SCIENTIFIC Publishing Limited

Journal of Intelligent Communication

http://ojs.ukscip.com/index.php/jic

Article

Fairness, Bias, and Ethics in AI: Exploring the Factors Affecting Student Performance

Doris Omughelli [1], Neil Gordon [2,*] and Tareq Al Jaber [2]

[1] Data Science, AI and Modelling, University of Hull, Hull HU6 7RU, UK
[2] School of Computer Science, University of Hull, Hull HU6 7RU, UK
[*] Correspondence: n.a.gordon@hull.ac.uk

Received: 26 June 2024; **Revised**: 19 July 2024; **Accepted**: 28 July 2024; **Published**: 30 July 2024

Abstract: The use of artificial intelligence (AI) as a data science tool for education has enormous potential for increasing student performance and course outcomes. However, the growing concern about fairness, bias, and ethics in AI systems requires a careful examination of these issues in an educational context. Using AI and predictive modelling tools, this paper explores the aspects influencing student performance and course success. The Open University Learning Analytics Dataset (OULAD) is analysed using several AI techniques (logistic regression and random forest) in this study to reveal insights about fairness, ethics, and potential biases. This dataset has been used by hundreds of studies to explore how educational data mining can provide information on students. However, potential bias or unfairness in that dataset could undermine the results and any conclusions made from them. To gain insights into the dataset's properties, this was analysed using a typical data science methodology, which included data collecting, cleaning, and exploratory data analysis using Python. By applying AI -based predictive models, this study aims to detect potential biases and their impact on student outcomes. Fairness and ethical considerations are central to the analysis as the representation of various demographic groups and any disparities are evaluated in course results. The goal is to provide useful insights on the proper use of AI in education, while also maintaining equitable and transparent decision-making procedures. The findings shed light on the complicated interplay between artificial intelligence, fairness, and ethics in the context of student performance and course success. As artificial intelligence continues to influence the educational landscape, this study will provide useful ideas for encouraging fairness and minimising biases, resulting in a more inclusive and equal learning environment.

Keywords: artificial intelligence; education; fairness; bias; ethics; predictive modelling

1. Introduction

Assessing student learning outcomes in higher education (HE) has gained significant importance globally over the past decade, with access to a quality education one of the United Nations Sustainable Development Goals, leading to the development of various assessment systems at national and international levels [1–3]. The ethical implications of AI, as discussed by the European Parliamentary Research Service, underscore the need for responsible and ethical AI implementations [4]. The International Framework for Principled AI emphasizes principles such as security and accountability to develop trustworthy algorithmic decision-making systems [5]. Machine learning algorithms, notably in sectors like banking, healthcare, and employment, are increasingly shaping daily life, but their complexity and reliance on historical data raise concerns about biases [6]. Biased data can perpetuate inequalities in AI systems, leading to mistrust and anxiety among the public [7]. For instance, biased hiring decisions based on race or gender can exacerbate societal inequities. Addressing these issues is crucial for ensuring fair and responsible AI deployment across various sectors to enable trustworthy data science outcomes.

As the use of algorithms and predictive models for student evaluation grows, a growing body of research has investigated the ethical implications. While online education has rapidly extended access to higher education, questions regarding equal student outcomes remain. Demographic characteristics have been identified that influence online course completion and achievement in previous research. For example, Park and Hee [8] found older students have higher online course drop-out rates. Studies have also shown worse attainment for minority students in online platforms [9].

One of the most crucial evaluation quality standards is fairness. The Standards were created by the American Educational Research Association [10], the American Psychological Association (APA), and the National Council on Measurement in Education (NCME). They provide teachers with helpful guidance as well as a complete overview of more than 240 standards for the creation, application, and assessment of educational and psychological tests.

Educational uses of AI include aspects such as personalised learning platforms to aid students in their learning, automated evaluation tools to support teachers, and learner behaviour analytics using facial recognition software. Such systems could be susceptible to bias and unfairness, for instance, an AI system that tracks student activity and uses facial recognition could be biased against students from certain racial or ethnic groups [11]. Holstein et al. [12] discovered significant differences in mistake rates of facial analysis algorithms across gender and ethnic groups, highlighting the importance of fairness considerations in deployment.

The application of AI tools and approaches to guarantee ethics, justice, and impartiality in education is an opportunity, for example, machine learning may evaluate a student's grade by removing human bias [13]. A report by UNESCO [14] shows how AI technology may assist educational institutions in using data to enhance learning outcomes, educational equity, and educational quality in underdeveloped nations. However, there are problems with bias and a lack of accountability and transparency in AI systems employed in education. For instance, AI systems used to assess student work or determine student outcomes may be biased against particular groups depending on criteria like race, gender, or socioeconomic status [15].

