Online-Journals.org

iJET International Journal of Emerging Technologies in Learning

iJET | eISSN: 1863-0383 | Vol. 18 No. 17 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijet.v18i17.41449

PAPER

Design of a Machine Learning Model to Predict Student Attrition

Tibor Fauszt[1](✉), Katalin Erdélyi[1], Dóra Dobák[1], László Bognár[2], Endre Kovács[1]

[1]Budapest Business School University of Applied Sciences, Budapest, Hungary

[2]University of Dunaújváros, Dunaújváros, Hungary

fauszt.tibor@uni-bge.hu

ABSTRACT

Higher education institutions are facing a major issue with student dropout rates, which is a global phenomenon that affects a significant portion of enrolled students, particularly those in their first year. The challenge is how to retain students who do not meet requirements during their first year and are at high risk of dropping out, which can have significant economic and social consequences as well as personal ramifications for the students themselves. Universities must prioritize identifying at-risk students and providing targeted assistance to prevent them from leaving the system. Machine learning (ML) models have proven effective in identifying students at risk of dropping out with a high degree of accuracy. In this study, we aim to construct a machine learning model using data extracted from the administration system (Neptun) to predict student dropout rates in the Business Informatics BSc course at the Faculty of Finance and Accounting of Budapest Business School.

KEYWORDS

student dropout, learning analytics, machine learning (ML)

1 INTRODUCTION

The BSc. in Business Informatics, offered by the Faculty of Finance and Accounting at Budapest Business School, was introduced in the academic year 2011–12 with an initial enrollment of 149 full-time students. Since its inception, the program has been consistently popular among students, with a steady increase in the number of both full-time and part-time students over the years. While approximately 400 students begin the program each year, only approximately half complete it, with the other half dropping out of the university. This dropout rate is not exceptional and is on par with the average rate for higher education institutions in Hungary [24]. However, reducing the dropout rate is a top priority for universities, as it is an important indicator of educational quality [17] [19] [22]. It is crucial to implement measures to prevent students from dropping out without compromising the quality of education.

Fauszt, T., Erdélyi, K., Dobák, D., Bognár, L., Kovács, E. (2023). Design of a Machine Learning Model to Predict Student Attrition. *International Journal of Emerging Technologies in Learning (iJET)*, 18(17), pp. 184–195. https://doi.org/10.3991/ijet.v18i17.41449

Article submitted 2023-05-16. Revision uploaded 2023-06-16. Final acceptance 2023-06-16.

Improving student performance is vital and should be the focus of interventions aimed at reducing dropout rates while meeting training and output requirements.

Defining the concept of dropout is challenging, and its lack of clarity has been a recurring theme in research [20] [25]. Broadly speaking, it refers to instances where a student exits higher education without earning a degree, either voluntarily or due to institutional factors [9]. However, leaving an educational program does not necessarily equate to ending one's higher education studies altogether, as the student may opt to enroll in another institution or course of study or pursue further education abroad. For the purposes of this study, dropouts are defined as those students who withdraw from the Business Informatics program without completing their studies.

Student dropouts can occur due to various factors, including family background, inadequate social connections, employment, poor relationships, lack of motivation, insufficient persistence or skills, previous academic struggles, stress, disengagement, self-evaluation issues, poor academic performance, financial hardships, unexpected external problems, classroom environment, teacher engagement, and course material, among other factors [1] [6] [7] [11] [18]. While these factors may indicate a student's risk of dropping out, gathering data, particularly sensitive data such as parental educational background and financial status, can be problematic. Additionally, certain indicators of dropout, such as motivation and engagement, are difficult to quantify. It is essential to note that these factors alone may not result in dropout, as the phenomenon is complex and multifaceted.

Student performance is a significant indicator of dropout [2]. Many studies, both national and international, have shown that students who perform poorly in their first semester are at a higher risk of dropping out, especially among first-year students [8] [12]. Therefore, it is crucial to identify students who are at risk of dropping out as early as possible. Once a student has decided to leave their studies, it becomes difficult to retain them. Early identification can take place even before the start of studies by analyzing student data from the time of admission using statistical procedures. However, these data may be sensitive and difficult to collect or measure, as mentioned earlier. Another way to identify students at risk is to monitor their activity and performance from the beginning of their studies. One effective method is to use a specific learning management system (LMS) such as Moodle and its machine learning (ML) models, to identify students who are lagging. It is possible to detect students who are not completing courses with high accuracy within a few weeks of starting their studies [4].

The cause of dropout can be difficult to pinpoint; however, it can be identified by monitoring student performance. The key to successful completion of a course is active learning and consistently good performance. Poor academic performance can indicate a student at risk of dropping out or a student who may lose interest in further study, even if there are no particular reasons for leaving the course. It is important to focus on student engagement and intervention in the learning process to help reduce drop-out rates and ensure successful course completion.

