UDC: 331.5:519.86
DOI: 10.33111/iep.eng.2023.39.08

Cluster Modelling of Labour Resources Employment in the Context of Globalisation

KATERYNA IVASHCHENKO[1],
ANDRIY MATVIYCHUK[2]

ANNOTATION. The article examines the essence of the phenomenon, the state and dynamics of unemployment in the structure of the modern economy, analyses the literature on forecasting the employment of labour resources. A new approach to solving the problem of analysing and forecasting the development of the labour market and indicators of labour force employment using the Kohonen self-organising maps is proposed. The basis of this approach is the limited data series for individual countries to obtain meaningful conclusions or forecasts. Therefore, to improve the accuracy of the modelling, it is advisable to expand the information base with relevant data for other countries. However, given the significant differences between different countries, there is a need to identify groups of countries that are similar in terms of the state and development of the labour market. This is where clustering methods come in handy.
The study selected more than 40 primary indicators that determine the level of unemployment, employment, labour market conditions, demographic and macroeconomic characteristics of 203 countries over the 12-year period from 2010 to 2021. As a result of the data filtering, 173 countries remained, on the basis of which further analysis and clustering are carried out. When filling in the gaps for these countries, the average values for the corresponding indicator for groups of countries with the same level of human development were taken.
The authors also argued for the expediency of using relative indicators in clustering to enable comparison of countries of different sizes. Accordingly, a number of relative indicators from the original list were selected for the final list of factors, and a number of new relative predictors were constructed on the basis of others. A total of 30 indicators were used to build the Kohonen self-organising map, which allowed segmenting countries by their level of socio-economic development and labour force potential. As a result of numerous experiments, it was found that the most effective distribution, in which the indicators of countries retain the greatest similarity in groups, is observed when dividing the world's countries into 12 clusters. In this case, Ukraine falls into a cluster with the following countries: Croatia, Czech Republic, Greece, Hungary, Poland, Slovenia, etc. Ukraine's position on the self-organising map indicates a high level of labour market development. Moreover, in 2018, Ukraine changed its position within the same cluster, moving closer to the group of more developed countries.

KEYWORDS: unemployment, employment, labour market, labour resources, clustering.

Introduction

One of the key obstacles to the country's economic development is the imbalance in the functioning of the labour market. High unemployment is a

[1] **Kateryna Ivashchenko** – student of the Department of Mathematical Modelling and Statistics at Kyiv National Economic University named after Vadym Hetman (Kyiv, Ukraine). Scientific interests: economic and mathematical modelling, socio-economic development, labour resources, employment. E-mail: kateryna.ivashchenko1@kneu.ua ORCID: https://orcid.org/0009-0000-8304-9026
[2] **Andriy Matviychuk** – Doctor of Economics, Professor of the Department of Mathematical Modelling and Statistics, Kyiv National Economic University named after Vadym Hetman (Kyiv, Ukraine). Scientific interests: modelling of complex systems, artificial intelligence, neural networks, fuzzy logic. E-mail: matviychuk@kneu.edu.ua ORCID: https://orcid.org/0000-0002-8911-5677

significant factor of social and economic instability. Growth in the number of unemployed leads to a number of negative consequences, creating a number of problems: underutilisation of the economic potential of society, reduction in tax revenues, additional costs for supporting the unemployed, and thus a deterioration in living standards and an increased risk of social tension. Forecasting and regulating the unemployment phenomenon at the macroeconomic level becomes an urgent task for the state. Accordingly, there is an urgent need for further development of theoretical provisions and practical solutions to create modern tools for forecasting the development of the labour market in the changing conditions of the modern global economy.

Unemployment, as a socio-economic phenomenon, slows down economic development and causes structural shifts in all economic sectors. In general, unemployment is a widespread problem in most countries, including Ukraine. Globalisation and the economic situation caused by the COVID-19 pandemic have shaken the stability of the labour market. During the pandemic, unemployment rates have risen to record levels across the world. As for the situation in the Ukrainian labour market, it is extremely complex and critical: before Ukraine could recover from the consequences of COVID-19, it faced new challenges – a full-scale war that made harsh "adjustments" to the socio-economic life of Ukrainian society and led to new challenges and tasks.

There are different interpretations of the essence of unemployment. According to the Law of Ukraine "On Employment of the Population": "Unemployment is a socio-economic phenomenon in which some people are unable to exercise their right to work and receive wages (remuneration) as a source of subsistence"[3]. Unemployment is caused by the imbalance of labour supply over demand in the labour market, resulting in the emergence of the category of "unemployed" who have the desire but not the ability to realise their labour potential. "Unemployed person is a person aged 15 to 70 who, due to lack of work, has no earnings or other income as a source of subsistence provided for by law, is ready and able to start work"[4].

Unemployment is to some extent an inevitable phenomenon, as no country in the world has managed to completely eliminate it, despite the enormous efforts made to overcome it. Looking at the dynamics of the number of registered unemployed in Ukraine over the period from 2019 to 2023, we note that there has been a gradual decrease. According to the State Employment Service of Ukraine[5], both the number of unemployed and the

[3] Verkhovna Rada, *On Employment of the Population (Law of Ukraine 5067-VI),* 2023, https://zakon.rada.gov.ua/laws/show/5067-17/page#Text [in Ukrainian].
[4] Ibid.
[5] State Employment Service of Ukraine, "Analytical and Statistical Information," 2023, https://www.dcz.gov.ua/analitics/view.

number of vacancies are on the decline. The ratio of these indicators is shown in Fig. 1.

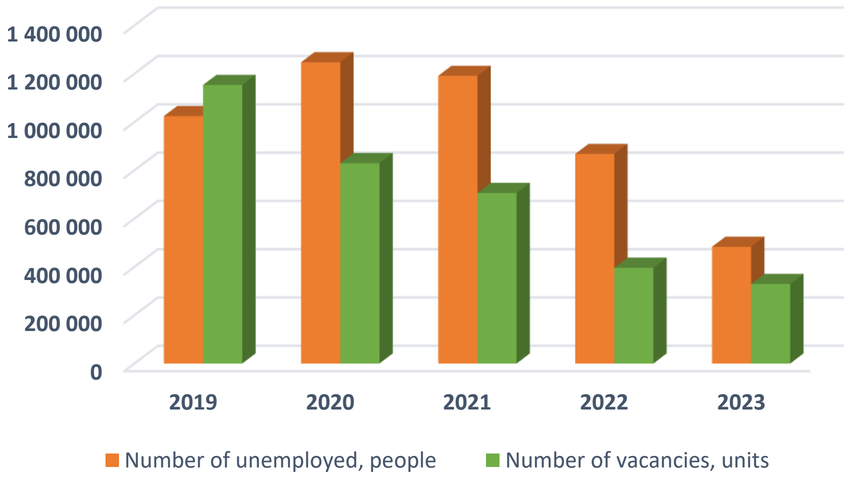

Figure 1. The ratio of the number of registered unemployed to the number of vacancies in Ukraine in 2019–2023

Source: developed by the authors on the basis of[6]

However, this trend has different causes and impacts on the labour market. For example, the decline in the number of officially registered unemployed (it is worth noting that the number has fallen to a historic low) is due to changes in the methods of accounting for unemployed people and a significant increase in the number of internally displaced persons abroad compared to the previous pre-war period. Another reason is that in order to obtain official unemployment status, persons liable for military service must provide a military registration certificate. Instead, the main reason for the decline in the number of vacancies is the closure of many businesses, organisations and institutions. In addition, "the decrease in the number of vacancies may be primarily due to the fact that the territory controlled by the Ukrainian authorities has decreased"[7].

---

[6] State Employment Service of Ukraine, "Analytical and Statistical Information," 2023, https://www.dcz.gov.ua/analitics/view; State Statistics Service of Ukraine, "Demographic and Social Statistics," 2023, http://www.ukrstat.gov.ua/.

[7] A. Diachkina, "Unemployment in Ukraine: Why the Number of Officially Unemployed People Has Changed and What Will Be the Situation on the Market After the War," *Slovo i Dilo (Word and Business)*, June 4, 2022, https://www.slovoidilo.ua/2022/06/04/stattja/suspilstvo/bezrobittya-ukrayini-chomu-kilkist-oficzijno-bezrobitnyx-zmenshylasya-ta-yaka-bude-sytuacziya-rynku-praczi-pislya-vijny [in Ukrainian].

Analysing the dynamics of the unemployment rate in Ukraine in 2000–2021, we can draw the following conclusions: unemployment remained relatively stable before the outbreak of large-scale hostilities with some minor fluctuations; the lowest unemployment rate of 6.9 per cent was observed in 2007-2008, and the highest in 2000 – 12.4 per cent. However, over the past two years, there has been an upward trend in unemployment caused by the global economic crisis, the pandemic, and now a full-scale war. "Due to the full-scale war, Ukraine [as of the end of 2023] is among the ten countries with the highest unemployment rates: [in 2022], 21.1 per cent of the country's population was unemployed – according to the National Bank, this is the highest since 2010"[8]. By the end of 2023, the unemployment rate is expected to fall to 19 per cent, which is still quite high[9].

Thus, high unemployment is a serious and urgent problem in a market economy that needs to be addressed immediately. A detailed analysis of this issue will help to create effective proposals that can be used to develop and implement effective socio-economic strategies aimed at reducing unemployment to a socially acceptable level and increasing employment among the economically active population.

