The data association algorithm for the formation of optional IT-courses list system

Andrii Roskladka[1[0000-0002-1297-377X]], Nataliia Roskladka[1[0000-0001-7333-4050]]
Ganna Kharlamova[2[0000-0003-3614-712X]] and Andriy Stavytskyy[[0000-0002-5645-6758]]

[1] Kyiv National University of Trade and Economics,
Kyoto st. 19, 02156, Kyiv, Ukraine
a.roskladka@knute.edu.ua, n.roskladka@knute.edu.ua
[2] Taras Shevchenko National University of Kyiv,
Vasylkivska st. 90A, 03022, Kyiv, Ukraine
akharlamova@ukr.net

Abstract. The article contains a study of the principles of student's educational trajectory formation by using modern technologies in data analysis. There is a mandatory requirement to have the selective component (optional to a student) among the curriculum educational components. This rule is legislated in the laws «On Education» and «On Higher Education» of Ukraine as well as in the normative documents on accreditation of educational programs, defined by the Standards and recommendations on quality assurance in the European Space of Higher Education (ESG) and the National Agency for Quality Assurance of Higher Education. However, adherence to the principles of the individual educational trajectory formation is mostly formal and is reduced to offering students a non-coherent list of courses. On the one hand, this leads to the disorientation of a student, who cannot see the systemic perspective of his future profession in the initial list of study courses, and therefore cannot consciously choose the optimal set of optional courses. On the other hand, the unknown choice of courses by students leads to situational management of the educational process at the HEI. A large number of courses create significant difficulties in managing the selection process. To analyse the process of individual educational trajectory formation, the authors propose to use methods of data association and, in particular, the apriori algorithm for the formation of associative rules. The procedure of popular sets of elective courses formation, the configuration of associative rules of educational courses choice is studied. The characteristics of these rules quality are calculated. The example of the procedure implementation in analytical platform Deductor Studio is considered.

Keywords: individual educational trajectory, selective study courses, Data Science, data association, associative rules, apriori algorithm.

1. Introduction

The prior goal of modern higher education is to prepare in HEIs competent specialists with a harmonious combination of personal and professional qualities, that are capable of self-development and self-realization in the future professional activity. In this regard, the orientation of the educational process towards the student becomes particularly relevant and valuable. It includes taking into account the individual opportunities and needs of the students. This orientation allows students to choose components of educational programs and to form an individual educational trajectory.

The educational trajectory determines the direction of the student in the educational space. Its existence is not a norm, but a fact of reality. Different scenarios for defining this trajectory are possible under different educational conditions. According to the traditional approach, the HEI defines a single trajectory of educational achievement results by students in accordance with the National standard. The more modern method offers several possible paths for groups of students with different capabilities. Both such approaches do not take into account the views of the student. Because of the goals, content, forms, methods, results are determined directly by the teacher. At the same time, a large proportion of students have a low level of motivation to achieve goals, as they do not see personal meaning in such educational activities. One way to overcome this problem in higher education is to form an individual educational trajectory.

Recognition of students' right to an individual educational trajectory is one of the progressive innovations of the «Law on Education». This law provides for «a personal way of realizing the student 's potential, is formed taking into account his abilities, interests, needs, motivation, opportunities and experience, is based on the student 's choice of types, forms and pace of education, courses of educational activities and their offered educational programs, educational courses and the level of their complexity, methods and means of education» [1].

The main tool for the realization of personal potential is in optional courses choice. According to the Law of Ukraine "On Higher Education" [2], students have the right to choose courses in the amount of at least 25% ECTS credits from the total educational program (paragraph 15, part 1, Article 62).

Requirements for individual educational trajectory are contained in the regulatory documents for the accreditation of educational programs defined by the Standards and Recommendations for Quality Assurance in the European Higher Education Space (ESG) [3] and the National Agency for Quality Assurance of Higher Education [4].

The authors of the article are experts of the National Agency for Quality Assurance of Higher Education and may argue that compliance with the principles of formation of individual educational trajectory in Ukrainian HEIs is mainly formal and is reduced to offering students to choose an unrelated list of courses. This leads to the disorientation of the students, who at the initial courses of study cannot see the systemic perspective of their future profession, and therefore cannot consciously choose the optimal set of optional courses. The unknown choice of courses by students leads to situational management of the educational process at the higher educational establishment. A large number of courses create significant difficulties in managing the selection process. System analysis of the process of the individual educational trajectory formation can be effectively carried out with the help of Data Science tools [5].

2. Research Methodology

The theoretical and methodological justification for research is the fundamental principles of the system approach, analysis and synthesis of information, dialectical method in the justification of the use of information technologies.

