UDC 004.9 https://doi.org/10.31891/csit-2023-1-11

## Khrystyna ZUB, Pavlo ZHEZHNYCH

Lviv Polytechnic National University

# MACHINE LEARNING BOOSTING METHODS FOR PREDICTION A HIGHER EDUCATION INSTITUTIONS ENTRANT'S ADMISSIONS IN UKRAINE

There is a constant and growing need for higher education institutions (HEI) to provide proper and high-quality support for the admissions campaign through information systems and technologies. Labor market trends, unreliability and low-quality sources, and a large volume of admission rules can complicate the admission process for an applicant. As a result, there is a risk that the applicant will not be able to make the right choice and quality assessment of the chances of admission. So, this paper considers increasing the entrant's chances of making an effective decision at the stage of education program selection through the implementation of an information system. The efficiency of such systems is largely based on the accuracy of its intelligent components. This article investigates the effectiveness of machine learning (ML) boosting methods to solve the problem of admission prediction through binary classification tasks. We evaluate the accuracy of such ML methods as Gradient Boosting, Adaptive Boosting (AdaBoost), and eXtreme Gradient Boosting (XGBoost). For a more detailed assessment of the studied methods, a comparison with Support Vector Machine (SVM) and Logistic regression is also presented. The simulation was performed using «Orange» software. The work of the studied methods was simulated based on a sample of archival data comprising 9,657 records of full-time entrants of two faculties of Lyiv Polytechnic National University. The sample was randomly divided into training and test sets in a ratio of 80% to 20%. To ensure the reliability of the obtained result, the work of each of the studied methods was subjected to 10-fold cross-validation. Classification accuracy (AUC), Precision, Recall and F1 score performace indicators was used to analyze the results. It has been experimentally established that the highest accuracy is achieved when using XGBoost. The obtained results shows high accurate. This makes it possible to use the researched methods in the subsequent stages of building an information system to support the decision-making of applicants.

Keywords: admission, entrant, higher education institution (HEI), prediction, machine learning, boosting, information system.

## Христина ЗУБ, Павло ЖЕЖНИЧ

Національний університет «Львівська політехніка»

## БУСТИНГОВІ МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ УСПІШНОСТІ ВСТУПУ АБІТУРІЄНТІВ ЗВО УКРАЇНИ

Існує постійна та зростаюча потреба закладів вищої освіти (ЗВО) у забезпеченні належного та якісного супроводу вступної кампанії за допомогою інформаційних систем та технологій. Тенденції на ринку праці, ненадійність і неякісність джерел), велика кількість правил прийому можуть ускладнити процес вступу абітурієнта. Як наслідок, є ризик того, що абітурієнт не зможе зробити правильний вибір та якісно оцінити шанси на вступ. Тож, у даній роботі розглядається завдання підвищення шансів абітурієнта прийняти ефективне рішення на етапі вибору освітньої програми. Ефективність таких систем значною мірою базується на точності їх інтелектуальних компонентів. У цій статті досліджується ефективність бустингових методів машинного навчання для вирішення проблеми прогнозування вступу за допомогою завдань бінарної класифікації. Ми оцінюємо такі точність роботи таких методів машинного навчання, як Gradient Boosting, Adaptive Boosting (AdaBoost) і eXtreme Gradient Boosting (XGBoost). Для більш детальної оцінки досліджуваних методів також представлено порівняння з методом опорних векторнів і логістичною регресією. Моделювання проводилось за допомогою програмного забезпечення «Orange». Роботу досліджуваних методиів було змодельовано на основі вибірки архівних даних, яка склала 9657 записів даних абітурієнтів денної форми навчання двох навчально-наукових інститутів Національного університету «Львівська політехніка». Вибірку випадковим чином було розподілено на навчальну та тестову вибірки у співвідношенні 80% до 20%. Для забезпечення достовірності отриманого результату роботу кожного з досліджуваних методів піддавали 10-кратній крос-валідації. Для аналізу результатів використано такі показники точності як Classification accuracy (AUC), Precision, Recall, F1 score. Експериментально встановлено, що найвища точність досягається при використанні XGBoost. Отримані результати досить точні. Це дає можливість використовувати досліджувані методи на наступних етапах побудови інформаційної системи підтримки прийняття рішень абітурієнтами.

