Bayesian Model for Academic Performance Prediction in Learning Analytics

Siti Salwa Salleh[1,3], Yamin Yassin[2]

[1]College of Computing, Informatics and Mathematics, Kampus Seremban, UiTM Negeri Sembilan, [2]Faculty of Applied Sciences, Kampus Kuala Pilah, UiTM Negeri Sembilan, [3]Malaysia Institute of Transport (MITRANS), Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
Corresponding Authors Email: ssalwa@uitm.edu.my

Abstract
The learning analytics dashboard (LAD) enables the prediction, tracking, and early recommendation of actions based on academic performance, student conduct, cognitive abilities, and personality traits. The creation of a questionnaire, which involved expert validation and a pilot test, came before the data gathering. Students independent traits, such as behaviour, cognitive skill, and personality, and academic data are the contingent variables. The study flow comprises knowledge acquisition, data collection and analysis, implementation that executes experiments, and evaluation. An exploratory data analysis (EDA) has been performed to investigate the patterns and trends within the data collected on students' characteristics. Simultaneously, a Bayesian prediction model has been created and trained using the gathered data. The forecasts yielded an accuracy of 81%. Additionally, the F1 scores of 0.74 indicate a moderate level of ability, while scores of 0.90 suggest a high level of performance, and scores of 0.98 indicate an excellent level of performance. The interpretation of the F1 score is conducted within the specific realm of the issue and compared to manual calculations. A satisfactory level of performance is considered acceptable. Within the scope of this investigation, a sensitivity score of 0.8 is deemed favourable due to its ability to accurately identify the related risk of misclassification.
Keywords: Bayesian Model, Learning Analytics, Prediction, Students' Traits, Visualisation

Introduction
Lecturers need to monitor students' performance constantly, as it saves time in analysis and reduces the long line of communication (Aldowah *et al.*, 2019; Matzen *et al.*, 2017). However, the traditional classroom approach normally observes students' behaviour, cognitive strategies, and personalities without the use of any quantitative measurement tools or systematic monitoring. This has made it challenging for educators to make an early intervention, measure, and predict students' academic progress, particularly in courses with multiple ongoing evaluations that do not have a final exam, such as quizzes, assignments, and exams (Ong and Singh, 2021). On top of that, open distance and blended learning modes present more challenges. As machine learning technology has advanced, a few learning

analytics dashboards (LAD) have been employed to aid in student monitoring. There are tools available to predict students' performances, but they are generic and do not represent specific parameters of where the students are studying, and the usability of LADs in advisory and assistive tools has not been thoroughly studied(Muñoz-Merino, 2019). Moreover, most LAD focused on parameters based on academic-related data only. There is a limited study that includes and emphasizes the interplay of personality traits, cognitive strategies, and behaviours on student academic performance and embeds it in the LAD as their major prediction indicator.

In response to that, this study has been conducted to bridge the knowledge gap by conducting the following activities to meet these objectives: (i) collecting data, analysing it, and providing information on behaviour, cognitive abilities, personality traits, demographic characteristics, and academic performance; and (ii) implementing predictions using the Bayesian probability model for tracking and monitoring students' academic progress in a progressive manner in order to begin early intervention. The LAD will implement visualization strategies to facilitate actionable insights from descriptive, diagnostic, and predictive analytics findings (Ong and Singh, 2021). This is to avoid the visual saliency in some tools that makes the prediction look poor with low abstraction and would result in insufficient information and parameter relationships (Matzen *et al.,* 2017*)*. To present the work, this article is organized as follows: Section 1 provides an introduction, Section 2 includes a literature review and Section 3 covers the methods and materials used. Followed with Section 4 that presents the results and discussion; and finally, a conclusion is provided to wrap up the article.

Review of Literature

This section discusses relevant previous studies and related analytics dashboards. It comprises the following three (3) sections that focus on learning analytics and a review of a five-factor model of personality description and corresponding analytics dashboard. It aims to provide a fundamental understanding of the related area and prepares for a review of related work in a subsequent section.

