I.J. Modern Education and Computer Science, **2025, 1, 28-46**
Published Online on February 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2025.01.03

Complex of Specialized Methods of Educational Data Mining for the Training of Vocational Education Teachers

Oleksandr Derevyanchuk
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, 58012, Ukraine
E-mail: o.v.derevyanchuk@chnu.edu.ua
ORCID iD: https://orcid.org/0000-0002-3749-9998

Zhengbing Hu
School of Computer Science, Hubei University of Technology, Wuhan, China
E-mail: drzbhu@gmail.com
ORCID iD: https://orcid.org/0000-0002-6140-3351

Serhiy Balovsyak
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, 58012, Ukraine
E-mail: s.balovsyak@chnu.edu.ua
ORCID iD: https://orcid.org/0000-0002-3253-9006

Serhii Holub
Cherkasy State Technological University, Cherkasy, 18006, Ukraine
E-mail: s.holub@chdtu.edu.ua
ORCID iD: https://orcid.org/0000-0002-5523-6120

Hanna Kravchenko
High State Educational Establishment «Chernivtsi transport college», Chernivtsi, 58000, Ukraine
E-mail: hannakravchenko81@gmail.com
ORCID iD: https://orcid.org/0009-0004-7609-0345

Iryna Sapsai*
Institute of Postgraduate Education, Borys Grinchenko Kyiv Metropolitan University, Kyiv, 02152, Ukraine
E-mail: i.sapsai@kubg.edu.ua
ORCID iD: https://orcid.org/0000-0002-7338-715X
*Corresponding Author

Received: 23 August, 2024; Revised: 13 October, 2024; Accepted: 17 November, 2024; Published: 08 February, 2025

Abstract: In the work, an analysis of modern methods of Educational Data Mining (EDM) was carried out, on the basis of which a set of methods of EDM was developed for the training of vocational education teachers. The basic methods of EDM are considered, namely Prediction, Clustering, Relationship Mining, Distillation of Data for Human Judgment, Discovery with Models. The possibilities of using artificial neural networks, in particular, networks of Long-Short-Term Memory (LSTM), to predict the results of the educational process are described. The main methods of clustering and segmentation of educational data are considered. The basic methods of EDM are complemented by specialized methods of digital image pre-processing and methods of artificial intelligence, taking into account the peculiarities of the training of future specialists in engineering and pedagogical specialties. As specialized methods of digital image pre-processing, methods of filtering, contrast enhancement and contour selection are used. As specialized methods of artificial intelligence, methods of image segmentation, object detection on images, object detection using fuzzy logic were used. Methods of object detection on images using convolutional neural networks and using the Viola-Jones method are described. To process data with a certain degree of uncertainty, it is proposed to apply the methods of EDM and Fuzzy Logic in a integral manner. Ways of integrating Fuzzy Logic with methods of data clustering, image segmentation and object detection on images are considered. The possibilities of applying the developed complex of specialized methods of EDM in the educational process, in particular, when performing STEM (Science, Technology, Engineering and Mathematics) projects, are described.

Index Terms: Educational Data Mining, Vocational Education Teachers, Clustering, Digital Image Processing, Image Segmentation, Object Detection, Fuzzy Logic

1. Introduction

Educational Data Mining (EDM) is becoming an important tool for improving learning and teaching. The use of information technologies in educational systems leads to the generation of a significant amount of data that contains valuable information about students and the educational process. Automatic analysis of such data by means of EDM allows to significantly optimize the educational process. EDM is an interdisciplinary field [1, 2], which combines Data Mining (DM) with machine learning, psychometrics, computer modelling, statistics, data visualization and web mining. EDM is an example of the effective use of digital technologies in the educational process. Unlike Data Mining, EDM takes into account the peculiarities of educational data (interrelationships between data sets, hierarchical organization of data) and the tasks of their processing for the purpose of application in the educational process [3]. The basic methods of EDM that can be applied in all areas of the educational process include: Prediction, Clustering, Relationship Mining, Distillation of Data for Human Judgment, Discovery with Models [4, 5].

However, the training of future vocational education teachers has certain features related to their field of professional activity. Such features include the processing of a significant amount of data not only in text, but also in numerical and graphic form [6]. Future engineers need to correctly construct, visualize and analyze various graphs, charts, diagrams, drawings, three-dimensional (3D) models and more [7]. In modern technical systems, computer vision and machine learning systems are increasingly used, accordingly engineers must understand the principles of digital image processing and possess the skills of their practical implementation [8, 9].

Therefore, the purpose of the work, which consists in the development of a complex of specialized methods of EDM, taking into account the peculiarities of the training of future vocational education teachers, is relevant. According to the purpose of the work, the existing basic methods of EDM are supplemented with specialized methods of digital image pre-processing and methods of artificial intelligence. As specialized methods of digital image pre-processing, methods of filtering, contrast enhancement and contour selection were used. As specialized methods of artificial intelligence, methods of image segmentation, object detection on images, object detection using fuzzy logic were used.

A certain degree of vagueness is characteristic of part of the data processed by specialists in engineering and pedagogical specialties [10]. For example, when students are divided into clusters, there may be a significant overlap between them, that is, one student may belong to several clusters to a certain extent. Therefore, in order to increase the processing capabilities of such data, it is proposed to apply the methods of EDM and Fuzzy logic in a complex (integral) manner [11-13]. Ways of integrating fuzzy logic with methods of data clustering, image segmentation and object detection on images were considered, as a result of which a number of specialized EDM methods were obtained.

Thus, the article proposes a complex of specialized methods of Educational Data Mining, which is intended to improve the system for training future vocational education teachers in institutions of higher education for the complex application of methods of EDM in professional activities.

1.1 Outline

An overview of the existing EDM methods is provided in Section 2; methods of Prediction, Clustering, Relationship Mining, Distillation of Data for Human Judgment, Discovery with Models are described. The methodology of the work is presented in Section 3, which describes a set of EDM methods for training future vocational education teachers. Section 4 presents the developed complex of specialized methods of fuzzy clustering, digital image pre-processing, image segmentation, object detection on images, object detection using fuzzy logic. The results of the work are described in Section 5.

2. Literature Review

EDM is used in several key areas. The paper [4] highlights four key areas of EDM application: improvement of student models, improvement of domain models, study of pedagogical support due to educational software, conducting scientific research on the educational process and its participants. In another study [14], the following areas of application of EDM are highlighted: evaluation of students' educational achievements, adaptation of the course for individualization of students' learning based on their behavior, evaluation of educational materials in online courses, analysis of students' educational behavior. That is, the classification of EDM application areas [14] complements the classification [4] taking into account the features of e-learning.

There are several classifications of the basic methods of EDM, which are used in various areas of the educational process. Researchers [4, 5, 15] distinguish the following main methods of EDM, which combined into groups #1-5 (Fig. 1).

The considered methods of Prediction (group #1), Clustering (group #2) and Relationship Mining (group #3) are considered categories of DM. The methods of Distillation of Data for Human Judgment (group #4) and Discovery with Models (group #5) are the most specific for EDM. Other authors structure the methods of EDM somewhat differently. For example, researchers [16] introduce new methods in comparison with the classification [4]: 1) Outlier Detection; 2) Process Mining; 3) Social Network Analysis; 4) Text Mining.

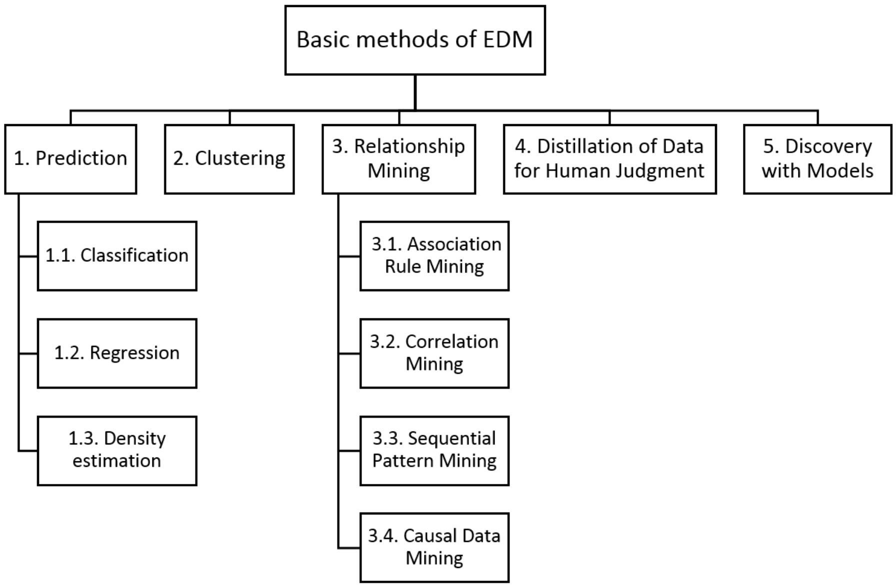


Fig. 1. Basic methods of EDM [4]

Outlier Detection consists of identifying data elements that are significantly different from other data. In EDM, the detection of outliers allows, in particular, to find students with learning outcomes or behavior characteristics that are significantly different from the average values. Detection of emissions can be performed by means of Correlation Mining, namely by method #3.2 (Fig. 1) [4].

Process Mining is related to the group of methods #3 Relationship mining (Fig. 1) [4].

Social Network Analysis specializes in obtaining data about social network users based on information posted by users, as well as based on connections in social networks.

Text Mining consists in the automatic analysis of the text in order to obtain information on a certain topic. This method uses other methods of EDM for data processing (for example, for text clustering).

The results of Social Network Analysis and Text Mining are used as input data for other methods of EDM. The authors [16] also indicate the need for pre-processing of educational data, which consists, for example, in the modification and deletion of confidential personal data. The work [17] highlights the following methods of EDM: Clustering, Classification, Outlier Detection, Association, Pattern Matching and Text Mining. The above methods repeat the main methods in the classifications [4] and [16]. Thus, the classifications of EDM methods [16] and [17] complement the classification [4], therefore, to describe the basic methods of EDM, we will use the widely known and quite detailed classification [4]. Let's consider in more detail the basic methods of EDM (Fig. 1).

Prediction methods (group of methods #1) allow to determine the potential results of the educational process based on existing data [15, 18, 19, 20]. Such methods can predict the academic performance of students or their difficulties in learning, as well as predict the effectiveness of the implementation of new educational methods. Predictions can take into account the individual characteristics of students, including their previous grades, learning style and other important factors. Forecasting allows educational institutions, teachers, and curriculum developers to predict future trends, outcomes, and student needs. This method uses historical and current data to model future events with a predetermined degree of accuracy. Forecasting in education is used to analyze data and identify patterns (trends) that can be used to predict future outcomes. The forecasting process typically involves collecting and pre-processing data, selecting an appropriate forecasting model, training the model on historical data, testing its accuracy on a validation data set, and finally using the model to predict future outcomes. Prediction in EDM is carried out by the methods of Classification, Regression, Density estimation. At the same time, the methods of Classification and Regression, along with Artificial Neural Networks (ANN), are integral parts of the direction of Machine Learning.

Forecasting by the Classification method is performed, for example, by a Naive Bayes classifier, the K-nearest neighbor method [18] or using machine learning [21]. Prediction of educational achievements of students based on their activity using machine learning (classification method according to the K-NN Algorithm) is described in [22].

