Ministry of Education
National University of Science and Technology
POLITEHNICA BUCHAREST
Doctoral School of
Industrial Engineering and Robotics

PhD THESIS

Research on the development of a prediction modelfor high school students' career decision-making

SUMMARY

PhD. Student,

Gabriel PETREA

Scientific coordinator,

Prof.univ.dr.eng. Cristian - Vasile DOICIN

Bucharest

- 2024 -

Table of contents

Introduction

In the current context, the younger generation is in constant interaction with the online environment, leading to a significant transformation in how they access information and make decisions. This continuous connection to the internet is redefining traditional sources of information and the ways in which young people are influenced in their educational and career choices. Unlike previous generations, who primarily relied on advice from parents, teachers, and mentors in the physical world, today's generation has access to a wide range of information sources that are quickly and instantly accessible.

Thus, young people today simultaneously access various sources of information. In addition to traditional advice from parents and teachers, they are significantly influenced by information found in the media and, more recently, on social media. Social networks have become a vast space where information circulates rapidly, providing access to a plurality of perspectives. While this exposure brings obvious benefits in terms of diversity of opinions and access to knowledge in multiple fields, it also comes with challenges, as it increases the risk of misinformation and influence from unauthorized or unreliable sources.

At the same time, the experiences of peers and friends have a considerable impact on young people's decisions. In this generation, the peer-to-peer community plays an essential role in social validation, and the opinions and experiences shared by peers can drive young people to follow certain educational paths or make specific career choices. These interactions are no longer limited to small circles but are amplified by digital platforms, where discussions can take place in large and diverse groups, from different geographical and cultural backgrounds.

Thus, in addition to traditional influences, today's generation is shaped by a series of modern factors specific to the digital era. The decisions young people make regarding their academic and professional futures are influenced by a combination of parental and teacher advice, information from the media and social media, as well as the experiences and opinions of peers and friends.

Justification for choosing the doctoral research topic

The purpose of implementing a mathematical model for predicting the career decisions made by high school graduates is to formalize an online prediction tool, design a platform or application that can anticipate the direction in which they will continue their studies, and later develop their professional careers.

The model is based on the analysis of data collected from students through a questionnaire that tracks various factors influencing their decisions regarding the university they plan to attend, the field of study they choose, or their professional orientations.

The main objective of this mathematical model was to identify and quantify the factors and the weights with which they influence students' decisions to choose a particular university. By thoroughly understanding these factors, universities can develop communication strategies and enhance their ability to attract more well-prepared students, better suited to the study programs offered by each faculty.

Part I. Curent state

Chapter 1. Opinion surveys

Research through opinion surveys refers to the method by which public opinion on a particular subject is assessed (Babbie, 2008).

The main characteristics of an opinion survey are: research on a large number of cases; cases are selected through sampling, following rigorous conditions; data collection takes place in normal life situations; the data is quantitatively measurable.

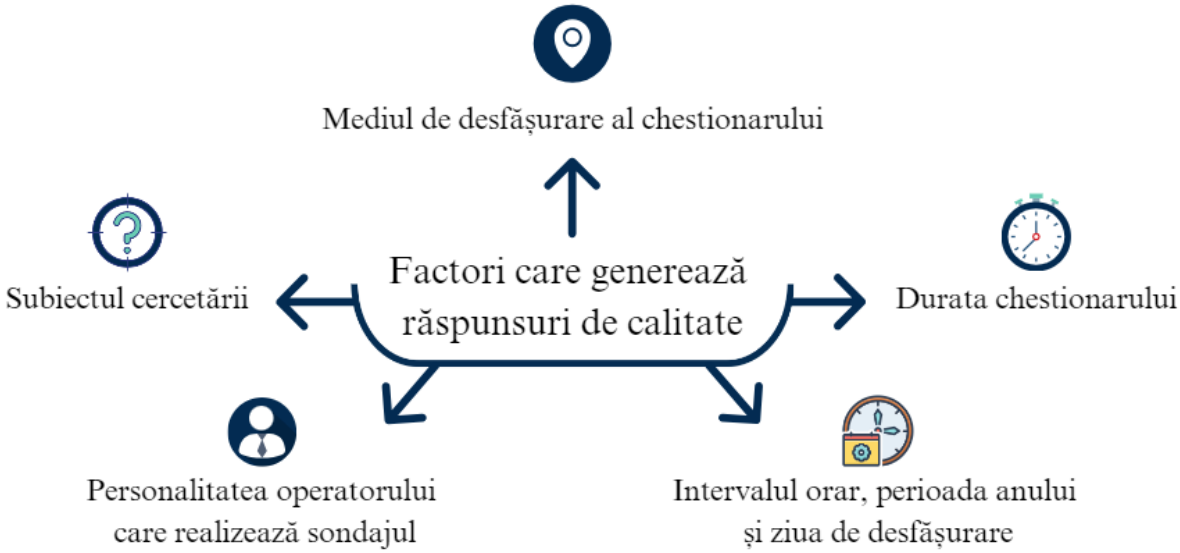

Fig. 1.4. *Factors generating quality responses*

The questionnaire is a data collection tool consisting of a set of questions designed to gather information from study participants (Alasuutari, Bickman & Brannen, 2008), (Brace, 2018).

Questionnaires can be classified into three main types (Fig. 1.5):

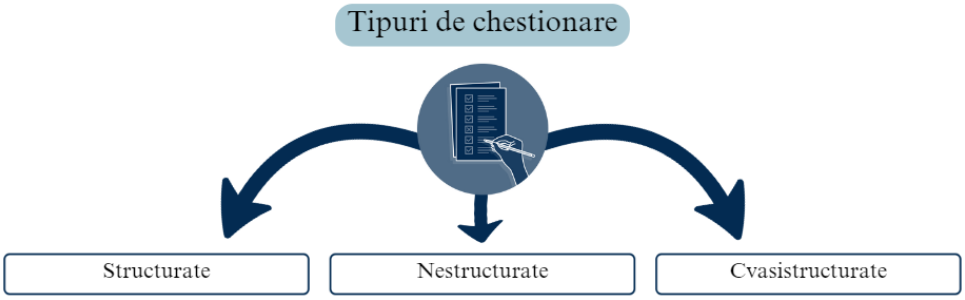

Fig. 1.5. *Classification of questionnaire types*

Over time, the techniques used for collecting data from opinion surveys have evolved alongside technological advancements. This evolution has transitioned from face-to-face

surveys and telephone surveys to email or online surveys. Each of these techniques has its own advantages and disadvantages. The following figure (Fig. 1.6) classifies the five techniques used in data collection (Babbie, 2008).

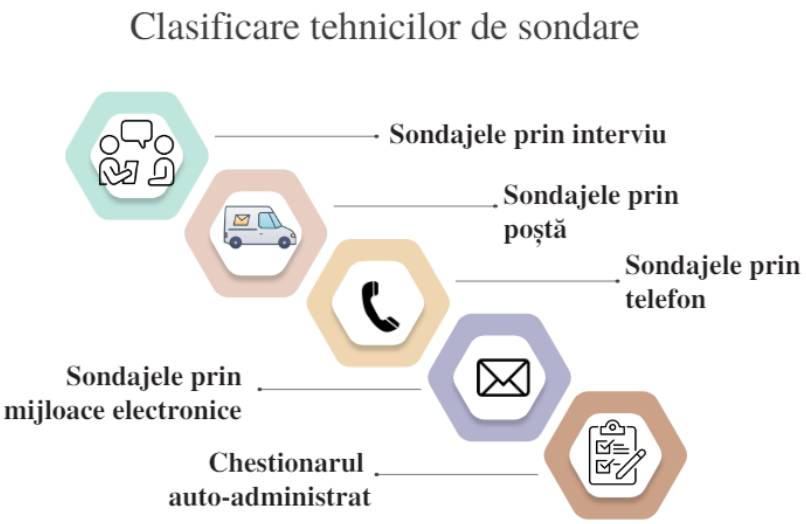

Fig. 1.6. *Survey techniques for conducting an opinion poll*

The design of the questionnaire can affect both the quality and quantity of the collected data and, consequently, the survey results. Questions are considered appropriate when two or more respondents interpret and understand the question in the same way (Foddy, 1993).

Field specialists recommend designing the questionnaire based on steps that contribute to achieving the best possible survey outcomes (Fig. 1.7).

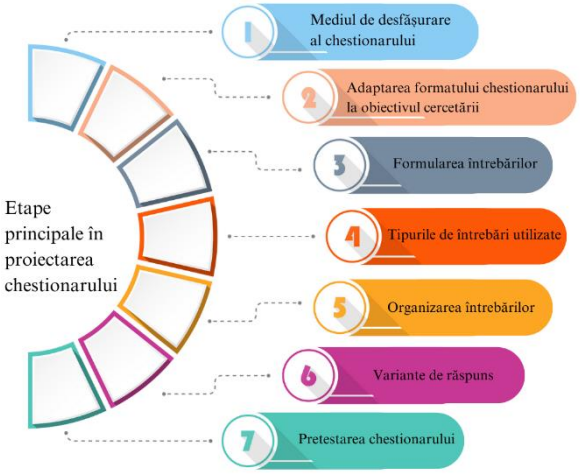

Fig. 1.7. *Main steps in questionnaire design*

An initial approach to career choice models was developed by Holland J. L. in 1976, proposing a new theory known as "Holland's Vocational Choice Theory." This theory posits that there is a correspondence between an individual's personality type and their work

environment, leading to career satisfaction and success. It represented a predictive model for career development and has been widely used in vocational counseling (Holland, 1997).

In the same year, Krumboltz and his collaborators developed a theory emphasizing the role of learning experiences and social interactions in shaping an individual's career options. The social learning model of career selection influenced how individuals can create and adjust their career preferences based on their life experiences (Krumboltz, Mitchell & Jones, 1976).

Later, in 2006, Thompson and Subich conducted a study to determine the connection between social status and the career decision-making process. The results demonstrated how socio-economic factors play an essential role in career choice decisions (Thompson & Subich, 2006).

In 2011, a study was conducted to model students' career preferences using neural networks. The predictions obtained were accurate and reliable, showcasing the ability of neural networks to capture the complexity of behavioral data (Paparrizos, Cambazoglu & Gionis, 2011).

Two years later, Guy and his collaborators utilized XGBoost, a data mining method, to identify students' career options. By employing the algorithm, they were able to model and predict students' behaviors regarding their plans after graduation (Guy, Avraham, Carmel, Jacovi & Ronen, 2013).

In 2015, a team of researchers employed another data mining method to determine the impact of behavioral data. It was shown that these real-time analyzed behaviors could reflect students' habits, skills, and preferences regarding the career choices they wish to pursue. The researchers succeeded in constructing predictive models for vocational counseling (Hadiji, Mladenov, Bauckhage & Kersting, 2015).

Also in the same year, several researchers used an innovative model titled "Approach Cluster Centers Based On XGBOOST" to collect data from various university campuses and predict students' career options. By utilizing Big Data for predicting career choices, the utility of this approach was demonstrated in comparison to the traditional techniques used up to that point. High-complexity data were employed to identify students' options in choosing their future careers (Xiong, Zhong, Deng & Yang, 2020).

Research through statistical surveys requires a considerable volume of research knowledge and experience, especially in the stages of sampling, questionnaire construction, and correct interpretation of the collected data. Depending on the encountered situations, sampling can be very simple or quite difficult, costly, and time-consuming.

The use of career choice modeling is an ongoing development. The utility and complexity of these approaches express valuable perspectives for creating new techniques in vocational counseling.

Chapter 2. Statistical Surveys

According to the Explanatory Dictionary of the Romanian Language (DEX), a statistical survey represents a "Method of researching public opinion, behavior, and attitudes of a social group by interviewing a representative sample."

Statistical methods are the tools and techniques used to analyze data and obtain relevant conclusions (Lawrence, 2014), (Mark, 2005).

Statistical methods include:

1. *Descriptive statistics*;
2. *Inferential statistics*;
3. *Relationship and association analysis*;
4. *Advanced analysis methods*;

Sampling is the process of selecting a group of subjects from a larger population that is being studied. Sampling in social research was developed alongside the implementation of political opinion polls, which were first applied in 1824 in the United States by a reporter to determine the favorite candidate in the presidential elections (Babbie, 2008).

The information collected from a sample can be used to draw conclusions about the entire population, as schematically represented in Fig. 2.1. The more representative a sample is (the quality of having the same distribution of characteristics as the population from which it is drawn), the more likely it is that all conclusions based on it will also be valid for the entire population (Lohr, 2021).

Fig. 2.1. *Schematic representation of sampling*

The sampling process can be divided into three stages: defining the population of interest, identifying the sampling frame, and selecting the sample.

1. Defining the population of interest;

2. Identifying the sampling frame;

3. Selecting the sample;

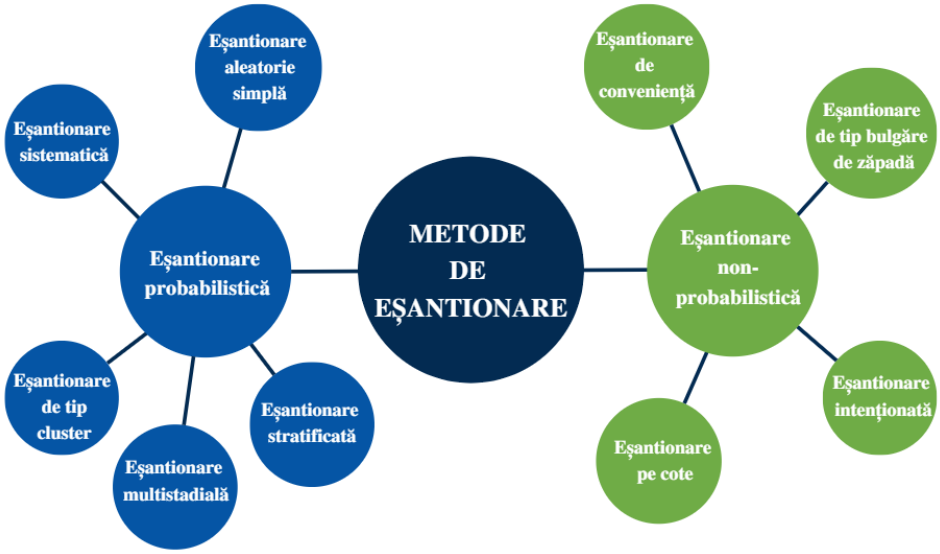

Fig. 2.2. *Classification of sampling methods*

Social research through opinion surveys plays an important role in collecting valuable data from individuals, groups, or populations, allowing researchers to gain insights into various social, economic, and scientific phenomena (Zhang, 2020).

Figure 2.5 presents the data preprocessing process:

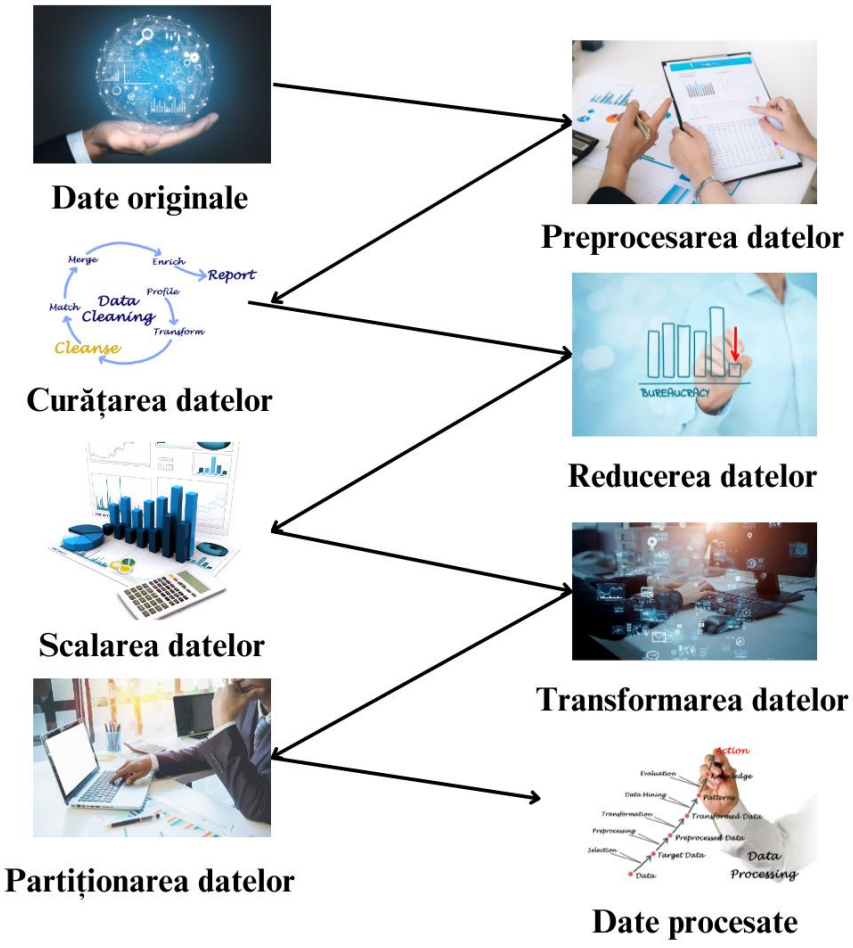

Fig.2.5. *Data preprocessing process (adapted to Zhang, 2020)*

Sampling errors occur when there is a discrepancy between the characteristics of the sample and the characteristics of the population from which the sample is drawn. Sampling errors can be of several types, as follows (De Leeuw, 2012):

1. Random sampling error;

2. Non-responses;

3. Coverage error;

4. Sampling frame error.

To assess the impact of sampling errors on the generalization of survey results, researchers rely on various measures that quantify the uncertainty introduced by the sampling process, among which the following can be mentioned (De Leeuw, 2012):

- *Confidence interval*;
- *Margin of error*;
- *Standard error*.