. Ключові слова: вступ, абітурієнт, заклад вищої освіти (ЗВО), прогнозування, машинне навчання, бустинг, інформаційна система.

### Introduction

The rapid development of information technologies, as well as means of artificial intelligence, contributes to their wide application, in particular in the field of education. In recent years ML has found larger and broader applications in HEI and is showing an increasing trend in scientific research that considers the increasing effectiveness of entrants' admission.

As the admission results directly affect young workers' professional trajectories, it is important to provide appropriate support for entrants at this stage. The world markets are developing rapidly and continuously looking for the best knowledge and experience among people. HEI rating, a wide range of educational programs, a constantly changing labor market, admission rules, incompetent recommendations, or career guidance activities could make this decision complicated. As a result, they could make unwise choices. To increase the entrant's chances of making an effective decision, this research aims to investigate the possibility of using ML techniques to predict their chances of admission.

HEI admission process could differ in all countries and requires investigation in each individual case. This research considers the Ukrainian HEI entrant's decision. Today, Ukrainian HEI provides the target audience with up-to-date and reliable information on official websites in a large volume and an accessible form. There are also open online services that provide additional information about rating lists and the education programs of admission. But still, there is a need to provide any free and open web resource that could support entrant's admission decisions.

According to the admission rules, everyone who intends to admit the HEI must register and pass an external independent assessment test aimed at determining the level of educational achievements of graduates of secondary educational institutions upon their admission to HEI in Ukraine. It is worth noting that the institution sets a minimum passing score for specific education programs. There is also a limit on the number of submitted applications. In every application submitted for the budget form of study, it is necessary to set its priority - from one to five, where one is the highest priority and five is the lowest. After submitting the first application, it will not be possible to change the priority. In accordance with the recommendations of the HEI, which has to be conducted based on the competition, rating lists are published, after which the entrant makes the final choice.

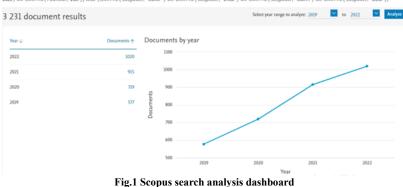
Such a complex admission system is aimed at raising the level of education of the population of Ukraine and ensuring the implementation of the constitutional rights of citizens to equal access to higher education, monitoring compliance with education standards, analyzing the state of the education system, and forecasting its development. However, at the same time, it is difficult for school graduates. There is no clear predictive vision of the most appropriate specialization for the student.

The wrong decision causes the entrant not to get state financing, enter to undesirable program, and could become a not success d student. Given the difference in the admission procedure of foreign higher education institutions, Ukraine cannot adopt the experience of supporting entrants. However, taking into account modern trends and existing solutions, in this study we will analyze the relevance of using ML methods on the example of an admission campaign of Lviv Polytechnic National University admission data.

The aim of the work is to apply and experimentally evaluate the accuracy of ML algorithms in solving the task of predicting the chances of admission to the HEI. The performance indicators will be consider as an indicator of model effectiveness. The most effective model could be used as a base intellectual component for an information system aimed to support Ukrainian entrants.

#### **Related works**

A huge amount of students and entrants data is available in many HEI, which may be utilized to make future decisions and improve future outcomes. As a result, there is a growing number of scientific research using ML methods in the education area tasks [1]. Fig. 1 shows the growing number of researches in the Scopus database related to the use of ML in HEI. This section summarizes the main directions and obtained results of existing latest studies aimed to predict HEI admission and describes some of them.



TLE-A85-KEY (('university' OR "HEI') AND "machine learning") AND (LIMIT-TO (PUBYEAR, 2023) OR LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR) OR LIMIT-

A common practice is to compare ML methods for solving classification problems. Applying such methods as Logistic Regression, KNN Classification, SVM, Naive Bayes Classification, Decision Tree Classification, and Random Forest authors aimed to predict the admission outcome of candidates with a set of known parameters. Comparing the performance metrics of these methods allowed to highlight the most effective solution for each data set [2-6].

