DOI https://doi.org/10.15407/csc.2021.01.035
UDC 364.2:331

YE.A. SAVCHENKO-SYNYAKOVA, PhD (Eng.), Senior Research Associate, International Research and Training Centre of Information Technologies and Systems of the NAS and MES of Ukraine, Glushkov ave., 40, Kyiv, 03187, Ukraine, savchenko_e@meta.ua

O.V. TUTOVA,PhD (Econ.), Senior Research Associate, International Research and Training Centre of Information Technologies and Systems of the NAS and MES of Ukraine, Glushkov ave., 40, Kyiv, 03187, Ukraine, olenatutova@gmail.com

THE TOOLS FOR INTELLIGENT DATA ANALYSIS, MODELING AND FORECASTING OF SOCIAL AND ECONOMICAL PROCESSES

The tools for intelligent data analysis, modelling, and forecasting were developed in MS Excel based on the procedure of analysis, modelling, and forecasting of complex processes by the data base of social and economical processes. This data base contains information about the indicators that characterize development of digital economy in Ukraine. It will give an opportunity to automate and improve the decision making process in the field of social and economical development of Ukraine at the regional level as well as in comparison with other countries.

Keywords: *tools, data analysis, modelling, forecasting, human development level, GMDH.*

Introduction

Nowadays many tools have been developed to solve the problems of analysis, modeling and forecasting, but the problem of building user-friendly tools in the economic field remains relevant.

Social and economic data often contain gaps. They may be unclear and incomplete. Often not all factors influencing the process under study are taken into account in the data. Expert data is often used. They have a subjective component, i.e. depend on the expert.

GMDH is one of the methods that allow building models under such conditions, which is confirmed by the wide practice of its application. This is a method of automatic model construction based on expert observational data [1-3]. This method has demonstrated its effectiveness in many fields of human activity.

Therefore, based on this method, a toolkit for data analysis, modeling and forecasting was developed. It focuses on solving the problems of the digital economy. One of the elements of such technology is a database (DB), which characterizes the development of Ukraine's digital economy. The structure of such a database will be shown below. It includes indicators that characterize the development of information and communication technologies (ICT) in Ukraine.

Review of the Application of Intelligent Modeling Methods to Solve Economic Problems

The methods of intelligent modeling are widely applied for scientific, research, educational, and commercial use.

There are many companies in the commercial field nowadays that proclaim that they develop intelligent models. They state that their mission is to create tools and methods that enable cost-effective application-oriented innovations about humans and their environments and blend together converging multi-disciplinary discoveries into actionable knowledge. They use tools and expertise to help their customers shorten the innovation cycle, develop offerings of high practical importance, and bring better products to the market faster.

They use advanced technologies, mathematics, and concepts to create powerful informational models.

For example, the Intelligent Modelling and Analysis (IMA) group of Nottingham University has established itself as a unique brand in the UK for end-to-end data modelling and analysis [4]. They represent themselves as a highly inter-disciplinary research group focusing on the development of models and techniques for real-world and multifaceted problems in data analysis. It includes researchers from a variety of backgrounds including computer science, the biomedical sciences, operational research, mathematics, statistics and complexity science.

They use a range of techniques including:

- Artificial Intelligence-based Data Mining;
- Bio-Inspired Algorithms;
- Computational Modelling;
- Discrete and Agent-Based Simulation;
- Fuzzy Methodologies;
- Multi-Sensor Data Fusion;
- Qualitative Methods including Structured Interviews.

IMA's main research objectives are:

- Modelling and representation of challenging problems, with particular emphasis on biomedical and digital economy application domains;
- Creating cutting-edge analysis methodologies, both for general purposes and specifically tailored to the main application domains;
- Focusing on difficult, challenging and important real-world problems, with particular emphasis on large and noisy data sets.

IMA fulfils projects in various fields including digital economy and sustainable development.

The aim of their programme Sustaining Urban Habitats is to develop a distinctively interdisciplinary approach to producing and evaluating scenarios for sustainable living in urban habitats. With two transition cities in Europe (Nottingham, Stuttgart) and two growth cities in China (Chengdu, Shanghai) as the empirical focus, ways of combining environmental and economic modelling with social and cultural ethnographic work are explored.

The Network for Integrated Behavioural Science is a partnership among the Universities of Nottingham, Warwick and East Anglia. It is funded by the Economic and Social Research Council. It is a cross-disciplinary group of researchers who develop and test models of human behaviour and behavioural change, and draw out their implications for the formulation and evaluation of public policy.

Prototyping Open Innovation Models for ICT-Enabled Manufacturing in Food Packaging Research intends to harness and develop ideas from the social media sphere, using the power of the 'crowd' to develop new products and types of packaging within the food industry. Involving a number of major industry partners the research will focus on designing a platform of ICT tools for state-of-the-art manufacturing processes allowing customers to be co-creators whilst products are being developed. The ICT tools and platform will gather content from users, organise those ideas, and integrate them with design and production systems. The advantage of using consumers in product development will be a shorter time to market for new products, and the ability to integrate the design and manufacturing process with people, so enabling customers to have 'conversations' with brands and manufacturers.

Horizon Digital Economy Hub will tackle the challenge of harnessing the power of ubiquitous computing for the digital economy in a way that is acceptable to our society and increases the quality of life for all. This will involve establishing a world-leading and sustainable centre of excellence for research and knowledge transfer for the ubiquitous digital economy.

Multiscale Modelling to Maximise Demand Side Management (DSM) is an answer how the significant potential benefits of DSM can be maximised through the provision of a unified, versatile and affordable digital infrastructure that allows us to reason across a whole energy system and supports new ways to exchange information between dynamic multiscale DSM models.

