Support for the personalization of master's studies based on a fuzzy competency model (on the example of data analysis disciplines)

T V Gaibova
1 Department of Business Informatics, Financial University under the Government of the Russian Federation, 49 Leningradsky Prospekt, 125993, Moscow, Russia
E-mail: tvgaibova@fa.ru

Abstract. Competence models for evaluating important teaching tools for personalizing educational content. This problem is of particular importance for studying in the master's program, the teacher of each discipline needs the diversity of the educational trajectories of the master's students and create an adaptive educational environment for both those who deepen their profiling, continuing their education in the previously chosen direction of training, and for those who have changed the vector of professional development. The aim is to demonstrate a fuzzy model for assessing competence using the example of this organization of training in data analysis. It is developed on the basis of the integration of cognitive modeling and fuzzy logic and is distinguished by the ability to analyze the mutual influence of the concepts of the considered area of knowledge of the target level of learning and is the basis for determining the area of acceptable educational trajectories. The implementation of the proposed model will provide high-quality personalization of semi-structured educational content. The author analyzed the modern skill sets of specialists in the field of data science and analytics, as well as methods and tools for competence assessment in the context of learning adaptation. The requirements for the developed concept of soft learning in the context of personalization of learning in the master's program are determined, the structure of a fuzzy cognitive map is proposed, which allows one to assess the semi-structured competencies of various groups and determine the priority of mastering the concept of the proposed educational content. The proposed approach will make it possible to systematize the process of identifying permissible educational trajectories at the discipline level, depending on the available knowledge and accumulated professional experience, as well as on the student's career aspirations, thereby expanding the range of possible educational strategies and increasing the level of organization of the educational process on the part of the teacher and the level of self-organization on the part student.

1. Introduction
The global trend of modern education is adaptive management and personalization of learning. This requires the creation of effective feedback and motivation tools - competence models for assessing students' knowledge, as well as mechanisms for using these models in the educational process. This problem is considered by a number of researchers as a problem of educational design [1]. Without setting up a high-quality measuring tool, it is not possible to organize the subsequent adaptation of training and the involvement of students in the educational process. This problem is of particular importance for the organization of training in the magistracy, when the teacher of each discipline must take into account the heterogeneity of the possible educational trajectories of undergraduates and create an adaptive
educational environment for both those who deepen their profiling, continuing their education in the previously chosen direction of training, and for those who changed vector of professional development. Education in the master's program is characterized by a high degree of independence of students, as well as a greater certainty of the goals of their professional development and motivation, therefore, the result of a particular student's training directly depends on the correct choice of an individual learning path, the effectiveness of the proposed organizational forms of acquiring knowledge (especially self-study), support and interaction.

Today, the problem of training specialists who own modern technologies of data science and analytics (DSA) is very urgent. Despite the fact that the potential of data analysis to create new knowledge and develop various spheres of society is recognized by various experts [2], companies lack qualified data science specialists [3]. Education and training institutions are often unable to meet the changing needs of employers, holding back growth in several sectors. DSA skills are in high demand, but supply is critically low and employers face significant challenges.

Obviously, these trends will intensify - the volume of data for analysis is constantly increasing, the performance of computers is growing significantly, algorithms are becoming more complex and efficient, not only data science tools are regularly updated, but also the possibilities of their use.

There is a growing demand from employers not only for data analysts directly, but also for workers in other fields who use data analysis to support decision-making in the performance of their job duties [3]. This is a fairly wide range of areas of training - public administration, manufacturing, financial services, marketing, healthcare, retail, service and others.

For the high-quality preparation of a specialist of the future, various areas of the master's program include disciplines related to data science and analytics in their curricula. But the gap between the requirements of the labor market and the level of training in higher education continues to remain large enough. Practitioners talk about the need to adjust educational programs in this area. For example, the authors of the study of industry trends and the state of data science [4] conducted an online survey of 2,360 people from more than 100 countries of different age groups, industries and job responsibilities. Almost 50% of those surveyed admitted that the key problem preventing them from demonstrating the full potential of data science is the lack of necessary skills: what is taught in universities and in courses often turns out to be ineffective in real work, so novice DSA specialists have to learn a lot on their own.

