63.04%	84.11%	3
Week 7 (MC3)	94.44%	36.17%	49.18%	3
Week 12 (MC3)	100.00%	57.14%	77.87%	3

At week 12, the comparison between predicted results and actual student performance using the ground truth and evaluation metrics, reaffirms that a cluster value of 3 provides a better recall at the compromise of accuracy and precision. The average accuracy across MC2, MS2, MC3 and MS3 early on (week 5 &week 7) Considering the early accuracy predictions for MC2, MC3, MS2 and MS3 was 71.52% with an average recall of 89.14%. The resulting clusters from the k-means algorithm were categorized into three groups: "at risk," "likely at risk," and "not at risk,". A combination of the "at risk" and "likely at risk" clusters named based on inspection of final course performances, was better at identifying all cases of students at risk early on.

Confusion Matrices were used to visualise the cluster validations based on True Positives, True Negatives, False Positives and False Negative values which are easily seen via the heatmap, as shown in the MC3 and MS3 visuals in Figure 8. Figure 8 shows comparison of predicted cluster categories with actual outcomes. The number of students in each box correspond to the True Positives, True Negatives, False Positives and False Negative predictions from the model upon which the metrics were calculated for MC3 and MS3.

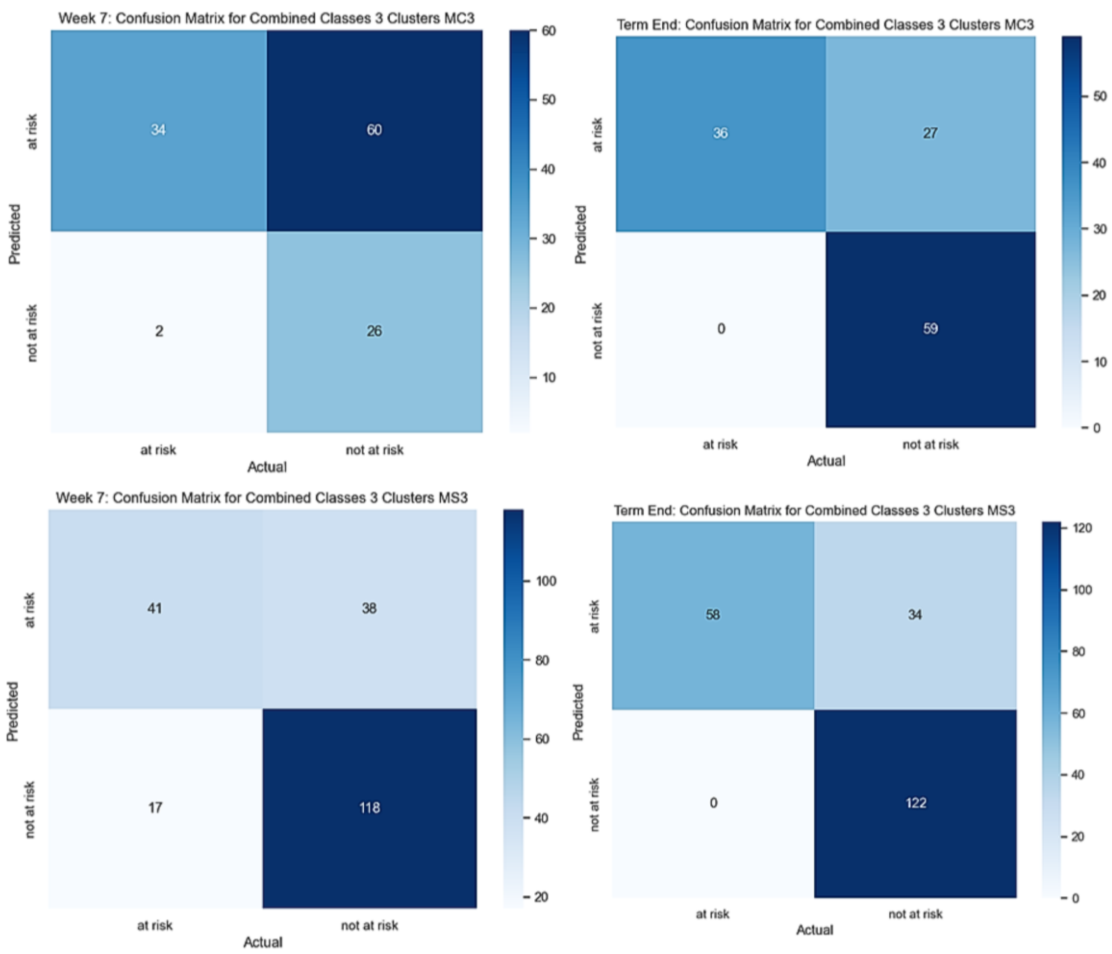

Figure 8. Heatmap Showing Confusion Matrices for MC3 and MS3

Conclusions

Can historical data (three years) uncover patterns and insights between successful and unsuccessful students based on their interactions and learning behaviour?

