

Molchanova K.,

PhD in Economics

Associate Professor at the Department of Logistics,

National Aviation University

Yakovenko Y.,

Student,

National Aviation University

PROSPECTS OF APPLICATION OF DATA MINING TOOLS IN FORECASTING THE DEVELOPMENT OF SOCIO-ECONOMIC SYSTEMS

This report explores important aspects of using data mining tools to predict the development of socio-economic systems. Machine learning methods and their application in the analysis of economic data, as well as the use of natural language processing technologies for understanding textual information are considered. Ethical aspects and challenges of ensuring confidentiality when using these tools are discussed. Examples of practical applications are provided, such as optimization of urban infrastructures and forecasting of regional economic development. The report emphasizes the importance of data mining in strategic decision-making and the development of modern socio-economic systems.

Keywords: data mining, forecasting, socio-economic systems, machine learning methods, natural language processing, ethics, privacy, urban infrastructures, economic development, strategic decisions.

In today's world, where data is becoming a key resource, data mining determines and optimizes the development of socioeconomic systems. Forecasting is becoming an essential tool for management and strategic decision-making. This area of research focuses on how intelligent tools contribute not only to understanding current trends, but also to actively shaping the future by helping to predict and adapt to changes in the socio-economic landscape.

Machine learning is widely used in the analysis of socioeconomic data. Classification methods can recognize patterns and predict categories, such as unemployment or economic growth risks. Cluster analysis groups similar objects to help identify population or market segments.

The graphs on Fig.1 illustrate an ideal and a real situation in terms of consumer clustering using conditional data. [1]

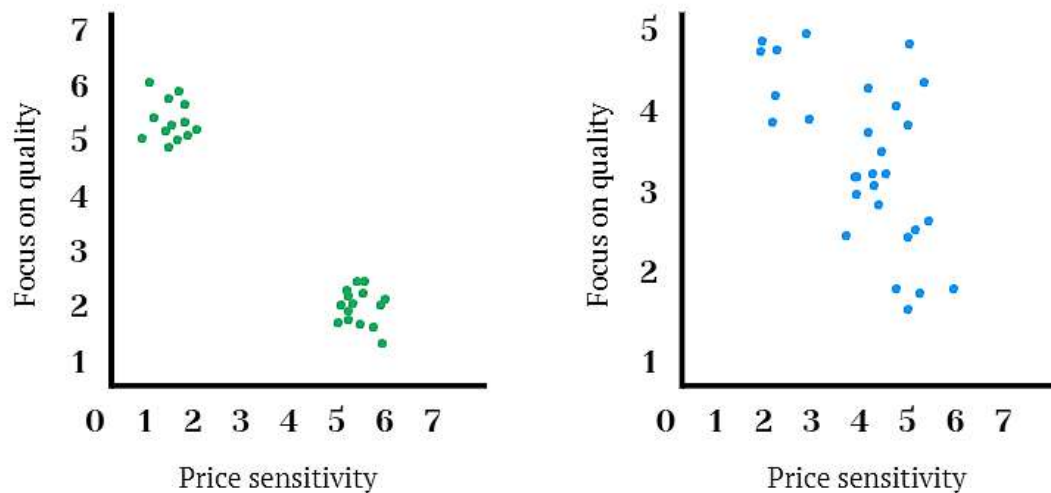


Fig. 1. Ideal and real situations from the point of view of dividing respondents into clusters

On the left graph we see that there are clearly two compact and far-flung "clusters" on the plane where the points corresponding to the respondents' answers to the two questions in the questionnaire are located. It is clear that the respondents are divided into two clusters. The right graph shows a situation that is most common in life. Different researchers may "see" different numbers of clusters here.

Regression analysis is used to predict numerical values such as GDP or income. Neural networks allow modelling complex dependencies in the economy and society. Natural language processing is used to analyze textual information from social networks, media, and other sources. Anomaly detection methods help to identify inconsistencies and unusual phenomena in economic data. All these approaches help in understanding complex socio-economic processes and making informed decisions. The application of machine learning in socioeconomic data analysis also includes image processing methods for analyzing graphs and cartography, where spatial and temporal patterns can be identified. Ensemble methods, such as random forest, allow combining several models to obtain more accurate forecasts. Dimensionality reduction methods, such as the principal component method, help to identify the main factors affecting socioeconomic processes. Machine learning is also used to optimize business processes, detect fraud, and improve the effectiveness of social programs. In general, these methods not only facilitate the processing of large amounts of data but also provide new opportunities for understanding and predicting socioeconomic phenomena.

