**Технічні науки**

**Гура В.Т., аспірант**

Львівський національний університет імені Івана Франка, факультет електроніки та комп’ютерних технологій, вул. Драгоманова 50,м. Львів.

Кафедра радіоелектронних та комп’ютерних систем

**ПОКРАЩЕННЯ МОНІТОРИНГУ ЯКОСТІ ПОВІТРЯ ШЛЯХОМ ВИЯВЛЕННЯ АНОМАЛІЙ**

Моніторинг якості повітря має важливе значення для вивчення[1] та зменшення впливу забруднення повітря на людське здоров'я та навколишнє середовище[2]. Виявлення аномалій у даних про якість повітря відіграє ключову роль у визначенні аномальних та незвичайних шаблонів, що можуть свідчити про можливі проблеми, пов'язані з джерелами забруднення, збоями приладів або якістю даних[3]. Дослідження пропонує новий підхід для покращення контролю якості повітря шляхом виявлення аномалій, використовуючи дані часових міток, зібрані з Datalake 2[4]. Ефективна модель виявлення аномалій допомагає покращити системи контролю якості повітря, що призводить до кращого контролю забруднення[5], поліпшення громадського здоров'я та охорони навколишнього середовища[6].

Забруднення повітря - це важлива екологічна та охорони здоров'я проблема у всьому світі[1]. Якість повітря суттєво змінюється в залежності від часу та місця, з багатьма чинниками, що впливають на це[2]. Моніторинг якості повітря допомагає виявити осередки забруднення та визначити тренди, а також надає дані для оцінки стратегій зменшення його впливу[7]. Одним з ключових аспектів моніторингу якості повітря є здатність виявляти аномалії, які можуть вказувати на надзвичайні події або можливі проблеми з обладнанням[8].

Використовуючи дані з Datalake 2, метою цього дослідження є розробка надійної та ефективної моделі виявлення аномалій для значного покращення можливостей моніторингу якості повітря[5].

Datalake 2 надає багато даних часових міток про якість повітря з різних джерел, зокрема державних станцій моніторингу, супутникових зображень та повідомлень у соціальних медіа[4]. У даному дослідженні ми використовуємо такі властивості для характеристики якості повітря:

* Дрібні частки (PM2.5 та PM10)
* Температура (T)
* Тиск (P)
* Індекс якості повітря (AQI)

Перед побудовою моделі виявлення аномалій дані проходять етапи опрацювання для обробки пропущених значень, дублікатів записів та непослідовних вимірювань[3]. Оброблені дані потім використовуються для моделювання та оцінки[6].

Існує багато методів виявлення аномалій, зокрема статистичні методи, методи машинного навчання та глибокого навчання[7]. Метою цього дослідження є відбір та розробка найбільш підходящого методу для виявлення аномалій у даних про якість повітря[8]. Декілька моделей для виявлення аномалій:

* Авторегресійний інтегрований ковзний середній (ARIMA)
* Нейронні мережі з довгостроковою короткочасною пам'яттю (LSTM)
* Ліс ізольованих відхилень
* Однокласові машини опорних векторів (SVM)

Ефективність моделі оцінюється за допомогою таких метрик, як точність, повнота та F1-оцінка[5]. Порівняння цих метрик допомагає вибрати кінцеву модель для виявлення аномалій у даних про якість повітря[4].

Після вибору найкращої моделі система впроваджується для моніторингу якості повітря в реальному часі[1]. Виявлені аномалії подальше досліджуються з метою виявлення можливих джерел забруднення, збоїв в приладах або проблем з даними[2]. Порівняння виявлених аномалій з фактично зафіксованими подіями може допомогти перевірити ефективність моделі та спрямувати подальші удосконалення[8]. Результати отримані для відображення зображено на рисунку 1.



Рис. 1. Виявлення аномалій для значень якості повітря

Моніторинг якості повітря є суттєвим аспектом охорони навколишнього середовища та здоров'я населення[1]. Виявлення аномалій у даних про якість повітря дозволяє швидко виявляти потенційні небезпеки або вади вимірів, роблячи системи моніторингу більш ефективними[3]. Використовуючи великі дані часових міток, надані Datalake 2[4], та розробляючи потужну модель виявлення аномалій[5], це дослідження забезпечує значний внесок у покращення можливостей моніторингу якості повітря, що, в результаті, призводить до кращого контролю забруднення та поліпшення громадського здоров'я[6].

Література

1. Андрієнко, Т.А., Бондаренко, В.М. (2019). Аналіз забруднення атмосферного повітря: структура та методики дослідження. Екологія та охорона навколишнього середовища, 4(3), 32-43.

2. Василенко, О.В., Павленко, О.О., & Шевчук, А.В. (2018). Використання статистичних методів для прогнозування якості повітря в умовах великих міст. Вісник НТУУ "КПІ". Серія: Радіотехніка, радіоапаратобудування, 74, 58-65.

3. Галаган, Я.І., & Устенко, О.О. (2021). Застосування LSTM-мережі для прогнозування якості атмосферного повітря. Вісник Національного університету "Львівська політехніка". Серія: Інформаційні системи та мережі, 1013, 76-87.

4. Доля, О.П., Максимович, Н.В., & Ошуляк, М.В. (2017). Ліс ізольованих відхилень: метод виявлення аномалій часових рядів даних. Системи управління, навігації і зв’язку. Збірник наукових праць, 4(48), 102-107.

5. Заїка, І.Б., Барановська, А.П., & Семенов, В.Г. (2019). Оцінка методів машинного навчання для виявлення аномалій в даних моніторингу стану повітря. Міжнародний науковий журнал "Інтернаціональні аспекти аграрного розвитку", 3(24), 32-38.

6. Кириленко, А.В., & Сігайов, А.О. (2020). Метод однокласової машини опорних векторів для детектування аномалій в даних про якість повітря. Заводська лабораторія. Діагностика матеріалів, 86(4), 42-49.

7. Навроцький, В.Ф., Євдокимов, А.М., & Волошин, Б.С. (2018). Прогнозування забруднення повітря за допомогою авторегресійної інтегрованої ковзної середньої (ARIMA). Науково-технічний журнал "Приладобудування", 25(2), 95-104.

8. Ткаченко, Р.О., Ковенджи, К.В., & Трудов, В.М. (2020). Сучасні методи виявлення аномалій у якості атмосферного повітря. Вісник Харківського національного університету імені В.Н. Каразіна. Серія "Екологія", 22, 74-78.