

A Merging Method to Discretizing and Grouping the Input Factors of ANOVA Model while Research of Time Dynamic of the Students Intelligence Quotient

Anastasiia Timofeeva^a, Tatiana Avdeenko^a and Olga Razumnikova^a

^a Novosibirsk State Technical University, 20, Karla Marksa ave., Novosibirsk, 630073, Russia

Abstract

In present work we study, with use of multivariate ANOVA model, the influence of independent factors such as year, faculty, gender, on the indicators of students' general intelligence (IQ) with a sample collected in 1991-2013 at the Novosibirsk State Technical University. The peculiarity of models of this type is that the response is a quantitative variable, and the input features must be qualitative. Therefore, first, the problem of converting quantitative features into categorical ones (discretization) arises, second, with a large number of levels of input qualitative features their grouping is required. If the variables are strongly correlated, then both tasks should be solved simultaneously. In this case, the optimal quality of the model should be ensured in accordance with a certain criterion. Existing methods for the features type conversion are limited to one of the tasks (discretization or grouping) and often do not take into account the relationships between the features. Therefore, an original approach is proposed that allows solving the problem and interpreting the results obtained.

Keywords ¹

intelligence, Flynn effect, analysis of variance, discretization, grouping, interaction effect

1. Introduction

The intelligence quotient (IQ) is associated with the quality of people's life and its duration. Thus, a study carried out in Scotland, and presented in [1], showed that the probability of surviving to 76 years depends significantly on the IQ level detected at the age of 11 years. The studies carried out were based on IQ measurements of 2792 children in 1932 in Scotland, born in 1921, the fate of 79.9% (2230) of which was subsequently tracked. One possible explanation for these findings is that intelligence enhances people's health care by helping them to acquire problem-solving skills that are useful for preventing chronic diseases, accidental injuries, and for adhering to complex treatment schemes.

There are other reasons for the influence of IQ on the quality of life, and, as a consequence, on its duration. Thus, in the article [2], based on a survey of 6870 participants living in England in 2007, a positive correlation was found between the level of verbal IQ and the feeling of happiness. People with lower IQ were found to be less happy than people with higher IQ.

On the other hand, recent studies show that high intelligence is associated with increased anxiety and stress, and can also cause chronic depression [3]. It is also noted that gifted people are more likely than others to suffer from asthma and allergies [4], and are also susceptible to autoimmune diseases [5].

SLET-2020: International Scientific Conference on Innovative Approaches to the Application of Digital Technologies in Education, November 12-13, 2020, Stavropol, Russia

EMAIL: a.timofeeva@corp.nstu.ru (Anastasiia Timofeeva); avdeenko@corp.nstu.ru (Tatiana Avdeenko); razoum@mail.ru (Olga Razumnikova)

ORCID: 0000-0001-9900-026X (Anastasiia Timofeeva); 0000-0002-8614-5934 (Tatiana Avdeenko); 0000-0002-7831-9404 (Olga Razumnikova)



© 2020 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

All of the above indicates the relevance of conducting research based on the accumulation and multivariate statistical analysis of intelligence indicators and its relationship with various time and demographic factors. A special place in these studies is played by the phenomenon of a gradual increase of IQ in the 20th century, known as the "Flynn effect". The effect was observed in different countries and for different categories of test subjects [6]. For example, in [7] it was concluded that a representative sample of Americans from 1932 to 1978 every year coped better and better with IQ tests, while the overall increase in average IQ over 46 years was 13.8 points. However, since the end of the 20th century, the reverse temporal dynamics of IQ (or the anti-Flynn effect) began to be observed, the reasons for which remain unclear [8, 9, 10].

At the Novosibirsk State Technical University for 23 years from 1991 to 2013, the intelligence of 1st year students was tested according to the Amthauer method. The sample consisted of 3,677 students of both sexes from various departments of the university in the natural science, technical, and humanitarian fields of knowledge. As a result of the analysis of these data, it becomes possible to establish the influence of factors such as gender and faculty on the IQ of students, as well as to study the temporal dynamics of changes in the IQ of students studying at a Russian university.

