Discovery of the predictive success factors of students in the Logistics Technology Course, in ENADE exams, through Educational Data Mining

Descubrimiento de los factores predictivos de éxito de los estudiantes en el Curso de Tecnología Logística, en los exámenes de ENADE, a través de la Minería de Datos Educativos

Ivonaldo Vicente da Silva[1,*], Márcia Terra da Silva[1], Pedro Luiz de Oliveira Costa Neto[1]

[1]Universidade Paulista. Programa de Pós-Graduação em Engenharia de Produção. Rua Doutor Bacelar, 1212- 4º Andar, Mirandópolis. São Paulo, Brasil.
*Autor de correspondencia: ivonaldosilva@yahoo.com.br

Abstract

The objective of this study was to discover, through Educational Data Mining, which factors were most associated with the best performances obtained by students of the Technology in Logistics course in the ENADE exams, of the 2018 edition. The data collected on the INEP website were treated and formatted in order to remove whites or nulls. The Educational Data Mining phase included the execution of the Decision Tree, Random Forest, Gradient Boosted Tree and Naive Bayes algorithms. After all tests were performed, the algorithm that showed the best performance was Naive Bayes with Accuracy = 98.21%, Kappa Index = 0.964, Recall = 83.32% and Precision = 82.40%. The results indicated that the factors related to the Number of hours for study, the level of education of the country, whether the Educational Institution provides adequate materials and equipment for classes, whether teachers use IT resources and whether the Course proposes Level Updated Knowledge, were more associated with better performance in ENADE assessments. The discovery of these factors can contribute to the development of action plans, by education professionals, that can propose improvements in the educational environment.

Keywords: Logistics, ENADE, Performance Factors, Educational Data Mining

Resumen

El objetivo de este estudio fue conocer, a través de la Minería de Datos Educativos, qué factores estaban más asociados a los mejores desempeños obtenidos por los estudiantes en la asignatura de Tecnología Logística en los exámenes de ENADE, de la edición de 2018. Se trataron los datos recolectados en la página web del INEP y formateado para eliminar blancos o nulos. La fase de Minería de Datos Educativos incluyó la ejecución de los algoritmos Decision Tree, Random Forest, Gradient Boosted Tree y Naive Bayes. Una vez realizadas todas las pruebas, el algoritmo que mostró el mejor rendimiento fue Naive Bayes con precisión = 98,21%, índice Kappa = 0,964, recuperación = 83,32% y precisión = 82,40%. Los resultados indicaron que los factores relacionados con la cantidad de horas a estudiar, el nivel de educación del país, si la institución educativa proporciona materiales y equipos adecuados para las clases, si los docentes utilizan recursos de tecnología de la información y si el curso propone el nivel de Conocimiento actualizado, se asociaron más con un mejor desempeño en las evaluaciones de ENADE. El descubrimiento de estos factores puede contribuir al desarrollo de planes de acción, por parte de los profesionales de la educación, que puedan proponer mejoras en el entorno educativo.

Palabras clave: Logística, ENADE, Factores de desempeño, Minería de datos educativos

1. Introdução

Production systems, in general, have evolved a lot since the 2000s. New production technologies, information systems and artificial intelligence are just a few examples of modernizations in the production área [1]. Another area that has also been improving in relation to its production techniques is that of Agriculture, as an example the forecast is for a new record for grain production in relation to the harvest of the period 2020/2021, with a total of approximately 268.7 million tons, 4.2% higher than the previous harvest [2].

Regarding the modernization of production systems, the Logistics area has not been forgotten, as intelligent tracking systems, the use of robotics in storage processes and automated inventory and distribution controls have been implemented in companies as alternatives to contain and control total costs of production [3].

Faced with a scenario where companies are looking for improvements in the entire production chain, including distribution, it becomes necessary to observe how future professionals in this area are being trained to act in this context. Specifically, about knowledge assessment, higher education courses are encouraged to direct the graduates of their courses to conduct to large scale assessments of the National Student Performance Exam-ENADE. Each higher education course is evaluated every three years through the application of tests for graduates. The results of these assessments, including the sociodemographic questionnaires applied to students, are made publicly available through the website of the National Institute of Educational Studies and Research Anísio Teixeira-INEP [4].

Understanding the level of knowledge of graduates of the Logistics Technology course is essential for actions in the educational environment to be developed in order to update educational curricula, practical internships and the implementation of more up-to-date content with a focus on the labor market work. In the 2015 edition of ENADE, the total number of students who took the exam was 10,597, while in the next edition in 2018, that total was 15,437, representing an increase of approximately 45.67% compared to the 2015 edition.

