Revista ARACÊ
ISSN: 2358-2472

EDUCATIONAL DATA MINING APPLIED IN THE ANALYSIS OF THE PROFILE OF GRADUATES OF PUBLIC TECHNICAL SCHOOLS IN THE STATE OF SÃO PAULO WHO PASSED ENTRANCE EXAMS

https://doi.org/10.56238/arev7n4-157

Submitted on: 03/15/2025 **Publication date:** 04/15/2025

Pamela Ferreira Alves Andrelo[1], Dacyr Dante de Oliveira Gatto[2], João Rafael Gonçalves Evangelista[3], Marcos Antônio Gaspar[4], Ricardo Pinto Ferreira[5] and Renato José Sassi[6]

ABSTRACT

The Public Technical Schools of the State of São Paulo (ETECS) have sought improvements aimed at the training of their students, through the analysis of data from their graduates who passed entrance exams. Educational Data Mining (EDM) that uses Artificial Intelligence and Statistics techniques can be used to analyze this data. The objective of this study was to identify and analyze the profile of graduates of public technical schools in the State of São Paulo who passed entrance exams, using Educational Data Mining, to implement changes that support educational management. The ETEC Paulistano located in Jardim Paulistano, object of study of this work, is located in the center of a highly populated region, but lacking in parks, schools, infrastructure and jobs. The school has a student body composed of 95% of students from the Vila Brasilândia region, 70% of whom come from the municipal public school. A questionnaire was applied, divided into two parts, to the graduates of the classes of the Etim Informatics and Etim Environment courses graduated

[1]Master in Informatics and Knowledge Management
Nove de Julho University
São Paulo State Technical School
pamelafaandrelo@uni9.edu.br
https://orcid.org/0000-0003-2252-1504
[2]Dr. in Informatics and Knowledge Management
Nove de Julho University
dacyr.gatto@uni9.pro.br
https://orcid.org/0000-0003-2146-4819
[3]Dr. in Informatics and Knowledge Management
Nove de Julho University
joao.rafael@uni9.pro.br
https://orcid.org/0000-0003-3541-8354
[4]Dr. of Business Administration
University of São Paulo
marcos.antonio@uni9.pro.br
https://orcid.org/0000-0002-2422-2455
[5]Dr. in Informatics and Knowledge Management
Nove de Julho University
somkohonem@gmail.com
https://orcid.org/0000-0003-1817-575X
[6]Dr. in Electrical Engineering
University of São Paulo
sassi@uni9.pro.br
https://orcid.org/0000-0001-5276-4895

Revista ARACÊ
ISSN: 2358-2472

in the years 2019 and 2020. To mine the results, five steps were applied: Selection and Data Collection; Data Pre-Processing; Educational Data Mining; Interpretation and Analysis of Discovered Knowledge and Implementation of Changes. The analysis of the graduate's profile made it possible to implement and predict changes aligned with the interest of those involved with the school. It was concluded that MDE is an important ally of ETEC Paulistano to support educational management in the development of educational strategies.

Keywords: Educational Data Mining. ETEC. Profile of Graduates. Paula Souza Center. Educational Management.

Revista ARACÊ
ISSN: 2358-2472

INTRODUCTION

Access to education in higher education can allow a young person to overcome the difficulties imposed by poverty and marginalization, getting a job or even becoming an entrepreneur, but this access is still difficult. To change this situation, the Public Technical Schools of the State of São Paulo (ETECS) have sought improvements in their decision-making aimed at attracting this young person, through the analysis of data from their students and graduates with the help of technology (CENTRO PAULA SOUZA, 2022).

These improvements aim to support educational management in the development of strategies such as the adaptation of educational curricula, which privilege both technical education and training for students to pass entrance exams. It is noteworthy that, despite the quality technical education offered by ETECS, graduates take entrance exams in higher education institutions to find better job opportunities and salaries (CORDAS, 2016).

Among the ETECS, the São Paulo State Technical School, called ETEC Paulistano, located in Jardim Paulistano, is located in the center of a highly populated region, but lacking in parks, schools, infrastructure and jobs. The student body of ETEC Paulistano is composed of 95% of young people from the Vila Brasilândia region. Of this public, 70% of them come directly from the municipal public school, that is, the profile of the student at ETEC Paulistano is from a public school and from low income (ETEC PAULISTANO, 2022).

Access to university through the technical courses offered by ETEC Paulistano can enable students to better conditions to enter the job market or even to undertake, which enables the financial independence necessary to study and improve the quality of life of themselves and their families (SIMÕES, *et al.,* 2018).

In this way, for ETEC Paulistano it is important not only to analyze the information of the students, but also to identify and analyze the profile of the graduates approved in the entrance exams, as it is an educational management instrument necessary to support teaching strategies to improve the quality of teaching, in addition to bringing more candidates to the institution's selection process.

Part of the immense volume of educational data generated by ETEC is not analyzed. In this way, valuable information is lost that could, for example, identify and monitor both the profile of the student and the graduate (ETEC PAULISTANO, 2022).

Educational Data Analysis with its tools can be applied to analyze student and graduate data generated in educational environments. One of the aspects of Educational Data Analysis is Educational Data Mining (EDM), a field of data exploration where intelligent

Revista ARACÊ
ISSN: 2358-2472

algorithms are applied in order to generate new knowledge aimed at supporting educational management (PRANAV, *et al.,* 2021).