Learning Analytics

Learning analytics is the application of machine learning algorithms in the measuring, analysing, and reporting of data about students and their contexts (Bodily & Verbert, 2017). It is used in education to inform instructional decision-making and improve student outcomes. Learning analytics involves the use of data from various sources, and there is no standardization in the way that data is collected, stored, and analysed, so it is dependent on the context of each study. Although many learning analytics have been deployed in recent years, there is still a lack of knowledge about predicting academic performance based on student behaviour (Bodily & Verbert, 2017). Learning analytics often involves the use of dashboards to display visuals of metrics related to student learning (Baker, 206). Examples of the display include the range of information included: (i) student performance, such as data on student grades, assessments, and course progress. (ii) Engagement data that covers student attendance, participation, and interaction with course materials (iii) Learning outcomes that include data on how well students are meeting learning objectives and achieving desired outcomes. Finally, (iv) predictive analytics that analyse student behavior and performance to predict future outcomes and identify areas where intervention may be needed. A learning analytics dashboard can be customized to meet the specific needs of a

particular institution or course (Khalil & Ebner, 2014). This is because its purpose is to provide a quick and easy way to access and analyse data that can be used to inform decision-making related to teaching, learning, and student support.

Five-Factor Model of Personality

The Five-Factor Model of Personality Research shows there is a link between personality traits and academic performance [5]. Personality traits like conscientiousness, openness, and emotional stability are positively associated with academic performance, whereas neuroticism and extraversion are negatively associated. Cognitive strategies such as metacognition and self-regulation, as well as behaviours such as attendance, study habits, and time management, can all have an impact on academic performance (Poropat, 2009). As a necessary consequence, this study suggests that these elements must be part of the ruling computation in the prediction so that the prediction is more in line with students' individual self-context.

Bayesian Model

In a Bayesian model, probabilities are used to show all the unknowns about the input and output parameters. The model's posterior predictive distribution generates simulated outcome values. It represents the distribution of future, unseen data given the observed data. This distribution can be utilized as a highly accurate predictor for forecasting future trends based on past data (Wulandari, 2020). One of the advantages of this approach is that it enables users to seamlessly integrate prior knowledge and data into a prediction framework. In addition, this model provides exact inferences that are based on the available information. The formula for Bayes' theorem is as follows:

$$P(A|B) = \frac{P(B|A)\,(P(A)}{P(B)} \qquad\qquad (1)$$

Where
- $P(A|B)$ is the probability of event A occurring, given event B has occurred.
- $P(B|A)$ is the probability of event B occurring, given event A has occurred, known as likelihood.
- $P(A)$ is the probability of event A, known as prior.
- $P(B)$ is the probability of event B, known as predictor probability.

In this study context, predictions are calculated based on student characteristics. Their performance in the coursework assessment and the teaching strategies employed in the course. It is noted from previous research that results from previous work showed that the Bayesian model outperformed the other models in terms of both accuracy and generalization ability (Alves *et al.*, 2021).

Related Research

Krueger *et al* [12] carried out research that is closely related to this study. They studied the current state of learning analytics, listed the ongoing challenges, and did a qualitative study to find out how the design features of LAD could be used to make users more engaged. The study pointed out important aspects that should be considered when designing LAD, although it has not been evaluated yet. A similar study was conducted by Carroll *et al* (2019), but it as equally limited in scope, as it needed to assess the efficiency of the design aspects. Koppelman

et al [14] suggested important design aspects should be incorporated in LAD for teachers, such as offering a clear overview of the data, allowing users to modify the dashboard, and providing relevant contextual information. It gives basic design concepts for constructing effective learning analytics dashboards but no specific instructions on how to create a dashboard tailored to a specific context. HIlliger *et al* (2021) constructed a learning analytics dashboard that provides lecturers with weekly response rates as well as the average number of weekly hours that students believe they spent on various subject activities. It consists of a web page with three major visuals. The results demonstrate that users regard the dashboard as a valuable tool for monitoring academic effort throughout the academic year. However, the dashboard needs to incorporate additional features to enable staff to monitor students' activity online rather than on weekly timesheets. Millecamp *et al* (2018) provides an assessment of a learning dashboard with limited analytics that supports the discussion between a student and a study adviser. It visualizes the student's grades, a summary of his or her progress during the year, his or her position in relation to peers, sliders to plan the next years, and a prognosis of the length of the bachelor's degree for this student in years based on historical data. The results show that the dashboard largely prompts factual, interpretive, and reflective insights but there is no prediction embedded in the LAD.

