**Title:** Artificial Intelligence in Evaluating the Impact of Tuition Fees on Students' Academic Decision-Making: A Case Study

**Running title:** Artificial Intelligence and Tuition Fees Impact

**Authors:**

Ahmad Keykha1, Rita Mojtahedzadeh2, Abdalsamad Keramatfar3, Ava Taghavi Monfared4, Aeen Mohammadi2\*

1 Postdoctoral Researcher at Sharif University of Technology, Sharif Policy Research Institute, Tehran, Iran.

2 Department of E-learning in Medical Education, Center of Excellence for E-learning in Medical Education, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran.

3 NLP Engineer, Stinas company, Tehran, Iran.

4 Department of Educational Administration and Planning, Faculty of Psychology and Education, University of Tehran, Tehran, Iran.

**Correspondent Author:**

Aeen Mohammadi, MD, MPH, PhD

Department of E-learning in Medical Education, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

Dolatshahi Alley, Naderi St, Keshavarz Blvd, Tehran, Iran

PO Box: 14166-14741

ORCID: 0000-0002-2745-3873

Email: aeen\_mohammadi@tums.ac.ir

**Artificial Intelligence in Evaluating the Impact of Tuition Fees on Students' Academic Decision-Making: A Case Study**

**Abstract**

**Background:** Tuition fees are a key factor in students’ academic decision-making, significantly impacting dropout rates, major changes, guest study programs, and university transfers. This study aims to analyze these effects and develop predictive models for students’ educational choices.

**Methods:** Students’ data at Tehran University of Medical Sciences (TUMS) from 2016 to 2023 were collected and analyzed. Machine learning methods, including logistic regression, neural networks, decision trees, and random forests, were employed. Additionally, various techniques were utilized to address the challenge of imbalanced data.

**Results:** The results demonstrated that the selected models could predict students' academic decisions with relatively high accuracy. For instance, the random forest model achieved 74% accuracy in predicting dropout decisions, with an F1-score of 52% for the minority class.

**Conclusion:** The innovation of this study lies in its pioneering application of diverse advanced machine learning algorithms and data imbalance mitigation techniques to analyze students' academic decision-making. This approach, unprecedented in this research domain, significantly enhanced the accuracy and comprehensiveness of the findings.

**Keywords:** Tuition fees, Academic decision-making, Artificial intelligence, Medical education

**Introduction**

Higher education institutions (HEIs) play a pivotal role in human and economic development by cultivating students as valuable human capital (1). This role has become particularly pronounced in developing nations amid the transition to knowledge-based economies. Such economies, grounded in knowledge production, drive job creation and societal prosperity. Within this framework, student performance holds significant developmental importance, as empirical evidence demonstrates that knowledge-based human capital plays a crucial role in strengthening national economic growth (2). However, financial crises and the recent COVID-19 pandemic have substantially impacted higher education financing (3). Concurrent with declining public funding, instructional costs have risen markedly, with notable global disparities in educational expenditures (4). In this context, tuition fees have emerged as a primary revenue stream for universities (5).

However, tuition fees exhibit paradoxical effects on students' academic decisions. Some studies demonstrate that tuition increases lead to enrollment declines, while others find negligible impacts (6). Similarly, financial aid and loans yield divergent outcomes: certain studies report positive effects, whereas others detect no significant influence (7). This complexity has prompted conceptual frameworks like the "tuition dilemma" and "tuition paradox" (8) to explain these contradictory phenomena. Numerous studies have investigated the application of artificial intelligence algorithms in predicting student dropout decisions. Mubarak et al. (2022) achieved 84% accuracy in predicting early student attrition in online learning environments using AI models (9). Gismondi and Huisman (2021) employed multilayer neural networks, reporting 98.97% prediction accuracy for student dropout at the National University of San Pedro (10).