The application of the MDE focuses on analyzing the information obtained to support decisions, considering the profile of the approved graduates and also enabling their monitoring, in order to support educational management. The applications of MDE are diverse in the area of Education, ranging from the identification and analysis of the factors that influenced student dropout in higher education (COSTA, 2021; BRITO, *et al*. 2019) to the identification of factors for improvement in the IDEB in municipal schools (PINTO, 2019).

In view of the above, the objective of this work was to identify and analyze the profile of graduates of public technical schools in the State of São Paulo who passed entrance exams, using Educational Data Mining, to implement changes that support educational management

In addition to the Introduction, this work was structured in five sections: in section 2 the theoretical foundation on the main themes is presented, in section 3 the materials and methods used in this work are addressed, in section 4, the results and analyses of the experiments carried out are presented and in section 5 the conclusion is presented.

THEORETICAL FOUNDATION

In this chapter, the theoretical foundation on Public State Technical Schools and Educational Data Mining is presented.

PUBLIC STATE TECHNICAL SCHOOLS

The management of the high school technical courses offered by ETECS is the responsibility of the Paula Souza Center (CPS) linked to the Secretariat of Economic Development of the government of the state of São Paulo with the purpose of expanding technical and professional education (CENTRO PAULA SOUZA, 2022).

According to the Paula Souza Center website, the State Technical Schools (ETECS) are characterized by: 223 ETECS, distributed across 165 municipalities in São Paulo. ETECS serve more than 208 thousand students in Technical Education, 151 of which are courses, aimed at all public and private productive sectors (CENTRO PAULA SOUZA, 2022).

Revista ARACÊ
ISSN: 2358-2472

ETEC Paulistano, located in Jardim Paulistano, a district of the Brasilândia neighborhood, is located in the center of a highly populated and needy region. The school's student body is composed of young people from the Brasilândia region, some of whom come from even more distant regions, such as Parada de Taipas and Perus, coming from public schools and low-income (ETEC PAULISTANO, 2022).

In view of this scenario, ETEC Paulistano's mission is to be a reference for a place for study, professionalization, culture and leisure for the population of the region with the objective of transforming the lives of students and their families. The school has done this through professionalization that helps many to get their first job and enter public and private universities. It has been playing this role well by ranking above the university access indexes proposed by the Basic Education Development Index IDEB (ETEC PAULISTANO, 2022).

In addition, ETEC Paulistano is ranked 27th among the best public schools in São Paulo, and 3,220th among the 27,780 schools analyzed by IDEB (ETEC PAULISTANO, 2022).

The first classes entered in 2009, in the High School, Etim of Informatics and Etim of Environment courses in the morning. The first classes of the morning period graduated in 2012, and in 2016 the offer of High School courses ended, so from that year on, only the Etim courses remained.

Although technical professional education is an interesting object of study, due to its various aspects, there is a need to identify and analyze the profile of graduates, since it focuses on educational development, school management and the development of strategies. The application of Educational Data Mining (EDM) can help in this identification and analysis Khelifi, *et al.* (2024).

EDUCATIONAL DATA MINING

In the educational context, data can be from different sources, such as virtual learning environments, questionnaires, teachers' websites, academic systems, intelligent tutoring systems, among others, and provide information about students, graduates, teachers and the educational contexts in which they are inserted (NASCIMENTO *et al.,* 2016).

For Cerezo, *et al.* (2024), Educational Data Analysis, is divided into three themes: *Academic Analytics*, *Learning Analytics,* and *Educational Data Mining* (MDE).

Revista ARACÊ
ISSN: 2358-2472

Academic Analytics (AA) seeks to understand the students' registration data, data that comes from the information systems of the Educational Institution itself, and relate it to the academic experience of students in the Institution (CEREZO, *et al.,* 2024)..

The term *Learning Analytics* (LA), according to Siemens (2013) is a process for intermediation, collection, analysis and relationship of student data and their learning environment, with the purpose of optimizing and understanding the teaching-learning in which this process takes place.

On the other hand, Educational Data Mining (EDM) aims to discover knowledge through the behavior of students and the scenario where they learn, enabling elements for the teacher, educational manager or even the student to evaluate any patterns discovered (CEREZO, *et al.,* 2024).

The objective of MDE is to assist in the teaching-learning process, school infrastructure, educational projects and other aspects that influence educational development. In this scenario for more assertive decision-making, a lot of work is being developed using educational data mining (NASCIMENTO, *et al.,* 2016).

For Couto (2017), both in face-to-face teaching and in distance education, the MDE can be applied. In both contexts, the application of techniques for knowledge discovery has distinctions regarding the sample of data and objectives.

MDE is an area that seeks to develop and apply algorithms for Knowledge *Discovery in DataBases* (KDD), in order to identify patterns from information related to education. The MDE was originally developed in the Data Mining area and uses the same phases and tasks provided for in the KDD process.

In Brazil, Brandão, *et al.* (2003) were one of the first to apply MDE in Brazil when they analyzed data from the National Program of Informatics in Education to identify patterns and verify critical factors and impacts on educational projects.

According to Fayyad, *et al.* (1996), the KDD process consists of five stages, with Data Mining being considered the main stage. It is noteworthy that the process for knowledge discovery is the same used for MDE and Data Mining, the main difference being that the first focuses on knowledge discovery in the area of education.

The phases of KDD are described below:

- **Data Selection:** After defining the objective to be achieved, the focus becomes the choice or selection of the data to be mined.

