Position Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 113–120
DOI: 10.15439/2022F271
ISSN 2300-5963 ACSIS, Vol. 31

Generative Adversarial Networks for students' structure prediction. Preliminary research

Agata Kozina, Krzysztof Nowosielski, Zdzisław Kes,
Olena Sidor, Marcin Hernes, Paweł Golec
Wroclaw University of Economics and Business
ul. Komandorska 118/120, 53-345 Wrocław, Poland
Email: {agata.kozina, krzysztof.nowosielski, zdzisław.kes,
olena.sidor, marcin.hernes, pawel.golec}@ue.wroc.pl

Korlan Zhanat
Almaty Management University
Rozybakiev Street 227, Almaty 050060, Kazachstan
Email: k.zhanatovna@gmail.com

***Abstract*— The effectiveness of the university's functioning and its organizational culture can be improved thanks to the use of machine learning. At Universities, the context of student anticipation is very important from the point of view of the fundamental planning and control functions associated with this specific form of management. The purpose of this study is to present the results of an experiment involving the prediction of student structure (attributes of students and their activities) based on the use of a machine learning solution and comparing them against real data obtained from a registry system of a European public institution of higher education in economic sciences. At universities, there is a clear need to support various components of system management. The experiments revealed that - for 11 out of the 48 examined datasets - the Percentage Similarity Index was in excess of 75% but was decidedly lower for the remaining sets (with 18 sets assessed below the margin of 50%).**

I. Introduction

A identifies a set of properties deemed important in the context of enrolment management are consisted from [1], [2]:

- factors that induce and incite university enrolment,
- proper understanding of reasons behind dropouts as well as incentives for persistence in students,
- forms of financing employed by students to cover the cost of their education,
- strategic planning of tasks related to the university's present and future financing needs,
- integration between enrolment management and retention management tasks.

In general, the principal function of enrolment management is to provide effective control over student characteristics and student population size.

As observed by Dixon (1995), enrolment management may be designed in pursuit of the following four objectives: (1) clear definition and propagation of institutional goals, (2) ensuring stakeholders' full support for marketing plans and activities made in relation to institutional goals, (3) making strategic decisions on the role and volume of financial aid required to reach and retain the desired size of the student population, and (4) making significant commitments for the realisation of the above. Enrolment management exerts a significant impact upon the structure of the student population and, consequently, the structure of university revenues, as forms of service and curricula are directly manifested in tuition costs, and thus determine the financial standing of the institution [3]. According to a European University Association report, several distinct trends can be observed in the development of public financing and student enrolment in the years 2008-2016 in Europe [4], [5]. In Poland, intensive efforts are underway at present to increase public support for higher learning to negate the effects of brain drain and the gradual decrease of the student population. The above aspects clearly emphasise the need for a more proactive design of the institutional enrolment policy as an essential determinant of future tuition revenues, resource allocation for subsequent academic years, and the creation of marketing plans, especially ensuring their adjustment to specific segments of the university's offer.

In the context of increased competition among the national universities, the strengthening of the influence of global educational processes on the domestic higher education, the need to change the management component of the system becomes obvious. In their development, universities face a large number of challenges, such as: the development of technologies, the commercialization of activities, the increase in the amount of information, the changing requirements of employers for graduates as potential employees. One of the research problems of higher education management is the prediction of students structure. Many higher-education institutions are now using data and analytics as an integral part of their processes. Whether the goal is to identify and better support pain points in the student journey, more efficiently allocate resources, or improve student and faculty experience, institutions are seeing the benefits of data-backed solutions.

The aim of the paper is to develop the method for students' structure prediction using machine learning.

II. BACKGROUND

Corporate culture in the organization arises regardless of whether it is planned from above or not. It exists as a given in any organization, even in a newly created company, it is created by the employees themselves. At the same time, it can either help in achieving the goals of the company, or slow down this process. Corporate culture determines how employees approach problem solving, interact with each other, behave in conflict situations, serve customers, deal with suppliers, and how they generally carry out their activities [6].

To organize effective work, it is necessary to use all available management methods. These methods, according to the authors of the book "Methods of personnel management" are divided into economic, administrative-legal and socio-psychological [7].