Regarding the performance of students, ENADE publishes the General Score that each student obtained in the exam, and this grade can vary from 0 to 100 points. A matter of concern is the level of performance presented by graduates of the Logistics Technology course, as in the 2015 edition the percentage of students who did not exceed 50 points in the exam was approximately 59.19%, in the 2018 edition this percentage increased to 75.77%. This increase may reflect a scenario in which students are in need of more support, be it from family, teachers, educational institutions, companies where they work, among other factors.

Discovering factors that are more associated with better student performance in large-scale assessments can help education professionals as well as educational institutions to be able to develop action plans that can propose improvements in the education system. In view of the large volume of data made available from the results of the ENADE exam, an analysis strategy that could be used for this purpose is necessary, which is why Educational Data Mining techniques are an alternative.

Educational Data Mining has become an important alternative for analyzing large volumes of data, as it uses algorithms and machine learning to identify patterns of behavior that may contribute to the decision-making process [5,6].

Based on this information, the objective of this study is to discover, through Educational Data Mining, which factors were most associated with the best performances obtained by students of the Logistics Technology course in the ENADE exams, of the 2018 edition.

Periodically, several studies are published focusing on the Logistics area. Just as an example, using the Google Scholar search system, when performing a search with the keyword "Business Logistics", the system returned 15,100 publications made in the period from 2017 onwards, and combining "Business Logistics" and "Teaching" , the return was only 1440 publications for the same period. And these studies, for the most part, focus on transportation, supply chain, storage and material acquisition. However, studies focusing on the preparation of professionals in this area or the dissemination of factors that may contribute to improving student performance in ENADE assessments are lacking.

Some observed studies used the results of the ENADE assessments and had as a proposal to understand which factors were present in students with better performances. As an example, the study by [7] aimed to investigate the determinants of academic performance of students from Higher Education Institutions in Brazil. The author used data from the 2013 edition of ENADE. The results obtained by econometric models of linear and quantile regression indicated that the age, sex, race and family income of the students and characteristics of the Educational Institution itself, such as the region where it is inserted and whether it is public or private, can influence the performances of students in assessments.

The study by [8] aimed to observe the factors that were associated with the performance of students in the Nutrition course. The authors used the simple and multiple linear regression technique in the ENADE data for the 2004,2007,2010 and 2013 editions and as a result found that students with lower performance in the evaluations were black students, who eventually work, with a family income of up to three salaries minimum, of fathers and mothers with no schooling, who attended half of high school in public school and half in private school, who received scholarships or funding and did not enter higher education through affirmative policies.

In order to observe whether the characteristics of the students' families were associated with performance in the ENADE assessments [9], used data from the 2017 edition for the Physical Education course. Through statistical techniques, the authors found as a result that students who have family members with training in higher education or in postgraduate courses, presented better performances in the ENADE assessments.

Guimarães [10] in his study based on the results of the 2012 ENADE evaluation, for the Logistics course, from an educational institution in the state of Tocantins, aimed to produce contributions to the development of this course. The author, through qualitative, descriptive and documentary research, obtained as a result that students have difficulties in the interpretation of questions in the evaluation of ENADE and still were unaware of some content present in the test. This may represent a situation in which the academic curricula are not up to date with the needs of the area, thus students obtain professional training with learning disabilities.

In another study [11] observed how the graduates of the Logistics Technology course of an Educational Institution located in Salvador were inserted in the job market. By applying questionnaires to 51 graduates and 12 companies, the graduates of the course indicated that important skills had developed during the course, but that more support would be needed in relation to practical internships.

The research by [12] had a proposal to analyze whether sociodemographic and academic characteristics were related to the performance of students in the Production Engineering course, in the state of Santa Catarina, in the evaluation of ENADE 2017. Using statistical techniques ANOVA and Chi-Square test , the results indicated that the monthly family income, the age group and the

RCI da Silva et al.

administrative category of the Higher Education Institution were related to the best performance of students at ENADE.

In addition to these studies, others were also collected and their results could contribute to the development of this research.

2. Material and Methods

To compose this study, a bibliographic search was initially carried out in scientific databases, with the objective of identifying in publications factors that are more associated with better educational performance. As a result of this step, a set of factors relevant to the students' performance was selected, and from this set, the data collection and selection step can be defined and planned.

The next step was to collect the data from the ENADE evaluations, from the 2018 edition, on the INEP website. Along with the results, INEP also makes available the results of the questionnaires completed by the students and the file containing a data dictionary, used to interpret the variables of the questionnaires [4].

The total number of records in the database made available by INEP, referring to students in the Logistics Technology course is 12,675. This total already includes students from courses offered in the face-to-face and distance (EAD) modality. The questionnaire completed by the students has 68 questions that aim to obtain information about the students' sociodemographic profile and how they perceive the quality of the course. From the data collected, the next phase was the pre-processing of the data.