For the research, a multivariate analysis of variance was chosen, in which the response (dependent variable) is the final IQ of students, measured on a scale of relationships. Categorical independent features, measured on a nominal scale, are student gender, faculty, and year of study. The aim of the study is to identify the influence of independent factors on the dependent variable - a quantitative indicator of IQ. It is important to assess not only the impact of factors separately, but also their interactions.

When conducting long-term studies of intelligence, it is not always possible to develop an experimental design that makes it possible to obtain optimal estimates of the effects in the ANOVA model, since it is difficult to ensure such conditions under which a similar sample population of individuals would be surveyed every year. In this regard, the analyzed sample is characterized by an uneven distribution of students across faculties and survey years, i.e. in one year, students from one subset of faculties were surveyed, and in the next year, students from another subset. To construct an acceptable analysis of variance model under these conditions, in present paper the method of agglomerative discretization and grouping of input features was developed, investigated and applied.

The article has the following structure. Section 2 provides an overview of the existing discretization methods, substantiates the development of a new method. Section 3 presents the quality criteria investigated in the article for constructing optimal discretization. In Section 4, we describe the ANOVA model used. Section 5 describes the developed discretization algorithm. Section 6 contains the results of the studies of the proposed approach, and section 7 contains their interpretation for solving the multifactor task of studying the IQ of students. In section 8 we provide a conclusion on the work.

2. Overview of discretization methods

A good overview of the current state of research on discretization methods is presented in [11, 12]. If the transformation of a quantitative attribute into a qualitative one is carried out in such a way as to ensure the best agreement with the response, then we are talking about supervised discretization. This task can be solved using top-down (divisive) discretization techniques or bottom-up (agglomerative) techniques. In the first case, a gradual division into intervals occurs, and in the second, the intervals are merged. At each step of such algorithms, an evaluation function is calculated that characterizes the quality of the division into intervals. In addition, the stop criterion is important, which determines that further partition (merging) does not make sense.

For example, an efficient recursive partitioning algorithm MDLP [13] evaluates the quality based on information gain based on entropy, and the stopping criterion is derived from the principle of minimum length description. The chi-square statistic is popular in the agglomerative merging problem. Algorithms such as ChiMerge [14], Chi2 [15] were built on its basis. Both approaches are designed for classification tasks, that is, they assume that the response is categorical. Therefore, their application to transform a set of input variables in the construction of ANOVA models requires discretizing the response, which can lead to the loss of significant information.

Another group of discretization methods, the so-called wrapping methods, focuses on the quality of the estimated model. Thus, these methods simultaneously solve the learning problem. The existing algorithms are built for classifiers, for example, such simple ones as a majority class voting classifier [16], or more general classifiers such as Naive Bayes [17].

Compared to the problem of discretization, the grouping problem has not been studied so deeply in the literature. A fairly complete overview of grouping methods is presented in [18]. Many commercial data mining packages suggest excluding variables that have too many categories. This approach, however, cannot be considered acceptable in cases where the research interest is to assess the effects of just such variables. Effective grouping methods allow for fewer, more informative categories. This can be done by Sequential Forward Selection method [19]. It is a greedy that initializes a group with the best category and then iteratively adds new categories to this first group. Decision tree algorithms often solve the grouping problem with a greedy heuristic based on bottom-up categorization. The CHAID algorithm [20] uses this greedy approach with a criterion close to the ChiMerge criterion [14]. In [18], a new method of grouping MODL based on the Bayesian approach was proposed, as well as the discretization method MODL [21]. It searches for the most likely grouping model for the given dataset. Optimization is done using a greedy bottom-up algorithm.