The importance of ethical norms in today's educational testing practises is growing. Concerns regarding potential biases and their impact on student achievements have grown as AI's influence in education has grown. This paper focuses on the use of AI-based predictive models to identify and eliminate biases in AI decision-making systems in order to address these concerns. The project intends to discover and correct any potential biases in AI models by analysing the OULAD dataset and utilising these predictive models, assuring fairness, openness, and accountability in the student assessment process. By examining the effectiveness of this well-used dataset, deep learning, and regression algorithms in forecasting student dropout and outcomes based on various groupings of factors, specifically student demographics, assessment scores, and VLE interaction data, this research builds on earlier research. The study will delve into various aspects, including gender, age, region, disability, and the highest education level, to examine their potential influence on student performance and course success. By incorporating ethical guidelines and frameworks for AI use in education, the research aims to contribute to the responsible and ethical adoption of AI technology in the education domain.

It is our responsibility to provide teachers and students with information about the ethical issues and consequences associated with using algorithms as academics, scientists, and citizens [16]. Frameworks and rules for the moral and responsible use of AI in education have also been developed. This includes the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, which has developed a set of moral standards for AI emphasising accountability, openness, and security [17].

This paper applies AI techniques—including logistic regression and random forest to explore a well-known dataset and identify some of the possible issues regarding fairness and bias. Following this introduction, the paper is structured with an overview of the approach (materials and methods) in Section 2. This includes an outline of the data collection processes and some of the initial analysis of the data to illustrate these. Section 3 includes the detailed results, alongside an interpretation of their implications. Section 4 provides a discussion of this and considers the limitations of this work. The final Section 5 provides some conclusions based on the research, with some recommendations for institutions to address fairness and equity issues in their courses and guidance on the proper use of AI.

2. Materials and Methods

This research adopts a quantitative approach, analysing a well-used dataset and exploring it for potential biases or unfairness. This could be considered an experiment or case study to illustrate the issues of fairness and bias in data used by machine learning and data science applications.

2.1. Data Collection and Cleaning

The Open University Learning Analytics (OULAD) dataset [18] was obtained. The data consists of 7 separate CSV files—see Table 1 for details—and was loaded into a Jupyter Notebook using the Pandas library [19]. It underwent significant processing and transformation to extract features before building prediction models. The OULAD dataset contains 10.9 million rows of 32,593 student data across 7 tables covering courses, assessments, demographics, registration, final grades, and online activity for courses offered by the Open University from 2013–2014. The merged OULAD DataFrame was pre-processed using Pandas and NumPy [20] in Jupyter Notebook. Key steps included handling missing values with .fillna(), dropping unnecessary columns like week_from and week_to due to a high percentage of missing entries (82%) was done using the .drop(), formatting dates, engineering new features like assessment scores, and one-hot encoding categorical variables using pd.get_dummies() [21]. Additionally, Scikit-learn [22] was used for standardization, normalization, and other crucial transformations necessary for accurate analysis.

Table 1. Summary of the OULAD dataset.

Data File (Filename)	No. of Records	Description	Attributes
Courses	22	Information about the modules and presentation	code_module, code_presentation, _module_presentation_length
studentInfo	325,893	Demographic information about the student with their results	code_module, code_presentation, id_student, gender, region, highest_education, imd_band, age_band, num_of_prev_attempts, studied_credits, disability, final_result
studentRegistration	32,593	The time for students to register for a course	code_module, code_presentation, id_student, date_registration, date_unregistration
assessments	196	Assessments for every module presentation	code_module, code_presentation, id_assessment, assessment_type, date, weight
studentAssessments	173,740	Containing the result of students' assessments	id_assessment, id_student, date_submitted, is banked, score
vle	6365	Containing information regarding online learning materials	id_site, code_module, code_presentation, activity_type, week_from, week_to
studentVle	1,048,575	Student interaction with the VLE resources	code_module, code_presentation, id_student, id_site, date, sum_click

2.2. Exploratory Data Analysis

This research aims to investigate student performance in the course and address specific research questions:

- Do gaps exist in online course completion rates and attainment based on gender, age, disability status, and prior academic performance?
- Are there differences in student performance across online courses?

Visual and statistical analyses were performed on the dataset to gain insights into student performance, using visualisations such as count plots, histograms, and bar charts. These depict distributions of parameters such as gender, age bands, region, and previous attempts. Statistical computations such as means and correlations were used to identify relationships between data such as submission dates, scores, and clicks. To get insights into student performance, the produced dataset was subjected to both visual and statistical analysis. This provides a clear view of the overarching patterns in the data. Figure 1 illustrates the count of students per highest education level showing that most of the students in the OULAD dataset are of A Level or equivalent background education.