Data pre-processing phase

This study focused on the completed questionnaires and the results obtained in the 2018 ENADE Assessment, by students of the Logistics Technology course at the national level. To compose the database for the Data Mining phase, called "*LogisticStudents*", initially the variables or attributes necessary for this study were separated, based on the observed literature and together with the questionnaires completed by the students. Of the total available questions, 14 were separated based on the factors mentioned in the bibliography. The list of factors and questions is described in Table 1.

With the "*LogisticStudents*" database ready, the next step was the task of treatment these data, that is, all the data present in that database were analyzed and the white, null or incorrectly filled data were eliminated. This task was essential to avoid errors of interpretation and analysis, in the Data Mining phase. After performing this task, of the total of 12,675 records available in the database, only 5,053 could be used. It is important to note that, in the Data Mining phase, only the records that had valid answers to the questions presented in Table 1 were used. As a next phase, the formatting of the data was carried out, which consisted of replacing the answers present in the records, with a word or acronym that provided greater gain of information, in the analysis phase. With that, and using the attribute "MOTHER_STUDIED" as an example, where the letter "A" was present, the value corresponding to the form was inserted, in this case the word used was "None".

Regarding student performance, grades obtained in evaluations can vary from 0 to 100 points, and these values are described in the "NT_GER" attribute. The performances obtained by the students were classified according to the quartile technique, with Quartile_1 comprising values between 0 and 24.99 points, Quartile_2 from 25 to 49.99 points, Quartile_3 from 50 to 74.99 points and finally

Quartile_4 from 75 points onwards. To store this classification, the "LEVEL" attribute was created in the database.

Finally, with the base properly prepared, the next phase consisted of the Data Mining process and the results achieved were analyzed based on the bibliography initially observed.

Table 1. Factors present in the Student Questionnaires

Attributes	Form questions
Questions related to Forms applied to Students	
MOTHER_STUDIED	By what stage of schooling did your mother complete?
FATHER_STUDIED	By what stage of schooling did your father complete?
IF_WORKS	Which of the following best describes your work situation (except internship or scholarships)?
WHO_ENCOURAGED_STUDY	Who gave you the greatest incentive to study?
FAMILY_COMPLETED_COLLEGE	Has anyone in your family completed a college degree?
STUDY_HOURS	Approximately how many hours per week did you dedicate to your studies, except for class hours?
WHY_CHOOSE_COURSE	What is the main reason why you chose this course?
WHY_CHOOSE_INSTITUTION	What is the main reason why you choose your institution of higher education?
COURSE_PROV_UPD_KNOWLEDGE	The course provided access to updated and / or contemporary knowledge in their area of training.
PROF_MASTERS_CONTENT	Professors demonstrated mastery of the content covered in the subjects.
PROF_USES_ICT	Professors used information and communication technologies (ICTs) as a teaching strategy (multimedia projector, computer lab, virtual learning environment).
SUITABLE_ENV_EQUIP	The environments and equipment for practical classes were suitable for the course.
SUITABLE_EQUIP_MATERIALS	The equipment and materials available for practical classes were suitable for the number of students.
NT_GER	Overall score in the ENADE assessment
LEVEL	Classification of the note according to the quartiles.

Educational Data Minig phase

With the database ready, the next step was to run the mining algorithms, for which the student version RapidMiner software was used. The algorithms chosen, as already described, were Naive Bayes, Decision Tree, Random Forest and Gradient Boosted Tree.

The Naive Bayes algorithm is considered a probabilistic classifier based on the Bayes Theorem created by the mathematician Thomas Bayes, whose logic is the conditional probability, that is, the probability of the event X occurring given the event Y. The formula that this algorithm uses to predicting the results is:

$$p\left(\frac{Ai}{B}\right) = \frac{P(Ai \cap B)}{P(B)} = \frac{P(Ai)\,P\left(\frac{B}{Ai}\right)}{\sum P(Ai)P(\frac{B}{Ai})}$$