Agrusti et al. (2020) achieved 93.4% accuracy in predicting student dropout by implementing deep neural networks (11). Kemper et al. (2020) demonstrated that decision tree algorithms outperformed logistic regression in dropout prediction, attaining 95% accuracy (12). Similarly, Behr et al. (2020) utilized random forest models to obtain 86% accuracy in early dropout prediction (13). Haiyang et al (2020) employed a time-series classification algorithm, reporting 84% accuracy in predicting student dropout decisions (14). Limsathitwong et al. (2018) emphasized that decision trees demonstrated superior performance among AI algorithms for dropout prediction, achieving 80% accuracy (15). Furthermore, Berens et al. (2018) demonstrated that dropout prediction accuracy reached 90% in public universities after the fourth semester and 95% in universities of applied sciences (16). Despite the significant impact of tuition fees on educational decisions, the authors of this article found no studies on this subject in Iran. Therefore, this study aims to employ AI algorithms to evaluate the impact of tuition fees on students' academic decisions. The study population consists of students from Tehran University of Medical Sciences (TUMS). Established in 1934, TUMS has consistently ranked first among Iranian medical universities in the Ministry of Health and Medical Education rankings. The university comprises 12 faculties across various medical disciplines.

The innovation of this study lies in its application of advanced AI algorithms and comparative analysis of multiple models' accuracy to evaluate the impact of tuition fees on academic decision-making among students at TUMS. Unlike previous studies that primarily focused on predicting dropout rates through individual or environmental factors, this research adopts a novel approach by examining the relationship between tuition fees and academic choices. By utilizing real-world data from TUMS and focusing on a key variable such as tuition fees, this study contributes new dimensions to the existing literature.

**Methods**

In this study, we evaluated the impact of tuition fees on students’ academic decisions using different AI algorithms.

**A) Data Collection**

The study encompassed the entire student population of TUMS across twelve schools from 2016 to 2023, resulting in an eight-year longitudinal dataset. This reflects the university's institutional data collection framework. The demographic distribution across academic units was as follows:

International Campus (2,822), School of Public Health (1,917), School of Nursing and Midwifery (2,460), School of Medicine (9,326), School of Rehabilitation Sciences (954), School of Pharmacy (1,468), School of Dentistry (1,276), School of Traditional Iranian Medicine (121), School of Paramedical Sciences (1,683), School of Nutrition and Dietetics (304), School of Advanced Medical Technologies (310), and School of Research and Technology (60). The gender distribution indicated that there were 9,868 male students (43.5%) and 12,832 female students (56.5%), with a total of 22,700 enrolled students during the study period.

**B) Variables**  
The independent variables of the study are detailed in Table 1.

**Table 1.** Independent Variables and Their Characteristics

|  |  |  |
| --- | --- | --- |
| Variable | Type | Description |
| Degree | Discrete (Nominal) | Academic programs: Associate Degree, Bachelor's (Continuous/Discontinuous), Master's (MSc, MPH), Professional Doctorates (MD, DDS, PharmD), MBBS, PhD, Residency/Fellowship |
| School | Discrete (Nominal) | Twelve constituent schools: International Campus, Public Health, Nursing & Midwifery, Medicine, Rehabilitation, Pharmacy, Dentistry, Iranian Medicine, Paramedical Sciences, Nutrition & Dietetics, Advanced Medical Technologies, Research and Technology |
| Age | Continuous | Student age in years |
| Course Type | Discrete (Nominal) | Tuition classification: • Tuition-paying (lower entrance exam ranks) • Non-tuition-paying (higher ranks, government-funded) |
| Nationality | Discrete (Nominal) | Categories: Iranian, Asian, European, Oceanian, African, American |
| Marital Status | Discrete (Binary) | Single, Married |
| Number of Children | Continuous | Count of children |
| Year | Discrete (Ordinal) | Academic years 2016–2023 |
| Birth Province | Discrete (Nominal) | Categorized by Iran’s official provincial borders |
| Semester | Discrete (Ordinal) | Biannual admissions (Fall/Spring) |
| Probationary Semesters | Continuous | Semesters with a GPA below the minimum threshold (e.g., <12/20 for undergraduates) |
| Native Status | Discrete (Binary) | Tehran residents vs. non-residents |
| Gender | Discrete (Binary) | Female, Male |
| GPA | Continuous | Grade Point Average (0–20 scale) |

The dependent variables include academic decisions: the decision to drop out (binary), the decision to become a guest student (binary), the decision to transfer, and the decision to change fields (binary).

