

Застосування алгоритмів машинного навчання для прогнозування температури та опадів на основі історичних метеоданих

Щербина Артем,

здобувач вищої освіти спеціальності 122 «Комп'ютерні науки»

Миколаївський національний аграрний університет

м. Миколаїв, Україна

Анотація: у тезах доповіді розглянуто використання алгоритмів машинного навчання для прогнозування температури повітря та кількості опадів на основі історичних метеорологічних даних. Проаналізовано традиційні та сучасні підходи до прогнозування погоди. Запропоновано використання інструментів Python для аналізу метеоданих та побудови прогнозних моделей.

Ключові слова: машинне навчання, прогнозування погоди, метеодані, Python, аналіз даних, штучний інтелект.

Прогнозування погодних умов є важливою міждисциплінарною задачею, що має критичне значення для широкого спектра сфер господарської діяльності та суспільного розвитку. Точні прогнози температури, опадів, швидкості вітру та інших метеорологічних параметрів безпосередньо впливають на ефективність функціонування сільського господарства (планування посівних і збиральних кампаній, управління зрошенням), транспорту (безпека авіаційних, морських і автомобільних перевезень), енергетики (балансування попиту та пропозиції, особливо у відновлюваній енергетиці), а також систем екологічного моніторингу та реагування на надзвичайні ситуації. В умовах зміни клімату та зростання частоти екстремальних погодних явищ значущість високоточного прогнозування лише посилюється.

Традиційні підходи до прогнозування погоди ґрунтуються на фізико-математичних моделях атмосфери, зокрема на розв'язанні систем диференціальних рівнянь динаміки рідини та термодинаміки (чисельні моделі прогнозування погоди – NWP, Numerical Weather Prediction). Такі моделі враховують складні взаємодії між атмосферою, океаном і поверхнею Землі, однак потребують значних обчислювальних ресурсів, високої якості початкових умов і часто демонструють обмеження у відтворенні локальних або короткострокових процесів [1].

З розвитком інформаційних технологій, поширенням сенсорних мереж, супутникових систем спостереження та накопиченням великих обсягів метеорологічних даних з'явилася можливість застосування алгоритмів машинного навчання для аналізу погодних процесів. Машинне навчання дозволяє ефективно виявляти приховані нелінійні закономірності у багатовимірних часових рядах, автоматично адаптуватися до нових даних та будувати прогностичні моделі без явного задання фізичних залежностей.

Зокрема, використовуються такі підходи, як регресійні моделі, дерева рішень, ансамблеві методи (Random Forest, Gradient Boosting), а також глибокі нейронні мережі (LSTM, CNN, трансформери), що здатні враховувати як часову, так і просторову структуру даних [2].

Історичні метеорологічні дані виступають ключовою основою для побудови таких моделей і можуть включати інформацію про температуру повітря, відносну вологість, атмосферний тиск, швидкість і напрямок вітру, кількість опадів, сонячну радіацію та інші параметри. Аналіз цих змінних у часовому та просторовому вимірах дозволяє формувати ознаки (features), що відображають сезонність, тренди та циклічні коливання, а також враховувати лагові залежності між показниками. Поєднання класичних статистичних методів (ARIMA, ETS) із сучасними підходами машинного навчання створює передумови для розробки гібридних моделей прогнозування, які поєднують інтерпретованість і високу прогностичну точність [2].

Таким чином, інтеграція фізичних моделей атмосфери з методами машинного навчання відкриває нові можливості для підвищення якості прогнозів, зменшення невизначеності та адаптації прогнозних систем до умов регіональної специфіки, що є особливо актуальним у контексті цифрової трансформації та забезпечення сталого розвитку.

У таблиці 1 розглянемо основні пакети та бібліотеки для попередньої обробки даних для подальшого навчання моделі.

Таблиця 1 Бібліотеки Python для аналізу метеоданих

Бібліотека	Функціонал
Pandas	Робота з табличними даними, очищення та аналіз
NumPy	Математичні обчислення та робота з масивами
Scikit-learn	Алгоритми машинного навчання
Matplotlib	Побудова графіків
Seaborn	Статистична візуалізація

Джерело: розроблено автором

Основні етапи прогнозування включають збір метеорологічних даних, їх очищення, аналіз параметрів та побудову моделей машинного навчання. Для прогнозування можуть використовуватися алгоритми регресії, дерева рішень та нейронні мережі. Наведемо приклад використання даних прогнозу погоди для побудови моделі випадкового лісу на Python:

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
import matplotlib.pyplot as plt

data = pd.read_csv("weather_data.csv")