As one of effective techniques authors used stacking to predict admission to a bachelor's program. It performs better while compared to other regression algorithms such us, Linear Regression, SVM, Decision Tree, Random Forest, and Stacking Regressor. To analyze all the models they presented evaluation metrics such as R2 Score, Mean Absolute Error, Mean Square Error, and Root Mean Square Error for each case [7].

In other paper, authors use ML methods to resolve classification task to identify the possibility of enrollment for a pool of applicants. There are two approaches presented in the study. The first one, based on SVM and LR models, uses a given set of features and defines the total number of enrollments. The second approach directly implements a semi-supervised probability model and a time series model and determines the number of applicants without identifying them individually. The results demonstrate that the presented models can predict enrollment with reliable accuracy using only a small set of features related to student and college features [8].

With the aim to rid of the accuracy limitation produced by existing web applications or consultancy services, this study uses a deep neural network (DNN) to predict the chance of getting admitted to a university according to the student's portfolio. The DNN model outperformed the results in comparing with existing methods in terms of different performance metrics in each benchmark [9].

Also, in one of the recent researches, the authors emphasize that there is a high level of unemployment and the absence of enough internship or full-time job opportunities for college graduates caused of the coronavirus. At the same time, the number of post-graduate applicants is growing. This makes the need to help applicants to understand their overall admission chances. They use the feasible ordinary least squared multiple linear regression models to analyze and predict the post-graduate school admission chance [10].

Current studies present experiment result based on datasets of various HEI. The problem-solving of entrants' admission chances evaluation is critical in each country. There are studies aimed to support Bangladeshi students intended to pursue higher studies abroad after completing their undergraduate degrees [11], to make recommendations for Indian students who apply for the admissions of overseas universities to study abroad [12], for foreign students came to USA [13].

They also analyze a student's academic achievements historical data to predict graduate program admission [14, 15, 16, 17.] or first-year admission [18, 19].

So today, various types of research consider the effectiveness of the application of ML methods in HEI admission procedures, particularly an admission prediction task The analyzed studies confirm the practicality and value of solving the problem of predicting admission. All the studies summarize that the presence of a decision support system will positively affect both the life trajectory of the future student and the activities of the HEI. They show high accuracy in terms of every settled task and used dataset. This set differs in individual cases according to the task at hand. But the existing studies concerned different HEI, education programs, and even countries, different levels - the first year, re-enrollment, master's, or Ph.D. It is obvious that there is a need to collect and use specific data in every separate case.

Since the research concerns the entrance of other countries and different admission rules, the task of researching the application of ML models for predicting Ukrainian applicants' admission remains open.

#### **Dataset description**

The selection of data was limited by the following characteristics: entry year - 2021, qualification level - bachelor's (entry to the first year), form of study - full-time, faculties - humanities and social sciences, computer sciences and technologies. In connection with the beginning of the Russian-Ukrainian war, the rules of reception have partially changed. As a return to typical procedures is envisaged, based on the external examination, the analysis was carried out on the basis of data from the 2021 entry campaign. Data source: vstup2021.lpnu.ua [20]. Total row number is 2057 for the first case and 5600 for the second. Data preprocessing is described below.

Exploratory data analysis shows some outliers. There are many different methods to deal with outliers - leave as is, drop them with dropna, fill missing values with test statistics like mean. In our case, there are few outliers. Therefore, the best option will be removing the rows that contain outliers - applicants who entered with low scores based on the privileges granted to them due to special cases of the admission rules. In order to removing redundancy we deleted row id because it does not correlate to the dependent variable; hence, it was removed from the dataset.