The project Potential of Psychological Information to Inform Credit Scoring utilizes psychological characteristics of Personality Profiles in order to guide Data Mining techniques that can explain economic behaviours and improve credit scoring models.

The Towards Data-Driven Environmental Policy Design research project is an inter-disciplinary collaboration between the Western Australian Government and the School for Computer Science at the University of Nottingham. It combines advances in uncertain data capture and computational intelligence with value-driven planning and policy making processes to provide data-driven conservation planning and policy making support mechanisms and tools which are evaluated in real-world conservation settings in Western Australia.

The Opening Developing World Markets by Using Personal Data and Collaboration project creates a platform that can provide businesses with information about those behavioural groups based on data streams that promote emerging economies through a combination of academic expertise; market intelligence expertise; and the combined experience of an extensive network of UK and international companies.

One of the most challenging issues in today's large-scale computational modeling and design is to effectively manage the complex distributed environments, such as computational clouds, grids, ad hoc, and P2P networks operating under various types of users with evolving relationships fraught with uncertainties. In this context, the IT resources and services usually belong to different owners (institutions, enterprises, or individuals) and are managed by different administrators. Moreover, uncertainties are presented to the system at hand in various forms of information that are incomplete, imprecise, fragmentary, or overloading, which hinders in the full and precise resolve of the evaluation criteria, subsequencing and selection, and the assignment scores. Intelligent scalable systems enable the flexible routing and charging, advanced user interactions and the aggregation and sharing of geographically-distributed resources in modern large-scale systems.

In [5] present ideas, theories, models, technologies, system architectures and implementation of applications in intelligent scalable computing systems. Several important Artificial Intelligence-based techniques, such as fuzzy logic, neural networks, evolutionary, and memetic algorithms are studied and implemented. All of those technologies have formed the foundation for the intelligent scalable computing.

The increasing complexity of logistic networks calls for a paradigm change in their modeling and operations. Centralized control is no longer a feasible option when dealing with extremely large systems. For this reason, decentralized autonomous systems are gaining popularity in providing robustness and scalability. The chapter Intelligent Modeling and Control for Autonomous Logistics of [5] focuses on the use of intelligent systems in autonomous logistics. Specifically, it describes issues related to knowledge management, a machine learning-based approach to adaptability and planning, and intelligent optimization by autonomous logistics entities.

It is undeniable the high cognitive load that modelers face to understand a multitude of complex abstractions and their relationships, and the urgent need to support tool builders to provide modelers with intelligent modeling assistance [6].

However, current Intelligent Modeling Assistants (IMAs) lack adaptability and flexibility for tool builders, and do not facilitate understanding the differences and commonalities of IMAs for modelers.

From a global perspective, a key challenge is to identify modeling activities where intelligent assistance is relevant, together with required models and data. An immediate concern for the modeling community is how to determine the, possibly evolving, requirements for an IMA to provide timely assistance based on confidence and relevance as well as modeler's skill and trust. The grand challenge is to make functionality for modeling assistance reusable across a wide range of domains.

Another example of use of intelligent modeling for commercial purposes is intelligent modeling of historical and asbuilt plant data for an efficient decision making [7]. The company that proposes these services facilitates an easy-to-access, centralized, repository data model, consisting of both structured and unstructured data, to owner operators and Energy Performance Contractors. Energy Performance Contractor means an engineering, procurement, and construction contractor with substantial experience in the engineering, procurement, and construction of power plants.

Over the entire period of a plant's life cycle, a large volume of data, drawings, and documents is collected. This data, both structured and unstructured, is classified into different groups, including numerical, categorical, historical, and real-time, and is then collected and modeled.

Analysis and visualization of this data helps companies enhance exploration and production, improve refining and manufacturing efficiency, and optimize global operations, while ensuring safety and environmental protection. Capital and operational expenditures can be reduced with analytics from modeled data.

The risks can be reduced by analyzing the data, modeled from a vast amount of plant operations collected data over many years. It requires the expertise in 2D drafting, CAD conversion, 3D modeling from 2D drawings, and TEF and Non-TEF data integration to transform the data into a strategic business intelligence repository, allowing enhance plant's efficiency and optimize its global operations, while ensuring safety and environmental protection.

Designing a Database in the Field of ICT

Data collection and systematic analysis of the subject area is the first and most important stage of database design [8]. Here it is necessary to make a detailed verbal description of the objects of the subject area and the real connections between the objects of this branch. The stages of designing a database of processes in the field of ICT will be considered.

The level of development of ICT is the basis for the digital transformation and the digital economy, as the digital economy is an economy based on digital computer technology. Digital transformations are one of the main factors of world economic growth.

Currently, there are several international rankings that directly or indirectly characterize the level of development of ICT and the maturity of e-government tools in different countries. There are ICT Development Index, Networked Readiness Index, E-government development rank, E-Participation Index, United Nations (UN) e-Government Readiness Index, Cybersecurity Index, Knowledge Economy Index, Global Innovation Index, The Digital Economy and Society Index.

The ICT Development Index is a composite index that combines 11 indicators into one benchmark measure [9]. It is used to monitor and compare developments in ICT between countries and over time. As the UN special agency for ICTs, International Telecommunication Union (ITU) is the official source for global ICT statistics. One of their core activities is the collection, verification and harmonization of telecommunication/ICT statistics for about 200 economies worldwide. There are two key sets of telecommunication/ICT data that ITU collects directly from countries:

- Telecommunication/ICT data collected from national telecommunication/ICT ministries and regulatory authorities: these include data on the fixed-telephone network, mobile-cellular services, Internet/broadband, traffic, revenues and investment; and prices of ICT services.
- Household ICT data collected from national statistical offices: these include data on household access to ICTs and individual use of ICTs.