In particular, the following scientific methods are used in research:

- association method – for combine courses into logical modules and create causal relationships between modules of optional courses;
- apriori algorithm – for the formation of frequent itemset sets of optional courses and a system of associative rules construction based on these sets;
- a graphic method – for creation of the formation scheme of frequent itemsets subsets of optional courses with various weights;
- a method of the quantitative analysis – for calculation of characteristics of support and confidence of associative rules and its sorting.

The information basis for the research is data on the content of the optional components of curricula set and results of course selection by students of the Faculty of Information Technologies of Kyiv National University of Trade and Economics.

3. Literature and Hypothesis Development

According to the National Strategy for the Development of Education in Ukraine for 2012-2021, the main direction of the educational space development is the introduction of the individualized learning concept [6]. Implementation of individualized learning is ensured by the determination of individual educational trajectory.

Basic concepts of individual educational trajectory are given, for example, in the work of I. Krasnoshchok [7]. Analysis of the advantages of its use in the educational process and review of the principles to construct an online information system supporting the student's educational trajectory is the subject of research by S. Sharov and T. Sharova in the article [8].

Various aspects of the information and communication technologies introduction in the educational process are constantly attracting the attention of many researchers.

Systematic studies of informatization of education and practical implementation of information and communication technologies in the educational sphere of Ukraine were conducted by Yu. Bykov, O. Burov, A. Gurzhii, M. Zhaldak and others in [9].
The needs of digital transformation, which require special flexibility from modern universities, the creation of a digital learning environment to support educational activities are discussed in the studies of O. Kuzminska, M. Mazorchuk, N. Morze, O. Kobylin [10]. The formation of competences in the field of information technologies was presented in the study by N. Morze and M. Umryk [11]. The development of innovative entrepreneurial universities in Ukraine on the platform of digitalization is assessed in [22]. The principles and structure of the information support system for

the individual educational trajectories as an open modular portal were investigated by A. Bogdanov, I. Chepovoy, P. Ukhan, L. Yurchuk in [12]. It analyzes Tools for Mobility, Tools for Quality, Tools for Transparency, and Portals and Databases, which can be useful for building and implementing an individual educational trajectory. The formation of an integrated quantitative assessment of the HEI activity is proposed by V. Bykov, A. Biloshchytskyi, O. Kuchanskyi, Yu. Andrashko, O. Dikhtiarenkoand S. Budnik in [13].

The problems of big data processing are extremely pressing today. Scientists D. Dietrich, B. Heller, B. Yang tell how it is effective to use Data Science tools in almost all areas of human activity [14]. An associative analysis is one of the main components of Knowledge Discovery in Databases and its main component, Data Mining [15]. Practical implementation of associative rules in various fields of scientific research is represented by C. Zhang and S. Zhang in [16] and G. Bhavani and S. Sivakumari in [17].

The main purpose of the vast majority of management systems is the integrated management of HEI. Among the well-known foreign systems of automation it is necessary to assign such as MyEdu – University Automation Software [18], Eifell Corp Services and Custom eLearning [19], CyberVision University Management System [20].

In Ukraine, at the state level, there is no more than a single state electronic database on education, which is an integrated information and telecommunications management system. The Ukrainian software market offers its solutions for the automation of the educational process. The most famous is the Automated control system of the higher educational institution ACS "University", developed by the Research Institute of Applied Information Technologies of the Cybernetics Center of the National Academy of Sciences of Ukraine [21], Computer Systems Packages "Dean's Office", "COLLOQUIUM", "PS-Staff" of the private enterprise "Politek-soft" [22], Program Complex "ALMA-MATER" of the company "Direct IT" [23] and others.

None of the control systems described above contains a module for automating the process of choosing courses. As a result of the analysis, the authors found only a few attempts of a scientific approach to solving this problem.

A. Kravets and R. Al-Shaebi in the article [24] offer a method of automated formation of the individualized curriculum. A. Kardan with co-authors describes using a neural network approach for the process of student behaviour modelling in choosing courses [25]. I. Ognjanovic, D. Gasevicand S. Dawson created a model for predicting the student course choice [26].

These approaches do not use the logical relationships between the courses that the student must choose. However, this is a necessary condition for the formation of a high-quality curriculum for the training of highly qualified specialists. The educational program must have a clear structure; educational components should be a logically interconnected system and together enable the achieving of learning goals and outcomes.

The search for associative rules is a good tool for solving the problem of establishing relationships between courses and a systematic approach to building a quality logical and structural system of educational components.

4. Objective and Context of Research

The purpose of this research is to provide a tool for students for a conscious choice of courses, to analyze the frequent itemsets of optional courses in the formation of professional qualities of future specialists. Along it is also to replace inefficient situational management with a systemic approach in the management of the educational process at the HEI.

The large volumes of modern databases have generated a steady demand for new scalable data analysis algorithms. Systematizing the complex structure of big data has led to the emergence of affinity analysis, one of the most common methods of Data Mining. The purpose of this method is to investigate the relationship between events that occur together.