2 A MACHINE LEARNING MODEL TO IDENTIFY STUDENTS AT RISK OF DROPPING OUT

As higher education becomes increasingly populated, tracking the performance of each student and identifying students at risk of failing or dropping out is almost impossible without IT systems. However, ML methods, which can be considered part of artificial intelligence (AI), can help address this problem [13–16] [21] [23] [26]. Predictive models can be used to effectively identify students at risk of dropping out, but this is only possible with well-designed models. To build an effective model, it is not enough to simply feed students' data into the model; this is a much more complex task, and many factors need to be considered [3]. The basic requirement for building models is that data be available in bulk. These include designing the indicators that are most correlated with the target function, determining the most efficient model structure, and ensuring that the models are not sensitive to changes that may occur (e.g., curriculum changes). Before building models, it is useful to conduct a learning analysis to identify indicators of dropout. Learning analytics (LA) is the collection and analysis of data on learners and their environments to understand learning processes and improve learner outcomes. LA is a multidisciplinary field that includes ML, AI, information retrieval, statistics, and data visualization [5].

The basis of effective learning and teaching is the precise definition of the requirements for the fulfillment of the subject [10]. The performance can be measured based on the requirements, which are generally measured on a five-point scale with grades in Hungarian higher education. The collection of data (grades) and the definition of indicators of student performance are simple in this case since the academic results are available in the student administration system (Academic Management System) (Neptun). Well-designed ML models can be built based on student performance indicators. With these models, student attrition can be predicted.

2.1 Data cleaning, statistical analysis, feature definition

For the model building, data extracted from the student administration system (Neptun Student Administration System) was used. These were data that were available for all students and that, according to Demcsákné and Huszárik's study "Attrition Studies in Higher Education," are the most prominent indicators of dropout in Hungary. According to the study, these are student demographic characteristics (gender, age), education characteristics (number of passive semesters, work schedule, type of financing), and regional characteristics (region of residence) [9]. Data were taken from the Neptun system from the first semester of 2011 to the second semester of 2022 in Excel spreadsheets.

Each table contained the following information: date of birth, gender, place of residence (city), degree grade, admission score, date of admission to the university, financial status (fee-paying, self-paying), number of active semesters, number of passive semesters, student status (in, out), and degree result. Next was the list of courses taken by the student in the semester, which included the following information for each course: course code, credit of the course, total number of times the student enrolled in the course, semester the course was taken, course requirement (teachers' signature, term mark, exam mark), and whether the course was passed or not.

The data was aggregated and then cleaned. In the first step, the records of students whose status were inactive, i.e., they had left the university with or without a degree, were deleted. The records of students whose date of admission had multiple values were deleted. This special case can occur when a student is admitted in a given academic year, has an active status for a period of time, completes some subjects, then drops out and reapplies to the university. After successful readmission, he or she accepts the previously completed courses and continues his or her studies. These cases led to anomalies in the model and were therefore removed from the list. In total, we ended up with data on 1851 students, for each of whom it was possible to determine in which semester they were admitted to the university and whether or not they had obtained a degree. The dropout rate of the participants in the study was 49.97%, with 925 students out of 1851 dropping out.

2.2 Descriptive statistical analysis of data

The dropout rates for the groups by different criteria are shown in the subsequent figures. The interpretation of the graphs is given below.

2.3 Gender and dropout rate

There is a significant difference between the genders when comparing dropout rates. The ratio of males to females is 1362 males and 489 females out of 1851 students. The dropout rate for males is 55%, while for females, it is slightly lower at 35%, as illustrated in Figure 1.

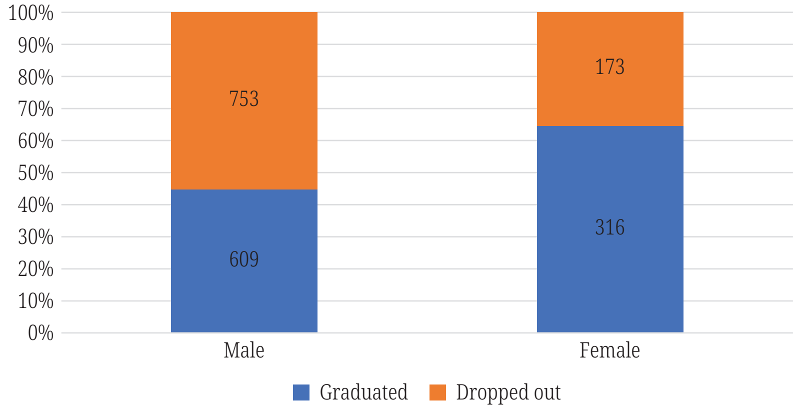

Fig. 1. Gender dropout rate

2.4 Type of finance

Based on the form of funding, students were divided into two groups: those receiving some type of financial subsidy and those paying a tuition fee. Figure 2 shows the dropout rates of the two groups. The dropout rate is 49% for subsidized students and 57% for those who are self-financed. The dropout rate is higher for self-financed students, but the difference is not significant.