Recognizing the importance of minimizing sampling errors, researchers adopt rigorous strategies and sampling techniques to optimize the selection process and improve the quality of data collection. Some of the strategies frequently considered for minimizing sampling errors are (Lohr, 2019):

- *Random sampling*;
- *Non-response management*;
- *Improving coverage*.

Survey reliability refers to the consistency and stability of the survey instrument or the measurement tool used for data collection. Some of the techniques for measuring survey reliability include (Mark, 2005):

- *Test-retest reliability*;
- *Inter-rater reliability*;
- *Split-half reliability*.

Sampling is the most efficient way to select the main elements of a study. Through probabilistic sampling, cases where less representative categories for the population are selected are avoided. If the majority of the elements of a population have an equal chance of being selected, the resulting sample will have significantly greater representativeness. Furthermore, applying probabilistic sampling allows for the calculation of sampling error estimates, so the researcher can estimate the degree of expected error.

Chapter 3. Artificial Intelligence – Machine Learning – Deep Learning

Currently, the terms artificial intelligence (AI), machine learning (ML), and deep learning (DL) are frequently used to describe systems or software that exhibit intelligent behavior (Sarker, 2021). By definition, deep learning is a component of machine learning and, at the same time, part of the broader field of artificial intelligence (Fig. 3.1). Essentially, artificial intelligence involves imitating human behavior and thinking ability in machines or systems, while machine learning is the process by which machines learn from data or experiences, automating the construction of analytical models (Sarker, 2021), (Goodfellow, Bengio & Courville, 2016).

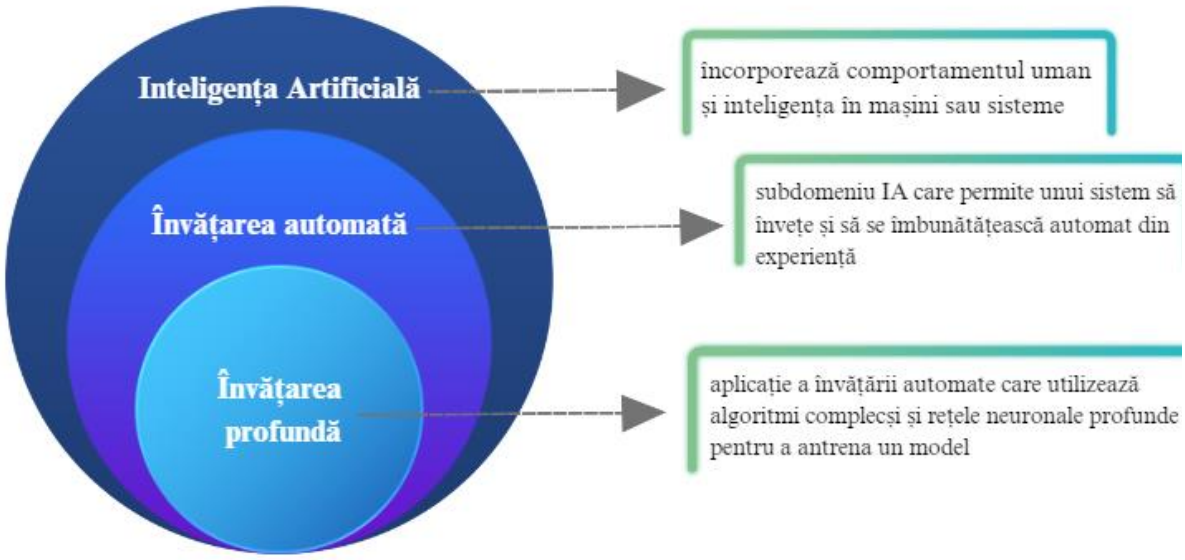

Fig. 3.1. *Artificial Intelligence, Machine Learning, and Deep Learning*

Artificial intelligence (AI) is a field of computer science that focuses on creating systems capable of performing tasks that require human intelligence. An artificial intelligence system is designed to process information, learn from data, and adapt its behavior based on new information received. Thus, AI not only performs predefined actions but can also improve its performance over time (Russell & Norvig, 2016).
For this study, 12 machine learning algorithms were analyzed on the same datasets to evaluate precision, accuracy, sensitivity, and other metrics to determine which of them provides the best results.
Regarding machine learning algorithms, there are several types, such as: supervised learning (Linear Regression, Decision Trees, Support Vector Machines, Random Forest), unsupervised learning (Principal Component Analysis), semi-supervised learning (Self-training, Graph Algorithms), reinforcement learning (Q-learning), deep learning (Convolutional Neural Networks), and so on.

***Tabel 3.1.**. Comparative analysis of machine learning algorithms*

Algorithm	Description	Learning Type	Model Type	Primary Problem	Data Type	Speed	Normalization	Advantages	Disadvantages
Linear Regression	A simple algorithm that models a linear relationship between inputs and a continuous numeric output variable	Supervised	Linear	Binary	Numeric	Low	No	• Explainable method • Interpretable results through its output coefficient	• Sensitive to outliers • Can underfit with small data and high dimensions
Logistic Regression	A simple algorithm that models a linear relationship between inputs and a categorical output (1 or 0)	Supervised	Linear	Binary	Numeric	Low	No	• Interpretable and explainable • Applicable for multi-class predictions	• Logistic Regression between inputs and outputs • Can overfit with small data and high dimensions
Ridge Regression	Part of the regression family—penalizes features with low predictive results by shrinking their coefficients	Supervised	Linear	Binary	Numeric	Low	No	• Less prone to overfitting • Best suited for data suffering from multicollinearity	• All predictors are retained in the final model • Does not perform feature selection

Algorithm	Description	Learning Type	Model Type	Primary Problem	Data Type	Speed	Normalization	Advantages	Disadvantages
Decision Tree	Decision tree models establish decision rules on features to make predictions	Supervised	Tree	Multi-class or binary	Numeric or categorical	High	No	• Explainable and interpretable • Can handle missing values	• Prone to overfitting • Sensitive to outliers
Random Forest	An ensemble learning method that combines the output of several decision trees	Supervised	Tree	Multi-class or binary	Numeric or categorical	High	No	• Reduces overfitting • Higher accuracy compared to other models	• Training complexity can be high • Not very interpretable
Gradient Boosting	Gradient boosting uses boosting to build predictive models from an ensemble of weak predictive learners	Supervised	Tree	Multi-class or binary	Numeric or categorical	Medium	Yes	• Better accuracy compared to other regression models • Handles multicollinearity and non-linear relationships	• Sensitive to outliers and can cause overfitting • Computationally expensive and complex
k-NN	k-NN is the most used clustering approach—determines K clusters based on Euclidean distances	Unsupervised	Clustering	Multi-class or binary	Numeric or categorical	Medium	Yes	• Can be used with large datasets • Simple to implement and interpret	• Requires the expected number of clusters beforehand • Issues with clusters of different sizes and densities

Algorithm	Description	Learning Type	Model Type	Primary Problem	Data Type	Speed	Normalization	Advantages	Disadvantages
Hierarchical Clustering	A "bottom-up" approach where each data point is treated as its own cluster, and the closest clusters are iteratively merged	Unsupervised	Clustering	Multi-class or binary	Numeric or categorical	Medium	Yes	• No need to specify the number of clusters • Resulting dendrogram is informative	• Does not always lead to the best clustering • Not suitable for large datasets due to high complexity
Apriori	A rule-based approach that identifies the most frequent itemsets in a dataset, using prior knowledge of frequent itemset properties	Unsupervised	Rules	Multi-class or binary	Numeric or categorical	Low	Yes	• Results are intuitive and interpretable • Exhaustive approach as it finds all rules based on confidence and support	• Generates many uninteresting itemset • Computationally and memory intensive • Results in many overlapping itemsets

Deep Learning (DL) is a powerful machine learning paradigm that relies on deep neural networks to solve complex problems and extract high-level representations of input data (Sarker, 2021).

The role of deep learning in machine learning is to enable neural networks to learn layered and abstract representations of data. These representations allow artificial intelligence systems to understand deep features and patterns in the input data, enabling them to make more accurate predictions and solve complex problems (Geron, 2019).

Neural networks are mathematical models inspired by the functioning of the human brain. They consist of a collection of computing units called artificial neurons, organized in layers.

Regarding future directions and challenges for Artificial Intelligence and Deep Learning, the focus should be on two aspects:

- Automating data labeling;
- Preparing data to ensure its quality.

Machine learning encompasses a range of techniques that allow computers to learn from data and improve their outcomes without being explicitly programmed. In the context of statistical surveys, supervised learning stands out as a key approach, where ML models are trained on labeled data to make predictions on new data.

The use of Artificial Intelligence in the educational system highlights its potential by facilitating certain processes, such as student admission to a specific university or faculty, predicting student performance, and optimizing both time and human resources.

Random Forest is the ideal algorithm for determining the weights of each coefficient in the prediction model due to its robustness, ability to capture complex relationships, feature importance, and efficient generalization. Compared to other algorithms, it provides an optimal balance between accuracy, interpretability, and performance.

Chapter 4. Analysis of the university admission process

Each year, high school graduates embark on an exciting journey to choose the university where they want to study (Soutar & Turner, 2002).

When discussing admissions, universities may consider aspects such as (Table 4.1):

Tabel 4.1. *The context of university admissions*

No.	*Criterion*	*Characteristic*
1.	Admission criteria	High school grades
		Standardized test scores
		Recommendation letters
		Personal statements
		Interviews
2.	Competition	High level of selectivity
		Increased demand for prestigious universities
		Growing pressure on applicants
		Increasing admission standards
		Completing study applications
		Searching for alternatives
3.	Diversity policy	Promoting equal opportunities
		Creating an inclusive learning environment
		Developing intercultural skills
		Enriching the educational experience
		Preparing for the real world
		Fair representation
		Culture of tolerance and respect
4.	International admission	Language requirements
		Academic evaluation
		Cultural criteria and adaptability
		Application documents
		Visas and legal regulations
		Pre-university preparation programs
		Benefits of international admissions
		Challenges
5.	Recruitment strategies	Participation in educational fairs
		Presentations in schools and online
		Educational marketing
		Partnerships with educational agencies
		Personalized offers
		On-campus events
		Personal communication
		Use of technology
6.	Involved expenses	Decision factor
		Impact on indebtedness

No.	*Criterion*	*Characteristic*
		Financial options
		Work schedule
		Financial aid programs
		Informative
		Promoting investment in education
		Solutions for financial diversity
7.	Technological evolution	Online admission platforms
		Virtual communication
		Social media
		Virtual tours
		Webinars and online informational sessions
		Mobile applications
		Artificial intelligence and chatbots
		Personalized communication

From the analysis of the specialized literature, the following key factors emerged, highlighting the importance of analyzing the student's journey (Fig. 4.1):

- *Fairness and transparency*;
- *Success prediction;*
- *Personalized guidance;*
- *Diversity and inclusion;*
- *Long-term success.*

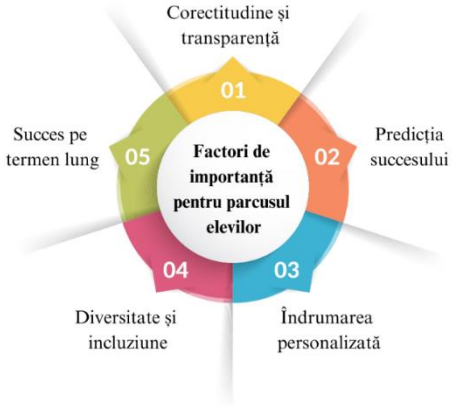

***Fig. 4.1**. Key factors in the analysis of students' journey*

Admission procedures in higher education show considerable diversity depending on the country, region, and private institutions. This diversity extends to every aspect of the admission process, encompassing the types of documents required for evaluation and the

methods used to reach final admission decisions. Moreover, some admission decisions are often centralized at the governmental level, while others are considered locally or institutionally (Selingo, 2020; Michael, 2020).

The influence of family on a student's university choice is a major factor that can significantly shape their decision-making process (Pimpa, 2003). In a study conducted in northeast Scotland, Gibbon concluded that parents are the primary influence on university selection, followed by friends, teachers, and school counselors (Gibbon, 1998).

Renowned universities are often synonymous with high-quality education, prestigious faculty, and exceptional resources (Briggs, 2006). Employment opportunities are a key factor that profoundly influences a student's choice of university (Soutar, 2002).

The geographical distance between a student's home and a university is a substantial factor that significantly shapes their decision-making process when choosing a higher education institution (Drewes & Michael, 2006). Universities offering substantial merit-based scholarships can attract high-achieving students (Shanka, 2006).

In education, the effective use of data has emerged as a transformative solution, reshaping how educators, decision-makers, and researchers approach improving learning outcomes and educational systems.

The context of university admissions is highly complex and influenced by a wide range of factors. The admission process is not merely a formality but reflects the evolving nature of higher education and society as a whole. Universities must strike a balance between admission criteria that ensure the quality of education and student diversity, while candidates must navigate the many aspects of the process to achieve their academic and professional goals.

Academic institutions must adapt their recruitment strategies and admission criteria to meet the current demands of society and the economy. They must promote diversity and inclusion, ensuring that every student has access to equitable educational opportunities. At the same time, universities need to find innovative ways to respond to technological challenges and connect with candidates in an increasingly virtual environment.

Thus, the context of university admissions is not merely a technical aspect of higher education but a field where educational policies, demographics, technology, and the individual needs of students intersect. Adaptability and innovation are essential to ensuring that the admission process remains relevant, inclusive, and equitable, offering students the chance to fulfill their aspirations and contribute to society's development.

Part II. Research on predicting students' university choice

Chapter 5. Doctoral thesis objectives and research methodology

The main objective of the doctoral thesis is to develop a predictive model of students' university choice based on information gathered from the conducted research.

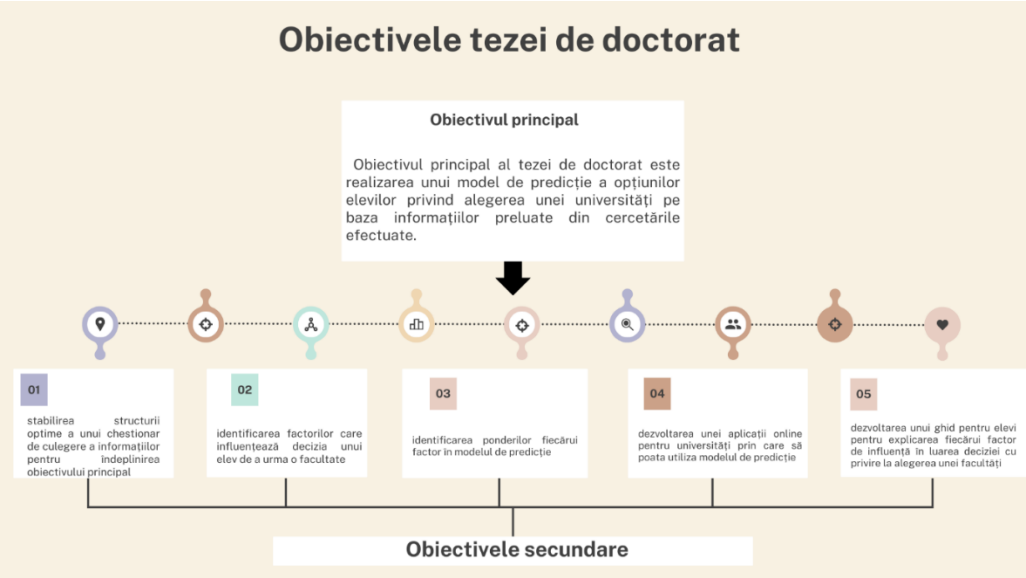

Fig. 5.1. *Doctoral thesis objectives*

The working methodology for developing the mathematical model to determine students' university choice follows several stages:

1. Critical Analysis of Specialized Literature: Review of key concepts and previous studies on the university admission process, focusing on analyzing students' decision-making pathways.
2. Initial Survey (Survey No. 1): Design of the first questionnaire aimed at identifying the factors influencing university choice.
3. Identification of Mathematical Methods: Research and identification of mathematical models and methods used in educational data analysis within the specialized literature.
4. Initial Data Collection (Dataset No. 1): Collection of initial data, including high school grades, standardized test scores, academic and extracurricular interests, and geographical preferences, to analyze student decisions.
5. Construction of the Mathematical Model: Development of a predictive mathematical model for students' university choice based on the identified influencing parameters.
6. Additional Surveys (Surveys No. 2 and 3): Design of two more questionnaires aimed at further identifying factors influencing university selection.
7. Collection of Second and Third Data Sets: Collection of additional data sets to gather more relevant information on high school graduates and their career decisions.
8. Individual Factor Analysis: Examination of each influencing factor individually to understand distributions, trends, and potential anomalies.