This article proposes a new approach to ensuring the personalization of educational content based on a fuzzy cognitive map, which will allow adapting the learning process of data analysis specialists in the master's program to different educational levels and practical experience of students, as well as to the modern requirements of potential employers.

2. Competence of data analysts

Since the proposed approach is based on a competency model, the author analyzed the modern requirements for the competence of data analysts described in the literature. It should be noted that the problem of assessing the competence of an analyst is important both for the academic community and for potential employers and future employees. Competencies help you pinpoint the skills you need and can help you learn to develop those skills.

Many researchers have participated in the formation of both terminology and the main objectives of data science and analytics, as well as professional knowledge and skills. A detailed overview of the main trends in the development of ideas about DSA competence is given in [5]. In [6-8], the concept of "computational thinking" is proposed and explored to describe the basic skills that data scientists will need. Computational thinking capabilities include data collection, presentation and analysis, abstraction, automation, model testing and validation, algorithms and procedures, problem decomposition, control structures, parallelization and modeling.

The required technical skills are detailed - mathematics and statistics, machine learning, predictive analytics, computer science and programming, in addition to basic information systems, databases, data warehouses and data mining, as well as the ability to develop innovative algorithms for data analysis [9]. These same skills are most often mentioned in employers' ads [10]. You also need to be able to work...
with data sources of various structures [11]. Since the arsenal of tools and software for data analysis is quite extensive and continues to grow, technical skills should help analysts choose the most effective models and algorithms or their combinations for solving a specific business problem. Requirements for ethical use and data security are being strengthened.

It is also noted that data analysts must have an understanding of the business or subject areas for which data analysis is applied — that is, they must be able to view business problems from a data perspective [12-14]. Many sources conclude that subject competence may be more significant than deep technical competence alone [15]. Competencies of this type require not only deep knowledge of the subject area and setting of data analysis tasks, but also systems thinking. Technologies of fast implementation in the considered subject area are often required.

An important role in the competence of data analysts belongs to the competence of communication - the ability to verbally and visually present the results of analysis to various categories of stakeholders [16-17]. Such competencies are formed mainly using group interaction methods and active learning technologies.

The analysis showed that despite the small number of studies of procedures for assessing the competence of data analysts, the problems associated with the multi-parameter assessment have not been resolved - there is no information about what is the preferred sequence of competencies formation, how this affects the learning outcome, how the competencies of different groups are related. In [18], a map of student competencies in data science and analytics is proposed. Competencies are grouped into four levels - Personal Effectiveness Competencies, Academic Competencies, Workplace Competencies, Industry-Wide Technical Competencies. The opinions of more than 50 DSA experts from business, government and higher education were used to determine the academic-level competencies. The complex nature of the map and the interrelated nature of competencies are emphasized. But the model does not reflect the sequence of competencies development or their significance.

The analysis of modern professional requirements for data analysts led to the following conclusions:

- To measure the competence of a data analyst, several groups of parameters (sets of competencies) are required.
- Subject competencies and soft skills are of a semi-structured nature, therefore they will require fuzzy metrics to measure.
- To support decision-making on the personalization of educational content in accordance with the learning objectives and the current level of competence, knowledge is required about the relationships and mutual influence of competencies both within different groups of competencies and between groups, as well as their impact on the final indicator of professional competence of the data analyst.

Paragraph 3 provides an overview of research methods described in the literature related to fuzzy measurements in decision-making in dynamic systems, as well as the application of these methods in the context of adaptation of the educational process.

3. Research methods and Related Works

Fuzzy sets and fuzzy logic are successfully used to solve problems with weakly formalized and unreliable (incomplete, inaccurate) initial data.

The strengths of this approach are the description of the formulation and solution of the problem in a language close to natural, as well as its versatility - the possibility of approximation using fuzzy logic of any mathematical system. Currently, among the most popular applications of fuzzy logic, one can single out the assessment, classification and analysis of data, inference under uncertainty, decision support.

The review was carried out among scientific studies related to the problem of adaptive management and personalization of the learning process based on the results of a student's competence assessment without a relative level of automation of this process. A significant number of studies are devoted to
specific aspects of building online learning systems. There are several options for using fuzzy logic for this area - generating the adaptive learning trajectory [19] and forecasting the learning style in online systems [20-21], modeling the student profile [22], developing assessment tests [23] and other tasks of diagnosing students' knowledge [24-26].