Revista ARACÊ
ISSN: 2358-2472

- **Pre-Processing:** The objective is to ensure the quality of the data by cleaning the data.
- **Data transformation:** The goal is to standardize data by applying the process of normalization or discretization of attributes, improving the understanding of the data.
- **Data Mining:** is characterized by the existence of a mining algorithm that must be able to efficiently extract implicit and useful knowledge from a database.
- **Interpretation and Evaluation of Results:** it is the phase in which the most interesting criteria are identified, among the patterns extracted in the Data Mining stage, according to the criteria established by the user. At the end of the evaluation, the knowledge discovered can be implemented and incorporated.

Several studies have applied EDM, with emphasis on Couto (2017), Digiampietri, *et al.* (2016) and Brandão, *et al.* (2003) that they used in the prediction of school dropout in higher education, Silva (2021) used in the prediction of school dropout in high school, Arun (2021), Pranav, *et al.* (2021) and Soni *et al.* (2018) used Sarker, *et al.* (2024) used in the analysis of the academic performance of students (2024), Nascimento, *et al.* (2016) applied in the analysis of education indicators in INEP (National Institute of Educational Studies and Research Anísio Teixeira) databases. Filho and Brum (2024) applied Artificial Intelligence, Mathematical Modeling, and Problem Solving as enhancers in the process of teaching and learning mathematics in the seventh year of elementary school

Kampff (2009) used the generation of alerts in virtual learning environments to support teaching practice, Lima (2020) applied it to the redistribution of the education districts of the city of Fortaleza, Okewu, *et al.* (2021) investigated the application of artificial neural networks for MDE in Higher Education through a systematic review of literature, Barbeiro, *et al.* (2024) and Khelifi, *et al.* (2024) conducted literature reviews on DEM. Cerezo, *et al.* (2024) formed five groups: the general purposes of application and examples of application in a school focused on problem solving, with the support of KDD tasks for MDE.

Goldschmidt, *et al.* (2015) report that the Brazilian federal government created the "One Computer per Student" (UCA) project, inspired by the *One Laptop per Child* (OLPC) initiative of the Media Laboratory of the *Massachusetts Institute of Technology* (MIT) in which it distributed *laptops* to about 300 schools in the country, giving rise to the *Memore* project, which aimed to apply the MDE in the data generated by the UCA project, to support

Revista ARACÊ
ISSN: 2358-2472

the decisions of teachers and educational managers in the teaching-learning process of these students.

According to Goldschmidt, *et al.* (2015) there are different groups and examples of interests related to the application of the DEM, as shown in Table 1.

Table 1: Interest Groups related to the implementation of the EAW

Group	Examples of Interests
Students Apprentices	Customizing environments
	Identifying learning strategies
	Supporting Resource Recommendations
	Suggested reinforcement tasks
	Identifying and addressing learning gaps
Instructors Teachers Tutors	Obtaining *objective* feedback
	Student learning and behavior analytics
	Identify students who require differentiated support
	Apprentice grouping
	Characterization of successful teaching strategies
	Improved customization of courses and teaching materials
	Prediction of student performance
Researchers Course and Teaching Materials Developers	Evaluation of teaching material
	Course Evaluation
	Improvement of student learning
	Automatic construction of student and tutor models
	Comparative study of mining techniques
	Development of mining tools
Organizations Sponsors	Course recommendation for student groups
	Improvement of the decision-making process in investments in Education
	Support in the student admission selection process
Managers Directors Administrators	Supporting investment in educational resources
	Efficient use of educational resources
	Instructor and curriculum evaluation
	Support in the configuration of websites and services

Source: Goldschmidt, *et al.* (2015).

Chart 1 shows the diversity of interest groups in the MDE, ranging from students and apprentices, through teachers, researchers to management.

MATERIALS AND METHODS

The research methodology adopted in this work was defined as applied, exploratory and qualitative research. Experimental research is also present, as it determines an object of study, selects the variables that would be able to influence it, defines the forms of control and observation of the effects that the variable produces on the object (YIN, 2016).

A questionnaire divided into two parts was applied to two groups of graduates. Yin (2016) defines a questionnaire as the investigation technique composed of a set of questions that are submitted to people with the purpose of obtaining information about

Revista ARACÊ
ISSN: 2358-2472

knowledge, beliefs, feelings, values, interests, expectations, aspirations, fears, present or past behavior.

LITERATURE RESEARCH

A bibliographic research was carried out, considering theses, dissertations, books, articles and conferences in the following databases: *Scholar Google, IEEExplore, Scielo, Scient Direct*, *Scopus and ProQuest,* on the central themes related to the work: Paula Souza Center, ETEC, Student Profile, Educational Data Mining, Public Technical School, MDE, Alumni Monitoring and Identification of the Profile of Alumni.

All the central themes were researched in Portuguese and English. Duplicate publications that did not fit the themes addressed in this study were disregarded. 144 publications related to the central themes were selected.

It was verified in the bibliographic research carried out that the main focus of the application of the MDE was the problem of student dropout. It was also observed that there was no application of the MDE in the monitoring of the graduate, current employability and relationship with the institution, most of the studies were carried out on the application of the MDE to analyze the learning and behavior of students on online teaching platforms.

It was concluded that the application of MDE in identifying the profile of ETECS graduates, for any purpose, is scarce, which shows the importance of this work and that there are few studies carried out in the context of face-to-face technical courses, with the exception of the works of Torres (2021) and Lima (2016), who investigated how much technical education influences the continuity of studies in higher education.