Analyzing textual information using natural language processing methods can help monitor public sentiment and determine attitudes toward economic issues. For example, social media analysis can detect changes in the tone and emotional colour of

messages, which will allow analyzing public reaction to economic reforms or other events. This approach provides insights into how public opinion can influence the economic climate and decision-making in this context.

The usage of data mining in socio-economic forecasting raises a number of ethical issues that are becoming relevant as these technologies develop. First, it is important to consider the confidentiality and privacy of individual data when it is used for modelling and forecasting. The collection and use of personal information may violate the rights of individuals, and measures must be taken to protect this data.

Additionally, there is a risk of algorithmic bias, where models may misinterpret or take into account differences between social groups, leading to incorrect or unequal decisions. This can exacerbate existing social inequalities and affect equity.

Table 1

Ethical principles in data intelligence for socio-economic forecasting

Ethical Principles	Potential Issues
Data Privacy	Unauthorized use of personal information
Transparency in Algorithm Use	Influence of external interests on analysis results
Avoiding Bias	Exacerbation of social inequality through analysis
Socio-cultural Context Awareness	Unequal impact of technologies on different socio-cultural groups

Subsequently, it is important to consider aspects of transparency and explainability of algorithms. The clarity and accessibility of the analysis can play a key role in how society accepts data mining in socioeconomic forecasting. A clear understanding of how decisions are made and how algorithms affect socio-economic issues fosters open dialogue and supports interaction between developers, authorities, and the public.

Thus, the ethical use of data mining in socio-economic forecasting requires a careful balance between the power of technology and the protection of the rights and interests of individuals and groups in society.

In the process of data analysis, one of the key issues is ensuring the confidentiality and protection of personal data. Personal information that can be collected for analysis is often sensitive and private. This may include personal data about health, finances, location, and other aspects of personal life. There is a risk that if this data is not adequately protected, it may be misused, disclosed, or abused. This is especially true in the era of big data and rapidly evolving analysis technologies. In recent years, technical and legal mechanisms to protect data privacy have improved significantly. The use of anonymization, encryption, and other methods makes it

Секція 2. Інструменти інтелектуального аналізу даних та прогнозування розвитку соціально-економічних систем

harder to access personal information. However, there is an ongoing struggle between protecting privacy and the need to use this data to gain valuable insights. In addition to technical aspects, it is also important to consider ethical and legal issues. As a rule, the definition of what constitutes confidential information can be subjective and context-specific. Legislation and standards in this area are changing, and organizations conducting data analytics should constantly update their practices and policies to meet current privacy and data protection standards.

The use of data mining in urban infrastructure development and planning allows cities to use resources efficiently and improve the quality of life of their residents. For example, geo data analysis can include the use of satellite imagery to determine the condition of roads and other facilities, which can serve as a basis for developing optimal strategies and repair work. In managing the economic development of regions and countries, data mining plays an important role in forecasting and making informed decisions. For example, econometric models can be used to forecast economic growth, allocate resources, and identify effective strategies to stimulate economic development. One example of the use of econometric models and data analysis to forecast the development of regional economies and identify strategies to stimulate economic growth is associated with Oxford Economics. Oxford Economics is a consulting company that specializes in economic data analysis and forecasting. [2] They develop econometric models for different regions and countries to help governments and businesses understand economic trends and make informed decisions. For example, a company may use econometric models to forecast GDP growth, employment, and other key economic indicators for a particular region. Based on these forecasts, they can develop strategies to stimulate economic growth, including recommendations for investments, infrastructure development, and economic development programs. This approach helps to identify potential scenarios for regional development and develop specific strategies to maximize economic impact. On the Fig. 2 is presented World GDP Forecast by Oxford Economics. [3]

In the real world, where data is a key resource, examples include using analytics to identify trends in the consumer market, determine the needs of regions for the development of new industries, and analyze the impact of infrastructure projects on regional economies. Such data can serve as a basis for strategic planning and effective management decisions on a regional or even national scale. Therefore, I propose to consider a specific example of the use of data mining for urban infrastructure development and planning related to Sidewalk Labs, a subsidiary of Alphabet Inc. (Google's parent company). [4] Sidewalk Labs focuses on the development and implementation of innovative technologies to create smart cities. One of their projects is Quayside in Toronto, Canada. Sidewalk Labs uses big data to create smart urban infrastructure. For example, they analyze data on pedestrian traffic, traffic flow, air quality, and other parameters to develop optimal urban planning (see Fig.3).