The unemployment rate has been a matter of serious concern and has attracted the attention of scholars since the 1970s, when stagflation, high unemployment and constant economic shocks first appeared[10]. Issues related to the functioning and development of the labour market in Ukraine, in particular the problems of employment and unemployment, have repeatedly been the subject of in-depth research and analysis by leading national scholars. The results of their research have contributed to a deeper understanding of the processes taking place in the Ukrainian labour market and formed the basis for developing effective mechanisms for regulating employment and overcoming unemployment in our country. Most domestic studies use statistical and econometric methods to measure unemployment.

For example, the paper by A. Kolot et al.[11] presents a scientific and applied substantiation of current trends and promising trajectories of employment development in the context of the digital economy. The authors

---

[8] Analytical portal 'Slovo i Dilo', "Ukraine is Among the Ten Countries with the Highest Unemployment Rates in the World," October 12, 2023, https://www.slovoidilo.ua/2023/10/12/infografika/svit/ukrayina-vxodyt-desyatky-krayin-svitu-najvyshhymy-pokaznykamy-bezrobittya [in Ukrainian].

[9] NBU, "Inflation will continue to decelerate and the economy will recover – NBU Inflation Report," 2023, https://bank.gov.ua/ua/news/all/inflyatsiya-i-dali-spovilnyuvatimetsya-a-ekonomika-vidnovlyuvatimetsya--inflyatsiyniy-zvit-nbu [in Ukrainian].

[10] K. Brunner, A. Cukierman, and A. Meltzer, "Stagflation, Persistent Unemployment and the Permanence of Economic Shocks," *Journal of Monetary Economics* 6, no. 4 (1980): 467–492, https://doi.org/10.1016/0304-3932(80)90002-1.

[11] A. Kolot, O. Herasymenko, A. Shevchenko, and I. Ryabokon, "Employment in the Coordinates of Digital Economy: Current Trends and Foresight Trajectories," *Neuro-Fuzzy Modelling Techniques in Economics* 11 (2022): 78–123. http://doi.org/10.33111/nfmte.2022.078.

have proved the hypothesis that the scale and structure of employment will change intensively under the influence of digitalisation. The authors also present the results of a study of the impact of employment in high-tech industries on the dynamics of GDP and gross value added in the Ukrainian economy by calculating the elasticity in multivariate regression models. The article presents the results of forecasting the relevant indicators with the indication of upper and lower confidence limits in the trend of one-factor regression, which allowed to model optimistic, realistic and pessimistic scenarios of the development of the studied macroeconomic indicators.

The paper by M. Nikulina et al.[12] examines unemployment trends in Ukraine in the long and medium term, taking into account the impact of the COVID-19 pandemic and the war, factors of development of the domestic labour market in 2002–2019, as well as the prospects for post-war reconstruction and digitalisation of the economy. Using the methods of correlation and regression analysis, the determinants of the long-term impact on the unemployment rate in the Ukrainian economy in the pre-pandemic period are identified. The main trends that will have an impact on the domestic labour market in the future are analysed, including technological transformations, changes in economic and social models, the spread of globalisation, environmental changes, and relevant areas of digital transformation.

A significant contribution to the study of the labour market was also made by M. Oliskevych, I. Lukianenko, O. Bazhenova, V. Tokarchuk. Thus, in the paper[13], a comparative analysis of the cyclical behaviour of the labour markets of Ukraine and Poland was carried out. Econometric analysis was carried out, and an autoregressive Markov switching model was developed, which reveals different modes of unemployment rate behaviour over time associated with recession and growth. The authors argue that the opposite features of the Ukrainian and Polish labour markets lead to significant labour migration from Ukraine to Poland. The paper[14] examines the dynamics of unemployment rates in a number of Eastern European countries, including Bulgaria, Estonia, Latvia, Lithuania, Poland, Romania, Slovakia, Hungary, and Ukraine in 2000-2017. A similar methodological approach was applied: econometric regression models with nonlinearities and a Markov switching model were built. The article estimates 2 modes of unemployment behaviour

[12] M. Nikulina, I. Sotnyk, O. Derykolenko, and I. Starodub, "Unemployment in Ukraine's Economy: COVID-19, War and Digitalisation," *Mechanism of an Economic Regulation* 1-2, no. 95-96 (2022): 25–32, https://doi.org/10.32782/mer.2022.95-96.04.

[13] M. Oliskevych, and V. Tokarchuk, "Dynamic Modelling of Nonlinearities in the Behaviour of Labour Market Indicators in Ukraine and Poland," *Economic Annals-XXI* 169, no. 1-2 (2018): 35–39, https://doi.org/ 10.21003/ea.V169-07.

[14] I. Lukianenko, M. Oliskevych, and O. Bazhenova, "Regime Switching Modelling of Unemployment Rate in Eastern Europe," *Ekonomický chasopis* 68, no. 4 (2020): 380-408, https://www.sav.sk/journals/uploads/ 0426175204 %2020 %20Oliskevych %20a %20kol. %20 %20+ %20SR.pdf.

associated with its high and low levels. The probabilities of transitions between the regimes and the average expected duration of stay in each of them are predicted.

In[15], S. Kozlovskyi et al. investigated the relationship between labour migration and economic growth in a number of countries in the context of the COVID-19 pandemic and the Russian-Ukrainian war. The aim of this paper[16] is to develop a system dynamics model to assess the impact of labour migrants on the economic growth of the host country during the COVID-19 pandemic.

At the same time, in recent years, researchers have increasingly focused on the use of machine and deep learning models in the study of unemployment.

For example, in the above-mentioned paper by A. Kolot et al.[17], in addition to econometric modelling methods, neural network tools, namely Kohonen self-organising maps, were used to cluster the 8 sectors of the economy under study. This made it possible to analyse the state of development of each industry in terms of employment structure and economic development based on a set of characteristic indicators.

C. Stasinakis et al.[18] investigated the effectiveness of radial basis function neural networks (RBFNN) in forecasting unemployment in the United States. They also tested the effectiveness of the Kalman filter and support vector regression as combined forecasting methods. The results show that the RBFNN statistically outperforms the individual performance of all the models studied.

A. Davidescu et al.[19] used univariate and multivariate methods to forecast the unemployment rate in Romania. Several models were considered in the study, including the Seasonal Autoregressive Integrated Moving Average (SARIMA), SETAR, Holt-Winters, ETS, and the Neural Network Autoregressive (NNAR) models. In terms of root mean square error (RMSE) and mean absolute error (MAE), the NNAR model showed the highest forecast accuracy.

[15] S. Kozlovskyi, D. Bilenko, M. Kuzheliev, R. Lavrov, V. Kozlovskyi, H. Mazur, and A. Taranych, "The System Dynamic Model of the Labour Migrant Policy in Economic Growth Affected by COVID-19," *Global Journal of Environmental Science and Management* 6, no. SI (2020): 95-106, https://doi.org/10.22034/GJESM.2019.06.SI.09.

[16] Ibid.

[17] A. Kolot, O. Herasymenko, A. Shevchenko, and I. Ryabokon, "Employment in the Coordinates of Digital Economy: Current Trends and Foresight Trajectories," *Neuro-Fuzzy Modelling Techniques in Economics* 11 (2022): 78–123. http://doi.org/10.33111/nfmte.2022.078.

[18] Stasinakis, G. Sermpinis, K. Theofilatos, and A. Karathanasopoulos, "Forecasting US Unemployment with Radial Basis Neural Networks, Kalman Filters and Support Vector Regressions," *Computational Economics* 47, no. 4 (2016): 569–587, https://doi.org/10.1007/s10614-014-9479-y.

[19] A. Davidescu, S.-A. Apostu, and A. Paul, "Comparative Analysis of Different Univariate Forecasting Methods in Modelling and Predicting the Romanian Unemployment Rate for the Period 2021-2022," *Entropy* 23, no. 3 (2021): Article 325, https://doi.org/10.3390/e23030325.

C. Katris[20] conducted a comparative analysis of different models for forecasting unemployment rates in 22 countries of a number of geographical regions such as the Mediterranean, Baltic, Balkans and Benelux. Autoregressive integrated moving average model (FARIMA), FARIMA with generalised autoregressive conditional heteroscedasticity (FARIMA/GARCH), artificial neural network (ANN) models based on support vector machine (SVM) and multivariate adaptive regression splines (MARS) were built. The results showed that there is no single universal model that would be suitable for all cases, and the choice of the most optimal method depends on both the forecasting interval and the region. In particular, FARIMA models proved to be the best for short-term forecasts of one step ahead, while neural network models showed optimal results for long-term forecasts (12 steps ahead). For intermediate forecasting intervals (3 steps ahead), the Holt-Winters model proved to be the most effective.

T. Chakraborty et al.[21] built models to forecast unemployment in several countries, including Canada, Germany, Japan, the Netherlands, New Zealand, Sweden, Switzerland and the United States. Several models have been used for forecasting, including ARIMA, ANN, autoregressive neural network (ARNN), SVM, ARIMA with support vector machine (ARIMA-SVM), ARIMA with artificial neural network (ARIMA-ANN) and ARIMA with autoregressive neural network (ARIMA-ARNN) for monthly and quarterly seasonally adjusted unemployment data for the mentioned countries. First, the ARIMA model is fitted to the data, and then the ARNN model is trained on the ARIMA residuals. Then, the ARIMA forecast results and the ARNN residuals are combined to obtain the final forecast values. The results are compared with other models, and performance metrics such as RMSE, MAE, and MAPE are used to evaluate the accuracy of the forecasts. The results show that the proposed hybrid ARIMA-ARNN model combines the ability of ARIMA models to detect seasonality with the ability of ARNN models to capture nonlinearities in the ARIMA residuals, which leads to better forecast accuracy compared to other standalone and hybrid models.