Prediction using Regression Analysis is based on the fact that the dependence of the predicted values (dependent variables) from the initial parameters (independent variables) is mathematically described by a regression equation (for example, a polynomial) [10, 18]. The coefficients of the regression equation are calculated in most cases by the Least

Squares Method. Prediction of placement of college students based on academic performance using machine learning approaches (in particular, regression analysis) is described in [23].

In the method of Density Estimation [17] the predicted values of the variable are described by the Probability Density Function, which depends on the initial data. Certain kernel functions are used to estimate the density. Random variables in EDM are often described by a normal distribution, so in such cases a Gaussian function is used as a kernel function.

However, in modern EDM systems, in addition to the described forecasting methods, forecasting by Artificial Neural Networks (ANN) is often used [24, 25, 26, 27, 28]. The use of ANN to predict the results of the educational process is effective, because such results are often influenced by a significant quantity of factors (which are difficult to take into account by other methods). ANN is a field of Machine Learning. Important advantage of ANN is the ability to train on specially selected data sets. The known parameters of the educational process are applied to the inputs $X = (x_1, x_2,..., x_N)$ of the ANN, and the predicted values are obtained at the outputs $Y = (y_1, y_2,..., y_M)$. In the case of forecasting time dependences (series), data for the previous period are applied to the inputs of the ANN, and data for the next period are received at the outputs. In the process of ANN training, the training error ε is calculated by the difference between the network outputs Y and the true outputs Y^T. The learning error ε is minimized by changing the weights of the neurons. In most cases, ANN training is performed by the method of backpropagation. ANN capabilities increase when the quantity of hidden layers increases, but this leads to an increase in training time.

The work [29] considered the use of ANN for predicting the educational achievements of students. The prediction of students' academic performance was performed using a Deep Neural Network (DNN), which made it possible to automatically determine the behavioral characteristics of students. Long-Short-Term Memory (LSTM) networks have been applied to predict time dependencies for student behavior. A Two-Dimensional Convolutional Neural Network (2DCNN) was used to establish a correlation between different types of student behavior.

The use of Convolutional Neural Networks (CNN) for the recognition of educational images in order to identify at-risk students is considered in [9]. It is shown that one-channel and three-channel recognition of educational images by the proposed methods allows to more accurately identify students in the risk group, compared to other machine learning algorithms (for example, compared to support vector machine, random forest).

Prediction of the academic success of students in Chinese-foreign cooperation in management schools using a Graph Convolutional Network was performed in work [30]. An undirected graph was used to connect students with similar characteristics. Student grades from previous semesters were used as input data for the neural network. It is shown that the graph convolutional network provides higher prediction accuracy compared to the support vector machine and the random forest model. Prediction of educational courses using ANN is described in [31].

Therefore, the Prediction method using ANN (taking into account its popularity and effectiveness) was added to the list of basic methods used in the EDM (Fig. 2).

Clustering methods (group of methods #2, Fig. 1) are widely used for the analysis of educational data, as they ensure the division of the initial set of large data into parts (clusters) [32, 33]. In the clustering method, learning without a teacher is used. Clustering allows to structure and purposefully process training data, identify patterns and connections useful for practice. In addition, sequential processing of objects of individual clusters is much simpler than processing all the studied data. In most studies, clustering is performed in the space of two parameters. Clustering is the process of grouping objects in such a way that objects in one group (cluster) have more characteristics in common with each other than with objects in other groups. In the context of EDM, clustering is used to identify groups of students with similar learning needs, behaviors or achievements, allowing for the development of targeted learning strategies.

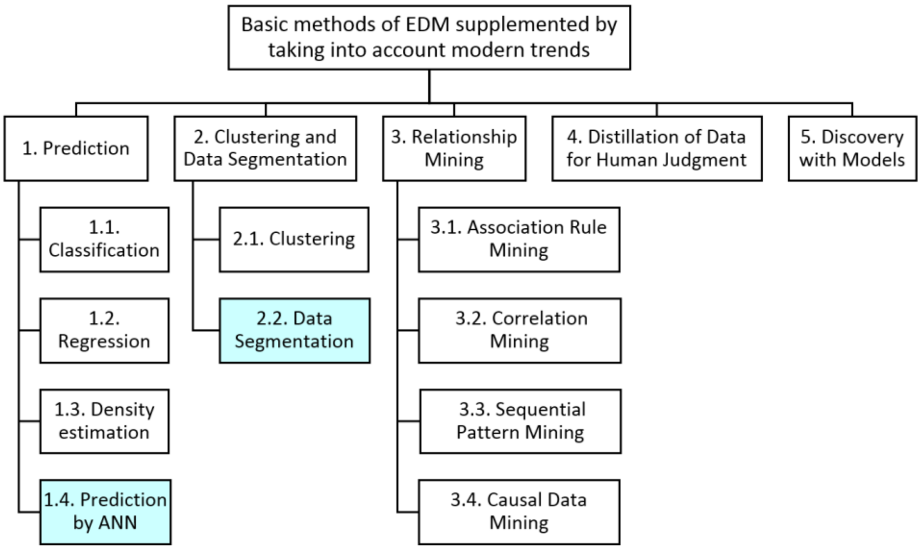


Fig. 2. Basic methods of EDM [4], supplemented by taking into account modern trends; added methods are shown in color

Due to clustering, individualized learning, identification of students with learning difficulties, identification of groups of students with similar interests or needs for targeted development or adaptation of educational materials is ensured. Clustering helps group students by similar characteristics or behaviors, facilitating the development of personalized learning approaches. Based on the analysis of the results of clustering, recommendations for training are provided both in the case of face-to-face and distance learning. Cluster analysis of intermediate and final learning results allows adjusting the level of difficulty of learning tasks. In the field of education, clustering is used to identify student behavior patterns and take appropriate actions to optimize the learning process. Cluster analysis can be used to assess the complexity of test tasks (identifying a group of the most difficult tasks).

In particular, such clustering methods as the K-means method, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) and hierarchical clustering are used [34]. The most common is the classical K-means clustering method, which consists in iterative division of objects into *K* clusters [35]. The number of clusters is selected by the user or determined by the minimum ratio of intra-cluster to inter-cluster distances. The DBSCAN method is efficient when processing compact clusters and noisy data. Hierarchical clustering builds nested clusters by sequentially merging or splitting clusters based on similarity measures.

The possibilities of cluster analysis for Student Behavioral Patterns were demonstrated in the work [36]. For the purpose of deeper analysis, the same educational data can be divided into clusters using different methods, for example, DBSCAN and k-means. Examples of evaluating student learning outcomes using the DBSCAN clustering method in modern educational platforms are given in [37]. Possibilities of correction of student learning in electronic systems using clustering methods are described in [38]. Work [39] shows that clustering allows for effective grouping of educational data objects even if they are connected by several complex relationships.

When processing educational data, they can be divided into clusters and data segments [40]. A feature of data segmentation is the preliminary establishment of the number and characteristics of segments; the criteria by which data elements are distributed between different segments are indicated. Setting criteria for data segmentation is usually performed by a person. An example of segmentation is the grouping of students by group number, specialty, level of education, etc. Segmentation of educational data is carried out in order to better understand and address the individual needs of students. Therefore, in order to expand the possibilities of processing educational data, it is proposed to name the group of methods #2 "Clustering" as "Clustering and Data Segmentation" and to supplement it with the method "2.2. Data segmentation" (Fig. 2).

The purpose of relationship mining (group of methods #3, Fig. 1) is to identify relationships and associations between characteristics of objects in data sets [15, 17]. Such an analysis allows to identify the variables that are most strongly related to each other. The relationships found through relationship mining must satisfy the criteria of statistical significance and utility. In the context of EDM, the resulting relationships provide insight into how different factors (for example, participating in group projects) affect learning outcomes. The following methods of relationship mining are distinguished: Association Rule Mining, Correlation Mining, Sequential Pattern Mining, Causal Data Mining.

The Association Rule Mining consists in the fact that the relations between educational data are described by the rules "if...then...". For example, if average score of the student is low and the student has a job, the student is more likely to drop out. Associative rules allow to establish how one event leads to another, and to identify characteristic sets of data or groups of items that often appear together in transactions.

Correlation Mining allows to establish the degree of relationship (positive or negative) between two variables. Due to this, it is possible to identify the characteristics of objects that are most strongly related to each other. In educational data, such analysis can reveal, for example, the relationship between the amount of time spent on independent study and student grades.

Sequential Pattern Mining makes it possible to discover the sequence in which students usually study certain topics or perform tasks. Due to this, the correct sequence of actions is established, for example, the order of execution and control of the stages of STEM projects and other types of educational activities.

Causal Data Mining is used to find causal relationships between events in the educational process. Such analysis leads to improved curricula, individualized learning, effective allocation of resources, and increased student motivation and engagement.

Methods of Distillation of Data for Human Judgment (group of methods #4, Fig. 1) are designed to transform complex data into forms that are easily interpreted by people, which contributes to quick understanding of data and decision-making [5, 15]. Data presentation uses visualization (Visual Mining), in particular methods of Infographics and Image Processing. Transformation of data into easy-to-understand forms is performed, in particular, with the use of Pivot Tables and dimensionality reduction algorithms. Methods of Distillation of Data for Human Judgment are particularly effective in the educational process, as they simplify the perception of even large volumes of data.

Data Visualization is the use of graphs, charts, images, diagrams, maps, three-dimensional (3D) models, and other visual tools to present data in an understandable form. Visualization helps to quickly identify regularities, trends, patterns and outliers (anomalies) in the data. As a preliminary stage of visualization, image segmentation can be used, which allows to automatically divide the image into meaningful areas and focus students' attention on such areas. In the educational process, 3D data visualization is quite effective, which helps to quickly stimulate interest in learning among students, better assimilation of the material, and focus attention on key areas.

Pivot Tables are consolidated reports that summarize large volumes of data in a form that is convenient for analysis. This allows users to see key metrics and indicators at a high (summarized) level.

Infographic is a combination of a visual (graphic) form of information presentation with a textual one, which is used to present complex data or ideas in an easily digestible form.

Discovery with Models (group of methods #5, Fig. 1) uses created models to generate new knowledge about learning processes, student behavior, and the effectiveness of learning strategies, which can be applied to improve the educational process [5, 15]. This approach may include modeling learning trajectories, designing adaptive learning systems, or predicting the outcomes of new learning strategies. Model discovery includes a wide range of techniques, including clustering and prediction.

One of the areas of Discovery with Models is the Simulation of Educational Processes by developing computer models of the system that simulate the interaction between students and the educational environment. Such modeling allows analyzing the influence of various factors on the educational process. For example, the development of a logical model of a system can be performed by Fuzzy Cognitive Map (FCM) and the Mental Modeler package.

Another area of Discovery with Models is the Optimization of Learning Strategies, which consists in using models to assess potential effectiveness and optimize learning approaches. For example, models for optimizing the learning style of students can be obtained as a result of testing.