**C) Evaluation Metrics**

Three metrics were used to evaluate the performance of the regression model: mean absolute error (MAE), root mean square error (RMSE), and mean square error (MSE). These metrics are defined as the average differences between the predicted and actual values.

Where y\_pred represents the predicted value, y\_actual refers to the actual value, and n is the number of data points. MAE measures the average of the absolute differences between the predicted and actual values:

Where |x| denotes the absolute value of x.

MSE measures the average of the squared differences between the predicted and actual values:

The square root of the MSE is the RMSE:

The RMSE metric offers superior interpretability compared to MSE as it maintains the same units as the dependent variable, facilitating direct comparison with observed values. While lower values of both metrics generally indicate better model performance, metric selection should ultimately align with the study's specific objectives and analytical requirements.

**D) Algorithms Used**

D-1) Logistic Regression

Logistic regression is a statistical modeling technique used to examine the relationship between a categorical dependent variable (e.g., the presence or absence of a condition) and one or more independent variables. Also referred to as logistic or logit models, this approach is widely applied in various domains, including medical diagnosis and predictive analytics (17).

D-2) Artificial Neural Network (ANN)  
ANNs are powerful modeling and prediction tools that learn system behavior from representative data without requiring explicit quantification of physical properties or strict mathematical formulations. This data-driven approach helps reduce uncertainties while enhancing predictive accuracy and decision-making efficiency (18).

D-3) Decision Tree  
A decision tree is a machine learning technique used for both classification and regression tasks. It offers clear interpretability, accommodates various data formats (such as numerical and categorical values), and generates predictions by formulating decision rules from the training dataset. Its hierarchical structure enables a transparent approach to understanding how decisions are made (19).

D-4) Random Forest (RF):  
RF is an ensemble learning technique that aggregates multiple decision trees, each built with random variations, to perform classification or regression tasks. Its inherent randomness helps mitigate overfitting, as every tree independently contributes unique insights to the overall prediction (20).

Given the imbalanced nature of the dataset, we employed multiple strategies to mitigate class imbalance. In classification problems where one class (the minority class) is significantly underrepresented compared to the other (the majority class), predictive models often exhibit poor performance for the minority class. The following approaches were implemented in this study to address class imbalance:

SMOTE (Synthetic Minority Oversampling Technique) is a technique designed to mitigate class imbalance within datasets. It operates by producing artificial samples for the underrepresented class, selecting random instances, and applying interpolation to generate new data points. This method helps prevent overfitting from duplicate samples and enhances classifier decision boundaries.

SMOTE Tomek is a combined resampling method designed to handle class imbalance in datasets by integrating SMOTE and Tomek links. Initially, SMOTE creates synthetic samples to enhance the representation of the minority class. Afterward, Tomek's links identify overlapping nearest-neighbor pairs from different classes, allowing for the removal of ambiguous instances to refine the dataset.

SMOTEENN (Synthetic Minority Oversampling Technique and Edited Nearest Neighbors) is a combined approach that integrates oversampling and data refinement to tackle class imbalance in datasets. It begins with SMOTE, which generates artificial samples to enhance the minority class representation. Following this, Edited Nearest Neighbors (ENN) refines the dataset by eliminating uncertain or misclassified instances from both classes, based on their surrounding data points.

Random under-sampling is a method for managing class imbalance by decreasing the number of samples in the dominant class. It achieves this by randomly selecting a portion of the majority class instances to align with the minority class size or a predetermined ratio, resulting in a more evenly distributed dataset.

**Ethical consideration**

The Ethics Committee of the TUMS approved the study, ethical code number (IR.TUMS.MEDICINE.REC.1402.691). All methods were carried out following relevant guidelines and regulations. The confidentiality of the participants' information was assured.

**Results**

For each academic decision category (dropout, guest student transfer, major change, and university transfer), multiple models were developed. Three top-performing models were selected for each decision based on rigorous evaluation criteria. Below, we detail the selected models for each decision category.