Thus, most of the existing supervised discretization algorithms are designed to solve classification problems, that is, for categorical response. They are mainly aimed at improving the quality of predicting the response (quality of classification) [[22],[23]]. Moreover, they are usually univariate. In this regard, it seems relevant to develop an algorithm for the optimal categorization of input features, taking into account their interrelationships, to build a model of analysis of variance. Here categorization includes two tasks: discretization of quantitative variables and grouping of nominal features. Due to the specifics of the practical task, the construction of response predictions is secondary, therefore, the use of criteria such as cross-validation in order to assess the quality of the model and avoid overfitting is limited. The main task was to obtain and interpret estimates of the effects of influencing factors. As a result, we had to resort to goodness-of-fit criteria.

3. Goodness of fit criteria

Most often, the quality of a regression model is judged by the coefficient of determination, calculated as

$$R^2 = 1 - \frac{ESS}{TSS},$$

where ESS is residual sum of squares of the model, TSS is total sum of squares of the model. However, this indicator has an obvious drawback. With increasing complexity of the model (including new variables), it is possible to better describe the response, thereby decreasing ESS and increasing R^2 . However, the number of the degrees of freedom decrease, which is in no way taken into account when calculating the coefficient of determination.

To check the significance of the model, the F-statistic is used, calculated as

$$F = \frac{R^2}{1 - R^2} \frac{N - m}{m - 1},$$

where N is the number of observations, m is the number of estimated parameters. It takes degrees of freedom into account, so the increase in model complexity must be offset by a sufficient decrease in the residual sum of squares.

Akaike information criterion is often used in the problem of feature selection, for example, in the stepwise regression procedure. It provides a trade-off between goodness of fit and complexity of the model (number of parameters). The Akaike criterion is calculated as follows.

$$AIC = 2m + N \log ESS.$$

It should be borne in mind that with a very large number of categories, building good groupings is difficult because of the risk of overfitting the model. In the extreme case, to avoid overfitting, efficient grouping methods can combine all values into one group, thereby excluding the variable from consideration. In order to prevent such a situation, the stopping criterion must include a condition for the minimum number of categories (for example, two).

4. ANOVA model

For research, the following model of analysis of variance was formulated:

$$y_{kji} = \mu + \alpha_k + \beta_t + \gamma_j + (\alpha\beta)_{kt} + (\alpha\gamma)_{kj} + (\beta\gamma)_{tj} + (\alpha\beta\gamma)_{kjt} + \varepsilon_{kji}, \quad (1)$$

where y_{kji} – i -th observed value corresponding to the IQ level for a student of the k -th sex of the j -th faculty in the year t , α_k is the effect of the k -th sex ($k=1$ for male, $k=2$ for female), β_t is the effect of the year t , $t=1991, \dots, 2013$, γ_j is the effect of the j -th faculty, $j=1, \dots, 11$, $(\alpha\beta)_{kt}$ is the interaction effect of the k -th sex and t -th year, $(\alpha\gamma)_{kj}$ is the effect of the interaction of the k -th year and the j -th faculty, $(\beta\gamma)_{tj}$ is the effect of the interaction of the j -th faculty and the t -th year, $(\alpha\beta\gamma)_{kjt}$ is the effect of the interaction of the k -th sex, t -th year and j -th faculty, ε_{kji} is a random error.

It is impossible to estimate all the effects in model (1). Usually they resort to reduction. This estimates paired comparisons with some baseline, for example, $(\alpha_2 - \alpha_1)$ is the influence of female versus male. The first levels of factors are taken as the baseline levels.

The distribution of the studied students is uneven over the years (see table 1). There is a close relationship between the variables Faculty and Year. The chi-square statistic is 8092.9, which indicates a significant correlation at 0.1% significance level. Nevertheless, it should be borne in mind that the original contingency table has a very large dimension (220 degrees of freedom), and, as a consequence, cells with a small number of observations, which negatively affects the correctness of the chi-square test. For confirmation, the correlation ratio was calculated, showing the influence of the faculty for the year. It is 0.192 (F-statistic is equal to 86.9), which also speaks of a significant connection at 0.1% significance level.