3. Methodology

The analysis of scientific articles showed that the effectiveness of the educational process can be significantly increased through the use of modern methods of EDM [1, 2]. At the same time, methods for training future vocational education teachers need visualization and computer processing skills of graphs, diagrams, schemes, drawings, images, three-dimensional (3D) models, as well as knowledge of Intelligent Image Analysis [6, 7, 9]. Methods of Distillation of Data for Human Judgment in classification [4] (Fig. 2) are mainly focused on data visualization. Therefore, the following structure is proposed for group of methods #4 (Distillation of Data for Human Judgment) (Fig. 3):

4.1. Data visualization.
4.2. Digital Image Pre-processing.
4.3. Image Segmentation.
4.4. Object Detection in Images.

Specialized methods of Digital Image Pre-processing (method 4.2) and methods of Artificial Intelligence (methods 4.3, 4.4) have been added to the existing methods of EDM (Data Visualization), which take into account the peculiarities of training future vocational education teachers. Specialized methods of digital image pre-processing are used, which implement filtering (in the spatial or frequency domain), contrast enhancement (local or global) and selection of image contours. Due to filtering, the level of noise in images is reduced, which leads to an increase in their visual quality. Increasing the local or global contrast also helps improve the visual quality of images. Due to the selection of contours, the boundaries of the studied objects are emphasized on the images, which simplifies their visual perception for students. Thus, specialized methods of digital image pre-processing increase the efficiency of Data Visualization methods.

Specialized methods of artificial intelligence, which implement image segmentation and object detection in images, have also been added to the existing EDM methods. Such intelligent image processing leads to the selection of researched objects on the images, which contributes to their better perception by students due to purposeful focusing of attention. The considered methods of intelligent image processing can be implemented when students perform certain stages of STEM projects, which contributes to the effective deepening of their knowledge and skills.

In the process of EDM, some parameters of objects take fuzzy values [41]. For example, when clustering educational subjects, there is a significant overlap of the clusters, that is, one subject may partially belong to several clusters. The same situation is observed when clustering students according to their educational achievements. Segments or areas of the image can belong to the object under study with varying degrees of certainty. Therefore, in order to increase the possibilities of processing such data, it is proposed to integrate the methods of fuzzy logic into the methods of EDM by specialists of engineering and pedagogical specialties. Due to the complex (integral) application of fuzzy logic to the existing methods of EDM, a number of specialized methods were obtained, namely (Fig. 3):

2.3. Fuzzy clustering.
4.5. Object detection using Fuzzy Logic

Thus, for the training of future vocational education teachers, it is proposed to apply the following methods of EDM (Fig. 3). In fig. 3 red rectangles show methods borrowed from digital image processing and artificial intelligence. Complex methods obtained by applying fuzzy logic to existing methods of IAOD are shown by blue rectangles.

The integration of the proposed methods of artificial intelligence, digital image processing and fuzzy logic in the EDM opens up new opportunities for improving the quality of education for the training of vocational education teachers. Let's consider in more detail such specialized methods of EDM.

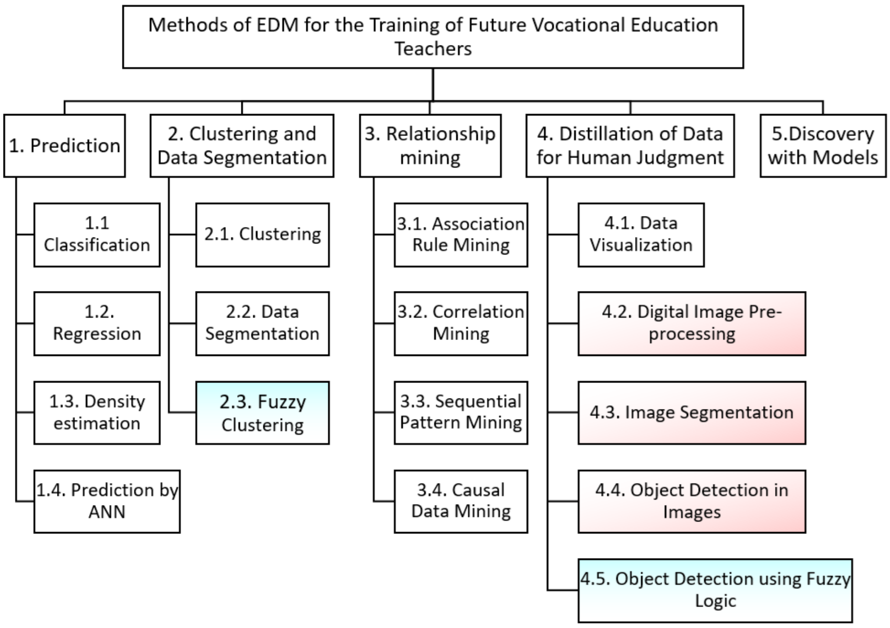


Fig. 3. Complex of EDM methods for the training of future vocational education teachers; added specialized methods are shown in color (compared to the classification in Fig. 2)

4. Results and Discussion

4.1 Fuzzy clustering

Fuzzy clustering (method 2.3 in Fig. 3) significantly expands the possibilities of ordinary (exact) clustering [35, 36]. The integration of clustering and fuzzy logic in educational data analysis creates a powerful tool for understanding complex educational processes, which allows adapting educational processes to the individual needs of students with greater accuracy [12]. This combination not only allows data to be grouped based on similar characteristics, but also effectively handles the uncertainty and subjectivity often present in educational data. The use of fuzzy logic is particularly effective when the clusters are not clearly separated. Due to the fuzzy membership function, it is possible to more accurately analyze objects that are within clusters. In an ideal case, the inter-cluster distance is much larger than the intra-cluster distance, so such clusters are clearly separated. However, when processing real educational data, there are often no clear boundaries between clusters. Because of these objects within the clusters, it is difficult to attribute them to only one cluster. Therefore, clustering methods are supplemented by fuzzy logic.

The task of clustering is to divide Q objects into Q_k clusters. The number of clusters is set manually or calculated automatically (for example, based on compactness criteria). The paper considers clustering in the space of two parameters (coordinates) x and y, which describe each object under study. The most common Euclidean distance between clustering objects was used, as the use of the Mathetian distance gives similar results. In the process of clustering, the ratio of intra-cluster distance to inter-cluster distance is minimized. Each cluster is described by a radius R, which is calculated as the maximum distance from the center of the cluster to its objects. Clusters are also described by the radius R_x in the x-coordinate and the radius R_y in the y-coordinate.

When using fuzzy logic, clusters of data are considered as fuzzy sets, so the belonging of objects to clusters is described by fuzzy membership functions $\mu_x(k, dx)$, $\mu_y(k, dy)$, $\mu_{xy}(k, dx, dy)$ [41, 42]. The fuzzy function $\mu_x(k, dx)$ describes the belonging of objects to clusters with numbers k by x coordinate, where dx is the distance of the object to the center of the cluster by x coordinate. Similarly, the fuzzy function $\mu_y(k, dy)$ describes the belonging of objects to clusters with numbers k along the y coordinate, where dy is the distance of the object to the center of the cluster along the y coordinate. The fuzzy function $\mu_{xy}(k, dx, dy)$ describes the belonging of objects to clusters with numbers k by coordinates x and y. The value of the fuzzy functions is in the range from 0 to 1, where 0 means that the object does not belong to the cluster at all (the distance of the object to the center of the cluster exceeds its radius), 1 means that it belongs to the cluster completely (the object is in the center of the cluster), and intermediate values indicate partial ownership. Due to the fuzzy membership functions $\mu_x(k, dx)$, $\mu_y(k, dy)$, $\mu_{xy}(k, dx, dy)$ it is possible to correctly describe the membership of a certain object, which is on the border of clusters, to several clusters at the same time. For example, triangular or Gaussian functions can be used as fuzzy membership functions. Triangular membership functions are easy to implement in software, and Gaussian functions provide a smooth change in the value of an object's membership to a cluster.

Fuzzy clustering can be performed by: 1) fuzzy clustering methods; 2) usual methods (for example, K-Means) followed by analysis using fuzzy membership functions. The second case is appropriate if you want to obtain fuzzy functions $\mu_x(k, dx)$ and $\mu_y(k, dy)$ of objects belonging to clusters separately for each parameter x and y.

Fuzzy clustering algorithms (Fuzzy clustering, soft clustering) include, for example, Fuzzy K-Means (fuzzy analogue of the K-Means algorithm) and Fuzzy C-Means (FCM) algorithms [43]. In such clustering algorithms, it is considered that one object can belong to several clusters at the same time with different degrees of belonging. The Fuzzy K-Means algorithm is a fuzzy analogue of the K-Means algorithm. In the Fuzzy C-Means algorithm, the number of clusters is calculated automatically, and in the Fuzzy K-Means algorithm, it is set by the user. Fuzzy clustering allows the identification of non-obvious groups of students with similar needs or problems, allowing educational institutions to adapt their curricula and resources more effectively.

Consider an example of fuzzy clustering of educational data [44]. Based on the initial data about each student, the following parameters are known: Average grade in previous courses "Previous Grades" (coordinate *x*) and Average number of study hours per week "Study Hours per Week" (coordinate *y*). Accordingly, in the space of such parameters *x* and *y*, each student corresponds to a certain point (Fig. 4). First, clear clustering of the data is performed using the K-Means method, and subsequent analysis is performed using fuzzy membership functions $\mu_x(k, dx)$, $\mu_y(k, dy)$, $\mu_{xy}(k, dx, dy)$. Clustering was performed by a Python program using the Google Colab cloud service (in the Jupyter Notebook). As a result of K-Means clustering, $Q = 200$ objects are divided into $Q_k = 4$ clusters (Fig. 4). Each obtained cluster means a characteristic group of students with a certain level of preparation and learning style:

- Cluster #1 is High value of the "Previous Grades" parameter and average value of the "Study Hours per Week".
- Cluster #2 is Average value of the "Previous Grades" parameter and high value of the "Study Hours per Week".
- Cluster #3 is Average value of the "Previous Grades" parameter and low value of the "Study Hours per Week".
- Cluster #4 is Low value of the "Previous Grades" parameter and medium value of the "Study Hours per Week".

Therefore, for each cluster of students, it is possible to separately analyze their behavior in the educational process, the admissibility of the average number of hours of study per week. Such clustering allows forming effective groups of students for joint educational activities.

However, the obtained clusters of students are not clearly separated, which creates problems for the analysis of students at the boundaries of the clusters. Therefore, fuzzy triangular membership functions $\mu_x(k, dx)$ (Fig. 5), $\mu_y(k, dy)$ and $\mu_{xy}(k, dx, dy)$ were used to describe cluster membership, which describe the degree of student membership in clusters with numbers *k* depending from *x* and *y* coordinates. In this case, the parameter "Previous Grades" is used as the *x* coordinate, and the parameter "Study Hours per Week" is used as the *y* coordinate.

The values of the degree of membership $\mu_{xR}(k)$ (Fig. 6a) of the student with number i_v (Fig. 4) to clusters *k* are calculated on the basis of fuzzy functions of membership $\mu_x(k, dx)$ and distances $dx = dx0$ to the center of the cluster according to the parameter "Previous Grades" (Fig. 5). For example, the student with number i_v (Fig. 6a) according to the parameter "Previous Grades" most belongs to cluster 2 (the value of the degree of belongingness $\mu_{xR}(2) = 0.98$), which corresponds to the fuzzy set "Average value of the parameter "Previous Grades" and a high value of the parameter "Study Hours per Week". To a lesser extent, such a student belongs to cluster 4 ($\mu_{xR}(4) = 0.82$) and the corresponding fuzzy set "Low value of the parameter "Previous Grades" and average value of the parameter "Study Hours per Week"". Similarly, the value of the student's degree of membership $\mu_{yR}(k)$ to clusters *k* is calculated based on $\mu_y(k, dy)$ and distances $dy = dy0$ to the center of the cluster according to the "Study Hours per Week" parameter.