Instead of introducing separate properties that correspond to the results of the external examination for each subject, we used the Competitive score - this is a comprehensive assessment of the entrant's achievements, which includes the results of entrance tests and other indicators calculated in accordance with the rules of admission to HEI. In addition to the results of external examinations, the formula also provides for integral weighting factors determined for each specialty and additional points for successfully completing the preparatory courses of a HEI in the year of admission. In addition, corrective coefficients are also taken into account - rural (can be calculated for entrants registered in the village and who graduated from a rural school in the year of admission), regional (for those entering regional universities) and branch (for specialties that receive special support). These data are not listed separately in the data source, but they are taken into account in the competitive score.

#### **Proposed technique**

To obtain more accurate results when applying ML methods, the algorithm of their ensembles is used, which consists of the simultaneous application of several basic algorithms. When using them, the algorithms learn simultaneously and can correct each other's errors. The ensemble itself is a supervised learning algorithm because it can be trained and then used to make predictions, so the trained ensemble represents a single hypothesis. This hypothesis, however, does not necessarily lie in the hypothesis space of the models from which it is constructed. Thus, ensembles can have more flexibility in the functions they represent.

There are several approaches to building ensemble algorithms, including stacking, boosting, and bagging. Boosting consists of the gradual application of each model while each subsequent one corrects the errors of the

previous one. During bagging, the basic algorithms are trained in parallel, and the final results are aggregated. In stacking, several different algorithms are trained, and at their outputs, the metamodel produces the final result.

There are a number of key advantages that the boosting method provides when used for classification or regression problems, such as ease of implementation, bias reduction, and computational efficiency. So, in this work, we investigated the accuracy of solving the binary classification task using a number of existing methods of ML, in particular: AdaBoost, XGBoosting, and Gradient Boosting. We aim to evaluate the accuracy and efficiency of boosting methods and the possibility of their future implementation. For better evaluation, we compare its result with other different ML methods, such as SVM and Logistic Regression, for two separate datasets.

AdaBoost was the first really successful boosting algorithm developed for the purpose of binary classification. AdaBoost is a very popular boosting technique that combines multiple "weak classifiers" into a single "strong classifier". Building on the work of Leo Breiman, Jerome H. Friedman developed gradient boosting, which works by sequentially adding predictors to an ensemble, with each one correcting for the errors of its predecessor. However, instead of changing the weights of data points like AdaBoost, the gradient boosting trains on the residual errors of the previous predictor. The name, gradient boosting, is used since it combines the gradient descent algorithm and boosting method.

XGboost is a decision-tree-based ensemble ML algorithm that uses a gradient boosting framework. It is one of the gradient boosting implementations that is acknowledged as one of the best-performing algorithms used for supervised learning. Its main advantage is high execution speed out of core computation. XGBoost algorithm was developed as a research project at the University of Washington. Since its introduction, this algorithm differentiates itself in a wide range of applications - can be used to solve regression, classification, ranking, and user-defined prediction problems. XGboost preferred by data scientists because its high execution speed out of core computation. More detailed description can be found in [21].

The simulations were performed using "Orange" software [22], an open-source online data visualization, ML, and data analysis tool. It is equipped with a visual programming interface for fast qualitative analysis and interactive visualization of data. The block diagram of this process is presented in Fig. 2.

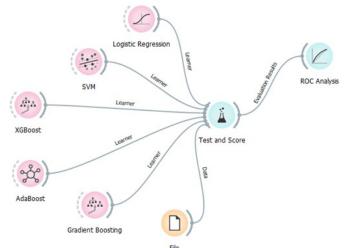


Fig. 2 Research simulation in the "Orange" software

Due to the sufficient size of the data sample, it was divided into training and test samples in the ratio of 80% to 20%. Classification accuracy was assessed using 10-fold cross-validation. The essence of such a check is to compare the results of the classification accuracy of the test and training sets. It is considered that the studied methods has passed the test, provided that the classification of the test set gives approximately the same results in terms of accuracy as the classification of the training set.

To evaluate the performance of the ML method, we use following performance indicators:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} , \tag{1}$$

$$Precision = \frac{TP}{TP + FP},$$
(2)

$$Recall = \frac{TP}{TP + FN},$$
(3)

$$F1\_score = 2 \times \frac{(Recall \times Precision)}{(Recall + Precision)},$$
(4)

where: TP are true positive observations, TN are true negatives observations, FP are false positive observations, FN are false negatives observations.

