Decision-making on Human Individual Capital Investment*

Natalia V. Apatova [1 [0000-0003-4066-3821]], Irina N. Ostapenko [1 [0000-0003-1946-7586]], Roman S. Usenko [1 [0000-0003-0095-1183]]

[1]V.I. Vernadsky Crimean Federal University, Simferopol, Russia

i.n.ostapenko@mail.ru

Abstract. This article presents an approach to the feasibility of investing in human capital in the framework of the development of socio-economic systems of intellectual capital. In conditions of partial or complete uncertainty in the presence of incomplete input data, the use of tools based on artificial intelligence methods (neural networks, genetic algorithms) allows us to evaluate a wide range of indicators. The article presents non-linear mathematical models for assessing the feasibility of investing in the intellectual capital of a university, in particular, in its structural component - individual human capital, namely, in the formation of an individual. Models are based on logistic regression and neural networks. The authors proposed several key factors affecting the binary resulting variable "investment effect": the level of knowledge and skills in the specialty; self-education (in terms of regularity of raising the level of information culture); employee age. The developed methodology is the basis of a decision support system that allows for the strategic development of the university based on effective investment in the human capital of its employee.

Keywords: human capital, intellectual capital, investment portfolio, logistic regression, neural networks, evolutionary modeling.

1 Introduction

In the conditions of the digital transformation of all spheres of human life, his creative and intellectual potential acquires a new significance. The future of humanity never then earlier depends on competent balanced decisions, for every person's potential, which is unique, must be more fully disclosed and optimally used. But the reality is that the streaming development of the digital economy is ahead of the willingness of each individual to comply with the newest technologies - it needs permanent growth, continuing education.

Human capital development may be a more important factor in the long-term success of a country (region or any large-scale system) success. Ingenuity, creativity, and an imaginative approach, which has humanity, allow us not only to solve the main problems of our time but to build a future, which is oriented on a human being [1].

According to the report about human capital in 2017, which is developed by the World economic forum, a global index of human capital is estimated for 130 countries by the scale from 0 to 100 by four sub-indexes (potential, development, deployment, and know-how). It aims to provide a holistic assessment of the human capital of the country, both current and expected values. It allows doing an effective comparison according to regions, age groups, the income of people, etc. [2]. Russian Federation is included in the top-20 countries, which have the highest index of human capital, it takes 16 places. Except “know-how”, it has a high rate by three sub-indexes, which is explained by the high levels of primary, secondary, and higher education in all age groups. Nevertheless, in the frames of sub-index “know-now”, it can be said, that the training of personal and its efficiency are insufficiently high. That points to the necessity for additional efforts in the field of labor force development and preparing the country’s population for the fourth industrial revolution, which is conditioned by modern trends in robotics and the introduction of cyber-physical systems.

According to the national educational doctrine in the Russian Federation up to 2025 year, one of the main aims is “lifelong persons steadily education” [3]. And if the free public education (FPE) stages implements as a continuous process at the state level (from preschool FPE, onto basic generally, secondary (complete) education, initial and secondary vocational education, higher education, and postgraduate free education in graduate and doctorate schools, education on a competitive basis), then further stages, that assume self-education, still fully depends on the person himself, his willingness to changes, and from the employer, ho acting as an investor, can carry on this process through motivation his workers to lifelong studying.

2 Main Part

The problems of developing human capital and evaluating the effectiveness of investments in human capital, in particular in education, were studied by such scientists: G.S. Becker [4], J. Mincer [5], J.J. Heckman, L.J. Lochner, P.E. Todd [6], N.G. Mankiw, D. Romer, D.N. Weil [7], R.J. Barro [8], V.M. Porokhnya [9], N.R. Kelchevskaya, E.V. Shirinkina [10] and others.

The intellectual capital (IC) market cannot satisfy the investor according to somewhat algorithm. Except for unique genius, many talent specialists or simply obligatory specialists in a certain area, all other potential employees only with certain probability can impact on the organization profitability, by investing capital in them, or else, by investing in intellectual recourse.

However, the feasibility of investing in individual human capital is relevant.