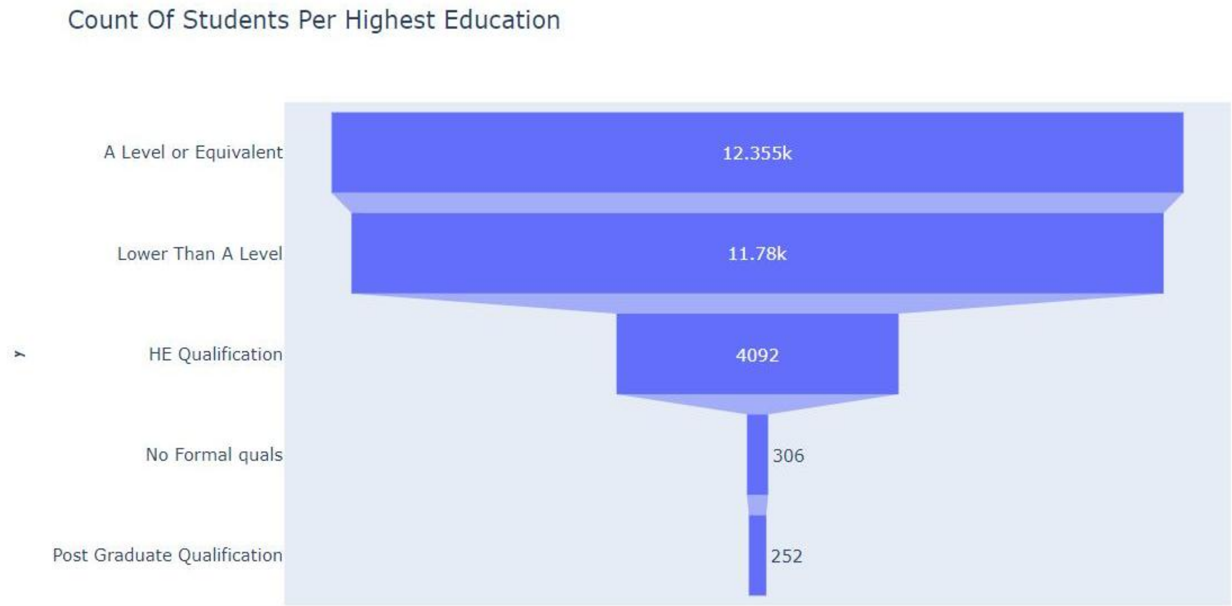

Figure 1. Count of students per higher education.

Figure 2 shows a statistical computation, with means and correlations used to identify relationships between data such as the submission deadline (date), the actual submission date (date_submitted), the grade (score), and the interactions as illustrated by clicks. Diverse data sources, such as student information, assessments, and interaction data, are integrated into a single complete dataset to undertake this research.

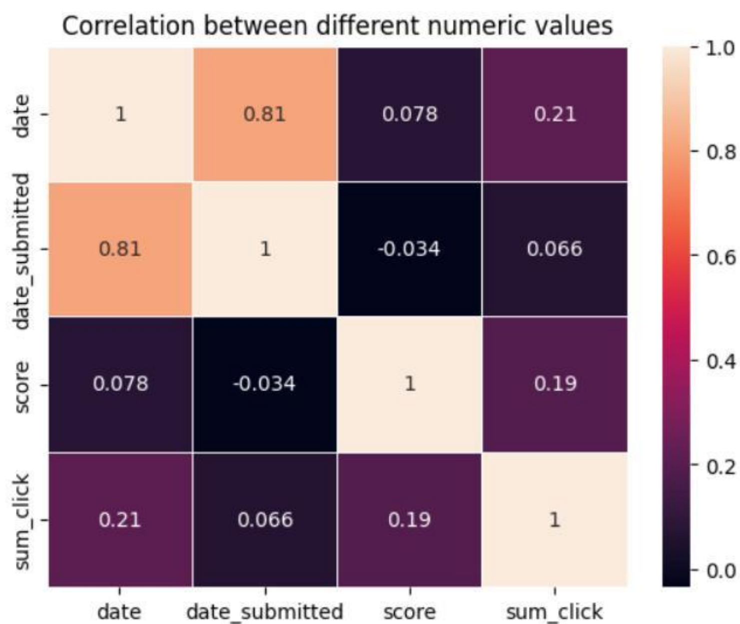

Figure 2. Correlation matrix between date, score, and clicks.

2.3. Statistical Analysis and Modelling

A statistical analysis was performed on the cleansed numeric dataset. Categorical variables were encoded to numerical representations to facilitate the application of statistical models. The best-performing model was carefully examined for potential demographic biases. To reduce biases and discrepancies between groups, fairness requirements and sampling approaches were used. The data is separated into training and testing sets, which is an important stage in model building to ensure accurate evaluation and validation.

Random forest and logistic regression models were developed using Scikit-learn to predict student outcomes from variables like demographics, disability status, and past academic performance. Confusion matrix analysis and classification metrics were used to evaluate model performance. Grid search cross-validation was employed to fine-tune model hyperparameters. Feature importance analysis was conducted to determine the most relevant variables for forecasting the target and assess the model's fairness and robustness across various demographic groups.