Corporate governance always relies on both formal and informal structures. The formal structure is based on the norms that are mandatory for the organization's personnel: hierarchy of subordination, unity of command, sanctions, coercion. The informal structure is based on norms associated with values: sympathy, authority, collegiality, initiative. It is obvious that the use of one formal management leads to rigidity, lack of flexibility in the organization, hinders the development of initiative, which hinders the further development of the organization and leads to loss in competition, while the predominance of the informal structure will lead to chaos, loss of control over the entire hierarchical structure. chain. Therefore, a necessary condition for successful management is the fulfillment of two requirements:

- decentralization of powers and responsibilities within the company to certain limits;
- formation of a single team of employees of the organization.

Currently, universities are mainly characterized by "Club culture", which allows them to work efficiently and smoothly [1].

There are a number of phases of employee interaction with the corporate culture of the university [8]:

1. At the first stage (orientation phase), the employee gets acquainted with the mission, values, symbols of the university, using Internet sites and other information materials;
2. At the second stage (adaptation phase), adaptation to the corporate culture of the university takes place;
3. At the stage of interaction, immersion into the value system of the university takes place, a wide range of communicative interactions with various groups is carried out;
4. The phase of integration involves the value unity of the university with the employee as a bearer of corporate culture.

Those at the forefront of this trend are focusing on harnessing analytics to increase program personalization and flexibility, as well as to improve retention by identifying students at risk of dropping out and reaching out proactively with tailored interventions. Indeed, data science and machine learning may unlock significant value for universities by ensuring resources are targeted toward the highest-impact opportunities to improve access for more students, as well as student engagement and satisfaction [10].

Yet higher education is still in the early stages of data capability building. With universities facing many challenges (such as financial pressures, the demographic cliff, and an uptick in student mental-health issues) and a variety of opportunities (including reaching adult learners and scaling online learning), expanding use of advanced analytics and machine learning may prove beneficial.

Below, we share some of the most promising use cases for advanced analytics in higher education to show how universities are capitalizing on those opportunities to overcome current challenges, both enabling access for many more students and improving the student experience[11].

Data science and machine learning may unlock significant value for universities by ensuring resources are targeted toward the highest-impact opportunities to improve access for more students, as well as student engagement and satisfaction.

Advanced analytics—which uses the power of algorithms such as gradient boosting and random forest—may also help institutions address inadvertent biases in their existing methods of identifying at-risk students and proactively design tailored interventions to mitigate the majority of identified risks. For instance, institutions using linear, rule-based approaches look at indicators such as low grades and poor attendance to identify students at risk of dropping out; institutions then reach out to these students and launch initiatives to better support them. While such initiatives may be of use, they often are implemented too late and only target a subset of the at-risk population [4]. This approach could be a good makeshift solution for two problems facing student success leaders at universities. First, there are too many variables that could be analyzed to indicate risk of attrition (such as academic, financial, and mental health factors, and sense of belonging on campus). Second, while it's easy to identify notable variance on any one or two variables, it is challenging to identify nominal variance on multiple variables.

III. MATERIALS AND METHOD

A. Input data characteristic

As already established, input data includes two groups of variables:

- dependent variables (set of 9883 records, with each record described by 15 attributes) obtained from the registry system of the examined university for the years 2016-2020;
- independent variables (3 attributes) obtained from the national statistical records published by the Central Statistical Office

The variables were coded as follows:

- X1 - work_name – entities employing the candidates were divided according to the type of business;
- X2 - code – fields of studies were grouped by subject;
- X3 - work_city – places of student residence were coded by their physical distance from the university (in km);
- X4 - nationality – the nationality of students;
- X5 - gender – gender of students was coded as follows: 0 for males, 1 for females;
- X6 - status – codes of student status;
- X7 - finished_university – the enrolment data provides details of each candidate's previous education. The recorded institutions of higher learning were assigned codes from 0 to 364.
- X8 - work_years – work experience of candidates registered in the database (in years of service).

B. The model's learning method

GANs are a relatively new method in the field of machine learning. These networks, which were introduced in 2014 by Ian Goodfellow and his collaborators, are designed to create new data that in some form mimics the statistical properties of a given set of training data. Given a target dataset, such as celebrity faces or categories from the ImageNet dataset, a GAN can be trained that generates new, unseen data that (ideally) fit comfortably and indistinguishably in the dataset. Since the introduction of GANs, several variations of the architecture and many theories to help train these inherently unstable networks have been developed[11].