M. Ahmad et al.[22] conducted a study on forecasting unemployment in European countries: Belgium, France, Germany, Italy, Spain, Turkey and France. They considered classical models such as ARIMA and machine

---

[20] C. Katris, "Prediction of Unemployment Rates with Time Series and Machine Learning Techniques," *Computational Economics* 55 (2020): 673–706, https://doi.org/10.1007/s10614-019-09908-9.

[21] T. Chakraborty, A. K. Chakraborty, M. Biswas, S. Banerjee, and S. Bhattacharya, "Unemployment Rate Forecasting: A Hybrid Approach," *Computational Economics* 57 (2021): 183-201, https://doi.org/10.1007/s10614-020-10040-2.

[22] M. Ahmad, Y. Khan, C. Jiang, S. Kazmi, and S. Abbas, "The Impact of COVID-19 on Unemployment Rate: An Intelligent Based Unemployment Rate Prediction in Selected Countries of Europe," *International Journal of Finance & Economics* 28, no. 1 (2023): 528-543, https://doi.org/10.1002/ijfe.2434.

learning models such as ANN and SVM. In addition, hybrid approaches such as ARIMA-ARNN, ARIMA-ANN and ARIMA-SVM were proposed. The data was divided into training and test sets, and several hybrid models and approaches were applied to forecast the unemployment rate in these 6 countries. The Akaike Information Criterion (AIC) was used to select the best ARIMA models for each country. Hybrid models combined linear ARIMA forecasts with non-linear predictions from neural network or SVM models. The ARIMA-ARNN hybrid model performed well in predicting the unemployment rate in Belgium, Germany, Turkey and France, while the ARIMA-SVM hybrid model performed better for Spain and Italy. These results indicate that the choice of the most efficient model depends on the specific conditions of a particular country and the characteristics of the unemployment data.

A review of the above-mentioned developments of scientists further confirms the relevance and importance of the problem under study. However, the analysis of existing methods of forecasting employment indicators has revealed that there is a need to create a set of economic and mathematical models that would take into account the limited statistical data, the specifics of the development of individual countries, etc. A variety of mathematical methods and models can be used to build forecast models, ranging from traditional econometric methods to machine learning and artificial intelligence technologies. However, the choice of the most appropriate mathematical tools for the task at hand will depend on the available statistical data for analysis.

Thus, the above analysis of the essence of the phenomenon of unemployment, the state and dynamics of its development in Ukraine, the latest scientific works in this area, allowed to formulate the purpose of this study, which is to increase the efficiency of analysis and management of labour resources using modern mathematical tools. Achievement of this goal led to the solution of the following tasks: formation of an information base for the study; construction of models for assessing labour force employment indicators; conducting model experiments to find the most rational cluster structure of countries in terms of employment indicators; analysis of migration of the country under study (Ukraine) on the basis of the cluster model during the study period to determine trends in the development of the labour market and formulation of conclusions and recommendations to improve the position towards the most developed countries in terms of indicators of economic development and employment of population.

The object of the study is the employment of labour resources in the world, including Ukraine. The subject of the study is modelling of labour force employment indicators in the context of globalisation. To achieve this goal, the following tools were used in the study: statistical methods, neural networks, Kohonen self-organising maps.

Methodological foundations of labour force employment modelling

When modelling labour force participation rates for a particular country (in our study, Ukraine), a number of difficulties arise due to the limited data series available to identify functional dependencies of employment rates on influencing factors. However, in order to use machine learning technologies, especially deep learning, for forecasting, a sufficiently large data set covering observations over a long period of time is required.

If we study a dataset that includes annual indicators of employment in Ukraine, the total number of observations will be too small to identify significant patterns. This is even more so if we take into account the constant changes in the political and socio-economic life of our country, which introduces significant heterogeneity in the processes under study in different time periods. And this makes it irrelevant to use, say, ten-year old indicators for employment modelling. In this regard, a decision was made to increase the number of observations to build an adequate model for estimating employment indicators.

To improve the accuracy of the modelling, it is advisable to take into account not only the data of a single country, but also information on a number of other countries, thus expanding the information base for the study. However, this raises the problem of significant differences in the state and patterns of labour market development in different countries. For example, it is impossible to build a single adequate forecast model for employment indicators based on statistics from countries with different socio-economic potentials, such as Sweden, Zimbabwe and Ukraine. It is quite clear that these countries have significant differences in labour market development, so each of them will have different forecast models based on specific factors and trends.

The development of such employment forecasting models should take into account the heterogeneity of labour market development in different countries. One-size-fits-all approaches may not be effective, so it is advisable to group countries according to certain criteria and create separate models for each group of countries, thus enriching the training set to ensure greater accuracy of the modelling.

Thus, in order to build models for estimating labour force participation rates, it is necessary to create homogeneous data sets by country group. For example, in order to model the level of employment in Ukraine, it is necessary to create a database that includes information on other countries whose labour market conditions and development are similar to the situation in Ukraine. To this end, it is advisable to cluster countries based on the indicators studied over a certain period of time. Those countries that fall into the same cluster as Ukraine as a result of segmentation will be included in the training sample for further study of the state and modelling of labour market dynamics in Ukraine.

It should be noted that in the field of Data Mining today there are a large number of different clustering algorithms and their modifications. Despite the wide choice of such algorithms, in economic and mathematical modelling, the most commonly used algorithms are those that are easy to use and at the same time provide adequate clustering results. Such cluster analysis algorithms include:

1. Centroid algorithms (k-means, fuzzy c-means, mean shift, etc.) or medoid algorithms (k-means, subtractive clustering, etc.)[23, 24]. The k-means method and subtractive clustering should be used when there is noise or outliers in the data. K-means and fuzzy c-means are effective for fast processing of large data sets. Subtractive clustering and mean shifting should be used when the number of clusters is unknown.

2. Models of mixture of distributions (EM-algorithm, etc.)[25]. These methods are used in cases where clusters are of different sizes and the set of objects can be described by a mixture of distributions (e.g., Gaussian distributions).

3. Dense algorithms (OPTICS, DBSCAN) that are effective in the presence of noise, random outliers, clusters of different shapes and sizes, and an unknown number of clusters[26].

4. Divisive (DIANA) or agglomerative – centroid linkage, single linkage, full linkage, Ward, group average[27]. They are used for small datasets, building dendrograms based on intercluster distances.

5. Connectionist algorithms that use artificial neural networks (SOM – Kohonen self-organising maps[28, 29], ART family, etc.) to form clusters.

6. Metaheuristic methods[30, 31], which use, for example, a genetic algorithm to form clusters.

---

[23] M. Brusco, E. Shireman, and D. Steinley, "A Comparison of Latent Class, K-means, and K-median Methods for Clustering Dichotomous Data," *Psychological Methods* 22, no. 3 (2017): 563–580, https://doi.org/10.1037/met0000095.

[24] C. Aggarwal and C. Reddy, *Data Clustering: Algorithms and Applications* (1st ed.; Chapman and Hall/CRC, 2014), https://doi.org/10.1201/9781315373515.

[25] Z. Fu, and L. Wang, "Colour Image Segmentation Using Gaussian Mixture Model and EM Algorithm," *Multimedia and Signal Processing, Communications in Computer and Information Science* 346 (2012): 61–66, https://doi.org/10.1007/978-3-642-35286-7_9.

[26] C. Aggarwal and C. Reddy, *Data Clustering: Algorithms and Applications* (1st ed.; Chapman and Hall/CRC, 2014), https://doi.org/10.1201/9781315373515.

[27] Ibid.

[28] T. Kohonen, *Self-Organising and Associative Memory* (3rd ed.; Springer, 2012), https://doi.org/10.1007/978-3-642-88163-3.

[29] T. Kohonen, "Essentials of the Self-Organising Map," *Neural Networks* 37 (2013): 52–65, https://doi.org/10.1016/j.neunet.2012.09.018.

[30] J. Radosavljević, *Metaheuristic Optimization in Power Engineering* (IET, 2018), https://doi.org/10.1049/pbpo131e.

[31] E. Alba, A. Nakib, and P. Siarry, *Metaheuristics for Dynamic Optimization* (Springer, 2013), https://doi.org/10.1007/978-3-642-30665-5.

Given the possibilities of different clustering methods and the specifics of the task, the Kohonen self-organising mapping tool was chosen to group countries by labour force employment. This method not only allows for the formation of homogeneous groups of objects under study, but also provides a convenient tool for visual analysis of clustering results. Unlike other algorithms, the location of an object on the Kohonen map immediately gives the analyst a visual representation of how well the property under study is developed compared to other objects. After all, the best and worst performers in terms of the analysed indicator are located in opposite corners of the self-organising map. This makes it easier to interpret and visualise the results of cluster analysis.

In the Kohonen self-organising map, each neuron is connected to all components of the *n*-dimensional input data vector $x = (x_1, x_2, x_n)$ (Fig. 2).