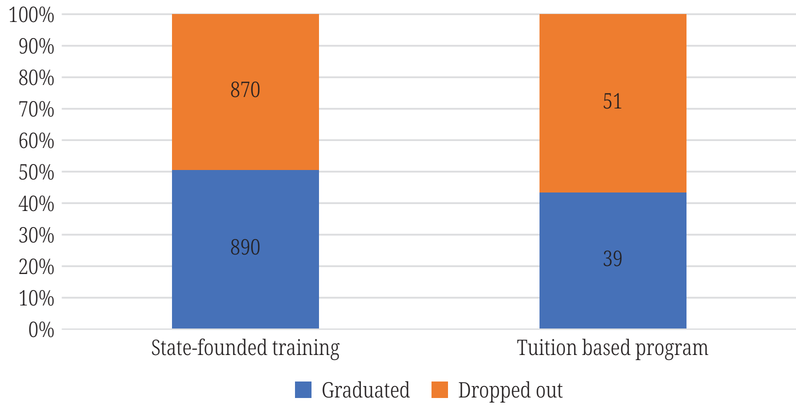

Fig. 2. Dropout rate by funding of training

2.5 Type of training

Figure 3 shows the dropout rates of students grouped by type of training (full-time, part-time). The dropout rate for full-time students is 47%, while for part-time students, it is 62%. Therefore, the dropout rate of correspondence students is higher than that of full-time students. The majority of correspondence students study while working, which puts extra pressure on them as they have much less time to study effectively. Without a supportive family background, it is difficult to complete a correspondence course.

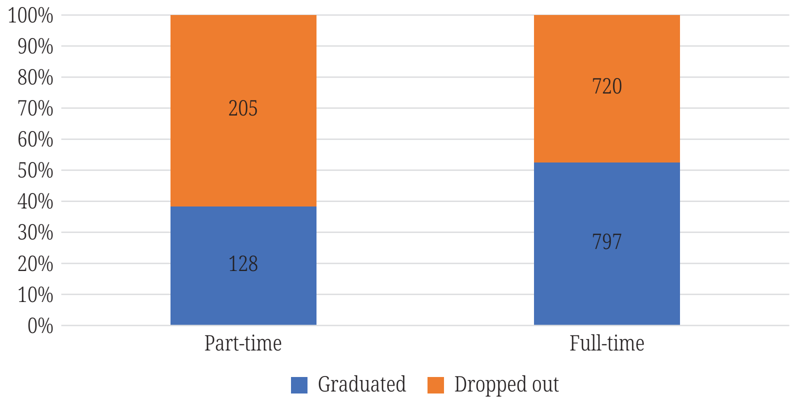

Fig. 3. Dropout rate by type of training

2.6 Regional characteristics

To examine the role of dropout in terms of region, students were divided into two groups: Capital and noncapital students. Budapest Business School has a campus in Budapest. No significant difference was found between the two categories, with a slight difference in favor of the provincial group. The dropout rate for students living in rural areas is 48%, while the dropout rate for students in Budapest is 54% as shown in Figure 4.

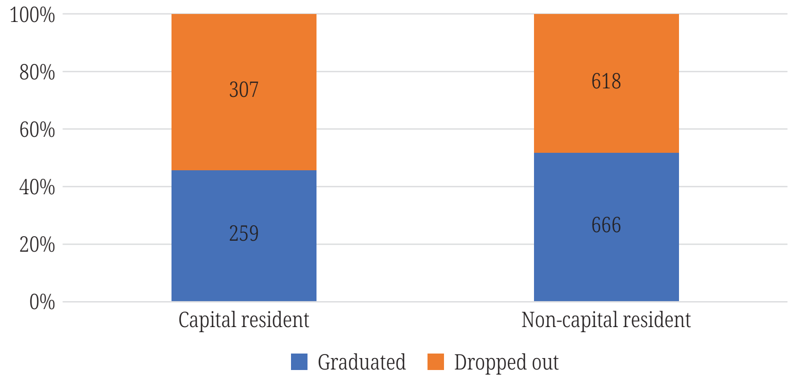

Fig. 4. Dropout rate by place of residence

2.7 Female and male dropout rates among correspondence students

In this grouping, an interesting pattern is observed. The dropout rate of men who are correspondence students is significantly higher than that of women who are correspondence students. Figure 5 shows that only 34% of male correspondence students and 49% of female correspondence students obtain a degree. Male correspondents, therefore, require special attention in terms of dropout rates.

