

СИСТЕМНИЙ АНАЛІЗ ТА ІНТЕЛЕКТУАЛЬНЕ ПРОГНОЗУВАННЯ КІЛЬКОСТІ ВОДИ У РІЧКОВОМУ БАСЕЙНІ ПІВДЕННОГО БУГУ

¹Вінницький національний технічний університет

Анотація

У роботі розглянуто задачу інтелектуального прогнозування рівня води водопосту Олександрівка у басейні річки Південний Буг. Для дослідження використано багаторічні гідрологічні дані. Виконано підготовку та аналіз даних, формування ознак часового ряду, побудову та порівняння моделей прогнозування. Досліджено моделі Prophet, ARIMA та алгоритми машинного навчання. Для оцінювання точності використано метрики RMSE та MAPE. За результатами проведеного порівняння найменшу похибку прогнозування показала модель KNeighbors Regressor.

Ключові слова: інтелектуальне прогнозування, рівень води, машинне навчання, часовий ряд, Південний Буг, KNeighbors Regressor.

Abstract

The paper considers the problem of intelligent forecasting of the water level at the Oleksandrivka hydrological station in the Southern Bug River basin. Long-term hydrological data were used for the study. Data preparation and analysis, time series feature engineering, development and comparison of forecasting models were performed. Prophet, ARIMA and machine learning algorithms were investigated. RMSE and MAPE metrics were used to evaluate forecasting accuracy. According to the comparison results, the KNeighbors Regressor model demonstrated the lowest forecasting error.

Keywords: intelligent forecasting, water level, machine learning, time series, Southern Bug River, KNeighbors Regressor.

Вступ

Рациональне управління водними ресурсами потребує своєчасного отримання інформації про їх кількісні характеристики. Проблема ефективного використання водних ресурсів є актуальною через зміни клімату, антропогенне навантаження та необхідність забезпечення сталого водокористування [1, 5].

Для басейну річки Південний Буг прогнозування рівня води має практичне значення для планування використання водних ресурсів та оцінювання гідрологічної ситуації. Використання методів машинного навчання дає можливість аналізувати великі масиви гідрологічних даних і виявляти закономірності, які можуть бути складними для традиційних методів аналізу [3, 4].

Метою роботи є дослідження методів машинного навчання для прогнозування рівня води водопосту Олександрівка та визначення моделі з найменшою похибкою прогнозування.

Результати дослідження

Для проведення дослідження використано датасет із багаторічними спостереженнями за рівнем води водопосту Олександрівка у басейні річки Південний Буг. Використання довготривалих гідрологічних спостережень дає можливість досліджувати зміни показників у часі та формувати прогнозні моделі [2, 5].

На початковому етапі проведено підготовку даних, перевірку їх структури, аналіз пропущених та аномальних значень. Для дослідження закономірностей зміни рівня води виконано розвідувальний аналіз часового ряду.