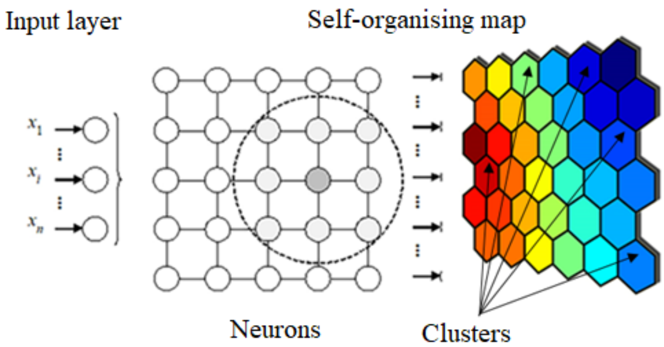

Figure 2. Visual representation of clusters on the Kohonen map

Source: [32].

The input vector contains a description of each individual object (in our case, these are the characteristics that determine the employment of labour resources in a particular country for a certain year) that will be used for clustering. Each *j*-th neuron of the map is characterised by its own vector of weights:

$$w_j = (w_{1j}, w_{2j}, w_{nj}),$$

where *n* is the number of parameters of each neuron, which corresponds to the number of features of the input vectors.

Kohonen self-organising maps use an unsupervised learning procedure based on competitive processes between neurons. When a data vector is

[32] A. Matviychuk, O. Lukianenko, and I. Miroshnychenko, "Neuro-Fuzzy Model of Country's Investment Potential Assessment," *Fuzzy Economic Review* 24, no. 2 (2019): 65-88, https://doi.org/10.25102/fer.2019.02.04.

input, a "competition" between neurons takes place. The winner is the neuron whose weighting vector is most similar to the input vector by some metric, usually the Euclidean distance. This neuron is called the winning neuron and must satisfy the condition:

$$c = \underset{j}{\operatorname{argmin}}\{\|x - w_j\|\}.$$

After determining the winning neuron for a given input vector, the weights of both the winning neuron and all other neighbouring neurons are adjusted according to the Kohonen learning function:

$$w_j(t+1) = w_j(t) + v(t) \cdot h_{cj}(t) \cdot [x(t) - w_j(t)],$$

where $v(t)$ – is the learning rate coefficient, which decreases with each training epoch t, $h_{cj}(t)$ – is the topological neighbourhood function to the winning neuron, which is determined by the Gaussian function:

$$h_{cj}(t) = exp\left[-\frac{\|r_c - r_j\|}{2 \cdot \sigma^2(t)}\right],$$

$\sigma(t)$ is the effective width of the topological neighbourhood (a specially chosen time function that monotonically decreases during training), r_c and r_j – coordinates of the geometric location of the winning neuron c and other neurons j on the map.

The result of the Kohonen map creation is a visual representation of a two-dimensional lattice of neurons that reflect the organisational structure of countries, forming clusters in which countries are similar to each other in terms of a group of employment indicators (similar to Fig. 2).

Thus, at the first stage of modelling the country's labour force employment indicators, it is necessary to identify relevant sources of information, collect data on employment, demographic indicators, macroeconomic indicators and other factors that may affect the level of labour force employment. After collecting the data, it should be structured, filtered, filled in the gaps, constructed new significant predictors, and presented in an understandable form while maintaining information content.

At the second stage of the modelling, countries are clustered according to the selected set of indicators using an artificial neural network built on the basis of the Kohonen self-organising map principle. The result of building a self-organising map is a visual representation of the cluster structure in the form of a two-dimensional hexagonal lattice of neurons (Fig. 2) and a set of groups of countries with similar indicators that determine the development of the labour market and employment. At the end of the second stage, the location on the map of the main country to assess the level of employment (Ukraine) and the cluster to which it belongs are determined. Also, the migration of this country is analysed using the Kohonen map during the period under study in order to identify trends in the labour market development and to formulate conclusions and recommendations for correcting certain indicators to improve the position towards the most developed countries in terms of economic development and employment.

Creation of an information base for modelling labour force employment

The circumstances caused by russia's full-scale invasion of Ukraine have led to dramatic transformations in all aspects of Ukrainian society, with the greatest impact on the labour market. Moreover, the assessment of these radical changes in the labour market is complicated by the lack of statistical data, as due to russia's armed aggression against Ukraine, since February 2022, state statistics authorities have not been conducting sample surveys of the population (households), and therefore are unable to obtain information on their results[33]. Accordingly, the study of the state and development of the labour market based on data for Ukraine alone becomes quite limited, so it will be advisable to use data from different countries of the world to enrich the training sample, as justified in the section "Methodologies for modelling labour force employment".

In order to adequately identify patterns, it is necessary to form a certain set of attributes that will describe the most important aspects that affect the labour force participation rate. After clustering the countries of the world, the result will be clusters, each of which will unite countries with more or less similar values of the studied indicators, and, as a result, a similar level of labour force employment.

In the course of data collection, an analysis of publicly available information sources containing indicators that could potentially affect labour force employment was conducted. Among them is the World Development Indicators database, which is available on the World Bank's website at[34]. This dataset contains about 1500 indicators for 217 countries, including wages, births, deaths; labour force, unemployment and the share of employers in total and by gender; total employment, by industry, vulnerable employment, by age; literacy, by age and by education level; a wide range of macroeconomic and demographic indicators, etc. Indicators are calculated annually; the range of availability is 1960–2023. However, some indicators in this database may not be calculated for certain countries (for example,

---

[33] Verkhovna Rada, *On the Protection of the Interests of the Subjects of the Presentation of Information and Other Documents During the Period of War (Law of Ukraine 2115-IX)*, 2023, https://zakon.rada.gov.ua/laws/show/2115-20#Text. [in Ukrainian].

[34] The World Bank, *World Development Indicators* [Data set], 2023, https://databank.worldbank.org/source/world-development-indicators.

with a low level of development), and statistical information by country is limited for 2022–2023.

In selecting indicators for assessing the level of labour force participation, the leading source of labour statistics in the world, ILOSTAT, an information database of the International Labour Organisation that collects labour statistics and metadata for more than 200 countries, was also analysed[35]. The databases cover statistics on household income and expenditure, economically active population, employment, unemployment, detailed occupational groups and status from censuses, comparative estimates of employment and unemployment, public sector employment, hours worked, wages, productivity and costs, labour income and inequality, social protection, informal economy, as well as information on youth, women, migrant workers and volunteers. However, a large number of indicators duplicate those in the World Development Indicators database[36] and, in addition, most indicators lack values for more than 100 countries. In addition to the general list of primary indicators for assessing labour force employment, one could add, for example, "Statutory nominal gross monthly minimum wage, USD" and "Average number of weekly hours worked per employed person". However, these two indicators are calculated for only 138 and 136 countries, respectively, and if we take statistics for the last 10 years, the list of countries will be even shorter. Therefore, it is better not to include ILOSTAT indicators in the overall dataset to avoid the need to fill in a large number of missing values, as this may distort the calculations based on statistical information.

Another information resource is the UNDP database. The platform [30] contains a wide range of data, including socio-economic indicators, environmental data, demographic statistics, etc. There are also datasets related to development projects. In addition, tools for data analysis and visualisation are available. Currently, the platform's tools automatically extract more than 400 indicators from more than 60 reliable sources (including the International Labour Organisation, World Resources Institute, World Bank, World Economic Forum, etc.).

As a result of the analysis of information sources and databases of various international organisations, it was decided to cluster countries by employment indicators based on the World Bank statistics, in particular the World Development Indicators database[37]. The primary list of indicators includes those that, in the authors' opinion, have the greatest impact on the labour force employment rate, namely: "Compensation of employees (per cent of expense)", "Birth rate, crude (per 1000 people)", "Death rate, crude

[35] International Labour Organisation, *Labour Force Participation Rate* [Data set], 2023, https://ilostat.ilo.org/data/.

[36] The World Bank, *World Development Indicators* [Data set], 2023, https://databank.worldbank.org/source/world-development-indicators.

[37] Ibid.

(per 1000 people)", "Children out of school (per cent of primary school age)", "Employers, total (per cent of total employment)", "Employment in agriculture (per cent of total employment)", "Employment in industry (per cent of total employment)", "Employment in services (per cent of total employment)", "Employment to population ratio, 15+ (per cent)", "Inflation, GDP deflator (annual per cent)", "Labour force participation rate, total (per cent of total population ages 15–64)", "Labour force with advanced education (per cent of total working-age population with advanced education)", "Labour force with basic education (per cent of total working-age population with basic education)", "Population growth (annual per cent)", "Unemployment, female (per cent of female labour force)", "Unemployment, male (per cent of male labour force)", "Unemployment, total (per cent of total labour force)", "Vulnerable employment, total (per cent of total employment)", "Wage and salaried workers, total (per cent of total employment)", "Ratio of female to male labour force participation rate (per cent)", "Net migration", "Population, total", "Labour force, total", "Labour force, female (per cent of total labour force)", "GDP, total (current USD)", "Exports of goods and services (current USD)", "Imports of goods and services (current USD)", "Foreign direct investment, net inflows (current USD)", "Foreign direct investment, net outflows (current USD)", "Individuals using the Internet (per cent of population)", "Government expenditure on education, total (current USD)", "Population ages 15-64", "Population ages 65 and above", "Population ages 0–14".

This list of indicators was used to create a database for 203 countries for the period 2010–2021. It should be noted that if the study reveals, for example, that certain indicators do not significantly affect the clustering process or introduce imbalances, they will be excluded from the set of explanatory variables of the model.

