

О. Л. Мельник, М. В. Драгований, к.т.н., доц., О. М. Козачко, к.т.н., доц.

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ АНАЛІЗУ ТА ПЕРЕДБАЧЕННЯ ОПАДІВ

Анотація. В статті розглянуто проблему аналізу опадів, а також шляхи впровадження автоматизованої технології їх передбачення.

Ключові слова: опади, автоматизація, машинне навчання.

Вступ. Атмосферні опади – атмосферні явища, пов'язані з присутністю в атмосфері води в рідкому або твердому стані, яка випадає з хмар або осідає з повітря на земну поверхню і будь-які предмети.

Систематично спостерігати за погодою люди почали ще за часів Аристотеля. Тоді були виявлені перші закономірності, правда лише на рівні прикмет, типу «вітер з півдня приносить дощ». Мореплавці прокачали техніку передбачення погоди, але вони все одно залишалися швидше мистецтвом, а не наукою. Вони навіть не знали про атмосферний тиск – головну складову сучасної погоди.

Сучасна гідрометеорологія почалася, коли фізики відкрили термодинаміку і гідродинаміку. американець Клівленд Еббе в 1873 році опублікував роботу «Фізична основа довгострокового прогнозування погоди» («The physical basis of long-range weather forecasting»), в якій вперше закликав всіх задуматися над математичним описом погоди. В основу передбачення погоди лягли диференційні рівняння Л. Річардсона, які були запропоновані вченим ще в 1922 році, проте тоді зазнали краху. З появою комп'ютерів було реалізовано математичні моделі числового передбачення погоди, в основі яких лежать рівняння Л. Річардсона.

З розвитком інформаційних технологій зросла і якість передбачення погодних умов. Одним з найпопулярніших напрямів сьогодення є машинне навчання (штучний інтелект).

Вибір моделі прогнозування. Для реалізації системи обрано одну з найпоширеніших бібліотек машинного навчання Scikit-learn. Вона надає широкий вибір алгоритмів навчання з учителем і без вчителя. Навчання з учителем передбачає наявність розміченого датасета, в якому відомо значення цільового показника. У той час як навчання без вчителя не передбачає наявності розмітки в датасеті – потрібно навчитися отримувати корисну інформацію з довільних даних. Основною перевагою даної бібліотеки є те, що вона працює на основі декількох поширених математичних бібліотек, і легко інтегрує їх між собою.

При побудові моделі прогнозування будуть використовуватись три класифікатора: DecisionTreeClassifier, ExtraTreesClassifier, RandomForestClassifier.

Дерева рішень використовують як моделі класифікаторів чи регресії, адже побудована структура дерева дозволяє розбивати набір даних на менші підмножини, що веде до аналізу та передбачення даних. ExtraTrees та RandomForest досить схожі між собою моделі, оскільки вони будують кілька дерев та розбивають вузли, використовуючи випадкові підмножини функцій, але з двома ключовими відмінностями: ExtraTrees не завантажує спостереження (тобто це вибірки без заміни), а вузли розбиваються на випадкові розбиття, в той час як RandomForestModels знижує ризик перевитрати.

Приклад використання розробленої технології. Вхідними даними для виконання є датасет зі статистикою щоденних погодних умов протягом 10 років в м. Вінниця. Він ґрунтується на погодних умовах, на основі яких можливо визначити ймовірність опадів. Набір містить наступні атрибути: дата спостереження; назва локації погодної станції; мінімальна температура в градусах Цельсія; максимальна температура в градусах Цельсія; рекордна кількість опадів в день, в мм; кількість сонячних годин в день; напрямок найсильнішого пориву; швидкість найсильнішого пориву вітру; вологість повітря (у відсотках); атмосферний тиск; хмарність.

Висновки. Розроблена технологія аналізу передбачення опадів є достатньо точною та може бути використана професійно фахівцями в галузі метеорології в цілях вдосконалення сучасних прогнозних систем опадів.

Література

1. Атмосферні опади [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/Атмосферні_опади. – Назва з екрану.
2. Введення в Scikit-learn [Електронний ресурс]. – Режим доступу: <https://neurohive.io/ru/osnovy-data-science/vvedenie-v-scikit-learn/>. – Назва з екрану.
3. Decision Tree Classifier in Python using Scikit-learn [Електронний ресурс]. – Режим доступу: <http://benalexkeen.com/decision-tree-classifier-in-python-using-scikit-learn/>. – Назва з екрану.
4. ExtraTreesClassifier [Електронний ресурс]. – Режим доступу: <https://medium.com/@namanbhandari/extratreesclassifier-8e7fc0502c7>. – Назва з екрану.