(1)

where Ai are the causes and B the observed results

Decision Trees are normally used in classification and regression models, can be supervised or not and are based on algorithms such as J48, C.45, ID3, CART, Random Forest, Gradient Boosted Tree. When using the supervised method, it is necessary that there is a prediction target attribute for the development of the trees, and in this model, the algorithm is in charge of finding the best pruning points to assemble the Tree, with the attributes that showed the best gains in information based on the Entropy calculation. After all calculations are made, the main node with the highest gain of information is indicated, and from that node all other branches or leaves are added. To calculate the information gain it is necessary to choose one of the criteria which are: *Accuracy, Information_gain, Gain_ratio, Least_Square*. In this way, it can be understood that the most important attribute, based on Entropy and Information Gain, is at the top of the tree. The formula used to calculate Entropy and:

$$\text{Entropy (S)} = \sum_{i=1}^{c} - p_i \, log_2 \, p_i \tag{2}$$

where p_i is the proportion of data in S that belong to class i

The Random Forest and Gradient Boosted Tree trees are considered as more robust algorithms and their results are usually through a set of trees, which can be analyzed based on the factors with the greatest gain in information or the tree's importance. They are considered a joint learning method (Ensemble) where, to create the forest of trees, the algorithm uses the method of Bootstrap Agregating, or Bagging, which generates a set of data by sampling bootstrap from the database used in the process, and the result set can provide a better overall result. Several small trees with low performance are created, but when analyzed together they become a good alternative for prediction of attributes, since it uses the average between trees to indicate the attributes of greatest importance.

In this study, the classification models were conducted by the supervised method, that is, the algorithms were executed using the target attribute "LEVEL" and all the criteria used in Data Mining are described in Table 2.

Table 2. Criteria for implementing the algorithms in the Educational Data Mining phase

Criteria	Algorithms			
	Decision Tree	*Random Forest*	*Gradient Boosted Tree*	*Naive Bayes*
Data Split	70/30	70/30	70/30	70/30
CrossValidation	k=10	k=10	k=10	K=10
Depth	n=10	n=10	n=10	Not applicable
Total of Trees	1	100	100	Not applicable

3. Results and Discussion

After the methodological phase involving data collection, treatment and selection, the next phase was to migrate the ready data to the specific software for Data Mining. For this study, RapidMiner was used in the educational version.

After running all the algorithms, the one that showed the best performance was Naive Bayes. The performance metrics were Accuracy = 98.21%, Kappa Index = 0.964, Recall = 83.32% and Precision = 82.40%. Recalling that the target Attribute used for all the algorithms was the "Level" of

RCI da Silva et al.

performance of students graduating from the Logistics course. The results for the factors used in Data Mining are illustrated in Figures 1 to 6.

A=None, I just attend classes, B=One to three, C=Four to seven, D=Eight to twelve, E=More than twelve
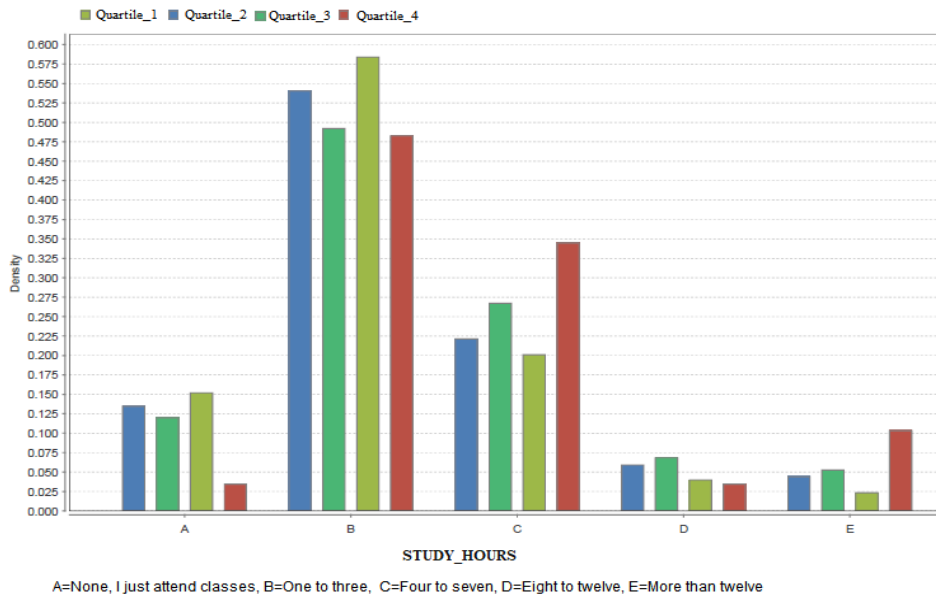
Fig.1 Result for the STUDY_HOURS attribute

The result presented for this attribute indicates that students who seek to dedicate a portion of their time to studies, are better able to obtain better performances in the ENADE assessments. This result converges with the studies by [13]. In Figure 1, it is observed that the students who dedicated "From four to seven" and "More than twelve" hours of study per week achieved better results in the evaluations. Students who dedicate themselves more and devote more time to their studies can develop more skills and competencies necessary to understand the contents of their course, such as the development of logical and critical reasoning, skills with technological resources, skills in carrying out critical analysis through reading and interpretation of texts, among others, in addition, they may be more prepared to carry out large-scale evaluations.