In general, GANs are composed of generator and discriminator neural networks (Figure 1), which, for image data, are typically convolutional networks.

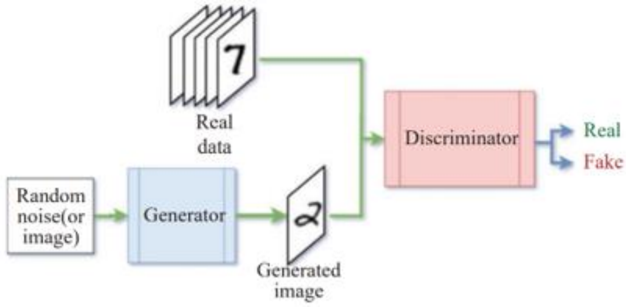

Fig. 1 A diagram of a generic generative adversarial network. The network shown here is designed to produce new images of handwritten MNIST digits. The generator converts random noise into images that attempt to match the data from the target dataset. The discriminator distinguisges between real and generated data

Training is accomplished by repeatedly presenting the networks with data from a target dataset. The generator is tasked with learning to convert random n-dimensional vectors to data matching the dataset. The discriminator, in turn, is tasked with distinguishing between data from the dataset and the generator's output. In a descriptive analogy offered by Goodfellow et al., the generator can be likened to an art forger, the goal of which is to create undetectable forgeries of the world's great artists. The discriminator plays the role of a detective, trying to discover which pieces are real and which are fakes. The loss function for a GAN is given by

$$\begin{aligned}\min_G \max_D L(D,G) &= \mathbb{E}_{x\sim p_r(x)}[\log(D(x))] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))] \\ &= \mathbb{E}_{x\sim p_r(x)}[\log(D(x))] + \mathbb{E}_{x\sim p_g(x)}[\log(1-D(x))]\end{aligned} \quad (1)$$

where $G(z)$is the output of generator network, $D(x)$ is the output the discriminator network, z is a multidimensional random input to the generator, p_z is the distribution of z (usually uniform), and $p_g(x)$ and $p_r(x)$ are the probability manifold distributions for the generated data and the target dataset, respectively. Via backpropagation, these objectives direct the generator to create data that fits well with the dataset, while simultaneously increasing the distinguishing power of the discriminator. It can be shown[11] that, for a GAN with sufficient capacity, this training objective minimizes the Kullback-Leibler divergence between $p_g(x)$ and $p_r(x)$. This divergence metric describes how similar two probability distributions are, with low values denoting greater similarity. In other words, training a GAN creates a generator that is able to mimic the distribution of data in the given dataset at some level.

Models of the generator and the discriminator are presented below. The Sequential model utilises the following layers: Dense, LeakyReLU oraz BatchNormalization. All the layers and the model itself are derived from the Keras library.

Tables VI and VII present the structure of both models. The generator model includes five layers of 'Dense', two layers of batch normalisation, and two functions Leaky ReLU, serving as activation functions. The discriminator model employs seven layers of 'Dense', four functions Leaky ReLU as activation functions for the neurons defined above. The 'Output Shape' column reports a number of nodes for each layer. The loss function used in the discriminator model was developed on the basis of the Binary Cross Entropy function defined as follows (Eq. 1):

$$Loss(y_i, z_i) = \begin{cases} z_i - z_i y_i + \log(1 + e^{-z_i}) & if\ z_i \geq 0 \\ -z_i y_i + \log(1 + e^{z_i}) & if\ z_i < 0 \end{cases} \quad (2)$$

Both models (the generator and the discriminator) are fed with discriminator responses (argument z_i):

- The generator's loss function is the Binary Cross Entropy function with values
$$y_i = 1, for\ \forall_{i=0}^{N} z_i$$
- Discriminator: sum of Binary Cross Entropy functions with values
$$y_i = \begin{cases} 1, & for\ real\ z \\ 0, & for\ artificial\ z \end{cases}$$

Both models utilise the 'Adam' algorithm with a learning step: 0.0001. This step value has already been employed in GANs procedures for TensorFlow. At the same time, as evidenced by research presented in [12], the value yields much better results compared to other algorithms, offering the added benefit of facile and simple implementation in Keras.