In [27], it is shown that the use of fuzzy logic is one of the important ways to improve the performance of an adaptive learning system while supporting decisions about the proposed learning materials. This is due to the fact that it allows solutions to the problem of uncertainty associated with the learner's characteristics' fuzziness.

The complex nature of educational phenomena is one of the main reasons for using artificial intelligence tools to ensure adaptation of learning. The most common option is fuzzy inference systems that combine artificial neural networks and fuzzy logic, for example, using the method ANFIS (Adaptive Neuro-Fuzzy Inference System) [28]. ANFIS network training involves mapping input data through input membership functions and mapping output through output membership functions. The parameters associated with each membership function will change during the learning process.

Another area of work related to this study is the use of fuzzy cognitive mapping in education. Fuzzy cognitive maps (FCM) are a combination of neural networks and fuzzy logic that predict the change of concepts and are presented as a network model. They are fuzzy feedback graphs consisting of various nodes and directed arcs that reflect the causal relationship between these nodes. Fuzzy Cognitive Mapping (FCM) has been applied in a variety of theoretical and applied contexts spanning hard and soft sciences [29] to define a range of inferences about the belief systems of individuals and communities. This approach is recognized as effective in modeling systems with simple connections. A number of studies have shown that when modeling real dynamic systems consisting of a large set of concepts that change over time, the FCM approach does not adequately describe the hidden relationships between concepts. In such cases, it is advisable to use the Rules-based Fuzzy Cognitive Maps (RBFCM) approach, which extends the capabilities of the traditional FCM approach. In [30], evidence of the usefulness of Rule Based FCM for the formation of decisions in education is presented, as well as an approach to decision making on the adaptation of the proposed lectures, implemented within the framework of an intelligent adaptive web educational system. As shown in [31], Rule Based FCM has two important differences from traditional concept maps - the ability to represent vague or weak knowledge using fuzzy logic, as well as the detection of hidden relationships between concepts. The authors also presented the general structure of fuzzy cognitive mapping in relation to the learning process.

Since this study examines the hypothesis of the influence of the composition and structure of educational content on the learning outcome, I was primarily interested in the issue of taking into account the content aspects of the studied disciplines in the proposed fuzzy models.

It was found that despite a fairly large number of studies in which fuzzy numbers, fuzzy sets, fuzzy rules and fuzzy inference systems are used for various educational assessment and adaptation systems, the model inputs characterize mainly the organizational aspects of learning, and the content is not sufficiently covered. Also promising for research is the question of the mutual influence of the concepts of educational material and the determination of the best sequence of their development within the framework of personalization of training.

4. Results and discussion
This article presents the following results of the study:

- The concept of assessing the competence of the DSA in training.
- Rules-based Fuzzy Cognitive Map for assessing DSA competence.

4.1. Assessment concept
To describe the concept of assessing the competence of a data analyst, it is necessary to define the scope and type of task for which the meter should be designed.
The proposed approach to support the personalization of learning data analysis is one of the tools for solving the problem of adaptive learning management.

This is the task:
- dynamic - since it is necessary to track changes in the level of the learner's competencies and accordingly adjust the proposed educational content within the duration of training in the discipline;
- multi-criteria - since the concept of DSA competence includes several sets of competencies;
- semi-structured - since it is not completely amenable to detailed quantitative analysis, especially when considering the process of forming the subject component of DSA competence, as well as soft skills;
- which requires a solution in conditions of incompleteness and inaccuracy of information - since the feedback procedures in the learning process are still imperfect, despite the active use of modern information technologies and do not always allow obtaining unambiguous information about the current state of the student.

The developed meter should provide the ability to:
- Taking into account heterogeneous competencies.
- Research of the potential professional competence of a student in conditions of a lack of information.
- Adaptation to different stages of training within the discipline.
- Search for ways to adjust educational content at the level of meaningful concepts.
- The visibility and information content of the assessment results for both the teacher and the student.