The World Economic Forum's Networked Readiness Index, also referred to as Technology Readiness, measures the propensity for countries to exploit the opportunities offered by ICT [10]. It is regarded as the most authoritative and comprehensive assessment of how ICT impacts the competitiveness and well-being of nations. It ranks a total of 134 economies based on their performance across 60 variables. Recognizing the pervasiveness of digital technologies in today's networked world, the Index is grounded in four fundamental dimensions: Technology, People, Governance and Impact. This holistic approach means that the NRI covers issues ranging from future technologies such as Artificial Intelligence (AI) and the Internet of Things (IoT) to the role of the digital economy in reaching the Sustainable Development Goals (SDG).

Technology is at the heart of the network economy. This group of indices is meant to assess the level of technology that is necessary for a country's participation in the global economy. The following three sub-indices have been identified for that purpose:

- *Access*: The fundamental level of ICT in countries, including on issues of communications infrastructure and affordability.
- *Content*: The type of digital technology produced in countries, and the content/applications that can be deployed locally.
- *Future Technologies*: The extent to which countries are prepared for the future of the network economy and new technology trends such as AI and IoT.

The availability and level of technology in a country is only of interest insofar as its population and organizations have the access, resources, and skills to use it productively. The group of indices People is therefore concerned with the application of ICT by people at three levels of analysis: individuals, businesses, and governments.

- *Individuals*: How individuals use technology and how they leverage their skills to participate in the network economy.
- *Businesses*: How businesses use ICT and participate in the network economy.
- *Governments*: How governments use and invest in ICT for the benefit of the general population.

The group of indices Governance describes how safe individuals and firms are in the context of the network economy. This does not only relate to actual crime and security, but also to perceptions of safety and privacy. Regulation sub-indices show the extent to which the government promotes participation in the network economy through regulation. Inclusion sub-indices characterize the digital divides within countries where governance can address issues such as inequality based on gender, disabilities, and socioeconomic status.

Ultimately, readiness in the network economy is a means to improve the growth and well-being in society and the economy. This pillar therefore seeks to assess the economic, social, and human impact of participation in the network economy.

The group of indices Impact includes three sub-indices:

- *Economy*: The economic impact of participating in the network economy.
- *Quality of Life*: The social impact of participating in the network economy.
- *SDG Contribution*: The impact of participating in the network economy in the context of the SDGs — the goals agreed upon by the UN for a better and more sustainable future for all. The focus is on goals where ICT has an important role to play, including such indicators as health, education, and environment.

The E-Government Development Index presents the state of E-Government Development of the UN Member States [11]. Along with an assessment of the website development patterns in a country, it incorporates the access characteristics, such as the infrastructure and educational levels, to reflect how a country is using information technologies to promote access and inclusion of its people. This index is a composite measure of three important dimensions of e-government, namely: provision of online services, telecommunication connectivity and human capacity. It is not designed to capture e-government development in an absolute sense; rather, it aims to give a performance rating of national governments relative to one another.

Promoting participation of the citizenry is the cornerstone of socially inclusive governance. The goal of e-participation initiatives should be to improve the citizen's access to information and public services; and promote participation in public decision-making which impacts the well-being of society, in general, and the individual, in particular.

The e-participation index is derived as a supplementary index to the UN E-Government Survey [12]. It extends the dimension of the Survey by focusing on the use of online services to facilitate provision of information by governments to citizens ("e-information sharing"), interaction with stakeholders ("e-consultation"), and engagement in decision-making processes ("e-decision making").

E-Participation Framework:

- *E-information*: Enabling participation by providing citizens with public information and access to information without or upon demand
- *E-consultation*: Engaging citizens in contributions to and deliberation on public policies and services
- *E-decision-making*: Empowering citizens through co-design of policy option and co-production of service components and delivery modalities.

The UN e-government readiness index is an internationally agreed-upon composite index that measures the capacity of governments to develop and implement e-government services [13]. Constructed within the framework of the UN global e-government survey, the indicator consists of three sub-indices: the web measure index, the telecommunication infrastructure index and the human capital index. The UN's e-government readiness index is a combined indicator of the supply of, potential demand for and maturity of e-government services. This is generally characterised by an extensive broadband infrastructure; a repository of electronic information on government laws and policies, including links to archived information and downloadable forms; and a high level of comfort with ICT by citizens and businesses.

The Global Cybersecurity Index measures the commitment of countries to cybersecurity at a global level [14]. As cybersecurity has a broad field of application, cutting across many industries and various sectors, each country's level of development or engagement is assessed along five pillars – Legal Measures, Technical Measures, Organizational Measures, Capacity Development, and Cooperation – and then aggregated into an overall score. The ITU Global Cybersecurity Agenda provides the general foundation and framework for the initiative. For each of the five pillars country commitment was assessed through a question-based online survey, which further allowed for the collection of supporting evidence.

The "knowledge economy" is a concept of economic development, in which innovation and access to information drive productivity growth [15]. New trends, such as the IoT or digitalisation, are examples of the transition towards to the knowledge economy. To measure knowledge economy development, the European Bank for Reconstruction and Development has constructed its Knowledge Economy Index, spanning 46 economies. It uses indicators like institutional and legal frameworks (as a basis for patents etc.), number of technical graduates, research spending number of patents, some measure of collaboration, and amount of venture capital. In total there are 38 contributing indicators described in the index methodology.