In relation to the question of the form "The equipment and materials available for practical classes were adequate for the number of students", the results for the attribute SUITABLE_EQUIP_MATERIALS illustrate that the Teaching Institutions that provide adequate equipment and materials for the exercise of their pedagogical practices, can provide better learning conditions, and with that, students can get better performances in ENADE assessments. It is observed that students who perceived a good level of quality in the supply of equipment and materials showed relatively better results in the evaluations.
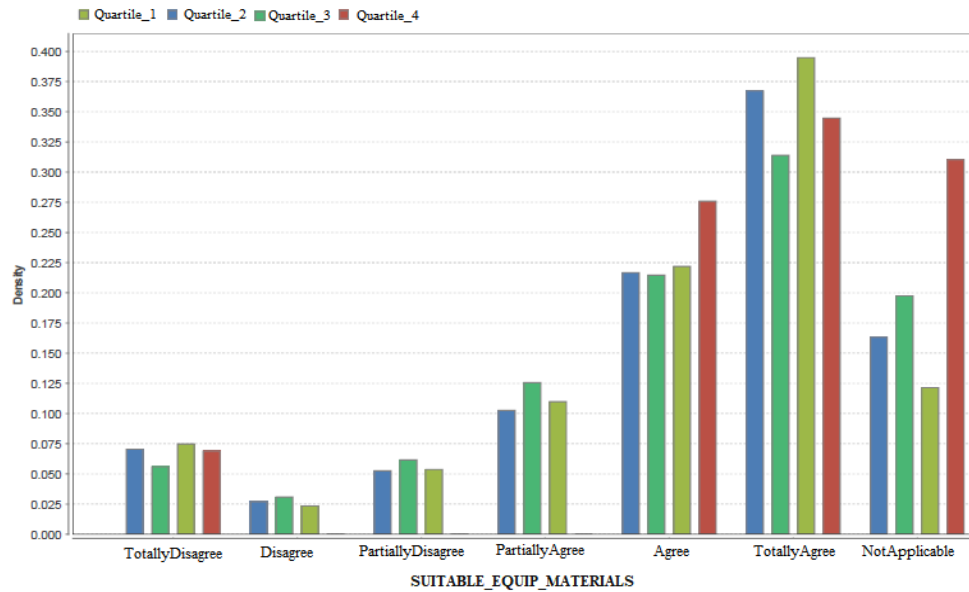
RCI da Silva et al.

Fig.2 Result for the SUITABLE_EQUIP_MATERIALS attribute

The result of this attribute can provide a very important gain of information, since the Education Institutions, from this information, can create strategies for obtaining and offering more modern resources and equipment and updated with the reality of the labor market, with the objective of improve the learning of their students and consequently increase the performance of graduates in ENADE assessments.
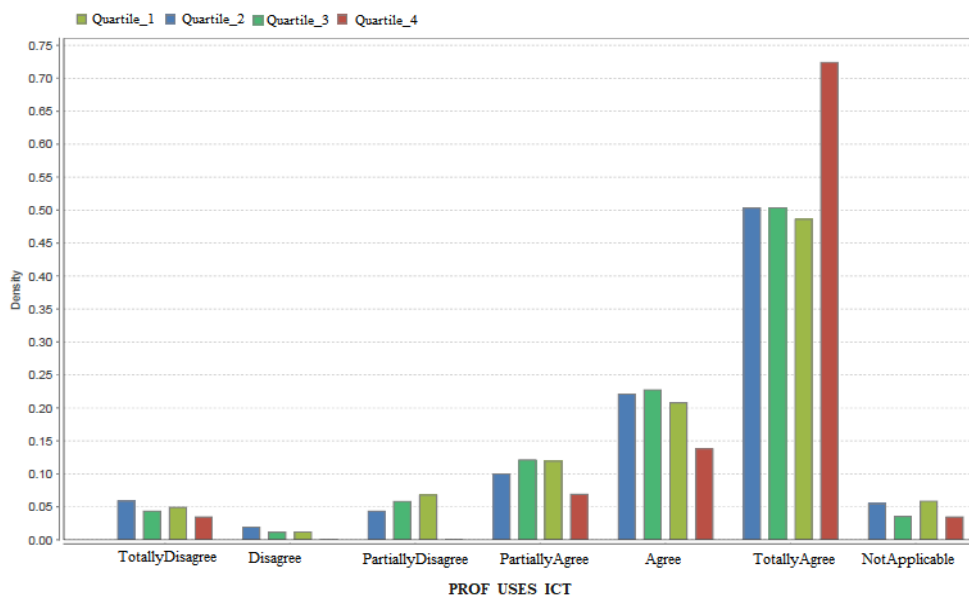