For the entire duration of the learning process, examples were fed randomly. The network gained knowledge of the student patterns based on the entire set of input data. The training procedure was set at 55 000 iterations. One epoch was represented by one packet of data holding information on 16 students. Figure 2 provides a plot of the training history.

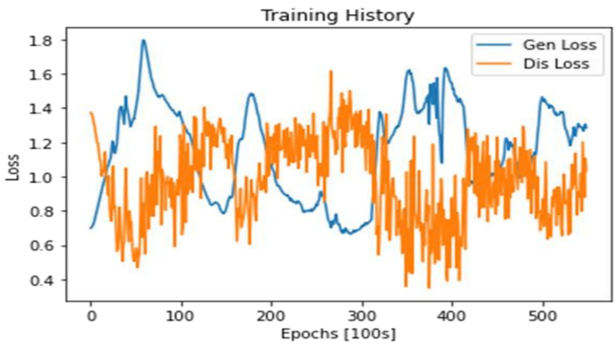

Fig. 2 Training history

As evidenced by the above, the discriminator was able to recognise between fake and real data at a relatively early stage of the training process. This was accompanied by deterioration in the quality of the generator's output over time. This phenomenon can be explained by differences in the number of layers. As the generator utilised fewer layers than the discriminator, its training processes were more immediate. However, the discriminator was, at the same time, more effective in its long-term predictions, owing to the benefit of more layers. The continual learning phenomenon was established.

TABLE VI.
GENERATOR. MODEL: "SEQUENTIAL"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 4)	20
batch_normalization_3	(Batch (None, 4)	16
leaky_re_lu_9 (LeakyReLU)	(None, 4)	0
dense_1 (Dense)	(None, 5)	25
dense_2 (Dense)	(None, 6)	36
dense_3 (Dense)	(None, 7)	49
dense_4 (Dense)	(None, 7)	56
batch_normalization_1	(Batch (None, 7)	28
leaky_re_lu_1 (LeakyReLU)	(None, 7)	0
dense_5 (Dense)	(None, 8)	64
Total params: 294 Trainable params: 272 Non-trainable params: 22		

TABLE VII.
DISCRIMINATOR. MODEL: "SEQUENTIAL_1"

Layer (type)	Output Shape	Param #
dense_6 (Dense)	(None, 7)	63
dense_7 (Dense)	(None, 6)	48
leaky_re_lu_2 (LeakyReLU)	(None, 6)	0
dense_8 (Dense)	(None, 5)	35
leaky_re_lu_3 (LeakyReLU)	(None, 5)	0
dense_9 (Dense)	(None, 4)	24
leaky_re_lu_4 (LeakyReLU)	(None, 4)	0
dense_10 (Dense)	(None, 3)	15
leaky_re_lu_5 (LeakyReLU)	(None, 3)	0
dense_11 (Dense)	(None, 2)	8
dense_12 (Dense)	(None, 1)	3
Total params: 196 Trainable params: 196 Non-trainable params: 0		

C. Methods of output data verification

The Percentage Similarity Index (PSI) was adopted to verify the established similarities between structures of individual variables and those generated by GANs. The index was calculated for equinumerous sets of structural indices based on formula (Eg. 2).

$$PSI = \sum_{k=1}^{n} min(I_{1k}; I_{2k}) \qquad (2)$$

where:

- PSI - percentage similarity index,
- I_{1k} - percentage share of k-th component in the structure of set 1,
- I_{2k} - percentage share of k-th component of the structure of set 2,
- n - number of elements in set 1 (both sets need to be equinumerous).

The following similarity ranges were defined for the evaluation of the sets:

- 100% - 90% - sets are similar,
- 90% - 75% - sets are moderately similar,
- 75% - 50% - similarity between sets is marginal,
- 50% - 0% - sets are not similar

IV. Research results and discussion

A. Output data

The results obtained from the trained generator in the course of the experiment, complete with student characteristics, statistical properties and examples derived from the output data set, are presented below. Table VIII presents a segment of output data generated by GANs for the year 2021.

Each column of the generated output corresponds to specific information items stored in the university database of student records. Parts of the output data were rounded off, as dictated by the specificity of information stored therein. An example of such procedure is the 'Status' column, with domain defined by $x \in \langle 0;4 \rangle \wedge x \in Z$.