Random Forest is a machine learning technique that builds multiple decision trees and combines their results to improve the accuracy of the output. The trees are built from random subsets of the training data, and the predictions of the set of trees are aggregated to produce a final output. As noted above, the random forest for this work used the sklearn model from the Scikit-learn tool, with parameters of a test size of 0.2, and 42 random states. A logistic regression model was also used. This binary classification approach estimated the probability of inputs belonging to a specified class by fitting a logistic function. The model calculates a weighted sum of the input features of the dataset, mapping a probability of 0 to 1 that values lie within it based on a threshold value. For this work, the parameters for the logistic regression were a tolerance of 0.00001. All classes were given a weight of 1, and the lbfgs solver was used, with a maximum iteration value of 100.

This was carried out on a Windows 11 computer, with a Nvidia RTX 3090 GPU.

3. Results

In the online course data considered, an analysis of student demographics, course completion rates, and academic performance provides insightful data on a variety of student behaviours and results. This section highlights key findings, shedding light on biases and patterns observed within the dataset.

Gender Disparities in Performance

The data indicated an unexpected gender gap in evaluations, with females regularly outperforming males by about 5% on average. This finding is consistent with previous research indicating that females outperform males in traditional face-to-face undergraduate courses [23], as well as showing equivalent or higher performance online [24,25].

Age and Course Completion

A noteworthy trend emerged regarding age and course completion rates. Younger students falling within the 0–35 age band displayed a 10% higher likelihood of withdrawing or failing compared to their more mature counterparts. The significance of modifying support systems to address the particular difficulties faced by younger learners in online education is highlighted by this research. Additionally, more than half (52.8%) of the students failed or dropped out of class. Figure 3 below illustrates this. Notably, certain courses, such as BBB and FFF, had greater rates of withdrawal and failure. BBB: 23.5% withdrawn and 25.1% failed, and FFF: 24.2% withdrawn and 23.6% failed. Course AAA has the lowest withdrawal and failure percent. 1.2% withdrawn and 1.3% failed. This emphasizes the significance of examining course design, content, and support methods in order to solve difficulties and improve student success rates [26].

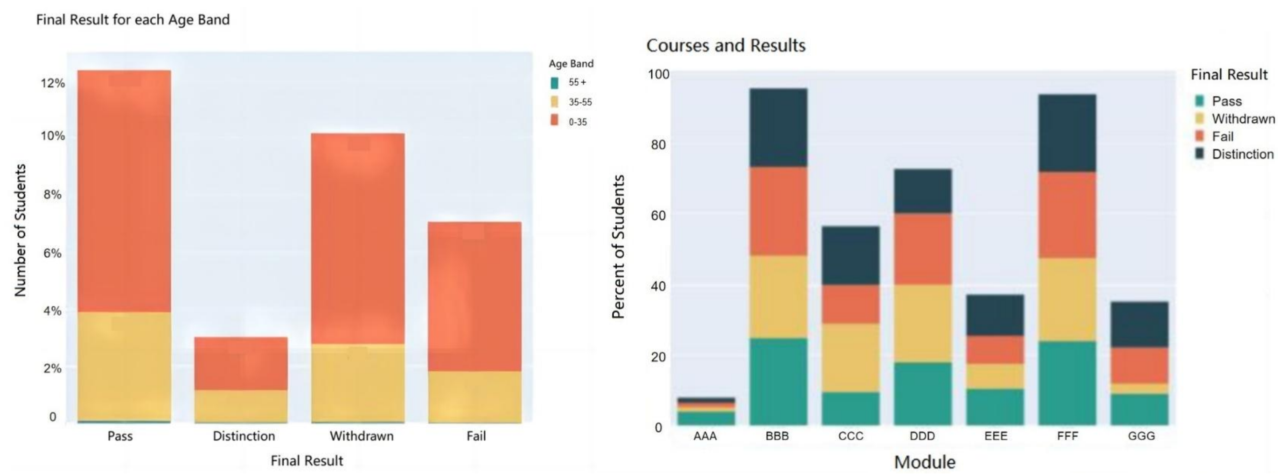

Figure 3. Graphical representation between age and final result.

Ethnic Disparities

The study discovered a significant achievement disparity between ethnic groupings. When compared to white students, black students had a 20% lower achievement rate. This raises concerns regarding the possible impact of institutional prejudices on minority students' academic development, emphasising the importance of inclusive and culturally responsive educational practices [27].

Impact of Disabilities

Students with impairments performed slightly worse than their non-disabled peers, with a 2% lower pass rate. This research emphasises the need of accessibility and accommodations in online learning settings for ensuring fair educational experiences for students with varying requirements [28].