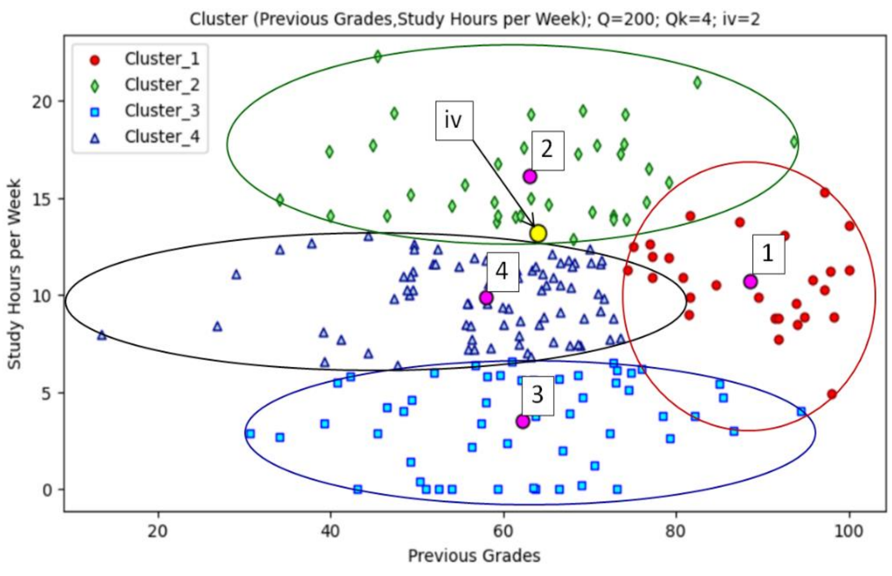


Fig. 4. An example of student data clustering according to the parameters "Previous Grades" (Average grade on previous courses), "Study Hours per Week" (Average number of study hours per week) [44]; cluster boundaries are marked with ellipses; the centers of clusters #1-4 are signed with the corresponding numbers; the student with number iv =2 is highlighted with an arrow

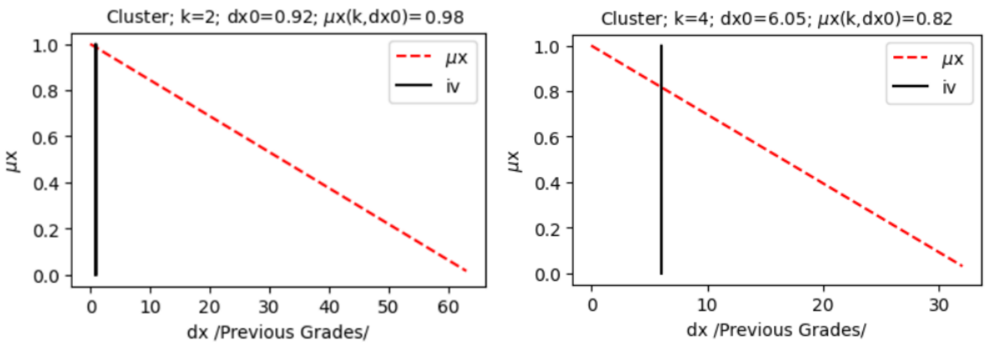
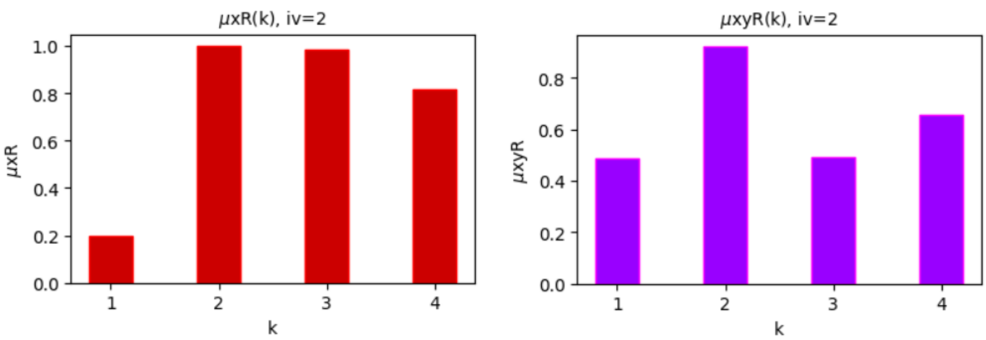


Fig. 5. An example of fuzzy functions $\mu_x(k, dx)$ of students belonging to clusters with numbers k = 2, 4 according to the parameter x "Previous Grades", where dx is the distance of the object to the cluster center by x coordinate, and the radius of the cluster is equal to the maximum value dx; the values of the distance $dx0$ to the cluster center for the student with number i_v = 2 are indicated by a vertical line (Fig. 4)

The values of the degree of membership $\mu_{xR}(k)$ (Fig. 6a) of the student with number i_v (Fig. 4) to clusters with numbers k are calculated on the basis of fuzzy functions of membership $\mu_x(k, dx)$ and distances $dx = dx0$ to the center of the cluster according to the parameter "Previous Grades" (Fig. 5). For example, the student with number i_v (Fig. 6a) according to the parameter "Previous Grades" most belongs to cluster 2 (the value of the degree of belongingness $\mu_{xR}(2)$ = 1.0), which corresponds to the fuzzy set "Average value of the parameter "Previous Grades" and a high value of the parameter "Study Hours per Week". To a lesser extent, such a student belongs to cluster 4 ($\mu_{xR}(4)$ = 0.82) and the corresponding fuzzy set "Low value of the parameter "Previous Grades" and average value of the parameter "Study Hours per Week"". Similarly, the value of the student's degree of membership $\mu_{yR}(k)$ to clusters #k is calculated based on functions $\mu_y(k, dy)$ and distances $dy = dy0$ to the cluster center according to the "Study Hours per Week" parameter.

The resulting values of the degree of membership $\mu_{xyR}(k)$ of the student with number i_v to clusters with numbers k are calculated on the basis of fuzzy membership functions $\mu_{xy}(k, dx, dy)$ and distances $dx = dx0$, $dy = dy0$ to the center of the cluster according to the parameters "Previous Grades" and "Study Hours per Week" respectively. However, since in the general case the range of changes of the clustering parameters (x, y) is different, and the shape of the clusters is asymmetric, therefore the resulting values of the degree of membership $\mu_{xyR}(k)$ are calculated more correctly based on the average value of the degrees of membership $\mu_{xR}(k)$ and $\mu_{yR}(k)$. The obtained values of the degree of belonging $\mu_{xyR}(k)$ make it possible to estimate the belonging of the student with number i_v to different clusters k according to both parameters "Previous Grades" and "Study Hours per Week" (Fig. 6b).

For example, the student with number i_v (Fig. 6b) by both parameters belongs to cluster 2 the most ($\mu_{xyR}(2)$ = 0.93), and to a lesser extent belongs to the neighboring cluster 4 ($\mu_{xyR}(4)$ = 0.67). Such information can be used to evaluate the educational behavior of a certain student, to transfer students to project groups with common interests, to check the admissibility of his average number of hours of study per week. For example, the considered student with number i_v is proposed to be included in groups (teams) of students with an average value of the parameter "Previous Grades" and a high value of the parameter "Study Hours per Week" (Cluster #2), and in the other case, in groups with a low value of the parameter " Previous Grades" and the average value of the parameter "Study Hours per Week" (Cluster #4).



a) b)

Fig. 6. Values of degrees of membership of a student with number iv = 2 (Fig. 4) to clusters with numbers k: a) $\mu_{xR}(k)$ values calculated on the basis of fuzzy membership functions $\mu_x(k, dx0)$ (Fig. 5); b) values of $\mu_{xyR}(k)$, calculated on the basis of values of degree of belonging $\mu_{xR}(k)$ and $\mu_{yR}(k)$

Fuzzy clustering makes it possible to effectively group students according to their characteristics of the educational process: educational behavior, test results in various subjects, etc. Due to this, it is possible to personalize the learning trajectories of students, to offer them materials and tasks that best match their level of knowledge and learning style [45]. Defining vague clusters of students based on similar interests and skills helps to form effective groups for joint projects or discussions. The integration of clustering and fuzzy logic in educational data analysis opens up new opportunities for deeper understanding and improvement of the educational process.

4.2 Digital Image Pre-processing

Specialized methods of digital images pre-processing (methods 4.2 in Fig. 3) take into account the peculiarities of tasks that arise during the training of future specialists in engineering and pedagogical specialties. Digital pre-processing of images can improve the accuracy and quality of Data Distillation for Human Judgment techniques that are associated with image processing. Such processing consists in removing noise from images by filtering, increasing global or local contrast, calculating and visualizing image profiles, highlighting contours. This allows to increase the accuracy of visualization, segmentation, detection and subsequent analysis of images.

Filtering methods make it possible to reduce the level of noise in images and increase their visual quality, which significantly improves students' perception of such images in the educational process. In addition, image filtering is an important stage of many STEM projects [46] for engineering and pedagogical students, as it is widely used in their future professional activities. Image filtering is performed in the spatial or frequency domain (using forward and inverse Fourier transforms) [47]. Common filtering methods in the spatial domain include the Gaussian filter and the median filter. The Gaussian filter leads to blurring of contours, so it is advisable to use it for images with smooth changes in brightness. The median filter preserves contours better, so it can be applied to remove noise in images with sharp contours. In most cases, non-oriented filtering of images is used, however, oriented filtering, which is performed along the contours, is particularly effective for maintaining the clarity of contours [48].

Image contrast enhancement methods, as well as filtering methods, make it possible to improve the visual quality of images [49]. By increasing the contrast of the images of the studied objects, students' perception of such objects is simplified, which is important for increasing the efficiency of the educational process. Increasing the contrast of images is an important stage of many STEM projects for students of vocational education teachers related to the construction of computer systems for intelligent image processing, in particular, images of railway transport [50] and cars (Fig. 7).

a) b)

Fig. 7. Initial fRGB color image (a) and gRGB image with non-uniform background removed and local contrast increased (b) [51]

The initial color image is read (from a file or from a video camera) as a matrix $f_{RGB}(i, k, c)$, where i = 0, ..., M-1; k = 0, ..., N-1; c = 0,..., Q_C-1; M is the quantity of pixels by height; N is the quantity of pixels in width; Q_C = 3 is the quantity of color channels. Local image contrast enhancement, which is performed in windows w (size $M_w \times N_w$ pixels; in most cases $M_w = N_w$), is particularly effective. This allows to increase the contrast for low-contrast areas even if there are high-contrast areas in the image. When reducing the size of the w window, the contrast of the resulting g_{RGB} image increases, but at the same time artifacts (defects) may appear in the image. By changing the size of the window w, the contrast of the resulting g_{RGB} image is adjusted: when the size of the window w is reduced, the contrast increases.

Image profile analysis is useful for accurate evaluation of filtering results and image contrast enhancement [52]. Analysis of image profiles before and after processing makes it easier for students to understand the principles of the applied processing method. For example, noisy image profiles have a significant random (high-frequency) component, and after filtering, the image profiles are more smoothed.