9. Factor Relationship Analysis: Investigate relationships between pairs of factors to identify correlations and patterns.
10. Fourth Survey (Survey No. 4): Creation of a fourth questionnaire to refine the understanding of factors influencing university choice.
11. Collection of Fourth Data Set: Gathering of data relevant to students' career decisions for further model development.
12. Feature Engineering: Identification and creation of relevant features to improve model performance, which may include generating new variables or combining existing ones.
13. Referential Survey: Conduct a referential survey to further identify influencing factors in students' university selection.
14. Collection of Referential Data Set: Gather referential data to enhance the model with more comprehensive insights into student career decisions.
15. AI Algorithm Training: Train an artificial intelligence algorithm to determine the weights of each influencing factor in the predictive model.
16. Factor Weight Determination: Calculate the weights of each influencing factor within the predictive model.
17. Model Refinement: Refine the predictive model for high school students' career decisions.
18. Model Testing and Validation: Test the model to ensure its validity and accuracy in predictions.
19. Performance Monitoring: Continuously monitor the model's performance and application to ensure prediction accuracy and correct operation.
20. Development of an Online Application: Create an online application for universities to use the predictive model for decision-making.
21. Development of a Student Guide: Develop a guide for students explaining the importance of each influencing factor in choosing a university.

Thus, the research process began with the hypothesis that the current generation is constantly connected to the online environment, uses multiple sources of information in the decision-making process, and is subject to a large number of influencing factors when deciding to attend a particular university.

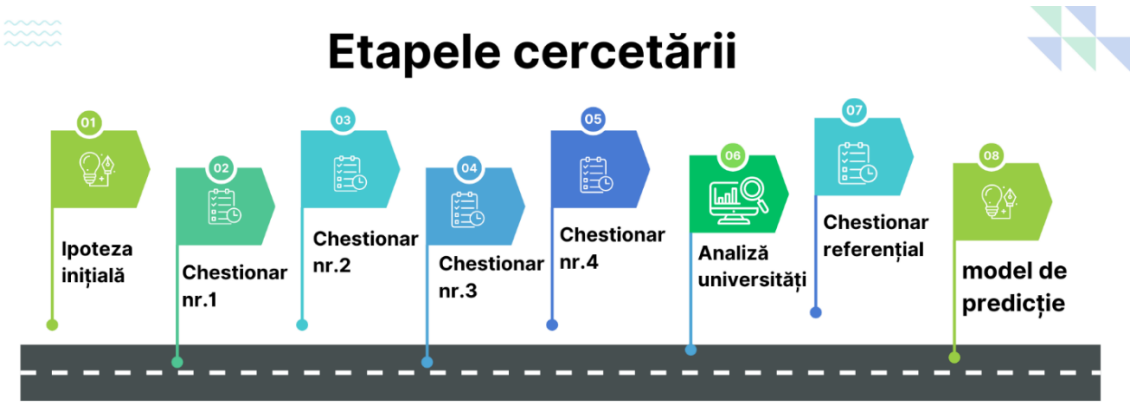

Fig. 5.3. *Research stages*

Chapter 6. Questionnaire design

After critically analyzing the literature in the field and identifying the parameters that form the basis of the mathematical model predicting how students choose a university, the process of developing four initial questionnaires was initiated. These questionnaires were designed to address all the parameters taken into account in the model.

A first questionnaire was developed to gather responses from candidates regarding various aspects such as:

- Sources of information,
- Opinions on the factors influencing educational decisions,
- Awareness of universities in Romania,
- The degree and manner of Internet use,
- The role of social networks in the choice of a university.

The questionnaire included the following sections:

1. General information;
2. Information about employment after graduation;
3. Information about the educational offer of the faculty the candidate wishes to attend;
4. Information about the candidate's complementary skills;
5. Information about accommodation at the university;
6. Factors influencing the choice of a faculty;
7. Awareness of universities in Romania.

Initial Research 1, conducted on a sample of 500 graduates, provided insights into how students make decisions regarding their choice of faculty. This research can be used to identify and validate the influencing factors for students and will contribute to the development of the final questionnaire aimed at identifying these factors and their respective weights in influencing university choice.

The results of Initial Research 1 were used to define and refine the questions in the final questionnaire. These findings offered a deeper understanding of how various factors, from social and economic influences to personal aspirations and family circumstances, contribute to students' educational decision-making processes.

Following the analysis of responses from **Initial Questionnaire 1**, four key factors were identified as influencing students' decisions in choosing a higher education institution:

1. University strengths,
2. Influencing factors,
3. Distance coefficient,
4. Labor market trends.

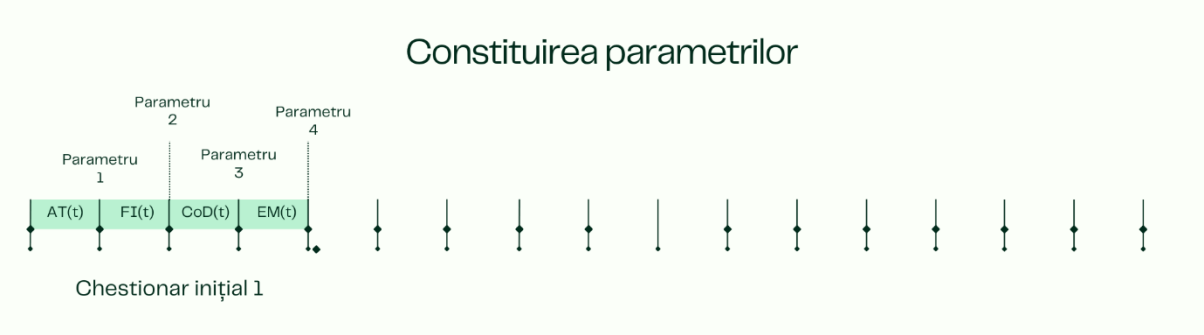

Fig. 6.2. *Establishing parameters - Initial questionnaire 1*

Based on these findings, the predictive model was formulated as a function dependent on these four factors:

$$C(t)=f(AT(t), FI(t), CoD(t), EM(t)) \quad (6.1)$$

A second initial questionnaire was conducted to determine how various factors influenced young people's perception of selecting a technical university. Information from the first questionnaire was used. The questionnaire was divided into 9 sections to group information by topic, as follows:

1. Section 1: Information regarding the use of the Internet and social media networks;
2. Section 2: Information about the faculty a candidate wishes to attend;
3. Section 3: Factors influencing the choice of a faculty;
4. Section 4: Awareness of universities in Romania;
5. Section 5: The academic situation of a candidate;
6. Section 6: The profession of the candidates' parents;
7. Section 7: Family income level;
8. Section 8: Labor market trends;
9. Section 9: Respondent identification data.

A total of 399 people responded to this questionnaire.

The additional parameters identified are: Expected salary level after graduation (SA(t)), Parents' profession (PP(t)), Immediate employability of graduates (A(t)), Family income (VE(t)), Level of scholarships offered by the university (NB(t)), and Influence of information available from various media sources (II(t)).

By adding these 6 new parameters to the predictive model, a more comprehensive and reality-adapted approach to students' decision-making processes is ensured. The revised model is:

$$C(t)=f(AT(t), FI(t), CoD(t), EM(t), SA(t), PP(t), A(t), VE(t), NB(t), II(t)) \quad (6.2)$$

This extended formula reflects the factors influencing high school students' career decisions, as derived from the analysis of Initial Questionnaire Number 2.

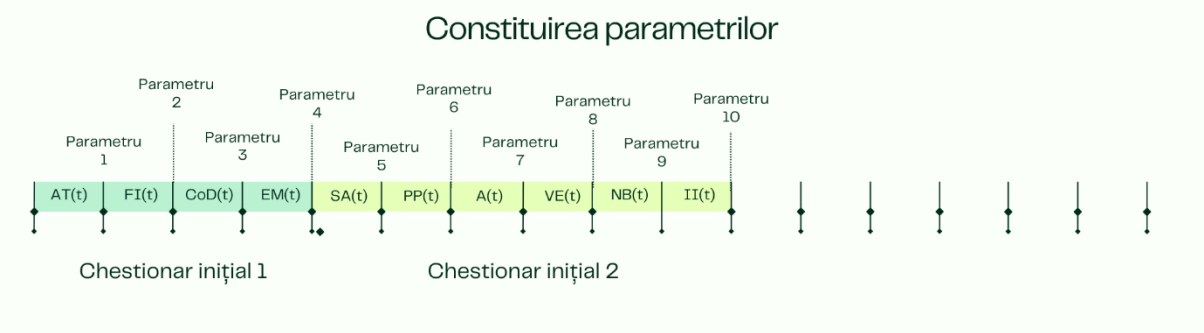

Fig. 6.4. *Establishing parameters - Initial questionnaire 2*

A third initial questionnaire was developed to identify the factors influencing students in choosing their educational path. The questionnaire consisted of 51 questions and was completed by 6,843 respondents.

Following the analysis of responses from Initial Questionnaire 3, in addition to confirming the factors already identified in Initial Questionnaires 1 and 2, a new factor contributing to high school students' career decisions was identified. The new factor is **Educational Path Test Results (REZ(t))**.

As a result, the predictive model formula was extended to include this new factor. Thus, the new formula is expressed as follows:

$$C(t)=f(AT(t), FI(t), CoD(t), EM(t), SA(t), PP(t), A(t), VE(t), NB(t), II(t), REZ(t)) \quad (6.3)$$

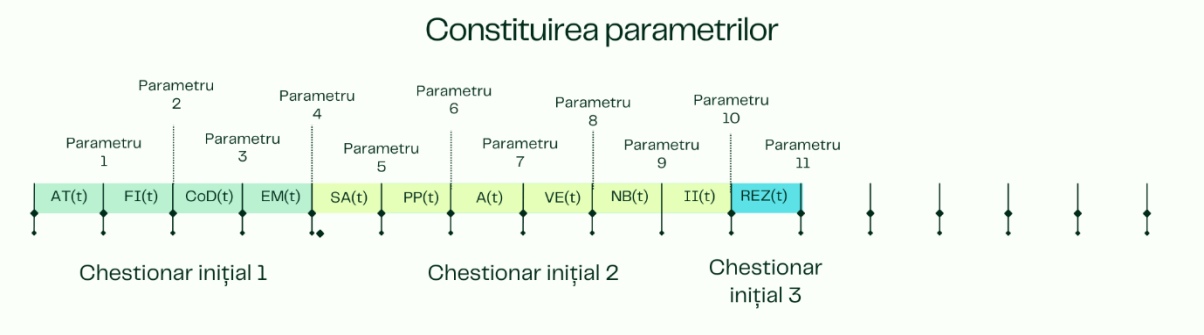

Fig. 6.6. *Establishing parameters - Initial questionnaire 3*

A fourth initial questionnaire was developed to identify the factors influencing students in choosing their educational path. The questionnaire consisted of 23 questions, grouped into 6 sections based on the topic they addressed.

As the research progressed, by applying Initial Questionnaire 4 to 4,084 students, an online validation of the already identified parameters was conducted, confirming their importance in the decision-making process for selecting a university. Additionally, the analysis of the results led to the identification of 2 new relevant parameters that contribute to shaping students' decisions. These 2 new parameters are **Economic Trends (EE(t))** and **Group Membership (AG(t))**, which are added to the existing mathematical model.

Thus, the updated formula of the predictive model becomes:

$$C(t)=f(AT(t), FI(t), CoD(t), EM(t), SA(t), PP(t), A(t), VE(t), NB(t), II(t),$$
$$REZ(t), AG(t), EE(t)) \quad (6.4)$$

After completing the analysis of responses from the referential questionnaire, two additional factors were identified that contribute to high school students' career decisions: **Academic Performance (SS(t))** and **Student Profile (PR(t))**. These factors provide added value in understanding the academic and personal influences on students' decisions to choose a university.

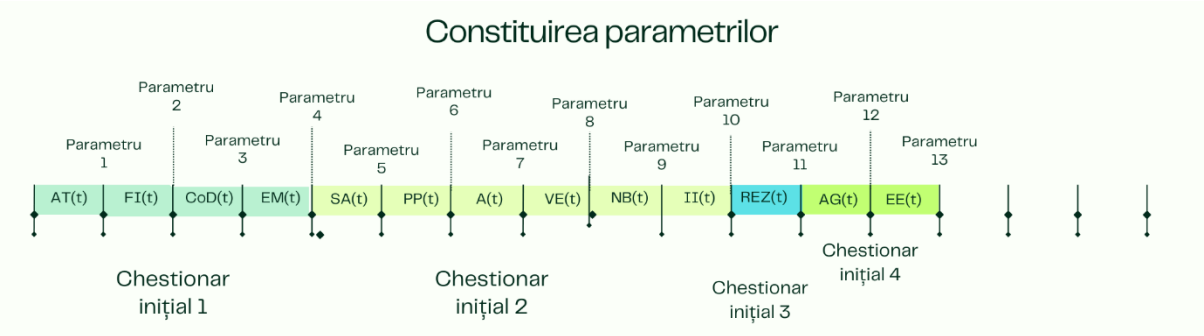

Fig. 6.7. *Establishing parameters - Initial questionnaire 4*

By integrating these 2 new factors, the predictive model becomes:

$$C(t)=f\,(AT(t), FI(t), CoD(t), EM(t), SA(t), PP(t), A(t), VE(t), NB(t), II(t), REZ(t), AG(t), EE(t), PR(t), SS(t)) \qquad (6.5)$$

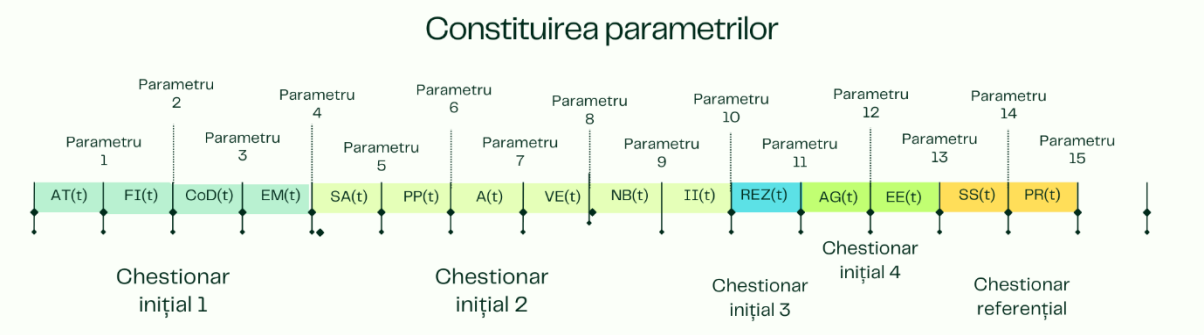

Fig. 6.10. *Establishing parameters - Referential questionnaire*

For the questionnaire of this study, the questions were addressed and completed by high school students, as they represent the specific category of respondents for the research. It was distributed to over 100,000 individuals, and 62,109 responses were received, of which 53,438 were valid.

Based on the centralization of the responses provided for each question, the response rate for each was calculated and presented in Table 6.2.

Tabel 6.2. *Response rate for each question*

Question	Question Variant	Response Rate
COVID-19 influenced your decision regarding	Choice of university	84,64%
	Subject preparation	85,27%
	School attendance	84,97%
	University studies decision	84,74%
	Academic results	85,24%
Factors contributing to university choice	Well-paid job	85,33%
	Feeling proud	85,16%
	Profession in demand	84,80%
	Field I like	84,77%
	Same profession as parents	84,55%

	Working in the desired company	84,53%
	Financial independence	84,42%
	Going abroad	84,29%
	Finding a job easily	84,85%
Influencers in university choice	Parents	85,28%
	High school teachers	84,77%
	High school classmates	84,63%
	Friends	84,38%
	Alumni of chosen university	84,59%
	Students of chosen university	84,38%
	Employers	84,25%
	Celebrities/public figures	84,41%
	Private tutors	84,16%
	Homeroom teacher	84,50%
Factors for choosing a university	Good professional training	77,38%
	More knowledge	77,22%
	Modern facilities at the university	76,94%
	University's international ranking	76,80%
	Budget (no tuition fee) seats	76,64%
	Low competition for admission	76,41%
	High scholarship level	76,70%
	Academic performance	76,57%
	Early admission	76,48%
Factors for choosing a university	New people and fun	77,19%
	A better future than my parents	77,07%
	Doing what I love	77,13%
	Desire to follow my dream	77,04%
	Proving that I can succeed in life	77,06%
	Earning a higher salary	77,01%
Factors for choosing a university	Overcoming material conditions	76,92%
	Keeping up with friends/classmates	76,88%
	Being part of a team	76,85%
	Contributing to society	76,82%
	Pleasing parents	76,66%
	Moving abroad more easily	76,66%
	Moving to a bigger city	76,69%
	Doing what society expects	76,60%
	Learning a future-proof profession	76,93%
Factors for choosing a university	High salary after graduation	77,12%
	High salary in chosen field	77,02%
	Job sustainability	76,60%
	Employment after graduation	76,89%
	Employment during studies	76,67%
	Labor market trends	76,60%
	Economic situation	76,60%
	Job abroad	76,78%
Factors for	Family income level	76,90%
	Distance from home to university	76,57%

choosing a university	Vocational test result	75,72%
	Parents' profession	76,57%
	Media information about the university	76,56%
	Opinions of alumni	76,44%
	Opinions of employers	76,28%
	Accommodation provided	76,32%
	Social media information	76,51%
True values	Well-prepared student	76,85%
	I choose my university by myself	77,45%
	I have constant internet access	76,69%
	The main source of information is the internet	76,54%
	The main source of information is television	76,38%
	I consult with parents when making decisions	76,67%
	Foreign universities are better than local ones	76,48%
	Pay attention to discussions about the field	76,40%
	I consult with private tutors	76,37%
True values	I know which university I'll attend	76,82%
	I will study abroad	76,54%
	I will attend a university just for the degree	76,44%
	University gives you a profession	76,34%
	Prefer universities with online studies	76,37%
	I've heard of early admission	75,80%
	Educational counseling is important	76,01%
	I did volunteer work in high school	76,15%
True values	Online learning during the pandemic was beneficial	76,96%
	Online learning during the pandemic affected knowledge	76,95%
	Online learning during the pandemic changed my university decision	76,64%
	Prefer online learning over face-to-face	76,69%
	It is difficult to understand during online learning	76,84%

It can be observed that the highest value exceeds 85%, while the lowest value is over 75%, specifically 75.80%. From this, we can conclude that the response rate was significant, considering that this questionnaire contained over 100 questions to which the young participants responded. This was due to the fact that there were questions where, in practice, responses were required for 10 options.