#### **Result and discussion**

This section presents the results of work and a comparison of such ML methods as Gradient Boosting, AdaBoost, and XGBoost. Table 1 and Table 2 describe the results obtained for each data set according to the accuracy evaluation measures described in the previous section.

Table 1

| <b>Results of boosting ML methods, dataset 1</b> |       |       |           |        |  |
|--|-------|-------|-----------|--------|--|
| Model  | AUC   | F1    | Precision | Recall |  |
| XGBoost  | 0.890 | 0.919 | 0.919     | 0.924  |  |
| Gradient Boosting                                | 0.888 | 0.915 | 0.917     | 0.921  |  |
| AdaBoost   | 0.804 | 0.890 | 0.889     | 0.891  |  |

Table 2

| Results of boosting ML methods, dataset 2 |       |       |           |        |  |
|---|-------|-------|-----------|--------|--|
| Model                                     | AUC   | F1    | Precision | Recall |  |
| XGBoost                                   | 0.972 | 0.962 | 0.962     | 0.963  |  |
| Gradient Boosting                         | 0.967 | 0.962 | 0.961     | 0.962  |  |
| AdaBoost                                  | 0.891 | 0.930 | 0.930     | 0.931  |  |

Taking into account the fact that the competitive scores of entrants of humanities majors can be higher on average than those of mathematics and natural sciences, therefore, in this study, two separate cases of different faculties are cursed and evaluated. This will allow for a more detailed evaluation of the effectiveness of the selected methods and to avoid discrimination of technical specialties of higher education institutions.

The findings of this study clearly show that the XGboost method gives the most accurate results. The obtained numerical values of the metrics were also confirmed by visual analysis. An error curve, the so-called ROC curve, was used to visualize the research results (Fig. 3, Fig. 4). This is one of the most commonly used methods of demonstrating binary classification results. The ROC curve shows the dependence of the number of true positive values on the number of false positive values for each individual class. Accordingly, the studied classifier, whose ROC curve is located above and to the left of the graph, demonstrates greater accuracy.

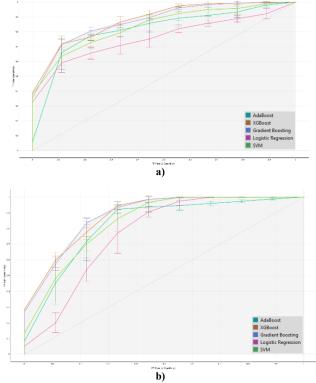


Fig. 3 ROC-curves for the boosting methods, dataset 1: a) class 1 (admitted); b) class 2 (not admitted)

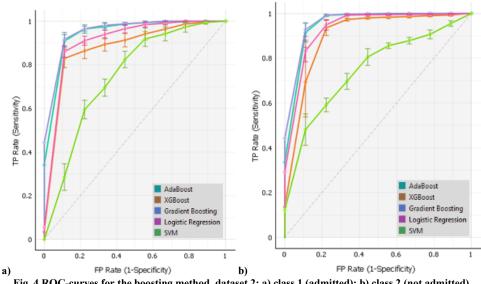


Fig. 4 ROC-curves for the boosting method, dataset 2: a) class 1 (admitted); b) class 2 (not admitted)

For a better assessment of the researched methods, an analysis of the work of such methods as SVM and Logistic Regression was also carried out. The results are presented in Figure 5.

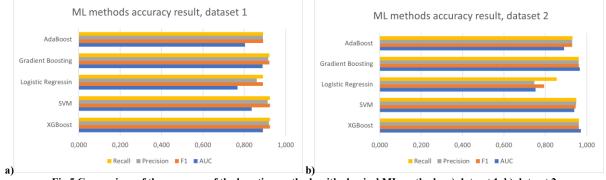


Fig.5 Comparison of the accuracy of the boosting methods with classical ML methods: a) dataset 1, b) dataset 2

As can be seen from both graphs of Figs. 3 and Fig. 4, ROC-curves several algorithms almost overlap. This indicates that they are approximately equally effective. This is confirmed by the results presented in Table 1, and Table 2. However, XGBoost shows slightly better results. This is also confirmed by numerical estimates of comparison with other methods presented at Fig. 5. This provides the possibility of applying this method when building a real system for predicting the success of the entrant's entry to HEI.

