**Introduction:**

Blended learning, which integrates both online and face-to-face instruction, has emerged as a popular pedagogical approach in higher education. Finding the right balance between these components is critical for promoting student engagement and achieving better learning outcomes. This study aims to examine the most effective blend of instruction for Malaysia Higher Education students, as well as identify the factors that impact this optimal mix.

To attain these goals, this research leverages a variety of previous studies and scholarly works. Notably, several references have been consulted to inform and steer this investigation. Salas-Rueda's (2020) research evaluates the effect of the WampServer tool in blended learning, analyzing topics such as data science, machine learning, and neural networks. Musabirov, Pozdniakov, and Tenisheva's (2019) study explores predictors of academic success in blended learning, with a particular emphasis on a data science minor.

Additionally, Wang, Huang, and Omar (2021) scrutinize the employment of blended learning models through text mining techniques. Park, Yu, and Jo (2016) conduct an institutional case study in Korea, clustering blended learning courses according to online behavior data. Banyen, Viriyavejakul, and Ratanaolarn (2016) present a blended learning model designed to promote learning outcomes among Thai undergraduates. Finally, Yapici and Akbayin (2012) examine into the effectiveness of blended learning in developing students' problem-solving skills, focusing on a Turkish university.

By drawing on the insights and discoveries of these sources, this study seeks to add to our comprehension of the appropriate combination of face-to-face and online instruction for University of Malaya students. Its objectives are to pinpoint the primary factors that influence the optimal blend, analyze the impact of various proportions of online and face-to-face instruction on student engagement and academic achievement, and furnish research-backed suggestions for crafting and executing successful blended learning approaches.

All in all, this investigation aims to bridge the current research gap by examining the most effective composition of instruction in blended learning environments, with a particular focus on University of Malaya students. The results of this inquiry will provide valuable guidance and suggestions for educators, instructional designers, and policymakers seeking to enhance the efficiency and efficacy of blended learning approaches, resulting in better student engagement and academic performance.