Even a return from geniuses and talents depends on the vagaries of the "human factor" (for example a person loses his motivation at a certain stage of his intellectual potential growth, or transmission of values exists). In the human capital market is observed a constant rivalry: the income of each investor often depends on how many other employers are willing to invest in a certain moment of an uncertain future in human capital. But no one can control or even predict the behavior of a large number of other investors with a sufficient degree of reliability. However, the employers, who are investing in the human recourse, or generally, in the intellectual capital, can carry on a risk, which they take on themselves.

The research will focus on the evaluation of an investment in teachers of higher education organizations, when educational institutions seek to build capacity for strategic development, given the dynamics changes especially.

During the formation of the "investment portfolio" IC of the organization appears a problem of evaluation of return probability from the investment or the task of evaluating risk. In the frames of this task conclusion, it might use logistic binary regression.

Let F_1, F_2, ..., F_n – be a factor sign, Y – resulting binary sign: $Y \in \{0;1\}$. We set ourselves the task of determining the object of study belonging to one of two classes.

Let give examples of binary regression using:

- issuance of credit – refusal to issue a credit – the solving of a credit scoring task based on the values of the factors F_1, F_2, ..., F_n, specified in the application. The resulting sign value, which takes only two values: $Y=1$, if the client is trustworthy, and $Y=0$ otherwise, is predicting;
- link operators – the solving of problems like payable account: customer payment, wherein $Y=1$, if the customer is not going to change the operator, and $Y=0$ – client transition to contestants;
- equipment replacement task: according to the instrument readings F_1, F_2, ..., F_n, the task of predicting an uninterrupted operation of equipment for a certain time is solving;
- advertising response: wherein $Y=1$, if the customer has bought some product or service, and $Y=0$ – the potential client hasn't bought anything.

In the case of solving an investment in the intellectual capital task, particularly in human capital, it necessary to check concrete factor signs, which characterize an investment object. It must be noted, they are chosen purely for a particular organization, taking into account its strategic objectives and tactical measures. Among that factors maybe sex, intellect level, creativity level, participation in conferences, marital status, number of children, housing conditions, etc.

We can recognize several main factors that influence the parameter being evaluated:

- the level of knowledge and skills in the specialty (F_1);
- self-education (in terms of regularity of information culture level increasing - F_2);
- age of employee (F_3).

There are presently only three factors among the main ones in this research, affecting the impact of return on investment in the staff of the department. There may be much more factors, their quantity is determined by the researcher. The meaning of the selected factor F_1 is shown in Fig. 1.

The meaning of the selected factor F_2 is shown in Fig. 2.

It can be noted, that the evaluation of the level of knowledge and skills in the specialty and self-education can be carried out by an expert method (experts are department staff himself, his colleagues, the head of the department).

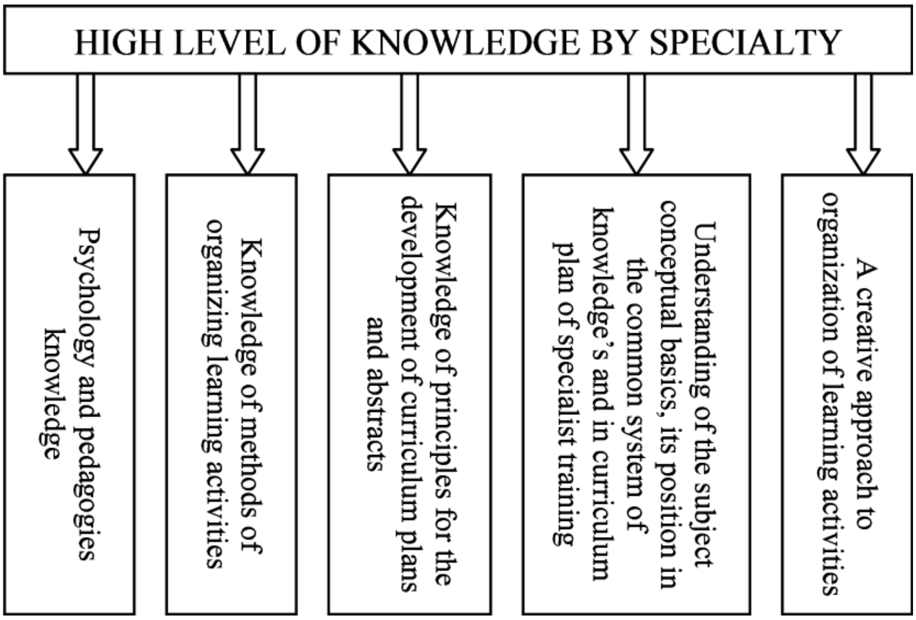

Fig. 1. Main signs of the high level of specialty knowledge.

