

Прогнозування показників ринку праці в Україні з використанням методів машинного навчання та часових рядів

Вінницький національний технічний університет;

Анотація

У сучасних умовах нестабільності економіки особливої актуальності набуває задача прогнозування показників ринку праці, зокрема рівня безробіття. Зміни макроекономічних факторів, таких як інфляція та зростання валового внутрішнього продукту, суттєво впливають на зайнятість населення та потребують детального аналізу.

У даній роботі досліджено можливість використання відкритих даних Світового банку для аналізу та прогнозування показників ринку праці України. Проведено обробку часових рядів, включаючи інтерполяцію річних даних до квартальної частоти, що дозволило підвищити точність моделювання. Розглянуто застосування моделей SARIMA та Random Forest для прогнозування рівня безробіття.

Отримані результати показали, що використання комбінованого підходу, який поєднує класичні економетричні методи та алгоритми машинного навчання, дозволяє ефективно моделювати динаміку ринку праці та отримувати достовірні прогнози.

Ключові слова: ринок праці, безробіття, прогнозування, SARIMA, Random Forest, часові ряди, машинне навчання.

Abstract

In modern conditions of economic instability, forecasting labor market indicators, particularly unemployment rate, becomes highly relevant. Changes in macroeconomic factors such as inflation and GDP growth significantly affect employment levels and require detailed analysis.

This paper investigates the use of World Bank open data for analyzing and forecasting labor market indicators in Ukraine. Time series preprocessing was performed, including interpolation of annual data into quarterly frequency to improve modeling accuracy. SARIMA and Random Forest models were applied for unemployment forecasting.

The results demonstrate that a combined approach integrating econometric methods and machine learning algorithms provides effective modeling of labor market dynamics and reliable forecasts.

Keywords: labor market, unemployment, forecasting, SARIMA, Random Forest, time series, machine learning.

Вступ

Ринок праці є важливою складовою економічної системи країни, що відображає рівень зайнятості населення та ефективність економічної політики. У сучасних умовах економічних коливань та глобальних викликів прогнозування показників ринку праці стає необхідним інструментом для прийняття управлінських рішень.

Традиційні методи аналізу не завжди дозволяють врахувати складні взаємозв'язки між економічними показниками, тому актуальним є використання методів машинного навчання разом із класичними моделями часових рядів.

Метою даної роботи є дослідження показників ринку праці України та побудова моделей прогнозування рівня безробіття на основі відкритих даних.

Результати дослідження

У ході дослідження використано відкриті дані Світового банку, що включають показники рівня безробіття, інфляції та зростання ВВП за період 2000–2024 років. Дані були попередньо оброблені, зокрема виконано інтерполяцію до квартальної частоти за допомогою кубічного сплайну.

Проведений аналіз показав наявність трендових та сезонних компонентів у часовому ряді безробіття. Для перевірки стаціонарності використано тест Дікі–Фуллера, після чого застосовано диференціювання.

Для прогнозування було використано дві моделі:

- SARIMA — для врахування сезонності та часових залежностей;
- Random Forest — для моделювання нелінійних взаємозв'язків між показниками.

Оцінка якості моделей проводилась за метриками RMSE, MAE та MAPE. Результати показали, що SARIMA добре описує часову структуру даних, тоді як Random Forest демонструє кращу адаптацію до складних залежностей між змінними.

Висновки

У результаті проведеного дослідження було підтверджено ефективність використання методів машинного навчання та часових рядів для прогнозування показників ринку праці.

Встановлено, що:

- показники ринку праці мають виражену часову структуру;
- макроекономічні фактори впливають на рівень безробіття;
- комбіноване використання SARIMA та Random Forest підвищує точність прогнозування.

Отримані результати можуть бути використані для подальших досліджень та практичного застосування у сфері економічного аналізу та планування.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice. – 3rd ed. – OTexts, 2021.
2. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning. – 2nd ed. – Springer, 2021.
3. McKinney W. Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter. – 3rd ed. – O'Reilly Media, 2022.
4. World Bank World Development Indicators. [Електронний ресурс] – Режим доступу: <https://data.worldbank.org/>
5. Shumway R. H., Stoffer D. S. Time Series Analysis and Its Applications. – 4th ed. – Springer, 2021.

Охріменко Анна Володимирівна – студентка групи СА-22Б, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: cossmann123@gmail.com

Варчук Ілона Вячеславівна – к.т.н., доцент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, e-mail: ivvarchuk@vntu.edu.ua

Okhrimenko Anna V. – student group SA-22B of Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: cossmann123@gmail.com

Varchuk Ilona V. – Varchuk Ilona V. - Ph.D., Assistant Professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: ivvarchuk@vntu.edu.ua