Critical Analysis Table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Reference | Research Objective | Research Methodology | Model Used | Research Outcome | Research Implication | Research Gap |
|  |
| Salas-Rueda (2020) | To examine the impact of the WampServer application in blended learning using data science, machine learning, and neural networks. | Quantitative Method | Neural Network Model | Highest accuracy of the model is 71.43%. To conlude, the use of WampServer in blended learning can improve the teaching and learning conditions in computer science education | (1)The study suggests that the use of WampServer in blended learning can improve the teaching and learning process by facilitating the assimilation of knowledge and developing students' skills under this hybrid modality.  -This finding can be useful for educators who want to incorporate blended learning into their teaching practices.  (2) The study highlights the importance of using machine learning techniques to analyze educational data.  -The predictive models used in this research can help educators identify factors that positively or negatively influence student performance and adjust their teaching strategies accordingly.  (3)The study suggests that future research can explore the use of various ICTs in blended learning such as social networks, web 2.0 tools, and audiovisual content in computer science education.  -This can help educators to design more effective and engaging learning experiences for their students. | (1)The size of the sample used in this study. The sample size was relatively small, which may limit the generalizability of the findings to other contexts.  -Future research can address this limitation by using a larger and more diverse sample.  (2)The study only focused on the impact of WampServer in blended learning in computer science education.  -Future research can explore the use of WampServer and other ICTs in blended learning across different disciplines to determine their effectiveness. |  |
| Musabirov et al. (2019) | To examine predictors of academic achievement in blended learning for a data science minor. | Quantitative Method | Logistic Regression | The research analyzed the main predictors of academic achievement for the 2nd year non-STEM undergraduates from a blended data science course. The factors related to initial mathematics knowledge, specific traits of educational programs, online and off-line learning engagement, and connections with peers were considered. The grade gap with friends demonstrated a significant positive connection with academic achievement. Students who got higher grades than their friends did in the first semester tend to have an additional increase in academic achievement. | The research provides insights into the factors that influence academic achievement in blended-learning settings. The findings suggest that peer effects and engagement in learning are important predictors of academic success. The results can be used to inform the design of educational programs and interventions aimed at improving student outcomes in blended-learning environments. Additionally, the study highlights the importance of considering individual differences in students' backgrounds and learning behaviors when designing such programs. | (1)Lack of consideration for other potential factors that may influence academic achievement, such as socioeconomic status, motivation, and study habits.   (2)The study only focuses on non-STEM undergraduates from a blended data science course, so the generalizability of the findings to other populations or disciplines may be limited. |  |
| Chango et al. (2019) | To predict academic performance of university students in blended learning using multi-source data. | Qualitative Method | Decision Tree, KNN | Accuracy for Nearest-Neighbourhood Method = 82.4561 Accuracy for Decision Tree = 82.4561  A predictive model for academic performance of university students using multi-source data in blended learning environments. The model was created by applying data fusion and white box classification algorithms to the gathered data. The best prediction model was obtained using the PART algorithm with categorical data. The model identified variables with well-known predictive value in learning achievement and dropout, as well as emerging sources of information like student behavior at Moodle forums. This model can help instructors take remedial actions with students at risk of dropout or failing. | (1)The predictive model developed in this study can help instructors identify students at risk of dropout or failing and take remedial actions to improve their academic performance.  -This can lead to better retention rates and higher graduation rates, which are important indicators of the quality of education provided by universities.   (2)The use of data fusion and white box classification algorithms can help researchers gain insights into the factors that affect student performance in blended learning environments.  -This can inform the development of more effective teaching strategies and personalized learning experiences for students. | (1) This study is the limited sample size of only 65 university students from a single institution.  -While the authors acknowledge this limitation and suggest using more courses and subjects in future experiments, a larger and more diverse sample size could provide more generalizable results.   (2)The study focuses on predicting academic performance using multi-source data, but does not explore the underlying reasons for student success or failure.  -Future research could investigate the factors that contribute to student performance in blended learning environments, such as motivation, engagement, and learning strategies.  (3)Other machine learning techniques like deep learning could be explored to improve prediction accuracy. |  |
| Van Goidsenhoven et al. (2020) | To predict student success in blended learning. | Qualitative Method | Random Forest | Accuracy for model = 0.84 Leave-One-Out Cross-Validation accuracy remains the most important metric to gauge the accuracy of the predictive models.  The research outcome of this work is that success vs. fail predictions could be made as early as in the middle of the course, which could be used in the future for timely interventions, both for failure prevention and for supporting the positive learning behavior of students. The study also discusses the results, as well as the limitations of the study, and provides several possible directions for future research. | (1)The study shows that it is possible to predict student success or failure in an online course as early as in the middle of the course, which could be used for timely interventions to prevent failure and support positive learning behavior.   (2)The study also provides insights into the factors that contribute to student success or failure in an online course, which could be used to design more effective online courses.   (3)The study highlights the importance of using appropriate predictive models and metrics to evaluate their performance.   (4)The study suggests several directions for future research, including exploring the generalizability of the predictive models to other courses and contexts. | (1) The study was conducted on only one data set from one specific course, which limits the generalizability of the findings.  - future studies could explore the generalizability of the predictive models to other courses and contexts.   (2)The study lacks a high enough amount of data points to independently verify the prediction model, which suggests that a data set from another year is needed to verify the prediction model as proposed in this paper.  (3) It does not explore how these factors interact with each other or how they change over time.  - future studies could explore these interactions and changes to gain a more nuanced understanding of student success and failure in online courses. |  |
| Wang et al. (2021) | To analyze the application of blended learning using text mining methods. | Qualitative Method | Text-Mining | The research outcome of the paper is an analysis of blended learning practices in 17 countries using text mining method. The study compares the implementation models, platforms, and practical courses adopted and clarifies the root causes of the difficult problems in blended learning practice to increase the effectiveness of blended learning implementation in the future. It also identifies factors that hinder the implementation of blended learning and provides insights into how schools can select a sustainable blended learning model. | (1)It provides insights into the implementation of blended learning practices in different countries.  -The study identifies factors that hinder the implementation of blended learning and suggests ways to increase the effectiveness of blended learning implementation in the future.  -It can be useful for educators, policymakers, and researchers who are interested in improving the efficiency of blended learning practices. | The research identifies factors that hinder the implementation of blended learning, which can be considered as a research gap that needs to be addressed in future studies. |  |