Due to the selection of contours [53] the borders of the studied objects are emphasized on the images, which simplifies their visual perception and diagnosis for students. Extraction of contours of images in most cases is performed by the methods of Sobel, Prewitt and Canny. The Canny method provides the highest accuracy of contour selection, but this method has a relatively low speed (compared to the Sobel, Prewitt methods).

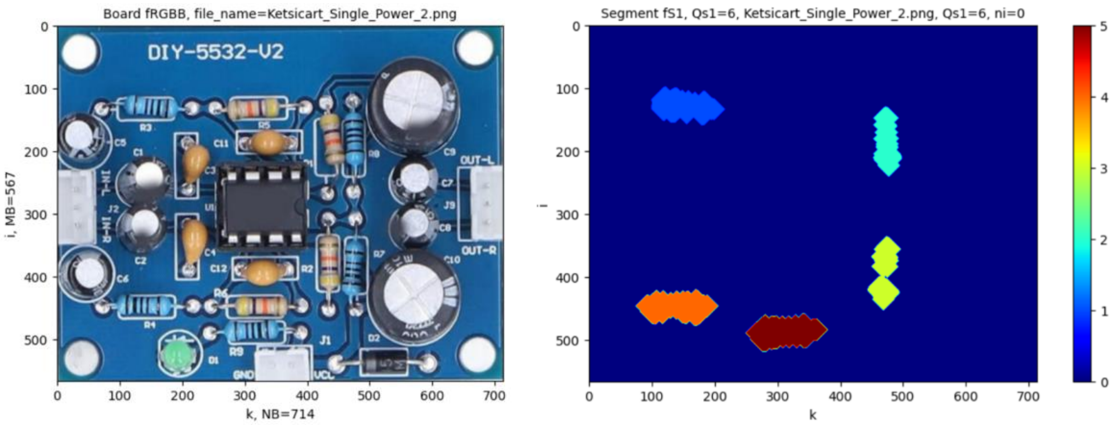
Thus, specialized methods of digital image pre-processing increase the efficiency of data visualization methods, simplify the perception and analysis of the studied images. In this way, image processing achieves the main tasks of methods of Distillation of data for human judgment.

4.3 Image Segmentation

Specialized image segmentation methods (methods 4.3 in Fig. 3) are a component of intelligent image analysis methods [54]. Such methods take into account the peculiarities of the tasks that arise before specialists of engineering and pedagogical specialties in the educational process and in practical activities. In the educational process, segmentation allows you to effectively focus students' attention on the studied objects (segments) by purposefully increasing their contrast and reducing the contrast of the background, which corresponds to the tasks of methods of Distillation of data for human judgment. Practical application of image segmentation methods allows automating a number of tasks in industry, medicine, science, transport, and other fields [55].

Image segmentation allows to highlight meaningful areas (for example, people's faces, car parts, electronic device components) in images that have similar brightness, color, or texture. Common image segmentation methods include the method of building up homogeneous regions, watersheds, and contour lines [8, 50]. In the method of growing homogeneous areas, the initial pixels of the segments are first selected, to which new pixels with close values of brightness and/or color are joined. The principle of the watershed method consists in highlighting contours on the image, on the basis of which segments are determined. Therefore, the watershed method is effective in segmenting images with clear contours, for example, symbol images. Segmentation is performed for both grayscale and color images. The segmentation process is the simplest for binarized images, so in some cases before segmentation images are binarized with a certain threshold (calculated, for example, by the Otsu's method). On the binarized images, their contours are determined, on the basis of which the images are divided into segments by the method of contour lines.

Modern segmentation methods include the method of Semantic Image Segmentation using a convolutional neural network (CNN) [56]. The initial images are fed to the CNN, and the segmented images are obtained at the outputs. CNN training is performed on the images of the training sample (train dataset), and the images of the control sample (validation dataset) are used to prevent overtraining. In the educational process, segmentation can be applied before image visualization, which allows to automatically divide the image into meaningful areas (regions of interest (ROI)) [57] and focus students' attention on such areas (Fig. 8, Fig. 9). After segmentation, the resulting segments are selected according to height, width, orientation, area, color and brightness, which allows to select only the investigated objects, for example, components of electrical circuits (resistors, capacitors, etc.) in the image (Fig. 8).



a) b)

Fig. 8. The initial image of the electronic device (a) and the image after segmentation using the method of contour lines and selection of the resulting segments by size and color (b)

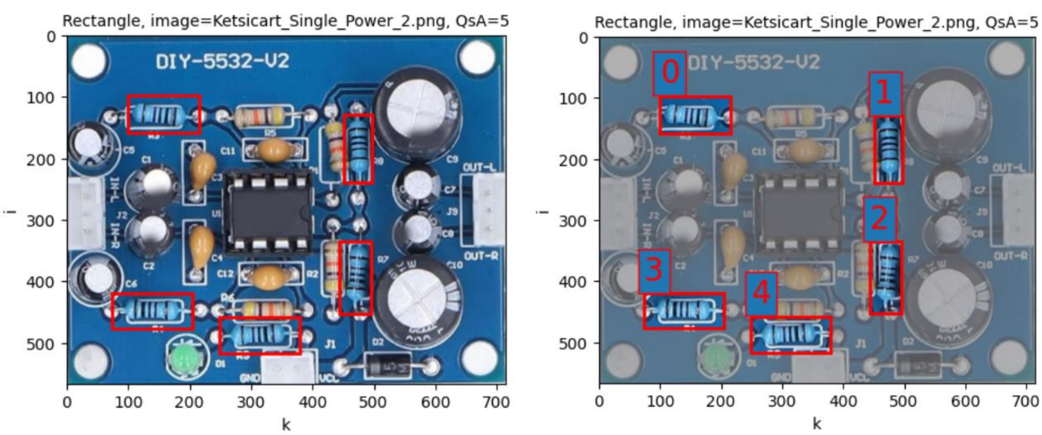


Fig. 9. Visualization of the device components (resistors #0-4 of the specified rating), selected by segmentation (Fig. 8b); focusing students' attention on the studied objects is achieved by highlighting them with frames, reducing the contrast and/or saturation of other objects

In the process of visualization, the obtained segments are highlighted with frames and marks (infographic method) (Fig. 9), which allows students to focus their attention on the studied objects. Reducing the contrast of the image outside the selected areas of the segments makes it possible to visually highlight the objects under study.

Image segmentation simplifies their further computer processing by automatic analysis of segment parameters (rotation angles, sizes, areas, perimeter length, etc.). Image segments often correspond to objects of interest in the image, such as symbols or electronic components. Therefore, image segmentation simplifies the following stages of its intelligence analysis: recognition and detection of objects.

4.4 Object Detection in Images

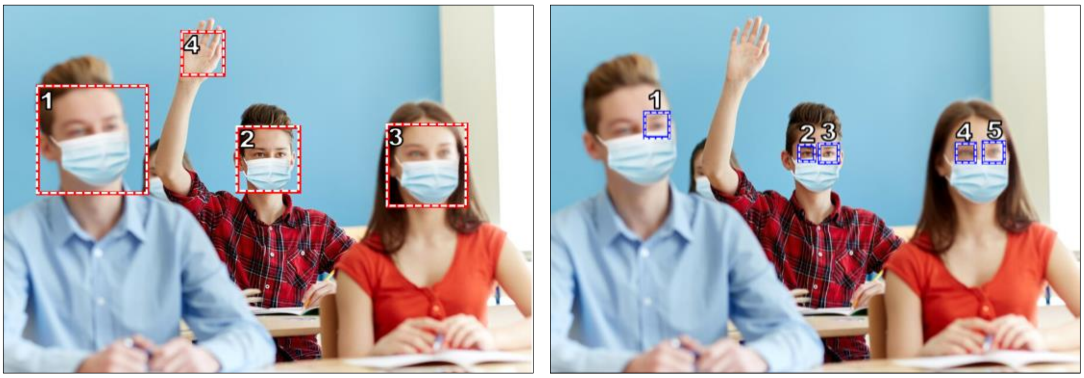
Specialized methods of object detection in images (methods 4.4 in Fig. 3) are used to determine the spatial position of the studied objects in the images (the position of the object in the image is often indicated by a rectangular frame) [8, 58]. Such detection is used, for example, to determine the spatial position of participants in the educational process (determining the number of students, their location near computers or near laboratory equipment, etc.). Image recognition is used as a detection stage, i.e. the presence or absence of an object of a certain class in the investigated area of the image or video is established by recognition methods. Object detection is used in many areas of computer vision, for example, in image retrieval systems and video surveillance. In the case of processing a series of images or videos, it is possible to determine the trajectories of the movement of objects (Object Tracking), for example, students during practical tasks. Object detection in images is performed with or without the use of ANN. Input signals of such methods are initial images, and output signals are coordinates of rectangles with localized objects. A common method for object detection in images that does not use ANN is the Viola-Jones method.

Neural network methods use convolutional neural networks (CNN), such as R-CNN and YOLO (You Only Look Once) [59]. R-CNN networks (Region Based Convolutional Neural Networks) are designed to detect and localize objects in certain areas (regions) of images. Such networks belong to two-level networks: at the first level, the areas of the image with the likely location of the object are determined, and at the second level, the presence of the object is checked (recognition is performed, that is, the class of the object is determined). YOLO networks divide the image into a set of rectangular areas (cells) and identify the areas with the most likely location of the object. A feature of YOLO is its high speed, which allows it to detect objects on video in real time.

An important advantage of the Viola-Jones method is the use of trained Haar cascades, which allows to avoid a complex and lengthy training procedure (which is necessary in the case of ANN). The Viola-Jones method is one of the best in terms of the ratio of detection accuracy/speed of work. The Viola-Jones detector also has a low probability of falsely detecting objects (e.g., faces). There are specialized Haar cascades designed to detect frontal images of faces and images of faces in profile, as well as to detect facial features (eyes, mouth). Haar features are calculated for rectangular regions (windows) that overlap the entire image. The presence or absence of an object in the window is determined by the difference between the value of the Haar feature and the set threshold. The Viola-Jones method uses a large number of Haar features, so the features are organized into a cascade classifier. The result of the Viola-Jones method is the position of the objects on the image, which are described by the following parameters: x and y are the coordinates of the upper right corner of the rectangle (which bounds the desired object, for example, a face), w is the width, h is the height.

Before object detecting using the Viola-Jones method, the values of the scaleFactor and minNeighbors parameters are pre-set (otherwise, the parameter values are set by default). The scaleFactor parameter specifies how much the scan window size is reduced for each image scale (for example, scaleFactor = 1.15). The minNeighbors parameter, which indicates how many neighbors each rectangle (detected part of the object) should have in order to save it during selection (the value of minNeighbors varies from 1 to 6).

Consider an example of detecting faces (Fig. 10a) and eyes (Fig. 10b) of students using the Viola-Jones method with fixed values of the scaleFactor and minNeighbors parameters [58]. As a result, area #4 of the image is mistakenly detected as a face (Fig. 10a). Eye detection in an image is performed using the appropriate Haar cascade (Fig. 10b).



a) b)

Fig. 10. Objects detection in the image of students using the Viola-Jones method [60]: a) detection of face, scaleFactor = 1.1, minNeighbors = 3; b) detection of eye, scaleFactor =1.05, minNeighbors = 2

The accuracy of object detection using the Viola-Jones method can be increased by choosing the optimal values of the scaleFactor and minNeighbors parameters (for certain types of images). In addition, in order to increase the accuracy of detection, it is advisable to apply fuzzy logic and comprehensively take into account the confidence of detecting faces and their features (eyes, mouth).