B. Output data verification

As suggested by the statistical properties of data, the generated output records are well contained in the brackets defined by the real data. It was assumed that the structure of output data generated by the GEN network for the years 2016-2020 should take up values similar to those of the real records stored for the period. The PSI was used to verify the similarity between the structures of individual variables and the output data generated by the GEN network. Results of the output data verification procedure are provided in Table IX. The PSI exceeded 75% for eleven cases among all tested variables, which suggests their similar or moderately similar character. Thus, it may be concluded that in those cases, the training turned out to be consistent with these segments of data. The best results were obtained for variable X4, nationality – PSI for all tested years was over 90%, and the overall value was 93.5%. Quite satisfactory levels of PSI, between 60 and 90%, were received for X3, city, X5, gender, and X7, work experience, overall PSI reached respectively 73.5%, 70.4%, and 67.2%. The worst fit, overall PSI 30.9%, was found for X8, university. For each variable, histograms were produced to observe the similarities between the representations of both datasets. Figures below present percentage shares in the categories represented in the real dataset and the set generated by the GEN network. Fig. 3 illustrates the structure of the 'employer' variable (X1) for real and generated data for the tested years. The overall PSI index calculated for variable X1 was at 44.5%, suggesting a dissimilarity between the structures generated by GENs and those of the real data. The structure of the 'study field' variable (X2) for real and generated data for 2016-2020 is shown in Fig. 4. The PSI index calculated for variable X2 was 52.3%, suggesting a

TABLE VIII.
CHARACTERISTICS OF REAL DATA

Parameter	Employer	Study field	City	Nationality	Gender	Status	Work experience	Finished University
mean	2.456	4.973	53.954	0.020	0.674	1.051	3.386	44.747
std	2.261	2.857	123.136	0.331	0.469	0.843	6.215	71.159
min	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
25%	0.000	3.000	0.000	0.000	0.000	0.000	0.000	2.000
50%	3.000	4.000	0.000	0.000	1.000	1.000	0.000	3.000
75%	4.000	6.000	70.200	0.000	1.000	1.000	4.000	78.000
max	6.000	12.000	3 349.000	12.000	1.000	4.000	42.000	364.000

TABLE IX.
PERCENTAGE SIMILARITY INDEX (PSI) FOR TESTED VARIABLES

Variable		Years					All Years
		2016	2017	2018	2019	2020	
Employer	X1	46.2%	47.5%	41.3%	44.7%	37.9%	44.5%
Study field	X2	49.7%	49.3%	51.4%	52.4%	5.8%	52.3%
City	X3	68.9%	74.0%	72.8%	78.5%	71.9%	73.5%
Nationality	X4	90.1%	90.9%	95.9%	95.3%	99.5%	93.5%
Gender	X5	68.7%	66.9%	71.2%	71.3%	87.7%	70.4%
Status	X6	43.6%	90.3%	71.0%	15.6%	0.0%	54.0%
Work exp.	X7	77.1%	82.1%	59.4%	53.8%	50.7%	67.2%
University	X8	25.6%	33.3%	30.6%	31.8%	6.7%	30.9%

medium similarity between the structures generated by GENs and those of the real data. Data stored in the real records of the university identify management as the most attractive field of study. In contrast, output data generated by the network reported Audit and financial control as the dominant area.

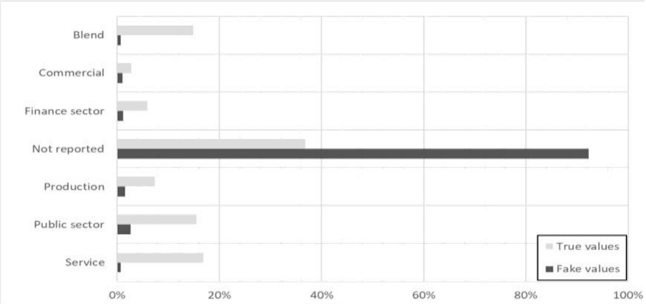

Fig. 3 Histogram of the 'employer' variable (X1)

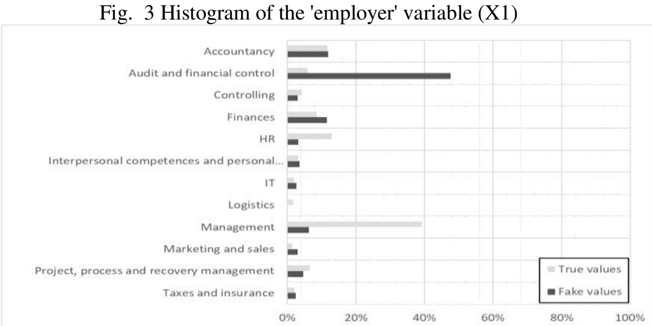

Fig. 4 Histogram of the 'study field' variable (X2)

Fig. 5 presents the structure of the 'city' variable (X3) for real and generated data.