Enrolment and Course Preferences

The majority of students were male (54.84%), with 45.16% female. In terms of age, the largest age band was 0 to 35 years old, with 68.34% of students within that band. A significant percentage (90.56%) did not report having a disability. Course enrolment trends varied, with course BBB attracting the most students (18.85%) and course AAA attracting the least (2.34%). This information can help course designers and educational institutions customise offers to student interests and needs.

Impact of Online Interactions

A significant positive correlation emerged between student interactions (clicks) on the online platform and their results. This underscores the importance of student engagement and active participation in online courses. However, it's important to note that while interactions are beneficial, weak correlations between clicks and scores suggest that success is not solely determined by the volume of interactions.

Model Biases and Predictive Analysis

Figure 4 shows the analysis of the logistic regression model's performance in predicting outcomes. Here, the axis values are mapped to {'Distinction': 0, 'Fail': 1, 'Pass': 2, 'Withdrawn': 3}. The logistic regression model revealed gender discrepancies in predicting outcomes. The model's overall accuracy was 38.52%, but the accuracy was lower for male students (37.65%) than for female students (39.6%), indicating potential biases within the predictive model. Similarly, students with impairments had greater rates of failure and withdrawal, highlighting the model's inadequacies in accounting for the requirements of varied student populations.

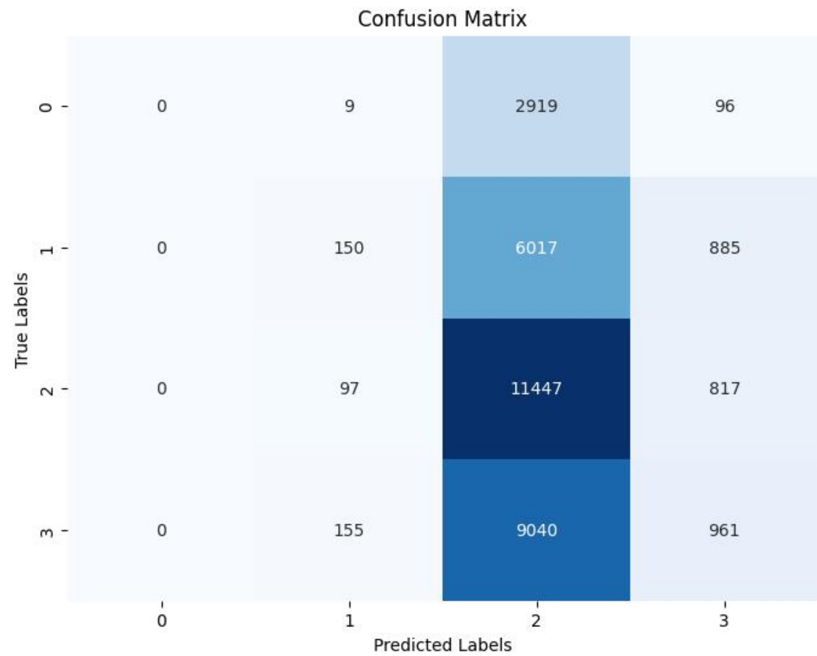

Figure 4. Confusion matrix for logistic regression.

Submission Timeliness and Academic Performance

The importance of following deadlines in achieving favourable academic outcomes was demonstrated by a high association between timely assignment submissions and final scores. This emphasises the importance of time management and discipline in achieving success in online learning. Feature importance was done using a random forest approach to identify the most important features, as illustrated in Figure 5. This predicted the target variable, and showed that the most effective predictive features are a number of previous attempts and being in the age band 0 to 35. Conversely, gender and older age (over 55) are less informative. These results could guide future data collection and model development to focus on the most influential.