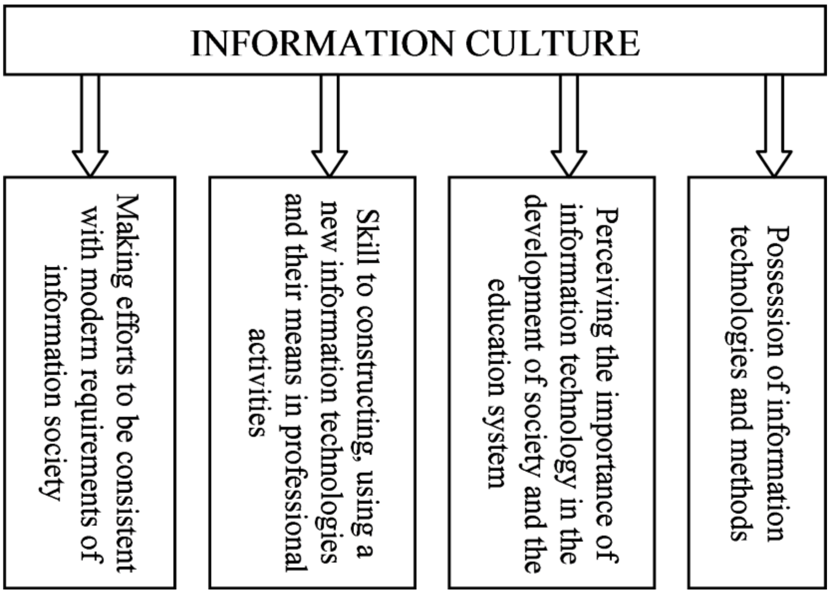

Fig. 2. Main signs of teacher's information culture.

As a binary resulting variable in this task can be chosen an effectiveness of department staff activity (“effect” – “return”). This variable will take the meaning Y=1, in case it appears effect from investments (copyright certificates, patents, publications in Scopus, Web of Science, etc.), and Y=0 if the investment effect is absent.

For evaluation of the impact of mentioned above signs on the resulting binary, the variable can be used the mathematical modeling method, based on the using of the logistic regression equation. The logistic model allows evaluating a probability $P(Y \mid F)$ of dependent variable (Y) binary result independence from the independent variable (*F*) meanings according to the next formula [11, 12].

$$P(Y \mid F) = \frac{1}{1 + e^{-(b_0 + b_1 F)}} \tag{1}$$

In the case of existing of three independent variables it may use the next formula:

$$P(Y \mid F) = \frac{1}{1 + e^{-(b_0 + b_1 F_1 + b_2 F_2 + b_3 F_3)}} \tag{2}$$

The coefficients of the model are calculated by using instruments of the STATISTICA package (the module Nonlinear Estimation can be used for these purposes), presented in Fig. 3.