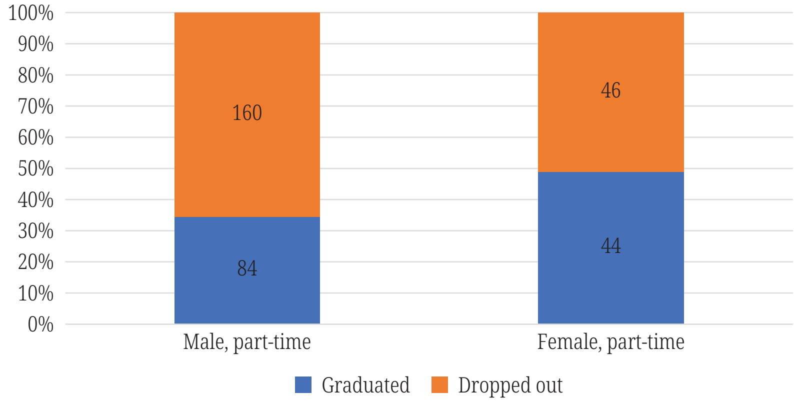

Fig. 5. Male and Female dropout rates among correspondence students

2.8 Correlation between the number of passive semesters and dropouts

The relationship between the number of passive semesters and dropout rates was very strong. A passive semester refers to a student taking a break from his or her studies, being placed in a passive semester, not taking a course, or not taking exams. Students are divided into five groups. The first group is made up of those who did not have any passive semesters. The next groups are those with 1, 2-3, 4-5-6, and 7 or more passive semesters (see Figure 6). The data show that the dropout rate is lowest for students with 0 and 7 or more passive semesters. For students with 0 passive semesters, the dropout rate is only 33%, while for students with 7 or more passive semesters, the rate is even better, with just under 14% of students dropping out. The dropout rate for students with 1 passive semester increases from 33% to 72% compared to the dropout rate for students with 0 passive semesters and to 96% compared to students with 2-3 passive semesters. For students with 4-5-6 passive semesters, the dropout rate is 100%, with none of the students in this group graduating. The number of passive semesters is therefore a very important factor in the dropout rate. These results suggest that persistent students who do not dropout can graduate after seven or more passive semesters.

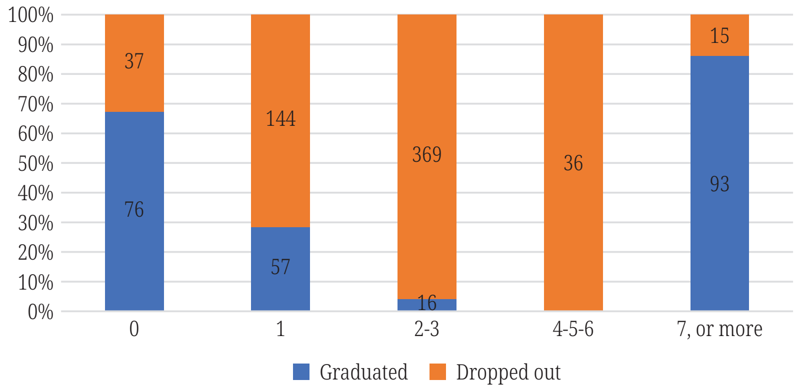

Fig. 6. Dropout rate as a rate of passive semesters

2.9 Age and dropouts

Based on age, students were divided into four groups: 18–21, 22–23, 34–37, and 38 and above. Students in the 18–21 and 38+ age groups performed best, with graduation rates of 55% and 61%, respectively. Only 36.5% of students in the 22–33 age group graduated. Students aged 34–37 are the worst performers, with those in this group being the most at risk, with just 22% of them obtaining a degree (see Figure 7). The exact background of this phenomenon has not been analyzed due to a lack of information, but family formation and specific life situations are likely the reasons for the high dropout rates.

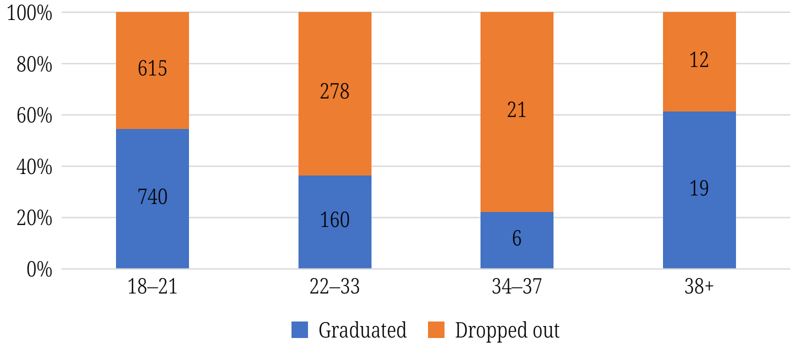

Fig. 7. Attrition rates by age

3 DEFINING MACHINE LEARNING MODEL INDICATORS

In the present study, we aim to build a predictive model that predicts with high probability whether a student is at risk of dropping out based on the data available at the time of enrollment and the student's performance in the first two semesters after enrollment.