Table 1

The ratio of faculties and survey years in the sample

Faculty	Abbreviation	Survey years
automation and computer engineering	ACEF	1994, 1995, 1997-1999, 2001, 2002, 2004
mechanical engineering and technologies	MTF	1993, 1995-2001, 2013
radio engineering and electronics	REEF	1992-2002, 2006, 2008-2010
business	FB	1997-1999, 2001-2005
humanity education	HEF	1994, 1998-2004, 2006-2008, 2010-2012
aircraft engineering	AEF	1993, 1994, 1997, 2000, 2002-2006
mechatronics and automation	MAF	1994, 1995, 1997, 1998
applied mathematics and computer science	AMCSF	1995-2004
physical engineering	PEF	1991, 1994-1996, 1998-2000, 2003, 2004, 2006, 2009
power engineering	PEF	1994, 1995, 1997-1999, 2001, 2002, 2004
natural sciences	NSF	2007

Consequently, it is impossible to assess all the effects of faculty and year interactions in order to separate the effect of student specialization from the time trend. Therefore, it is necessary to discretize the Faculty and Year variables in such a way as to ensure the optimal quality of estimation of the model, which includes interaction effects.

5. The developed algorithm

The algorithm is developed for the case when it is required to discretize one quantitative variable and group one categorical variable, and the variables are highly correlated. It can be extended to the case when there are more than two variables, but with a large number of variables and levels the curse of dimension arises.

The pseudocode of the algorithm for the optimal categorization of input features, taking into account their interrelationships for constructing an analysis of variance model, is shown in Figure 1.

```

Q = quality( $x_1, x_2$ )

repeat
  for  $k = 0$  to  $(T_0 - 1)$  do
    if  $k \neq 0$  &  $T_0 > 2$  then  $x'_1 = \text{merge}(x_1, k, k + 1)$  else  $x'_1 = x_1$ 
    if  $k \neq 0$  then  $Q_1(0, 0, k) = \text{quality}(x'_1, x_2)$ 
    for  $i = 1$  to  $(K_0 - 1)$  do
      for  $j = i + 1$  to  $K_0$  do
        if  $K_0 > 2$  then  $x'_2 = \text{merge}(x_2, i, j)$  else  $x'_2 = x_2$ 
         $Q_1(i, j, k) = \text{quality}(x'_1, x'_2)$ 
      end for
    end for
  end for
   $Q' = \underset{i, j, k}{\text{opt}} Q_1$ 
  if  $Q' \leq Q$  then break
   $Q := Q'$ 
   $(i^*, j^*, k^*) = \underset{i, j, k}{\text{arg opt}} Q_1$ 
  if  $k^* \neq 0$  then  $x_1 := \text{merge}(x_1, k^*, k^* + 1)$ ,  $T_0 := T_0 - 1$ 
  if  $i^* \neq 0$  then  $x_2 := \text{merge}(x_2, i^*, j^*)$ ,  $K_0 := K_0 - 1$ 
end repeat
return  $x_1, x_2$ 

```

Figure 1: Pseudocode of the developed algorithm

Input: raw data including response values, quantitative factor x_1 with T_0 levels, and qualitative factor x_2 with K_0 levels.

The thresholds were selected simultaneously for two variables by the agglomerative merging method. The initial model was built taking into account all available levels of factors. Further, one boundary between the levels was successively removed. For a categorical variable, all possible pairs

of factor levels were considered, for a quantitative variable, only adjacent values. In addition, such an option was considered when the levels were not combined. It was assigned an index 0 according to the variable for which the levels were not combined. This is done in case the optimal solution is to combine levels in only one of the variables. If the best value of the quality index corresponding to the optimal solution was achieved, the levels were combined. Then the procedure was repeated until an improvement was obtained.

The function $quality(x_1, x_2)$ returns an indicator of the quality of fitting an ANOVA model of the form (1) (determination coefficient, F-statistic, AIC) depending on the input data.

The function $merge(x, i, j)$ combines the levels i, j of a variable x so that the number of levels is reduced by one. If the input variable included K levels, then the function returns the transformed variable with $(K - 1)$ levels numbered from 1 to $(K - 1)$.