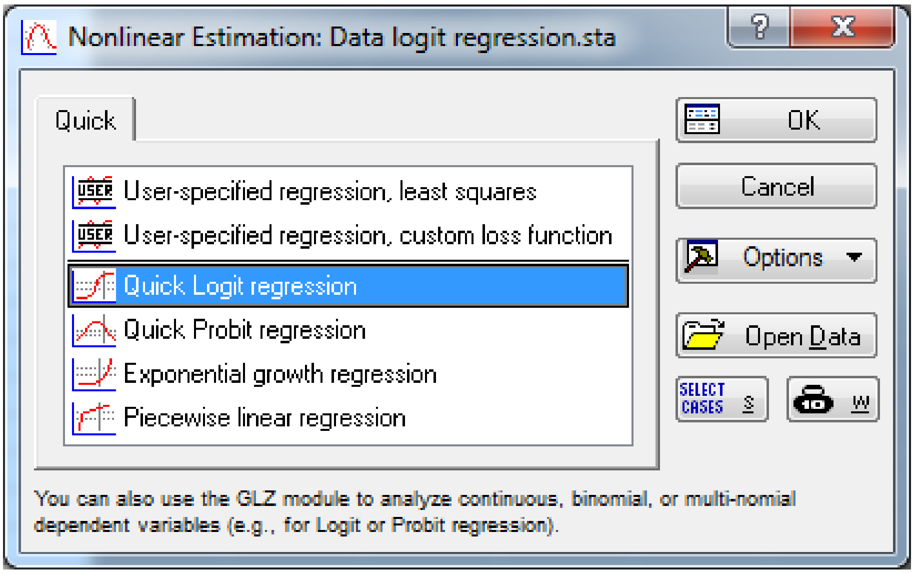

Fig. 3. Nonlinear Estimation: Data logit regression.

The calculation of the model parameters is presented in Figure 4, and evaluation of the model parameters – is in Figure 5.

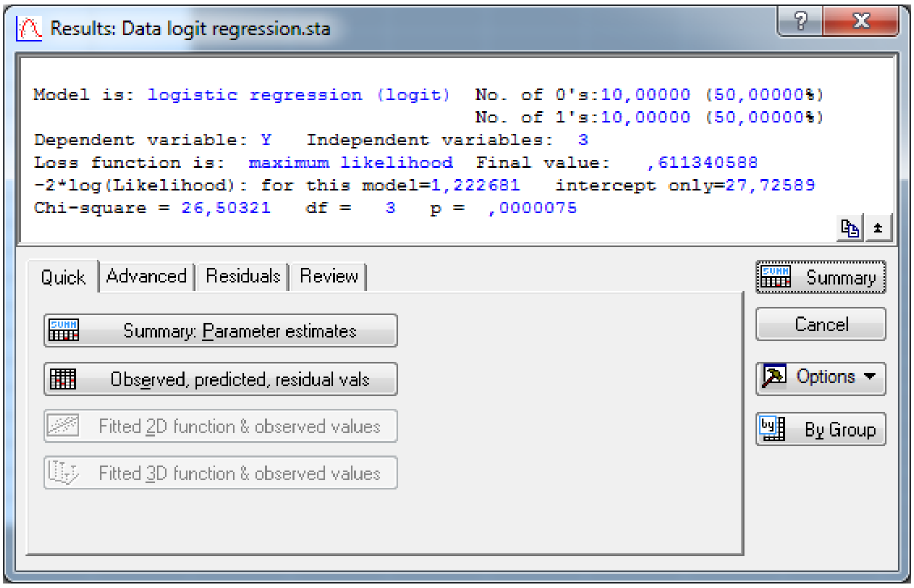

Fig. 4. Calculation of logistic regression parameters in the package Statistica.

	Model: Logistic regression (logit) N of 0's: 10 1's: 10 (Data logit regression.sta) Dep. var: Y Loss: Max likelihood Final loss: ,611340588 Chi?(3)=26,503 p=,00001			
N=20	**Const.B0**	F1	F2	F3
Estimate	-35,3750	0,792834	3,70021	-0,334255
Odds ratio (unit ch)	0,0000	2,209650	40,45594	0,715871
Odds ratio (range)			40,45594	0,000002

Fig. 5. The ratings of the logistic regression parameters.

Accordingly, for the initial data, one can use a formula with the following logistic regression coefficients, the calculation results of which are presented in Table 1.

$$P(Y \mid F) = \frac{1}{1 + e^{-(-35{,}375 + 0{,}793 F_1 + 3{,}700 F_2 - 0{,}334 F_3)}} \qquad (3)$$

A conclusion is made by the calculated probability value $P(Y \mid F) > 0{,}5$, that the predicted value of a dependent variable must to be equal *Y=1*. The otherwise conclusion (*Y=0*) is made in the case of the calculated probability value $P(Y \mid F) \le 0{,}5$.

Based on data, outlined in Table 1 (a factual value of dependent variable Y and its predicted values), it can be resumed, the accuracy of the developed logistic model is about 95%, it has correctly predicted 19 values of the dependent amount.

Table 1. The results of calculation according to logistic regression.