Through this questionnaire, considering the size and distribution of the dataset, it was structured in such a way as to obtain relevant information regarding the identification of decisions and preferences students have when choosing a university.

Chapter 7. Development of a mathematical model for predicting students' university choices

As a result of the analysis conducted in Chapter 4, regarding previous studies on students' decision-making process in university admission, a series of parameters have been identified that researchers have considered in this regard: family influence, university reputation, employability after graduation, or geographical distance from home (Bastedo, 2021), (IvyWise, 2023), (Michel & Pollard, 2020), (Selingo, 2020).

The initial parameters are considered to form the set $P_i = \{P_{i1}, P_{i2}, \ldots, P_{in}\}$, cu n = 7, as presented below:

P_{i1}: students' previous academic performance, highlighted by the grades obtained in high school. This was an important criterion, as it reflected their academic achievements and level of preparation;

P_{i2}: students' performance in standardized tests, another significant indicator of academic abilities;

P_{i3}: academic and extracurricular interests were also included, as there was a desire to better understand the fields that attract students and how these influence their choices;

P_{i4}: geographical preferences, through which an analysis of university locations was carried out and how this matched the students' preferences;

P_{i5}: students' motivation and confidence in the choice they made;

P_{i6}: the influence of family and close individuals in the decision-making process;

P_{i7}: the reputation of the university and employability after graduation.

The parameter selection process involved a careful evaluation of the relevance of each parameter and how it could influence students' choices. Consideration was also given to the possible interactions between these parameters to ensure that the proposed model could capture the complexity of students' decision-making.

The selected parameters included in the model form the set $P_f = \{P_{f1}, P_{f2}, \ldots, P_{fm}\}$, with m = 21, as presented below:

P_{f1}: Positive information, Ip, represents the number of positive pieces of information that appear across all media sources (print media, online media, TV, radio, outdoor), weighted by the credibility of the source.

P_{f2}: Negative information, In, refers to the number of negative pieces of information that appear across all media sources (print media, online media, TV, radio, outdoor), weighted by the credibility of the source.

P_{f3}: The influence of information, II = $\frac{\textbf{Positive information Ip}}{\textbf{Negative information In}}$ epresents the ratio of positive information to negative information.

P_{f4}: Positive influence, Fp, represents the set of all elements and individuals that can, voluntarily or involuntarily, change the evolution, thinking, or opinion of a person in relation to the institution.

P_{f5}: Negative influence, Fn, express the set of all elements and individuals that can, voluntarily or involuntarily, change the evolution, thinking, or opinion of a person in relation to the institution.

P_{f6}: Influence factors, FI = $\frac{\textbf{Positive influence Fp}}{\textbf{Negative influence Fn}}$, represents the ratio of positive influence factors to negative influence factors.

P_{f7}: Strengths, S, refer to the set of strong points of an institution in relation to direct competition**.**

P_{f8}: Weaknesses, W, are the set of weak points of an organization in relation to direct competition.

P_{f9}: The university's advantages, AT = $\frac{\textbf{Strenghts S}}{\textbf{Weaknesses S}}$, represents the ratio of the strengths to the weaknesses of a university.

P_{f10}: Academic situation, SS – a less well-prepared student will choose a university considered easier, while a student with a better academic standing will opt for a prestigious university.

P_{f11}: Educational progress test result, REZ – the test results can influence the decision-making process.

P_{f12}: Parents' profession, PP – for example, a student from an engineering family is more easily convinced to choose a faculty in the field of engineering if they have interacted with their parents' profession and enjoyed it.

P_{f13}: Labor market evolution, EM – in-demand (future) professions are much more attractive to young people than traditional ones.

P_{f14}: Economic evolution, EE – depending on the economic situation, certain professions are in demand.

P_{f15}: Family income level, VE – a family with modest income is less likely to choose a private university or one located far from home, just as a student from a low-income family will seek a field where they can find employment while in college.

P_{f16}: Student profile – scientific/humanities, PR – a student who has followed a scientific profile will be very difficult to convince to pursue a humanities faculty, just as a student who has followed a humanities profile will struggle to pass an admission exam for a scientific faculty.

P_{f17}: Distance coefficient, CoD – the distance from home can influence the choice between faculties with the same profile.

P_{f18}: Immediate employability, A – is a factor considered by students, given that there are fields with many job openings and others with fewer job opportunities or that are regulated.

P_{f19}: Salary, SA – is a factor considered by students.

P_{f20}: Group membership, AG – a student's belonging to a certain ethnic, social, etc., group can influence their choice of faculty.

P_{f21}: Scholarship level, NB – represents the value of the scholarships that a university can offer to prospective students.

The next step involved identifying the methods and mathematical models used in educational data analysis to understand how to correctly approach this specific field. This stage helped us select an initial set of questions to include in the working questionnaire.

Considering that a student makes a decision in the early years of high school regarding which faculty to attend, and that this decision requires time depending on various factors, we developed a mathematical model based on the selected parameters, allowing us to forecast each student's option. This relationship formed the core of our working application.

Based on the research conducted, we considered that the degree of conviction C(t) represents the level of certainty that a student will choose a field of study. The degree of conviction is 0 at the beginning of the study years (5th grade) and tends toward 1 on the day of the admission exam. The degree of conviction belongs to the interval (0,1]. It can be 1 at the moment when the student is decided and chooses a faculty. The value 0 is never reached, as at any moment a student has a desire to pursue a profession or have a career.

Thus, we defined the Degree of Conviction, C(t), as:

$$\{C(t) = 0, t < fifth\ grade \qquad (7.1)$$

$$C(t) = 1, t > first\ year\ of\ university$$

$$C(t) = f(\,II(t), FI(t), AT(t), VE(t), PP(t), REZ(t), SS(t),$$

$$EE(t), EM(t)p, PR(t), CoD(t), A(t), SA\,(t), AG(t), NB(t))$$

Developing a mathematical model for predicting high school graduates' career decisions represents an innovative and pragmatic approach in the complex context of education and academic choices. The direct consequence of this enhanced understanding is the ability of educational institutions to adapt their offerings to meet labor market demands, as well as the desires and aspirations of students.

The mathematical model for predicting high school graduates' career decisions is a significant initiative with the potential to improve the career decision-making process and higher education as a whole. By focusing on understanding and analyzing data, the model becomes a useful tool for shaping future professionals and optimizing their educational and professional pathways in an ever-changing world.

Chapter 8. Defining the machine learning algorithm

To determine the most suitable machine learning algorithm for the proposed solution, we conducted a detailed critical analysis of several machine learning algorithms. We analyzed 9 machine learning algorithms using the same dataset.

In Fig. 8.9, the source code for data modeling is presented. To achieve this, a series of steps must be followed, starting from splitting the dataset into two subsets for training and testing, through initializing the parameters and the Random Forest machine learning algorithm, and up to the prediction and evaluation of the model.

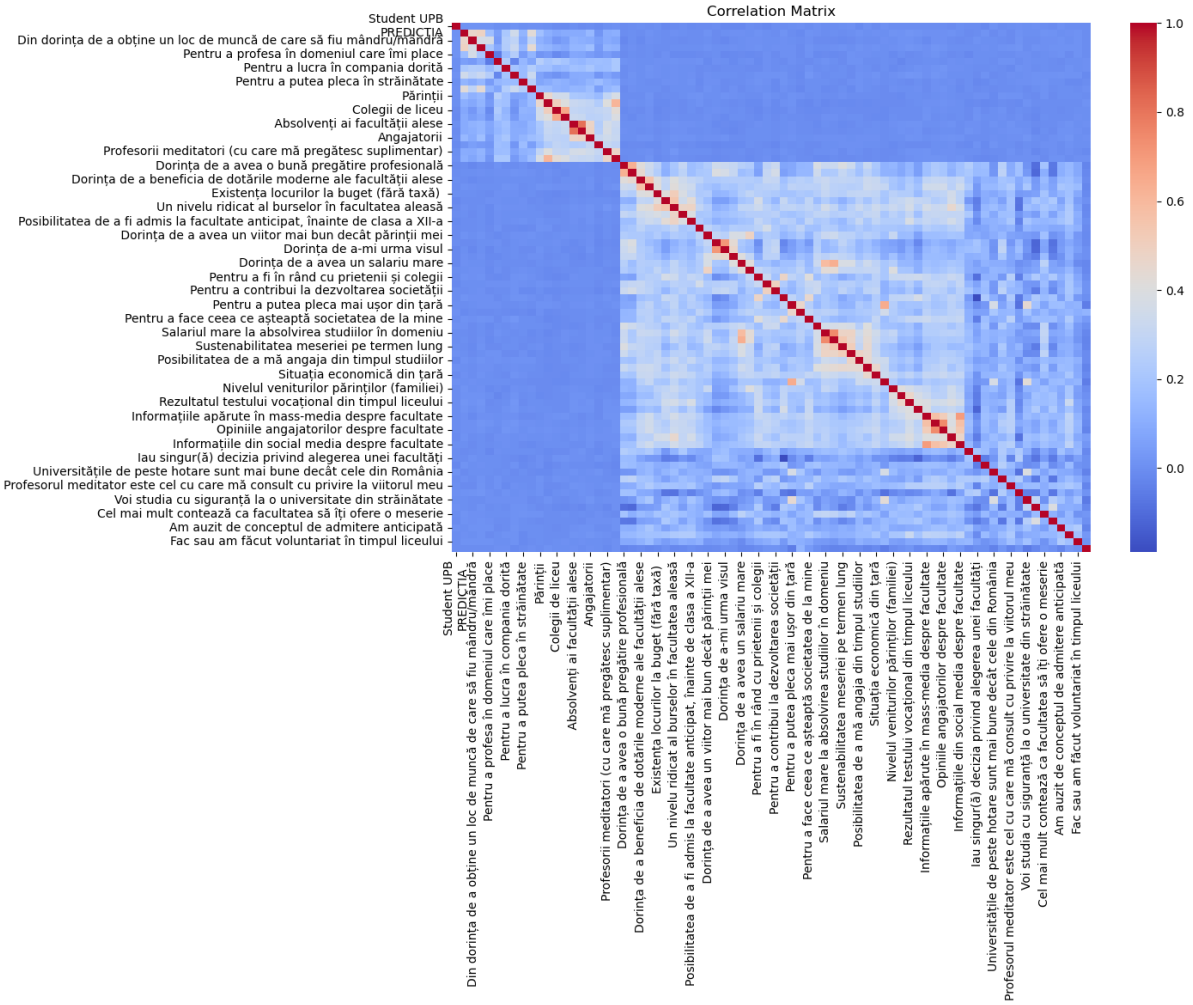

Fig. 8.9. *Extensive data analysis in the correlation matrix*

In Fig. 8.12, the graph indicating the weight values for each individual feature is represented.

From the analysis of the features in the performance matrix, two additional parameters emerged that were not included in the initial mathematical model: **Aspirational Desire (DA(t))**, for which 17 features were assigned, and **Desire to Emigrate (DE(t))**, for which 5 features were assigned.

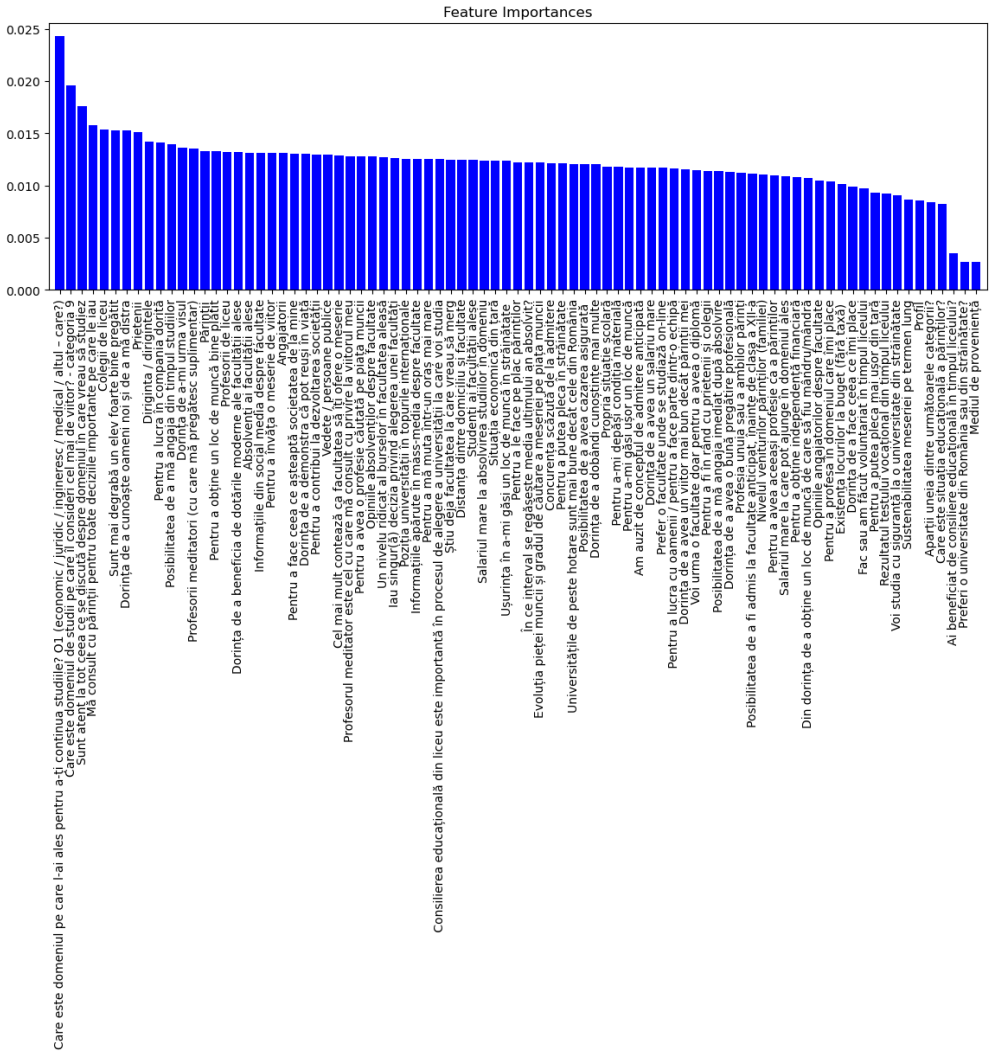

Fig. 8.16. *Classification of indicators by importance – Performance matrix*

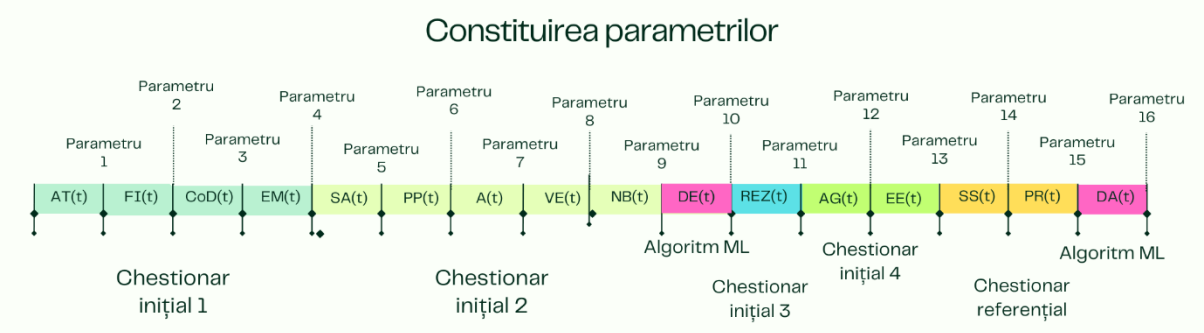

Fig. 8.17. *Finalization of the parameters*

Thus, as a result of analyzing the data from the questionnaire and inputting it into the machine learning algorithm, the mathematical model was revised, starting from the initially proposed version:

{Ct=0,t<first year of high school

Ct=∞, t>first year of university

Ct=FIt+ATt+VEt+PPt+NBt+REZt+SSt+EEt+EMt+PRt+CoDt+At+SAt+AGt+DAt+DEt,

t is between the first year of high school and the moment of admission,, first year of high school <t<admission

C(t) = F(FIt,ATt,VEt,PPt,NBt,REZt,SSt,EEt,EMt,PRt,CoDt,At,SAt,AGt,DAt, DEt)

for witch:

Tabel 8.2. *Concise description of the parameters*

PARAMETER NAME	CONCISE DESCRIPTION OF THE PARAMETER
II(t)	Influence of Information: positive and negative information appearing in all media sources.
Fi(t)	Influence Factor: the totality of elements and individuals that can change a person's evolution, thinking, and opinion regarding an institution.
AT(t)	Strengths and Weaknesses: the set of strengths and weaknesses of an institution.
VE(t)	Family Income Level.
PP(t)	Parents' Profession.
NB(t)	Scholarship Level: the level of scholarships that an institution offers to its students.
REZ(t)	Influence of Educational Orientation Tests: the impact that educational orientation tests have on the choice of an institution.
SS(t)	Student's Academic Situation.
EE(t)	Economic Evolution: the relationship of economic evolution to job professions in the labor market.
EM(t)	Labor Market Evolution: the dynamics of job professions in the labor market.
PR(t)	Student Profile: the inclination of a student towards a scientific or humanities profile.
CoD(t)	Distance Coefficient: the extent to which the distance from home can influence the choice of a university.
A(t)	Immediate Employability: the employability of students immediately after graduation.
SA(t)	Salary Level: the salary level that students have after completing their studies.
AG(t)	Group Membership: the student's belonging to a specific group.
DA(t)	Aspirational Desires: the aspirations and goals of the student.
DE(t)	Desire to Emigrate: the students' desire to emigrate after graduation.