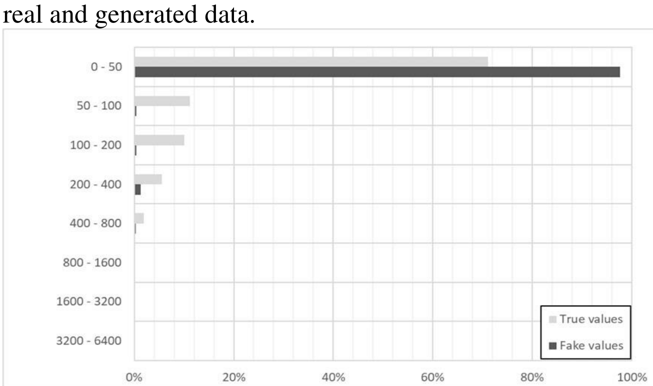

Fig. 5 Histogram of the 'city' variable (X3)

The overall PSI index calculated for variable X3 was 73.5%, suggesting a significant similarity between the structures generated by GENs and those of the real data. Real data shows that Wroclaw (including the city outskirts) is the place of residence for the overwhelming majority of the student population. The structure of GENs output data suggests that ca. 71% of students commute over a distance of 0 to 50 km, which is compatible with real data.

The structure of the 'nationality' variable (X4) for real and generated data is illustrated in Fig. 6. The overall PSI index calculated for variable X4 was 93.5%, suggesting a strong similarity between the structures generated by GENs and those of the real data. The real data shows that more than 99% of postgraduate students come from Poland. The structure of GENs output data suggests that ca. 7% represent a foreign nationality.

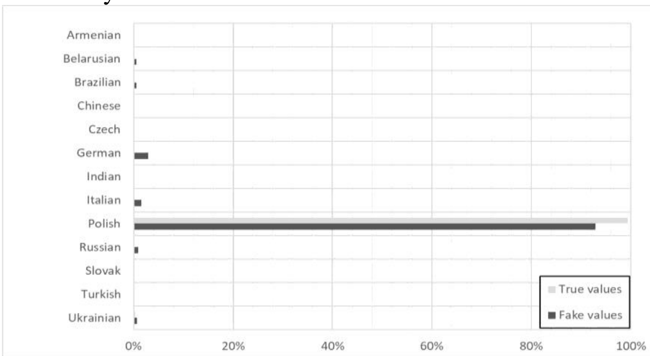

Fig. 6 Histogram of the 'nationality' variable (X4)

Fig. 7 presents the structure of the 'gender' variable (X5) for real and generated data for the tested years. The overall PSI index calculated for variable X5 was 70.4%, suggesting a significant similarity between the structures generated by GENs and those of the real data. The actual data shows that female students account for two-thirds of the student population, while the data predicted by GENs shows complete female dominance.

The structure of the 'status' variable (X6) for real and generated data is shown in Fig. 8. The PSI index calculated for variable X6 was 54.0%, suggesting a medium similarity between the structures generated by GENs and those of the real data. Generated data indicate that most of the students are promoted, while the real data show that many students are still studying, have been deleted or resigned.