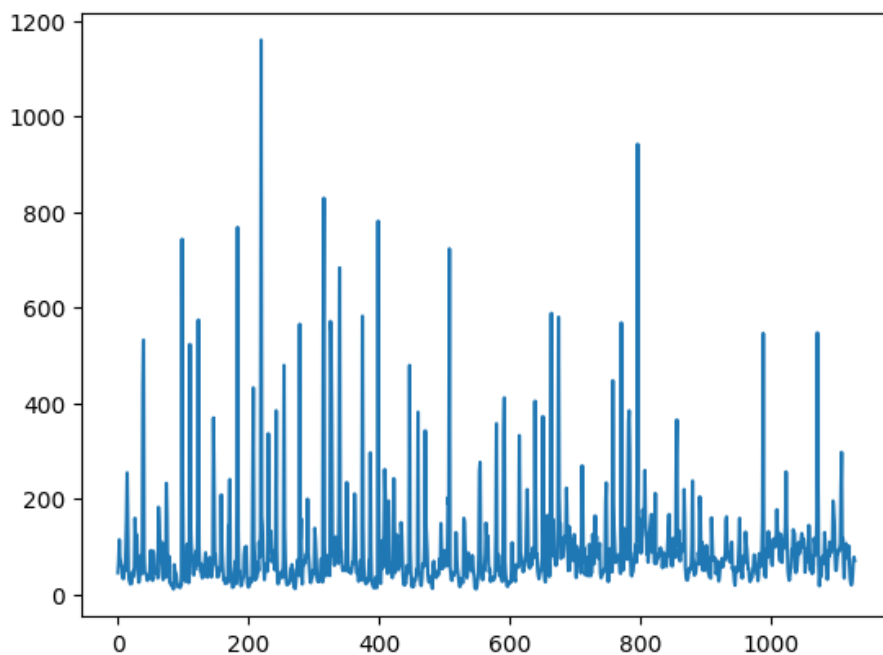


Рис. 1. Графіки часового ряду рівня води

На основі вихідного часового ряду сформовано додаткові ознаки, які використовувалися для навчання моделей машинного навчання. Автоматичне формування характеристик часового ряду дозволяє отримати додаткову інформацію для побудови прогнозних моделей [4].

У роботі досліджено декілька підходів до прогнозування. Серед них моделі Prophet та ARIMA, а також алгоритми машинного навчання: Linear Regression, KNeighbors Regressor, Support Vector Regressor, Random Forest Regressor, XGB Regressor, Bagging Regressor та MLP Regressor.

Застосування різних алгоритмів дозволило порівняти їхню здатність прогнозувати зміни рівня води. Використання машинного навчання для аналізу гідрологічних даних є перспективним напрямом, оскільки такі методи можуть враховувати складні залежності між даними [3, 4].

Для оцінювання якості прогнозування використано метрики RMSE та MAPE. RMSE характеризує величину похибки прогнозу, а MAPE визначає середню відносну похибку у відсотках. Менші значення цих показників відповідають точнішому прогнозуванню.

Для порівняння результатів проведено перевірку побудованих моделей на валідаційній вибірці.

	name_model	type_data	rmse	mape
6	KNeighbors Regressor	valid	8.10754	9.639
1	Prophet_12_days_12_order	valid	12.634104	21.944868
0	Prophet_12_days_3_order	valid	16.004665	25.872943
9	Random Forest Regressor	valid	19.430447	28.876471
10	Bagging Regressor	valid	20.994404	31.780217
4	ARIMA_auto	valid	17.732891	33.548737
2	Prophet_180_days_3_order	valid	23.29503	41.010707
8	Linear SVR	valid	26.098617	41.122685
7	Support Vector Machines	valid	26.50231	41.800386
11	XGB Regressor	valid	29.587776	45.266344
12	MLP Regressor	valid	30.972919	50.057966
5	Linear Regression	valid	31.458928	53.349717
3	Prophet_180_days_12_order	valid	35.881258	67.272247

Рис. 2. Результати перевірки моделей на валідаційній вибірці

За результатами порівняння встановлено, що найменшу похибку прогнозування показала модель KNeighbors Regressor. Отриманий результат свідчить про можливість ефективного використання алгоритмів машинного навчання для прогнозування рівня води на основі історичних спостережень.

Практична перевірка прогнозування також виконувалась на тестовому наборі даних.