Fig.3 Result for the PROF_USES_ICTs Attribute

Regarding the attribute if Teachers use Information Technology resources in their classes, it is observed that teachers who have more skills in learning technologies through computer science, are able to provide a higher level of learning for their students. The use of technology can provide a higher level of learning to students, through the use of specific software for learning and the internet, teachers and students can share content, make simulations and propose tests in a practical and real time. In addition, the use of technological tools can develop intellectual skills in students, allowing them to reach a higher level of creativity and inventiveness [14].
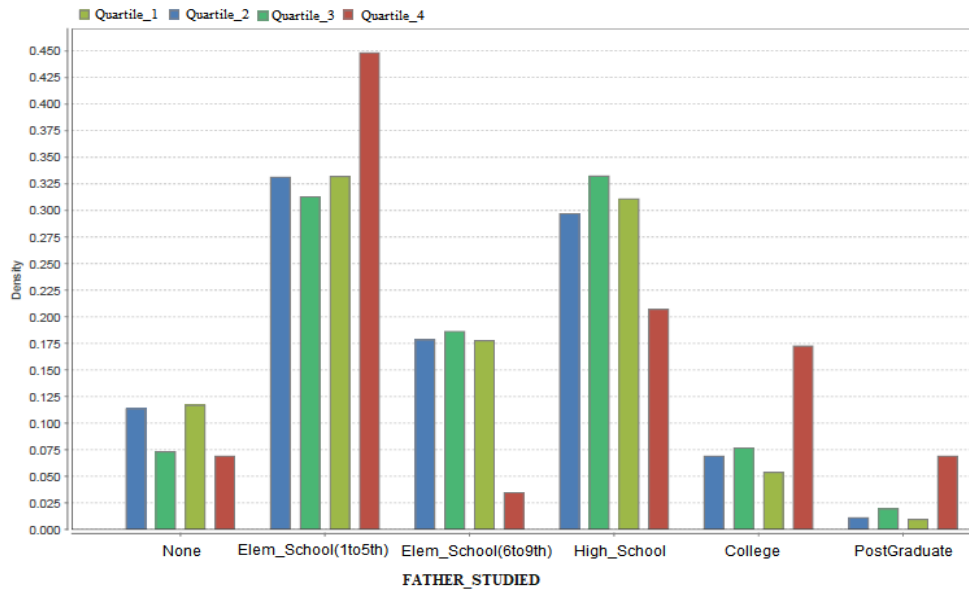
da Silva et al.

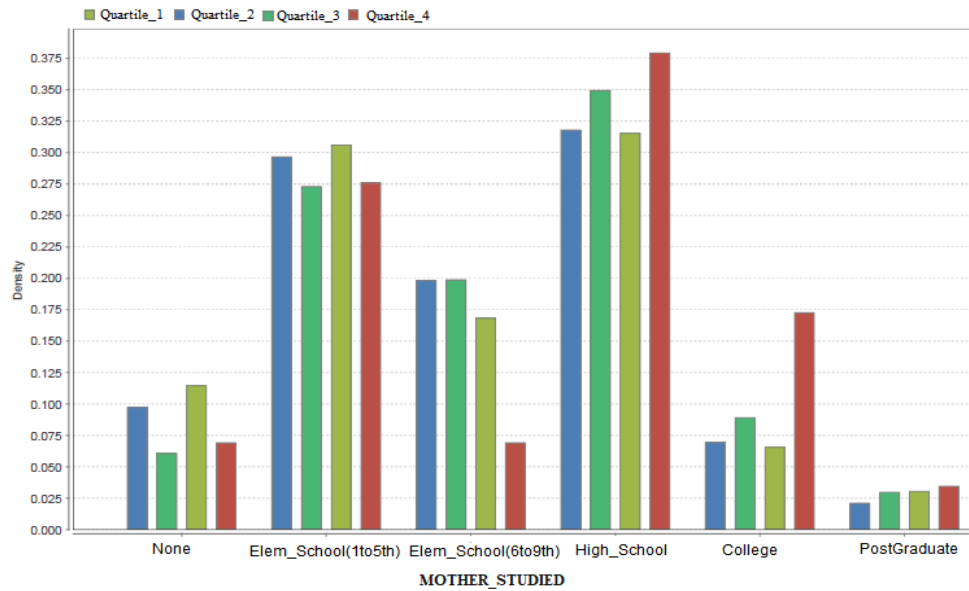**Fig.4** Result for the FATHER_STUDIED attribute

Fig.5 Result for the MOTHER_STUDIED attribute

The attributes FATHER_STUDIED and MOTHER_STUDIED can represent the importance of the educational background of parents in the students' school life. It is observed that when the parents have a higher education or postgraduate course, the students presented better performances in the evaluations. Likewise, those students in which their mothers have completed high school or college, also showed better results. These results were also observed in the studies by [9, 13].