A) Dropout Prediction

Three machine learning models were evaluated to predict student dropout at TUMS. Performance was compared using accuracy, precision, recall, and F1-score. While all models were trained on the same data, Model 3 performed best with 74% overall accuracy and 52% F1-score for minority class (dropout) prediction. This model provides more accurate identification of at-risk students, enabling timely interventions to reduce dropout rates. The analysis used random under-sampling to address class imbalance. The model provides more accurate predictions and improves student status identification. It effectively identifies students at risk of dropping out, enabling universities to implement timely intervention programs and reduce dropout rates. These analyses used random under sampling to address class imbalance.

**Table 2: The Selected Model for Predicting Student Dropout at TUMS**

|  |  |  |
| --- | --- | --- |
| Sample | Criteria | Value |
| 6811 | Accuracy | 0.74 |
| 6811 | Macro Avg | 0.55 (Precision), 0.73 (Recall), 0.52 (F1-Score) |
| 6811 | Weighted Avg | 0.94 (Precision), 0.74 (Recall), 0.81 (F1-Score) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| support | f1-score | recall | precision | Class |
| 6495 | 0.84 | 0.74 | 0.98 | 0 |
| 316 | 0.20 | 0.73 | 0.12 | 1 |

Table 2 shows the selected model for predicting student dropout at TUMS. According to this table, the model's accuracy is 74%, and the macro average performance for the class of students’ dropout (class 1) is as follows: precision of 0.55, recall of 0.73, and F1-score of 0.52. For the model's weighted average performance, the precision is 0.94, the recall is 0.74, and the F1-score is 0.81. The model's accuracy of 74% means that out of every 100 predictions made by the model, 74 are correct. This number indicates that the model has performed relatively well in identifying the overall status of students.

B) Major Change Prediction

Three distinct machine learning models were developed and evaluated to predict major change decisions among students at TUMS. Based on the results, one of the models is selected as the best due to its very high overall accuracy (96%) and superior performance compared to the other models. Although performance in the minority class can still be improved, this model has performed stronger in providing overall predictions for the decision to change majors. The selected model can be used to identify students who are likely to change their major. This analysis was conducted using the SMOTE technique. (Table 3)

**Table 3: The Selected Model for Predicting Major Change Among Students at TUMS**

|  |  |  |
| --- | --- | --- |
| Sample | Criteria | Value |
| 6811 | Accuracy | 0.96 |
| 6811 | Macro Avg | 0.50 (Precision), 0.48 (Recall), 0.49 (F1-Score) |
| 6811 | Weighted Avg | 1.00 (Precision), 0.96 (Recall), 0.98 (F1-Score) |

|  |  |  |  |
| --- | --- | --- | --- |
| support | f1-score | recall | precision |
| 6808 | 0.98 | 0.96 | 1.00 |
| 3 | 0.00 | 0.00 | 0.00 |

According to Table 3, the model's accuracy is 96%, and its macro average performance for the class of students who have decided to change their major (class 1) is as follows: precision of 0.50, recall of 0.48, and F1-score of 0.49. For the model's weighted average performance, the precision is 100%, the recall is 0.96, and the F1-score is 0.98. The model's precision for this class is 50%, meaning that half of the predictions regarding students who intend to change their major were correct. The recall for this class is 48%, indicating that the model correctly identified only 48% of the students who decided to change their major. The F1-score for this class is 49%, reflecting a relatively weak balance between precision and recall, as the model has a low precision while maintaining a recall close to 50%.

C) Guest Student Enrollment Prediction

To analyze the decision to enroll as a guest student, three different statistical models were fitted, evaluating their characteristics and performance in both the minority class (students deciding to enroll as guests) and overall performance. The selected model, in addition to having a very high overall accuracy (99%), demonstrates a balanced performance in identifying the minority class (F1-score = 82%). Furthermore, the precision, recall, and F1-score for the minority class are superior and more reliable compared to the other models. The third model is identified as the most suitable for predicting guest enrollment decisions among students at TUMS due to its exceptionally high overall accuracy (99%) and strong performance in identifying students who do not intend to enroll as guests (majority class). This model is well-suited for monitoring the general status of students. The analysis was conducted using the SMOTEENN technique (Table 4).