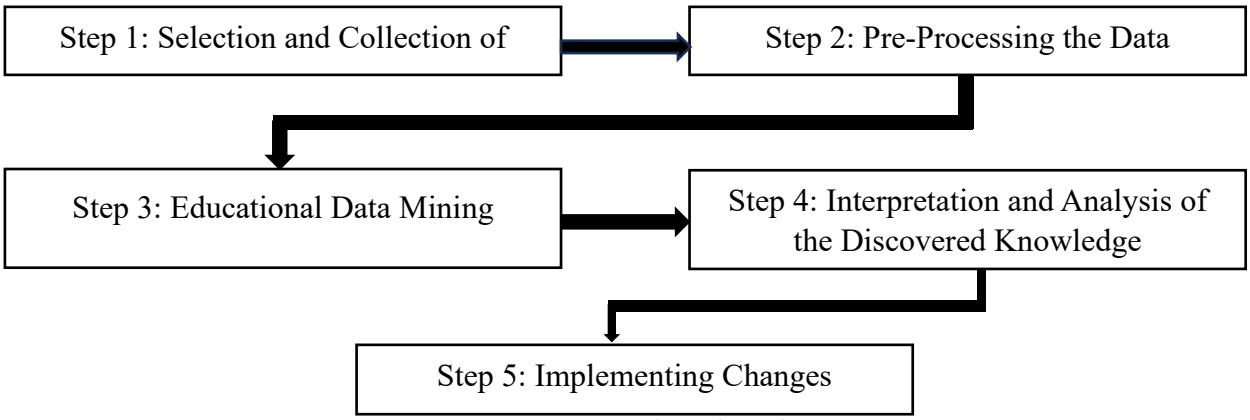
It is worth noting that the realization of this work found fertile ground, due to the real difficulty that ETEC Paulistano has with the organization of the information of the graduates, which makes it impossible to identify and monitor him to adapt the courses to the needs of approval of the entrance exams. It is also worth noting that the Director of ETEC Paulistano authorized the use of the school's data to carry out this work.

STAGES OF CONDUCTING COMPUTATIONAL EXPERIMENTS

The computational experiments for MDE were divided into five steps adapted from the phases of the KDD process, as shown in Figure 1

Revista ARACÊ
ISSN: 2358-2472

Figure 1: Stages of Conducting Computational Experiments



Source: Authors (2025)

The stages of computational experiments for the DE are described below.

Step 1 - Selection and Collection of Data:

The information was collected through the application of a questionnaire divided into two parts for graduates of face-to-face courses at ETEC Paulistano graduated in the years 2019 and 2020, in the Etim Informatics and Etim Environment courses, approved in entrance exams.

It should be noted that although the ETEC Paulistano courses are all face-to-face, the Etim Informática and Etim Environment classes of the 2020s studied in the *online* system, due to the COVID-19 pandemic.

The questionnaire contains open and closed questions. For some closed questions, the *Likert* Scale (LIKERT, 1932) was used. The scale adopted in the closed questions was as follows: 1 - Strongly disagree; 2 - Partially disagree; 3 – Indifferent; 4 - Partially Agree and 5 Totally Agree.

Thus, graduates who graduated from the Etim Informatics and Etim Environment courses were asked to answer the questions on *Google Forms.*

The questionnaire was divided into two parts:

- **Part 1:** composed of eight questions related to the identification of the profile of ETEC Paulistano graduates.

In this part, the graduates answered questions related to: age, area of approval in the entrance exam and types of scholarships received.

- **Part 2:** composed of fifteen questions related to the relationship with the course and with ETEC Paulistano.

Revista ARACÊ
ISSN: 2358-2472

In this part, the graduates answered questions related to their current occupation, salary range, whether the technical course at ETEC Paulistano helped them get a job, in addition to giving opinions about the course taken. The questionnaire also includes two descriptive questions that address their opinion about ETEC Paulistano and the course(s) that the school could offer. The questionnaire can be found in the Appendix of this work.

The questionnaire was sent to 69 graduates from 2019 and to 60 graduates from 2020, during the period from 08/28/2020 to 04/20/2021 for 2019 graduates and 04/20/2021 to 09/20/2021 for 2020 graduates.

Step 2 - Data Pre-Processing:

After tabulating the results of the questionnaire in spreadsheets, there was a need to pre-process the data. The pre-processing was carried out in five stages, described below:

a) **Anonymization of data.** The data used in this study from the answers to the questionnaire applied were anonymized, that is, they lost the possibility of direct or indirect association with an individual when their personal information was deleted.
b) **Removal of Spaces:** some values in the students' answers contained duplicate spaces between words, which were removed.
c) **Removal of Special Characters:** Removal of accents and other punctuation.
d) **Standardization of Similar Values:** Standardization of values that described the same result, such as "USP" and "University of São Paulo".

Step 3 - Educational Data Mining:

Power BI https://powerbi.microsoft.com/pt-br/ software was used to visualize the results to support Step 4.

Step 4 - Interpretation and Analysis of the Discovered Knowledge:

The visualization of the results seeks to enable the interpretation and evaluation of the knowledge discovered by a group of teachers, coordinators and the principal gathered with a view to implementing changes that support educational management.