Processing the data set under study

At the first stage of processing the data set under study, we applied filtering, sorted and distributed the data by country. A visual inspection of the data reveals that there are gaps in a number of indicators for certain countries, such as Cayman Islands, Channel Islands, Djibouti, Eritrea, French Polynesia, Gibraltar, Greenland, Grenada, Isle of Man, Kiribati, Kosovo, Libya, Liechtenstein, Marshall Islands, Monaco, Nauru, New Caledonia, Northern Mariana Islands, Palau, San Marino, Seychelles, etc. For these countries, a significant part of statistical information is not available at all. Accordingly, these countries were not included in the dataset. It should be noted that this does not reduce the adequacy of the modelling, as the labour markets of these small countries are not similar to Ukraine, so they would almost certainly not be included in the same cluster as Ukraine if comprehensive data were available. After their exclusion, the total number of countries studied is 173, which is the basis for further analysis and clustering.

The next step in processing statistical information is to fill in the gaps in the data for the countries selected for the training sample, because, as mentioned earlier, there should be no gaps in the input data when constructing Kohonen self-organising maps.

A common solution in this case is to fill in the missing elements in the dataset by finding the average value of the relevant indicator[38]. However, in order to avoid the possibility of distortion, which is likely to lead to inaccuracies in the modelling results, it would be more appropriate to take the average values for a particular indicator not for all countries of the world, but for some groups of countries that will generally have a similar level of Human Development Index (HDI)[39], and, as a result, similar socio-economic indicators. Therefore, based on the principle of grouping countries into 4 classes depending on the level of human development, all 173 countries in the data set under study were ranked. For each of the groups of countries, the average values of all indicators were calculated. If a country had a gap in a particular indicator, this gap was replaced by the corresponding average value.

Another important step in preparing the dataset for further clustering is to convert all indicators to the required form. The fact is that a number of indicators in the dataset under study are presented in percentage form (employment by industry, unemployment by gender/total, share of the labour force with higher education, employee remuneration (per cent of total expenditure), etc.) But there are also absolute indicators, such as population, labour force, net migration, GDP, exports and imports of goods and services, and population by age. When data is presented in absolute terms, some indicators may have a greater impact on clustering simply because of their large numerical values. Using such indicators can distort the results and lead to incorrect clustering. Relative values (such as the ratio of education expenditure to GDP) can be more informative because they allow for comparisons across countries of different scales.

Therefore, the preparation of new predictors in a relative format based on absolute indicators will facilitate the comparison of data between different countries and reduce the number of indicators to be analysed (as the primary data set includes more than 30 indicators), while maintaining the essence of the information. Table 1 provides a list of new additional predictors in relative form that will be used for clustering experiments.

---

[38] D. Lukianenko, A. Matviychuk, L. Lukianenko, and I. Dvornyk, "University Competitiveness in the Knowledge Economy: A Kohonen Map Approach," *CEUR Workshop Proceedings* 3465 (2023): 236–250, https://ceur-ws.org/Vol-3465/paper23.pdf.

[39] United Nations Development Programme, *Human Development Reports* [Data set], 2023, http://hdr.undp.org/en/content/download-data.

Table 1

A SET OF ADDITIONAL PREDICTORS IN RELATIVE FORM

Name of the indicator	Formula	Explanation
Exports of goods and services (% of GDP)	(Exports of goods and services / GDP) * 100	Indicates the relative volume of exports of goods and services compared to GDP.
Imports of goods and services (% of GDP)	(Imports of goods and services / GDP) * 100	Indicates the relative volume of imports of goods and services compared to GDP.
Foreign direct investment, net inflows (% of GDP)	(Foreign direct investment, net inflows / GDP) * 100	Indicates the relative amount of net foreign direct investment that entered the country compared to GDP.
Foreign direct investment, net outflows (% of GDP)	(Foreign direct investment, net outflows / GDP) * 100	Indicates the relative amount of net foreign direct investment that has left the country compared to GDP.
Individuals using the Internet (% of the population)	(Individuals using the Internet / Population, total) * 100	Indicates the percentage of the population that uses the Internet.
Government expenditure on education, total (% of GDP)	(Government expenditure on education, total / GDP) * 100	Indicates the relative amount of Government expenditure on education compared to GDP.
Labour force, female (% of total labour force)	(Labour force, female / Labour force, total) * 100	Indicates the percentage of women in the total workforce.
Net migration (% of population)	(Net migration / Population, total) * 100	Indicates the relative volume of net migration compared to the total population.
Population ages 15–64 (% of total population)	(Population ages 15–64 / Population, total) * 100	Indicates the proportion of the population aged 15 to 64 compared to the total population.
Population ages 65 and above (% of total population)	(Population ages 65 and above / Population, total) * 100	Indicates the proportion of the population aged 65 and over compared to the total population.
Population ages 0–14 (% of total population)	(Population ages 0–14 / Population, total) * 100	Indicates the proportion of the population aged 0–14 years compared to the total population.

As can be seen from Table 1, a total of 11 new predictors were added in relative terms. From the original list of indicators, 19 were selected, which are also presented in relative terms and can be used to compare countries of different scales. Together, these two sets of indicators form a single body of information that will be used to cluster countries into groups characterised by similar labour market conditions and dynamics.

Clustering the countries by employment indicators

The analytical platform Deductor Studio Academic was used for clustering. This platform offers the most extensive functionality compared to similar products when it comes to data integration, visualisation and cleaning, as well as the use of data mining methods.

As a result of numerous experiments, a hexagonal grid of 22×16 neurons was determined as the optimal structure of the self-organising map for clustering countries of the world by 30 selected employment indicators, which included some of the primary indicators listed in the section “ Creation of an information base for modelling labour force employment” and all the relative ones from Table 1. The clustering process involved 5000 training epochs. The Gaussian function was chosen to determine the neighbourhood of neurons. As a result, we obtained the distribution of the world’s countries among the three clusters, as can be seen in Fig. 3.

As a result of this clustering, Ukraine was placed in the cluster of countries with the best labour force participation rates. Although the distribution of values of indicators in this group is more even than in the whole world, it is still unlikely that this cluster is completely homogeneous. This necessitated a number of experiments, during which Kohonen maps were constructed with different input variables and the least significant indicators were removed one by one. However, despite the exclusion of certain factors and changes in the settings, the quality of the distribution did not improve significantly. In most cases, the clustering results placed Bosnia and Herzegovina, Bulgaria, Croatia, Greece, Hungary, Montenegro, Poland, Romania and Serbia in the same group as Ukraine, but often the neighbouring countries in the cluster were also Finland, Japan, South Korea, and the United States. Such results of clustering countries cannot be considered adequate, so it was decided to normalise the indicators of the data set under study and compare the results of clustering with each other.

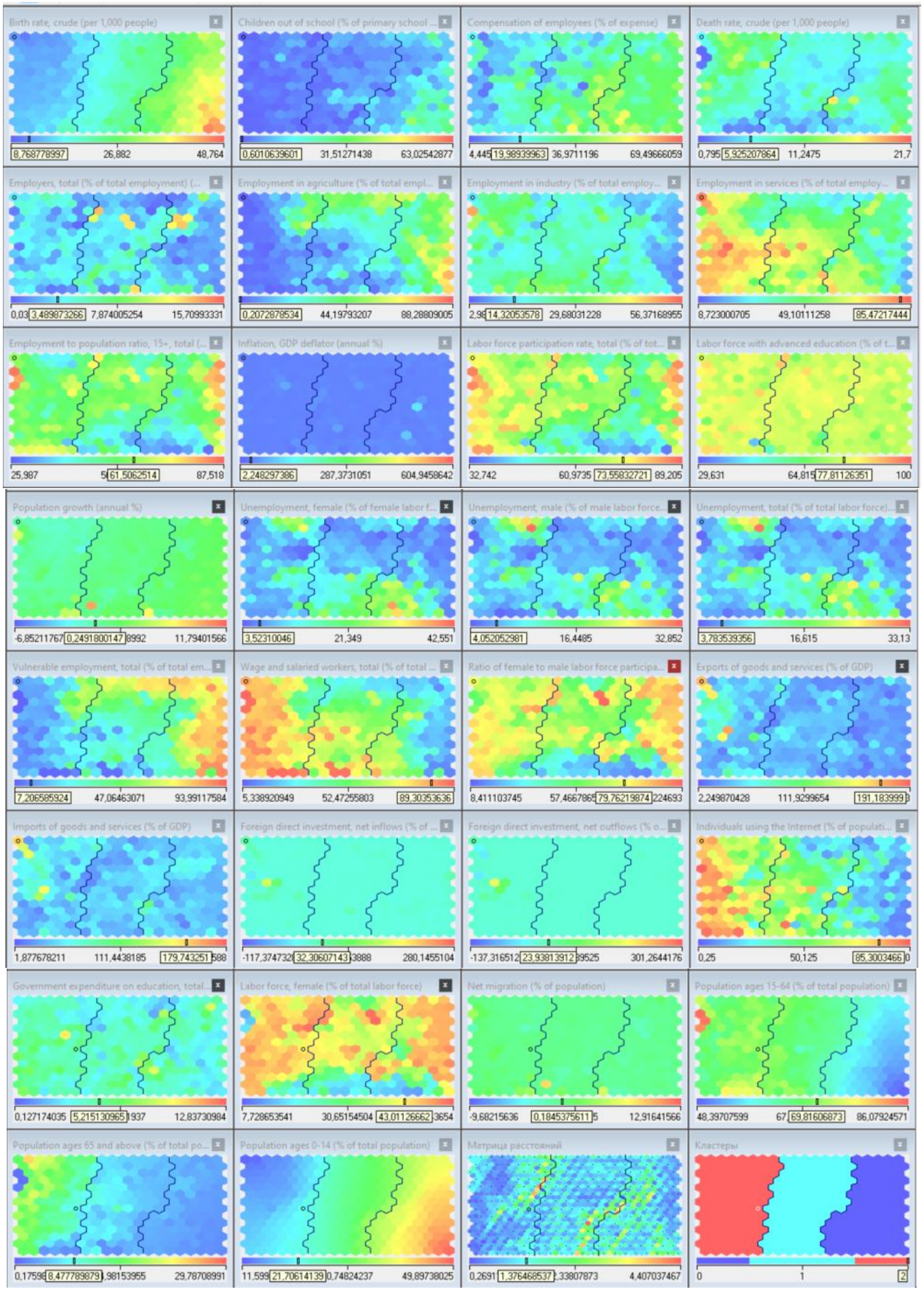

Figure 3. Clustering of countries in the world by employment indicators

Empirical studies have shown that the most effective distribution, in which the indicators of countries retain the greatest similarity in groups, is observed when dividing the world's countries into 12 clusters. The result of this division is shown in Fig. 4.