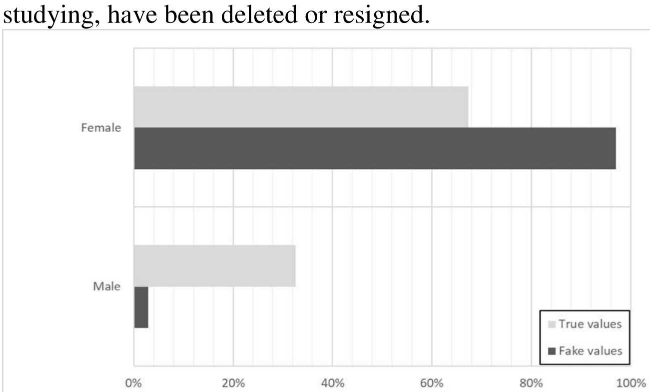

Fig. 7 Histogram of the 'gender' variable (X5)

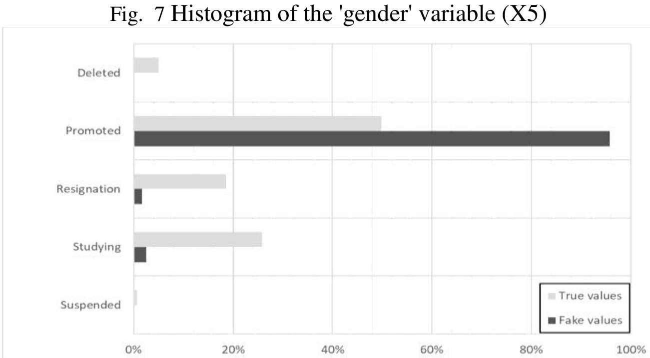

Fig. 8 Histogram of the 'status' variable (X6)

Fig. 9 presents the structure of the 'work experience' variable (X7) for real and generated data for the tested years. The PSI index calculated for variable X8 was at 67.2%, suggesting a moderate similarity between the structures generated by GENs and those of the real data. Generated data show that most of the students have worked for two years or less, while the real data show also that the work experience in many cases is longer.

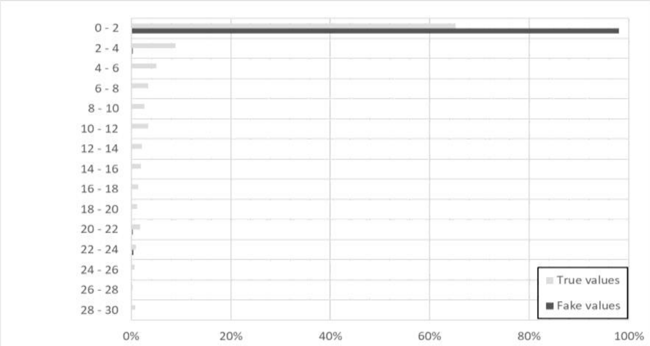

Fig. 9 Histogram of the 'work experience' variable (X7)

The structure of the 'finished university' variable (X8) for real and generated data for the years 2016-2020 is illustrated in Fig. 10. The PSI index calculated for variable X8 was at 30.9%, suggesting a dissimilarity between the structures generated by GENs and those of the real data.

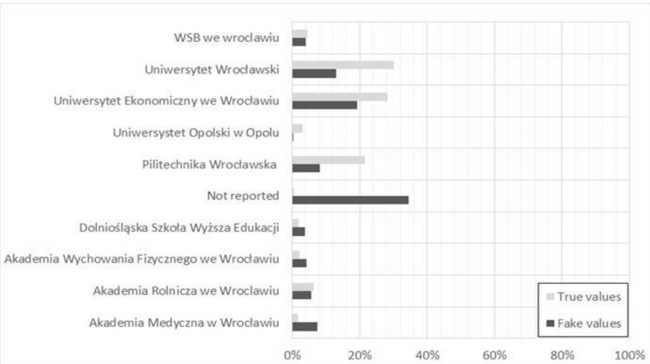

Fig. 10 Histogram of the 'finished university' variable (X8)

Based on analytical evaluations, it may be concluded that the topic of student structure prediction (SSP) is not adequately represented in the professional literature. At the same time, the studied concept serves important practical purposes and presents a major challenge for the managerial cadres of public institutions of higher education. Another important issue of this paper is the use of GANs in predictions. The use of such tools is presented in the literature. However, from the viewpoint of this research and the instrumental utilisation of this technique in SSP, this particular segment of knowledge should not be treated as a point of reference in our discussion. There remains a key area for the purpose of the work, i.e. the use of GANs in the SSP. In this context, scarcity of reference material can be observed, similar to that of the SSP. The available literature is limited to the prediction of the number of students or students' performance. Naturally, these aspects are in close association with SSP, but they are too far detached from the nature of SSP to be of any rational significance in the studied context. Their practical benefits may be important from a broader perspective of using machine learning solutions as a form of management support in university administration. Because of the observed scarcity of reference material related to the studied context, the authors of this study propose to treat the presented research results as a contribution to the discussion on the use of GANs in the SSP.