Revista ARACÊ
ISSN: 2358-2472

Step 5 - Implementation of Changes:

It is expected that the analysis of the knowledge discovered will support the implementation of changes in the courses of ETEC Paulistano treated in this work. Implementation will depend on approval in course and pedagogical meetings, with the director of ETEC Paulistano, course coordinators and professors.

PRESENTATION, DISCUSSION AND ANALYSIS OF RESULTS

The number of responding graduates, considering the two classes, was 44 for the year 2019 and 31 for the year 2020, totaling 75 graduates.

The questionnaire was divided into two parts: Part 1: Identification of the Profile of ETEC Paulistano Graduates and Part 2: Relationship with the ETEC Paulistano Course.

PART 1: IDENTIFICATION OF THE PROFILE OF ETEC PAULISTANO GRADUATES

The analysis of the identification of the profile of the graduates of the Etim considered in the study revealed that part of the graduates of the Etim Informatics course continued their studies in higher education in courses related to the area of informatics and some graduates chose to continue their studies in different areas, such as Business Administration, Psychology, Law, Advertising and Propaganda, Pharmacy and Biomedicine.

Some graduates of Etim Environment chose to continue their studies in higher education in areas such as Biological Sciences, Letters, Journalism, Law, Accounting Sciences and some of them opted for Environmental Management, continuing in the same area of technical education training.

In the two courses evaluated, approval was observed in public universities such as USP, UFSCAR, Federal University of Santa Catarina, FATEC, Unesp and Unicamp, and in private universities such as Uninove, FMU, Impacta, Anhembi Morumbi, São Judas University, Mackenzie, FAAP and Fiap, among others.

The percentage of discount obtained with scholarships was diversified from 100% to 20%. The analysis shows that in relation to the type of scholarship obtained, the University for All Program (ProUni) had the largest amount, followed by scholarships provided by the educational institution itself, in third place comes the Sisu (Unified Selection System).

Revista ARACÊ
ISSN: 2358-2472

PART 2: RELATIONSHIP WITH THE ETEC PAULISTANO COURSE

With the analysis of the results obtained in part 2 of the questionnaire, it can be identified which profile of the ETEC Paulistano graduates manage to win their job vacancies and part of them in their area of training, the salary range was between R$ 998.00 and R$ 3,000.00, but most still find it difficult to enter the job market. Still, this majority agrees that the course helped them win their jobs, college, and even increase their salary. They have the necessary knowledge to perform their tasks in the job market and would recommend the course and ETEC Paulistano.

The analysis of the answers to the two essay questions applied, "What is your opinion about ETEC Paulistano?" and "Given the current scenario, what course(s) could ETEC Paulistano offer?" revealed that in general, the graduates believe that ETEC Paulistano is a good institution, which prepares the student for the job market, in addition to infrastructure being a relevant point; The faculty also stands out for being considered good professionals.

Regarding improvements, they highlight the curricular components that need to be updated, that the school is located in a place that offers danger and communication between the board and the student is insufficient. Among the courses that ETEC Paulistano could offer, according to the answers of the graduates are Pharmacy, Marketing, Arts, Accounting, Distance Learning Courses, Sports, Gastronomy.

ANALYSIS OF THE RESULTS

The analysis of the results obtained made it possible to implement changes in the courses of ETEC Paulistano divided into:

a) Implementation of Changes Based on the Analysis of the Results of the Questionnaire applied to graduates of the years 2019 and 2020 and;
b) Implementation of Changes that can be made based on the Analysis of the Results of the Questionnaire for the years 2019 and 2020.

The implementation of these changes is described below.

Revista ARACÊ
ISSN: 2358-2472

Implementation of Predicted Changes Based on Analysis of Questionnaire Results (2019 and 2020)

4 meetings were held with the institution's council, formed by the director, pedagogical coordinator, 4 course coordinators and 30 professors of the courses selected in this work to discuss the implementations and approve the implementations of the changes portrayed in Chart 2.

Chart 2 presents, based on the analysis of the results, the implementations of changes and their justifications that were included in the Multi-Year Management Plan (PPG) of ETEC Paulistano. The implementations were added to the PPG.

Chart 2: Implementation of Changes Based on Analysis of Questionnaire Results

Implementations	Justifications
Training of course coordinators and teachers	The teacher is one of the strengths of ETEC Paulistano. Therefore, it is necessary to train him more and more and also to encourage him.
Practical classes requested by the Environment class	They would like to have a greater number of practical classes in the curriculum. Among the classes, the Environment was the one that most requested this type of class.
Projects, intercourse and socio-emotional skills	They informed that the course helped in their personal life. Thus, encouraging greater interaction between students and greater interaction with different experiences, opinions and relationships is important for their personal lives.
Development of Projects to assist graduates in the provision of entrance exams.	The school can encourage and assist with projects to achieve better results. One of these projects is in the development phase and will aim to guide the student to write a quality essay for ENEM, ENADE.

Source: Authors (2025).

It should be noted that the institution's help is essential to encourage teachers to acquire more knowledge and update in the market, which, consequently, will bring the student a better education. Practical knowledge in classes, which aims to present experience in professional practice, can be acquired with intercourse projects to stimulate socio-emotional skills.

The analysis of the results continues, presenting the alignment and contribution to the interest groups of ETEC Paulistano, according to Goldschmidt et al. (2015), through the implementation of changes, presented in Chart 3.