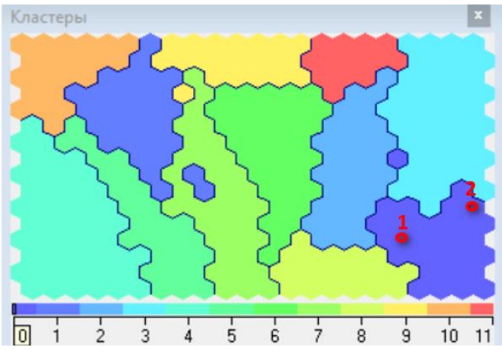

Figure 4. Clustering of countries
in the world by employment measurement indicators

Fig. 4 shows the location of Ukraine in red: under the label "1" Ukraine's position on the Kohonen map during 2010–2017, "2" indicates the position in the period from 2018 to 2021. It can be seen that Ukraine has moved higher and to the right, approaching cluster 3 (the cluster number can be identified by its colour and the ruler with numbers under the map). Table 2 shows the content of each cluster by country and year.

Table 2

DISTRIBUTION OF COUNTRIES IN THE WORLD BY CLUSTERS

Cluster	Countries
0	belarus (2010–2015), Bosnia and Herzegovina, Bulgaria (2010–2021), Croatia (2010–2017; 2020–2021), Cuba (2010–2021), Czech Republic (2010–2014), Estonia (2010–2011), Greece, Hungary, Italy (2010–2021), Latvia (2010–2012), Lithuania (2010–2013), Mauritius (2020–2021), Montenegro (2015–2021), North Macedonia (2010–2021), Poland (2010–2017), Portugal (2010–2016), Romania (2010–2021), russia (2010), Serbia (2010–2021), Slovakia, Slovenia, Spain (2010–2016), Ukraine (2010–2021).
1	Belize (2010–2014), Botswana (2010–2021), Egypt (2010–2018), Eswatini, Gabon, Guatemala (2010–2021), Honduras (2010–2014), Jordan (2010–2020), Kyrgyzstan (2017–2021), Lesotho (2010–2021), Namibia, Samoa, Tajikistan, Tonga (2010–2021).
2	Barbados (2010–2021), Brazil, Brunei (2016–2021), Chile, Colombia (2010–2021), Costa Rica (2014–2021), Cyprus (2010–2016), Ireland, Iceland, Jamaica (2010–2021), Mauritius (2010–2019), Montenegro (2010–2014), New Zealand, Saint Lucia (2010–2021), The Bahamas (2012–2021), Trinidad and Tobago, Uruguay (2010–2021).

Completion of the table 2

Cluster	Countries
3	Australia, Austria, Belgium (2010–2021), belarus (2016–2021), Canada (2010–2021), Croatia (2018–2019), Czech Republic (2015–2021), Cyprus (2017–2021), Denmark (2010–2021), Estonia (2012–2021), Finland, France, Germany, Japan (2010–2021), Latvia (2013–2021), Lithuania (2014–2021), Luxembourg, Malta, Netherlands, Norway (2010–2021), Poland (2018–2021), Portugal (2017–2021), russia (2011–2021), Singapore (2010–2021), Slovakia, Slovenia (2017–2021), Sweden, Switzerland, The Republic of Korea, The United Kingdom, The United States of America (2010–2021).
4	Angola, Benin, Burkina Faso, Burundi, Cameroon, Chad, Congo, DR Congo, Ethiopia, Gambia, Guinea, Guinea–Bissau (2010–2021), Kenya (2010–2016), Liberia, Madagascar, Malawi, Mali, Mozambique, Niger, Nigeria (2010–2021), Rwanda (2010–2015), Sierra Leone (2010–2021), Solomon Islands (2010–2018), South Sudan (2010–2021), Tanzania (2010–2021), Timor–Leste (2010–2015), Togo (2010–2019), Uganda (2010–2021), Vanuatu (2010–2016), Zambia, Zimbabwe (2010–2021).
5	Equatorial Guinea, Ghana, Haiti (2010–2021), Kenya (2017–2021), Laos, Nepal, Papua New Guinea (2010–2021), Rwanda (2016–2021), Solomon Islands (2019–2021), Togo (2020–2021), Timor–Leste (2016–2021), Vanuatu (2017–2021).
6	Argentina (2010–2021), Belize (2018–2021), Brazil, Brunei (2010–2015), Cape Verde (2016–2021), Colombia (2010–2019), Costa Rica (2010–2013), Dominican Republic (2010–2021), Ecuador (2016–2021), El Salvador (2015–2021), Fiji (2017–2021), Guyana (2019–2021), Israel (2010–2021), Indonesia (2017–2021), Jamaica (2010–2016), Kazakhstan, Malaysia, Mexico, Panama (2010–2021), Paraguay (2019–2021), Peru (2020–2021), South Africa, Suriname (2010–2021), The Bahamas (2010–2011).
7	Bangladesh (2010–2021), Belize (2015–2017), Bhutan (2010–2015), Bolivia, Cambodia (2010–2021), Cape Verde, Ecuador (2010–2015), El Salvador (2010–2014), Fiji (2010–2016), Guyana (2010–2018), Honduras (2015–2021), India (2010–2021), Indonesia, Kyrgyzstan (2010–2016), Myanmar (2010–2017), Mongolia, Nicaragua (2010–2021), Paraguay (2010–2018), Peru (2010–2019), Philippines, Turkmenistan, Uzbekistan (2010–2021).
8	Albania, Armenia, Azerbaijan (2010–2021), Bhutan (2016–2021), China, Georgia, Moldova (2010–2021), Myanmar (2018–2021), Thailand, Vietnam (2010–2021).
9	Algeria (2010–2021), Egypt (2019–2021), Iran (2010–2021), Jordan (2021), Lebanon (2010–2021), Maldives (2010–2016), Morocco (2010–2021), Saudi Arabia (2010–2020), Sri Lanka (2010–2021), Syria (2020–2021), Tunisia, Turkey (2010–2021).
10	Afghanistan, Comoros, Gaza, Iraq, Mauritania, Pakistan, Sao Tome and Principe, Senegal, Somalia, Sudan (2010–2021), Syria (2010–2019), Yemen (2010–2021).
11	Bahrain, Kuwait (2010–2021), Maldives (2017–2021), Oman, Qatar (2010–2021), Saudi Arabia (2021), UAE (2010–2021).

As can be seen from Fig. 4 and Table 2, the clustering results put Ukraine in cluster 0. The values of the indicators in each of the 12 groups are characterised by a higher level of homogeneity than when the countries were divided into 3 clusters (or other clustering options tested). Fig. 5 shows the heat maps for each of the explanatory indicators (17 most significant factors were selected from the base set of 30).

We will pay special attention to the zero cluster, as we are most interested in the group of countries that Ukraine has fallen into. Therefore, let us describe it in more detail, and give a brief description of the others.

Cluster 0 includes countries characterised by a low birth rate, low compensation of employees relative to total expense, an average death rate; the share of employers in total employment is below average compared to other countries, most of the labour force is involved in the service sector, the percentage of employment in agriculture is low, the level of employment in industry is average, the ratio of employment to the population aged 15+ and unemployment is also average, and vulnerable employment at a low level; share of wage earners, percentage of the population using the Internet above average; high share of female in the total workforce; countries in this cluster also have average or above average shares of the population aged 15–64, 65+ years and a small share of the population aged 0–14 years.

The main characteristics of the countries in Cluster 1 are a large share of female in the total workforce and a small share of the population aged 65+; the rest of the indicators in this cluster are average or slightly below average.

Cluster 2 includes countries with a small share of employers; low birth rates, employment in agriculture; the bulk of the population engaged in the service sector; unemployment and vulnerable employment below average; above average share of employees; a large share of female in the total workforce; and a low percentage of the population aged 0–14.

Cluster 3 includes countries with low values for the following indicators: birth rate, compensation of employees (per cent of total expense), share of employers in total employment, employment in agriculture, unemployment rate, vulnerable employment, share of the population aged 0–14; high values for the following indicators: employment in the service sector, share of employees, internet users, female in the total labour force, and population aged 65+.