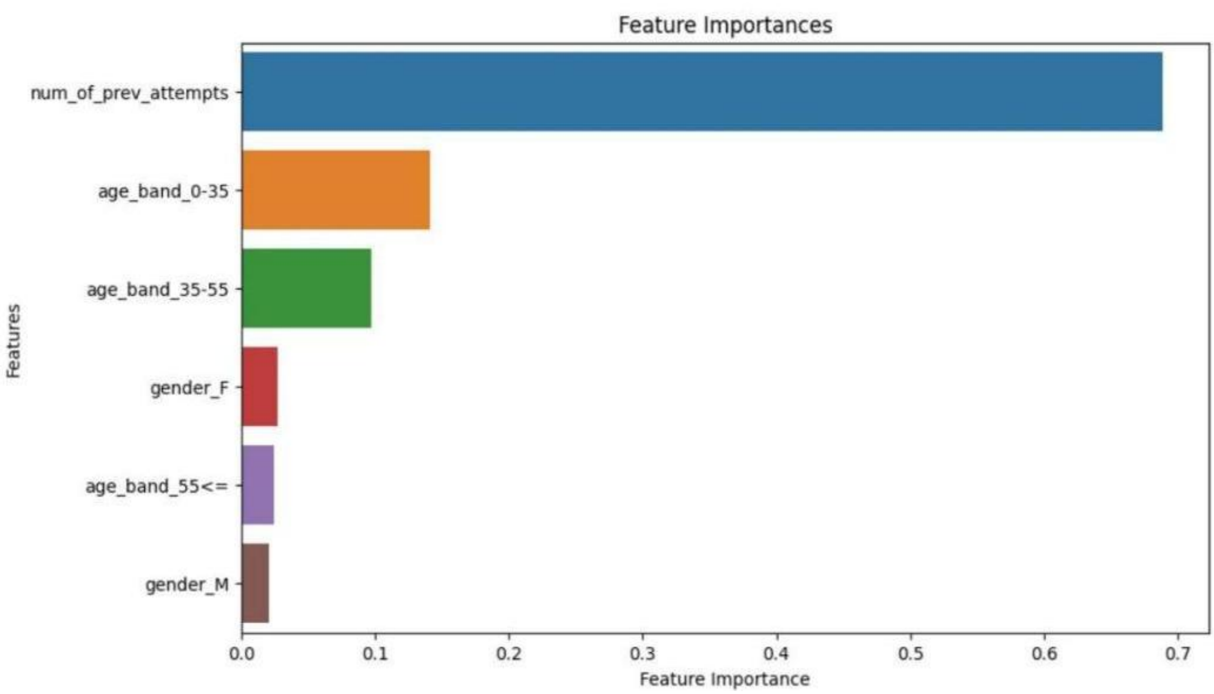

Figure 5. Plot showing feature importance in descending order.

Finally, the findings of the study provide a thorough insight into student demographics, course completion, and academic performance in online courses. To promote equitable and successful online education, the identification of biases and inequities necessitates targeted interventions and inclusive practices.

4. Discussion

The findings provide a comprehensive overview of the complex terrain of online education, shedding light on discrepancies in course completion and academic achievement across demographic and academic dimensions. This discussion goes deeper into the implications of these findings and how they correspond with previous literature, while also noting areas for improvement and future research concerns. The analysis revealed significant gaps in online course completion and academic performance based on gender, disability status, and prior academic achievement. Notably, female students outperformed males in assessments by around 5%, and younger students (0–35 years old) were 10% more likely to withdraw or fail than mature students. Additionally, Black students exhibited a 20% lower attainment rate compared to white students, and students with disabilities had a 2% lower pass rate than their peers. These findings underscore the presence of biases in the AI system's predictions, indicating potential systemic inequities.

Examining the broader demographic landscape, most students were female (45.16%), a majority of students fell within the 0–35 age band (68.34%), and a majority did not report having a disability (90.56%). The majority of students were also enrolled in their first attempt at the course (75.86%) and pursued education up to A levels or equivalent (43.21%). Notably, course BBB had the highest enrolment (18.85%), while course AAA had the lowest (2.34%). Over half of the students withdrew or failed courses (52.8%), with the highest rates observed in course BBB and course FFF. Moreover, students who engaged more in the online platform tended to achieve better final results.

While the statistical analysis and modelling offered valuable insights, certain disparities across student groups emerged. The accuracy of the logistic regression model was lower for predicting outcomes of male students (37.65%) compared to females (39.6%). Students reporting disabilities had lower failure and withdrawal rates (53.6%) than non-disabled peers (60.2%). Furthermore, strong correlations were observed between timely assignment submissions and final scores. In contrast, weak correlations between interactions (clicks) and scores suggested that interaction alone does not guarantee academic success.

The findings also highlighted some noteworthy factors that warrant further investigation. For instance, age did not emerge as a significant predictor of student success in this analysis. This contrasts with certain studies showing age-related performance differences in online education [29] but aligns with others finding minimal age -based performance differences [30].

A significant outcome of the analysis was the variation in pass rates across different courses, with modules BBB and FFF exhibiting particularly high withdrawal and failure percentages. This discrepancy suggests potential inconsistencies in the quality and suitability of course design, potentially disadvantaging students enrolled in modules with suboptimal design.

The strong positive association between student participation (clicks) and final grades corroborates prior research emphasizing the pivotal role of engagement for online achievement [31]. To enhance student involvement further, the incorporation of techniques like gamification and collaborative assignments could prove to be beneficial [32,33].

Limitations

While the research has produced useful insights into the relationship of AI, fairness, and student success, some limits must be addressed. To begin, the analysis mainly depends on the Open University Learning Analytics Dataset (OULAD), which may limit the implications of the findings to larger educational settings. Different institutions' unique student demographics, course structures, and teaching approaches may result in diverse patterns that are not reflected by this dataset.

Furthermore, the study focuses on specific demographic criteria such as gender, age, and disability status while ignoring other relevant variables such as race, language competence, socioeconomic level, and cultural diversity. This shortcoming may impede a thorough understanding of many influences on student outcomes.