Innovation is widely recognized as a central driver of economic growth and development. The aim of the Global Innovation Index is to provide insightful data on innovation and, in turn, to assist economies in evaluating their innovation performance and making informed innovation policy considerations [16]. It is now considered a criterion for measuring innovation by the UN General Assembly, as noted in its resolution on Science, Technology and Innovation for achieving SDGs at its 74th session in 2019. There are 5 pillars in the Input Sub-index and 2 pillars in the Output Sub-index. Each pillar is divided into three sub-pillars, each of which is composed of individual indicators, a total of 80 this year.

The Digital Economy and Society Index is a composite index that summarises relevant indicators on Europe's digital performance and tracks the evolution of EU Member States in digital competitiveness [17]. European Commission services selected various indicators, divided into thematic groups, which illustrate some key dimensions of the European information society (Telecom sector, Broadband, Mobile, Internet usage, Internet services, eGovernment, eCommerce, eBusiness, ICT Skills, Research and Development). These indicators allow a comparison of progress across European countries as well as over time.

Based on an analysis of those indices the ICT indicators that affect the development of Ukraine's digital economy were selected in order to design a database that helps to study the interrelations between those chosen ICT indicators and social and demographic factors and to understand how the development of digital economy contributes to the human well-being in Ukraine.

Such a database includes nine main groups of indicators:

- Indicators that characterize the level of development of the infrastructure for efficient use of ICT;
- Indicators that characterize the level how population uses ICT:
- Indicators that characterize the level how business uses ICT:
- Indicators that characterize the level how the authorities use ICT and promote their development;
- The level of education of population that helps to use ICT efficiently;
- Indicators that characterize how favourable the conditions for business activity are;
- Indicators that characterize the level of efficiency of the law machinery;
- Social and demographic indicators;
- Economic indicators.

The block diagram and affiliation of these indicators with a separate block is shown in Fig.1.

Building the relationships between the indicators of the database and analyzing them, it can be determined how the change of some indicators affects the change of others. For example, how individual macroeconomic indicators can affect Ukraine's GDP growth.

The following is a description of the methodology and tools for data analysis, modeling and forecasting of the digital economy.

Methods of Data Analysis, Modeling and Forecasting of Indicators that Characterize the Development of the Digital Economy

Methods of data analysis, modeling and forecasting of the social and economical indicators are described in detail in [18]. The deve-lopment of this technique began with the solution of real applied problems of modeling and finding relationships in the data.

Correlation-regression analysis and combinatorial algorithm of GMDH are used to build models here. The advantage of the combinatorial algorithm is that it does not require convergence, as it performs a complete search of all possible models. This algorithm solves the problem of structural-parametric identification, i.e. it allows finding not only the parameters of the model, but also its optimal structure.

Fig. 2 shows a block diagram of the successive stages of the developed technique.

The developed methodology includes the following separate stages:

The data sample contains information on indicators that describe the level of human development of countries or one of the selected regions, the state of development of the digital economy, the level of ICT use.

Restore gaps in data. There are many different ways to recover gaps in data, depending on the form of the gap itself, as well as the availability of sufficient information to restore it. In the proposed method, the combinatorial algorithm of GMDH was used to recover gaps [19]. To recover each gap, a sample of data is compiled and a model is built. That is, how many gaps in the data, so many models need to be built.

Data scaling. Economic data are usually measured in different units and ranges. Standardization or scaling of data is used to reduce them to one format. To do this, the value of each indicator is divided by the maximum value of this indicator.

Extending the data sample. In order to possibly increase the accuracy of the model, a method of expanding the sample is used. New additional indicators are introduced. They can be products, sums and other functions of input indicators [20].

Data analysis. This block includes two methods of analysis: by correlation analysis and GMDH. Correlation analysis is used to analyze the degree of influence of each of the indicators on the output value and the selection of informative indicators, which is carried out by calculating the pair correlation coefficient of one of the indicators with the output value. Using the combinatorial algorithm of GMDH for data analysis, a model can be built and analyzed: what indicators and with what coef-