Methods and Materials

This study has four parts: making instruments and collecting data, analysing the data, putting the plans into action, and evaluating the results. Figure 1 illustrates the activities involved.
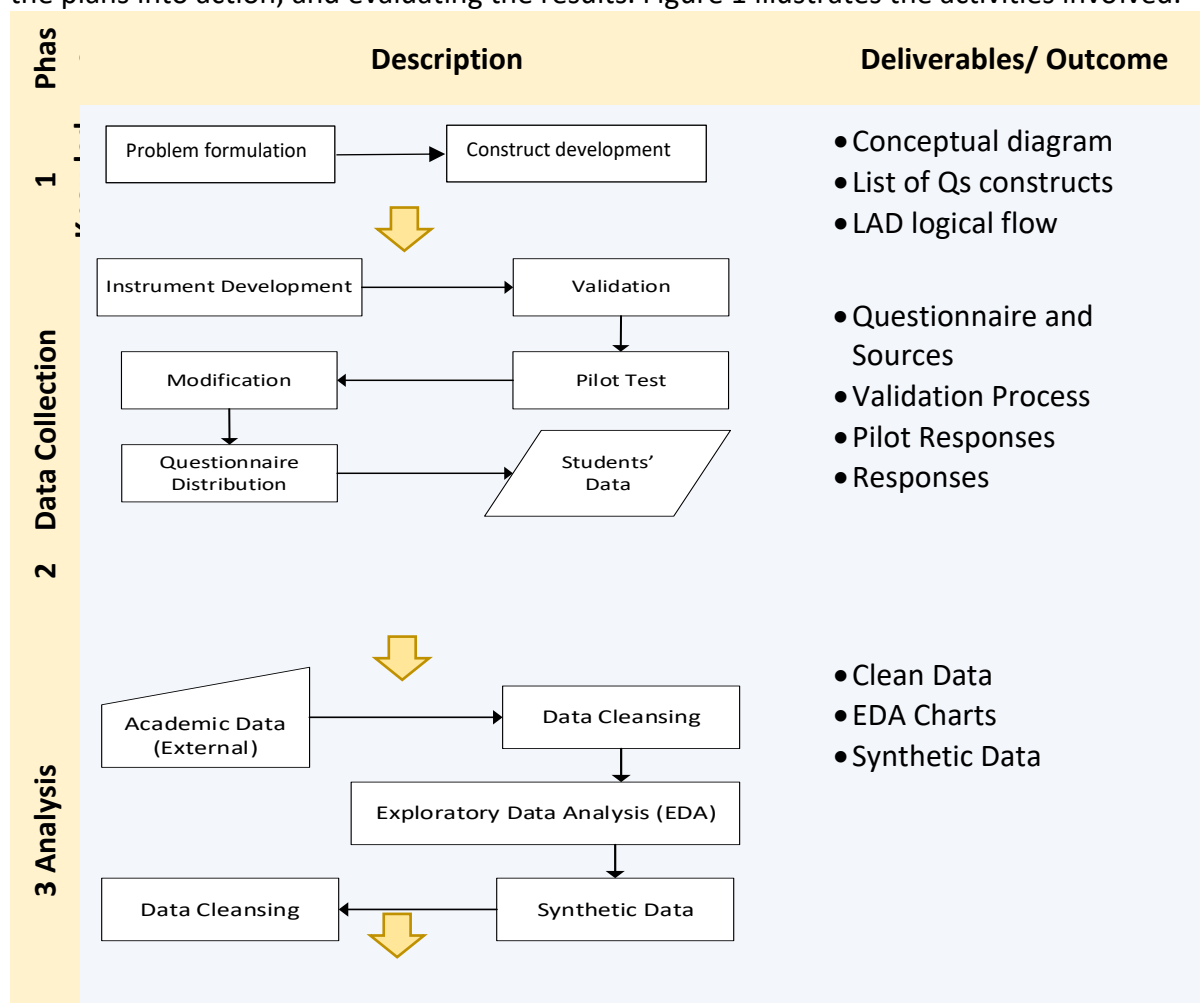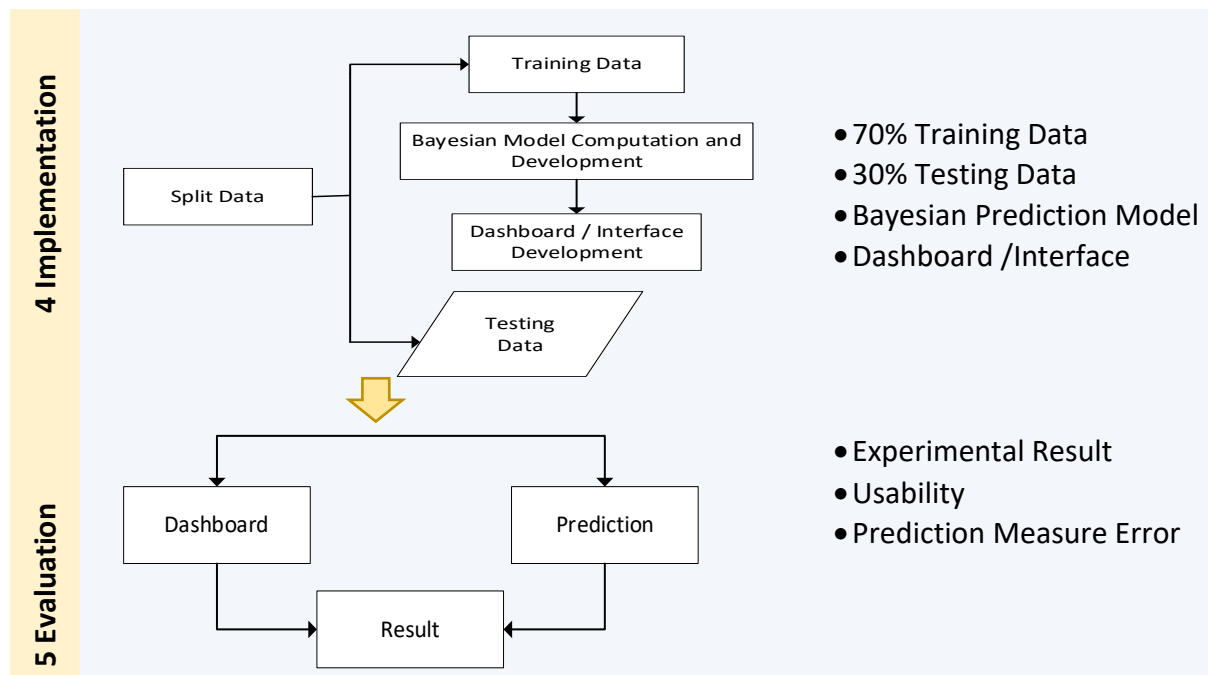
Figure 1. Research Flow