One popular method of knowledge discovery is the algorithms for associative rules finding. For the first time, the associative rules finding task has been proposed to find typical shopping patterns made in supermarkets, so it is sometimes also called market basket analysis.

Associative rules are now applied to solve problems in various areas:

- identifying sets of goods in supermarkets that are often bought together or never are bought together;
- identification of a part of clients who are positive about innovations in their services;
- determining the profile of visitors to the web resource.
- identification of a part of cases where new drugs cause dangerous side effects, etc.

The authors of the article aimed to embody the idea of affinity analysis and associative rules to optimize the educational process in the formation of the individual educational trajectory of the student.

5. Results

5.1. Structure of Initial Data to Form Associative Rules

Although the National Agency for Quality Assurance of Higher Education promotes a wide selection of courses for students and claims not to limit their choice to separate blocks, the division of optional courses into logical units is a prerequisite for analyzing the structural and logical relationship between courses.

As initial data, we take the recommended optional courses of the professional training at the Faculty of Information Technologies at the Kyiv National University of Trade and Economic (Table 1). Once again, it should be stressed that there is no division of optional courses into blocks in the curricula of the Faculty of Information Technology. Courses are combined into logical units only as part of this study.

Table 1: Division selective courses into logical units

Programming Unit

Enterprise programming Java	Web-design and web-programming	Cross-platform programming	Functional and logical programming
Database and Knowledge Unit			
The technology of design and administration of databases and data storage	Technology for distributed databases and knowledge creating	Knowledge representation and processing technologies in intelligent systems	Distributed systems and parallel computing technologies
Data Processing Unit			
Data analysis technologies	Tools of business intelligence	Business analytics of an enterprise	Computer technologies of data processing
Intellectual Analysis Unit			
Expert systems	Intelligent systems	Machine learning	Artificial Intelligence
Internet Technologies Unit			
Cloud and GRID technologies	Internet technologies in business	Digital technologies in business	Digital systems and technologies
Information Security Unit			
Cryptographic methods of information security	Biometric authentication technologies in information systems	Methods and means of protecting information in computer networks	Security of Internet resources

Source: formed by the authors based on the real study curriculas

5.2. Formation and intellectual analysis of relationships between units of courses

According to the Regulation on the Organization of the educational process of students, the applicants choose educational courses for the next academic year in February. During the study, students are asked to choose one course from each logic unit. It should be noted that before the survey, students had the opportunity to familiarize themselves with presentations of optional courses to raise the consciousness of their choice.

A total of 100 faculty students enrolled in the second, third or fourth study years were interviewed. Such a sample is appropriate, as it takes into account not only the desire of students of junior courses to gain knowledge in certain areas of information technology in the future but also a certain experience of students of senior courses who have already studied some of the offered courses. Results of the courses selection among students of one of the groups of the Faculty of Information Technology are presented in Table 2. In this table, the ID is the serial number of the student who participated in the survey. Each row contains the results of the student-specific selection of six of the 24 courses offered in Table 1.

First, we find single-element course sets by presenting the transaction database from Table 2 in normalized form (Table 3). In this Table U_{ij} – the course name with the course sequence number *j* belongs to the logical unit number *i*. At the intersection of the transaction row and the course, the column is 1 if the course is present in the transaction and 0 otherwise. The column amount will be the frequency at which each course appears in the selection results.

Table 2: Results of students' choice of courses from six logical units

ID	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Unit 6
S1	Enterprise programming Java	Technology for distributed databases and knowledge creating	Data analysis technologies	Expert systems	Internet technologies in business	Security of Internet resources
S2	Cross-platform programming	Distributed systems and parallel computing technologies	Business analytics of an enterprise	Expert systems	Digital systems and technologies	Biometric authentication technologies in information systems
S3	Cross-platform programming	Distributed systems and parallel computing technologies	Data analysis technologies	Machine learning	Cloud and GRID technologies	Biometric authentication technologies in information systems
S4	Enterprise programming Java	Distributed systems and parallel computing technologies	Business analytics of an enterprise	Artificial Intelligence	Internet technologies in business	Security of Internet resources
S5	Enterprise programming Java	Technology for distributed databases and knowledge creating	Data analysis technologies	Machine learning	Cloud and GRID technologies	Security of Internet resources
S6	Web-design and web-programming	Technology for distributed databases and knowledge creating	Data analysis technologies	Expert systems	Internet technologies in business	Cryptographic methods of information security
S7	Cross-platform programming	Technology for distributed databases and knowledge creating	Data analysis technologies	Intelligent systems	Digital technologies in business	Biometric authentication technologies in information systems
S8	Enterprise programming Java	The technology of design and administration of databases and data storage	Computer technologies of data processing	Expert systems	Digital technologies in business	Biometric authentication technologies in information systems
S9	Web-design and web-programming	Technology for distributed databases and knowledge creating	Computer technologies of data processing	Expert systems	Cloud and GRID technologies	Biometric authentication technologies in information systems
S10	Enterprise programming Java	Technology for distributed databases and knowledge creating	Data analysis technologies	Artificial Intelligence	Internet technologies in business	Methods and means of protecting information in computer networks