Countries in Cluster 4 have high values for the following indicators: birth rate, agricultural employment, vulnerable employment, share of female in the labour force, population aged 0–4; low values for the following indicators: share of employers and employees, employment in industry and services, Internet users, population aged 15–64 and 65+.

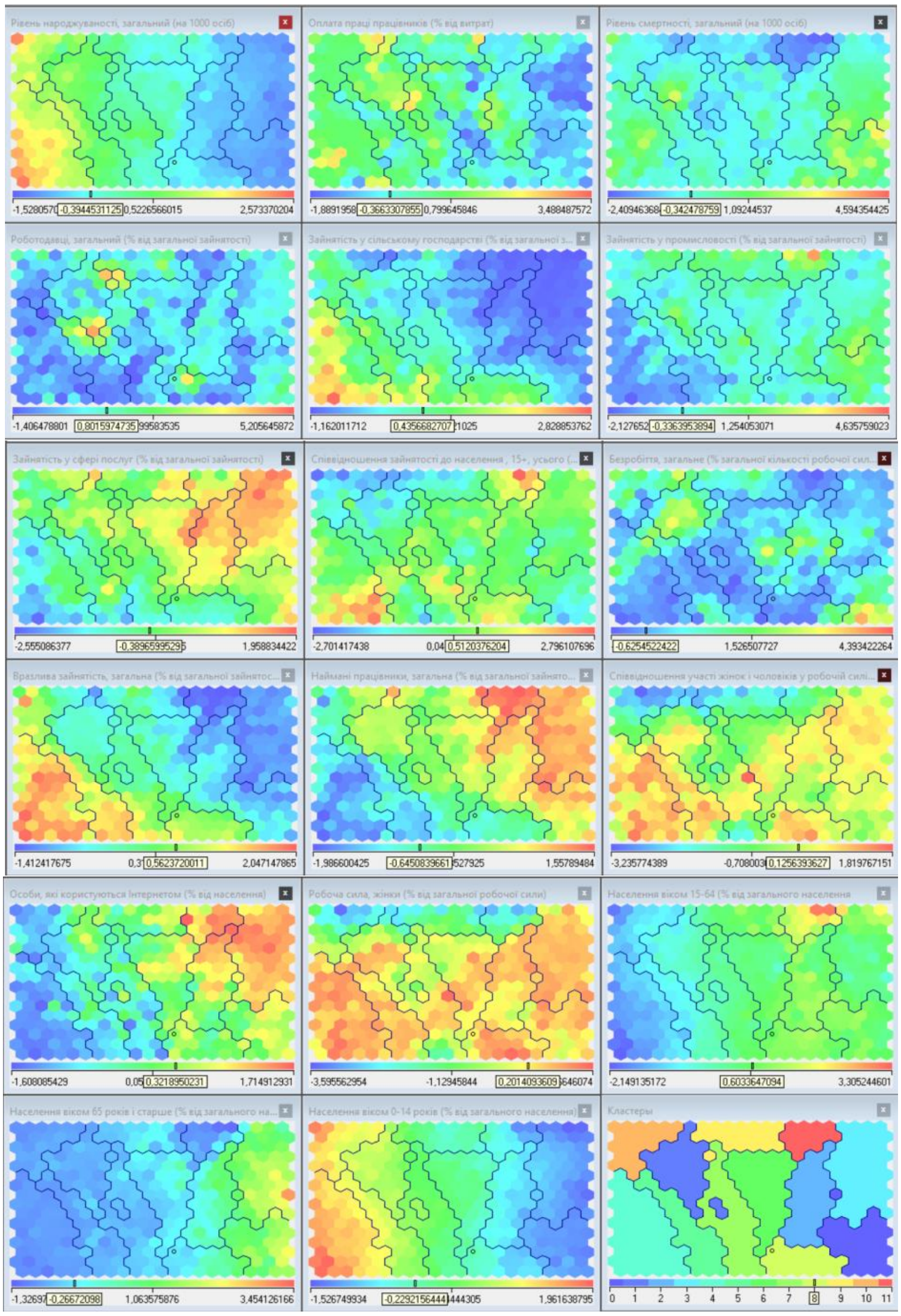

Figure 5. Topological planes of the Kohonen map by indicators of labour force participation

Cluster 5 includes countries characterised by low values for the following indicators: employment in industry, services, wage earners, Internet users, population aged 15–64 and 65+; high or above average values are inherent in the following indicators: birth rate, vulnerable employment, female labour force participation rate, population aged 0-14.

Cluster 6 includes countries with low agricultural employment, below-average industrial employment, as the bulk of the workforce is employed in the service sector; below-average vulnerable employment and percentage of the population aged 0–14; and above-average shares of service sector employment, wage earners, and Internet users.

Cluster 7 is characterised by a low unemployment rate, a small share of the population aged 65+, and a small percentage of employers in total employment; the rest of the indicators are average.

Cluster 8 includes countries whose main characteristics are low birth rates, unemployment, industrial employment, and population aged 0–14 years; a high proportion of female in the total workforce; and other indicators studied are generally average.

Cluster 9 includes countries with low values for the following indicators: death rate, agricultural employment, and share of the population aged 65+; other indicators are mostly average.

The main characteristics of the 10th cluster are higher than average values of the following indicators: birth rate, compensation of employees (per cent of total expense), employment in the service sector, vulnerable employment, population aged 0–14; average unemployment rate; low percentage of Internet users, population aged 15–64, 65+.

Cluster 11 includes countries that have low birth rates, death rates, unemployment, vulnerable employment, a small share of employers in total employment, female in the total labour force, population aged 0–14, 65 and above; low employment in agriculture, but above average employment in services and industry, and a high share of wage earners, Internet users, and population aged 15–64.

Thus, the clustering by employment indicators allowed us to identify characteristic groups of countries with similar socio–economic and demographic characteristics of labour markets. The obtained results clearly demonstrate significant differences in the structure and peculiarities of labour resources use.

The cluster analysis shows that the labour market in Ukraine did not undergo significant changes during the study period of 2010–2021, as the country remained in the same cluster all this time. Moreover, Ukraine's location on the right side of the Kohonen map, where the most developed countries are grouped, indicates a high level of labour market development. This is further confirmed by the grouping of the most problematic countries in terms of employment and labour resources in the opposite cluster, cluster

10, which indicates the greatest differences between countries in these clusters. It is also worth noting that in 2018, Ukraine changed its position within one cluster, moving closer to the group of more developed countries, such as: Canada, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland, The United Kingdom, The United States, etc.

Conclusions

This study analyses the essence of the unemployment phenomenon, the state and dynamics of its development in Ukraine, as well as the latest scientific works in this area. It allowed formulating the research objective, which is to improve the efficiency of analysis and management of labour resources using modern mathematical tools.

The first step in achieving this goal was to form an information base for the study. For this purpose, a number of databases were analysed from both Ukrainian agencies, including the State Employment Service of Ukraine and the State Statistics Service of Ukraine, and a number of international organisations, including the World Bank, the International Labour Organisation, UNDP, including data from the World Resources Institute, the World Economic Forum, etc. It has been shown that for modelling labour force employment for a particular country on the basis of previous years' data for that country, the number of observations is too small to identify significant patterns. This is especially true for Ukraine, given the constant changes in political and socio-economic life, which introduces significant heterogeneity in the processes under study in different time periods.

To improve the accuracy of employment modelling, it was decided to take into account not only data from a single country, but also information from a number of other countries. However, this raises the problem of significant differences in the state and patterns of labour market development in different countries. Therefore, in order to model labour force participation rates, it was decided to create homogeneous data sets by country group. For example, in order to model the level of employment in Ukraine, it is necessary to collect a database that will include information on other countries whose labour market conditions and development are similar to the situation in Ukraine. To this end, it is advisable to cluster countries based on the indicators studied over a certain period of time. Those countries that fall into the same cluster as Ukraine as a result of segmentation will be included in the training sample for further study of the state and modelling of labour market dynamics in Ukraine.

The study selected more than 40 primary indicators from the World Development Indicators database of the World Bank, which measure unemployment, employment, labour market conditions, demographic and macroeconomic characteristics of 203 countries. During the processing the gaps, a number of small countries, such as Nauru, New Caledonia, The Channel Islands, The Isle of Man, etc., were excluded from the overall database due to a significant number of blank fields for a large number of indicators. It is argued that this will not affect the adequacy of the study, since the labour markets of these countries are not similar to Ukraine, and, accordingly, would hardly be included in the same cluster as Ukraine even if there were comprehensive data on them. The gaps for the remaining countries in the sample were filled with the average values for the corresponding indicator for groups of countries with the same level of human development.

The authors also argued for the expediency of using relative indicators in clustering to enable comparison of countries of different sizes. Accordingly, a number of relative indicators from the original list were selected for the final list of factors, and a number of new relative predictors were constructed on the basis of others. As a result, a database of 30 indicators for 173 countries over the 12-year period from 2010 to 2021 was formed for further analysis and clustering.

The next stage of the study was the clustering of countries by the level of socio-economic development and labour resources potential. To this end, a wide range of clustering methods were analysed and the Kohonen self-organising mapping tool was chosen, which not only allows for the formation of homogeneous groups of objects under study but also provides a convenient tool for visual analysis of clustering results. Unlike other algorithms, the location of an object on the Kohonen map immediately gives the analyst a visual representation of how well the property under study is developed in comparison with other objects.