Chart 3: Alignment and Contributions to the Interest Groups of ETEC Paulistano

Group	Examples of Interests
Apprentices	Identifying learning strategies
	Supporting Resource Recommendations
	Suggested reinforcement tasks
	Identifying and addressing learning gaps

Revista ARACÊ
ISSN: 2358-2472

Instructors Teachers Tutors	Obtaining *objective* feedback
	Characterization of successful teaching strategies
	Prediction of student performance
Researchers Developers of Managing Directors Administrators	Course Evaluation
	Improvement of student learning
	Efficient use of educational resources
	Instructor and curriculum evaluation

Source: Adapted from Goldschmidt *et al.* (2015).

Of the groups highlighted in gray, it is worth mentioning that in the group of Students and Learners, support resources, suggestions for reinforcement tasks, identification and treatment of gaps in the curricula of the courses and the addition of new courses implemented in the school and in the group of Researchers, Course Developers and Didactic Materials the application of the five phases of the computational experiments developed in this work were recommended.

Implementation of Changes that may be made based on the Analysis of the Results of the Questionnaire (2019 and 2020)

4 meetings were held with the institution's council formed by the director, pedagogical coordinator, 4 course coordinators and 30 professors of the courses covered in this work to discuss the implementations and approve the implementations of the changes portrayed in Chart 4.

It was observed that the infrastructure was one of the negative points of ETEC Paulistano. Therefore, the institution needs more investments from the CPS and also seek partnerships with private companies, since the intention is to train students to work in the job market, in addition these partnerships could also be a source of dissemination of internships and the possibility of workshops and even technical visits.

Chart 4 presents the implementations of changes that can be made based on the analysis of the results of the questionnaire, as well as their justifications.

Chart 4: Implementations of Changes that can be carried out based on the Analysis of the Results of the Questionnaire

Implementations that can be carried out	**Justifications**
Infrastructure – Increase CPS investment in ETEC Paulistano and seek partnerships with private companies to obtain financial resources and equipment	The graduates informed that the infrastructure of ETEC Paulistano is not the institution's strong point. The increase in the investment of the CPS in ETEC is fundamental and should be discussed again with the Center. Another option in parallel is to seek partnerships with private companies that are interested in investing in the school's infrastructure.

Revista ARACÊ
ISSN: 2358-2472

Creation of an Internship Center for internal disclosure of vacancies in companies, which may also be the companies mentioned in the item above.	The creation of an Internship Center, in addition to identifying vacancies, can refer the student based on the profile desired by the company.
Conduct technical visits with students in companies and workshops with professionals in the area.	In a competitive world, there is a need to prepare and understand the market, both to update the curriculum, and for the student to know how to direct their studies, in addition to contact with professionals in the area to bring market experience.
Partnerships with private colleges and universities.	It seeks to present to students teaching, **research** and **extension,** which form a mandatory tripod in universities and inseparable as pillars of a higher education institution. This partnership can motivate students to enter higher education teaching and research development, working as a researcher in stricto-sensu programs and in companies.

Source: Authors (2022).

The analysis of the results continues, presenting the alignment and contribution to the interest groups of ETEC Paulistano, according to Goldschmidt, *et al.* (2015), through the implementation of changes that can be made, presented in Chart 5.

Chart 5: Alignments and Contributions that may be made to the interest groups of ETEC Paulistano.

Group	Examples of Interests
Students Apprentices	Customizing environments
	Identifying learning strategies
	Supporting Resource Recommendations
	Suggested reinforcement tasks
	Identifying and addressing learning gaps
Instructors Teachers Developer Researchers	Obtaining *objective* feedback
	Characterization of successful teaching strategies
	Course Evaluation
	Improvement of student learning
Organizations Managing Directors Administrators	Course recommendation for student groups
	Efficient use of educational resources
	Instructor and curriculum evaluation

Source: Adapted from Goldschmidt *et al.* (2015).

With the investments in ETEC Paulistano and partners, it is possible to customize the environments, make new recommendations, invest in an appropriate dissemination for each course area, properly using the institution's resources.

CONCLUSION

The results obtained with the application of the questionnaire allowed the identification and analysis of the profile of the graduates, then starting to implement important changes that covered several sectors of the school.

Revista ARACÊ
ISSN: 2358-2472

It can now be considered that MDE is an important ally of ETEC Paulistano to support educational management in the development of educational strategies. In addition, the analysis of the results made it possible for the implementations of changes to be aligned and contribute to the interest of all those involved with the school.

Thus, by identifying the profile of the graduate, the institution will be able to evidence what impacts have been provided, such as: access to higher education courses, level of employability, and consequently, it will be able to assess whether the planning of its courses is obtaining the expected results. Thus, improvements can be applied in the strategic management of the courses and the institution.

The results obtained in this work confirm that the MDE can be successfully applied both in the identification and analysis of the profile of graduates and students, corroborating the studies of Sarker, *et al.,* (2024), Arun (2021), Costa (2021) and Pranav, *et al.* (2021).

The results of the work have several consequences, such as, for example, for the graduate and his family with the conquest of the job, coming from a quality technical course; for companies, the possibility of having candidates prepared to act in their positions; for society by making sure that the taxes collected revert to the improvement of the education offered by ETEC Paulistano; for academic research, with regard to the phases of carrying out computational experiments, become a roadmap for the use of the MDE in other ETECS and schools and for ETECs.

It can be highlighted as limitations of the work, the fact that much of the development was carried out during the pandemic period, which limited several activities, meetings, conversations and interviews, which could have been carried out with students, graduates, coordinators and professors of ETEC Paulistano.