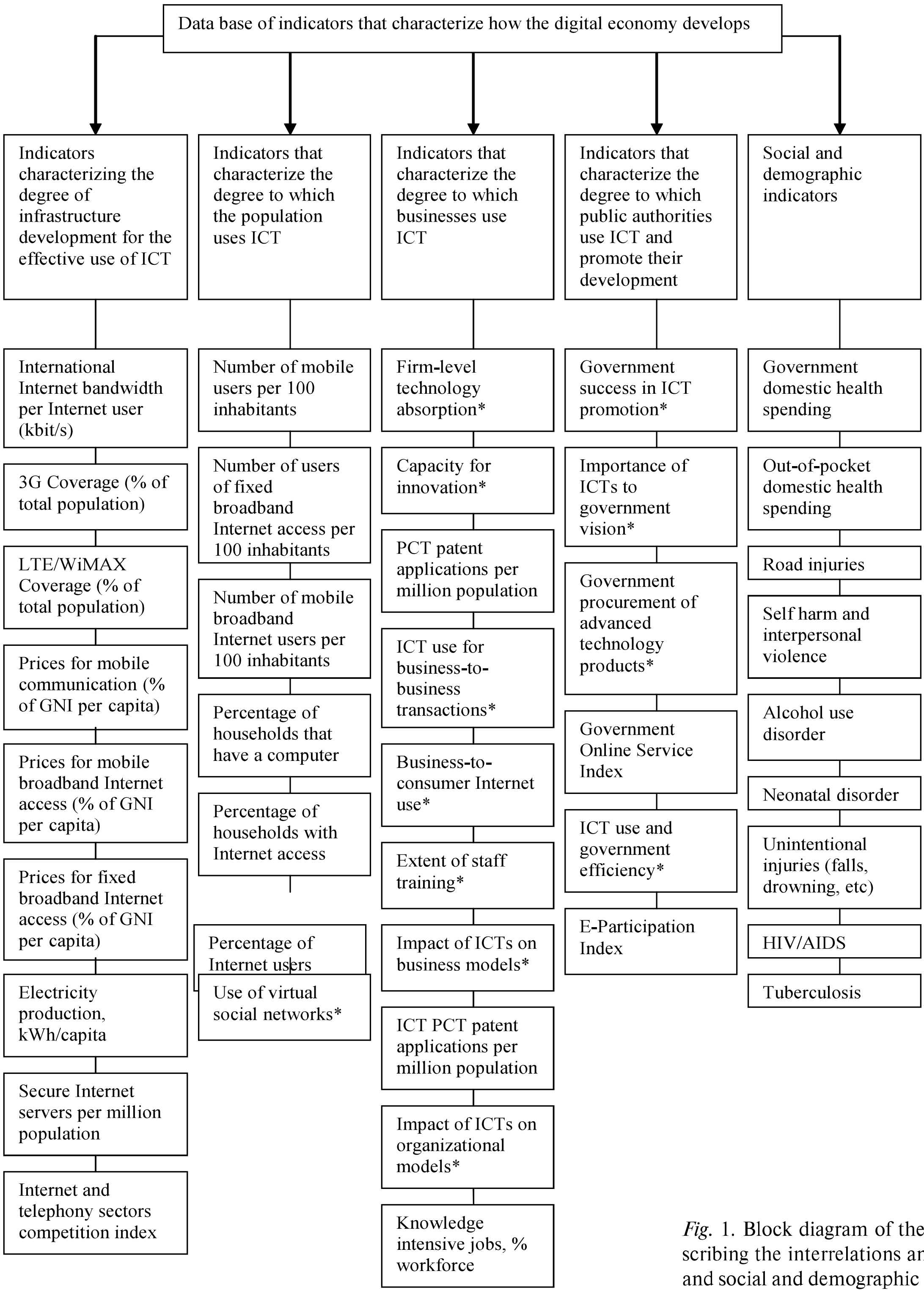

Fig. 1. Block diagram of the database describing the interrelations among the ICT and social and demographic indicators

* Marked indicators obtained with the help of expert assessments

ficients were included in it, and then draw conclusions about the existing dependencies in the data.

Building a model. Carried out in two ways:

1) Regression models are constructed, ie coefficients are determined using the least squares method (LSM); at the same time the structure of the model is given.

2) The models are found using the combinatorial algorithm of GMDH according to the given external criteria. Thus we define not only parameters of model, but also its structure.

Depending on which of the external criteria is set as the main criterion for selecting a model, models with different properties will be obtained. The main external criterion that is responsible for the accuracy of the model is the external criterion of regularity. The bias criterion is a criterion that reflects the requirement that the best models obtained in different parts of the sample differ minimally. In the theory of GMDH the criterion of unbiased decisions and errors is known [21].

Comparison of models on an independent sample. The obtained models are compared on some independent (examination) data sample, which was not used to obtain them. The results of research confirm the fact known in inductive modeling that regression models give a good result only on the data on which they are obtained, in contrast to the models of GMDH, which give quite good results on new data [22].

Analysis of the obtained models and development of recommendations. Models are analyzed both in terms of the obtained parameters, how each of them included in the model affects the given output value (this is indicated by the value of the coefficient at the parameter and its sign), and in terms of model accuracy. It can be analyzed, how a change in one or more parameters (decrease or increase) affects the value of the original value (by how many percent it will change it).

Based on the analysis of the models, it is possible to develop recommendations for specialists in this field on the relationship between input and output parameters. Based on the analysis of the coefficients of the model and their signs, it can be defined what contribution each indicator makes to the output value.

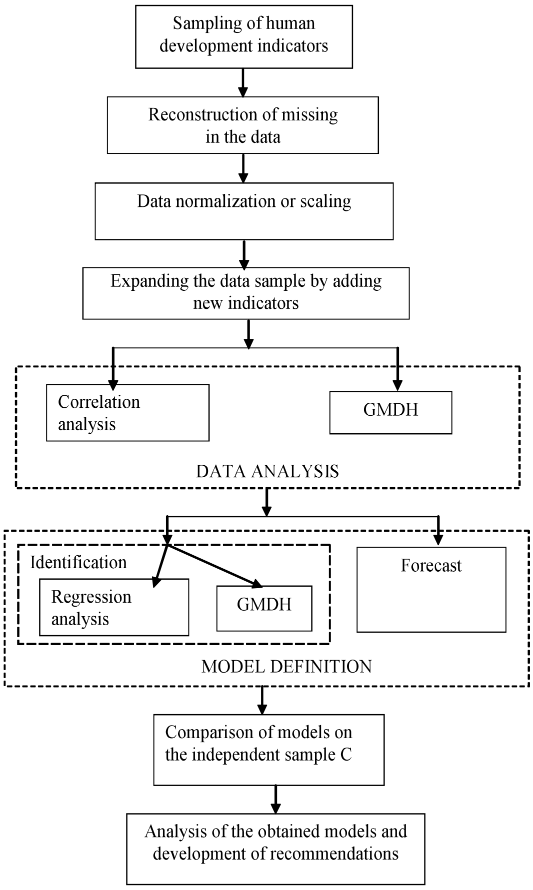

Fig. 2. Flowchart for the solving of problems of data analysis, modeling and forecasting

Development of Tools for Data Analysis, Modeling and Forecasting of Digital Economy Indicators

The toolkit designed to solve the problems of the analysis of dependences in data is realized on the basis of the developed technique. Thus it is meant for opening of new knowledge, construction of dependences in data and forecasting of variables.