In the model, 16 indicators were defined, which can be divided into two groups according to their type. One type includes indicators that are available at the time of enrollment. These indicators are defined based on the data listed in the descriptive statistical analysis. These indicators are:

1. *Type of study* (full-time, part-time) (*TS*),
2. *Gender* (male, female) (*G*),
3. *Residence* (Budapest, province) (*R*),
4. *Admission score* (*AS*),
5. *Financial status* (fee-paying, self-paying) (*FS*),
6. *Age* (*A*).

The next group includes indicators whose values are calculated based on the number of subjects taken during the two semesters and the number of subjects completed. Before listing each indicator, some concepts are clarified. According to the requirements of the core curriculum, there are three main ways to complete a subject. Requirements are (a) teacher signature, where the subject requires a signature to be completed, the value of which may be a signature or a signature may be refused; (b) term mark, where the subject requires a practical grade to be obtained during the semester, the value of which may be 1, 2,..., 5; and (c) exam mark, where the student has to pass an examination to complete the subject in the exam period, also worth 1, 2,..., 5. Each subject is assigned a credit value that, if passed, increases the total credit value for the semester. On this basis, the following time-variant indicators of student performance are defined for the model for the first two semesters:

7. *Total number of cources taken (TNTC)*
8. *Signature rate* (Number of signatures obtained for subjects with a signature requirement/Total number of subjects with signature requirement) (*SR*)
9. *Exam mark rate* (Number of subjects with exam mark requirements passed with at least satisfactory level/Total number of subjects with a exam mark requirement) (*EMR*)
10. *Exam mark average* (average of the results of the subjects with the requirements of the exam mark) (*EMA*)
11. *Term mark rate* (Number of subjects with term mark requirements passed with at least satisfactory level/Total number of subjects with term mark requirement) (*TMR*)
12. *Term mark average* (Average of the results of the subjects meeting the requirements of the term mark) (*TMA*)
13. *Credit rate* (Sum of credits of subjects completed/Sum of credits of subjects taken) (*CR*)
14. *Number of subjects taken more than once (NOSTMTO)*
15. *Number of successfully completed subjects taken more than once (NOSCSTMTO)*
16. *Number of passive semesters (NOPS)*

The model included the above 16 indicator values for a total of 1851 students over two semesters.

4 RESULTS

The relative importance of each predictor is shown in Figure 8.