Since the optimization of the goodness-of-fit criteria can go in different directions (for the determination coefficient and F-statistic it is maximization, for AIC it is minimization), we denoted the optimal value as opt . Wherein $Q' p Q | Q' = Q$ means that Q' is no better than Q .

6. Results

The choice of the determination coefficient as an evaluation criterion did not give any results, since the original partition provided the minimum residual sum of squares. As expected, any merging of intervals led to a decrease in the determination coefficient.

The use of the F-statistic, on the contrary, led to the fact that at each step there was an improvement in the values of the evaluation function. Thus, the work of the algorithm ended only when the intervals could no longer be combined, that is, when there were two categories left for each feature. The faculty of AMCSF stood out in a separate group, as well as 1991. The results of evaluating such a model indicate one significant effect - on the AMCSF, compared to the rest of the faculty, IQ is 6.3 points higher (significant at the 1% level).

The use of the Akaike information criterion made it possible to obtain more interesting results. When applying the algorithm, three groups of faculties were distinguished. From table 1 it is clearly seen that there are years in which some faculties were not covered by the study. This problem was partially solved by discretizing the variable Year. Table 2 shows the proportions of students of faculties of three groups studied in a given range of years. For example, for the first group, there were no periods left when the faculties of this group were not covered by the study. Nevertheless, there is a gap for the second group of faculties in 2009, and for the third - in 2008 and 2010-2013. Therefore, it was not possible to estimate the corresponding effects.

Table 2

Shares of students of faculty groups in the total number of students studied in a given range of years

Survey years	1 st group	2 nd group	3 rd group
1991-1996	0.697	0.073	0.230
1997	0.556	0.148	0.296
1998-1999	0.661	0.195	0.144
2000	0.203	0.228	0.568
2001	0.296	0.245	0.460
2002	0.579	0.274	0.147
2003-2005	0.306	0.320	0.373
2006-2007	0.394	0.518	0.088
2008	0.600	0.400	0
2009	0.759	0	0.241
2010-2013	0.556	0.444	0

After discretizing the variables, a model was estimated describing the dependence of IQ on gender, faculty, and year and on their interactions. It turned out that gender has an insignificant effect on the level of intelligence. Therefore, the gender factor was eliminated and the model was re-estimated.

Table 3 provides a summary table with the values of the F-statistic and p-value for 1% significance level. Almost all the effects of the variable Year turned out to be significant at the 5% or 10% level. For the base year, the effect of the faculties of the second group compared to the first was -5.3 and is significant at the 1% level. The effect of faculties of the third group compared to the first for the base year is estimated as 1.2 and is significant at the 10% level. Most of the interactions between the year and the faculty were significant. The general average is estimated at 112.3 .

Table 3
The significance of factors

Factor	Degrees of freedom	F-statistics	Critical F-value	p-value
Year	10	14.88	0.55	$<2 \cdot 10^{-16}$
Faculty	2	244.41	0.63	$<2 \cdot 10^{-16}$
Faculty:Year	17	11.38	0.53	$<2 \cdot 10^{-16}$

Figure shows the predicted IQ values by year and depending on the group of faculties.