Phase 1 knowledge acquisition that comprises of problem formulation, LAD requirements, and constructs development. Phase 2 is data collection as comprises activities of instrument development, where the constructs for the questionnaire were obtained from Pintrich *et al.* (1993), validation by experts, and a pilot test that has been executed. The independent attributes are sociodemographic, motivation, cognitive strategy use, and personality data obtained directly through the questionnaire. While academic results, attendance, and other academic continuous assessment like quizzes, assignments, tests, and project progress were synthetic data. Table 1 lists the construct, its purposes, and their outcome.

Table 1
Constructs and Outcome

Item	Constructs & Purposes		Outcome
A1 -A10	Student socio demography data		Students profile
Q1 – Q5	**Motivation (M)** reward toward their accomplishment, value each task given to them, intrinsic motivation, self-efficacy, anxiety.		• Intervention • Monitoring • Disciplinary action • Recommendation to Student Affairs takes charge of helping with assistance or counseling.
Q6 – Q10	**Cognitive Strategy Used (C)** studying strategy, memory recalling, ability, prioritising strategy, learning experience, holistic thinking.		
Q11 – Q15	**Personality (P):** Mindfulness, ability to change, resilience, ability to change with high desire, diligence and linking ability.		
Continuous assesment Quiz, tutorial, assignment, project and test			**Prediction:** Final grade

In phase 3, data treatment, cleansing and exploratory data analysis (EDA) have been conducted to identify the significant patterns of each attribute, relationships, and data anomalies. All outliers were replaced with the corresponding values or the closest approximation of those values. Phase 4 focuses on the design and development of the prediction model and dashboard. The prediction model has been trained iteratively to achieve higher training precision, and weighted elements have been added to the computation. In the final phase 5, the evaluation of the prediction model using the confusion table has been done, and simulations of the whole process have been executed.

LAD Logical Flow
Based on the interviews with three academics in Phase 1, a set of requirements for the LAD has been made up of a set of features and conditions. It consists of five basic features and provides access to two types of users. The description of the basic features obtained from the requirements is as follows
i.   Link, update, and view socio-demographic and profile data.
ii. Link, update, and view behaviour, cognitive skill used, and personality data.
iii. Update data for ongoing assessments, such as quizzes, assignments, on-going projects, and tests.
iv. View grade, interventions, and descriptive, diagnostic, and predictive analytics outcomes. View the benchmarking descriptive analysis of each student against his or her group member.