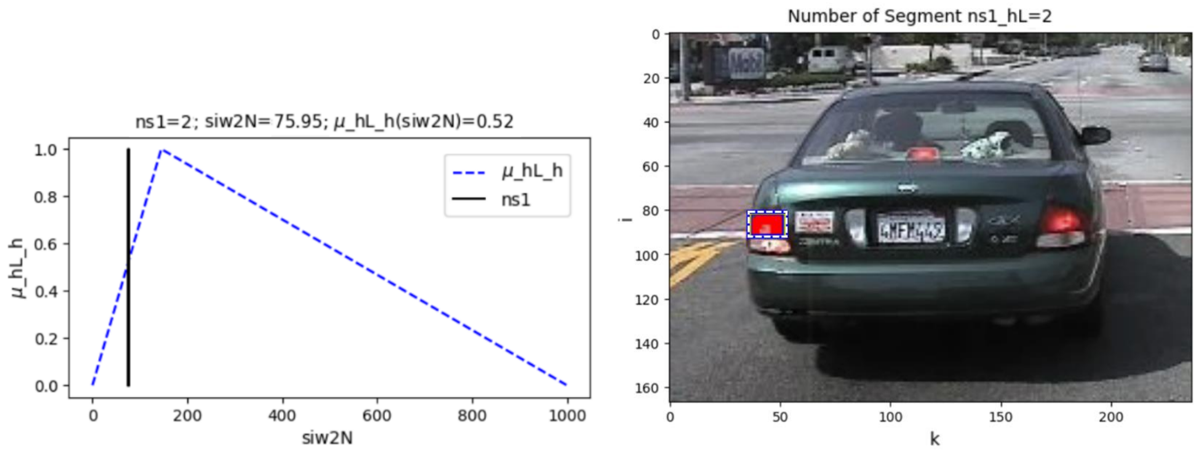
4.5 Object detection using Fuzzy Logic

Specialized methods of object detection in images using fuzzy logic (methods 4.5 in Fig. 3) are used to determine the confidence of finding the object under study in a certain spatial position on the image. Such object detection is performed by integrating fuzzy logic with segmentation methods and with methods for clear object detection in the image (for example, with the Viola-Jones method).

Image Segmentation and Fuzzy Logic make it possible to automatically detect complex-shaped objects that correspond to image segments [50]. Since the shape of the segment can only partially coincide with the shape of the object under study (due to uneven lighting and other experimental conditions), fuzzy membership functions are introduced to describe the degree to which image segments belong to the object under study. Based on the obtained degree of belonging, spatial object detection and determination of detection confidence is performed. Such processing is effective, for example, for analyzing images of vehicles when performing STEM projects. In such cases, fuzzy logic methods reveal segments belonging to the investigated parts of the vehicle: windows, license plates, car light, etc. (Fig. 11). The detection process consists of two stages.

At the first (training) stage, the segmentation of the training set of images is performed and the parameters of the fuzzy membership functions are calculated, which describe the membership of the segments to the object under study. On each image, the parameters of the segment that corresponds to the studied object (for example, the area of headlights and license plates) are determined. For each of these segments, the normalized height siw2N is calculated (for an image height of 1000 pixels, and for other heights the values of siw2N change proportionally), the normalized width of the segment skw2N (for an image width of 1000 pixels), the normalized coordinates of the center of the segment by height sic2N by width skc2N. Based on the calculated parameters of segments siw2N, skw2N, sic2N, skc2N (for all images of the training set), their average height siw2A, average width skw2A, average coordinates of the center of the segment by height sic2A and width skc2A are determined. That is, the parameters siw2A, skw2A, sic2A, skc2A describe the average values of the parameters of the segments that correspond to the object under study in the images of the training set. The value of the average height of siw2A is used as the coordinate of the maximum for the fuzzy triangular function of the segment μ_{hLh}(siw2N) belonging to the object under study, taking into account its height (Fig. 11a). That is, the more precisely the normalized height of the segment siw2N corresponds to the average height of siw2A, the greater the value of the μ_{hLh} function (the values of which are in the range from 0 to 1) is obtained. Similarly, the coordinates of the maximum are determined for the membership function μ_{hLw}(skw2N) taking into account the width of the segment skw2N, the membership function μ_{hLic}(sic2N) taking into account the coordinates of the segment center by height sic2N, the membership function μ_{hLkc}(skc2N) taking into account the coordinates of the segment center by width skc2N.

At the second (working) stage of detection, the segmentation of the investigated image is performed, the parameters siw2N, skw2N, sic2N, skc2N and values of fuzzy functions μ_{hLh}(siw2N), μ_{hLw}(skw2N), μ_{hLic}(sic2N), μ_{hLkc}(skc2N) are calculated for all its segments. The values of the resulting membership function $\mu_{hLs}(n_{s1})$ of the segment with number n_{s1} to the object under study, taking into account all parameters of the segment (height, width, coordinates of the center by height and width), are determined as the product of the values of the corresponding membership functions. Based on the maximum of the μ_{hLs} membership function, the segment number n_{s1_hL}, which most completely belongs to headlights and license plates, is calculated (Fig. 11b). Other objects in the image are detected similarly.



a) b)

Fig. 11. Object detection using image segmentation [61] and fuzzy logic: a) fuzzy membership function μ_{hLh}(siw2N), which describes the membership of the segment with the number n_{s1} to left rear car light depending on the normalized height of the segment siw2N; b) the result of detecting left rear car light for the gRGB image (Fig. 7b)

Object detection in images using the Viola-Jones method and fuzzy logic allows not only to establish the fact of the presence or absence of the object under study in a certain spatial position (rectangle), but also to determine the confidence of the object's localization [58]. The Viola-Jones method uses Haar cascades designed to detect the face and its features (eyes, mouth), so the complex use of the results of various Haar cascades allows to increase the accuracy of object detection. For example, if a face, eyes and mouth are detected with high confidence in a certain area of the image, then a face without a mask is also detected in such an area with high confidence. If the face and eyes are detected with high confidence in the area of the image, and the mouth is detected with low confidence, then a mask on the face is detected in such an area with high confidence.

The confidence of object detection by the Viola-Jones method is defined as the maximum value of the minNeighbors parameter (the number of neighboring rectangles for the detected area) at which a face is still detected. Therefore, the face detection confidence D_f is in the range from 1 to 6 (the higher the D_f, the more confident the face is detected; $D_f = 1$ means that the face is detected with minimal confidence). Similarly, the eye detection confidence value D_e is calculated, which is in the range from 0 to 5 ($D_e = 0$ means that the eye is not detected in a certain area). The confidence value of mouth detection D_s is in the range from 0 to 5.

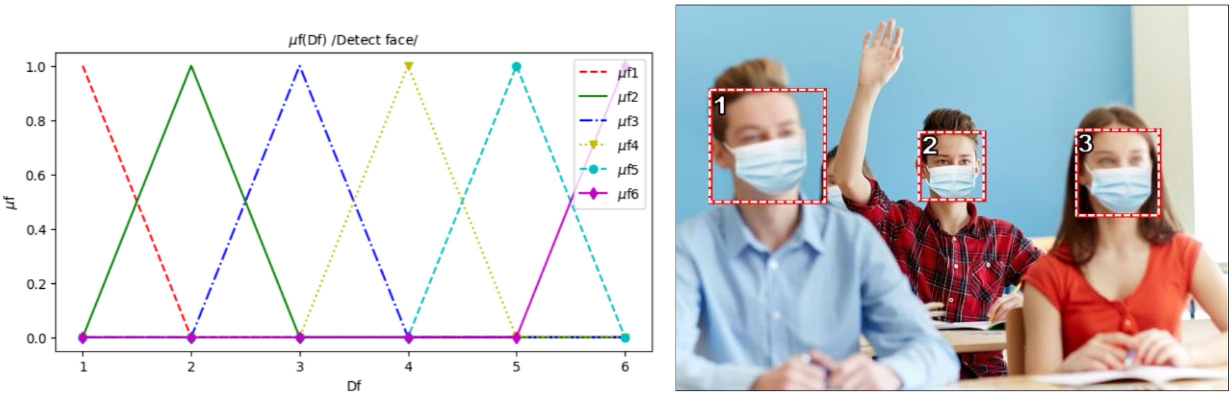
Face detection based on face and eye regions is performed by a fuzzy system whose input parameters are the face detection confidence D_f and eye detection confidence D_e (for a certain image region), and the output parameter is the eye detection face confidence D_{fr} (which ranges from 1 to 6). Mamdani fuzzy knowledge base #1 [58] has been developed for face detection, which contains 36 rules, for example:

Rule #1; If DfL= Low and DeL = Low, then DfrL = Low.
Rule #6; If DfL= Below Average and DeL = Below Average, then DfrL = Below Average.
Rule #14; If DfL= Above Average and DeL = Below Average, then DfrL = Average.
Rule #20; If DfL= Average and DeL = Average, then DfrL = Above Average.
Rule #30; If DfL= Below High and DeL = Above Average, then DfrL = Below High.
Rule #33; If DfL= High and DeL = High, then DfrL = High.

In the rules of the knowledge base, the notation DfL describes the linguistic variable "Confidence of face detection", DeL – "Confidence of eye detection", DfrL – "Confidence of face detection with eyes in mind".

The terms of the linguistic variable D_{fL} "Face detection confidence" are the values "low", "below average", "medium", "above average", "below high", "high", and the numerical linguistic variable is described by the face detection confidence D_f. Similar terms describe other linguistic variables. The terms of the linguistic variable D_{fL} "Face detection confidence" are described by fuzzy sets with triangular membership functions $\mu_{f1}(D_f)$, $\mu_{f2}(D_f)$... $\mu_{f6}(D_f)$ and the carrier D_f (Fig. 12a). The terms of other linguistic variables are similarly described by fuzzy sets with triangular membership functions.

Based on the parameters of the detected face and eyes using the Mamdani knowledge base, an eye-aware face detection (with D_{fr} confidence) is performed by fuzzy logic inference. A face is considered detected if confidence $D_{fr} > 0$. Therefore, faces #1-3 are correctly detected in the resulting image (Fig. 12b), and section #4 (for which $D_{fr} = 0$) is rejected (Fig. 10a).



a) b)

Fig. 12. Detection of faces in the image using the Viola-Jones method and fuzzy logic: a) fuzzy membership functions of the linguistic variable DfL "Confidence of face detection": μf1 (red graph) is the confidence of face detection is low; μf2 is lower than average, μf3 is average; μf4 is above average; μf5 is lower than high; μf6 is high; b) detection of faces taking into account the eyes

Face mask detection based on the detected face and mouth regions (obtained by the Viola-Jones method) is performed by a fuzzy system (based on Mamdani fuzzy knowledge base #2), whose input parameters are the face recognition confidence D_{fr} and the mouth detection confidence D_s, and the output parameter is confidence of D_m mask recognition [58]. The developed system for detecting faces and masks on faces can be practically used not only for accounting and determining the state of participants in the educational process, but also in STEM education [62, 63] when studying the principles of image detection.

5. Conclusion

A complex of methods of Educational Data Mining (EDM) has been developed, which take into account the peculiarities of the training of vocational education teachers. The complex of methods was developed by adding specialized methods to the basic methods of EDM. Prediction, Clustering, Relationship Mining, Distillation of Data for Human Judgment, Discovery with Models are used as basic methods of EDM.