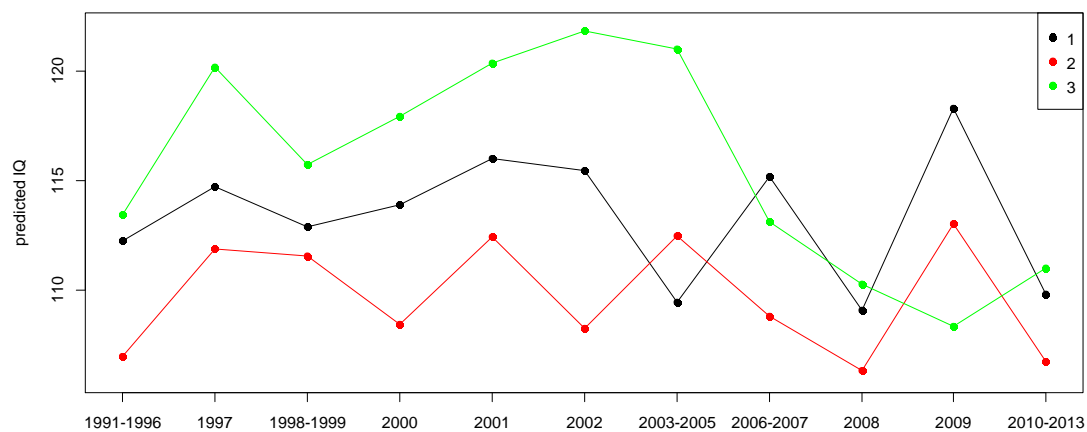


Figure 2: Model estimation results

7. Interpreting the Results

From the point of view of specialization, the distinguished groups of faculties can be divided as follows. The first group is technical and economic faculties, the second is humanitarian and applied faculties, and the third is physics and mathematics. The latter group, on average, is characterized by the highest level of intelligence. Although since 2006 the IQ has dropped and has become comparable to the level of intelligence of students in other faculties. But during this period, a group of students with a physical and mathematical specialization was observed very little (see Table 1): only PEF 2006 (22 students) and 2009 (14 students). Therefore, the decline in IQ may be due to the non-representativeness of the sample.

In the 2000s, there was instability of IQ indicators among students of technical and economic specialization. Growth period 2006-2007 can be explained by the fact that in 2007 only the NSF was observed from this group, which was characterized by higher IQ indices.

For students of humanitarian and applied specialties from 2000 to 2005 in general, there was an increase in intelligence indicators, and then a sharp decline began in 2006-2008. In 2009, the faculties

of this group were not studied, so the interaction effect could not be estimated, and the IQ forecast is based only on the main effects. This explains the sharp increase in the IQ forecast in 2009, which cannot be considered reasonable.

8. Conclusion

Thus, in this work, an analysis of variance model was constructed to study the influence of input factors on the IQ of students. To build a qualitative model, taking into account the specifics of the collected data, a new agglomerative method for discretizing and grouping input features was developed and tested. The interpretation of the obtained estimation results is carried out. In practice, the obtained results of interpretation can be used in the construction of individual educational trajectories, which is one of the key problems of the modern digital educational environment [24].

Future work involves the improvement of the developed algorithm in terms of finding the optimal solution, as well as the development of alternative models for the study of students' IQ with subsequent comparison of the results.

9. Acknowledgements

The research is supported by Ministry of Science and Higher Education of Russian Federation (project No. FSUN-2020-0009).

10. References

- [1] L.J. Whalley, Longitudinal Cohort Study of Childhood IQ and Survival up to Age 76, *Bmj*. 322 (2001) 819–819. doi:10.1136/bmj.322.7290.819.
- [2] A. Ali, et al., The Relationship between Happiness and Intelligent Quotient: the Contribution of Socio-Economic and Clinical Factors, *Psychological Medicine*. 43 (2012) 1303–1312. doi:10.1017/s0033291712002139.
- [3] A. M. Penney, et al., Intelligence and Emotional Disorders: Is the Worrying and Ruminating Mind a More Intelligent Mind? *Personality and Individual Differences*. 74 (2015) 90–93. doi:10.1016/j.paid.2014.10.005.
- [4] E. A. Hildreth, Some Common Allergic Emergencies, *Medical Clinics of North America*. 50 (1966) 1313–1324. doi:10.1016/s0025-7125(16)33127-3.
- [5] C. P. Benbow, Intellectually Gifted Students Also Suffer from Immune Disorders, *Behavioral and Brain Sciences*. 8 (1985) 442. doi:10.1017/s0140525x00001059.
- [6] J. R. Flynn, Searching for Justice: The Discovery of IQ Gains over Time, *American Psychologist*. 54 (1999) 5–20. doi:10.1037/0003-066X.54.1.5.
- [7] J. R. Flynn, The Mean IQ of Americans: Massive Gains 1932 to 1978, *Psychological Bulletin*. 95 (1984) 29–51. doi:10.1037/0033-2909.95.1.29
- [8] B. Bratsberg, O. Rogeberg, Flynn effect and its reversal are both environmentally caused, *PNAS*. 115 (2018) 6674–6678. doi:10.1073/pnas.1718793115
- [9] E. Dutton, D. van der Linden, R. Lynn, The negative Flynn effect: A systematic literature review, *Intelligence*. 59 (2016) 163–169. doi:10.1016/j.intell.2016.10.002
- [10] J.R. Flynn, M. Shayer, IQ decline and Piaget: Does the rot start at the top? *Intelligence*. 66 (2018) 112–121. doi:10.1016/j.intell.2017.11.010
- [11] S. Kotsiantis, D. Kanellopoulos, Discretization techniques: A recent survey, *GESTS International Transactions on Computer Science and Engineering*. 32 (2006) 47–58. doi:10.1.1.109.3084.
- [12] S. Garcia, J. Luengo, J. A. Sáez, V. Lopez, F. Herrera, A survey of discretization techniques: Taxonomy and empirical analysis in supervised learning, *IEEE Transactions on Knowledge and Data Engineering*. 25 (2012) 734–750. doi: 10.1109/TKDE.2012.35