In Table 8.3, the parameter names along with their acronyms are listed in columns, while the identified questions from the questionnaire are presented in rows. The weight of each parameter for each question has been identified in the columns. Finally, in the last row, the percentage weight for each parameter has been calculated from the Performance Matrix.

Tabel 8.3. *Performance Matrix*

Parameter / Question	Influence factors FI(t)	Aspirational desire DA(t)	Desire to emigrate DE(t)	University advantages AT(t)	Family income level VE(t)	Parents' profession PP(t)	Scholarship level – NB(t)	Educational progress test result REZ(t)	Student's academic situation SS(t)	Economic evolution EE(t)	Labor market evolution EM(t)	Student profile – scientific /human – PR(t)	Distance coefficient - CoD (t)	Immediate employability - A (t)	Salary - SA(t)	Group membership - AG(t)	Weight
Do you prefer a university in Romania or abroad?	0	0	0,0083	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0083
What field have you chosen to continue your studies in? (economic/legal/engineering/medical/other – which?)	0	0	0	0	0	0	0	0	0	0	0	0,0027	0	0	0	0	0,0027
What field do you consider to be the future?	0	0	0	0	0	0	0	0	0	0	0,0243	0	0	0	0	0	0,0243
To obtain a well-paid job.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0154	0	0,0154
Out of a desire to obtain a job I can be proud of.	0	0,0176	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0176
To have a profession that is in demand in the labor market.	0	0	0	0	0	0	0	0	0	0	0,0121	0	0	0	0	0	0,0121
To work in the field I enjoy.	0	0,0115	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0115
To have the same profession as my parents.	0	0	0	0	0	0,0126	0	0	0	0	0	0	0	0	0	0	0,0126
To work in the desired company.	0	0,0108	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0108

Parameter / Question	Influence factors FI(t)	Aspirational desire DA(t)	Desire to emigrate DE(t)	University advantages AT(t)	Family income level VE(t)	Parents' profession PP(t)	Scholarship level – NB(t)	Educational progress test result REZ(t)	Student's academic situation SS(t)	Economic evolution EE(t)	Labor market evolution EM(t)	Student profile – scientific /human – PR(t)	Distance coefficient - CoD (t)	Immediate employability - A (t)	Salary - SA(t)	Group membership - AG(t)	Weight
To achieve financial independence.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0121	0	0,0121
To be able to go abroad.	0	0	0,0131	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0131
To easily find a job.	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0122	0	0	0,0122
Parents.	0,0132	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0132
High school teachers.	0,0151	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0151
High school classmates.	0,0107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0107
Friends.	0,0109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0109
Graduates of the chosen faculty.	0	0	0	0,0124	0	0	0	0	0	0	0	0	0	0	0	0	0,0124
Students of the chosen faculty.	0	0	0	0,0118	0	0	0	0	0	0	0	0	0	0	0	0	0,0118
Employers.	0,0135	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0135
Celebrities/public figures.	0,0132	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0132
Tutors (with whom I prepare additionally).	0,0133	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0133
The class teacher.	0,0141	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0141
The desire to have a good professional training.	0	0,0104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0104
The desire to gain more knowledge.	0	0,0087	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0087
The desire to benefit from the modern facilities of	0	0	0	0,0132	0	0	0	0	0	0	0	0	0	0	0	0	0,0132

Parameter / Question	Influence factors FI(t)	Aspirational desire DA(t)	Desire to emigrate DE(t)	University advantages AT(t)	Family income level VE(t)	Parents' profession PP(t)	Scholarship level – NB(t)	Educational progress test result REZ(t)	Student's academic situation SS(t)	Economic evolution EE(t)	Labor market evolution EM(t)	Student profile – scientific /human – PR(t)	Distance coefficient - CoD (t)	Immediate employability - A (t)	Salary - SA(t)	Group membership - AG(t)	Weight
the chosen faculty.																	
The university's position in international rankings.	0	0	0	0,0152	0	0	0	0	0	0	0	0	0	0	0	0	0,0152
The existence of budget places (without tuition).	0	0	0	0,0091	0	0	0	0	0	0	0	0	0	0	0	0	0,0091
Low competition in admission.	0	0	0	0,0131	0	0	0	0	0	0	0	0	0	0	0	0	0,0131
A high level of scholarships in the chosen faculty.	0	0	0	0	0	0	0,0118	0	0	0	0	0	0	0	0	0	0,0118
My own academic situation.	0	0	0	0	0	0	0	0	0,0114	0	0	0	0	0	0	0	0,0114
The possibility of being admitted early, before the 12th grade.	0	0	0	0,0132	0	0	0	0	0	0	0	0	0	0	0	0	0,0132
The desire to meet new people and have fun.	0	0,0104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0104
The desire to have a better future than my parents.	0	0	0	0	0	0,0092	0	0	0	0	0	0	0	0	0	0	0,0092
The desire to do what I love.	0	0,0115	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0115
The desire to follow my dream.	0	0,0117	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0117

Parameter / Question	Influence factors FI(t)	Aspirational desire DA(t)	Desire to emigrate DE(t)	University advantages AT(t)	Family income level VE(t)	Parents' profession PP(t)	Scholarship level – NB(t)	Educational progress test result REZ(t)	Student's academic situation SS(t)	Economic evolution EE(t)	Labor market evolution EM(t)	Student profile – scientific /human – PR(t)	Distance coefficient - CoD (t)	Immediate employability - A (t)	Salary - SA(t)	Group membership - AG(t)	Weight
The desire to prove that I can succeed in life.	0	0,0113	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0113
The desire to have a high salary.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0123	0	0,0123
To overcome my financial situation.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0139	0	0,0139
To keep up with friends and classmates.	0	0,0126	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0126
To work with people/to be part of a team.	0	0,0124	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0124
To contribute to the development of society.	0	0,012	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,012
To please my parents.	0	0,0122	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0122
To leave the country more easily.	0	0	0,0117	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0117
To move to a bigger city.	0	0	0	0	0	0	0	0	0	0	0,0099	0	0	0	0	0	0,0099
To do what society expects of me.	0	0,0142	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0142
To learn a future profession.	0	0	0	0	0	0	0	0	0	0	0,013	0	0	0	0	0	0,013
The high salary upon graduation in the field.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0114	0	0,0114
The high salary I can reach in the chosen field.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0158	0	0,0158

Parameter / Question	Influence factors FI(t)	Aspirational desire DA(t)	Desire to emigrate DE(t)	University advantages AT(t)	Family income level VE(t)	Parents' profession PP(t)	Scholarship level – NB(t)	Educational progress test result REZ(t)	Student's academic situation SS(t)	Economic evolution EE(t)	Labor market evolution EM(t)	Student profile – scientific /human – PR(t)	Distance coefficient - CoD (t)	Immediate employability - A (t)	Salary - SA(t)	Group membership - AG(t)	Weight
The sustainability of the profession in the long term.	0	0	0	0	0	0	0	0	0	0	0,0097	0	0	0	0	0	0,0097
The possibility of getting a job immediately after graduation.	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0128	0	0	0,0128
The possibility of getting a job during my studies.	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0121	0	0	0,0121
The evolution of the labor market and the demand for the profession.	0	0	0	0	0	0	0	0	0	0	0,0125	0	0	0	0	0	0,0125
The economic situation in the country.	0	0	0	0	0	0	0	0	0	0,0116	0	0	0	0	0	0	0,0116
The ease of finding a job abroad.	0	0	0	0	0	0	0	0	0	0	0	0	0	0,013	0	0	0,013
The income level of my parents (family).	0	0	0	0	0,0152	0	0	0	0	0	0	0	0	0	0	0	0,0152
The distance between home and university.	0	0	0	0	0	0	0	0	0	0	0	0	0,0117	0	0	0	0,0117
The result of the vocational test during high school.	0	0	0	0	0	0	0	0,0117	0	0	0	0	0	0	0	0	0,0117
The profession of one or both parents.	0	0	0	0	0	0,012	0	0	0	0	0	0	0	0	0	0	0,012

Parameter / Question	Influence factors FI(t)	Aspirational desire DA(t)	Desire to emigrate DE(t)	University advantages AT(t)	Family income level VE(t)	Parents' profession PP(t)	Scholarship level – NB(t)	Educational progress test result REZ(t)	Student's academic situation SS(t)	Economic evolution EE(t)	Labor market evolution EM(t)	Student profile – scientific /human – PR(t)	Distance coefficient - CoD (t)	Immediate employability - A (t)	Salary - SA(t)	Group membership - AG(t)	Weight
Information appearing in the media about the faculty.	0,0116	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0116
Opinions of graduates about the faculty.	0,0127	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0127
Opinions of employers about the faculty.	0,0128	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0128
The possibility of having accommodation provided.	0	0	0	0,0109	0	0	0	0	0	0	0	0	0	0	0	0	0,0109
Information from social media about the faculty.	0,0131	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0131
I am rather a very well-prepared student.	0	0	0	0	0	0	0	0	0,0128	0	0	0	0	0	0	0	0,0128
I make the decision regarding my choice of faculty on my own.	0,0111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0111
I consult with my parents for all important decisions I make.	0,0133	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0133
Universities abroad are better than those in Romania.	0	0	0,0124	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0124
I pay attention to what is discussed about the field I study in.	0,0196	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0196

Parameter / Question	Influence factors FI(t)	Aspirational desire DA(t)	Desire to emigrate DE(t)	University advantages AT(t)	Family income level VE(t)	Parents' profession PP(t)	Scholarship level – NB(t)	Educational progress test result REZ(t)	Student's academic situation SS(t)	Economic evolution EE(t)	Labor market evolution EM(t)	Student profile – scientific /human – PR(t)	Distance coefficient - CoD (t)	Immediate employability - A (t)	Salary - SA(t)	Group membership - AG(t)	Weight
The tutor is the one I consult regarding my future.	0,0101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0101
I already know which faculty I want to attend.	0	0,0131	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0131
I will definitely study at a university abroad.	0	0	0,0125	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0125
I will attend a faculty just to have a diploma.	0	0,0113	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0113
What matters most is that the faculty offers you a profession.	0	0	0	0	0	0	0	0	0	0	0,0126	0	0	0	0	0	0,0126
I prefer a faculty where online study is available.	0	0	0	0,011	0	0	0	0	0	0	0	0	0	0	0	0	0,011
I have heard of the concept of early admission.	0	0	0	0,0123	0	0	0	0	0	0	0	0	0	0	0	0	0,0123
Educational counseling in high school is important in the process of choosing the university I will study at.	0	0	0	0	0	0	0	0,0136	0	0	0	0	0	0	0	0	0,0136
I do or have done volunteering during high school.	0	0,0129	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0129
Profile.	0	0	0	0	0	0	0	0	0	0	0	0,0086	0	0	0	0	0,0086

Parameter / Question	Influence factors FI(t)	Aspirational desire DA(t)	Desire to emigrate DE(t)	University advantages AT(t)	Family income level VE(t)	Parents' profession PP(t)	Scholarship level – NB(t)	Educational progress test result REZ(t)	Student's academic situation SS(t)	Economic evolution EE(t)	Labor market evolution EM(t)	Student profile – scientific /human – PR(t)	Distance coefficient - CoD (t)	Immediate employability - A (t)	Salary - SA(t)	Group membership - AG(t)	Weight
In what range does the average of the last year graduated fall?	0	0	0	0	0	0	0	0	0,0093	0	0	0	0	0	0	0	0,0093
Do you belong to one of the following categories?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0122	0,0122
What is the educational situation of your parents?	0	0	0	0	0	0,0027	0	0	0	0	0	0	0	0	0	0	0,0027
Have you benefited from educational counseling during high school?	0	0	0	0	0	0	0	0,0084	0	0	0	0	0	0	0	0	0,0084
Background.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0035	0,0035
Weight value.	**0,2083**	**0,2046**	**0,058**	**0,1222**	**0,0152**	**0,0365**	**0,0118**	**0,0337**	**0,0335**	**0,0116**	**0,0941**	**0,0113**	**0,0117**	**0,0501**	**0,0809**	**0,0157**	**0,9992**
	20,83%	20,46%	5,80%	12,22%	1,52%	3,65%	1,18%	3,37%	3,35%	1,16%	9,41%	1,13%	1,17%	5,01%	8,09%	1,57%	

The weight for Influence Factors was determined by summing the entries in the Performance Matrix (Table 8.4) for each characteristic assigned to it, as follows:

Tabel 8.4. *Weights for Influence factors*

Parents	High school teachers	High school classmates	Friends	Employers	Celebrities/ public figures	Tutors (with whom I prepare additionally)	The class teacher	Information appearing in the media about the faculty	Opinions of graduates about the faculty	Opinions of employers about the faculty	Information from social media about the faculty	I make the decision regarding my choice of faculty on my own	I consult with my parents for all important decisions I make	I pay attention to what is discussed about the field I study in	The tutor is the one I consult regarding my future	**Weight value**
0,0132	0,0151	0,0107	0,0109	0,0135	0,0132	0,0133	0,0141	0,0116	0,0127	0,0128	0,0131	0,0111	0,0133	0,0196	0,0101	**0,2083**

Similarly, the weights for the other factors were calculated (Tables 8.5 – 8.19).