Such prediction methods as Classification, Regression, Density Estimation are considered. Based on the analysis of modern EDM systems, a conclusion was made about the prospects of prediction of educational data using Artificial Neural Networks, in particular long-short-term memory (LSTM) and convolutional neural networks (CNN).

The main methods of clustering educational data are described, in particular, the K-means method, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) and hierarchical clustering. To expand the possibilities of processing educational data, clustering methods are supplemented with the method of data segmentation, the feature of which is the preliminary establishment of the number and characteristics of segments. The integration of fuzzy logic with data clustering methods was performed. The methods of Relationship Mining, namely, Association Rule Mining, Correlation Mining, Sequential Pattern Mining, Causal Data Mining, were considered.

The existing methods of Distillation of Data for Human Judgment, namely methods of Data Visualization (Visual Mining), in particular, methods of infographics, image processing and pivot tables, are analyzed. Distillation of data for human judgment are complemented by specialized methods of digital image pre-processing and artificial intelligence methods: image segmentation, object detection in images. Such methods of digital images pre-processing as filtering, contrast enhancement, contour selection were used. Methods of object detection in images using convolutional neural networks and the Viola-Jones method are considered. Integration of fuzzy logic with methods of image segmentation and object detection in images was performed.

Methods of Discovery with Models are described, which include the areas of Simulation of Educational Processes and Optimization of Learning Strategies. Simulation of Educational Processes is performed by developing computer models of the system that simulate the interaction between students and the educational environment. Optimizing of Learning Strategies consists in using models to assess potential effectiveness and optimize learning approaches.

The novelty of the work is the integration of fuzzy logic with methods of data clustering, image segmentation and object detection in images. The novelty of the work also includes the integration of methods of Distillation of Data for Human Judgment with methods of digital image pre-processing and methods of artificial intelligence.

Thus, the main results of the work are the development of EDM methods, which are focused on the training of vocational education teachers. Such methods take into account the need to use digital image processing methods and artificial intelligence in the educational process. The integration of fuzzy logic with the methods of data clustering, image segmentation, and image object detection has made it possible to effectively process educational data even with a certain degree of uncertainty. The developed methods are practically applied in the educational process, in particular, for performing STEM projects.

References

[1] H. Aldowah, H. Al-Samarraie, and W.M. Fauzy, “Educational data mining and learning analytics for 21st century higher education: A review and synthesis,” *Telemat. Inform.*, vol. 37, pp. 13-49, 2019. doi: 10.1016/j.tele.2019.01.007.

[2] Mina Fazlikhani, and Iman Bagheri, “Educational Computer Science Development Based on Data Mining,” In *5th International Conference on Science, Engineering, and role of Technology in new Busines*, Copenhagen-Denamrk, pp.1-8, December 16, 2023.

[3] M. Zhang, Y. Cheng, J. Gu, Y. Yang, L. Chen, and B. Cui, “Digital Evaluation of Innovative Logistics Talents Based on Improved SAGA-FCM Algorithm,” In: Hu, Z., Zhang, Q., He, M. (eds), Advances in Artificial Systems for Logistics Engineering, Springer, Cham., vol. 223, pp. 374-383, 2024. doi: 10.1007/978-3-031-72017-8_34.

[4] R. Baker, and K. Yacef, “The State of Educational Data Mining in 2009: A Review and Future Visions,” *Journal of Educational Data Mining*, vol. 1, no.1, pp. 3–17, 2009. doi: 10.5281/zenodo.3554658.

[5] R. Baker et al., “Data mining for education,” International encyclopedia of education, vol. 7, pp. 112–118, 2010.

[6] Saurabh Bilgaiyan, Kunwar Aditya, Samaresh Mishra, Madhabananda Das, "A Swarm Intelligence based Chaotic Morphological Approach for Software Development Cost Estimation", International Journal of Intelligent Systems and Applications, Vol.10, No.9, pp.13-22, 2018.

[7] Adlen Kerboua, Mohamed Batouche, "3D Skeleton Action Recognition for Security Improvement", International Journal of Intelligent Systems and Applications, Vol.11, No.3, pp.42-52, 2019.

[8] DaeEun Kim, Dosik Hwang, *Intelligent Imaging and Analysis*. Switzerland, Basel, MDPI, 2020.

[9] Z. Yang, J. Yang et al. “Using convolutional neural network to recognize learning images for early warning of at-risk students,” *IEEE Trans Learn Technol*, vol. 13, no. 3, pp. 617–630, 2020. doi: 10.1109/TLT.2020.2988253.

[10] Nataliya Mutovkina, Zhongfeng Pan, "Solving the Problem of Pre-Selection of Personnel Based on Fuzzy Sets", International Journal of Information Engineering and Electronic Business, Vol.16, No.5, pp. 85-94, 2024.

[11] Ahmed Fahim, "Finding the Number of Clusters in Data and Better Initial Centers for K-means Algorithm", International Journal of Intelligent Systems and Applications, Vol.12, No.6, pp.1-20, 2020.

[12] Serhiy Balovsyak, Oleksandr Derevyanchuk, Hanna Kravchenko, Yuriy Ushenko, and Zhengbing Hu, "Clustering Students According to their Academic Achievement Using Fuzzy Logic," *International Journal of Modern Education and Computer Science*, vol. 15, no. 6, pp. 31-43, 2023.

[13] O.V. Derevyanchuk, and S.V. Balovsyak, Certificate of copyright registration for the work, No. 123369, 31.01.2024. ID CR3206310124. Computer program "Data clustering using fuzzy logic" ("ClusterFuzzy23"). Ukrainian National Office for Intellectual Property and Innovations (IP Office). URL: https://sis.nipo.gov.ua/uk/services/original-document/

[14] F. Castro, A. Vellido, A. Nebot, and F. Mugica, "Applying data mining techniques to e-learning problems," in *Evolution of teaching and learning paradigms in intelligent environment*, Springer, pp. 183–221, 2007.

[15] Abdulmohsen Algarni, "Data Mining in Education," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 7, no. 6, pp. 456-461, 2016.

[16] C. Romero, and S. Ventura, "Data mining in education," *Wiley Interdisciplinary Reviews: Data mining and knowledge discovery*, vol. 3, no. 1, pp. 12-27, 2013. doi: 10.1002/widm.1075.

[17] Mais Haj Qasem, Raneem Qaddoura, and Bassam Hammo, "Educational Data Mining (EDM): A Review," In *Conf. New Trends in Information Technology (NTIT)-2017*, 25-27 April 2017, The University of Jordan, Amman, Jordan, pp. 149-156, 2017.

[18] Q. Zhou, W. Quan et al, "Predicting high-risk students using internet access logs," *Knowl, Inf, Syst*., vol. 55(2), pp. 393–413, 2018. doi: 10.1007/s10115-017-1086-5.

[19] S. Batool, J. Rashid, M.W. Nisar, et al., "Educational data mining to predict students' academic performance: A survey study," *Educ. Inf. Technol*, vol. 28, pp. 905-971, 2023. doi: 10.1007/s10639-022-11152-y.

[20] Zareen Alamgir, Habiba Akram, Saira Karim, and Aamir Wali, "Enhancing Student Performance Prediction via Educational Data Mining on Academic data," *Informatics in Education*, vol. 23, no. 1, pp. 1-24, 2024. doi: 10.15388/infedu.2024.04.

[21] Bassam M. Kanber, Naglaa F. Noaman, Amr M. H. Saeed, Mansoor Malas, "DDoS Attacks Detection in the Application Layer Using Three Level Machine Learning Classification Architecture", International Journal of Computer Network and Information Security, Vol.14, No.3, pp.33-46, 2022.

[22] Nyme Ahmed, Sultanul Arifeen Hamim, Dip Nandi, "A Comprehensive Study to Analyze Student Evaluations of Teaching in Online Education", International Journal of Modern Education and Computer Science, Vol.16, No.5, pp. 105-117, 2024.

[23] Maxwell Kontor Owusu, Bakari Yusuf Dramanu, Mark Owusu Amponsah, " Classroom Management Strategies and Academic Performance of Junior High School Students ", International Journal of Education and Management Engineering (IJEME), Vol.11, No.6, pp. 29-38, 2021.

[24] N, Ahmad, N. Hassan, H. Jaafar, and N.I.M. Enzai, "Students' Performance Prediction using Artificial Neural Network," *IOP Conf. Ser.: Mater. Sci. Eng*., vol. 1176, pp.1-8, 2021. doi: 10.1088/1757-899X/1176/1/012020.

[25] Carlos Felipe Rodríguez-Hernández, Mariel Musso, Eva Kyndt, and Eduardo Cascallar, "Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation," *Computers and Education: Artificial Intelligence*, vol. 2, pp.1-14, 2021. doi: 10.1016/j.caeai.2021.100018.

[26] S. Sarsa, J. Leinonen, and A. Hellas, "Empirical Evaluation of Deep Learning Models for Knowledge Tracing: Of Hyperparameters and Metrics on Performance and Replicability," *Journal of Educational Data Mining*, vol. 14, no. 2, 2022. doi: 10.5281/zenodo.7086179.

[27] P. Nayak, S. Vaheed, S. Gupta, et al., "Predicting students' academic performance by mining the educational data through machine learning-based classification model," *Educ. Inf. Technol*, vol. 28, pp. 14611-14637, 2023. doi: 10.1007/s10639-023-11706-8.

[28] M. Delianidi, and K. Diamantaras, "KT-Bi-GRU: Student Performance Prediction with a Bi-Directional Recurrent Knowledge Tracing Neural Network," *Journal of Educational Data Mining*, vol. 15, no. 2, pp. 1-21. 2023. doi: 10.5281/zenodo.7808087.

[29] X. Li, Y. Zhang, H. Cheng, et al., "Student achievement prediction using deep neural network from multi-source campus data," *Complex Intell. Syst*., vol. 8, pp. 5143–5156, 2022. doi: 10.1007/s40747-022-00731-8.

[30] P. Hai-tao, F. Ming-qu, Z. Hong-bin et al., "Predicting academic performance of students in Chinese-foreign cooperation in running schools with graph convolutional network," *Neural Comput & Applic*, vol. 33, pp. 637–645, 2021. doi: 10.1007/s00521-020-05045-9.

[31] Sadi Mohammad, Ibrahim Adnan Chowdhury, Niloy Roy, Md. Nazim Hasan, Dip Nandi, "Investigation of Student Dropout Problem by Using Data Mining Technique", International Journal of Education and Management Engineering, Vol.13, No.5, pp. 43-61, 2023.

[32] R. Ahuja, A. Jha, R. Maurya, and R. Srivastava, "Analysis of Educational Data Mining," In *Harmony Search and Nature Inspired Optimization Algorithm*s, Springer: Singapore, pp. 897–907, 2019. doi: 10.1007/978-981-13-0761-4_85.

[33] T. Omar, A. Alzahrani, M. Andzohdy, "Clustering approach for analyzing the student's efficiency and performance based on data," *Journal of Data Analysis and Information Processing*, vol. 8, no. 03, pp. 171–182, 2020. doi: 10.4236/jdaip.2020.83010.

[34] Moez Ali, Clustering in Machine Learning: 5 Essential Clustering Algorithms. URL: https://www.datacamp.com/blog/clustering-in-machine-learning-5-essential-clustering-algorithms.