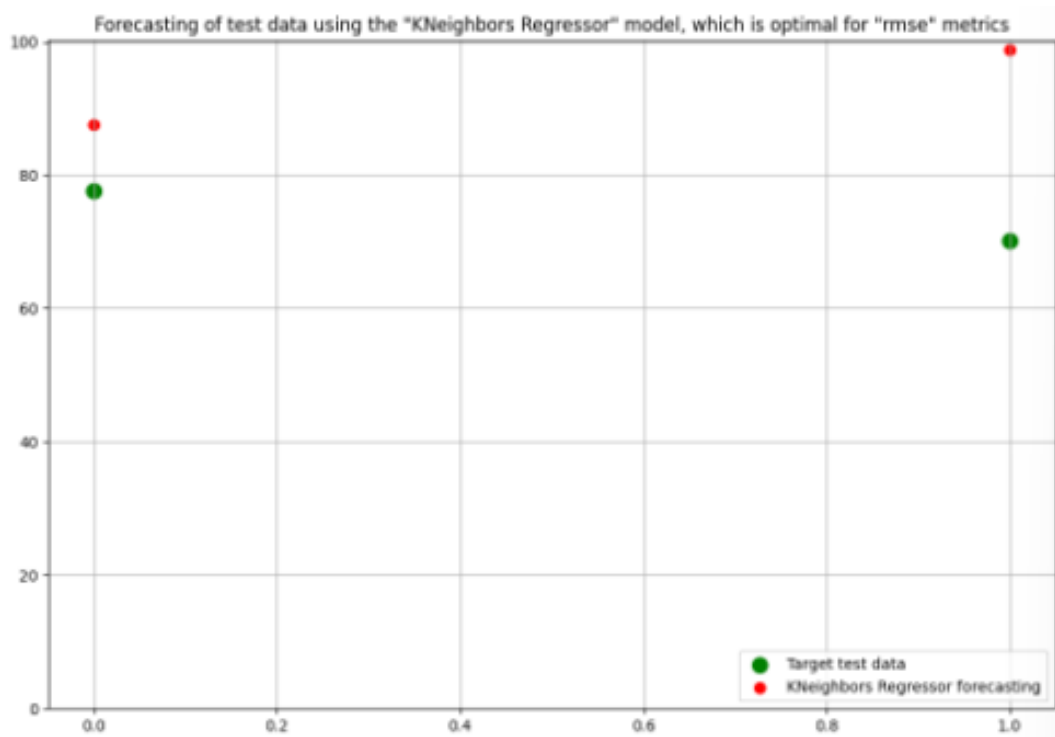


Рис. 3. Результати 5-денного прогнозування рівня води на тестовому наборі даних за допомогою моделі Prophet_90_days_12_order

Отриманий підхід може бути використаний як основа для подальшого створення систем прогнозування рівня води. Результати прогнозування можуть бути корисними для аналізу гідрологічної ситуації, планування водокористування та підтримки прийняття рішень у сфері управління водними ресурсами [5].

Висновки

У роботі досліджено методи інтелектуального прогнозування рівня води водопосту Олександрівка у басейні річки Південний Буг. Проведено підготовку та аналіз багаторічних гідрологічних даних, сформовано ознаки часового ряду та побудовано декілька моделей прогнозування.

Для оцінювання моделей використано метрики RMSE та MAPE. Порівняння результатів показало, що найменшу похибку прогнозування серед досліджених моделей забезпечила модель KNeighbors Regressor.

Отримані результати підтверджують доцільність застосування методів машинного навчання для прогнозування гідрологічних часових рядів. Подальше дослідження може бути спрямоване на використання додаткових гідрометеорологічних показників та розширення періоду прогнозування.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Masood A., Niazkar M., Zakwan M., Piraei R. A Machine Learning-Based Framework for Water Quality Index Estimation in the Southern Bug River. *Water*. 2023. Vol. 15, no. 20. P. 3543. URL: <https://doi.org/10.3390/w15203543>.
2. Science Explorer. USGS. URL: <https://www.usgs.gov/science/science-explorer>.
3. Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. *Hydrology and Earth System Sciences*. 2019. URL: <https://hess.copernicus.org/articles/23/5089/2019/>.
4. Мокін В. Б., Дратовий М. В. Наука про дані: машинне навчання та інтелектуальний аналіз даних : електронний навчальний посібник комбінованого (локального та мережевого) використання. Вінниця : ВНТУ, 2024. 258 с.
5. Мокін В. Б., Гребінь В. В., Крижановський С. М. Методи оцінювання та засоби автоматизації розрахунку складових водогосподарського балансу районів річкових басейнів України : монографія. Вінниця : ВНТУ, 2023. 168 с.

Янченко Каміла Олексіївна – студентка групи 2ІСТ-216, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: kamila4252k@gmail.com

Євгеній Миколайович Крижановський – канд. техн. наук, доцент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: kruzhan@gmail.com

Ігор Миколайович Штельмах – канд. техн. наук, старший викладач кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: igor.shtelmakh@vntu.edu.ua.

Yanchenko Kamila O. - student of Faculty of Intelligent Information Technology and Automation, 2IST-21b,

Vinnytsia National Technical University, Vinnytsia, e-mail: kamila4252k@gmail.com

Kryzhanovsky, Evgeniy M. - candidate of technical sciences, associate professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: kruzhan@gmail.com

Shtelmah, Igor M. – candidate of technical sciences, Senior Lecturer of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: igor.shtelmakh@vntu.edu.ua.