**Table 4: The Selected Model for Predicting Guest Student Enrollment Among Students at TUMS**

|  |  |  |
| --- | --- | --- |
| Sample | Criteria | Value |
| 6811 | Accuracy | 0.99 |
| 6811 | Macro Avg | 0.84 (Precision), 0.80 (Recall), 0.82 (F1-Score) |
| 6811 | Weighted Avg | 0.99 (Precision), 0.99 (Recall), 0.99 (F1-Score) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| support | f1-score | recall | precision | Class |
| 6717 | 1.00 | 1.00 | 0.99 | 0 |
| 94 | 0.64 | 0.61 | 0.68 | 1 |

According to Table 4, the model's accuracy is 99%, and its macro average performance for the class of students who have decided to enroll as guest students (class 1) is as follows: precision of 0.84, recall of 0.80, and F1-score of 0.82. For the model's weighted average performance, the precision is 0.99, the recall is 0.99, and the F1-score is 0.99.

D) Student Transfer Prediction

The selected model for predicting student transfers at TUMS has an accuracy of 68% and outperforms the second model in identifying transfer students (class 1). This model is useful for broader analyses and prioritizing transfer programs. Due to its higher accuracy in the majority class (class 0) and relatively appropriate recall in class 1, it can be utilized for predicting infrastructure needs and assessing the impact of changes in university capacity. However, the low F1-score for the transfer class (42%) indicates that the model requires improvements for applications that demand more precise identification of transfer students. This analysis was conducted using the SMOTE Tomek technique (Table 5).

**Table 5: The Selected Model for Predicting** **Student Transfers at TUMS**

|  |  |  |
| --- | --- | --- |
| Sample | Value | Criteria |
| 6811 | Accuracy | 0.68 |
| 6811 | Macro Avg | 0.51 (Precision), 0.65 (Recall), 0.42 (F1-Score) |
| 6811 | Weighted Avg | 0.99 (Precision), 0.68 (Recall), 0.80 (F1-Score) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| support | f1-score | recall | precision | Class |
| 6747 | 0.81 | 0.68 | 0.99 | 0 |
| 64 | 0.03 | 0.61 | 0.02 | 1 |

According to Table 5, the model's accuracy is 68%, and its macro average performance for the class of students who have decided to transfer (class 1) is as follows: precision of 0.51, recall of 0.65, and F1-score of 0.42. For the model's weighted average performance, the precision is 0.99, the recall is 0.68, and the F1-score is 0.80.

**Discussion and Conclusion**

This study explored the applications of AI in evaluating the impact of tuition fees on students' academic decisions at TUMS. These decisions included dropout, guest enrollment, major change, and transfer. Various models were fitted for each of these decisions, followed by a comparative analysis of the models. The third model outperforms the other two across all evaluation metrics. Notably, a recall of 73% and an F1-score of 52% for Class 1 highlight its superior ability to identify students at risk of dropout. Empirical evidence supports the relationship between tuition fees and student dropout. Bäulke et al. (2022) concluded that financial resources are among the most critical factors influencing students' academic persistence (21). Similarly, Huo et al. (2022), in their analysis of the demographic characteristics of students who opted to drop out, identified the inability to pay tuition fees as a significant determinant of this decision (22). The findings of Stoyanova and Goranova (2021) also indicate that most students perceive the current tuition fees as excessively high. Moreover, a tuition increase of more than 30% is projected to lead to a 68.8% decrease in the number of students willing to continue their studies, ultimately resulting in their decision to drop out (23).

Regarding the relationship between the decision to transfer, some studies have found that the type of academic program (school) is also influential in the decision to transfer (24). In another study, the decision to transfer across five schools: Business, Education, Fine Arts and Humanities, Mathematics and Science, and Social Sciences was assessed. They concluded that students in Education, Fine Arts and Humanities, and Social Sciences who decide to transfer experience an improvement in their GPA after the transfer. However, students transferring from the Business and Mathematics, and Science fields face a transfer shock and achieve lower GPAs. Additionally, their evidence showed that female students are less likely than their male counterparts to earn a degree in engineering after transferring (25, 26).