[35] M. Wang, and X. Wei, "Research on Logistics Center Location-Allocation Problem Based on Two-Stage K-Means Algorithms," *Advances in Computer Science for Engineering and Education*, Hu, Z., Petoukhov, S., Dychka, I., He, M. (Eds.), Springer International Publishing, vol. 1247, pp. 52-62, 2021. doi: 10.1007/978-3-030-55506-1_5.

[36] X. Li, Y. Zhang, H. Cheng, F. Zhou and B. Yin, "An Unsupervised Ensemble Clustering Approach for the Analysis of Student Behavioral Patterns," *IEEE Access*, vol. 9, pp. 7076-7091, 2021, doi: 10.1109/ACCESS.2021.3049157.

[37] C. Maithri, H. Chandramouli, "Parallel DBSCAN Clustering Algorithm Using Hadoop Map-reduce Framework for Spatial Data", *International Journal of Information Technology and Computer Science*, vol. 14, no. 6, pp. 1-12, 2022.

[38] Mohamed Nafuri, Ahmad Fikri, Nor Samsiah Sani, Nur Fatin Aqilah Zainudin, Abdul Hadi Abd Rahman, and Mohd Aliff, "Clustering Analysis for Classifying Student Academic Performance in Higher Education," *Applied Sciences*, vol. 12, no. 19, pp. 9467, 2022. doi: 10.3390/app12199467.

[39] Denon Arthur Richmond Gono, Bi Tra Goore, Yves Tiecoura, and Kouame Abel Assielou, "Multi-relational Matrix Factorization Approach for Educational Items Clustering," *International Journal of Information and Education Technology*, vol. 13, no. 1, pp. 42-47, 2023. doi: 10.18178/ijiet.2023.13.1.1778.

[40] Aydin Sinan, and Ozturk Aylin, “Segmenting Learners in Online Learning Environments,” In *The Online, Open and Flexible Higher Education Conference*, pp. 29-39, 2015.
[41] D.S. Hooda, and Vivek Raich, *Fuzzy Logic Models. An Introduction*, U.K., Oxfird: Alpha Science International Ltd, 2017.
[42] A.R. Fayek, “Fuzzy Logic and Fuzzy Hybrid Techniques for Construction Engineering and Management,” *Journal of Construction Engineering and Management*, vol. 146, no. 7, pp. 1-12, 2020. doi: 10.1061/(ASCE)CO.1943-7862.0001854.
[43] A. Heni, I. Jdey and H. Ltifi, “K-means and fuzzy c-means fusion for object clustering,” In *8th International Conference on Control, Decision and Information Technologies (CoDIT)*, Istanbul, Turkey, pp. 177-182, 2022. doi: 10.1109/CoDIT55151.2022.9804078.
[44] Student Performance Prediction. URL: https://www.kaggle.com/datasets/souradippal/student-performance-prediction.
[45] Tengku Zatul Hidayah Tengku Petra, and Mohd Juzaiddin Ab Aziz, “Analysing Student Performance In Higher Education Using Fuzzy Logic Evaluation,” *I. J. scientific & technology research*, vol. 10, no. 01, pp. 322-327, 2021.
[46] Riska Ekawita, Nori Wirahmi, Elfi Yuliza, "Project Based Learning of Free-Falling Objects in Physics: Microcontroller-Based Multi-Sensor Test in Granular Flow Rate Measurement", International Journal of Engineering and Manufacturing, Vol.14, No.5, pp. 1-13, 2024.
[47] S.V. Balovsyak, and Kh. S. Odaiska, “Automatic Determination of the Gaussian Noise Level on Digital Images by High-Pass Filtering for Regions of Interest,” *Cybernetics and Systems Analysis*, vol. 54, no. 4, pp. 662-670, 2018.
[48] S.V. Balovsyak, O.V. Derevyanchuk, I.M. Fodchuk, O.P. Kroitor, Kh.S. Odaiska, O.O. Pshenychnyi, A. Kotyra, and A. Abisheva, “Adaptive oriented filtration of digital images in the spatial domain,” *Proc. SPIE 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments*, vol. 11176, pp 111761A-1–111761A-6, 2019. doi: https://doi.org/10.1117/12.2537165.
[49] S.V. Balovsyak, O.V. Derevyanchuk, H.O. Kravchenko, O.P. Kroitor, and V.V. Tomash, “Computer system for increasing the local contrast of railway transport images,” *Proc. SPIE, Fifteenth International Conference on Correlation Optics*, vol. 12126, pp. 121261E1-7, 2021. doi: 10.1117/12.2615761.
[50] Serhiy Balovsyak, Oleksandr Derevyanchuk, Vasyl Kovalchuk, Hanna Kravchenko, Yuriy Ushenko, and Zhengbing Hu, “STEM Project for Vehicle Image Segmentation Using Fuzzy Logic,” *International Journal of Modern Education and Computer Science*, vol. 16, no. 2, pp. 45-57, 2024.
[51] Stanford Cars Dataset. URL: https://www.kaggle.com/code/pamhohhgkgm/stanford-cars-dataset-a-quick-look-up.
[52] S.V. Balovsyak, O.V. Derevyanchuk, and I.M. Fodchuk, “Method of calculation of averaged digital image profiles by envelopes as the conic sections,” *Advances in Intelligent Systems and Computing*, Hu Z., Petoukhov S., Dychka I., He M. (Eds.), Springer International Publishing, vol. 754, pp. 204–212, 2019. doi: 10.1007/978-3-319-91008-6_21.
[53] O.V. Derevyanchuk, H.O. Kravchenko, Y.V. Derevianchuk, and V.V. Tomash, “Recognition images of broken window glass,” *Proceedings of SPIE*, vol. 12938, pp. 210-213, 2024. doi: 10.1117/12.3012995.
[54] A. K. M. Fahim Rahman, Mostofa Rakib Raihan, and S.M. Mohidul Islam, “Pedestrian Detection in Thermal Images Using Deep Saliency Map and Instance Segmentation,” *International Journal of Image, Graphics and Signal Processing*, vol. 13, no. 1, pp. 40-49, 2021.
[55] O. Berezsky, M. Zarichnyi, and O. Pitsun, “Development of a metric and the methods for quantitative estimation of the segmentation of biomedical images,” *Eastern-European Journal of Enterprise Technologies*, vol. 6, no. 4(90), pp. 4–11, 2017. doi: 10.15587/1729-4061.2017.119493.
[56] Ihor Tereikovskyi, Zhengbing Hu, Denys Chernyshev, Liudmyla Tereikovska, Oleksandr Korystin, and Oleh Tereikovskyi, “The Method of Semantic Image Segmentation Using Neural Networks,” *International Journal of Image, Graphics and Signal Processing*, vol. 14, no. 6, pp. 1-14, 2022.
[57] Kama Ramudu, Gajula Laxmi Bhavani, Manabolu Nishanth, Akula Prakash Raj, and Vamshika Analdas, “Level Set Segmentation of Images using Block Matching Local SVD Operator based Sparsity and TV Regularization,” International Journal of Image, Graphics and Signal Processing, vol. 15, no. 2, pp. 47-58, 2023.
[58] Serhiy Balovsyak, Oleksandr Derevyanchuk, Vasyl Kovalchuk, Hanna Kravchenko, and Maryna Kozhokar, “Face Mask Recognition by the Viola-Jones Method Using Fuzzy Logic,” *International Journal of Image, Graphics and Signal Processing*, vol. 16, no. 3, pp. 39-51, 2024.
[59] Deepa Indrawal, and Archana Sharma, “Multi-Module Convolutional Neural Network Based Optimal Face Recognition with Minibatch Optimization,” *International Journal of Image, Graphics and Signal Processing*, vol.14, no.3, pp. 32-46, 2022.
[60] Freepik. URL: https://www.freepik.com.
[61] Caltech Vision Lab. Datasets. URL: http://www.vision.caltech.edu/datasets.
[62] O.V. Derevyanchuk, V.I. Kovalchuk, V.M. Kramar, H.O. Kravchenko, D.V. Kondryuk, A.V. Kovalchuk, and B.V.Onufriichuk, “Implementation of STEM education in the process of training of future specialists of engineering and pedagogical specialties,” *Proceedings of SPIE*, vol. 12938, pp. 214-217, 2024. doi: 10.1117/12.3012996.
[63] V. Kovalchuk, A. Androsenko, O. Derevyanchuk, N. Volkova, and Y. Piven, “Development of pedagogical skills of students of technology and pedagogical specialties using STEM technologies,” *Edelweiss Applied Science and Technology*, vol. 8, no. 4, pp. 498-506, 2024. https://doi.org/10.55214/25768484.v8i4.1125

Authors' Profiles

Oleksandr Derevyanchuk: Received the Master of Engineering degree (1999) and Ph.D. of Physics and Mathematics (2014) at the Yuriy Fedkovych Chernivtsi National University. He is a Candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Professional and Technological Education and General Physics, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine.

Research Interests: Education Technology, Educational Data Mining, Digital Processing of Signals and Images, Programming, Pattern Recognition, Artificial Neural Networks.

Zhengbing Hu: Prof. , Deputy Director, International Center of Informatics and Computer Science, Faculty of Applied Mathematics, National Technical University of Ukraine "Kyiv Polytechnic Institute", Ukraine. Adjunct Professor, School of Computer Science, Hubei University of Technology, China. Visiting Prof., DSc Candidate in National Aviation University (Ukraine) from 2019. Major research interests: Computer Science and Technology Applications, Artificial Intelligence, Network Security, Communications, Data Processing, Cloud Computing, Education Technology.

Serhiy Balovsyak: Graduated from Chernivtsi State University (1995). In 2018, he defended his doctoral dissertation in the specialty "Computer systems and components".

Currently position – associate professor at the Department of Computer Systems and Networks of Yuriy Fedkovych Chernivtsi National University, Ukraine.

Research Interests: Education Technology, Data Mining, Digital Processing of Signals and Images, Programming, Pattern Recognition, Artificial Neural Networks.

Serhii Holub: Doctor of technical sciences, Professor, Head of the department of software of automated systems of the Cherkasy State University of Technology.

Research Interests: Agent programming, Data Mining and Analysis, Computer Vision and Pattern Recognition.

Hanna Kravchenko: Teacher of the State Educational Establishment "Chernivtsi transport college", Chernivtsi, Ukraine.

Research Interests: Education Technology, Educational Data Mining, Digital Processing of Signals and Images, Programming, Pattern Recognition, Artificial Neural Networks.

Iryna Sapsai: M.Sc. of Physics M. Dragomanov National Pedagogical University of Kyiv (2009). PhD in Pedagogic "Physics", M. Dragomanov National Pedagogical University of Kyiv (2014). Current position – Lecturer at the Department of Science and Mathematics Education and Technology, Institute of Postgraduate Education, Borys Grinchenko Kyiv Metropolitan University.

Research Interests: Physics education, Optics, Teaching and Learning, Pedagogy and Education, E-learning.

How to cite this paper: Oleksandr Derevyanchuk, Zhengbing Hu, Serhiy Balovsyak, Serhii Holub, Hanna Kravchenko, Iryna Sapsai, "Complex of Specialized Methods of Educational Data Mining for the Training of Vocational Education Teachers", International Journal of Modern Education and Computer Science(IJMECS), Vol.17, No.1, pp. 28-46, 2025. DOI:10.5815/ijmecs.2025.01.03