Additionally, the study is based on data from a certain period (2013–2014), which may not fully depict the dynamic character of the evolving online education scene and the ongoing improvements in AI technologies. A multi-year longitudinal method could show temporal trends and changes in student behaviour and results.

The anticipated addition of qualitative insights via surveys or student interviews was hampered by time restrictions and the vastness of the Open University Learning Analytics Dataset (OULAD). As a result, this study focuses primarily on quantitative examination of demographic and performance data. Future studies with more time allocation could explore these qualitative features, enriching our understanding of AI, justice, and student achievement in online education. While making substantial contributions to this field of study, it is critical to recognise these limitations. Future research should aim to solve these limits to gain a fuller knowledge of the complex difficulties and opportunities presented by AI integration in education.

5. Conclusions

This study on online education demographics, completion rates, and academic performance reveals disparities based on gender, disability, and prior achievement. Greater student interaction correlates with higher grades, highlighting engagement's importance. Varied pass rates across courses suggest instructional design inconsistencies. Biases in completion rates underscore the need for targeted interventions, especially for male and disabled students. Standardized and evidence-based course design is crucial. Recommendations include inclusive design practices, tailored support services, and fostering student participation. While age isn't a significant predictor, recognizing study limitations is vital. Overall, this research contributes to understanding variables impacting student success in online education, aiming for a fairer and more efficient learning ecosystem. It also identifies some of the potential issues with the underlying data, including differences in the sample sizes by gender and disability, which indicate areas where bias in the models and outcomes may be influenced.

Our final recommendations are that institutions should prioritize evidence-based and inclusive course design to accommodate diverse learning needs. Tailored support services should target at risk groups, especially males and students with disabilities, fostering their success. Instructors should employ engaging techniques like gamification and collaborative assignments to enhance student participation. Recognizing the study's limitations, future research should incorporate qualitative input from students and encompass a broader range of demographic and contextual variables. This holistic approach will contribute to a fairer and more efficient online learning environment, enabling all learners to reach their full potential.

Author Contributions

Conceptualization, N.G. and D.O.; methodology, D.O.; software, D.O.; validation, D.O., N.G. and T.A.J.; formal analysis, D.O.; investigation, D.O.; resources, D.O.; data curation, D.O.; writing—original draft preparation, D.O.; writing—review and editing, N.G. and T.A.J.; visualization, D.O.; supervision, N.G. and T.A.J. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

The research was considered and approved by the Faculty of Science and Engineering Ethics Committee. This study did not involve humans or animals, but used published anonymized data.

Informed Consent Statement

Not applicable.