X = data[["Humidity", "Pressure", "WindSpeed"]]
y = data["Temperature"]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
model = RandomForestRegressor()
model.fit(X_train, y_train)

pred = model.predict(X_test)
```

Наведений фрагмент коду реалізує базовий пайплайн побудови моделі прогнозування температури на основі метеорологічних параметрів із використанням бібліотек `pandas` та `scikit-learn`. Дані зчитуються з CSV-файлу, після чого формується матриця ознак (вологість, тиск, швидкість вітру) та цільова змінна (температура). Дані розділяються на тренувальну і тестову вибірки, що забезпечує коректну оцінку узагальнювальної здатності моделі. Для навчання використовується ансамблевий алгоритм `RandomForestRegressor`, який добре працює з нелінійними залежностями та є стійким до шуму в даних. Після навчання модель застосовується для прогнозування значень температури на тестовій вибірці.

Для задач прогнозування погодних умов доцільно використовувати комбінацію статистичних та сучасних алгоритмів машинного навчання залежно від горизонту прогнозу, обсягу даних і просторово-часової структури вибірки. Для короткострокових прогнозів (*short-term forecasting*) ефективними залишаються класичні моделі часових рядів, зокрема `ARIMA/SARIMA` (для врахування сезонності) та `ETS (Exponential Smoothing)`, які забезпечують високу інтерпретованість і стабільність на стаціонарних рядах. Водночас ансамблеві методи, такі як `Random Forest` та `Gradient Boosting (XGBoost, LightGBM)`, демонструють кращу здатність моделювати нелінійні залежності між метеорологічними параметрами та можуть ефективно працювати з табличними даними середнього обсягу [3].

Для більш складних задач, особливо при наявності великомасштабних та багатовимірних даних, доцільно застосовувати методи глибокого навчання. Рекурентні нейронні мережі, зокрема `LSTM` та `GRU`, добре підходять для моделювання часових залежностей і довгострокових ефектів у погодних рядах. `Convolutional Neural Networks (CNN)` можуть використовуватися для обробки просторових даних (наприклад, супутникових зображень), а сучасні `Transformer-based` моделі (наприклад, `Temporal Fusion Transformer`) забезпечують високоточні прогнози завдяки механізмам уваги [4]. Найбільш перспективним напрямом є побудова гібридних моделей, які поєднують фізичні (`NWP`) та `data-driven` підходи, що дозволяє підвищити точність прогнозування та адаптувати моделі до регіональної специфіки, зокрема для умов України.

Узагальнюючи, поєднання історичних метеорологічних даних із методами машинного навчання дозволяє створювати ефективні прогностичні моделі, здатні виявляти складні закономірності погодних процесів. Використання алгоритмів на кшталт випадкового лісу демонструє практичну реалізацію підходів, описаних вище, і підтверджує доцільність інтеграції `data-driven` моделей у задачі прогнозування погоди. Це відкриває перспективи для підвищення точності прогнозів, особливо на локальному рівні, та формує основу для подальшого розвитку гібридних і більш складних інтелектуальних систем аналізу кліматичних даних.

Список використаних джерел:

1. Thiessen A. H. Weather glossary. Washington: U.S. Dept. of Commerce, Weather Bureau, 1946. 299 p.
2. Narvekar M., Fargose P. Daily Weather Forecasting using Artificial Neural Network. International Journal of Computer Applications. 2015. Vol. 121, No. 22. P. 9–13.
3. Шмиг Р.А., Боярчук В.М., Добрянський І.М., Барабаш В.М. Термінологічний словник-довідник з будівництва та архітектури / за ред. Р. А. Шмига. Львів: Арал, 2010. 222 с.
4. Jain H., Jain R. Big data in weather forecasting: Applications and challenges. 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), Chirala, Andhra Pradesh, India, 23– 25 March, 2017. P. 233-274.

Abstract: *The thesis considers the use of machine learning algorithms for forecasting air temperature and precipitation based on historical meteorological data. Traditional and modern approaches to weather forecasting are analyzed. The use of Python tools for analyzing meteorological data and building forecast models is proposed.*

Keywords: *machine learning, weather forecasting, meteorological data, Python, data analysis, artificial intelligence.*

Науковий керівник:

Жебко О.О.,

*асистент кафедри економічної кібернетики,
комп'ютерних наук та інформаційних технологій,
Миколаївський національний аграрний університет*

УДК 004.932.2:004.932.72:004.91

Аналіз метаданих у системах управління інформаційними процесами аграрного сектора

Ярченко Дмитро,

здобувач вищої освіти спеціальності 122 «Комп'ютерні науки»

Миколаївський національний аграрний університет,

м. Миколаїв, Україна

Анотація: *Сучасний етап розвитку світового сільського господарства характеризується переходом від традиційних методів господарювання до інтелектуальних моделей, об'єднаних концепцією AgriTech. Ця трансформація базується на тотальному зборі даних, проте справжня цінність інформації полягає не в її обсязі, а в можливості швидкої інтерпретації та включення у контури управління. У цьому контексті аналіз метаданих – даних, що описують контекст, структуру, походження та умови збору первинної інформації – стає фундаментом для побудови систем управління інформаційними процесами*