Historical data analysis effectively uncovered distinct behavioural patterns between successful and unsuccessful students, which led to a better understanding of student behaviour. There were also key indicators of student success found in data analysis. Analysis of features showed that higher interaction frequency, shorter gaps between interactions, and better attendance consistently correlated with better academic performance. Potential times for targeted interventions were also uncovered from patterns in student engagement. Monitoring these indicators on an ongoing basis would be a valuable tool for identifying struggling students early leading to more proactive retention strategy and learner support from the institution.

Can machine learning models predict risk/struggling students with good accuracy? What are the early indicators that predict students' success or disengagement?

Unsupervised machine learning models evaluated, demonstrated promising results in predicting students' success and risk of disengagement. The results indicated that the models, particularly the k-means clustering algorithm, showcased significant potential in identifying at-risk students early in the academic term.

Features such as quiz activity, attendance, journal activity, and interaction frequency emerged as critical indicators of student success. Experimenting with cluster values of 2 and 3 showed reliable identification of at-risk students with early recall accuracy exceeding 70% as early as week 5. These findings highlight the usefulness of machine learning models for proactive student support and intervention based on the early indicators. Information from these can assist educational institutions to best cater for student needs.

How effective is automated and personalised feedback as an intervention method?

This research demonstrated the efficacy of automated and personalised feedback as an intervention method. The overwhelming positive perception from students on receiving this form of feedback and their preference for receiving it via email highlights the need for building such systems at scale. 93% of students (n=365) found the feedback style very useful with comments from student showing great positive perception and crucial in fostering positive learning experiences and engagement.

Recommendations

Based on the promising results from this study, it is recommended that the areas of Natural Language Processing (NLP), Large Language Models (LLMs), and AI Agents be further evaluated to determine their impact on Educational Learning Analytics. As technology in the field of Artificial Intelligence continues to develop rapidly, there is a need for extensive research to understand how these technologies affect learning analytics.

Declaration of Conflicting Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. Please insert relevant information here

Acknowledgement/Funding

I would like to acknowledge my supervisors for their expert guidance and support throughout this research: Dr. Etain Kiely (principal supervisor), Dr. Cormac Quigley, Donal McGinty, Dr. Konrad Mulrennan, and Dr. John Donovan. I also gratefully acknowledge the RISE scholarship funding from ATU and thank Dr. Vida Awinime Bugri for her assistance with proofreading and reviewing the paper prior to submission.

References

Adiguzel, T., Kamit, T., & Ertas, B. (2020). Teaching and learning experiences with enhanced books in engineering math and science courses. *Contemporary Educational Technology*, *11*(2), 143–158. https://doi.org/10.30935/cet.660725

Anthony, W. et al. (2021). Adopting Course Completion Tracking and Conditional Activities to Enhance Engagement in eLearning for University Students. *IST-Africa Conference (IST-Africa)*. ISBN: 978-1-905824-66-3

Attard, C., & Holmes, K. (2022). An exploration of teacher and student perceptions of blended learning in four secondary mathematics classrooms. *Mathematics Education Research Journal*, *34*(4), 719–740. https://doi.org/10.1007/s13394-020-00359-2

Aziz, S. R., Khan, T. A., & Nadeem, A. (2021). Inheritance metrics feats in unsupervised learning to classify unlabeled datasets and clusters in fault prediction. *PeerJ Computer Science*, *7*. https://doi.org/10.7717/PEERJ-CS.722

Banihashem, S. K., Aliabadi, K., Pourroostaei Ardakani, S., Delaver, A., & Nili Ahmadabadi, M. (2018). Learning Analytics: A Systematic Literature Review. *Interdisciplinary Journal of Virtual Learning in Medical Sciences*, *9*(2). https://doi.org/10.5812/ijvlms.63024

Baolin, Y. et al. ,. (2017). Learning Analytics-Based Evaluation Mode for Blended Learning and Its Applications. *International Symposium on Educational Technology (ISET)*. https://dx.doi.org/10.1109/ISET.2017.42

Boulton, C. A., Hughes, E., Kent, C., Smith, J. R., & Williams, H. T. P. (2019). Student engagement and wellbeing over time at a higher education institution. *PLOS ONE*, *14*(11), e0225770. https://doi.org/10.1371/journal.pone.0225770