As a continuation of the work, it is recommended to use the MDE in other areas of ETEC Paulistano, such as in the services directorate with data from teachers and school employees and in the academic data of the secretariat. It is also recommended to analyze the reasons that led some students of the two courses considered in this study to opt for areas different from the area of training in the technical course in the entrance exams. Finally, it is recommended to return to the analysis of the results of the questionnaire to compare the impact of the pandemic on the performance of the classes in the entrance exams and on the achievement of a job vacancy.

To conclude, it is hoped that this work will serve as a stimulus for the development of new research involving graduates of technical schools.

Revista ARACÊ
ISSN: 2358-2472

ACKNOWLEDGMENTS

To ETEC Paulistano and Universidade Nove de Julho for their support in carrying out this work.

Revista ARACÊ
ISSN: 2358-2472

REFERENCES

1. Arun, D. K., Kumar, S., & Rajesh, M. (2021). Student academic performance prediction using educational data mining. In 2021 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1–9). IEEE. https://doi.org/10.1109/ICCCI50826.2021.9457021

2. Barbeiro, L., Santos, J., & Silva, M. (2024). A review of educational data mining trends. Procedia Computer Science, 237, 88–95. https://doi.org/10.1016/j.procs.2024.05.083

3. Brandão, M. F. R., Silva, J. B., & Costa, A. G. (2003). Análise de agrupamento de escolas e núcleos de tecnologia educacional: Mineração na base de dados de avaliação do programa nacional de informática na educação. In Anais do Simpósio Brasileiro de Informática na Educação (SBIE) (Vol. 1, pp. 366–374). https://br-ie.org/pub/index.php/sbie/article/view/42

4. Brito, J., Costa, A. G., & Silva, M. P. (2019). Uso de mineração de dados educacionais para a classificação e identificação de perfis de evasão de graduandos em sistemas de informação. In Anais do VIII Congresso Brasileiro de Informática na Educação (CBIE) (pp. 159–168). https://br-ie.org/pub/index.php/wcbie/article/view/8796

5. Centro Paula Souza. (n.d.). Sobre o Centro Paula Souza. https://www.cps.sp.gov.br/sobre-o-centro-paula-souza/

6. Cerezo, R., Sánchez-Santillán, M., & Núñez, J. C. (2024). Reviewing the differences between learning analytics and educational data mining: Towards educational data science. Computers in Human Behavior, 154, Article 108155. https://doi.org/10.1016/j.chb.2024.108155

7. Cordas, V. F. C. (2016). Fundamentos e práticas em educação profissional e tecnológica: Maturidade vocacional e formação tecnológica (4th ed.). CEETEPS.

8. Costa, A. G. (2021). Aplicação de técnicas de mineração de dados e learning analytics para predição de evasão de alunos nos cursos de ciência da computação e engenharias da UFPel [Master's thesis, Universidade Federal de Pelotas]. http://repositorio.ufpel.edu.br/handle/123456789/1234

9. Couto, D. C. (2017). Mineração de dados educacionais aplicada à busca de perfis de alunos em casos de evasão ou retenção: Uma abordagem através de redes bayesianas [Master's thesis, Universidade Federal do Pará].

10. Digiampietri, L. A., Cerri, R., & Costa, A. G. (2016). Mineração de dados para identificação de alunos com alto risco de evasão: Um estudo de caso. Revista de Graduação USP, 1(1), 17–23. https://doi.org/10.11606/issn.2525-376X.v1i1p17-23

11. ETEC Paulistano. (2016). Plano pedagógico 2016. http://www.etecpaulistano.com.br/wp-content/uploads/2015/05/ppg_2016etecpaulistano.pdf

Revista ARACÊ
ISSN: 2358-2472

12. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI Magazine, 17(3), 37–54. https://doi.org/10.1609/aimag.v17i3.1230

13. Filho, J. K. Y., & Brum, V. F. M. C. (2024). O uso da inteligência artificial, modelagem matemática e resolução de problemas como potencializadores no processo de ensino e aprendizagem de matemática no sétimo ano do ensino fundamental. Revista Aracê, 6(4), 18299–18324. https://doi.org/10.56238/arev6n4-427

14. Goldschmidt, R., Passos, E., & Bezerra, E. (2015). Data mining: Conceitos, técnicas, algoritmos, orientações e aplicações (2nd ed.). Elsevier.

15. Kampff, A. J. C. (2009). Mineração de dados educacionais para geração de alertas em ambientes virtuais de aprendizagem como apoio à prática docente [Master's thesis, Universidade Federal do Rio Grande do Sul].

16. Khelifi, T., Amrani, R., & Bensaid, A. (2024). A comprehensive review of educational datasets: A systematic mapping study (2022-2023). Procedia Computer Science, 246, 1780–1789. https://doi.org/10.1016/j.procs.2024.09.682

17. Likert, R. (1932). A technique for the measurement of attitudes. Archives of Psychology, 140, 1–55.

18. Lima, L. J. (2016). Em busca do ensino superior: Trajetórias e estratégias de estudantes de uma ETEC da cidade de São Paulo [Master's thesis, Pontifícia Universidade Católica de São Paulo].