**Conclusion:**

In this study, various models were evaluated to predict students' academic decisions, including dropout, guest enrollment, major change, and transfer. The evaluation criteria included accuracy, recall, precision, and F1-score. The primary objective was to select models with the highest efficiency for both the minority class (students who made specific academic decisions) and the majority class. These models were selected based on the balance among evaluation metrics and their ability to accurately identify the minority class. To address data imbalance challenges, various techniques such as SMOTE, SMOTEENN, and Random Under Sampling were employed. These methods enhanced data balance and improved model performance in minority classes. The final model selection was based on criteria such as predictive accuracy, balance among evaluation metrics, and alignment with the research objectives.

This study demonstrated that tuition fees play a significant role in students' academic decisions. The selected models indicated that an increase in tuition fees could lead to higher dropout and major change rates, while in some cases, students may opt for guest enrollment or transfer to reduce costs. Overall, these findings suggest that tuition policy should be carefully designed based on analytical data to prevent negative impacts on students' academic experiences. For universities, these insights can aid in the early identification of students at risk of dropping out or changing their majors, enabling the provision of financial support and counseling services. Additionally, designing tuition plans that align with students' financial capacities can help mitigate adverse effects.

**Limitations**

Despite its innovations, this study has certain limitations. First, the data used are exclusively from students at TUMS, which may limit the generalizability of the findings to other universities or countries. Second, data imbalance and the small sample size in specific categories (e.g., transfer students) posed challenges in prediction accuracy. Additionally, the study did not account for external factors such as macroeconomic conditions, government policies, and inflation rates, which may also influence students' academic decisions.

**Research Recommendations**

Future studies are encouraged to incorporate macroeconomic factors such as inflation rates, unemployment, and government support policies into their analyses. Additionally, a comparative study between Iranian and international universities could provide valuable insights for policymaking, helping to develop more effective and data-driven strategies.

**Ethical Considerations**

Compliance with ethical guidelines

The Ethics Committee of the TUMS approved the study, ethical code number (IR.TUMS.MEDICINE.REC.1402.691). All methods were carried out following relevant guidelines and regulations. The confidentiality of the participants' information was assured.

**Funding**

This research was supported by the Tehran University of Medical Sciences and Health Services (Grant No: 1402-4-101-69607). The funders had no role in study design, data collection and analysis, the decision to publish, or preparation of the manuscript.

**Authors' contributions**

All authors contributed equally to the conception and design of the study, data collection and analysis, interpretation of the results, and drafting of the manuscript. Each author approved the final version of the manuscript for submission.

**Conflict of interest**

The authors declared no conflict of interest.

**Acknowledgments**

The authors would like to thank the Vice Chancellor of Education Affairs of Tehran University of Medical Sciences for providing secure access to the data.

**References**

1. Rehman WU, Degirmen S, Waseem F. Propensity for and quality of intellectual capital divulgence across the BRICS banking sector: A knowledge-based perspective from emerging economies. Journal of the Knowledge Economy. 2021:1-28.

2. Ngepah N, Saba CS, Mabindisa NG. Human capital and economic growth in South Africa: A cross-municipality panel data analysis. South African Journal of Economic and Management Sciences. 2021;24(1):1-1.

3. Civera A, Cattaneo M, Meoli M, Paleari S, Seeber M. Universities’ responses to crises: The influence of competition and reputation on tuition fees. Higher Education. 2021 Jul;82:61-84.

4. Mughan S, Sherrod Hale J, Woronkowicz J. Build it and will they come?: the effect of investing in cultural consumption amenities in higher education on student-level outcomes. Research in Higher Education. 2022 Feb;63(1):60-91.

5. Jaquette O, Kramer DA, Curs BR. Growing the pie? The effect of responsibility center management on tuition revenue. The Journal of Higher Education. 2018 Sep 3;89(5):637-76.