Data Availability Statement

The dataset used is available via the reference provided. Access to the software and code used in the research can be arranged by contacting the authors.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Zlatkin-Troitschanskaia, O.; Shavelson, R.J.; Pant, H.A. Assessment of Learning Outcomes in Higher Education. In *Handbook on Measurement, Assessment, and Evaluation in Higher Education*, 2nd ed.; Secolsky, C., Denison, D.B., Eds.; Routledge: New York, NY, USA, 2017; pp. 686–698. [CrossRef]
2. Michelsen, S.; Sweetman, R.; Stensaker, B.; Bleiklie, I. Shaping Perceptions of a Policy Instrument: The Political-Administrative Formation of Learning Outcomes in Higher Education in Norway and England. *Higher Educ. Policy* **2016**, *29*, 399–417. [CrossRef]
3. OECD. *Education at a Glance 2017: OECD Indicators*; OECD Publishing: Paris, France, 2017; pp. 1–456. [CrossRef]
4. Bird, E.; Fox-Skelly, J.; Jenner, N.; Larbey, R.; Weitkamp E.; Winfield, A. *The Ethics of Artificial Intelligence: Issues and Initiatives*; Panel for the Future of Science and Technology: Brussels, Belgium, 2020; pp. 1–128. [CrossRef]
5. Varona, D.; Suárez, J.L. Discrimination, Bias, Fairness, and Trustworthy AI. *Appl. Sci.* **2022**, *12*, 5826. [CrossRef]
6. Zhao, J.; Wang, T.; Yatskar, M.; Ordonez, V.; Chang, K. Men Also Like Shopping: Reducing Gender Bias Amplification Using Corpus-level Constraints. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 7–11 September 2017. [CrossRef]
7. Zhou, J.; Chen, F.; Holzinger, A. Towards Explainability for AI Fairness. In *xxAI-Beyond Explainable AI: International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers*, 1st ed.; Holzinger, A., Goebel, R., Fong, R., Moon, T., Müller, K.R., Samek, W., Eds.; Springer: Cham, Switzerland, 2020; pp. 375–386. [CrossRef]
8. Park, J.H.; Choi, H.J. Factors Influencing Adult Learners' Decision to Drop out or Persist in Online Learning. *J. Educ. Technol. Soc.* **2009**, *12*, 207–217.
9. Xu, D.; Jaggars, S. *Adaptability to Online Learning: Differences Across Types of Students and Academic Subject Areas*; Community College Research Center, Teachers College, Columbia University: New York, NY, USA, February 2013. [CrossRef]
10. AERA. Code of Ethics. *Educ. Res.* **2011**, *40*, 145–156. [CrossRef]
11. Akgun, S.; Greenhow, C. Artificial Intelligence in Education: Addressing Ethical Challenges in K-12 Settings. *AI Ethics* **2022**, *2*, 431–440. [CrossRef]
12. Holstein, K.; Vaughan, J.W.; Daumé, H.; Dudik, M.; Wallach, H. Improving Fairness in Machine Learning Systems: What do Industry Practitioners Need? In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019. [CrossRef]
13. Danijel, K.; Vedran, J.; Goran, D. Machine Learning in Education—A Survey of Current Research Trends. In Proceedings of the 29th DAAAM International Symposium on Intelligent Manufacturing and Automation, Zadar, Croatia, 24 – 27 October 2018. [CrossRef]
14. UNESCO. Beijing Consensus on Artificial Intelligence and Education. In Proceedings of the International Conference on Artificial Intelligence and Education, Planning Education in the AI Era: Lead the Leap, Beijing, China, 16–18 May 2019.
15. Pedro, F.; Subosa, M.; Rivas, A.; Valverde, P. *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*; UNESCO: Paris, France, 2019.
16. Ali, S., Payne, B.H.; Williams, R.; Park, H.W.; Breazeal, C. Constructionism, Ethics, and Creativity: Developing Primary and Middle School Artificial Intelligence Education. In Proceedings of the EduAI19: International Workshop on Education in Artificial Intelligence K-12, Macao, China, 28 April 2019.
17. Autonomous and Intelligent Systems. Available online: https://standards.ieee.org/initiatives/autonomous-intelligence-systems/ (accessed on 13 April 2023).
18. Kuzilek J.; Hlosta M.; Zdrahal, Z. Open University Learning Analytics Dataset. *Sci. Data* **2017**, *4*, 1–8. [CrossRef]
19. McKinney, W. Data Structures for Statistical Computing in Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; pp. 51–56. [CrossRef]
20. Harris, C.R.; Millman, K.J.; Van Der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array Programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef]
21. Brownlee, J. *Better Machine Learning: Model Evaluation, Hyperparameter Tuning, and Error Analysis with Scikit-Learn*. Machine Learning Mastery, 2020.

22. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830. [CrossRef]
23. Bayrak, T.; Gulati, A. The Role of Gender on Student Success. *Int. J. Inf. Commun. Technol. Educ.* **2015**, *11*, 1–17. [CrossRef]
24. Cavanaugh, J.K.; Jacquemin, S.J. A Large Sample Comparison of Grade Based Student Learning Outcomes in Online vs. Face-to-Face Courses. *Online Learn.* **2015**, *19*, 25–32. [CrossRef]
25. Vella, E.J.; Turesky, E.F.; Hebert, J. Predictors of Academic Success in Web-Based Courses: Age, GPA, and Instruction Mode. *Qual. Assur. Educ.* **2016**, *24,* 586–600. [CrossRef]
26. Eom, S.B.; Ashill, N.; Wen, H.J. The Determinants of Students' Perceived Learning Outcomes and Satisfaction in University Online Education: An Empirical Investigation. *Decis. Sci. J. Innovative Educ.* **2006**, *4*, 215–235. [CrossRef]
27. Patton, M.Q. A Transcultural Global Systems Perspective: In Search of Blue Marble Evaluators. *Can. J. Program Eval.* **2016**, *30*, 374–390. [CrossRef]
28. Burgstahler, S.; Moore, E. Making Student Services Welcoming and Accessible Through Accommodations and Universal Design. *J. Postsecond. Educ. Disability* **2009**, *21*, 155–174.
29. Gregori, P.; Martínez, V.; Moyano-Fernández, J.J. Basic Actions to Reduce Dropout Rates in Distance Learning. *Eval. Program Plann.* **2018**, *66*, 48–52. [CrossRef]
30. Amro, H.J.; Mundy, M.A.; Kupczynski, L. The Effects of Age and Gender on Student Achievement in Face-to-Face and Online College Algebra Classes. *Res. Higher Educ. J.* **2015**, *27*.
31. Henrie, C.R.; Halverson, L.R.; Graham, C.R. Measuring Student Engagement in Technology-Mediated Learning: A Review. *Comput. Educ.* **2015**, *90*, 36–53. [CrossRef]
32. Buckley, P.; Doyle, E. Gamification and Student Motivation. *Interact. Learn. Environ.* **2016**, *24*, 1162–1175. [CrossRef]
33. Subhash, S.; Cudney, E.A. Gamified Learning in Higher Education: A Systematic Review of the Literature. *Comput. Hum. Behav.* **2018**, *87*, 192–206. [CrossRef]