№	F1	F2	F3	Y	Logistic regression $P(Y \mid F)$
1	45	1	52	0	0,000
2	62	2	32	1	1,000
3	52	2	36	1	0,772
4	92	2	34	1	1,000
5	78	2	41	1	1,000
6	30	1	50	0	0,000
7	48	1	60	0	0,000
8	74	2	67	1	1,000
9	64	1	38	1	0,998
10	56	2	56	0	0,092
11	75	1	57	1	1,000
12	36	1	48	0	0,000
13	45	2	28	0	0,160
14	98	2	37	1	1,000
15	42	1	36	0	0,000
16	30	2	44	0	0,000
17	58	2	46	1	0,933
18	86	2	62	1	1,000
19	35	2	42	0	0,000
20	54	1	47	0	0,990

Recently, artificial intelligence methods have become widely used [13-18]. Another tool that approximates nonlinear relationships well is artificial neural networks. [19]. As for the neural network that unites a large number of neutrons, it represents a powerful modeling tool, which allows reproducing quite complex dependencies. The software package STATISTICA tools (module STATISTICA Automated Neural Networks (SANN)), which is presented in Figure 6, can be also used for the solution of neural network tasks.

Among the obvious advantages of using neural networks, it can be highlighted a large sector of the use of this tool in various branches of science and production (data analysis, optimization, forecasting, marketing, and advertising, work with images, audio, and video resources, etc.).

Also, it must be highlighted the ability to filter input data, fast learning of neural networks, detection of symptoms when approaching a critical point. The calculated neural networks allow fulfilling the modeling according to incoming data accurate to 100% (as shown in Fig. 7).

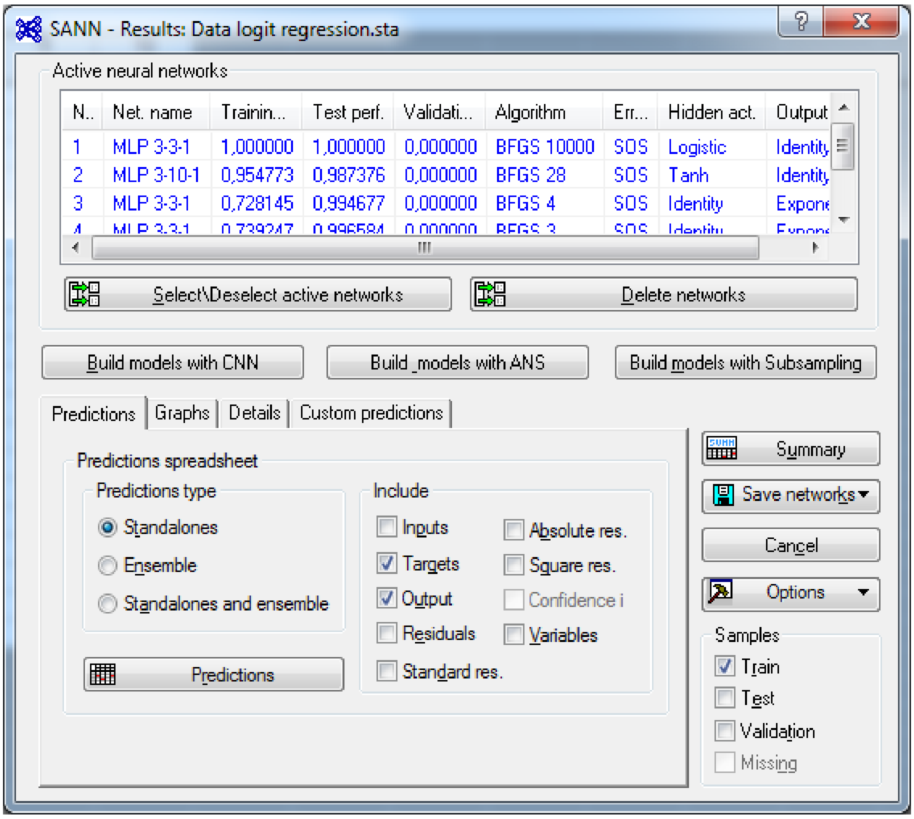

Fig. 6. A neural network parameters calculation.