This toolkit was developed using MS Excel. The combinatorial algorithm is carried out by G.O. Ivakhnenko using VBA MS Excel.

The main steps of how to use this tool should be considered in more detail.

Assessment of the arguments' comprehension. Suppose that the data sample $X\,[nxm]$ and the initial value $Y\,[nx1]$ are given. An example of them is shown in Fig. 3.

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
x1	27,8	28,7	28,6	28,5	28,1	26,8	26,2	24,6	21,6	19,8
x2	119,6	121,1	119,1	117,1	123	130,3	138,1	144,1	144	132,6
x3	1,72	3,46	4,15	6,45	6,96	8,04	8,86	9,31	11,81	12,22
x4	0,6	1,8	2,95	4,1	4,4	5,4	6,7	7,5	8,1	22,6
x5	16,2	21,2	23,2	25,2	30,7	40,5	40,5	52,4	59,2	65,1
x6	6,8	10,3		22,2	26	35,6	43,7	43	50,2	54,8
x7	6,55	11,00	17,90	23,30	28,71	35,27	40,95	46,24	48,88	53,00
x8	3154	5477	10194,5	14912	11332	49818	52883	40704	73425	79885
y	8374	8666	7330	7715	8089	8195	8195	8147	7375	7593

Fig. 3. Example of a data sample for assessment of arguments' comprehension

X1	1									
X2	0,9748	1								
X3	0,7832	0,9035	1							
X4	0,6799	0,8263	0,9863	1						
X5	0,7876	0,9052	1,0000	0,9873	1					
X6	0,3995	0,5939	0,8750	0,9439	0,8796	1				
X7	0,3907	0,5861	0,8703	0,9407	0,8749	1,0000	1			
X8	0,4319	0,6221	0,8917	0,9550	0,8959	0,9994	0,9990	1		
X9	0,3397	0,5408	0,8419	0,9206	0,8471	0,9979	0,9985	0,9950	1	
Y	*0,3346*	*0,5363*	*0,8390*	*0,9185*	*0,8442*	*0,9976*	*0,9982*	*0,9944*	*1,0000*	*1,0000*

Fig. 4. Example of calculation of a pairwise correlation coefficient

Main	criterion	=	0.7989045	Model	1	(	8)			
Additiona	criterion	=	0.75741415E+04							
Mean	square	dev.	(MSE)	=	0.27763714E+03					
Mean	absolute	dev.	(MAE)	=	0.20739999E+03					
Correlatio	coefficien	=	0.326056829							
R^2	=	0.106313056								
Mean	relative	dev.	(MAPE)	=	2.536	%				
	Coefficient		Variable							
a	0		8369.2271503834							
a	1		-0.0048856053		X	8				
Errors	on	validation	sample:	MAPE	=	6,8	%	MSE	=0.2764E+06	
Model	values	calculatec	on	validation	subsample	(not	normilized):			
	No.		Calculated		Real	value		Abs.deviation		
	9		0.80105E+04		0.73750E+04		-.63550E+03			
	10		0.79789E+04		0.75930E+04		-.38594E+03			

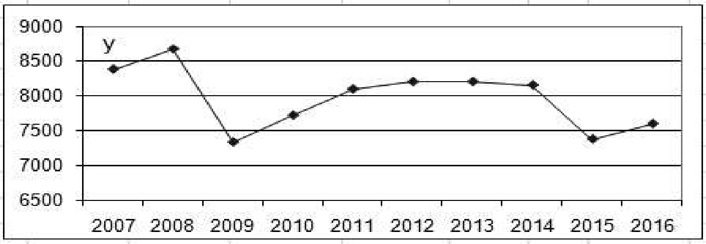

Fig. 5. Example of the results of modeling by the combinatorial algorithm of *GMDH*

Performance is assessed using correlation analysis. An example of how to calculate pairwise correlation between variables is shown in Fig. 4.

Then, based on the selected variables, models are built according to the combinatorial algorithm of GMDH.

When building a model, part of the observations (rows) can be left for verification and automatically calculate the value of the accuracy criterion on this data, which will check the adequacy of the obtained models.

An example of the obtained results by the combinatorial algorithm of GMDH is shown in Fig. 5.

Such models can be obtained to analyze interrelations in data as well as to recover data gaps. Moreover, in order to restore each gap, an own sample of data is compiled and a new model is built each time. Furthermore, the values of statistical criteria such as MAPE, MSE, R2 are calculated to verify the adequacy of the obtained models.

Conclusion

These tools were developed in order to build dependencies among the indices in social and economic problems, to analyze them, and to develop recommendations. Therefore, the level and evolution of information and communication technology developments in Ukraine and its experience relative to other Eastern European and post-Soviet countries based on the Information and Communication Technology Development Index over 2002-2017 were examined with help of these tools [23]. Also the connection between gross national income and this index, and how the components of the main ICT indices influence gross national income in these countries was studied by these methods [24]. These tools will be applied in the further research for modeling, forecasting, and decision-making to evaluate the interconnection between the health care system and the overall economical and social progress of Ukraine in comparison with the neighbouring countries.