S11	Enterprise programming Java	Knowledge representation and processing technologies in intelligent systems	Data analysis technologies	Intelligent systems	Digital systems and technologies	Methods and means of protecting information in computer networks
S12	Cross-platform programming	Distributed systems and parallel computing technologies	Business analytics of an enterprise	Expert systems	Digital systems and technologies	Cryptographic methods of information security
S13	Functional and logical programming	Knowledge representation and processing technologies in intelligent systems	Business analytics of an enterprise	Machine learning	Digital technologies in business	Methods and means of protecting information in computer networks
S14	Web-design and web-programming	The technology of design and administration of databases and data storage	Data analysis technologies	Machine learning	Cloud and GRID technologies	Cryptographic methods of information security
S15	Enterprise programming Java	The technology of design and administration of databases and data storage	Business analytics of an enterprise	Machine learning	Internet technologies in business	Methods and means of protecting information in computer networks
S16	Web-design and web-programming	Distributed systems and parallel computing technologies	Business analytics of an enterprise	Expert systems	Internet technologies in business	Methods and means of protecting information in computer networks
S17	Cross-platform programming	Distributed systems and parallel computing technologies	Computer technologies of data processing	Machine learning	Digital systems and technologies	Methods and means of protecting information in computer networks
S18	Web-design and web-programming	Distributed systems and parallel computing technologies	Business analytics of an enterprise	Expert systems	Cloud and GRID technologies	Methods and means of protecting information in computer networks
S19	Cross-platform programming	Distributed systems and parallel computing technologies	Business analytics of an enterprise	Expert systems	Digital systems and technologies	Security of Internet resources
S20	Web-design and web-programming	Distributed systems and parallel computing technologies	Business analytics of an enterprise	Machine learning	Digital technologies in business	Methods and means of protecting information in computer networks

Source: formed by the authors

Table 3: Formation of a set of single-element frequent itemsets

	Unit 1				Unit 2				Unit 3				Unit 4				Unit 5				Unit 6			
	U11	U12	U13	U14	U21	U22	U23	U24	U31	U32	U33	U34	U41	U42	U43	U44	U51	U52	U53	U54	U61	U62	U63	U64
S1	1					1			1				1					1						1
S2			1					1			1		1							1		1		
S3			1					1	1						1		1					1		
S4	1							1			1					1		1						1
S5	1					1			1						1		1							1
S6		1				1			1				1					1			1			
S7			1			1			1					1					1			1		
S8	1				1							1	1						1			1		
S9		1				1						1	1				1					1		
S10	1					1			1							1		1					1	
S11	1						1		1					1						1			1	
S12			1					1			1		1							1	1			
S13				1			1				1				1				1				1	
S14		1			1				1						1		1				1			
S15	1				1						1				1			1					1	
S16		1						1			1		1					1					1	
S17			1					1				1			1					1			1	
S18		1						1			1		1				1						1	
S19			1					1			1		1							1				1
S20		1						1			1				1				1				1	
Sum	**7**	**6**	**6**	**1**	**3**	**6**	**2**	**9**	**8**	**0**	**9**	**3**	**9**	**2**	**7**	**2**	**5**	**6**	**4**	**5**	**3**	**5**	**8**	**4**

Source: formed by the authors

In the practical implementation of associative rule search systems different methods are used, which allow reducing search space to dimensions, providing acceptable computational and time costs, for example, Apriori algorithm [14-17]. The Apriori algorithm is based on the concept of frequent itemsets, i.e. sets with high frequency in a given number of transactions.

In the classic Apriori algorithm, a popular subject set is a subject set with support, equal to or greater than a given threshold. This threshold is called minimum support. However, the problem of establishing relationships between modules imposes limitations on the selection of single-element sets, which must contain courses from different logical units. Therefore, we form the set F_1 of single-element subsets from the most frequent itemsets of each logical unit.

Next, we find two-element sets, forming all possible combinations from F_1 two courses (Table 4). In the future, the most popular representative logical unit U_i will indicate the name of this unit.