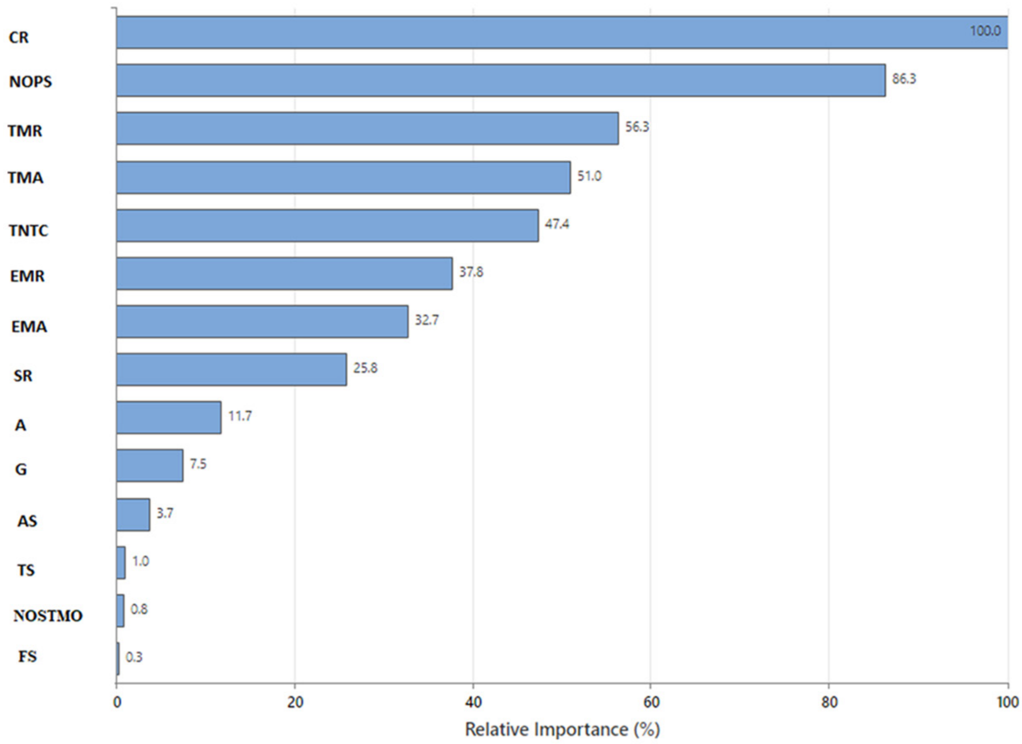

Fig. 8. Relative importance of predictors

The most important indicator for dropout was the *Credit rate*. This indicator expressed the percentage of successful completion of the subjects taken. The next very important indicator is the number of *Passive semesters*, which was expected from the results of the descriptive statistics. The importance of the indicators *Term mark rate, Term mark average, Number of subjects taken, Exam mark rate,* and *Exam mark average* is also significant, but not as much as *Credit rate* or the *Number of passive semesters*. This suggests that students who do not have an outstanding academic record and who do not earn good grades have a similar chance of graduating as students whose academic record is outstanding and who do well on exams. Other relatively important indicators are *Age* and *Gender*, *Admission score, Method of study, Number of subjects taken more than once,* and *Number of successfully completed subjects taken more than once* are of relatively low importance, and *Residence* and *Financial status* are of negligible importance in the model. The relatively high importance of the indicator the *Number of subjects taken* may be related to academic performance. Students who took a few subjects in the first semester took more subjects in the following semester to meet their academic obligations. For many students, however, this led to a failure to meet requirements, with few students being able to cope with the pressure of taking many subjects.

The model was created in MATLAB, and the decision tree was chosen as the ML algorithm for the model. The decision tree is a supervised ML algorithm best suited for classification problems. In our case, this algorithm proved to be the best for predicting attrition. The trained model was used for predictions on test data. The error matrix generated from the predictions on test data is shown in Figure 9.

	Predicted Class (Training)				Predicted Class (Test)			
Actual Class	**Count**	**0**	**1**	**% Correct**	**Count**	**0**	**1**	**% Correct**
0 (Event)	651	546	105	83.9	275	234	41	85.1
1	646	59	587	90.9	279	26	253	90.7
All	1297	605	692	87.4	554	260	294	87.9

Statistics	Training (%)	Test (%)
True positive rate (sensitivity or power)	83.9	85.1
False positive rate (type I error)	9.1	9.3
False negative rate (type II error)	16.1	14.9
True negative rate (specificity)	90.9	90.7

Fig. 9. Confusion matrix

The decision tree model built on training data has a classification accuracy of 87.9% when applied to test data. The model was primarily built to predict cases where a student is expected to drop out and not graduate. The model has a true positive rate (sensitivity) of 85.1% for the test cases, which means that students who do not graduate are identified with a probability of 85%. Students in the group who are expected to graduate (True negative rate, specificity) and who have actually graduated are identified with an accuracy of 90.7%. The first type of error (false-positive rate) means that the model incorrectly predicts that a student will not graduate when, in reality, they did. The value of this indicator is 9.3%. The false-negative rate means that the model incorrectly predicts that the student will graduate when, in fact, he or she did not. The value of this indicator is 14.9%. These indicators have an acceptable value, and the model has a high probability of identifying students at risk of dropping out, but in 15% of cases, it fails to identify those who then go on to drop out. The indicators can be further improved if more than two semesters of data are considered, but as mentioned in the introduction, early identification is essential for managing and controlling dropout.

5 SUMMARY

Artificial intelligence and ML tools have made incredible progress in recent years and are becoming indispensable tools in our everyday lives. This technology has progressed in automotive, image processing, social media, commerce, medicine, education, and many other fields. This technology can be used effectively primarily in environments where conditions are considered constant. In changing environments, such as education, the use of technology requires careful planning and analysis. After all, much can change over the years in an educational process. Curriculums, instructors, students' backgrounds, subject content, the number of subjects taken per semester, and many other factors can change. The only way to develop a predictive ML model in this context is to build predictors that perform well in this changing environment and are closely related to the outcomes. In the model built in our study, we verified that the changing environment does not affect the predictive ability of the model, which ultimately results in its good performance. The ML model can identify students at risk of dropping out with acceptable accuracy. Based on the data from the student administration system and complemented by appropriate IT improvements, it can be an effective tool for training development, primarily in the area of student learning support activities at the university.

