

ПОБУДОВА ІНТЕЛЕКТУАЛЬНОЇ МОДЕЛІ ДЛЯ ПРОГНОЗУВАННЯ ВЕЛИКОДИСПЕРСНОГО ПИЛУ (PM10)

Вінницький національний технічний університет;

Анотація

У роботі розглянуто побудову інтелектуальної моделі для прогнозування концентрації великодисперсного пилу (PM10) в атмосферному повітрі міста Вінниці. Застосовано підхід системного аналізу з використання методів машинного навчання на прикладі Random Forest. Реалізацію проведено на платформі Kaggle з використанням мови Python та бібліотек Scikit-learn, Pandas, Matplotlib. Отримана модель дозволяє здійснювати автоматизоване прогнозування рівня PM10, яка може бути використана у системах моніторингу якості повітря.

Ключові слова: системний аналіз, PM10, прогнозування, якість повітря, машинне навчання, Random Forest, Python, Kaggle.

Abstract

In this article presents the development of an intelligent model for forecasting the concentration of coarse particulate matter (PM10) in the atmospheric air of Vinnytsia city. The study applies a systems analysis approach in combination with machine learning methods, specifically using the Random Forest algorithm. The model was implemented on the Kaggle platform using the Python programming language and libraries such as Scikit-learn, Pandas, and Matplotlib. The resulting model enables automated forecasting of PM10 levels and can be used in air quality monitoring systems to support decision-making in the field of environmental safety.

Keywords: systems analysis, PM10, air quality, forecasting, machine learning, Random Forest, Python, Kaggle.

Вступ

Стан атмосферного повітря є важливим показником рівня життя населення, особливо в урбанізованих територіях. Забруднення повітря шкідливими домішками, зокрема аерозольними частками великої дисперсності (PM10), є серйозною екологічною та медичною проблемою. Високі концентрації даних частинок можуть негативно впливати на органи дихання, серцево-судинну систему, а також провокувати хронічні захворювання.

У місті Вінниця, як і в інших містах України, джерелами пилового забруднення є транспортна діяльність, промислові підприємства, активне будівництво та побутові джерела. Оцінка рівня забруднення ускладнюється через нестачу автоматизованих систем моніторингу, тому виникає потреба у створенні прогнозних моделей, які дозволяють оцінити майбутній рівень концентрації шкідливих речовин на основі аналізу історичних даних.

Дана робота полягає в поєднанні методів системного аналізу з інструментами машинного навчання для побудови надійної інтелектуальної моделі прогнозування PM10. Такий підхід дозволяє враховувати множинні фактори впливу, виявляти приховані закономірності та забезпечувати прийнятну точність моделювання

Метою роботи є розробка програмної реалізації інформаційної системи прогнозування рівня PM10 в атмосферному повітрі міста Вінниці на основі відкритих даних і алгоритмів машинного навчання.

Результати дослідження

Було виконано системний аналіз джерел пилового забруднення та факторів впливу в урбанізованому середовищі. Визначено найбільш значущі параметри, що формують рівень PM10, зокрема: концентрації супутніх забруднювачів (NO2, CO, SO2), погодні умови, час доби, місяць та сезонність.

Для реалізації інтелектуальної моделі використано відкриті екологічні дані, зібрані з платформи Kaggle. Застосовано попередню обробку: очищення від пропусків, генерацію часових ознак, нормалізацію параметрів. Побудовано навчальну вибірку яку зображено на рис. 1, на основі якої здійснено тренування моделі машинного навчання – Random Forest Regressor.

```
In [4]: df_sampled = df.sample(n=100000, random_state=42)

df[['CO', 'NO2', 'SO2', 'O3']] = df[['CO', 'NO2', 'SO2', 'O3']].fillna(df[['CO', 'NO2', 'SO2', 'O3']].mean())

features = ['CO', 'NO2', 'SO2', 'O3', 'month', 'day', 'hour']
X = df[features]
y = df['PM10']

In [5]: from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_absolute_error, r2_score

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

y_pred = model.predict(X_test)
print("MAE:", mean_absolute_error(y_test, y_pred))
print("R²:", r2_score(y_test, y_pred))

MAE: 52.96460047694496
R²: 0.63419669218696
```

Рис. 1. Фрагмент коду реалізації моделі

Здійснено візуалізацію результатів у вигляді графіків: розподіл значень який зображений на рис. 2 та порівняння реальних і прогнозованих даних які зображені на рис. 3.