Table 4: Formation of a set of two-elementt frequent itemsets

	U1&U2	U1&U3	U1&U4	U1&U5	U1&U6	U2&U3	U2&U4	U2&U5	U2&U6	U3&U4	U3&U5	U3&U6	U4&U5	U4&U6	U5&U6
S1			1	1									1		
S2						1	1			1					
S3															
S4	1	1		1		1		1			1				
S5															
S6													1		
S7															
S8			1												
S9															
S10				1	1										1
S11					1										
S12						1	1			1					
S13												1			
S14															
S15		1		1	1						1	1			1
S16						1	1	1	1	1	1	1	1	1	1
S17									1						
S18						1	1		1	1		1		1	
S19						1	1			1					
S20						1			1			1			
Sum	**1**	**2**	**2**	**4**	**3**	**7**	**5**	**2**	**4**	**5**	**3**	**5**	**3**	**2**	**3**

Source: formed by the authors

In Table 4 $U_i \& U_j$ defines a two-element set of two courses U_i and U_j that simultaneously chose (1) or did not chose (0) a certain student. Since in many two-element frequent itemsets F_2 we have representatives of all logical units, we can set the threshold of minimum support, which in conditions of this task we will accept, for example, equal to 3. Thus, a set of two courses will be considered popular if at least three students have chosen it.

Using frequent itemsets from F_2, we generate a set F_3 of three-element itemsets (Table 5). To do this, we need to associate a set F_2 with itself by selecting binding itemsets. Note, k-itemsets are binders if they have common *k-1* items. For example:

$$\{U_2 \& U_5\} + \{U_3 \& U_5\} = \{U_2 \& U_3 \& U_5\}$$

To reduce the search for associative rules, the Apriori algorithm uses the anti-monotony property: if *Z* is not a popular set, then adding some new subject *A* to the *Z* set does not make it more popular, meaning the set $Z \cup A$ will not be popular either. For example, the $U_2 \& U_3 \& U_5$ set will not be popular because the set $U_2 \& U_5$ is not popular.

Table 5: Formation of a set of three- and four-element frequent itemsets

	U1&U5&U6	U2&U3&U4	U2&U3&U6	U3&U4&U5	U3&U4&U6	U3&U5&U6	U2&U3&U4&U6
S1			1	1			0
S2						1	0
S3							0
S4	1	1		1		1	0
S5							0
S6							0
S7							0
S8			1				0
S9							0
S10				1	1		0
S11					1		0
S12						1	0
S13							0
S14							0
S15		1		1	1		0
S16						1	1
S17							0
S18						1	1
S19						1	0
S20						1	0
Sum	**2**	**5**	**3**	**1**	**2**	**2**	**2**

Source: formed by the authors

For a given set F_3, we form the set of possible four-element sets. According to the rule of formation of many popular sets, we have only one option – the set $U_2 \& U_3 \& U_4 \& U_6$.

This set is not frequent itemset according to the anti-monotony property (itemset $U_4 \& U_6$ is not frequent). The same is shown in the last column of Table 5, which is less than the minimum support threshold. So, the set $F_4 = \varnothing$.

A graphical model of the formation of sets F_1, F_2, F_3 is presented in Fig. 1

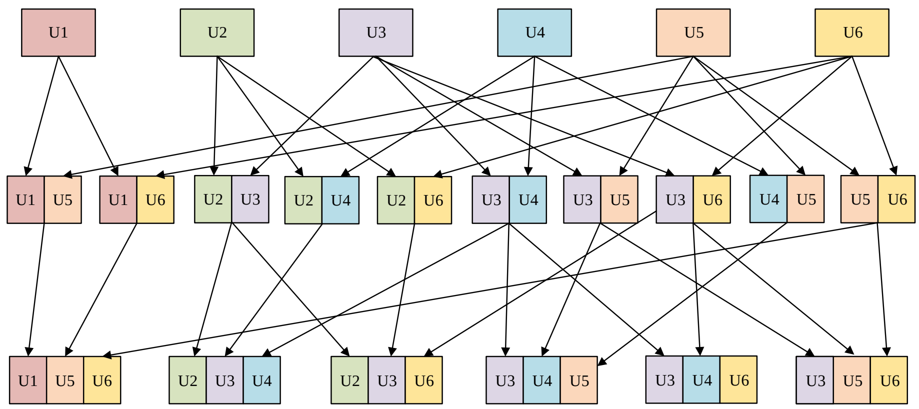

Fig.1. Graphical model of formation of sets of frequent itemset for different power. Source: formed by the authors

The next step in the Apriori algorithm is to generate associative rules. Each set of frequent itemset F_i requires a two-step procedure:

1. 1. All possible subsets of the set F_i are generated. We denote $\bar{F_i}$ the set of all subsets F_i.
2. If a subset f is a non-empty subset of $\bar{F_i}$, then the association. $R: f \to \left(\bar{F_i} \setminus f\right)$, where $\bar{F_i}$ is the same subset without f, is considered.

This procedure repeated for each subset f of $\bar{F_i}$.

The associative rule consists of two sets of objects called a condition (antecedent, left-hand side – LHS) and consequence (consequent, right-hand side – RHS), which are signed as $X \to Y$: «If *antecedent* then *consequent*». Association rules describe the relationship between item sets, which is characterized by two main indicators – support S (support) and C (confidence).