19. Lima, V. A. (2020). Aplicando mineração de dados educacionais para a redistribuição dos distritos de educação de Fortaleza. RENOTE: Revista Novas Tecnologias na Educação, 18(2), Article 110253. https://doi.org/10.22456/1679-1916.110253

20. Nascimento, R. L. S. D., Silva, J. B., & Costa, A. G. (2016). Mineração de dados educacionais: Um estudo sobre indicadores da educação em bases de dados do INEP. RENOTE: Revista Novas Tecnologias na Educação, 14(2), Article 1679-1916. https://doi.org/10.22456/1679-1916.68484

21. Okewu, E., Adewole, P., & Misra, S. (2021). Artificial neural networks for educational data mining in higher education: A systematic literature review. Applied Artificial Intelligence, 35(13), 983–1021. https://doi.org/10.1080/08839514.2021.1922847

22. Pinto, G. D. S., Costa, A. G., & Silva, J. B. (2019). Identificação dos fatores de melhorias no IDEB pelo uso de mineração de dados: Um estudo de caso em escolas municipais de Maceió. In Anais do VIII Congresso Brasileiro de Informática na Educação (CBIE) (pp. 1828–1837). https://br-ie.org/pub/index.php/wcbie/article/view/8854

23. Pranav, D., Kumar, S., & Singh, R. (2021). Educational data mining for predicting students' academic performance using machine learning algorithms. Materials Today: Proceedings, 47, 5260–5267. https://doi.org/10.1016/j.matpr.2021.05.646

Revista ARACÊ
ISSN: 2358-2472

24. Sarker, S., Islam, M. S., & Khan, M. A. (2024). Analyzing students' academic performance using educational data mining. Computers and Education: Artificial Intelligence, 7, Article 100263. https://doi.org/10.1016/j.caeai.2024.100263

25. Silva, A. S. (2021). Mineração de dados aplicada à predição da evasão escolar no ensino médio [Doctoral dissertation, Universidade Estadual Paulista].

26. Simões, C. F., Costa, A. G., & Silva, M. P. (2018). A relação entre a escolaridade e a pobreza: Uma análise das políticas para democratização do acesso ao ensino superior no Brasil. Revista Brasileira de Ensino Superior, 4(2), 21–43. https://doi.org/10.18256/2447-3944.2018.v4i2.2392

27. Siemens, G. (2013). Learning analytics: The emergence of a discipline. American Behavioral Scientist, 57(10), 1380–1400. https://doi.org/10.1177/0002764213498851

28. Soni, A., Kumar, S., & Sharma, R. (2018). Predicting student performance using data mining techniques. International Journal of Pure and Applied Mathematics, 119(12), 221–227.

29. Torres, N. A. (2021). Educação profissional técnica em administração de nível médio: Contribuições para o ingresso no ensino superior [Master's thesis, Universidade do Oeste Paulista].

30. Yin, R. K. (2016). Pesquisa qualitativa do início ao fim (2nd ed.). Penso.

Revista ARACÊ
ISSN: 2358-2472

APPENDIX: QUESTIONNAIRE APPLIED TO THE CLASSES OF 2019 AND 2020

Part 1: Identification of the Profile of Etec Paulistano Graduates

1) How old are you?
2) In which institution did you pass the entrance exam?
3) In which area did you pass the entrance exam?
4) Are you a public college scholar?
5) Specify the percentage of scholarship obtained.
6) Are you a private college scholar?
7) Specify the percentage of scholarship obtained.
8) What type of scholarship did you get?

Part 2: Relationship with the Etec Paulistano course

1) Do you currently work?
- Yes, in the same area where I graduated
- Yes, but in a different area from the one I graduated from
- No, I don't currently work

2) What salary range are you in?
- Less than R$998.00
- R$999,00 to R$.1,200.00
- R$1.201,00 to R$1.996,00
- R$1.997,00 to R$3.000,00
- Greater than 3,000.00
- No income

3) The technical course at ETEC Paulistano helped me get a job.
1. Strongly disagree
2. Partially disagree
3. Indifferent
4. Partially Agree
5. I totally agree

4) The technical course at ETEC Paulistano contributed to increase my salary.
1. Strongly disagree
2. Partially disagree
3. Indifferent
4. Partially Agree
5. I totally agree

5) I am able to professionally execute what I have learned, after completing the course
1. Strongly disagree
2. Partially disagree
3. Indifferent
4. Partially Agree
5. I totally agree

6) Taking the course helped in my personal life.

Revista ARACÊ
ISSN: 2358-2472

1. Strongly disagree
2. Partially disagree
3. Indifferent
4. Partially Agree
5. I totally agree

7) Taking the course helped in my professional life.
1. Strongly disagree
2. Partially disagree
3. Indifferent
4. Partially Agree
5. I totally agree

8) The curriculum, that is, the subjects offered are in line with what the job market is looking for in a professional.
1. Strongly disagree
2. Partially disagree
3. Indifferent
4. Partially Agree
5. I totally agree

9) Indicate what was the strong point of the course taken.
- Content of the curricular components
- School infrastructure
- Laboratories
- Teaching staff
- Other

10) Among the curricular components, mark the one that helped to win the place in college
- High School Subjects
- Technical subjects

11) Among the curricular components, mark the one that helped to win the vacancy in the current job.
- High School Subjects
- Technical subjects

12) How likely are you to recommend the course?
- Discharge
- Average
- Low

13) How likely are you to recommend ETEC?
- Discharge
- Average
- Low

14) What is your opinion about ETEC Paulistano?

Revista ARACÊ
ISSN: 2358-2472

15) Given the current scenario, what course(s) could ETEC Paulistano offer?