Numerous experiments have revealed that the most effective distribution, in which the indicators of countries retain the greatest similarity in groups, is observed when the world's countries are divided into 12 clusters. In this case, Ukraine falls into a cluster with the following countries: Croatia, Czech Republic, Greece, Hungary, Poland, Slovenia, etc., which are characterised by a mostly low birth rate, low compensation of employees relative to total expense, and an average death rate; the share of employers in total employment is below average compared to other countries; most of the labour force is involved in the service sector; the percentage of employment in agriculture is low; the level of employment in industry is average; the ratio of employment to the population aged 15+ and unemployment is also at an average level, vulnerable employment is low; the share of employees, the percentage of the population using the Internet is above average; the share of female in the total labour force is high; the countries of this cluster also have average or above average shares of the population aged 15–64, 65+, and a small share of the population aged 0–14.

It can be concluded that Ukraine's position on the self-organising map indicates a high level of labour market development. Moreover, in 2018, Ukraine changed its position within one cluster, moving closer to the group of more developed countries.

A promising development of this study will be the construction of a mathematical model for forecasting employment and unemployment for Ukraine based on the countries of the same cluster as Ukraine (zero) and the study of labour force employment in Ukraine during the war (when the relevant statistics are made public).

** This article was translated from its original in Ukrainian*

References

1. Aggarwal, C. C., and C. K. Reddy. *Data Clustering: Algorithms and Applications*. 1st ed. Chapman and Hall/CRC, 2014. https://doi.org/10.1201/9781315373515.

2. Ahmad, M., Y. Khan, C. Jiang, S. Kazmi, and S. Abbas. "The Impact of COVID-19 on Unemployment Rate: An Intelligent Based Unemployment Rate Prediction in Selected Countries of Europe." *International Journal of Finance & Economics* 28, no. 1 (2023): 528–543. https://doi.org/10.1002/ijfe.2434.

3. Alba, E., A. Nakib, and P. Siarry. *Metaheuristics for Dynamic Optimisation*. Springer, 2013. https://doi.org/10.1007/978-3-642-30665-5.

4. Analytical portal 'Slovo i Dilo'. "Ukraine is Among the Ten Countries with the Highest Unemployment Rates in the World." October 12, 2023. https://www.slovoidilo.ua/2023/10/12/infografika/svit/ukrayina-vxodyt-desyatky-krayin-svitu-najvyshhymy-pokaznykamy-bezrobittya. [in Ukrainian].

5. Brunner, K., A. Cukierman, and A. H. Meltzer. "Stagflation, Persistent Unemployment and the Permanence of Economic Shocks." *Journal of Monetary Economics* 6, no. 4 (1980): 467–492. https://doi.org/10.1016/0304-3932(80)90002-1.

6. Brusco, M., E. Shireman, and D. Steinley. "A Comparison of Latent Class, K-means, and K-median Methods for Clustering Dichotomous Data." *Psychological Methods* 22, no. 3 (2017): 563-580. https://doi.org/10.1037/met0000095.

7. Chakraborty, T., A. K. Chakraborty, M. Biswas, S. Banerjee, and S. Bhattacharya. "Unemployment Rate Forecasting: A Hybrid Approach." *Computational Economics* 57 (2021): 183–201. https://doi.org/10.1007/s10614-020-10040-2.

8. Davidescu, A., S.-A. Apostu, and A. Paul. "Comparative Analysis of Different Univariate Forecasting Methods in Modelling and Predicting the Romanian Unemployment Rate for the Period 2021-2022." *Entropy* 23, no. 3 (2021): Article 325. https://doi.org/10.3390/e23030325.

9. Diachkina, A. "Unemployment in Ukraine: Why the Number of Officially Unemployed People has Changed and What will be the Situation on the Market After the War." *Slovo i Dilo (Word and Business)*, June 4, 2022. https://www.slovoidilo.ua/2022/06/04/stattja/suspilstvo/bezrobittya-

ukrayini-chomu-kilkist-oficzijno-bezrobitnyx-zmenshylasya-ta-yaka-bude-sytuacziya-rynku-praczi-pislya-vijny. [in Ukrainian].

10. Du, K.-L., and M. Swamy. *Neural Networks and Statistical Learning*. Springer, 2013. https://doi.org/10.1007/978-1-4471-5571-3.

11. Fu, Z., and L. Wang. "Colour Image Segmentation Using Gaussian Mixture Model and EM Algorithm." *Multimedia and Signal Processing, Communications in Computer and Information Science* 346 (2012): 61–66. https://doi.org/10.1007/978-3-642-35286-7_9.

12. International Labour Organisation. "Labour Force Participation Rate." 2023. https://ilostat.ilo.org/data/.

13. Katris, C. "Prediction of Unemployment Rates with Time Series and Machine Learning Techniques." *Computational Economics* 55 (2020): 673–706. https://doi.org/10.1007/s10614-019-09908-9.

14. Kohonen, T. *Self-Organising and Associative Memory*. 3rd ed. Springer, 2012. https://doi.org/10.1007/978-3-642-88163-3.

15. Kohonen, T. "Essentials of the Self-Organising Map." *Neural Networks* 37 (2013): 52–65. https://doi.org/10.1016/j.neunet.2012.09.018.

16. Kolot, A., O. Herasymenko, A. Shevchenko, and I. Ryabokon. "Employment in the Coordinates of Digital Economy: Current Trends and Foresight Trajectories." *Neuro-Fuzzy Modelling Techniques in Economics* 11 (2022): 78–123. http://doi.org/10.33111/nfmte.2022.078.

17. Kozlovskyi, S., D. Bilenko, M. Kuzheliev, R. Lavrov, V. Kozlovskyi, H. Mazur, and A. Taranych. "The System Dynamic Model of the Labour Migrant Policy in Economic Growth Affected by COVID-19." *Global Journal of Environmental Science and Management* 6, no. SI (2020): 95–106. https://doi.org/10.22034/GJESM.2019.06.SI.09.

18. Lukianenko, I., M. Oliskevych, and O. Bazhenova. "University Competitiveness in the Knowledge Economy: A Kohonen Map Approach." *CEUR Workshop Proceedings* 3465 (2023): 236–250. https://ceur-ws.org/Vol-3465/paper23.pdf.

19. Lukianenko, I., M. Oliskevych, and O. Bazhenova. "Regime Switching Modelling of Unemployment Rate in Eastern Europe." *Ekonomickә časopis* 68, no. 4 (2020): 380–408.
https://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.cejsh-76a31cc8-9e67-4e21-a51e-e98181b4fbe5.

20. Matviychuk, A., O. Lukianenko, and I. Miroshnychenko. "Neuro-fuzzy Model of Country's Investment Potential Assessment." *Fuzzy Economic Review* 24, no. 2 (2019): 65-88. https://doi.org/10.25102/fer.2019.02.04.

21. National Bank of Ukraine (NBU). "Inflation Will Continue to Decelerate and the Economy Will Recover – NBU Inflation Report." 2023. https://bank.gov.ua/ua/news/all/inflyatsiya-i-dali-spovilnyuvatimetsya-a-ekonomika-vidnovlyuvatimetsya--inflyatsiyniy-zvit-nbu. [in Ukrainian].

22. Nikulina, M., I. Sotnyk, O. Derykolenko, and I. Starodub. "Unemployment in Ukraine's Economy: COVID-19, War and Digitalisation." *Mechanism of an Economic Regulation* 1-2, no. 95–96 (2022): 25–32. https://doi.org/10.32782/mer.2022.95-96.04.

23. Oliskevych, M., and V. Tokarchuk. “Dynamic Modelling of Nonlinearities in the Behaviour of Labour Market Indicators in Ukraine and Poland.” *Economic Annals-XXI* 169, no. 1-2 (2018): 35–39. https://doi.org/10.21003/ea.V169-07.
24. Radosavljević, J. *Metaheuristic Optimisation in Power Engineering*. IET, 2018. https://doi.org/10.1049/pbpo131e.
25. Stasinakis, C., G. Sermpinis, K. Theofilatos, and A. Karathanasopoulos. “Forecasting US Unemployment with Radial Basis Neural Networks, Kalman Filters and Support Vector Regressions.” *Computational Economics* 47, no. 4 (2016): 569–587. https://doi.org/10.1007/s10614-014-9479-y.
26. State Employment Service of Ukraine. “Analytical and Statistical Information.” 2023. https://www.dcz.gov.ua/analitics/view.
27. State Statistics Service of Ukraine. “Demographic and Social Statistics.” 2023. http://www.ukrstat.gov.ua/.
28. The World Bank. “World Development Indicators.” 2023. https://databank.worldbank.org/source/world-development-indicators.
29. United Nations Development Programme. “Data Futures Exchange.” 2023. https://data.undp.org/access-all-data.
30. United Nations Development Programme. “Human Development Reports.” 2023. http://hdr.undp.org/en/content/download-data.
31. Verkhovna Rada. *On Employment of the Population (Law of Ukraine 5067-VI)*. 2023. https://zakon.rada.gov.ua/laws/show/5067-17/page#Text. [in Ukrainian].
32. Verkhovna Rada. *On the Protection of the Interests of the Subjects of the Presentation of Information and Other Documents During the Period of War (Law of Ukraine 2115-IX)*. 2023. https://zakon.rada.gov.ua/laws/show/2115-20#Text. [in Ukrainian].

Received: November 13, 2023.
Accepted: November 29, 2023.
Released: December 15, 2023.