Tabel 8.5. *Weights for Aspirational desire*

Out of a desire to obtain a job I can be proud of	To work in the field I enjoy	To work in the desired company	The desire to have a good professional training	The desire to gain more knowledge	The desire to meet new people and have fun	The desire to do what I love	The desire to follow my dream	The desire to prove that I can succeed in life	To keep up with friends and classmates	To work with people/to be part of a team	To contribute to the development of society	To please my parents	To do what society expects of me	I already know which faculty I want to attend	I will attend a faculty just to have a diploma	I do or have done volunteering during high school	**Weight value**
0,0176	0,0115	0,0108	0,0104	0,0087	0,0104	0,0115	0,0117	0,0113	0,0126	0,0124	0,012	0,0122	0,0142	0,0131	0,0113	0,0129	**0,2046**

Tabel 8.6. *Weights for Desire to emigrate*

Do you prefer a university in Romania or abroad?	To be able to go abroad	To leave the country more easily	Universities abroad are better than those in Romania	I will definitely study at a university abroad	**Weight value**
0,0083	0,0131	0,0117	0,0124	0,0125	**0,058**

Tabel 8.7. *Weights for University advantages*

Graduates of the chosen faculty	Students of the chosen faculty	The desire to benefit from the modern facilities of the chosen faculty	The university's position in international rankings	The existence of budget places (without tuition)	Low competition in admission	The possibility of being admitted early, before the 12th grade	The possibility of having accommodation provided	I prefer a faculty where online study is available	I have heard of the concept of early admission	**Weight value**
0,0124	0,0118	0,0132	0,0152	0,0091	0,0131	0,0132	0,0109	0,011	0,0123	**0,1222**

Tabel 8.8. *Weights for Family income level*

The income level of the parents (family)	**Weight value**
0,0152	**0,0152**

Tabel 8.9. *Weights for Parents' profession*

To have the same profession as my parents	The desire to have a better future than my parents	The profession of one or both parents	What is the educational situation of your parents?	**Weight value**
0,0126	0,0092	0,012	0,0027	**0,0365**

Tabel 8.10. *Weights for Scholarship level*

A high level of scholarships in the chosen faculty	**Weight value**
0,0118	**0,0118**

Tabel 8.11. *Weights for Educational progress test result*

The result of the vocational test during high school	Educational counseling in high school is important in the process of choosing the university I will study at	Have you benefited from educational counseling during high school?	**Weight value**
0,0117	0,0136	0,0084	**0,0337**

Tabel 8.12. *Weights for Academic situation*

My own academic situation	I am rather a very well-prepared student	In what range does the average of the last year graduated fall?	**Weight value**
0,0114	0,0128	0,0093	**0,0335**

Tabel 8.13. *Weights for Economic evolution*

The economic situation in the country	**Weight value**
0,0116	**0,0116**

Tabel 8.14. *Weights for the Evolution of the labor market*

Which field do you consider to have a future?	To have a sought-after profession in the labor market	To move to a bigger city	To learn a future-oriented trade	The long-term sustainability of the profession	The evolution of the labor market and the demand for the profession in the labor market	The most important thing is for the university to provide you with a profession	**Weight value**
0,0243	0,0121	0,0099	0,013	0,0097	0,0125	0,0126	**0,0941**

Tabel 8.15. *Weights for the Student profile – science/humanities*

Which field did you choose to continue your studies in? (economic/legal/engineering/medical/other – which one?)	Profil	**Weight value**
0,0027	0,0086	**0,0113**

Tabel 8.16. *Weights for distance coefficient*

The distance between home and university	**Weight value**
0,0117	**0,0117**

Tabel 8.17. *Weights for Immediate employability*

To easily find a job	The possibility of getting a job immediately after graduation	The possibility of getting a job during studies	The ease of finding a job abroad	**Weight value**
0,0122	0,0128	0,0121	0,013	**0,0501**

***Tabel 8.18.** Weights for Salary*

To obtain a well-paid job	To achieve financial independence	The desire to have a high salary	To overcome my financial situation	High salary upon graduation in the field	High salary I can achieve in the chosen field	**Weight value**
0,0154	0,0121	0,0123	0,0139	0,0114	0,0158	**0,0809**

***Tabel 8.19.** Weights for Belonging to a group*

Do you belong to one of the following categories?	Background environment	**Weight value**
0,0122	0,0035	**0,0157**

Thus, starting from the mathematical model determined in Chapter 6, namely:

$$\{C(t) = 0, t < \text{First grade}$$

$$C(t) = 1, t > \text{First year of university}$$

$$C(t) = F(FI(t), AT(t), VE(t), PP(t), EV(t), REZ(t), SS(t), EE(t), EM(t)p, PR(t), CoD(t), A(t), SA\ (t), AG(t), NB(t))$$

With the amendments made in Chapter 7:

$$Ct=FIt+ATt+VEt+PPt+NBt+REZt+SSt+EEt+EMt+PRt+CoDt+At+SAt+AGt+DAt+DEt,$$

t is between first grade and the time of admission, first grade < t < admission

The mathematical model becomes:

$$C(t)=0,2083*FI(t)+0,2046*DA(t)+0,058*DE(t)+0,1222*AT(t)+0,0152*VE(t)+0,0365*PP(t)+0,0118*NB(t)+0,0337*REZ(t)+0,0335*SS(t)+0,0116*EE(t)+0,0941*EM(t)+0,0113*PR(t)+0,0117*CoD(t)+0,0501*A(t)+0,0809*SA(t)+0,0157*AG(t),$$

t is between first grade and the time of admission, first grade < t < admission

Thus, we can state that the prediction model is a function of 16 parameters with different weights, which is represented by the model:

$$C(t)= f(0,2083*FI(t), 0,2046*DA(t), 0,1222*AT(t), 0,0941*EM(t), 0809*SA(t), 0,058*DE(t), 0,0501*A(t), 0,0365*PP(t), 0,0337*REZ(t), 0,0335*SS(t), 0,0157*AG(t), 0,0152*VE(t), 0,0118,NB(t), 0,0117*CoD(t), 0,0116*EE(t), 0,0113*PR(t))$$

***Tabel 8.20.** The values of the weights for each parameter*

Parameter	Weight
Influence factors	0,2083
Aspirational desire	0,2046
University strengths	0,1222
Labor market evolution	0,0941
Salary	0,0809
Desire to emigrate	0,058
Immediate employability	0,0501
Parents' profession	0,0365

Educational path test result	0,0337
Academic performance	0,0335
Group belonging	0,0157
Family income level	0,0152
Scholarship level	0,0118
Distance coefficient	0,0117
Economic evolution	0,0116
Student profile – science/humanities	0,0113

For the validation of the model, the database resulting from the application of the Referential Questionnaire was used, in which we identified students who, between the time of completing the questionnaire and the present moment, became students of the National University of Science and Technology POLITEHNICA Bucharest. Thus, the database generated from the completion of the Referential Questionnaire was compared with the student database of POLITEHNICA Bucharest. From this comparison, 3,787 common entries resulted, generating a new database with 3,787 responses from students who, in the meantime, became university students. The prediction model was applied to this database, yielding 3,654 results with a conviction rate greater than 75%.

	BL	BM	BN	BO	BP	BQ	BR	BS	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC	CD	CG	CH
1	Posibilitatea de a avea cazarea asigură	Informațiile din social media despre faculta	Sunt mai degrabă un elev foarte bine pregăti	Iau singur(ă) decizia privind alegerea unei	Mă consult cu părinții pentru toate deciziil	Universitățile de peste hotare sunt mai bune	Sunt atent la tot ceea ce se discută despre	Profesorul meditator este cel mă	Știu deja facultatea la care vreau să merg	Voi studia cu siguranță la o universita te din	Voi urma o facultate doar pentru a avea o	Cel mai mult contează ca facultatea să îți	Prefer o facultate unde se studiază on-line	Am auzit de conceptul de admitere anticip	Consilierea educațion ală din liceu e	Fac sau am făcut voluntariat în timpul liceului	Profil	În ce interval se regăsește media ultimului an absolvit?	Mediul de proveniență	Ai beneficiat de consiliere educațion ală în	Gradul de convingere
45	0	0	7	10	0	1	9	10	10	0	0	8	0	2	0	0	Teoretic - Real	8,50-8,99	Urban	Nu	8,94
80	10	10	9	10	10	10	10	10	10	10	10	10	10	10	10	10	Tehnologic	8,50-8,99	Rural	Nu	9,70
94	0	0	3	10	5	0	10	0	10	0	5	10	5	4	0	0					8,32
100	0	5	5	3	9	0	9	5	9	0	4	8	3	6	4	0	Teoretic - Real	8,50-8,99	Urban	Nu	9,84
102	7	2	8	9	7	9	10	4	10	0	0	10	0	9	8	0	Teoretic - Real	9,00-9,49	Rural	Nu	7,88
110																	Teoretic - Real	7,50-7,99	Rural	Da	8,23
115	0	5	9	10	8	3	7	0	10	0	0	0	0	10	6	0	Tehnologic	9,50-10	Urban	Nu	9,11
151	7	7	6	10	3	8	8	0	1	5	0	4	0	8	0	0	Teoretic - Real	9,00-9,49	Urban	Nu	9,87
153	2	2	7	9	7	2	9		8	1	2	9	1	1	3	10	Teoretic - Uman	9,50-10	Urban	Nu	9,18
185	0	7	7	9	4	7	8	0	10	0	0	7	0	3	0	0	Teoretic - Real				9,33
191	3	7	8	9	9	1	10	0	10	0	0	2	1	3	4	0	Teoretic - Uman	9,50-10	Rural	Da	9,12
199	0	10	8	9	10	10	10	0	9	3	0	10	0	6	0	0	Teoretic - Uman	8,50-8,99	Urban	Nu	9,44
209	7	6	10	7	6	9	8	2	0	4	5	8	0	3	10	0	Teoretic - Real				8,98
218																	Teoretic - Uman	7,00-7,49	Urban	Da	8,77
225																	Teoretic - Uman	8,50-8,99	Rural	Nu	9,44
246	0	6	4	10	3	1	0	0	0	0	6	5		4	5	3	Teoretic - Real				9,34
257	6	0	8	8	10	10	10	0	9	8	5	10	4		10	0	Teoretic - Uman	9,50-10	Urban	Da	8.22
266	0	0	5	10	7	10	10	0	10	0	10	10	0	0	2	0	Teoretic - Uman	8,00-8,49	Urban	Nu	8,44
282		10	5	10	4	10	4	10	10	6	1	8	4	10	4	10	Teoretic - Uman	8,50-8,99	Urban	Nu	7,98

Fig. 8.18. *Database of 3,787 POLITEHNICA Bucharest students who completed the referential questionnaire*

Chapter 9. Development of a web application

In order to determine the degree of conviction of a student in choosing the university they wish to attend, we developed an online application.

Source: Web application

Fig. 9.1. *The web application interface – "Home menu"*

In Fig. 9.1, the web application interface is displayed. On the first page of the application, the three main buttons can be found, namely: "Home," "Mathematical Model," and "Form." The application is accompanied by a motto, which reads: "The Technical Future in Your Hands!" emphasizing the purpose of the application, which is to determine how convinced a student is about choosing a technical university.

Source: Web application

Fig. 9.3. *The actual form in the web application*

The main components of the logic diagram are presented below (Fig. 9.4):

1. **Client**: the entity that makes requests to the API;
2. **API**: the access point that mediates client requests to the database;
    - **HTTP Methods**: specifies the methods used, in this case, the GET method.

3. **Utility Function (connect)**: a utility function for connecting to the database;
4. **"Form" Database**: the database being connected to, containing the collections;
    - **"Questions" Collection**: one of the collections in the database that is being accessed;
    - **"Answers" Collection**: another collection mentioned in the database, but without further details on interactions.
5. **Data Processing**:
    - **Transformation into an Array**: The extracted data is transformed into an array structure.
6. **JSON Response**: The information is sent back to the client in the form of a JSON response.

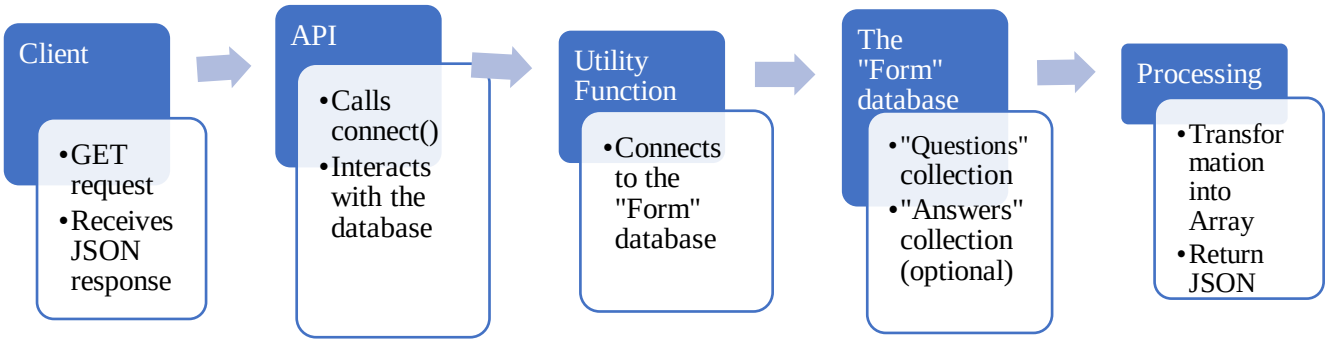

Fig. 9.4. *The components of the logic diagram*

In order to group the questions by the parameter to which they belong, it was necessary to centralize them. Table 9.1 shows the centralization of the parameters based on the questions they are associated with.

Tabel 9.1*. The distribution of questions by parameters*

No.	Parameter name	Question
1.	Influence factors	Parents
		High school teachers
		High school classmates
		Friends
		Employers
		Celebrities/public figures
		Tutors (with whom I prepare additionally)
		Homeroom teacher
		Information in the media about the university
		Opinions of graduates about the university
		Employers' opinions about the university
		Social media information about the university

No.	Parameter name	Question
		I make the decision about university choice on my own
		I consult with my parents for all important decisions I make
		I pay attention to what is being discussed about the field I study
		My tutor is the one I consult about my future
2.	Aspirational desire	Out of the desire to get a job I can be proud of
		To work in the field I enjoy
		To work for the company I desire
		The desire to have a solid professional education
		The desire to gain more knowledge
		The desire to meet new people and have fun
		The desire to do what I love
		The desire to follow my dream
		The desire to prove that I can succeed in life
		To keep up with my friends and classmates
		To work with people/to be part of a team
		To contribute to the development of society
		To please my parents
		To do what society expects of me
		I already know which university I want to attend
		I will attend university just to get a degree
		I do or have done volunteer work during high school
3.	Desire to emigrate	Do you prefer a university in Romania or abroad?
		To be able to go abroad
		To leave the country more easily
		Universities abroad are better than those in Romania
		I will definitely study at a university abroad
4.	University strengths	Graduates of the chosen university
		Students of the chosen university
		The desire to benefit from the modern facilities of the chosen university
		The university's position in international rankings
		The availability of budgeted (tuition-free) places
		Low competition for admission
		The possibility of being admitted early, before 12th grade
		The possibility of having guaranteed accommodation
		I prefer a university where studies are conducted online
		I have heard of the early admission concept
5.	Family income level	Parents' (Family's) Income Level
6.	Parents' profession	To have the same profession as my parents
		The desire to have a better future than my parents
		The profession of one or both parents
		What is your parents' educational background?
7.	Scholarship level	A high level of scholarships at the chosen university
8.	Educational path test result	The result of the vocational test during high school
		Educational counseling in high school is important in the process of choosing the university where I will study
		Did you benefit from educational counseling during high school?
9.		My own academic situation

No.	Parameter name	Question
	Academic performance	I am more of a well-prepared student
		In which range does the average grade of the last completed year fall?
10.	Economic evolution	The economic situation in the country
11.	Labor market evolution	Which field do you consider to have a future?
		To have a profession in demand on the labor market
		To move to a bigger city
		To learn a future-oriented trade
		The long-term sustainability of the profession
		The evolution of the labor market and the demand for the profession in the labor market
		The most important thing is for the university to provide you with a profession
12.	Student profile – science/ humanities	Which field did you choose to continue your studies in? (economic/legal/engineering/medical/other – which one?)
		Profile
13.	Distance coefficient	The distance between home and university
14.	Immediate employability	To easily find a job
		The possibility of getting a job immediately after graduation
		The possibility of getting a job during studies
		The ease of finding a job abroad
15.	Salary	To obtain a well-paid job
		To achieve financial independence
		The desire to have a high salary
		To overcome my financial situation
		High salary upon graduation in the field
		High salary I can achieve in the chosen field
16.	Group belonging	Do you belong to one of the following categories?
		Background environment

For each parameter, the weighted average of the values given for each question was calculated. The result was then multiplied by the weight value specific to the respective parameter. The model used for the calculations in the application is presented as follows:

Ct=20,83%*FI(t)+20,46%*DA(t)+5,8%*DE(t)+12,22%*AT(t)+1,52%*VE(t)+3,65%*PP(t)+
1,18%*NB(t)+3,37%*REZ(t)+3,35%*SS(t)+1,16%*EE(t)+9,41%*EM(t)+1,13%*PR(t)+
1,17%*CoD(t)+5,01%*A(t)+8,09%*SA(t)+1,57%*AG(t),

The weight values for these parameters are presented in Table 9.2:

Tabel 9.2. *The weight values of the parameters for the web application*

No.	Parameter name	Parameter acronym	Weight
1.	Influence factors	FI(t)	20,83%
2.	Aspirational desire	DA(t)	20,46%
3.	Desire to emigrate	DE(t)	5,8%
4.	University strengths	AT(t)	12,22%

No.	Parameter name	Parameter acronym	Weight
5.	Family income level	VE(t)	1,52%
6.	Parents' profession	PP(t)	3,65%
7.	Scholarship level	NB(t)	1,18%
8.	Educational path test result	REZ(t)	3,37%
9.	Academic performance	SS(t)	3,35%
10.	Economic evolution	EE(t)	1,16%
11.	Labor market evolution	EM(t)	9,41%
12.	Student profile – science/humanities	PR(t)	1,13%
13.	Distance coefficient	CoD(t)	1,17%
14.	Immediate employability	A(t)	5,01%
15.	Salary	SA(t)	8,09%
16.	Group belonging	AG(t)	1,57%
	Total		**99,92%**

The Career Guidance Guide presents information that helps students better understand the factors involved in the decision to choose a university.

The purpose of this guide is to provide a structured framework to help students clarify their information and make a well-informed decision that reflects both their personal desires and interests, as well as current economic and social realities. Each chapter addresses an essential aspect of the selection process, guiding them step by step to evaluate options and make a well-founded decision.