Figure 2 illustrates the flow of the LAD framework. As shown in the diagram, the outcome of the prediction leads to recommendations for either intervention, monitoring, disciplinary actions, or the last option of submitting the student's case to student affairs.
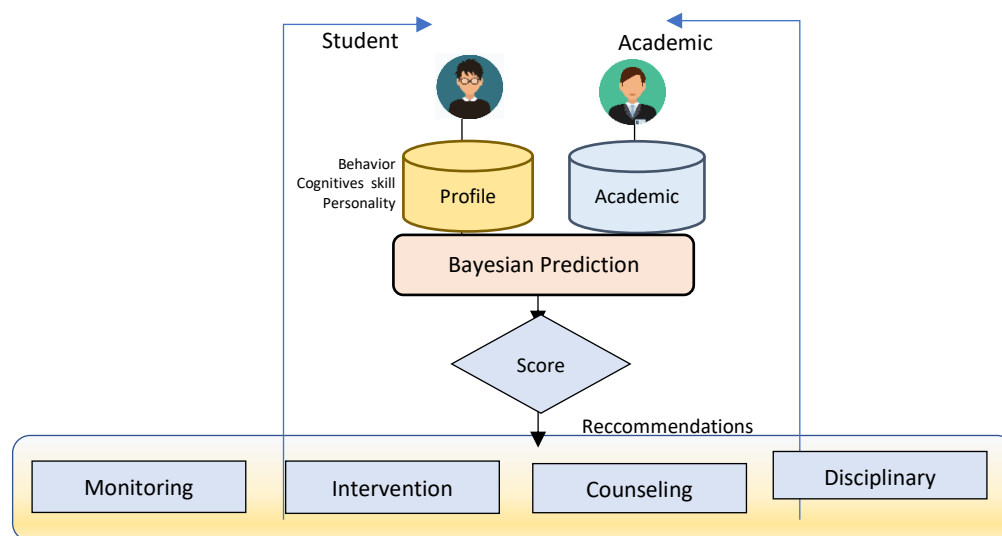

Figure 2. Learning analytics logical framework.

The recommendation is made based on the appropriateness calculated by the target class prediction. It is a formative procedure that is obtained based on a score indicator that has been set and tested iteratively. Further descriptions in Table 2 explain the details of the actionable indicators set in the model of the M, C, and P scores.

Table 2
Actionable Indicator

	M, C, P Average Score	Academic	Attendance record	Action to be taken
1	Poor Score: 1 - 2	Poor	Less than 80%	Disciplinary action and issue case to Students Affair Department for help
2		Below 69%	Above than 80%	Intervention
3	Moderate Score:3	Good		Close monitoring for immediate intervention
4	Good Score:4 - 5	Above 70%	Above 80%	Monitoring and maintaining performance

Bayesian Prediction Model

The logical flow of the Bayesian model started with calculations of prior and likelihood probability. As shown in Figure 3, the next step is selection of class target. Based on the prediction probability, the class target will be determined based on larger probability values. If the probability of PASS is higher than the probability of FAIL, then the student is predicted to PASS. Using the 70% training set, the model parameters were estimated from the data to train the Bayesian model. Specifying the prior distribution, figuring out the likelihood function, and using Bayes' theorem to update the prior distribution to get the posterior distribution are all parts of this process.
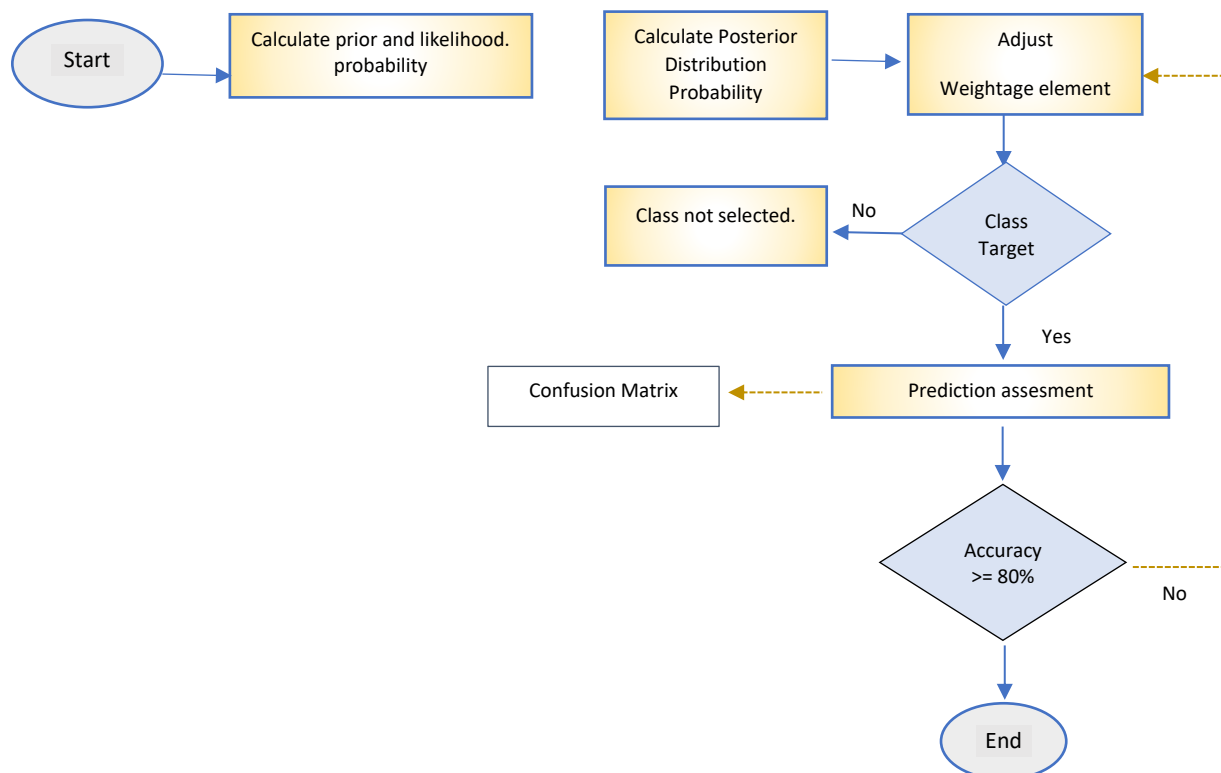
Fig. 3. Flow of Bayesian model implementation