6. Brooks R, Waters J. Fees, funding and overseas study: mobile UK students and educational inequalities. Sociological Research Online. 2011 Jun;16(2):19-28.

7. Lan Y, Winters JV. Did the DC tuition assistance grant program cause out-of-state tuition to increase. Economics Bulletin. 2011 Aug 29;31(3):2444-53.

8Heller DE, editor. The states and public higher education policy: Affordability, access, and accountability. JHU Press; 2001.‏

9. Mubarak AA, Cao H, Zhang W. Prediction of students’ early dropout based on their interaction logs in online learning environment. Interactive Learning Environments. 2022 Jul 4;30(8):1414-33.

10. Gismondi HE, Huiman LV. Multilayer neural networks for predicting academic dropout at the national university of santa-peru. In 2021 International Symposium on Accreditation of Engineering and Computing Education (ICACIT) 2021 Nov 4 (pp. 1-4). IEEE.

11. Agrusti F, Mezzini M, Bonavolontà G. Deep learning approach for predicting university dropout: A case study at Roma Tre University. Journal of e-learning and knowledge society. 2020 May 7;16(1):44-54.

12. Kemper L, Vorhoff G, Wigger BU. Predicting student dropout: A machine learning approach. European Journal of Higher Education. 2020 Jan 2;10(1):28-47.

13. Behr A, Giese M, Theune K. Early prediction of university dropouts–a random forest approach. Jahrbücher für Nationalökonomie und Statistik. 2020 Dec 1;240(6):743-89.

14. Haiyang L, Wang Z, Benachour P, Tubman P. A time series classification method for behaviour-based dropout prediction. In 2018 IEEE 18th international conference on advanced learning technologies (ICALT) 2018 Jul 9 (pp. 191-195). IEEE.

15. Limsathitwong K, Tiwatthanont K, Yatsungnoen T. Dropout prediction system to reduce discontinue study rate of information technology students. In2018 5th International Conference on Business and Industrial Research (ICBIR) 2018 May 17 (pp. 110-114). IEEE.

16. Berens J, Schneider K, Görtz S, Oster S, Burghoff J. Early detection of students at risk. Predicting Student Dropouts Using Administrative Student Data and Machine Learning Methods, CESIFO WP. 2018;7259..‏

17. Nick TG, Campbell KM. Logistic regression. Topics in biostatistics. 2007:273-301.

18. Coppola Jr EA, Rana AJ, Poulton MM, Szidarovszky F, Uhl VW. A neural network model for predicting aquifer water level elevations. Groundwater. 2005 Mar;43(2):231-41.

19. El Massari H, Gherabi N, Mhammedi S, Ghandi H, Qanouni F, Bahaj M. An ontological model based on machine learning for predicting breast cancer. International Journal of Advanced Computer Science and Applications (IJACSA). 2022;13(7):108-15.

20. Su Y, Weng K, Lin C, Zheng Z. An improved random forest model for the prediction of dam displacement. Ieee Access. 2021 Jan 6;9:9142-53.

21. Bäulke L, Grunschel C, Dresel M. Student dropout at university: A phase-orientated view on quitting studies and changing majors. European Journal of Psychology of Education. 2022 Sep;37(3):853-76.

22. Huo Y, Messenger RA, Miller D. Students' perspectives on why they drop out and possible retention strategies. Higher Education, Skills and Work-Based Learning. 2022 Oct 12;12(5):849-65.

23. Stoyanova R, Goranova S. The influence of a tuition fee increase on the drop-out rate of the nursing program. Journal of Economy Culture and Society. 2021 Jan 1(63):55-66.

24. Doyle WR. Effect of increased academic momentum on transfer rates: An application of the generalized propensity score. Economics of Education Review. 2011 Feb 1;30(1):191-200.

25. Mau WC. Characteristics of US Students That Pursued a STEM Major and Factors That Predicted Their Persistence in Degree Completion. Universal Journal of Educational Research. 2016;4(6):1495-500.

26. Wolniak GC. Examining STEM bachelor’s degree completion for students with differing propensities at college entry. Journal of College Student Retention: Research, Theory & Practice. 2016 Nov;18(3):287-309.