Students can feel more motivated to study when they perceive the presence and accompaniment of the parents, and when they have an academic background compatible with their child's studies, the level of dialogue and guidance can be fundamental for students to develop more secure learning and better and performance in large scale assessments.

It is important to note that higher education education has as one of its proposals to prepare the student to perform the function for which the course is intended. In this regard, the perception of students in relation to the contents taught during the course, including whether these contents are up

da Silva et al.

to date with the job market, can be fundamental for students to obtain better performance in assessments. Thus, it is observed that students who realized that the course proposed a more updated level of knowledge had better performances at ENADE.
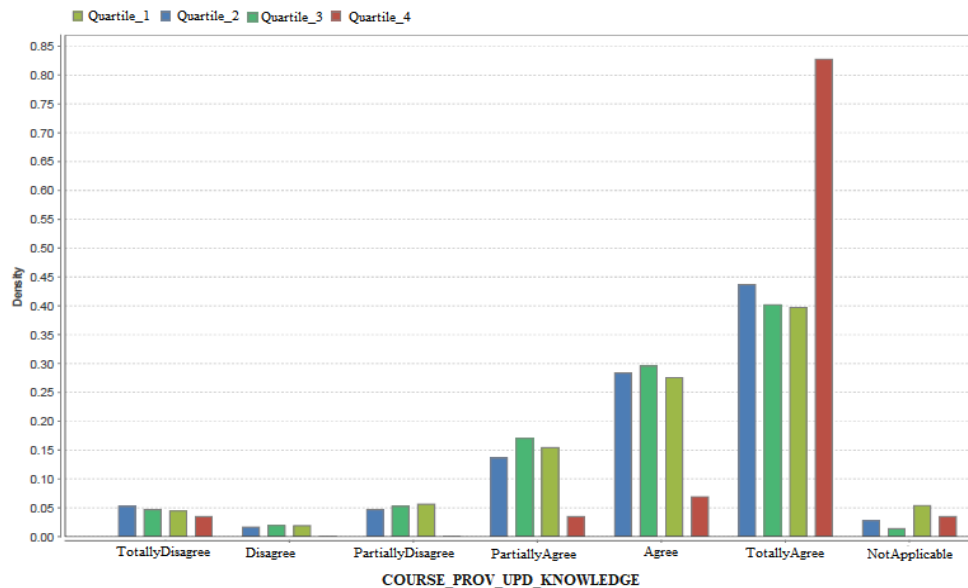

Fig.6 Result for the COURSE_PROV_UPD_KNOWLEDGE attribute

Offering a course that better prepares students in relation to their level of knowledge, without exceeding the maximum time for the formation of the individual, becomes a great challenge for Education Institutions, as the course plans must meet the National Curriculum Guidelines both in terms of in relation to the time of integration of the course regarding the contents considered as minimum requirements for professional training.

In view of this situation, it is up to the Teaching Institutions to create innovative teaching strategies and methodologies that are able to provide students with the most up-to-date level of knowledge possible, both in relation to the needs of the labor market and in relation to new literary approaches, always observing time maximum for the professional training of the individual.

4. Conclusions

The objective of this study was to discover, through Educational Data Mining, which factors were most associated with the best performances obtained by students of the Technology in Logistics course in the ENADE exams, of the 2018 edition. The data provided by INEP referring to the ENADE evaluation of the 2018 edition. After performing the data processing and formatting phase, to remove blank or null records, a new database was created, called "*LogisticStudents*", and supported by the RapidMiner software, the algorithms *Decision Tree, Random Forest, Gradient Boosted Tree and Naive Bayes were executed.*The algorithm that showed the best performance was Naive Bayes and the results indicated that the factors STUDY_HOURS, SUITABLE_EQUIP_MATERIALS, PROF_USES_ICT, FATHER_STUDIED, MOTHER_STUDIED and COURSE_PROV_UPD_KNOWLEDGE were more associated with the best performance of the students. The contribution of this research lies in the fact that by discovering the factors that are most associated with student performance, results of this study, professionals in the field of education and teaching institutions can develop debates that aim to promote actions for improvement for the teaching environment, mainly in relation to the provision of equipment and

materials necessary for classes, routine updating of course plans, as well as in the development of training for the faculty in relation to the use of Information Technology resources. The limitations found are primarily in the scarcity of studies on this topic, and in the difficulty of using the data collected because there are many blank or null records. As proposals for new studies it would be to observe the evolution of these factors from previous evaluations of ENADE, or even, to replicate the methodology for other courses evaluated by ENADE.