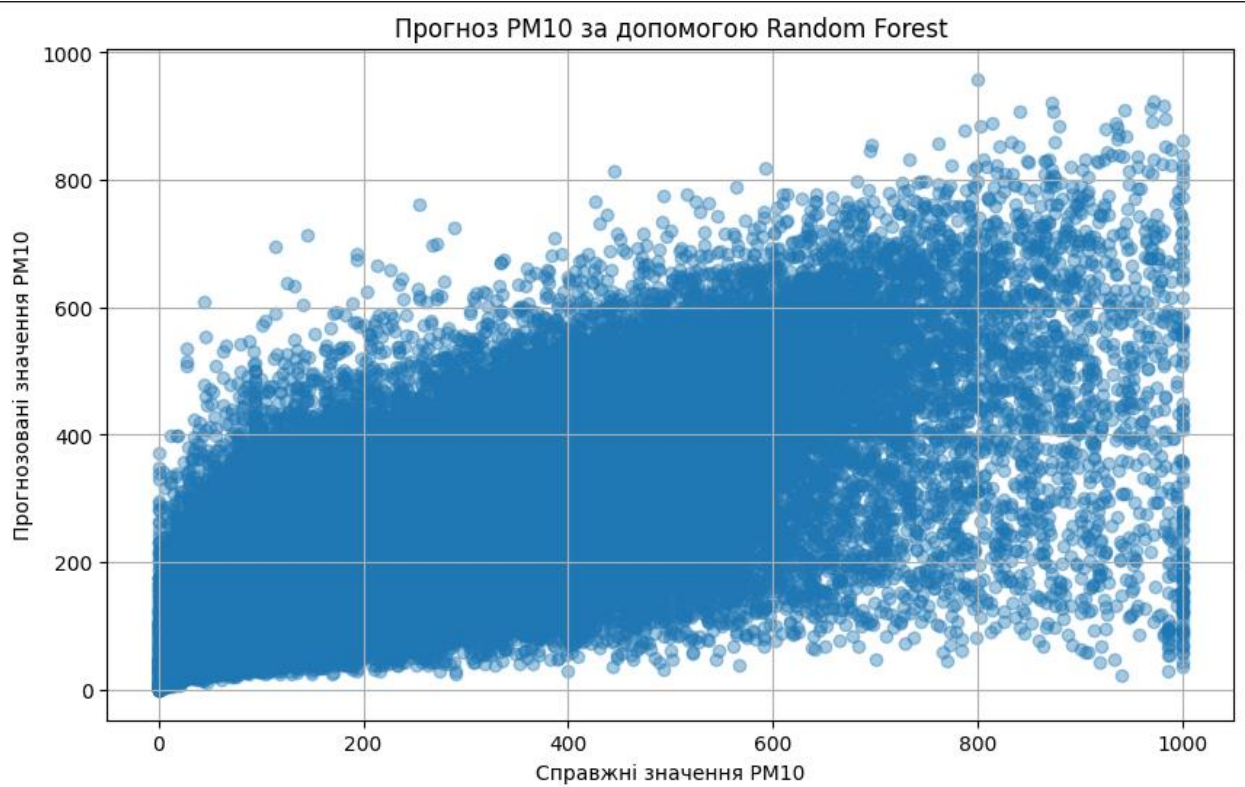


Рис. 2. Графік розподіл значень

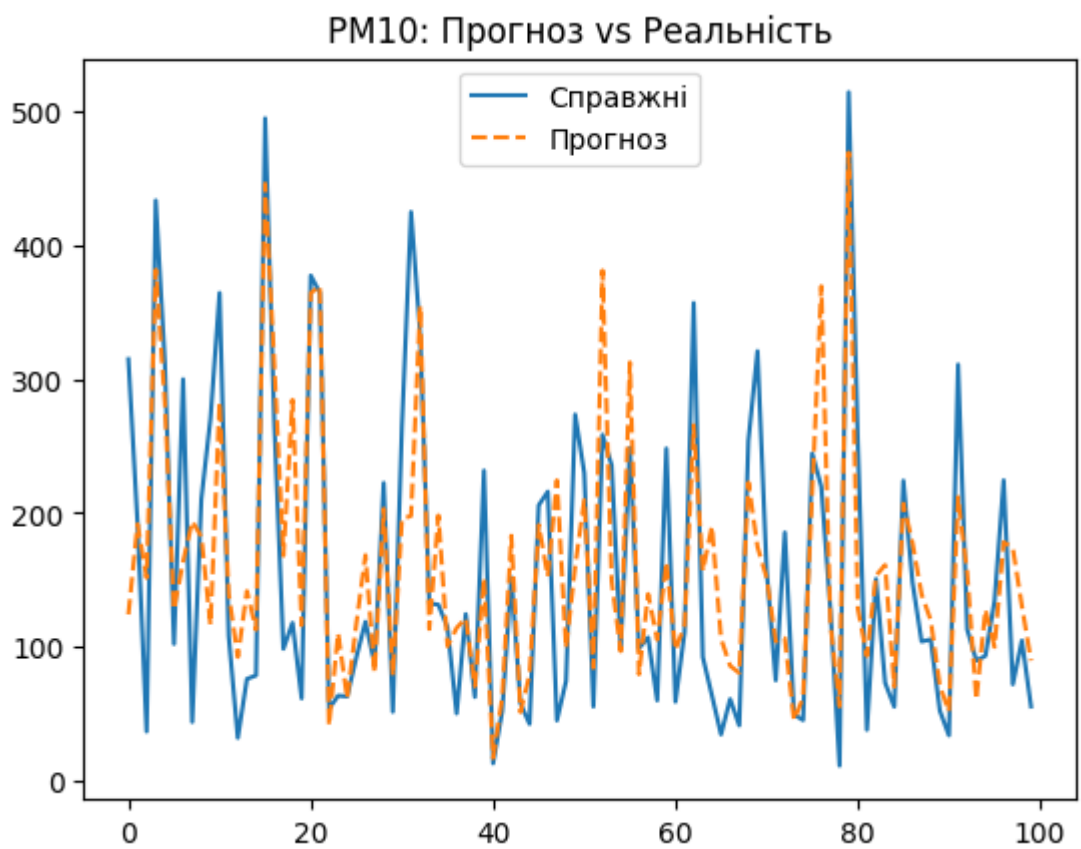


Рис. 3. Графік порівняння прогнозу з реальними даними

Висновки

Розроблену модель збережено у вигляді .joblib-файлу для подальшого використання або інтеграції в автоматизовану систему моніторингу якості повітря.

Таким чином, поставлені завдання дослідження виконано повністю. Отримана модель дозволяє своєчасно здійснювати прогнозування та може бути використана в системах екологічної безпеки. Повну опубліковану версію коду можна знайти за посиланням на платформі Kaggle: <https://www.kaggle.com/code/fah985/pm-10-vinnytsiad89de4258b>

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. URL: <https://www.kaggle.com> – Платформа для аналізу даних та машинного навчання.
2. URL: <https://www.openaq.org> – OpenAQ: Платформа відкритих екологічних даних.

Лучинський Владислав Анатолійович – студент групи СА-216, кафедра системного аналізу та інформаційних технологій, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м.Вінниця, e-mail: luchick919@gmail.com

Науковий керівник: *Варчук Ілона Вячеславівна* – кандидат технічних наук, доцент кафедри САІТ, Вінницький національний університет, м.Вінниця.

Luchinsky V. Anatoliyovych – student of group CA-216, Department of Systems Analysis and Information Technologies, Faculty of Intellectual Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: luchick919@gmail.com

Supervisor: *Varchuk I. Vyacheslavovna* – Ph.D. Sc. (Eng), Associate Professor of the Department of SAIT, Vinnytsia National University, Vinnytsia