| Park et al. (2016) | To cluster blended learning courses by online behavior data | Qualitative Method | Clustering | This study clustered 612 courses that were purposefully sampled with cases presenting relatively active online activity.   Latent Class Analysis identified four types of blended course with C-D-S-I (C: Communication or Collaboration, D: Delivery or Discussion, S: Sharing or Submission, and I: Inactive or Immature). While 50% of courses fell into Type I, the rest half of courses were divided into Type C (24.3%), Type D (18%), or Type S (7.2%). The results revealed that students in Type C courses more participated in small-group project online as well as conversations with instructor and/or peers in terms of course requirements. | (1)The data-driven approach adopted in this study could be a seminal contribution to university staffs that need to monitor the whole status and make better decisions based on the understanding of various contextual variables.   (2)The study's findings can help universities to improve their blended learning courses by identifying patterns and trends in student performance and engagement.   (3)The study's methodology can be applied to other higher education contexts for further research. | (1)Despite the increasing popularity of blended learning, there is still a lack of consensus on how to define and measure it. |  |
| Banyen et al. (2016) | To develop a blended learning model for enhancing learning achievement of Thai undergraduate students. | Qualitative Method | Linear Regression | Blended Learning Model has been used. The model aimed to improve the learning management conditions and methods of a Thai university undergraduate course in information, innovation, and technology in education.   The study also compared the learning achievements and information literacy of student groups using the blended learning method and traditional learning method with information literacy as the covariate.   The results showed that the blended learning method was more effective in enhancing students' learning achievements and information literacy than the traditional learning method. | (1)The blended learning model developed in this research can be applied to other courses and educational institutions to improve learning management conditions and methods.   (2)The study also highlights the effectiveness of blended learning in enhancing students' learning achievements and information literacy.  -This can encourage educational institutions to adopt blended learning as an effective approach to teaching and learning.   (3)The study provides a guideline for future courses on how to effectively improve the blended learning model and instructional media in accordance with the 80/80 criteria. | (1) This study focused only on undergraduate students in a Thai university and a specific course in information, innovation, and technology in education.  -Therefore, the generalizability of the findings to other educational contexts and courses may be limited.   (2)The study did not explore the challenges and limitations of implementing blended learning or the perspectives of teachers and students on the effectiveness of this approach.  -Further research is needed to address these gaps and provide a more comprehensive understanding of blended learning in different educational contexts. |  |
| Yapici and Akbayin (2012) | To examine the effect of a blended learning model on high school students' biology achievement and attitudes towards the internet. | Qualitative Method | Normal Statistical Analytics | The blended learning model has been used in this study to teach the course of human anatomy to 1st grade biology students from Pompeu Fabra University (Barcelona).  The study found that the blended learning model contributed more to the students' biology achievement than traditional teaching methods did and that the students' attitudes towards the internet developed statistically significantly. | The blended learning model can be an effective teaching method for improving high school students' biology achievement and attitudes towards the internet.  -This suggests that incorporating blended learning into classroom instruction can lead to positive outcomes for students. | The generalizability of the findings to other subjects or student populations as this study only focused on high school biology students from one university -further research is needed to determine if blended learning has similar effects in other subjects and with different student populations. |  |
| Salas et al. (2018) | To analyze the use of continuous improvement, technology, and flipped classroom in the teaching-learning process through data science, specifically in the field of mathematics education | Mixed Method | Decision Tree | Highest accuracy of the model is Predictive model 1 with 95%  The accuracy of Predictive Model 1 is 95.00%, and the accuracy of Predictive Model 2 is 85.00%. These models were used to determine the impact of FreeDFD simulator, flipped classroom, and YouTube videos on the educational process through learning machine.   The study found that the assimilation of knowledge by means of these tools positively influences the grade of the first partial related to Simple conditional structure, Multiple conditional structure, and While iterative structure. However, it is not specified what outcomes were obtained from Predictive Models 3 and 4 in relation to model accuracy. | (1)The study proposes a virtual environment using the FreeDFD simulator, flipped classroom, and YouTube videos to link theoretical topics in the Logical of Predicates Unit with computer science.  -The research shows that the assimilation of knowledge by means of these tools positively influences the grade of the first partial related to Simple conditional structure, Multiple conditional structure, and While iterative structure.  -This suggests that technology can be used as an effective tool to enhance the teaching-learning process in mathematics education.   (2)The use of predictive models can help educators identify factors that influence student performance and make data-driven decisions to improve educational outcomes. | It does not explore other potential factors that may influence student performance.  -For example, the study does not consider the impact of teacher quality or student motivation on academic achievement.  -Future research could investigate these factors to provide a more comprehensive understanding of how to improve mathematics education outcomes. |  |
| Salas et al. (2020) | To analyze the impact of the flipped classroom in the teaching-learning process on statistics considering data science and machine learning. | Quantitative Method | Decision Tree, Linear Regression | The results of machine learning indicate that the participation of the students  before, during and after the class positively influences the assimilation of knowledge and development of mathematical skills.  -Likewise, data science identifies 6 predictive models on the use of flipped classroom in the educational context through the decision tree technique.  The results of machine learning and decision tree technique indicate that student participation before, during, and after class positively influences their assimilation of knowledge and development of mathematical skills. Additionally, students are motivated and satisfied to use technological tools in the Introduction to Statistics Unit. | (1)The study suggests that flipped classroom can be an effective teaching strategy in statistics education.  -The use of technological tools such as YouTube videos, collaborative exercises, and spreadsheets can enhance student participation and motivation in learning statistics.   (2)Machine learning techniques such as linear regression and decision tree can be used to analyze the impact of flipped classroom on student learning outcomes. | (1) Future research can explore the impact of flipped classroom in teaching other statistical concepts beyond linear regression.   (2)Future studies can investigate the effectiveness of flipped classroom in different educational settings and with different student populations. |  |