References

1. Yamada, V.Y., Martins, L.M., *Indústria 4.0: um comparativo da indústria brasileira perante o mundo*. Rev. Terra Cult. Cad. Ensino E Pesqui., 2019. **34**: p. 1-15.
2. Brasil. Ministério da Agricultura, *Brasil deve ter novo recorde de produção na safra de grãos 2020/21*, 2020. Available at: https://www.gov.br/pt-br/noticias/agricultura-e-pecuaria/2020/10/brasil-deve-ter-novo-recorde-de-producao-na-safra-de-graos-2020-21.
3. Silva, V.R., Silca, A.C.C.J., Londero, L.B., Bianchi, R.C., Zanatta, J.M., *Análise dos impactos da logística 4.0 em uma empresa do ramo agrícola da cidade de Cambé, Estado do Paraná, Brasil*. Res. Soc. Dev., 2020. DOI: https://doi.org/10.33448/rsd-v9i8.5912.
4. INEP, *ENADE-Microdados*, 2020. DOI: http://inep.gov.br/web/guest/microdados.
5. Provost, F., Fawcett, T., *Data Science Para Negócios*, 2016. 1ra Ed. Alfa: Rio de Janeiro.
6. Romero, C., Ventura, S., *Data mining in education*. WIREs Data Min. Knowl. Discov., 2013. **3**(1): p. 12-27. DOI: https://doi.org/10.1002/widm.1075.
7. Araújo, F.H.R., *Determinantes do desempenho dos alunos das Instituições de Ensino Superior Brasileiras no ENADE 2013*, 2017. Dissertação de Mestrado. Universidade Federal do Ceará.
8. Rocha, A.L.P., Leles, C.R., Queiroz, M.G., *Fatores associados ao desempenho acadêmico de estudantes de Nutrição no Enade*. Rev. Bras. Estud. Pedagógicos, 2018. **99**(251): p. 74-94. DOI: https://doi.org/10.24109/2176-6681.rbep.99i251.3162.
9. Filho, A.E.C.M., Silva, L.S., Junior, J.A.F.P., *Características da família dos estudantes associadas ao desempenho acadêmico no exame nacional de desempenho dos estudantes (ENADE)*. Rev. Educ. Psicol. E Interfaces, 2019. DOI: https://doi.org/10.37444/issn-2594-5343.v3i3.163.
10. Guimarães, L.M., *Resultados do ENADE 2012: contribuições ao curso de logística da Universidade Federal do Tocantins*, 2020. Available at: http://repositorio.uft.edu.br/handle/11612/1832.
11. Handro, C.B., Pereira, C.S., Lordelo, S.N.B., Nogueira, T.B.R., Conceição, V.M., *Análise da Inserção Profissional e das habilidades do Técnico em Logística: Perspectiva de Egressos e Empresas*. Educ. Prof. E Tecnológica Em Rev., 2020. **4**(3): p. 194-217.
12. Cohen, M., Mourão, J.F., Kamlot, D., *Um Estudo dos Fatores Associados ao Desempenho na Prova do Enade de Engenharia de Produção em Santa Catarina*. 9 Simpósio de Integração Científica e Tecnológica do Sul Catarinense, 2020. Available at: https://eventoscientificos.ifsc.edu.br/index.php/sictsul/9-sict-sul/paper/view/3106.
13. Faria, E.O., *Uma análise dos fatores determinantes do desempenho dos alunos dos cursos superiores em Administração do Distrito Federal*, 2017. Dissertação de Mestrado. Universidade de Brasília.
14. Silva, A.R.F., Santos, A.P., Nogueira, V.S., *TICs na Educação: Uma Reflexão sobre a Importância da Educação à Distância para a Formação Continuada e Ressignificação das Práticas Pedagógicas*. IGAPÓ-Edições Especiais, 2020. Available at: http://igapo.ifam.edu.br/ojs/index.php/EDIES/article/view/782.

Conflict of interests

There are no conflicts of interest related to the development of this study.

Authors' contributions

Ivonaldo Vicente da Silva. ORCID: https://orcid.org/0000-0003-3364-1421
Conception, objective and method. Bibliographic review. Data collection, mining and analysis. Essay.

Márcia Terra da Silva. ORCID: https://orcid.org/0000-0002-5314-4978
Conceptualization of ideas, design of methodology, validation, Writing, Review and Editing.

Pedro Luiz de Oliveira Costa Neto. ORCID: https://orcid.org/0000-0002-6987-2996
Conceptualization of ideas, design of methodology.