6 REFERENCES

[1] M. A. Ali, N. Sahari, S. Mat, and S. Zainudin,"Identifying Students' Learning Patterns in Online Learning Environments: A Literature Review," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 17, no. 08, pp. 189–205, 2022. https://doi.org/10.3991/ijet.v17i08.29811

[2] R. Bennett, "Determinants of Undergraduate Student Drop Out Rates in a University Business Studies Department," *Journal of Further and Higher Education*, vol. 27, no. 2, pp. 123–141, 2003. https://doi.org/10.1080/030987703200065154

[3] L. Bognár and T. Fauszt, "Factors and Conditions that Affect the Goodness of Machine Learning Models for Predicting the Success of Learning," *Computers and Education: Artificial Intelligence*, vol. 3, pp. 1–16, 2022. https://doi.org/10.1016/j.caeai.2022.100100

[4] L. Bognár, T. Fauszt, and B. Nagy, "Machine Learning Model Building Techniques for Small and Medium-Sized University Courses," *International Journal of Artificial Intelligence*, vol. 19, no. 2, pp. 20–43, 2021.

[5] M. A. Chatti, A. L. Dyckhoff, U. Schroeder, and H. Thüs, "A Reference Model for Learning Analytics," *International Journal of Technology Enhanced Learning*, vol. 4, no. 5–6, pp. 318–331, 2012. https://doi.org/10.1504/IJTEL.2012.051815

[6] C. Csók, Á. R. Dusa, A. Hrabéczy, I. Novák, Z. Karászi, G. Ludescher, and D. Németh, "A hallgatói lemorzsolódás és háttértényezői egy kvalitatív kutatás tükrében," *Lemorzsolódás és perzisztencia a felsőoktatásban*, pp. 38–62, 2018. Avalilable: https://mek.oszk.hu/19000/19032/19032.pdf [Accessed: Feb. 20. 2023]

[7] F. Dalipi, A. S. Imran, and Z. Kastrati, "MOOC dropout prediction using machine learning techniques: Review and research challenges," In *2018 IEEE global engineering education conference (EDUCON)* IEEE, pp. 1007–1014, 2018. https://doi.org/10.1109/EDUCON.2018.8363340

[8] D. Delen, "A Comparative Analysis of Machine Learning Techniques for Student Retention Management," *Decision Support Systems*, vol. 49, no. 4, pp. 498–506, 2010. https://doi.org/10.1016/j.dss.2010.06.003

[9] O. Zs. Demcsákné and P. Huszárik, "Lemorzsolódási vizsgálatok a felsőoktatásban. Rendszerszintű fejlesztések és hozzáférés bővítését szolgáló ágazati programok a felsőoktatásban," *Oktatási Hivatal kiadványa*, 2020. Available: http://publicatio.bibl.u-szeged.hu/22710/1/EFOP345_Kompetenciameres_Zarotanulmany.pdf [Accessed: Feb. 20. 2023]

[10] É. Farkas, "Tanulási eredmények értékelése a felsőoktatásban," *University of Szeged*, Szeged, 2019. Available: https://mek.oszk.hu/19300/19399/19399.pdf [Accessed: Feb. 13. 2023]

[11] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth, "Active learning increases student performance in science, engineering, and mathematics," In *Proceedings of the national academy of sciences*, vol. 111, no. 23, pp. 8410–8415, 2014. https://doi.org/10.1073/pnas.1319030111

[12] A. S. Hoffait and M. Schyns, "Early Detection of University Students with Potential Difficulties," *Decision Support Systems*, vol. 101, pp. 1–11, 2017. https://doi.org/10.1016/j.dss.2017.05.003

[13] K. Oqaidi, S. Aouhassi, and K. Mansouri, "Towards a Students' Dropout Prediction Model in Higher Education Institutions using Machine Learning Algorithms," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 18, pp. 103–117, 2022. https://doi.org/10.3991/ijet.v17i18.25567

[14] G. Kovács, K. Erdélyi, Z. Pádár, P. Zenke, and M. Nogel, "Contamination of Forensic DNA Evidence in the Light of Hungarian Court Decisions: A Review of 25 Years," *Forensic Science International: Genetics Supplement Series*, vol. 8, p. 242–244, 2022. https://doi.org/10.1016/j.fsigss.2022.10.049

[15] D. Kučak, V. Juričić, and G. Đambić, "Machine Learning in Education, A Survey of Current Research Trends," *Annals of DAAAM & Proceedings*, vol. 29, pp. 0406–0410, 2018. https://doi.org/10.2507/29th.daaam.proceedings.059

[16] H. Mi, Z. Gao, Q. Zhang, and Y. Zheng, "Research on Constructing Online Learning Performance Prediction Model Combining Feature Selection and Neural Network," *International Journal of Emerging Technologies in Learning*, vol. 17, no. 7, 2022. https://doi.org/10.3991/ijet.v17i07.25587

[17] P. Miskolczi, F. Bársony, and G. Király, "Hallgatói lemorzsolódás a felsőoktatásban: elméleti, magyarázati utak és kutatási eredmények összefoglalása," *Iskolakultúra*, vol. 28, no. 3–4, pp. 87–105, 2018. Available: http://www.iskolakultura.hu/index.php/iskolakultura/article/view/22790 [Accessed: Feb. 13. 2023]

[18] S.A. Nonis and G.H. Fenner, "An Exploratory Study of Student Motivations for Taking Online Courses and Learning Outcomes," *Journal of Instructional Pedagogies,* vol. 7, 2012.