It is also necessary to provide for the possibility of adapting the meter for various master's programs, which directly affects the priority and content of various sets of DSA competencies.

According to the type of the problem under consideration, the Rules-based Fuzzy Cognitive Maps approach was chosen as a tool for its solution, allowing, using sets of rules, to describe the hidden relationships between the identified concepts and factors, as well as to visually present the results of the measurements and assess the competence of the data analyst for all participants educational process, facilitating mutual understanding and communication in the formation of trajectories for further professional development of students.

From the point of view of organizing effective group work of students with different levels of training, the developed fuzzy cognitive map is a tool for aligning knowledge.

For this, the period of mastering the discipline is divided into 3 phases:
- formation of the knowledge base (15-20% of the duration of study) - the possibility of primary alignment of students with different levels of initial training;
- basic training (60-70% of the duration of training) - second leveling;
- in-depth training (15-20% of the duration of training) - from the point of view of leveling, the purpose of this phase is to determine further trajectories of professional development and apply the acquired knowledge and skills to solve practical problems, not educational, but a real level of complexity.

According to the proposed measurement technology, at the beginning of training and at the end of each phase, it is necessary to assess the current professional competence of $PC_{DSA}$ based on Fuzzy Cognitive Maps.

Let us introduce the designations of the measurement results at each phase:
- $PC_{DSA}(0)$ - to assess the level of knowledge and experience at the beginning of training. At this stage, it is most especially important to identify the problem areas of students who have changed the vector of professional development when choosing a master's program. Analysis of the
assessment results will allow us to offer them in a timely manner the forms and methods of training that will allow them to quickly enter the area of knowledge under consideration and not feel outside the learning process. These can be both separate theoretical topics and practical tools. It is important to discuss the results obtained with each student, since an objective picture of the initial level of knowledge allows the teacher to form the correct goals for the current stage of training, estimate the necessary time resources and adjust motivation. As the proposed forms of education for this category of students, priority is given to online tools. It is equally important to assess the level of initial competence of students deepening profiling. The purpose of organizing training in this case is the systematization of existing knowledge of the mathematical level, as well as subject competence.

- $PC_{DSA}(1)$ - to determine the current level of knowledge and available experience based on the results of the assessment after the first stage - the formation of the knowledge base.
- $PC_{DSA}(2)$ – to determine the current level of knowledge and existing experience based on the results of the assessment after the second stage - basic training.
- $PC_{DSA}$ result – for the final assessment of professional competence after graduation.

The following question requires clarification: why is the RBFCM assessment proposed for individual stages of training instead of the apparently obvious option of mapping the entire learning process in the discipline under consideration? This is due to the complexity of the considered area of knowledge - data analysis, which includes many problem statements, as well as methods and tools for their solution.

Therefore, the procedure for assessing the current level of DSA competence, without taking into account the dynamics of learning - the previous state and mechanisms for adapting the educational content of the next phase of learning - is also complicated. When trying to take into account the dynamics and develop a map of the entire learning process, the concepts that form the learning strategy and educational content are unnecessarily generalized. This level of aggregation will not allow solving the task - to identify the problems of students and promptly offer assistance in eliminating them.

4.2. Rules-based Fuzzy Cognitive Maps to assess competence DSA

Within the framework of fuzzy inference terminology, the current professional competence of DSA is the target concept of Rules-based Fuzzy Cognitive Maps, which characterizes the level of training of a student to solve professional problems of data analysis in the subject area under consideration.

Based on the analysis of the above modern approaches to the formation of the competence profile of a data analyst, the structure of the proposed competency model was determined. The following levels are included:

- knowledge of the subject area;
- mathematical;
- technological;
- soft skills.

Table 1 presents the results of structuring the task of measuring the competence of the data analyst. The identification of concepts was carried out on the basis of comparing the content of the DSA competency levels with the content of the data analysis stages.