Chapter 1: Influence factors - This chapter presents the various influences on students' decisions: parents, high school teachers, classmates, friends, employers, celebrities, media, and social media information. The importance of consulting any of these factors and how students can use this information to make an informed decision is highlighted.

Chapter 2: Aspirational desire - The motivations for choosing a university are explored, such as the desire to get a respected job, work in a field they enjoy, gain knowledge, and contribute to societal development. Social influences, such as peer pressure and the desire to please parents, are also analyzed.

Chapter 3: Desire to emigrate - This chapter examines students' desire to study abroad and the perception that universities outside the country offer better education. The reasons some students might prefer studying at Romanian universities versus those abroad are also discussed.

Chapter 4: University strengths - This chapter focuses on elements that make a university attractive to students, such as modern facilities, position in international rankings, availability of tuition-free places, guaranteed accommodation, and the possibility of early admission.

Chapter 5: Family income level - It examines how family income influences students' decisions to attend a particular university, analyzing the impact of the costs associated with studies and the financial support provided by the family.

Chapter 6: Parents' profession - The influence of parents' professions on students' aspirations and educational choices is explored, including the desire to follow the same career as their parents or pursue a different profession.

Chapter 7: Scholarship level - This chapter presents the importance of scholarships and financial incentives offered by universities in attracting students and facilitating access to higher education.

Chapter 8: Educational path test result - The role of educational tests in guiding students toward the right university and the impact of these tests on educational choices are examined.

Chapter 9: Academic performance - This chapter analyzes the influence of academic performance on university choice, such as general grades and participation in competitions or olympiads.

Chapter 10: Economic evolution - The national economic situation and how it may affect students' decisions, including perceptions of long-term stability and opportunities in various fields of study, are examined.

Chapter 11: Labor market evolution - The trends in the labor market and how these influence university choice, the desire to have a future-oriented career, and immediate employment opportunities after graduation are addressed.

Chapter 12: Student profile - This chapter analyzes how preferences for science or humanities profiles influence study options and university selection.

Chapter 13: Distance from home - The importance of the distance between home and university, including preferences for staying close to home versus studying in a distant city, is explored.

Chapter 14: Immediate employability - The importance of immediate employment prospects after graduation and the possibility of working during studies as decisive factors in university choice are analyzed.

Chapter 15: Salary - This chapter explores students' salary expectations and the desire to obtain a well-paid job that offers financial independence and a better standard of living.

Chapter 16: Group belonging - The influence of belonging to different social or economic categories (e.g., students with parents working abroad, students from low-income families) on educational decisions is analyzed.

The online application was also used to validate the prediction model, and after recording over 400 responses from students at the National University of Science and Technology POLITEHNICA Bucharest, the model fell within validated limits, with a confidence level of over 95% and a career decision prediction accuracy of over 0.75.

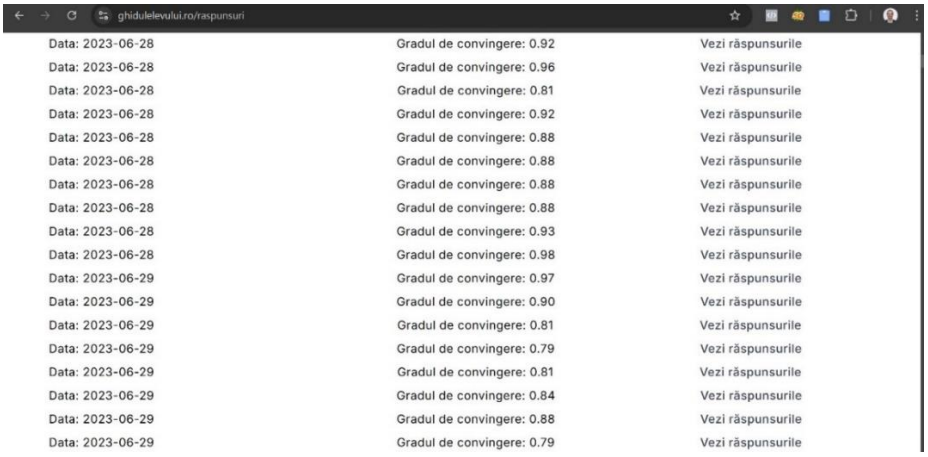

Fig. 9.11. *The students' responses in the web application*

Data: 9/5/2024 Gradul de convingere: 0.84

Lista răspunsurilor:

Factori de influență FI(t)	
Părinții:	8
Profesorii de liceu:	7
Colegii de liceu:	8
Prietenii:	7
Angajatorii:	4
Vedete / persoane publice:	9
Profesorii meditatori (cu care mă pregătesc suplimentar):	9
Diriginta / dirigintele:	9
Informații apărute în mass-media despre facultate:	8
Opiniile absolvenților despre facultate:	9
Opiniile angajatorilor despre facultate:	9

Source: Web application

Fig. 9.12. *The degree of conviction for a respondent in the web application*

The application is based on a form that must be completed by students (presented in Appendix 3), which includes the questions from the questionnaire developed in the previous chapters. The questions refer to the external influences that young people experience, their career opportunities, academic interests, etc. After completing the form, the Degree of Conviction, C(t), is determined. This value ranges between 0 and 1, indicating whether a student is less decided (values tending towards 0) or fully decided (value of 1) in choosing the university they wish to attend.

Regarding the simulations conducted within the application, by assigning extreme values to the questionnaire (form) responses, the correct functionality of the application was demonstrated. The resulting scores were accurately calculated, both in cases where maximum values were assigned and in cases where minimum values were assigned. Additionally, the displayed messages were appropriately assigned. From this analysis, it is concluded that the application can accurately assess the degree of conviction of students in choosing a university.

Chapter 10. Conclusions and proposals

The important factors in choosing a university include orientation towards a preferred profession, achieving financial independence, the desire to have a well-paid job, and obtaining employment in the desired field.

The Random Forest regression algorithm was chosen due to its simple architecture and ability to handle both regression and classification problems. Modifications made to the initial model included removing and adding certain factors to determine the degree of conviction of students in choosing a university. The Random Forest algorithm calculated the importance of each parameter, resulting in a ranking of influence factors, aspirational desires, university strengths, labor market evolution, and other relevant elements for high school graduates' career decisions.

The developed application offers a useful tool for both students and educational counselors, providing a structured and mathematical way to assess the degree of conviction students have in choosing a university.

As for the original contributions, the following were achieved:

- A comprehensive research on the specialized literature for social research conducted through opinion/statistical surveys;
- A study on the use of statistical surveys in social research, identifying the statistical methods used in social research and the data preprocessing techniques applied in the analysis of data from opinion surveys;
- A synthesis of information about what sampling represents, considering the sampling methods used, along with a classification of these methods. Sampling errors were described to determine the reliability of a survey;
- Descriptions of the concepts "Artificial Intelligence," "Machine Learning," and "Deep Learning";
- Analysis of the university admission process, developing a mathematical model for predicting high school graduates' career decisions, capable of anticipating or predicting the career path these graduates will follow, specifically: identifying trends in career choices among youth, predicting the preferences of future students, optimizing educational strategies, using the model as a guidance and counseling tool for youth, planning resources for preparing specializations or study programs, and improving graduation rates in universities;
- Definition of the primary objective, secondary objectives, research methodology, and the purpose of this thesis;
- Design and population of multiple databases necessary for administering the questionnaires: for the initial questionnaire 3, a database was developed for the top 100 high schools in Romania based on the average results of the baccalaureate exam. Similarly, for the referential questionnaire, a database was created with 497 high schools;
- Design and dissemination of initial questionnaire number 1, which contributed to the creation of the referential questionnaire. It contained 36 questions, was administered online to graduates, and gathered 556 valid responses;

- Design and dissemination of initial questionnaire number 2, which contributed to the creation of the referential questionnaire. It was administered online to high school students, divided into 9 sections, containing 30 questions, and gathered 399 valid responses;
- Design and dissemination of initial questionnaire number 3, which contributed to the creation of the referential questionnaire. It was physically administered to high school students and sent to the top 100 high schools based on baccalaureate exam results. A total of 10,000 questionnaires were sent, 100 for each high school, containing 42 questions, and 6,843 valid responses were collected;
- Design and dissemination of initial questionnaire number 4, which contributed to the creation of the referential questionnaire. It was administered online to students at Politehnica University of Bucharest, divided into 6 sections, containing 23 questions, and gathered 4,084 valid responses;
- Design and dissemination of the referential questionnaire, which contributed to the creation of the prediction model questionnaire. It was physically administered to high school students and sent to 497 high schools in Romania. A total of 99,600 questionnaires were sent, 200 for each high school, and 53,438 valid responses were collected;
- Research on the online presence of Romanian universities and the information available on their websites. A total of 109 universities were analyzed, and a database with 2,310 records was generated;
- Completion of a database with 53,438 valid responses, each containing 113 characteristics, resulting in over 5,200,000 records;
- Organization and management of the data collection process, which led to the creation of the mathematical model for determining the degree of conviction of students in choosing their desired faculties/universities, by presenting the context and importance of data collection;
- Description of the dataset by presenting the types of questions and the information included in the questionnaire, as well as the size and distribution of the dataset;
- Presentation of data collection methods through sample selection and geographic distribution, along with a description of how the questionnaires were administered and interpreted. Additionally, an evaluation and analysis of the response rate was conducted, achieving an excellent percentage of over 80%, exceeding the 75% standard in the specialized literature;
- Justification for choosing the Random Forest algorithm and its use to obtain the weights included in the mathematical model of the degree of conviction. The use of the algorithm was presented by outlining its characteristics, from implementation and testing to the evaluation of the model's data;
- Identification of 89 features of the dataset that were used in training the artificial intelligence algorithm;
- Generation of 89 histograms based on the dataset records for analyzing responses to each question;

- Design of a mathematical prediction model for high school students' career decisions, containing 16 defining parameters obtained from the analysis of over 5,200,000 records representing students' responses to 53,438 valid questionnaires out of 99,600 sent. The parameter values were generated by applying the Random Forest machine learning algorithm, ranked in order of importance based on the resulting weights;
- Determination and validation of the weights for each parameter included in the prediction model of high school students' career decisions through the training of the Random Forest machine learning algorithm;
- Design and implementation of a web application where students can access the developed questionnaire. The application also calculates the degree of conviction based on the scores provided by users for each question, as well as the weights for each factor, obtained through the previously presented machine learning algorithm. The web application builds a profile for each student who completes the questionnaire, making it a useful tool in the recruitment process;
- Design of a guide for students that analyzes each influence factor and offers useful advice for making university choices. The initial structure of the guide was generated by the OpenAI algorithm, and it was further modified and expanded.

Building on this research and highlighting the original contributions made, the following proposals and future research directions can be considered:

1. ***Redoing the existing research by restructuring the influence factors differently;***
2. ***Conducting a longitudinal analysis of the data;***
3. ***Adding new influence factors;***
4. ***Creating a new questionnaire;***
5. ***Developing the web application further;***
6. ***Utilizing Big Data technologies;***
7. ***Exploring other advanced artificial intelligence and machine learning algorithms;***
8. ***Implementing a Blockchain system in high schools.***

BIBLIOGRAPHY

1. Abbeel, P., Schulman, J. (2016). „Deep Reinforcement Learning through Policy Optimization.” Open AI / Berkeley AI Research Lab.
2. Abe, S. (2005). „Support vector machines for pattern classification (Vol. 2, p. 44). Springer, London.
3. Aggrawal, C. C. (2018). „Neural networks and deep learning: A textbook.” Springer Cham.
4. Agrey, L. & Lampadan, N. (2014). Determinant factors contributing to student choice in selecting a university. Journal of Education and Human Development. 3(2), 391-404.
5. Alasuutari, P., Bickman, L., Brannen, J. (2008). "The SAGE Handbook of Social Research Methods." Sage Publications, London, UK.
6. Alonderiene, R. & Klimavičiene, A. (2013). Insights into Lithuanian students' choice of university and study program in management and economics. Journal of Contemporary Management Issues.
7. Alpaydin, E. (2020). „Introduction to machine learning.” MIT press. Cambridge, USA.
8. Azunre, P. (2021). „Transfer learning for natural language processing.” Manning Publications Co., New York, USA.
9. Babbie, E. (2008). "The practice of social research." Cengage Learning, USA.
10. Babbie, E. R. (2016). The practice of social research. Cengage Learning.
11. Barber, D. (2012). „Bayesian reasoning and machine learning.” Cambridge University Press, USA.
12. Bastedo, M. (2021). Holistic admissions as a global phenomenon. In Higher education in the next decade (pp. 91-114). Brill.
13. Belle, V. (2020). Symbolic Logic meets Machine Learning: A Brief Survey in Infinite Domains. SpringerLink.
14. Bertsekas, D. (2019). „Reinforcement learning and optimal control.” Athena Scientific.
15. Bhandari, A. (2020). „Feature scaling for machine learning: Understanding the difference between normalization vs. standardization.” Analytics Vidhya, 34.
16. Bhatt, G. D., & Zaveri, J. (2002). The enabling role of decision support systems in organizational learning. Decision Support Systems, 32(3), 297-309.
17. Bishop, C. M. (2006). „Pattern Recognition and Machine Learning” Springer Science+Business Media, LLC, NY, USA.
18. Blenker, P., Dreisler, P., & Kjeldsen, J. (2006). Entrepreneurship Education - The New Challenge Facing the Universities: A framework for understanding and development of entrepreneurial university communities. (2006-02 ed.) Department of Management, Aarhus School of Business.
19. Bonaccorso, G. (2017). „Machine learning algorithms.” Packt Publishing Ltd.
20. Brace, I. (2018). "Questionnaire design: How to plan, structure and write survey material for effective market research." Kogan Page Publishers, London, UK.
21. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324.

22. Briggs, S. (2006). An exploratory study of the factors influencing undergraduate student choice: the case of higher education in Scotland. Studies in Higher Education, 31(6), 705-722.
23. Bryman, A. (2016). "Social research methods." Oxford University Press, Oxford, UK.
24. Buckley, J., Letukas, L., & Wildavsky, B. (2018). Measuring Success: Testing, Grades, and the Future of College Admissions. Baltimore, MD: Johns Hopkins University Press.
25. Bumblauskas, D., Gemmill, D., Igou, A., & Anzengruber, J. (2017). Smart Maintenance Decision Support Systems (SMDSS) based on corporate big data analytics. *Expert Systems with Applications, 90*, 303–317.
26. Buskirk, T. D., Kirchner, A., Eck, A., & Signorino, C. S. (2018). „An introduction to machine learning methods for survey researchers." Survey Practice, 11(1).
27. Cabrera, A. F. (1994). Logistic regression analysis in higher education: An applied perspective. Higher education: Handbook of theory and research, 10, 225-256.
28. Cattaneo, M., Horta, H., Malighetti, P., et al. (2017). „Effects of the financial crisis on university choice by gender. Higher Education."
29. Cattaneo, M., Horta, H., Malighetti, P., et al. (2017). „Effects of the financial crisis on university choice by gender. Higher Education."
30. Chaudhuri, A., Mukerjee, R. (2020). "Randomized response: Theory and techniques." Routledge, New York, USA.
31. Ciolacu, M., Tehrani, A. F., Binder, L., & Svasta, P. M. (2018, October). Education 4.0-Artificial Intelligence assisted higher education: early recognition system with machine learning to support students' success. In 2018 IEEE 24th International Symposium for Design and Technology in Electronic Packaging (SIITME) (pp. 23-30). IEEE.
32. Cohen, I., Huang, Y., Chen, J., Benesty, J., Benesty, J., Chen, J., & Cohen, I. (2009). Pearson correlation coefficient. Noise reduction in speech processing, 1-4.
33. Conerly, T., Holmes, K., Tamang, A. L. (2021). "Introduction to Sociology 3e." OpenStax, Rice University, Houston, USA.
34. Cook, T., & Rushton, B. S. (2010). How to Recruit and Retain Higher Education Students: A Handbook of Good Practice. Routledge.
35. Coulom, R. (2006). „Monte-Carlo Tree Search and Rapid Action Value Estimation in Computer Go. Artificial Intelligence." 170(11), 1856-1875.
36. Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage Publications.
37. Cummings, M. L. (2017). Automation bias in intelligent time critical decision support systems. In Decision making in aviation (pp. 289-294). Routledge.
38. Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. Ecology, 88(11), 2783-2792.
39. Daston, L., Lunbeck, E. (2011). "Histories of scientific observation." University of Chicago Press, Chicago, USA.
40. De Leeuw, E. D., Hox, J., Dillman, D. (2012). "International handbook of survey methodology." Routledge, New York, USA.
41. De Wit, H. (2007). European integration in higher education: The Bologna process towards a European higher education area. In International handbook of higher education (pp. 461-482). Dordrecht: Springer Netherlands.