[19] P. Nuankaew, "Dropout Situation of Business Computer Students, University of Phayao," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 14, no. 19, pp. 115–131, 2019. https://doi.org/10.3991/ijet.v14i19.11177

[20] G. Pusztai and F. Szigeti, "Lemorzsolódás és perzisztencia a felsőoktatásban [Dropout and persistence in higher education]," *Debreceni Egyetemi Kiadó,* Debrecen, Hungary, pp. 38–62, 2018. Available: https://mek.oszk.hu/19000/19032/19032.pdf [Accessed: Feb. 13. 2023]

[21] A. Ria, B.D. Harry, and A.H. Zanial, "Monitoring Learners' Performance By Modelling Learning Process Using Machine Learning," *Journal of Engineering Science and Technology*, Special Issue on SIMASE, pp. 30–39, 2021.

[22] M. Szemerszki, "Lemorzsólódási adatok és módszertani megfontolások," In *G. Pusztai & F. Szigeti (Eds.), Lemorzsolódás és perzisztencia a felsőoktatásban [Dropout and persistence in higher education] Debreceni Egyetemi Kiadó.* Debrecen, Hungary, pp. 15–27, 2018. Available: https://mek.oszk.hu/19000/19032/19032.pdf [Accessed: May. 08. 2023]

[23] H. Tagharobi and K. Simbeck, "Introducing a Framework for Code based Fairness Audits of Learning Analytics Systems on the Example of Moodle Learning Analytics," In *Proceedings of the 14th International Conference on Computer Supported Education. CSEDU (2)*, pp. 45–55, 2022. https://doi.org/10.5220/0010998900003182

[24] C. Vida and E. Németh, "Felsőoktatás a változások tükrében –verseny, minőség, teljesítmény," *Magyar Akkreditációs Szemle*. 2. szám, pp. 21–25, 2021. Available: https://www.penzugyiszemle.hu/hu/fokuszban/felsooktatas-a-valtozasok-tukreben-asz-elemzes-a-magyar-akkreditacios-szemleben. [Accessed: May. 08. 2023]

[25] S. Assami, N. Daoudi, and R. Ajhoun, "Implementation of a Machine Learning-Based MOOC Recommender System Using Learner Motivation Prediction," *International Journal of Engineering Pedagogy (iJEP)*, vol. 12, no. 5, pp. 68–85, 2022. https://doi.org/10.3991/ijep.v12i5.30523

[26] S. Sraidi, E.M. Smaili, S. Azzouzi, and M.E.H. Charaf, "A Neural Network-Based System to Predict Early MOOC Dropout," *International Journal of Engineering Pedagogy (iJEP)*, vol. 12, no. 5, pp. 86–101, 2022. https://doi.org/10.3991/ijep.v12i5.33779

7 AUTHORS

Tibor Fauszt is working as an Associate Professor of Information Technology at the Budapest Business School University of Applied Sciences. Buzogány u. 10-12. Budapest, Hungary, H-1149. https://uni-bge.hu/en. He is a member of the Informatics Sciences Committee of the Hungarian Rectors' Conference. http://www.mrk.hu/bizottsagok/informatikai-tudomanyok-bizottsaga/

Katalin Erdélyi Ph.D. is a student in Law Enforcement of University of Public Services, Üllői út. 82. Budapest, Hungary H-2400, https://rtk.uni-nke.hu/oktatas/rendeszettudomanyi-doktori-iskola/rolunk

Dóra Dobák is an Associate Professor of Information Technology at the Budapest Business School University of Applied Sciences. Buzogány u. 10-12. Budapest, Hungary, H-1149. https://uni-bge.hu/en. She is the head of the Department of Business Information Technology.

László Bognár is a Professor of Applied Statistics at the University of Dunaújváros, Táncsics M. u. 1. Dunaújváros, Hungary H-2400, www.uniduna.hu. He also works for APAVE Hungary and conductss trainings in different fields of industrial statistics, www.apave.hu

Endre Kovács is an Associate Professor of Information Technology at the Budapest Business School University of Applied Sciences. Buzogány u. 10-12. Budapest, Hungary, H-1149. https://uni-bge.hu/en. His research interests include business analytics and machine learning.