Once the model has been trained, it can be evaluated using the other 30% of the testing set. The testing set is used to see how well the model can predict data that it has never seen before. This is done by using the target class to make predictions for the testing set and comparing these predictions to the actual values. Adjustments can be made to improve its predictive ability. This involved changing the prior distribution, modifying the likelihood function, or using a different model altogether. The accuracy calculation is carried out using the following formula:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \qquad (2)$$

TP = True Positive   TN = True Negative, FP = False Postivie   FN = False Negative

True Positive (TP) meant that the prediction was right about a condition being present, while True Negative (TN) meant that the prediction was right about a condition not being present. False Positive (FP) is a test result that wrongly shows that a condition is present, while False Negative (FN) is a test result that wrongly shows that a condition is not present. Figure 4 shows the positive and negative inference.

Figure 4. Actual and predicted confusion matrix.

To cross checked the confusion matrix, the precision, recall and F1-score has also been calculated. Their formula is as the followings:

Precision = TP / (TP+FP)   (3)
Recall = TP / (TP+FN)   (4)
F1 score = 2TP / (2TP + FP + FN) (5)
TPR = TP / (TP + FN)   (6)

Interface of the Prototype

Microsoft Power BI was used to make the dashboard, and Microsoft Excel and Data Analysis Expressions (DAX) were used to figure out the Bayesian prediction probability values. There are three main pages in the LAD, which are: (a) the individual academics page, consisting of scales of behaviour, cognitive ability, and personality traits that are combined with suggestions for intervention (if any); (b) the individual academics page, consisting of grade prediction and average scales of behaviour, cognitive ability, and personality traits, as well as its average score in the Bubble chart. As indicated on the page, there is also student feedback on their weekly performance. Table 4 lists a few screen shots to illustrate the pages.