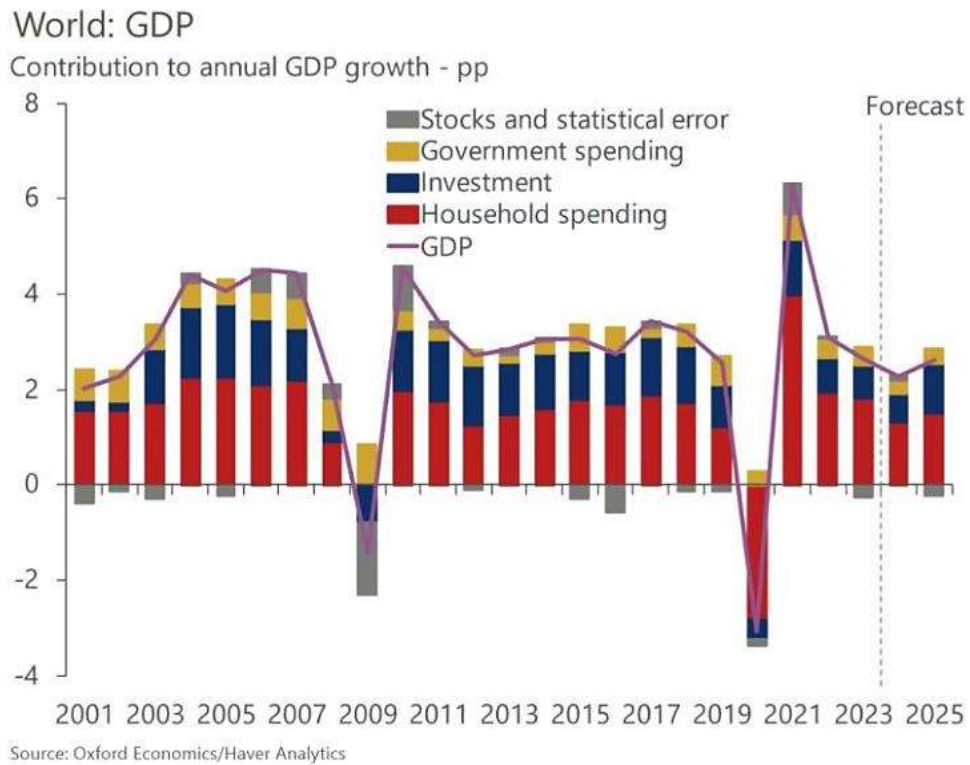


Fig. 2. An example of World GDP by Oxford Economics



Fig. 3. The city of the future: a block with transforming streets in Toronto

Data mining is becoming an essential tool for forecasting the development of socio-economic systems. Taking into account ethical aspects and addressing privacy issues, these tools are becoming key to effective management and strategic decision-

Секція 2. Інструменти інтелектуального аналізу даних та прогнозування розвитку соціально-економічних систем

making. The ability to not only understand and analyze the current state of affairs, but also actively shape the future makes them an integral part of modern socio-economic improvement. Thanks to these tools, societies are able to respond to changes in a timely manner, maximize their potential, and promote sustainable development.

References

1. Рясна М.В., Литвиненко Ю.О. Кластерний аналіз в маркетингових дослідженнях, *Актуальні питання розвитку економіки, харчових технологій та товарознавства* : збірник тез доповідей XXXVIII наукової студентської конференції (м. Полтава, 21–22 квітня 2015 року). Полтава: ПУЕТ, 2015. С.117-120 URL: <http://dspace.puet.edu.ua/handle/123456789/3448> (data of access 26.02.2024).
2. Oxford Economics. URL: <https://www.oxfordeconomics.com/> (data of access 26.02.2024).
3. Oxford Economics. *LinkedIn*. URL: https://www.linkedin.com/posts/oxford-economics_our-300-economists-have-updated-their-monthly-activity-7161725956083167232-CGyF/ (date of access 26.02.2024).
4. Ковальчук І. Місто майбутнього: в Торонто побудують квартал з вулицями-трансформерами. *Eco Technology*. URL: <https://ecotech.news/architecture/349-misto-majbutnogo-v-toronto-pobuduyut-kvartal-z-vulitsyami-transformerami.html> (date of access 26.02.2024).

Бриль І.,

*кандидат економічних наук,
старший науковий співробітник відділу проблем економіки підприємств
Інститут економіки промисловості НАН України, м. Київ*

ІНСТРУМЕНТИ ОЦІНЮВАННЯ ЛЮДСЬКОГО КАПІТАЛУ

В роботі проведено досліджувальну характеристику методик оцінювання людського капіталу. Розглянуто склад та структуру людського капіталу підприємства. Значення цифрових компетенцій для підвищення ефективності людського капіталу.

Ключові слова: людський капітал, методи оцінювання, доходи, структура людського капіталу, підприємство, інформаційні технології, цифрові компетенції.