Clow, D. (2012). The Learning Analytics Cycle: Closing the Loop Effectively. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 134–138. https://doi.org/10.1145/2330601.2330636

D. Pérez-Berenguer et al. (2016). A cross-platform interoperable component for course analytics. *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality. Salamanca, Spain, 3012622: ACM*, 887–892. https://doi.org/10.1145/3012430.3012622

Dass, S., Gary, K., & Cunningham, J. (2021). Predicting student dropout in self-paced mooc course using random forest model. *Information (Switzerland)*, *12*(11). https://doi.org/10.3390/info12110476

Daud, A., Lytras, M. D., Aljohani, N. R., Abbas, F., Abbasi, R. A., & Alowibdi, J. S. (2017). Predicting student performance using advanced learning analytics. *26th International World Wide Web Conference 2017, WWW 2017 Companion*, 415–421. https://doi.org/10.1145/3041021.3054164

Department of Further and Higher Education. (2022). *Funding the Future - Summary of the Department's Policy Response Funding higher education. Reducing the cost of education.* https://www.gov.ie/en/publication/28c9d-funding-the-future-an-annual-options-paper-on-reducing-the-cost-of-higher-education/

Dridi, S. (2021). *Unsupervised Learning-A Systematic Literature Review a Preprint*. https://doi.org/10.31219/osf.io/kpqr6

Embarak, O. (2020). Apply machine learning algorithms to predict at-risk students to admission period. *2020 7th International Conference on Information Technology Trends, ITT 2020*, 190–195. https://doi.org/10.1109/ITT51279.2020.9320878

Farag, A. A. (2020). Plenary Talk II: Measuring Student Engagement in Early Engineering Coursework. *2020 15th International Conference on Computer Engineering and Systems (ICCES)*, i–ii. https://doi.org/10.1109/ICCES51560.2020.9334628

Fidelia, O. et al. (2022). Using Machine Learning to Explore the Relation Between Student Engagement and Student Performance. *24th International Conference Information Visualisation (IV)*. http://dx.doi.org/10.1109/IV51561.2020.00083

Flores-Vivar, J.-M., & García-Peñalvo, F.-J. (2022). Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4). *Comunicar*, *31*(74). https://doi.org/10.3916/c74-2023-03

George S, & Phillip L. (2011). *LAK '11 : proceedings of the 1st International Conference on Learning Analytics and Knowledge, February 27 - March 1, 2011, Banff, Alberta, Canada*. https://dl.acm.org/doi/proceedings/10.1145/2090116

Gray, G. (2015). *Investigating the Efficacy of Algorithmic Student Modelling in Investigating the Efficacy of Algorithmic Student Modelling in Predicting Students at Risk of Failing in the Early Stages of Predicting Students at Risk of Failing in the Early Stages of*. https://arrow.tudublin.ie/ittthedoc

Hongxin, Y. (2020). Using Learning Analytics and Adaptive Formative Assessment to Support At-risk Students in Self-paced Online Learning. *IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*. http://dx.doi.org/10.1109/ICALT49669.2020.00125

Hu, M., Li, H., Deng, W., & Guan, H. (2016). Student Engagement: One of the Necessary Conditions for Online Learning. *2016 International Conference on Educational Innovation through Technology (EITT)*, 122–126. https://doi.org/10.1109/EITT.2016.31

Hu, Q., Huang, Y., & Deng, L. (2021). A framework to evaluate student engagement of programming course in blend learning environment. *2021 16th International Conference on Computer Science & Education (ICCSE)*, 915–919. https://doi.org/10.1109/ICCSE51940.2021.9569553

IBM. (2022). *What is Machine Learning?* Https://Www.Ibm.Com/Cloud/Learn/Machine-Learning.

Kaliisa, R., Kluge, A., & Mørch, A. I. (2022). Overcoming Challenges to the Adoption of Learning Analytics at the Practitioner Level: A Critical Analysis of 18 Learning Analytics Frameworks. *Scandinavian Journal of Educational Research*, *66*(3), 367–381. https://doi.org/10.1080/00313831.2020.1869082

Katerina, M. et al. (2019). Learning Analytics for Learning Design: A Systematic Literature Review of Analytics-Driven Design to Enhance Learning. *IEEE Transactions on Learning Technologies,* . http://dx.doi.org/10.1109/TLT.2018.2868673

Kew, S. N. et al. (2017). Identifying at-risk students in online learning by analysing learning behaviour: A systematic review. *IEEE Conference on Big Data and Analytics (ICBDA)*.