Summary of active networks (Data logit regression.sta)

Index	Net. name	Training perf.	Test perf.	Validation perf.	Training error	Test error	Validation error	Training algorithm	Error function	Hidden activation	Output activation
1	MLP 3-3-1	1,000000	1,000000	0,00	0,000000	0,000000	0,000001	BFGS 10000	SOS	Logistic	Identity
2	MLP 3-10-1	0,954773	0,987376	0,00	0,010197	0,007914	0,003843	BFGS 28	SOS	Tanh	Identity
3	MLP 3-3-1	0,728145	0,994677	0,00	0,067782	0,098539	0,135573	BFGS 4	SOS	Identity	Exponential
4	MLP 3-3-1	0,739247	0,996584	0,00	0,064663	0,094862	0,137521	BFGS 3	SOS	Identity	Exponential
5	MLP 3-3-1	1,000000	1,000000	0,00	0,000000	0,000000	0,000000	BFGS 112	SOS	Exponential	Exponential

Fig. 7. The results of calculated models

Let us consider the network, which is selected on the minimum error criterion. It has a formula MLP 3-3-1. This network represents a multilayer perceptron, consisting of 3 neurons in incoming and hidden layers and one neuron in the outlet layer. The calculated values of weight coefficients for this network are presented in Fig. 8.

As shown in Fig. 9, the values of input signs, output binary variable, and calculated values of network exit suggest, that the developed neural network has correctly predicted all 20 values of the dependent variable.

Weight ID	Network weights (Data logit regression.sta) Connections 1.MLP 3-3-1	Weight values 1.MLP 3-3-1
1	F1 --> hidden neuron 1	109,853
2	F2 --> hidden neuron 1	87,091
3	F3 --> hidden neuron 1	-62,360
4	F1 --> hidden neuron 2	-13,116
5	F2 --> hidden neuron 2	16,455
6	F3 --> hidden neuron 2	-111,520
7	F1 --> hidden neuron 3	350,591
8	F2 --> hidden neuron 3	53,106
9	F3 --> hidden neuron 3	-113,225
10	input bias --> hidden neuron 1	2,454
11	input bias --> hidden neuron 2	-15,014
12	input bias --> hidden neuron 3	-123,479
13	hidden neuron 1 --> Y	0,000
14	hidden neuron 2 --> Y	-5,269
15	hidden neuron 3 --> Y	1,000
16	hidden bias --> Y	-0,000

Fig. 8. Weight coefficients of the model

Case name	Predictions spreadsheet for Y (Data logit regression.sta) Samples: Train, Test, Validation Sample	F1 Input	F2 Input	F3 Input	Y Target	Y - Output 1. MLP 3-3-1
1	Train	45	1	52	0,000000	-0,000000
2	Train	62	2	32	1,000000	1,000000
3	Train	52	2	36	1,000000	1,000000
4	Train	92	2	34	1,000000	1,000000
5	Validation	78	2	41	1,000000	1,000000
6	Train	30	1	50	0,000000	-0,000000
7	Train	48	1	60	0,000000	-0,000000
8	Validation	74	2	67	1,000000	1,000000
9	Train	64	1	38	1,000000	1,000000
10	Train	56	2	56	0,000000	0,000000
11	Validation	75	1	57	1,000000	0,997321
12	Train	36	1	48	0,000000	-0,000000
13	Train	45	2	28	0,000000	-0,000000
14	Train	98	2	37	1,000000	1,000000
15	Train	42	1	36	0,000000	0,000000
16	Train	30	2	44	0,000000	0,000000
17	Test	58	2	46	1,000000	0,999966
18	Test	86	2	62	1,000000	1,000000
19	Test	35	2	42	0,000000	0,000000
20	Train	54	1	47	0,000000	0,000000

Fig. 9. Table of model results comparison

Also, it should be noted, that software package STATISTICA tools make it possible to calculate values of output binary dependent variable for various user's values of input variables, not only for those present in input data.

Conclusion

According to the comparison of the results, received by logistic regression, and that, received by the neural network, a conclusion is made, that using neural networks is a more accurate modeling method of complete nonlinear dependencies.

An approach with using neural networks may be difficult to a certain degree for those specialists, ho not able to use these tools. The use of a logistic regression equation is one of the more simple modeling methods of binary dependencies. Its characteristics can be obtained even as a result of calculations, for an instant, in an Excel spreadsheet.