Associative rule support $S(A \to B)$ is the part of transactions that contain both antecedent and consequent. For example, for association $A \to B$ support $S(A \to B)$ means the fraction of the number of transactions containing antecedent A and consequent B to the total number of transactions.

Associative rule confidence $C(A \to B)$ is a measure of the accuracy of a rule $A \to B$ and is defined as the fraction of the number of transactions that simultaneously contain antecedent A and consequent B to the number of transactions that contain only antecedent A.

We form a set of associative rules for the set F_3.

For set $\{U_2 \& U_3 \& U_4\}$ $\bar{F}_3 = \{U_2, U_3, U_4, U_2 \& U_3, U_2 \& U_4, U_3 \& U_4\}$.

For set $\{U_2 \& U_3 \& U_6\}$ $\bar{F}_3 = \{U_2, U_3, U_6, U_2 \& U_3, U_2 \& U_6, U_3 \& U_6\}$.

Because support and confidence give different evaluations of the quality of an associative rule, it is often the product $S \cdot C$ of these two quantitative characteristics that are used to rank associative rules by priority (Table 8).

Table 8: Ranking of associative rules in terms of support and confidence

Rule	S	C	$S \cdot C$	Rule	S	C	$S \cdot C$
U2→U3	35.0%	77.8%	0.2722	U2&U6→U3	15.0%	75.0%	0.1125
U3→U2	35.0%	77.8%	0.2722	U1→U6	15.0%	71.4%	0.1071
U2&U4→U3	25.0%	100.0%	0.2500	U6→U2	20.0%	50.0%	0.1000
U3&U4→U2	25.0%	100.0%	0.2500	U6→U1	15.0%	62.5%	0.0938
U2&U3→U4	25.0%	71.4%	0.1786	U3&U6→U2	15.0%	60.0%	0.0900
U5→U1	20.0%	83.3%	0.1667	U2→U6	20.0%	44.4%	0.0889
U6→U3	25.0%	62.5%	0.1563	U5→U3	15.0%	50.0%	0.0750
U1→U5	20.0%	71.4%	0.1429	U5→U4	15.0%	50.0%	0.0750
U2→U4	25.0%	55.6%	0.1389	U5→U6	15.0%	50.0%	0.0750
U4→U2	25.0%	55.6%	0.1389	U2&U3→U6	15.0%	42.9%	0.0643
U3→U4	25.0%	55.6%	0.1389	U6→U5	15.0%	37.5%	0.0563
U4→U3	25.0%	55.6%	0.1389	U3→U5	15.0%	33.3%	0.0500
U3→U6	25.0%	55.6%	0.1389	U4→U5	15.0%	33.3%	0.0500

Source: formed by the authors

Thus, the Apriori algorithm has found associative rules that show which sample courses from the initial set of transactions are most often selected by students together. Such indicators will allow to build logical links between different units of courses and to develop a better strategy for managing the educational process of the HEI.

5.3. Generate Associative Rules in the Deductor Studio Analytics Platform

When we formed the set F_1 of single-element frequent itemsets, we change the traditional scheme of the Apriori algorithm and set the task of identifying links between logical units. This approach is primarily related to the complexity of analyzing associative rules for all courses. In the case of a large number of transactions and a large number of courses, it is appropriate to use associative rule generation and processing software.

Such popular systems as Microsoft Power BI [27], R [28], RapidMiner [29], Deductor [30] and others have Associative rule tools. Next, we briefly describe the process of processing associative rules in the analytical platform Deductor Studio [31].

Deductor Studio uses a special unit «Associative rules» that implements the Apriori algorithm to solve such problems. When configuring the unit, it is possible to set minimum and maximum values of support and confidence of associative rules, calculate additional characteristics of the importance of rules (lift, leverage, improvement), create various visualizers (diagram, rule tree, OLAP cube).

The results of the survey of 100 students are processed. The Deductor system allowed the formation of 1,123 frequent itemsets with a capacity of up to five courses (Fig. 2) and found 8,594 associative rules (Fig. 3).