http://dx.doi.org/10.1109/ICBDAA.2017.8284117

Khalil, M., & Ebner, M. (2015). *Learning Analytics: Principles and Constraints*. https://doi.org/10.13140/RG.2.1.1733.2083

Khalil, M., Prinsloo, P., & Slade, S. (2022). A Comparison of Learning Analytics Frameworks: A Systematic Review. *ACM International Conference Proceeding Series*, 152–163. https://doi.org/10.1145/3506860.3506878

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M. D., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability (Switzerland)*, *13*(18). https://doi.org/10.3390/su131810424

Lee O'Farrell. (2017). *Using Learning Analytics to Support the Enhancement of Teaching and Learning in Higher*

Education Pre-Specified Nationally Co-Ordinated Project. https://dx.doi.org/10.6084/m9.figshare.5777580

Mark, T. W. (2021). Redesigning a First Year Physiology Course using Learning Analytics to Improve Student Performance. *International Journal of Learning Analytics and Artificial Intelligence in Education*, 4–19. http://dx.doi.org/10.3991/ijai.v3i1.21799

Mireilla, B. A. et al. (2017). The Potential of Learning Analytics in Understanding Students' Engagement with Their Assessment Feedback. *IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. http://dx.doi.org/10.1109/ICALT.2017.40

Mohamed Nafuri, A. F., Sani, N. S., Zainudin, N. F. A., Rahman, A. H. A., & Aliff, M. (2022). Clustering Analysis for Classifying Student Academic Performance in Higher Education. *Applied Sciences (Switzerland)*, *12*(19). https://doi.org/10.3390/app12199467

Mohd, I. H., Hussein, N., Aluwi, A. H., & Omar, M. K. (2016). Enhancing students engagement through blended learning satisfaction and lecturer support. *2016 IEEE 8th International Conference on Engineering Education (ICEED)*, 175–180. https://doi.org/10.1109/ICEED.2016.7856067

Nurfadhlina M. S. et al. (2021). *Learning Analytics of Online Instructional Design during COVID-19: Experience from Teaching Data Analytics Course*. https://doi.org/10.1109/ICADEIS52521.2021.9702058

O'Brien, C. (2022). *Student dropout rates climb across several universities amid concern over low engagement*. https://Www.Irishtimes.Com/Ireland/Education/2022/05/28/Student-Dropout-Rates-Climb-across-Many-Universities-amid-Concern-over-Low-Engagement

Oladipupo, T. (2010). Types of Machine Learning Algorithms. In *New Advances in Machine Learning*. InTech. https://doi.org/10.5772/9385

Omar, T., Alzahrani, A., & Zohdy, M. (2020). Clustering Approach for Analyzing the Student's Efficiency and Performance Based on Data. *Journal of Data Analysis and Information Processing*, *08*(03), 171–182. https://doi.org/10.4236/jdaip.2020.83010

Sghir, N., Adadi, A., & Lahmer, M. (2023). Recent advances in Predictive Learning Analytics: A decade systematic review (2012–2022). *Education and Information Technologies*, *28*(7), 8299–8333. https://doi.org/10.1007/s10639-022-11536-0

Shutaywi, M., & Kachouie, N. N. (2021). Silhouette analysis for performance evaluation in machine learning with applications to clustering. *Entropy*, *23*(6). https://doi.org/10.3390/e23060759

Tsai, Y. S., & Gasevic, D. (2017). Learning analytics in higher education - Challenges and policies: A review of eight learning analytics policies. *ACM International Conference Proceeding Series*, 233–242. https://doi.org/10.1145/3027385.3027400

Tsai, Y.-S., Moreno-Marcos, P. M., Jivet, I., Scheffel, M., Tammets, K., Kollom, K., & Gašević, D. (2018). The SHEILA Framework: Informing Institutional Strategies and Policy Processes of Learning Analytics. *Journal of Learning Analytics*, *5*(3). https://doi.org/10.18608/jla.2018.53.2

Tuti, P. et al. (2020). Data Analytics of Students' Profiles and Activities in a Full Online Learning Context. *Fifth International Conference on Informatics and Computing (ICIC)*. http://dx.doi.org/10.1109/ICIC50835.2020.9288540

Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*, *104*. https://doi.org/10.1016/j.chb.2019.106189

Yoong, C. H. (2014). Student engagement and learning using an Integrated Student-Lecturer Engagement Design Framework. *IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*. http://dx.doi.org/10.1109/TALE.2014.7

Author Information

Ikechukwu Ogbuchi

https://orcid.org/0000-0001-8752-7531

Atlantic Technological University

Ireland

Contact e-mail: *ikechukwu.ogbuchi@research.atu.ie*