**Table 1.** Content of competency levels at different stages of data analysis.

| Selection of features | Collection, cleaning and enrichment of data | Solving the problem of data analysis | Semantic interpretation |
|-----------------------|--------------------------------------------|-------------------------------------|-------------------------|
| Knowledge of the subject area | Identifying domain problems from a data analysis perspective | Determination of sources of collection, verification and enrichment of data | Formation of hypotheses for testing based on the needs of the subject area | Determining the required thresholds for model quality metrics |
Math skills
Mathematical formulation of the problem
Analysis of distributions, methods of smoothing and filtering
Mathematical algorithms for solving problems
Model quality metrics, hypothesis testing

Technological skills
Choosing software for solving the problem
Using data preparation tools
Using Modeling Tools
Using Model Evaluation Tools

Soft skills
Documenting proposals for the selection of features and formalization of the task for discussion
Justification of the acceptability of the risk level of the assessment of the general population by sample parameters
Explanation of the technical concepts of data analysis for non-technical audience
Presentation of the analysis results to various categories of stakeholders

The goal formed in relation to the chosen concept is expressed by a set of private goals $PC_i, i=1...4$, reflecting, respectively, the level of formation of subject, mathematical, technical and soft skills when analyzing data. It should be noted that the particular goals are fully consistent with the target concept.

Each particular goal is associated with a particular criterion described by a fuzzy set defined on $X_i$, and for $\forall x_i \in X_i$, the value $\mu_{PC_i}(x)$ characterizes the degree of its attainability.

The structure of the proposed RBFCM and the principle of forming links at the level of competence assessment is shown in figure 1.

RBFCM uses linguistic fuzzy rules to describe relationships between concepts. Each particular goal is determined based on a set of fuzzy production rules regarding input factors (in figure 1 - BRPC1, BRPC2, BRPC3, BRPC4, respectively). To solve this problem, you can use known methods, for example, Mamdani. The fuzzy rule base for the target concept $PC_{DSA}(t)$ can be used both to determine the resulting assessment of the DSA competence at the t-stage of training, and to determine the strategy for achieving a separate value from the term set characterizing the general target function when solving the problem of adapting educational content. In this case, each rule should be an operation of convolution of private factors of competence, taking into account the requirements of the current stage of training.

The proposed FCM competency measurement can be adapted for different master's programs, implying different levels of DSA competency requirements. This is done by changing concepts, fuzzy rules that determine the value of concepts and, as a result, the weights of the cognitive map. For example, for students enrolled in data analysis in specialized master's programs, the concepts of building technological competence will include the skills of writing data analysis code using modern statistical software (R, Python and SAS), and the concepts of building mathematical skills will have increased value require a more advanced level of development. For students not specialized in the context of DSA directions, using data analysis as one of the tools for solving basic professional problems, it will be sufficient to use high-level technologies and data analysis platforms that do not require coding skills. At the same time, the level of mastering mathematical skills should be sufficient to formulate business problems of data analysis and carry out a semantic interpretation of the results.

5. Conclusion and future work
The competence of a data analyst includes a set of different competencies that are often difficult to quantify and have different priorities in development depending on the current competence of the student.
Cognitive maps and fuzzy logic make it possible to model weakly formalized subject areas to improve the quality of the planned result, create possible scenarios for the development of the situation. This article discusses the possibility of using fuzzy cognitive maps in assessing the potential professional competence of DSA for organizing the training of students with different initial training in the magistracy.

Further studies are planned with the aim of developing the RBFCM structure to the level of control and measuring materials concepts (tests, practice-oriented tasks of various levels of complexity), forming fuzzy rule bases for concepts of all levels and weights of connections, conducting expert verification of FCM according to the criteria for the presence of a normal form in the selected factors and the criterion of the absence of duplicate links, the development of an algorithm for assembling the proposed educational content based on the results of the current assessment of the DSA competence.

This will improve personalization in the mastering of disciplines related to data science and analytics by graduate students, taking into account their previous educational experience and real career aspirations, increase motivation, involvement in the educational process of students of different training levels, the effectiveness of group learning in the formation of subject competencies and soft skills, will ensure the transfer of the most important knowledge in the adaptation of the proposed educational content.