42. Dietterich, T. G. (2000). Ensemble methods in machine learning. In International workshop on multiple classifier systems (pp. 1-15). Springer, Berlin, Heidelberg.
43. Dillman, D. A., Jolene, D. S., Leah, M. C. (2014). "Internet, phone, mail, and mixed-mode surveys: The tailored design method." John Wiley & Sons, New Jersey, USA.
44. Djafarova, E., & Bowes, T. (2021). 'Instagram made Me buy it': Generation Z impulse purchases in fashion industry. Journal of retailing and consumer services, 59, 102345.
45. Drăgoi, M. V., Puiu, R. A, **Petrea, G.**, Burtea, C., Spiridon – Mocioacă, T. M., Puiu, M., Suciu C. P., Improving The Learning Process For Students. Personalizing Learning Through Software Applications, Scientific Buletin of University POLITEHNICA of Bucharest, Series C – Electrical Engineering and Computer Science – în curs de publicare
46. Domingos, P. (2015). The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World. Basic Books.
47. Donges, N. (2024). Random forest: A complete guide for machine learning. Built In. https://builtin.com/data-science/random-forest-algorithm.
48. Drewes, T., & Michael, C. (2006). How do students choose a university?: an analysis of applications to universities in Ontario, Canada. Research in Higher Education, 47, 781-800.
49. Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling. American Journal of Theoretical and Applied Statistics, 5(1), 1-4. https://doi.org/10.11648/j.ajtas.20160501.11.
50. Ezugwu, A. E., Ikotun, A. M., Oyelade, O. O., Abualigah, L., Agushaka, J. O., Eke, C. I., & Akinyelu, A. A. (2022). A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. Engineering Applications of Artificial Intelligence, 110, 104743. https://doi.org/10.1016/j.engappai.2022.104743.
51. Fleacă, B., Fleacă, E., **Petrea, G.**, Ghiban A. (2013). Process management and project management – towards a theoretical analysis. 6th International Conference of Management and Industrial Engineering (ICMIE). ISBN: 2344-0937. WOS: 000448492200046.
52. Fleacă, E., **Petrea, G.** (2011). Approacihng marketing research based on the project management principles. 5th International Conference of Management and Industrial Engineering (ICMIE). ISBN: 978-973-748-658-5. WOS: 000325051100014.
53. Frankfort-Nachmias, C., Leon-Guerrero, A., Davis, G. (2019). "Social statistics for a diverse society." Sage Publishing, California, USA.
54. Fry, R. W. (2017). Data-driven decision-making and the challenges facing Pennsylvania school administrators, compliance vs. conviction.
55. Foddy, W. (1993). "Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research." Cambridge University Press, Cambridge, UK.
56. Fowler, J., Floyd, J. (2013). "Survey research methods." Sage Publishing, Boston, USA.
57. Friedrich-Ebert-Stiftung România. (2023). Analiza politicii sociale în România 2022-2023. https://library.fes.de/pdf-files/bueros/bukarest/21029.pdf
58. Garcia, E. (2016). Data Cleaning: A Practical Guide. Springer.
59. García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2016). Big data preprocessing: methods and prospects. Big Data Analytics, 1(1), 1-22.

60. Géron, A. (2022). „Hands-On Machine Learning with Scikit-Learn and TensorFlow – Concept, Tools, and Techniques to Build Intelligent Systems.” O'Reilly Media, Sebastopol, USA.
61. Gilchrist, R., Phillips, D. & Ross, A. (2005). Participation and potential participation in UK higher education. In Higher education and social class (pp. 87-108). Routledge.
62. Goodfellow, I., Bengio, Y., & Courville, A. (2016). „Deep learning.” MIT press, Cambridge, USA.
63. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., et al. (2020). „Generative adversarial networks.” Communications of the ACM, 63(11), 139-144.
64. Greenberg, J. (2003). Diversity, the university, and the world outside. Colum. L. Rev., 103, 1610.
65. Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). „Machine learning for social science: An agnostic approach.” Annual Review of Political Science, 24, 395-419.
66. Gu, M., Ma, J., & Teng, J. (2017). The Gaokao Experience of Chinese Students. In Portraits of Chinese Schools. Perspectives on Rethinking and Reforming Education (pp. 45-70). Springer.
67. Guy, I., Avraham, U., Carmel, D., Ur, S., Jacovi, M., & Ronen, I. (2013). Mining expertise and interests from social media. In Proceedings of WWW 2013 (pp. 515–526). International World Wide Web Conferences Steering Committee.
68. Hadiji, F., Mladenov, M., Bauckhage, C., & Kersting, K. (2015). Computer science on the move: inferring migration regularities from the web via compressed label propagation. In Proceedings of IJCAI 2015 (pp. 171–177). AAAI Press.
69. Harford, T. (2019). "Fifty things that made the modern economy." Little, Brown Book Group, London, UK.
70. Hastie, T., Tibshirani, R., Friedman, J. (2017). „The elements of statistical learning: Data mining, inference, and prediction.” Springer, New York, USA.
71. Henry, G. T. (1990). „Practical sampling.” Vol. 21. Sage Publications, London, UK.
72. Hirschauer, N., Grüner, S., & Mußhoff, O. (2023). Fundamentals of Statistical Inference: What is the Meaning of Random Error? Springer. https://doi.org/10.1007/978-3-662-66786-6.
73. Ho, T. K. (1995). Random decision forests. In Proceedings of 3rd international conference on document analysis and recognition (Vol. 1, pp. 278-282). IEEE.
74. Holland, J.L. (1997). Making Vocational Choices: A Theory of Vocational Personalities and Work Environments. Psychological Assessment Resources.
75. Horvath, J. (2019). „Machine learning with TensorFlow, Second Edition.” Manning Publications, Shelter Island, USA.
76. Institutul Național de Statistică. (2023). Revista Română de Statistică - Supliment, Nr. 11/2023.https://insse.ro/cms/sites/default/files/field/publicatii/revista_romana_statistica_supliment_11_2023.pdf
77. IvyWise. (2023). College Admission Secrets: How Colleges Read Applications. https://www.ivywise.com/blog/college-admission-secrets-how-colleges-read-applications/ (accesat la: 10.02.2023)

78. Janczyk, M., & Pfister, R. (2023). Understanding Inferential Statistics: From A for Significance Test to Z for Confidence Interval. Springer. https://doi.org/10.1007/978-3-662-66786-6.
79. James, G., Witten, D., Hastie, T., Tibshirani, R. (2021). „An introduction to statistical learning: with applications in R." Springer, New York, USA.
80. Jiang, H., & Smith, M. (2022). „Introduction to artificial intelligence: A comprehensive guide for beginners." Addison-Wesley, Boston, USA.
81. King, R. B. (2011). Multinomial logistic regression. SAS Data Analysis Examples. University of California.
82. Krumboltz, J.D., Mitchell, A.M., & Jones, G.B. (1976). A social learning theory of career selection. Counseling Psychologist, 6(1), 71-81.
83. Kubacki, K. & Siemieniako, D. (2007). A comparative analysis of the factors influencing the decision to study at university in Poland and Belarus. Journal of Business Economics and Management, 8(2), 25-31.
84. Kumar, T. S. (2020). Data mining based marketing decision support system using hybrid machine learning algorithm. Journal of Artificial Intelligence, 2(03), 185–193.
85. Leavy, P. (2020). "Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches." Guilford Publications, New York, USA.
86. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
87. Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2(3), 18-22.
88. Lin, D. (2023). „A practical guide to feature selection in machine learning." Towards Data Science.
89. Lin, J., & Reid, C. (2021). „Machine learning and human intelligence: The future of education in the 21st century." Springer, Singapore.
90. Lohr, S. L. (2019). Sampling: Design and Analysis (2nd ed.). CRC Press.
91. Lolli, F., Ishizaka, A., Gamberini, R., Rimini, B., Balugani, E., & Prandini, L. (2017). Requalifying public buildings and utilities using a group decision support system. Journal of Cleaner Production, 164, 1081–1092.
92. López, V., Fernández, A., García, S., Palade, V., & Herrera, F. (2013). An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. Information sciences, 250, 113-141.
93. Ma, X., & Baum, S. (2012). Trends in graduate student financing: Selected findings from the NPSAS: 2007-08. National Center for Education Statistics.
94. Mangai, S., & Sujatha, P. (2014). „A review on ensemble learning algorithms." International Journal of Computer Applications, 130(2), 7-11.
95. Martin, E. (2006). Survey questionnaire construction. Retrieved from https://www.census.gov/srd/papers/pdf/rsm2006-13.pdf
96. McArthur, R. C. (2005). Faculty-based advising: An important factor in community college retention. Community College Review, 32(4), 1-19.
97. Mechno R.A., Doicin C.V., **Petrea G.**, Scurtu D., Buzatu A.I. (2020), *Comparative Analysis of International Ranking Systems for Universities*, Proceedings of the 36th

International Business Information Management Association Conference, ISBN: 978-0-9998551-5-7, Granada, Spania.
98. Michel, R., & Pollard, S. (2020). Higher Education Admissions Practices: An International Perspective. Cambridge University Press.
99. Mitchell, T. M. (2006). „The discipline of machine learning." Vol. 9. Pittsburgh: Carnegie Mellon University, School of Computer Science, Machine Learning Department.
100. Mitchell, T. (1997). „Machine learning." McGraw-Hill, New York, USA.
101. Molina, M., & Garip, F. (2019). Machine Learning for the Social Sciences: An Introduction. Annual Review of Sociology, 45, 351-370. doi:10.1146/annurev-soc-073018-022453.
102. Moreno, J., & Calderón, F. (2020). „Machine learning in action: A practical guide for beginners." McGraw-Hill, New York, USA.
103. Muller, C., & Schiller, C. (2000). Leveling the playing field? Students' educational attainment and states' performance testing. Sociology of Education, 74(2), 107-128.
104. Mullins, D. (2023). „Advanced data visualization techniques in R." Packt Publishing Ltd.
105. Nagel, E. (2023). „The ethics of artificial intelligence: A comprehensive overview." Routledge, New York, USA.
106. National Association for College Admission Counseling. (2018). Guide to International University Admission. Arlington, VA: NACAC (accesat la: 10.02.2023)
107. Nelson, A. I. (2006). Evolution of the college admissions essay: Exploring the role of student experiences and educational backgrounds. Journal of Higher Education, 77(5), 675-688.
108. Ng, H. K. T., & Heitjan, D. F. (2022). Recent Advances on Sampling Methods and Educational Statistics: In Honor of S. Lynne Stokes. Springer. https://doi.org/10.1007/978-3-030-70638-0.
109. Norvig, P. (2014). „Paradigms of artificial intelligence programming: Case studies in Common Lisp." Morgan Kaufmann, San Francisco, USA.
110. OECD. (2013). Education at a glance 2013: OECD indicators. OECD Publishing.
111. O'Reilly, C. A. (1982). Variations in decision makers' use of information sources: The impact of quality and accessibility of information. Academy of Management Journal, 25(4), 756-771.
112. Olson, K. (2010). An examination of questionnaire evaluation by expert reviewers. Field Methods, 22(4), 295-318.
113. Pal, M. (2005). Random forest classifier for remote sensing classification. International journal of remote sensing, 26(1), 217-222.
114. Pal, S. (2023). „Transformers for Natural Language Processing: Concepts, Tools, and Techniques." Packt Publishing Ltd.
115. Pedregosa, F., Varoquaux, G., Gramfort, A., (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, *12*, 2825-2830.
116. Petersen, K. J., Handfield, R. B., & Ragatz, G. L. (2005). „Supplier integration into new product development: Coordinating product, process and supply chain design." Journal of Operations Management, 23(3-4), 371-388.

117. **Petrea, G.**, Puiu, R. A. (2023). Students' and high school students' preferences in choosing a university: a sociological analysis, The 11th International Conference of Management and Industrial Engineering. – în curs de indexare.
118. **Petrea, G.**, Puiu, R. A., Doicin, C. V. (2023). The impact of virtual instruments on Romanian education system, The 22nd International Conference of Nonconventional Technologies. – în curs de indexare.
119. **Petrea, G.**, Puiu, R. A., Mocanu, B. C., Al-Dulaimi O. M. K. (2024). Determining the Degree of Conviction of Students in University Selection Using the Random Forest Algorithm: An Approach for Adaptive and Personalized Decision Support System in Education. 23rd RoEduNet Conference. Networking in Education and Research. – în curs de indexare.
120. Poli, R., Langdon, W. B., McPhee, N. F., & Koza, J. R. (2008). A Field Guide to Genetic Programming. Lulu Enterprises.
121. Polya, G. (2014). „How to solve it: A new aspect of mathematical method." Princeton University Press, Princeton, USA.
122. Price, J. (2006). Does no child left behind really capture school quality? Evidence from an urban school district. Economics of Education Review, 25(5), 567-576.
123. Puiu (Mechno), R. A., **Petrea, G.**, Scurtu, D., Cristea, M. N., Dobrinoiu, R. A. (2020). The Impact of International Rankings on Students' Decision to Choose a University, Proceedings of the 36th International Business Information Management Association Conference. Journal of e-Learning and Higher Education. ISBN: 978-0-9998551-5-7.
124. Purcarea, A. A., Țigănoaia, B., **Petrea, G.** (2011). Considerations regarding the implementation and certification within an organization of an information security management system. 5th International Conference of Management and Industrial Engineering (ICMIE). ISBN: 978-973-748-658-5. WOS: 000325051100013.
125. Ramirez, E., Chen, R., & Pasupathy, R. (2015). „An exploration of the factors influencing undergraduate student choice of higher education institution." Research in Higher Education Journal, 30, 1-18.
126. Roberts, J. (2016). „The student engagement toolkit: A practical guide to engaging students in higher education." Higher Education Academy, York, UK.
127. Rouhani, S., Ghazanfari, M., Jafari, M., et al. (2017). „A survey on the applicability of data mining techniques to enhance effectiveness of enterprise resource planning systems." Enterprise Information Systems, 11(2), 219-259.
128. Rowley, J. (2005). The student experience: an approach to strategic development of the curriculum. Higher Education Academy.
129. Russell, S., & Norvig, P. (2016). Artificial Intelligence: A Modern Approach. Pearson.
130. Saito, T., & Rehmsmeier, M. (2015). „The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets." PloS one, 10(3), e0118432.
131. Salt, J., & Wood, P. (2014). Higher education, student migration and the internationalization of education. In Global skilled migration (pp. 115-136). Routledge.
132. Schmidhuber, J. (2015). „Deep learning in neural networks: An overview." Neural networks, 61, 85-117.

133. Scholz, R. W. (2011). „Environmental literacy in science and society: From knowledge to decisions." Cambridge University Press, Cambridge, UK.
134. Schütze, H., Manning, C. D., Raghavan, P. (2008). „Introduction to information retrieval." Vol. 39. Cambridge University Press, Cambridge, UK.
135. Scurtu, D., Puiu (Mechno), R. A., **Petrea, G.**, Ivan A. (2022). The Impact of Cross-Platform Applicationson the Digitalization of Educational Institutions, Scientific Buletin of University POLITEHNICA of Bucharest, Series C – Electrical Engineering and Computer Science, Number 3, Vol.84, Iss.3, ISBN: 2286-9540. WOS:000865792900007.
136. Selingo, J. (2020). Who Gets In and Why: A Year Inside College Admissions. Scribner.
137. Shawe-Taylor, J., & Cristianini, N. (2004). „Kernel methods for pattern analysis." Cambridge University Press, Cambridge, UK.
138. Siems, A. & Busler, M. (2012). Factors in student university choice: A comparative study. Journal of Business & Economics Research (JBER), 10(3), 529-532.
139. Silver, D., Schrittwieser, J., Simonyan, K., et al. (2017). „Mastering the game of Go without human knowledge." Nature, 550(7676), 354-359.
140. Smith, L. (2009). Factors influencing student choice in higher education. Perspectives: Policy and Practice in Higher Education, 13(3), 89-92.
141. Srivastava, N. (2018, May 15). Deep learning: Overview of neurons and activation functions. Medium. https://srnghn.medium.com/deep-learning-overview-of-neurons-and-activation-functions-1d98286cf1e4 (accesat la: 05.08.2024).
142. Sutton, R., & Barto, A. (2018). „Reinforcement learning: An introduction." MIT press, Cambridge, USA.
143. Thompson, M.N., & Subich, L.M. (2006). The relation of social status to the career decision-making process. Journal of Vocational Behavior, 69(2), 289-301.
144. Ting, Y., & Huang, C. (2018). „Deep learning for medical image analysis." Elsevier.
145. Weiner, M., & Metzger, K. (2021). „AI ethics: The ethical and social implications of artificial intelligence." Springer, New York, USA.
146. Weiss, S. M., & Indurkhya, N. (2015). „Predictive data mining: A practical guide." Morgan Kaufmann, San Francisco, USA.
147. Wu, C., & Thompson, M. E. (2020). Sampling Theory and Practice. Springer. https://doi.org/10.1007/978-3-030-43494-8.
148. Xiong, Z., Zhong, R., Deng, W., & Yang, G. (2020). Career Choice Prediction Based on Campus Big Data—Mining the Potential Behavior of College Students. Applied Sciences, 10(8), 2841. doi:10.3390/app10082841.
149. Yang, X., & Han, Z. (2018). „Data mining for business analytics: Concepts, techniques, and applications in R." Wiley, Hoboken, USA.
150. Yarkoni, T., & Westfall, J. (2017). „Choosing prediction over explanation in psychology: Lessons from machine learning." Perspectives on Psychological Science, 12(6), 1100-1122.