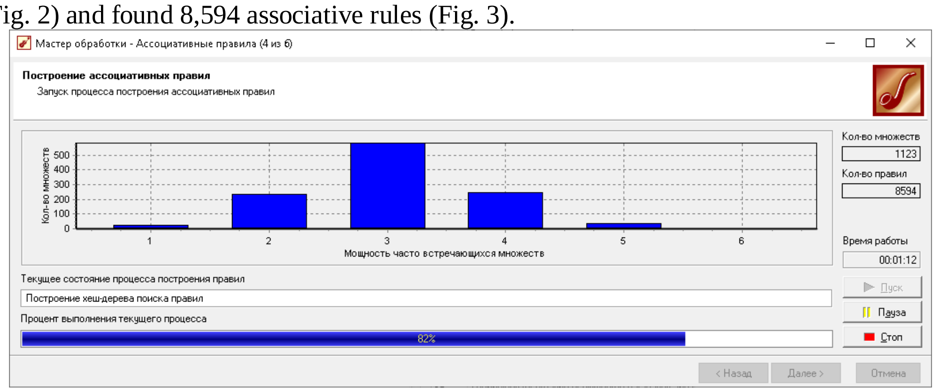

Fig.2. Formation of frequent itemsets in Deductor Studio. Source: formed by the authors

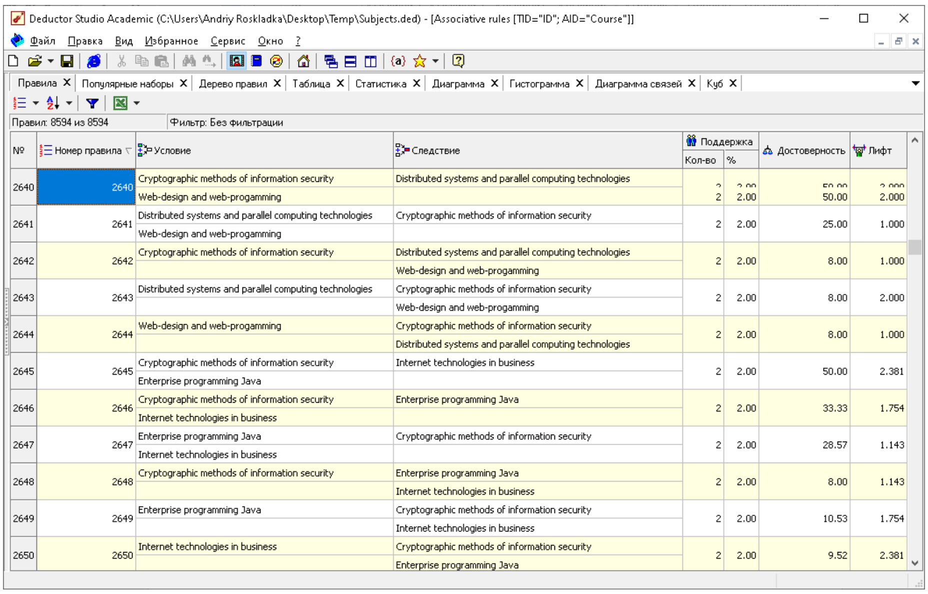

Fig.3. The result of the formation of associative rules in Deductor Studio. Source: formed by the authors

6. Conclusion

The formation of an individual educational trajectory is one of the main conditions identified in the guidelines for quality assurance in the European Space of Higher Education (ESG). A large number of courses creates significant difficulties in managing the choice process and often leads to situational management of the educational process in the HEI. The existing automation systems implement the complex management of a higher education institution and do not contain a module for automating the process of choosing academic courses.

As a systematic approach to manage the process of an individual student curriculum forming, we propose an algorithm based on the use of associative rules.

As a result of the Apriori algorithm using, we identified the associative rules of communication between educational courses chosen by the students together. The manual mode processing of 20 transactions revealed 18 popular course sets and 26 associated associative rules. The application of the Deductor system to process 100 transactions allowed the formation of 1,123 frequent itemsets and the identification of 8,594 associative rules. The number of rules that should be used in the management of the learning process depends on the minimum support threshold set, the confidence or their production.

The use of associative rules makes it possible to build logical links between different units of courses, to form an individual educational trajectory of the student and to develop a better strategy for managing the educational process of the HEI. It helps students to enhance their backround and knowledge by giving them the solid logical system of courses. This approach helps the HEI's management to be objective in the funds' allocation and staff management.

Obviously, the proposed approach requires further software development, which can be more effectively implemented using business intelligence tools such as Python, R, Microsoft Power BI.

References

1. The Law of Ukraine «On Education», 05.09.2017 № 2145-VIII // Supreme Council Information, 2017, № 38-39, art.380 (2019). https://zakon.rada.gov.ua/laws/show/2145-19, last accessed 2020/03/29.
2. The Law of Ukraine «On Higher Education», 01.07.2014 № 1556-VII // Supreme Council Information, 2014, № 37-38, art.2004 (2019). https://zakon.rada.gov.ua/laws/show/1556-18, last accessed 2020/03/29.
3. Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG) (2019). https://enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf, last accessed 2020/03/29.
4. Guidelines for experts of the National Agency on the application of the Educational Program Quality Assessment Criteria (2019). https://naqa.gov.ua/wp-

content/uploads/2019/09/Методичні-рекомендації_для-експертів.pdf, last accessed 2020/03/29.