#### Conclusion

ML algorithms are widely used in many fields of scientific and practical activity, including education. Ensemble learning has been commonly used in machine learning on various classification and regression tasks to improve performance by grouping individual algorithms. Boosting is one of the most popular ensemble learning techniques, where a set of so-called weak learners, i.e., models whose performance is slightly better than random guessing, is built.

The current development of decision-making support systems for applicants requires the search and application of the most effective and accurate methods of predicting admission success. The need for the entrants to decide on the choice of university and education program for admission arises every year during the admissions campaign. So, supporting applicants remains relevant for all educational institutions. From the research, we were able to verify the effectiveness of the application of ML methods and techniques to solve such a task.

This paper examines the task of supporting applicants' decision-making to HEI choosing an education program for admission using ML boosting methods. It has been experimentally established that the highest accuracy is achieved using the XGBoost method. The obtained results make it possible to consider the boosting method as the basic algorithm of the recommender system, i.e., the components and decision support of applicants to the University of Ukraine. Although the XGBoost method showed the highest accuracy, further research will be conducted in the more detailed investigation of boosting methods and a more full dataset.

#### References

Scopus - Elsevier's abstract and citation database. URL: <u>https://www.scopus.com</u> (revised on 10.01.2023).

2. Iman A., Tian X. A Comparison of Classification Models in Predicting Graduate Admission Decision. *Journal of Higher Education Theory and Practice*. 2021. Vol. 21, No. 7. P. 219-230.

3. Jeganathan S., Parthasarathy S., Lakshminarayanan A. R., Ashok Kumar P. M. and Khan M. K. A. Predicting the Post Graduate Admissions using Classification Techniques. 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2021. P. 346-350.

4. Kalakova A., Amanbek Y., Kairgeldin R. and Kalakova G. Classification and prediction of Student's Enrollment to Kazakhstanis Universities Using Characteristics of Applicant and Testing Results. *IEEE International Conference on Smart Information Systems and Technologies (SIST)*. Nur-Sultan, Kazakhstan, 2021. P. 1-7.

5. Kruthika CS., Apeksha B., Chinmaya GR., Madhumathi JB., Veena MR. University Admission Prediction using Machine Learning. *Global Journal of Research and Review*. 2021. Vol. 8, No. 7.

6. Zub K., Zhezhnych P. Performance evaluation of ML-based classifiers for HEI Graduate Entrants. International Workshop of IT-professionals on Artificial Intelligence (ProfIT AI 2021), Kharkiv, Ukraine, September 20-21, 2021. P. 92-97.

7. Ace C. Lagman, Lourwel P. Alfonso, Marie Luvett I. Goh, Jay-ar P. Lalata, Juan Paulo H. Magcuyao, and Heintjie N. Vicente. Classification Algorithm Accuracy Improvement for Student Graduation Prediction Using Ensemble Model. *International Journal of Information and Education Technology*. 2020. Vol. 10, No. 10. P/ 723-727.

8. Slim A., Hush D., Ojah T., Babbitt T. Predicting Student Enrollment Based on Student and College Characteristics. *Proceedings of the 11th International Conference on Educational Data Mining*, Raleigh, NC, Jul 16-20, 2018. P. 383-389.

9. Omaer Faruq Goni M., Matin A., Hasan T., Abu Ismail Siddique M., Jyoti O. and Sifnatul Hasnain F. M. Graduate Admission Chance Prediction Using Deep Neural Network. *IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering* (WIECON-ECE), Bhubaneswar, India, 2020. P. 259-262.