- [13] U. Fayyad, K. Irani, Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning, in: Proceedings of the 13th Int'l Joint Conf. Artificial Intelligence (IJCAI), 1993, pp. 1022–1029.
- [14] R. Kerber, ChiMerge: Discretization of Numeric Attributes, in: Proceedings of the Nat'l Conf. Artificial Intelligence Am. Assoc. for Artificial Intelligence (AAAI), 1992, pp. 123–128.
- [15] H. Liu, R. Setiono, Feature Selection via Discretization, IEEE Trans. Knowledge and Data Eng.. 9 (1997) 642–645. doi: 10.1109/69.617056
- [16] D. Ventura, T. R. Martinez, BRACE: A Paradigm for the Discretization of Continuously Valued Data, in: Proceedings of the Seventh Ann. Florida AI Research Symp. (FLAIRS), 1994, pp. 117–121.
- [17] M.J. Pazzani, An Iterative Improvement Approach for the Discretization of Numeric Attributes in Bayesian Classifiers, in: Proceedings of the First Int'l Conf. Knowledge Discovery and Data Mining (KDD), 1995, pp. 228–233.
- [18] M. A Boullé, Grouping method for categorical attributes having very large number of values, in: International Workshop on Machine Learning and Data Mining in Pattern Recognition, Springer, Berlin, Heidelberg, 2005. pp. 228–242.
- [19] G. Cestnik, I. Kononenko, I. Bratko, A knowledge-elicitation tool for sophisticated users, in: Progress in Machine Learning, SigmaPress, Wihnslow, England, 1987.
- [20] G. V. Kass, An exploratory technique for investigating large quantities of categorical data, Journal of the Royal Statistical Society: Series C (Applied Statistics). 29 (1980) 119–127. doi: 10.2307/2986296
- [21] M. Boullé, MODL: a Bayes optimal discretization method for continuous attributes, Machine learning. 65 (2006) 131–165. doi: 10.1007/s10994-006-8364-x
- [22] W. Huang, Y. Pan, J. Wu, Supervised discretization for optimal prediction, Procedia Computer Science. 30 (2014) C. 75–80. doi: 10.1016/j.procs.2014.05.383
- [23] J. L. Lustgarten, V. Gopalakrishnan, H. Grover, S. Visweswaran, Improving classification performance with discretization on biomedical datasets, in: AMIA annual symposium proceedings, American Medical Informatics Association, 2008. pp. 445–449.
- [24] D. Parfenov, V. Zaporozhko, M. Lapina, D. Sora, Development and Research of Algorithms for the Formation the Individual Educational Trajectories of Students in the Digital Educational Platform, CEUR Workshop Proceedings. 2494 (2019).