5. Roskladka, A., Roskladka, N., Kharlamova, G., Karpuk, A. & Stavytskyy, A. The data science tools for research of emigration processes in Ukraine. Problems and Perspectives in Management Volume 18, issue #1, 70-81 (2020).
6. The national strategy of education development in Ukraine for 2012-2021(2019). http://zakon3.rada.gov.ua/laws/show/344/2013, last accessed 2020/03/29.
7. Krasnoshchok I. Individual educational trajectory of the student: theoretical aspects of the organization. Pedagogy of formation of creative personality in higher and secondary schools, 60 (1), 101-107 (2018).
8. Sharov S., Sharova T. Formation of the individual educational trajectory of the student employing information system. Scientific Bulletin of Melitopol State Pedagogical University, 2(19), 149-154 (2017).
9. Theoretical and methodological foundations of education informatization and practical implementation of information and communication technologies in the educational sphere of Ukraine: monograph. Edited by V. Bykov, S. Litvinova and V. Lugovoi, Kyiv, Komprint, 214 (2019).
10. Kuzminska O., Mazorchuk M., Morze N., Kobylin O. Digital Learning Environment of Ukrainian Universities: The Main Components to Influence the Competence of Students and Teachers. International Conference on Information and Communication Technologies in Education, Research, and Industrial Applications. Springer, Cham , 210-230 (2019.
11. Morze N., Umryk M. Formation of Computing and Coding Competences of Computer Science Teachers in Ukraine Universities in the Networked Society. Springer, Cham, 49-69 (2019).
12. Bogdanov A., Chepovoy I., Ukhan P., Yurchuk L. The system of information support of the process of formation of the individual educational trajectory of students. Information technology and training tools, 54 (4), 119-138 (2016).
13. Bykov V., Biloshchytskyi A., Kuchanskyi O., Andrashko Yu., Dikhtiarenko O., Budnik S. Development of information technology for complex evaluation of higher education institutions. Information Technologies and Learning Tools, 73, 5 (2019).
14. Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data /EMC Education Services. Indianapolis: John Wiley & Sons, Inc, 432 (2015).
15. Pang-Ning T., Steinbach M., Karpatne A., Kumar V. Introduction to Data Mining (Second Edition). Pearson, 840 (2019).
16. Zhang C., Zhang S. Association Rule Mining: Models and Algorithms. Berlin, Springer, 238 (2002).
17. Bhavani G., Sivakumari S. Survey on Association Rule Hiding Techniques. International Journal of Scientific Research in Science, Engineering and Technology, 6(3), 300-305 (2019).
18. School Management System Software. https://www.myeducomm.com, last accessed 2020/05/15.
19. Eiffel Corporation. https://www.eiffelcorp.co.za, last accessed 2020/05/15.
20. CyberVision. http://cybervision.com.pk, last accessed 2020/05/15.
21. Automated control system of the higher educational institution ACS "University". https://vuz.osvita.net/ua/, last accessed 2020/05/15.
22. Private enterprise computer systems package "Politek-soft". http://www.politek-soft.kiev.ua/, last accessed 2020/05/15.
23. Program Complex "ALMA-MATER". http://direct-it.com.ua, last accessed 2020/05/15.

24. Kravets A., Al-Shaebi R. Iindividualized student training trajectory formation process computerized control. Modern problems of science and education. Volume 2 (2012). http:// www.science-education.ru/ru/article/view?id=5952, last accessed 2020/05/15.
25. Kardan A. Sadeghi H., Shiri-Ghidary S. &Fani Sani M. Prediction of student course selection in online higher education institutes using neural network. Computers & Education. Volume 65, 1-11 (2013)
26. Ognjanovic I., Gasevic D. &Dawson S. Using institutional data to predict student course selections in higher education. The Internet and Higher Education. Volume 29,49-62 (2016).
27. Microsoft Power BI Cookbook: Creating Business Intelligence Solutions of Analytical Data Models, Reports, and Dashboards. Birmingham: Packt Publishing Ltd, 802 (2017).
28. Shmueli G., Bruce P., Yahav I., Patel N., Lichtendahl K. Data Mining foR Business Analytics: Concepts, Techniques, and Applications in R. – New Jersey: Willey, 544 (2018).
29. RapidMiner: Data Mining Use Cases and Business Analytics Applications / Edited by-Markus Hofmann&Ralf Klinkenberg. Minneapolis : CRC Press, 518 (2004).
30. Deductor Studio (2019). https://basegroup.ru/deductor/description, last accessed 2020/03/29.
31. Stavytskyy, A., Dluhopolskyi, O., Kharlamova, G., Karpuk, A. and Osetskyi, V.: Testing the fruitfulness of the institutional environment for the development of innovative-entrepreneurial universities in Ukraine. Problems and Perspectives in Management, 17(4), 274-288. (2019). doi:10.21511/ppm.17(4).2019