Table 4
Screenshot of the LAD
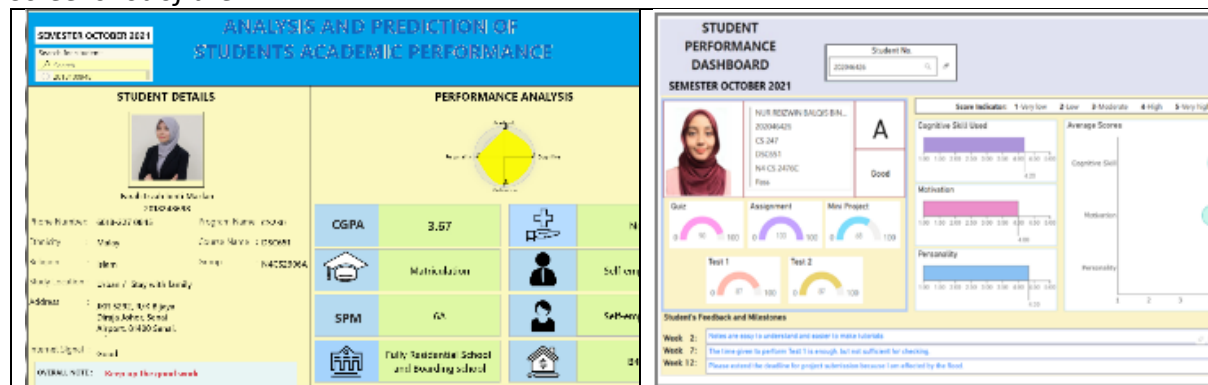
Link, update, view socio-demographic and profile data	
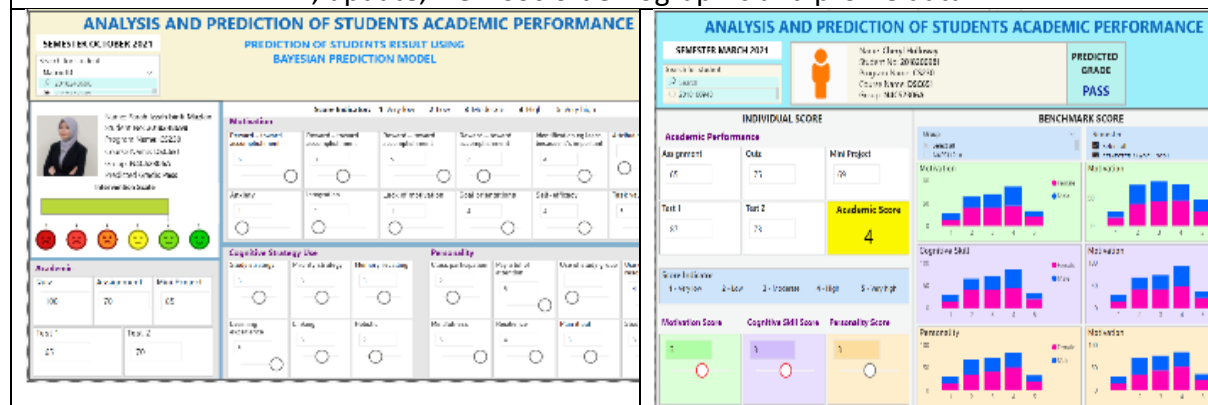
• Link, update, view behaviour, cognitive skill used, and personality data • Update data for ongoing assessments, such as quizzes, assignments, on-going projects, and tests	• View grade, interventions, and descriptive, diagnostic, and predictive analytics outcomes • View the benchmarking descriptive analysis of each student against his or her group member

Result and Discussion

Upon conducting a benchmark analysis using manual calculations, the model can be deemed satisfactory if the percentage of accuracy in probability results for both the testing and training data is 80% or above. The results of the formative test can be found in Table 5 and Table 6.

Table 5
Confusion Matrix For Individual Grade Prediction

	Class		Accuracy (%)
Prediction	**Pass**	**Fail**	
Pass	1316	67	80%
Fail	328	389	85%
Total of calss	1644	456	
Overall accuracy: 81%			

Thus, in the above table, the true positive (TP) is 1316 and the true negative (TN) is 389. The number for false positives (FP) is 328, and the value for false negatives (FN) is 67. Consequently, the following table yields an overall accuracy of 81%.

Table 6
Confusion Matrix For Suggestion For Intervention

Class target				
Prediction	**Poor**	**Moderate**	**Good**	**Excellent**
Poor	350	145	1	0
Moderate	106	548	27	0
Good	0	111	667	1
Excellence	0	0	5	139
Total of class	456	804	700	139
Overall accuracy: 81%				

Thus, in the above table, the TP is 350 (poor), 548 (moderate), 667 (good), and 139 (excellent).The confusion Table 6 shows an accuracy of 81%, it means that 81% of the predictions made by the model were correct. This level of accuracy is considered valid because it's higher than a random guess, which would achieve an accuracy of 50% in a binary classification problem. In this study, the model is proven to be useful if it is accurate more than 80% of the time. To validate the accuracy further, precision, recall, and the F1-score are also performed to provide a more nuanced evaluation of the model's performance and is shown in Table 7. For both calculation the sensitivity is 0.8 which means the model can properly identify 80% of the positive samples. This indicates that there may be some false negatives, in which the model misidentifies a positive sample as negative. In this study context, a sensitivity of 0.8 is generally considered to be a good score as the selection is depending on learning domain and the risk associated with misclassification.

Table 7
Precision, Recall and F1 Score

Class	Truth	Classified	A	P	R	F1 Score
Individual Grade Prediction						
Pass	1644	1383	81.19%	0.95	0.80	0.87
Fail	456	717	81.19%	0.54	0.85	0.66
Suggestion for Intervention						
Poor	456	496	87.99	0.71	0.77	0.74
Moderate	804	681	81.47	0.80	0.68	0.74
Good	700	778	93.14	0.86	0.95	0.90
Excellent	139	144	99.76	0.97	1.0	0.98