A set of factors, in addition to those chosen by the authors, may also contain such factors, as the level of culture of scientific activity, the level of methodical culture, the level of educational activity, sex, the presence of children in employees (their number), the gradation of children by age categories (for example, up to 3 years, from 3 to 7, primary school age, etc.), the availability of own housing, living conditions, etc.

The developed method is the basis of the decision support system that allows strategic development of high education organizations based on effective investment in the human capital of its employees.

References

1. Šlaus, I. and Jacobs, G. Human capital and sustainability // Sustainability, 3(1), 2011, pp. 97–154
2. The Global Human Capital Report 2017. Available at: https://www.weforum.org/reports /the-global-human-capital-report-2017 (accessed 20 February 2020).
3. «The National Doctrine of Education in the Russian Federation» dated October 4, 2000, No. 751. Available at: http://www.rg.ru/2000/10/11/doktrina-dok.html (accessed 20 February 2020). (In Russ.).
4. Becker, G.S. Human capital. A theoretical and empirical analysis, with special reference to education. Chicago, University of Chicago Press, 1993.
5. Mincer, J. Human capital and the labor market. A review of current research. Educational Researcher, 1989, vol. 18, no.4, pp.27–34.
6. Heckman, J.J., Lochner, L.J., Todd, P.E. Earnings functions and rates of return. Journal of Human Capital, 2008, vol. 2, no. 1, pp. 1–31.
7. Mankiw, N.G., Romer, D., Weil, D.N. A contribution to the empirics of economic growth. Quarterly Journal of Economics, 1992, vol. 107, no. 2, pp. 407–437.
8. Barro, R.J. Education, and economic growth. Annals of Economics and Finance, 2013, vol. 14, iss. no.2, pp. 301–328.
9. Porokhnya, V.M. Intelektualʹnyy kapital ekonomichnoho zrostannya: navch. posibn. [Intellectual capital of economic growth: textbook.]. Zaporizhzhya: KPU, 2012. (In Ukr.).
10. Kel'chevskaya, N. R., Shirinkina, E. V. Regional'nye determinanty effektivnogo ispol'zovaniya chelovecheskogo kapitala v cifrovoj ekonomike [Regional determinants of

effective use of human capital in the digital economy]. Ekonomika regiona [Economy of region]. 2019, № 15(2), pp. 465-482. (In Russ.).
11. Hosmer D.W., Lemeshow S., Sturdivant R.X. Applied logistic regression. A Wiley-Interscience Publication, 2000.
12. Harrel F.E. Regression modeling strategies. N.Y.: Springer, 2001.
13. Jevšček M. Competencies assessment using fuzzy logic // Journal of Universal Excellence. 2016, no.2, p. 187-202.
14. Houe R., Grabot B., Tchuente G. Fuzzy logic in competence management // European Society for Fuzzy Logic and Technology: 7th conference. 2011, p.651-656.
15. Macwan N., Srinivas S. Performance Appraisal using Fuzzy Evaluation Methodology // International Journal of Engineering and Innovative Technology. 2013, no.3, p. 324-329.
16. Krichevskiy, M. L., Martynova, Yu. A. Instrumenty iskusstvennogo intellekta pri otsenke effektivnosti investitsionnogo proekta [Instruments of artificial intelligence in assessment of effectiveness of investment project]. Kreativnaya ekonomika. 2018, № 12, pp. 1105-1118. DOI: 10.18334/ce.12.8.39265. (In Russ.).
17. Ostapenko, I. N. and Usenko, R. S. Modelirovaniye pokazatelya urovnya tvorcheskogo razvitiya lichnosti [Modeling an indicator of the level of creative development of a person]. Digital economy: information technology and models: monograph. 2018, pp.258-282. (In Russ.).
18. Gorbachevskaya, E. N., and Leonidov, A. V. Model' neyronnoy seti dlya reytingovoy otsenki kompetentnosti sotrudnikov [Neural network model for ranking employees competency]. Nauchnyj zhurnal «Vestnik Volzhskogo universiteta imeni V.N. Tatishcheva» [The scientific journal "Bulletin of the Volga University named after V.N. Tatishchev "]. 2015, № 1(23), pp.57-71. (In Russ.).
19. Haykin S. Neural Networks: a comprehensive foundation. Moscow, Publishing house "Williams". 2006, 1104 p. (In Russ.).