V. Conclusion and Implications

Working on a copious set of factors of potential impact upon the student structure prediction presented in this paper, the authors examined the perspective of applying 'intelligent solution' methods for the task performed based on Generative Adversarial Networks. The research was conducted on a dataset of records describing the real population of students of postgraduate studies over a period of 5 academic years, between 2016 and 2020. Individual properties and attributes of students were coded. The dataset was supplemented by a number of indices describing the general economic condition of the region proper for the studied university and the timeframe under study. The final design included 12 dependent variables and three independent variables to give a total of 15 variables. The experiment made use of artificial intelligence networks, specifically the GANs networks. The network was presented with the tasks or reproducing the structure of students to produce output adequately comparable with real data recorded for previous years. The Percentage Similarity Index (PSI) was calculated for each variable to illustrate the similarity between their real structure and that produced by GANs. The experiments revealed that – for 11 out of the 48 examined datasets – the PSI index was in excess of 75% but was decidedly lower for the remaining sets (with 18 sets assessed below the margin of 50%). This should be interpreted as evidence that only parts of data generated by GANs sufficiently reflect the real data. Additional tests may be required to provide grounds for more reliable predictions of student structure, including those involving different sets of independent variables. The need for extension of the set of variables is fairly evident. More effort should be placed to verify their information potential and activate a learning mechanism after verifying or exchanging variables. Other methods for selecting variables should also be examined, as the present set was established based on the expert method. Further research directions may involve the development of methods based on other neural network architectures (such as Recurrent Neural Networks, Convolutional Neural Networks) to predict postgraduate students' structure. The method applied by authors may not be an ideal solution to the problem at hand. However, since the attempt proved partially effective, the results are of scientific value and may serve as the basis for further examination of the SSP concept.

Funding

The project is financed by the Ministry of Science and Higher Education in Poland under the programme "Regional Initiative of Excellence" 2019 - 2022 project number 015/RID/2018/19 total funding amount 10 721 040,00 PLN

REFERENCES

[1] D. Hossler and B. Bontrager, *Handbook of strategic enrollment management*. San Francisco, CA: Jossey-Bass, A Wiley Brand, 2014.

[2] M. J. Denniss, ‘Anticipatory Enrollment Management: Another Level of Enrollment Management’, vol. 88, no. 1, pp. 10–16, 2012.

[3] D. Trusheim and C. Rylee, ‘Predictive modeling: linking enrollment and budgeting’, *Planning for Higher Education*, vol. 40, 2011.

[4] D. M. West, ‘Big data for education: Data mining, data analytics, and web dashboards’, *Brookings*, 2012. https://www.brookings.edu/research/big-data-for-education-data-mining-data-analytics-and-web-dashboards/

[5] X. Shacklock, ‘From bricks to clicks: the potential of data and analytics in higher education’, *VOCEDplus*, 2016. http://hdl.voced.edu.au/10707/411226

[6] E. N. Saribekyan, ‘Organizational culture and organizational culture’, *Culture: management, economics, law*, no. 4, p. 37, 2014.

[7] V. N. Fedoseev and S. N. Kapustin, ‘Methods of personnel management’, in *Analysis of the crop industry*, Almaty, 2014.

[8] A. V. Shelyakina, ‘Corporate culture of the organization’, *Young scientist*, no. 14, pp. 206–209, 2018.

[9] M. V. Shumeiko, ‘Typology of corporate culture’, *Cyberleninka*, p. 8.

[10] J.A. Gray and M. DiLoreto. "The effects of student engagement, student satisfaction, and perceived learning in online learning environments." *International Journal of Educational Leadership* Preparation 11.1 (2016): n1.

[11] S. Z. Gurbuz, Ed., ‘Machine learning techniques for SAR data augmentation’, in *Deep Neural Network Design for Radar Applications*, Institution of Engineering and Technology, 2020, pp. 163–206. doi: 10.1049/SBRA529E_ch6.

[12] J. Brownlee, ‘Gentle Introduction to the Adam Optimization Algorithm for Deep Learning’, *Machine Learning Mastery*, 2017. https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/