Summary:

The critical analysis table provides an overview of several research studies focusing on blended learning and its impact on various educational outcomes. These studies utilize different models, such as decision trees, logistic regression, neural networks, nearest-neighbour method, random forest, text mining, clustering, and linear regression, to investigate the research objectives.

While these studies contribute valuable insights to the field of blended learning, they also identify research gaps that need further exploration. Some common research gaps include:

1. Limited sample sizes: Many studies acknowledge the limitation of small sample sizes, which may restrict the generalizability of their findings to other contexts. Future research could address this gap by incorporating larger and more diverse samples.
2. Limited understanding on the main factor influenced blended learning in Malaysia: Given the differences in geographical, demographic, and background characteristics of Malaysian students compared to other foreign studies, it is necessary to understand the main factors influencing the optimal blend of online and face-to-face instruction in blended learning systems in Malaysian higher education to enable comparison with students from other countries in this context.
3. Lack of consideration for other influential factors: Some studies do not fully explore additional factors that may affect educational outcomes, such as socioeconomic status, motivation, and study habits. Future research could investigate these factors to gain a more comprehensive understanding of the complexities involved in blended learning.
4. Insufficient exploration of impact of different ratios of online and face-to-face instruction on student engagement and learning outcomes in blended learning: Most of the studies mentioned well Designed Model is important but did not mention about the impact of ratios for online and face to face instruction
5. Need for exploration of advanced machine learning techniques: While some studies utilize machine learning techniques, there is potential for further exploration of advanced methods such as deep learning to improve prediction accuracy and uncover deeper insights into the factors influencing student performance.

Reference:

Salas-Rueda, R. A. (2020). Impact of the WampServer application in Blended learning considering data science, machine learning, and neural networks. *E-Learning and Digital Media*, *17*(3), 199-217.

Musabirov, I., Pozdniakov, S., & Tenisheva, K. (2019). Predictors of academic achievement in blended learning: The case of data science minor. *International Journal of Emerging Technologies in Learning (Online)*, *14*(5), 64.

Chango, W., Cerezo, R., & Romero, C. (2019, December). Predicting academic performance of university students from multi-sources data in blended learning. In *Proceedings of the second international conference on data science, E-learning and information systems* (pp. 1-5).

Van Goidsenhoven, S., Bogdanova, D., Deeva, G., Broucke, S. V., De Weerdt, J., & Snoeck, M. (2020, March). Predicting student success in a blended learning environment. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 17-25).

Wang, L., Huang, Y., & Omar, M. (2021). Analysis of blended learning model application using text mining method. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(1), 172-187.

Park, Y., Yu, J. H., & Jo, I. H. (2016). Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute. *The Internet and Higher Education*, *29*, 1-11.

Banyen, W., Viriyavejakul, C., & Ratanaolarn, T. (2016). A Blended Learning Model for Learning Achievement Enhancement of Thai Undergraduate Students. *International Journal of Emerging Technologies in Learning*, *11*(4).

Yapici, I. U., & Akbayin, H. (2012). The Effect of Blended Learning Model on High School Students' Biology Achievement and on Their Attitudes towards the Internet. *Turkish Online Journal of Educational Technology-TOJET*, *11*(2), 228-237.

Salas-Rueda, R. A. (2018). Analysis on the use of continuous improvement, technology and flipped classroom in the teaching-learning process by means of data science. *Online Journal of Communication and Media Technologies*, *8*(4), 325-343.

Salas-Rueda, R. A. (2020). Flipped classroom: Pedagogical model necessary to improve the participation of the students during the learning process. *LUMAT: International Journal on Math, Science and Technology Education*, *8*(1), 271-296.