REFERENCES

1. Ivakhnenko A.G. Heuristic Self-organization in Problems of Engineering Cybernetics. Pergamon PRESS. Journal Automatica. 1970. №6. P. 207–219.
2. Madala H.R., Ivakhnenko A.G. Inductive Learning Algorithms for Complex Systems Modeling. Boca Raton: CRC Press Inc., 1994. 384 p.
3. Ivakhnenko, A.G, 1968. "Group Method of Data Handling as a Rival of Stochastic Approximation Method", Journal Soviet Automatic Control, 3, pp. 58-72. (In Russian).
4. Intelligent Modelling and Analysis. URL: https://www.nottingham.ac.uk/computerscience/research/ima.aspx
5. Advances in Intelligent Modelling and Simulation: Artificial Intelligence-Based Models and Techniques in Scalable Computing / Eds J. Ko odziej, S.U. Khan, T. Burczy ski. – Berlin–Heidelberg: Springer-Verlag, 2012. – 381 p.
6. Opportunities in Intelligent Modeling Assistance. URL: https://modeling-languages .com/opportunities-intelligent-modeling-assistance/.
7. Intelligent modeling of historical and as-built plant data for an efficient decision making. URL: https://www.avineon.com/en/intelligent-data-modeling
8. Kuznetsov, S.D., 2007. Basics of databases: a tutorial. 2nd ed. M .: Internet University of Information Technologies. Binomial. Knowledge laboratory.
9. The ICT Development Index (IDI): conceptual framework and methodology. URL: https://www.itu.int/en/ITU-D/Statistics/Pages/publications/mis/methodology.aspx
10. Networked Readiness Index. URL: https://networkreadinessindex.org/
11. E-Government Development Index (EGDI). URL: https://publicadministration.un.org/egovkb/en-us/About/Overview/-E-Government-Development-Index
12. E-Participation Index. URL: https://publicadministration.un.org/egovkb/en-us/About/Overview/E-Participation-Index
13. E-Government readiness. URL: https://www.oecd-ilibrary.org/docserver/9789264061651-33-en.pdf?expires=1621616606&id=id&accname=guest&checksum=A9A1AD23179D762594F1B0D1DC5211AB
14. Global Cybersecurity Index. URL: https://www.itu.int/en/ITU-D/Cybersecurity/Pages/global-cybersecurity-index.aspx

15. EBRD Knowledge Economy Index. URL: https://www.ebrd.com/news/publications/brochures/ebrd-knowledge-economy-index.html.
16. About the global innovation index. URL: https://www.globalinnovationindex.org/about-gii#keyfindings.
17. The Digital Economy and Society Index (DESI). URL: https://digital-strategy.ec.europa.eu/en/policies/desi.
18. Savchenko, Ye.A., Tutova, Ye.V., 2016. "A Procedure for the Human Potential Development Level Data Analysis, Modeling and Forecasting for Countries of the World and Ukrainian Regions". Upravl ie sistemy i ma iny, Issue 4 (264), pp. 86-92. DOI: https://doi.org/10.15407/usim.2016.04.086.
19. Savchenko, Ye.A., 2015. "Data Sample Preprocessing in the Tasks of Inductive Modelling". Upravl ie sistemy i ma iny, 2, pp. 82-87. (In Russian).
20. Savchenko, E.A., Stepashko, V.S., 2008. "Analyz selektyvnykh svoystv kryteryev MGUA pry ykh posledovatel nom prymenenyy. Modelyuvannya ta keruvannya stanom ekoloho-ekonomichnykh system rehionu. Zb. prats. K.: MNNTS IT ta S, pp. 199–210. (In Ukrainian).
21. Savchenko, E.A., Ivakhnenko, A.G. Investigation of efficiency of additional determination method of the model selection in the modeling problems by application of GMDH algorithm. Journal of Automation and Information Sciences, 2008, 40(3), pp. 47–58.
22. Koppa, Yu.V., Stepashko, V.S., 2002. "Measuring the efficiency of regression analysis and GMDH for predicting environmental processes". Pratsi Mizhnar. konf. zinduktivnogo modelyuvannya v 4-kh tomakh (MKIM-2002). T.3. L'v v: Derzhavniy NDI inf. infrastruktura, pp. 123–127. (In Ukrainian).
23. Tutova, O.V., Savchenko, Ye.A., 2019. "Ukraine in the Information and Communication Technology Development Ranking". Control systems and computers, 3, pp. 70-78.
24. Tutova, O., Savchenko, Ye., 2018. "Assessment of Influence of Digital Economy Development on Social and 6. Economic Environment", Upr. sist. ma ., 6 (278), pp. 81-92. (In Ukrainian).

Received 23.01.2021

ЛІТЕРАТУРА

1. Ivakhnenko A.G. Heuristic Self– organization in Problems of Engineering Cybernetics. Pergamon PRESS. Journal Automatica. 1970. №6. P. 207–219.
2. Madala H.R., Ivakhnenko A.G. Inductive Learning Algorithms for Complex Systems Modeling. Boca Raton: CRC Press Inc., 1994. 384 p.
3. Ивахненко А.Г. Метод группового учета аргументов - конкурент метода стохастической аппроксимации. Автоматика. 1968. №3. С. 58–72.
4. Intelligent Modelling and Analysis. URL: https://www.nottingham.ac.uk/computerscience/research/ima.aspx
5. Advances in Intelligent Modelling and Simulation: Artificial Intelligence-Based Models and Techniques in Scalable Computing / Eds J. Ko odziej, S.U. Khan, T. Burczy ski. – Berlin–Heidelberg: Springer-Verlag, 2012. – 381 p.
6. Opportunities in Intelligent Modeling Assistance. URL: https://modeling-languages
.com/opportunities-intelligent-modeling-assistance/.
7. Intelligent modeling of historical and as-built plant data for an efficient decision making. URL: https://www.avineon.com/en/intelligent-data-modeling
8. Кузнецов С. Д. Основы баз данных: учебное пособие. 2-е изд. М.: Интернет-Университет Информационных Технологий. Бином. Лаборатория знаний, 2007.
9. The ICT Development Index (IDI): conceptual framework and methodology. URL: https://www.itu.int/en/ITU-D/Statistics/Pages/publications/mis/methodology.aspx
10. Networked Readiness Index. URL: https://networkreadinessindex.org/
11. E-Government Development Index (EGDI). URL: https://publicadministration.un.org/egovkb/en-us/About/Overview/-E-Government-Development-Index
12. E-Participation Index. URL: https://publicadministration.un.org/egovkb/en-us/About/Overview/E-Participation-Index
13. E-Government readiness. URL: https://www.oecd-ilibrary.org/docserver/9789264061651-33-en.pdf?expires=1621616606&id=id&accname=guest&checksum=A9A1AD23179D762594F1B0D1DC5211AB
14. Global Cybersecurity Index. URL: https://www.itu.int/en/ITU-D/Cybersecurity/Pages/global-cybersecurity-index.aspx.
15. EBRD Knowledge Economy Index. URL: https://www.ebrd.com/news/publications/brochures/ebrd-knowledge-economy-index.html
16. ABOUT THE GLOBAL INNOVATION INDEX. URL: https://www.globalinnovationindex.org/about-gii#keyfindings
17. The Digital Economy and Society Index (DESI). URL: https://digital-strategy.ec.europa.eu/en/policies/desi