10. Bo Liu, Research on the Admission of Graduate Students Based on Multiple Regression Model. *Proceedings of the 2022* International Conference on mathematical statistics and economic analysis, December, 2023. P. 681-689.

11. Naimul Islam Suvon, Sadman Chowdhury Siam, Sadman Chowdhury Doctor of Philosophy admission prediction of Bangladeshi students into different classes of universities. *International Journal of Artificial Intelligence*. 2022. Vol. 11, No. 4. P. 1545-1553.

12. B. Wu, Z. Ke, M. Fu and Xia Y. SOUA: Towards Intelligent Recommendation for Applying for Overseas Universities. International Conference on Intelligent Computing, Automation and Systems (ICICAS), Chongqing, China, 2019. P. 124-128.

13. Nandal P. Deep Learning in diverse Computing and Network Applications Student Admission Predictor using Deep Learning. *Proceedings of the International Conference on Innovative Computing & Communications (ICICC)*, March 28, 2020.

14. Navoneel C., Siddhartha C., Srinibas, R. A. Statistical Approach to Graduate Admissions' Chance Prediction. *Part of the Lecture Notes in Networks and Systems book series* (LNNS), Vol. 103. 2020. P. 333-340.

15. Acharya M. S., Armaan A. and Antony A. S. A Comparison of Regression Models for Prediction of Graduate Admissions. International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019. P. 1-5.

16. Guabassi E., Bousalem I., Marah Z., Qazdar, A. A Recommender System for Predicting Students' Admission to a Graduate Program using Machine Learning Algorithms. *International Journal of Online and Biomedical Engineering (iJOE)*. 2021. Vol. 17., No. 2. P. 135– 147.

17. Callistus Obunadik Selection of best model to predict graduate admission into United States of American Universities

using University of California, Los Angeles as a Case Study. APSU Teaching Mathematics Conference, Clarksville, Tennessee, USA, 2022.

18. Protikuzzaman Md., Mrinal Kanti Baowaly, Maloy Kumar Devnath and Bikash Chandra Singh. Predicting Undergraduate Admission: A Case Study in Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Bangladesh. *International Journal of Advanced Computer Science and Applications (IJACSA).* 2020. Vol. 11, No. 12. P. 138-145.

19. Pavitha N., Ingale V., Verma V., Yeole A., Zawar S., Jamadar Z. Design Engineering Comparative analysis of Regression Algorithms for College Prediction. *Design Engineering Journal*. 2021.Vol.9. P. 6631-6643

20. Рейтингові списки вступників. URL: <u>https://vstup2021.lpnu.ua</u> (revised on 02.01.2023).

21. Ahmedbahaaaldin Ibrahem Ahmed Osman, Ali Najah Ahmed, Ming Fai Chow, Yuk Feng Huang, Ahmed El-Shafie, Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia, *Ain Shams Engineering Journal*. 2021. Vol. 12, No. 2. P.1545-1556.

22. Orange Visual Programming. URL: https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/index.html (revised on 02.01.2023).

| Khrystyna Zub<br>Христина Зуб   | Ph.D. student, Department of Social Communication and<br>Information Activities, Lviv Polytechnic National<br>University, Lviv, Ukraine.<br>e-mail: <u>khrystyna.v.zub@lpnu.ua</u><br><u>https://orcid.org/0000-0001-6476-7305</u>     | аспірант кафедри соціальних комунікацій<br>та інформаційної діяльності, Національний<br>університет «Львівська політехніка», Львів,<br>Україна       |  |
|---------------------------------|--|--|--|
| Pavlo Zhezhnych<br>Павло Жежнич | DScTech., Prof., Department of Social Communication<br>and Information Activities, Lviv Polytechnic National<br>University, Lviv, Ukraine.<br>e-mail: <u>pavlo.i.zhezhnych@lpnu.ua</u><br><u>https://orcid.org/0000-0002-2044-5408</u> | д.т.н., проф., кафедри соціальних<br>комунікацій та інформаційної діяльності,<br>Національний університет «Львівська<br>політехніка», Львів, Україна |  |

1.