A = accuracy, P= precision and R = recall

A classification model with an F1 score of 0.87 has a great overall performance in terms of both precision and recall. This suggests that the model is capable of accurately identifying positive samples while avoiding false positives and false negatives. This is regarded as an

excellent score, indicating that the model is doing well. The F1 score of 0.66, on the other hand, indicates that the model's performance is not as good as it may be. A score of 0.66 shows that the model is either struggling to balance precision and recall or is skewed toward one metric over the other. Further analysis may be required in this scenario to determine the source of the lower score and to improve the model's performance. An F1 score of 0.74 indicates that the model performs moderately well, with potential for improvement. It suggests that the model does a fair job of balancing precision and recall, although there may be some misclassifications that need to be corrected. If a model has several F1 scores of 0.74, it indicates that the model's performance is consistent across diverse subsets of data or classes in a multi-class problem. An F1 score of 0.90 suggests that the model performs well in terms of precision and recall. This implies that the model is capable of accurately identifying positive samples while avoiding false positives and false negatives. An F1 score of 0.98 indicates that the model performs exceptionally well, with very high precision and recall. This shows that the model can accurately identify positive samples while avoiding false positives and false negatives with a high level of accuracy.

Conclusion

This study demonstrates that the utilization of machine learning analytics can assist in the identification of students that require additional support, guidance, or early intervention by LAD. The Bayesian model was used for this study due to its ability to handle attribute variance effectively. It possesses unique strengths in conducting analytics to forecast academic progress. In general, the analytics effectively identify, evaluate, and convey important trends and insights in students' data narratives. The offered outcome, showcased through interactive visuals, provides viewers with a more lucid perspective of specific data compared to textual log data. The usability of the system facilitates early intervention by allowing timely execution of interventions based on many predictive indicators. The evaluation of the prediction model indicated that a dashboard may be developed to facilitate early intervention. This implies that behaviour, cognitive skills, and personality traits have a role in predisposition. Generally, the F1 scores, which range from modest to outstanding, suggest that the model has promise and could be further improved in the future.

Learning analytics using Bayesian prediction models aligns with constructivist theories by emphasizing the importance of personalized learning experiences. By analysing individual student data, educators can tailor instructional approaches to meet diverse learning needs, fostering active engagement and deeper understanding. Additionally, this approach resonates with socio-cultural theories by recognizing the social context of learning. Through data-driven insights, educators can create collaborative learning environments that promote social interaction and knowledge construction among students. Thus, learning analytics not only enhances educational practices but also reinforces key principles of constructivism and socio-cultural theories, ultimately enriching the learning experience.

Acknowledgment

References

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. Telematics and Informatics, 37, 13–49.

Alves, G. R., Carvalho, R. N., & Lima, M. V. (2021). A Bayesian model for predicting the academic performance of students in blended learning environments. Computers & Education, 173, 104319.

Baker, R. S. (2016). Learning analytics and education data mining. In Handbook of Educational Psychology (pp. 272-290). Routledge.

Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. IEEE Transactions on Learning Technologies, 10(4), 405–418.

Guan, J., & Wen, Z. (2020). The association between personality and academic achievement: A meta-analysis. Personality and Individual Differences, 155, 109723.

Hilliger, I., Miranda, C., Schuit, G., Duarte, F., Anselmo, M., & Parra, D. (2021). Evaluating a Learning Analytics Dashboard to Visualize Student Self-Reports of Time-on-task: A Case Study in a Latin American University. ACM.

Khalil, M., & Ebner, M. (2014). Learning analytics: Challenges and limitations. In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications.

Matzen, L., Haass, M., Divis, K., & Stites, M. C. (2017). Patterns of Attention: How Data Visualizations Are Read. In International Conference on Augmented Cognition, Lecture Notes in Computer Science.

Millecamp, M., Gutiérrez, F., Charleer, S., Verbert, K., & De Laet, T. (2018). A Qualitative Evaluation of a Learning Dashboard. LAK'18, Association for Computing Machinery.

Muñoz-Merino, P. J., Ruipérez-Valiente, J. A., & Alario-Hoyos, C. (2019). Challenges and opportunities for learning analytics and adaptive learning in higher education. Journal of Interactive Media in Education, 2019(1), 2.

Ong, K. X., & Singh, D. (2021). Development of Learning Analytics Dashboard based on Moodle Learning Management System. International Journal of Advanced Computer Science and Applications, 12(7), 838-843.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. Psychological Bulletin, 135(2), 322-338.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). Educational and Psychological Measurement, 53(3), 801-813.

Wulandari, D. A., Annisa, R., Yusuf, L., & Prihatin, T. (2020). Educational data mining for student academic prediction using k-means clustering and naïve bayes classifier. Jurnal Pilar Nusa Mandiri, 16(2), 155-160.