18. Савченко Е. А., Тутова Е. В. Методика анализа данных, моделирования и прогнозирования уровня развития человеческого потенциала стран мира и регионов Украины. Управляющие системы и машины. 2016. № 4. С. 86-92.
19. Савченко Е.А. Предварительная обработка данных в задаче индуктивного моделирования. Управляющие системы и машины. 2015. №2. С.82 – 87.
20. Савченко Е.А., Степашко В.С. Анализ селективных свойств критериев МГУА при их последовательном применении. Моделювання та керування станом еколого-економічних систем регіону. Збірник праць. Київ: МННЦ ІТС, 2008. С. 199–210.
21. Savchenko, E.A., Ivakhnenko, A.G. Investigation of efficiency of additional determination method of the model selection in the modeling problems by application of GMDH algorithm. Journal of Automation and Information Sciences, 2008, 40(3), pp. 47–58.
22. Коппа Ю.В., Степашко В.С. Порівняння ефективності застосування регресійного аналізу та МГУА для прогнозування екологічних процесів. Праці Міжнар. конф. з індуктивного моделювання в 4-х томах (МКІМ-2002). Т.3. Львів : Державний НДІ інф. інфраструктури, 2002. С. 123–127.
23. Tutova O.V., Savchenko Y.A. Ukraine in the Information and Communication Technology Development Ranking. Control systems and computers. 2019. № 3. P. 70-78.
24. Tutova, O., Savchenko, Ye., 2018. “Assessment of Influence of Digital Economy Development on Social and 6. Economic Environment”, Upr. sist. ma ., 6 (278), pp. 81-92. (In Ukrainian).

Надійшла 23.01.2021

Є.А. Савченко-Синякова, кандидат технічних наук, ст. науковий співробітник,
Міжнародний науково-навчальний центр інформаційних технологій та систем
НАН та МОН України, 03187, м. Київ, просп. Академіка Глушкова, 40, Київ, Україна,
savchenko_e@meta.ua
О.В. Тутова, кандидат економічних наук, ст. науковий співробітник,
Міжнародний науково-навчальний центр інформаційних технологій та систем
НАН та МОН України, 03187, м. Київ, просп. Академіка Глушкова, 40, Київ, Україна,
olenatutova@gmail.com

ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ, МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ПРОЦЕСІВ

Вступ. Макроекономічні дані часто містять невизначеності, можуть бути нечіткими й містити шум. Індуктивний підхід є одним із засобів моделювання та прогнозування за невеликими та зашумленими даними, що дає змогу з мінімальним втручанням людини у процес побудови моделі, отримати результат, який може допомогти в ухваленні людиною управлінських рішень.

За допомогою індуктивного підходу можуть бути розв'язані задачі аналізу даних, моделювання та прогнозування. Індуктивне моделювання побудовано на принципах індукції, тобто узагальнення на підставі деяких фактів. Один із основних методів індуктивного моделювання — метод групового урахування аргументів (МГУА). Найпопулярніший — комбінаторний алгоритм МГУА, який не потребує доведення збіжності, оскільки здійснює повний перебір усіх можливих варіантів моделей. Цей алгоритм розв'язує завдання структурно-параметричної ідентифікації, тобто дає змогу знайти не лише параметри моделі, а й її оптимальну структуру.

Мета — розробити на основі комбінаторного алгоритму МГУА та кореляційно-регресіоного аналізу методику розв'язання завдань аналізу даних, моделювання та прогнозування рівня людського розвитку країн світу та регіонів України. Ця методика допоможе вдосконалити процес прийняття рішень в сфері соціально-економічного розвитку України й на регіональному рівні, й у порівнянні з іншими країнами світу, що уможливить оцінку ситуації в сфері розвитку людського потенціалу і в Україні, й у світі загалом.

Методи. Кореляційно-регресійний аналіз, комбінаторний алгоритм МГУА.

Результати. Розроблено інструментальні засоби інтелектуального моделювання в MS Excel на основі методики аналізу, моделювання та прогнозування складних процесів за БД соціально-економічних процесів, яка містить інформацію про показники, що характеризують розвиток цифрової економіки України.

Висновок. Розроблені засоби інтелектуального моделювання дають змогу автоматизувати й удосконалити процес прийняття рішень в сфері соціально-економічного розвитку України й на регіональному рівні, й у порівнянні з іншими країнами світу.

Ключові слова: *інструментальні засоби, аналіз даних, моделювання, прогнозування, рівень людського розвитку, МГУА.*