References
[1] Russo D 2016 Competency Measurement Model *European Conference on Quality in Official Statistics* 29
[2] Vinuesa R, Azizpour H, Leite I and al 2020 The role of artificial intelligence in achieving the Sustainable Development Goals Nat Commun 11 233
[3] APEC HRDWG 2017 Data Science and Analytics Skills Shortage: Equipping the APEC Workforce with the Competencies Demanded by Employers 17
[4] Anaconda Inc 2020 The State of Data Science 2020: Moving from Hype Toward Maturity
[5] Strengell T 2017 Competitiveness from Data and Analytics: Required competency in Organization 91
[6] Papert S 1996 An exploration in the space of mathematics educations International Journal of Computers for Mathematical Learning 1(1) 95-123
[7] Wing J M 2006 Computational thinking Communications of the ACM 49 33–35
[8] Barr V and Stephenson C 2011 Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community? ACM Inroads 2(1) 48-54
[9] Miller S 2014 Collaborative Approaches Needed to Close the Big Data Skills Gap Journal of Organization Design 3(1) 26-30
[10] Pinola T 2015 Defining data scientist competencies 53
[11] Patil D J 2011 Building data science teams 25
[12] Dhar V 2013 Data science and prediction Communications of the ACM 56(12) 64-73
[13] Provost F and Fawcett T 2013 Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking 383
[14] Waller M A and Fawcett S E 2013 Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management Journal of Business Logistics 34(2) 77-84
[15] Patil T H and Davenport D J 2012 Data Scientist: The Sexiest Job of the 21st Century Harvard Business Review 90(10) 70-76
[16] Davenport T H and Dyché J 2013 Big data in big companies International Institute for Analytics 26
[17] Van den Driest F, Sthanunathan S and Weed K (2016) Building an insights engine Harvard Business Review 94(9) 15
[18] BHEF 2016 Competency Map for the Data Science and Analytics-Enabled Graduate BHEF DSA Competency Map 18
[19] Mohamed F, Abdeslam J and Lahcen E B 2017 Personalization of learning activities within a virtual environment for training based on fuzzy logic theory International Conference on E-Learning 5
[20] Deborah L J, Sathiyaseelan R, Audithan S and Vijayakumar P 2015 Fuzzy-logic based learning style prediction in e-learning using web interface information Sadhana 40 379–394
[21] Ennouamani S, Mahani Z and Akharraz L 2020 A context-aware mobile learning system for adapting learning content and format of presentation: design, validation and evaluation Education and Information Technologies 25 3919-3955
[22] Xu D, Wang H and Su K 2002 Intelligent student profiling with fuzzy models Proceedings of the 35th Annual Hawaii International Conference on System Sciences 8
[23] Goyal M, Yadav D and Choubey A 2012 Fuzzy logic approach for adaptive test sheet generation in e-learning IEEE International Conference on Technology Enhanced Education ICTEE 1-4
[24] Azar A T 2010 Fuzzy Systems in Education: A More Reliable System for Student Evaluation Fuzzy Systems 216
[25] Goodarzi M H and Amiri A 2009 Evaluating Students' Learning Progress by Using Fuzzy Inference System Sixth International Conference on Fuzzy Systems and Knowledge 6 561-565
[26] Grigoriadou M, Kornilakis H, Papanikolau K A and Magoulas G D 2002 Fuzzy Inference for Student Diagnosis in Adaptive Educational Hypermedia Methods and Applications of Artificial Intelligence (SETN 2002) Lecture Notes in Computer Science 2308 191-202
[27] Ennouamani S and Mahani Z 2019 Towards Adaptive Learning Systems Based on Fuzzy-Logic Intelligent Computing 1 625-640
[28] Al-Hmouz A, Shen J, Al-Hmouz R and Yan J 2012 Modeling and Simulation of an Adaptive Neuro-Fuzzy Inference System (ANFIS) for Mobile Learning 2012 IEEE Transactions on learning technologies 5(3) 226-237

[29] Gray S A, Zanre E and Gray S 2014 Fuzzy cognitive maps as representations of mental models and group beliefs Fuzzy Cognitive Maps for Applied Sciences and Engineering 54 29-48

[30] Peña-Ayala A and Sossa-Azuela J H 2014 Decision Making by Rule-Based Fuzzy Cognitive Maps: An Approach to Implement Student-Centered Education Fuzzy Cognitive Maps for Applied Sciences and Engineering 54 107-124

[31] Gouda N 2014 Fuzzy Cognitive Mapping: Applications in Education Rule Based FCM: Application in Education 50