

Дослідження впливу параметрів моделі Prophet на точність прогнозування якості атмосферного повітря

¹Вінницький національний технічний університет

Анотація

Було розглянуто вплив параметрів методів бібліотеки Prophet на точність прогнозування якості атмосферного повітря.

Ключові слова: системний аналіз, якість атмосферного повітря, часові ряди, Prophet.

Abstract

The purpose of the paper is to overview how input parameters of the Prophet model affect the quality of air condition prediction.

Keywords: system analysis, atmospheric air condition, time series, Prophet.

Вступ

Важливим елементом захисту стану атмосферного повітря є процес аналізу та передбачення його показників, що дозволяє завчасно вживати необхідних заходів та тим самим попереджувати екологічні погіршення. Важкість цього передбачення має місце, через сезонність змін якості атмосферного повітря та відсутністю повноцінних часових даних. Саме тому, використання моделі прогнозування часових рядів Prophet, є корисним інструментом, адже дана модель здатна показувати гарну точність в зазначених умовах, використовуючи параметри за замовчуванням. Проте, дана модель має широкі можливості налаштування і важливим є питання як різні набори параметрів впливають на точність передбачення часового ряду.

Ключовим елементом побудови моделей машинного навчання є наявність наборів даних. В даній роботі використовуються дані моніторингу якості атмосферного повітря від громадського проекту EcoCity [1] та їх сервісу «Кабінет дослідника» [2] - веб-система, яка дозволяє отримати доступ та використовувати у своїх дослідженнях інформацію, отриману від станцій моніторингу атмосферного повітря (рис. 1).

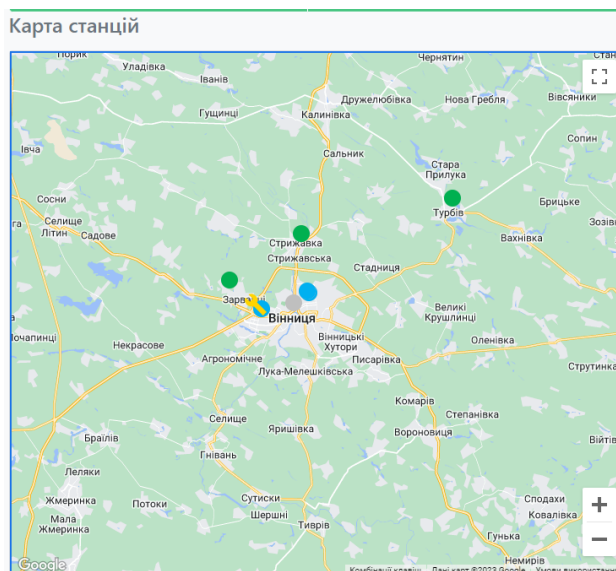


Рис. 1. Карта станцій моніторингу у Вінницькій області

Система надає інформацію по декількох основних показниках, серед яких: пил 2.5 мкм, пил 10 мкм, температура, вологість, атмосферний тиск. В цьому дослідженні використано дані показника PM10 з станції у селищі міського типу Турбів, Вінницька область.

Результати дослідження

Використовуючи раніше набутий досвід прогнозування з використанням моделі Prophet [3-6], було створено програму на Python. Для часового ряду обраного параметру відсікаються аномальні дані, використовуючи певні квантилі і для кожного такого ряду будується модель Prophet. Кожна модель будується з налаштуванням денної сезонності та параметру holidays, який є набором аномальних даних, обраних раніше та додатковим параметром prior_scale, що регулює вплив аномальних даних на загальний прогноз. Для кожного набору тестових даних робиться прогноз та обраховується показник точності. Також, використовуючи можливості бібліотеки Prophet, будуються графіки сезонності та компонентів сезонності. Таблиці з результатами для різних вхідних параметрів показані на рисунку 2.

	threshold	scale	r2
0	0.50	1.0	51.72
1	0.50	5.0	51.84
2	0.50	10.0	51.91
3	0.50	15.0	51.88
4	0.50	20.0	51.89
5	0.50	30.0	51.85
6	0.70	1.0	64.38
7	0.70	5.0	64.48
8	0.70	10.0	64.51
9	0.70	15.0	64.49
10	0.70	20.0	64.49
11	0.70	30.0	64.50
12	0.85	1.0	70.37
13	0.85	5.0	70.43
14	0.85	10.0	70.44
15	0.85	15.0	70.43
16	0.85	20.0	70.43
17	0.85	30.0	70.46
18	0.95	1.0	66.20
19	0.95	5.0	66.27
20	0.95	10.0	66.27
21	0.95	15.0	66.27
22	0.95	20.0	66.26
23	0.95	30.0	66.26

Рис. 2. Результати перевірки впливу вхідних параметрів на точність передбачення

Порівняємо графіки для набору параметрів, які дали найкращий результат та для набору який показав гіршу точність. З рисунків 3, 4 можна зробити такі висновки:

- 1) тренд постійно зменшується на 5-6 місяцях року, що говорить про регулярне зменшення забруднення в цей проміжок;
- 2) значення тренду змінюються в діапазоні [1,3; 6], тобто амплітуда складає приблизно 4,7 одиниць;
- 3) зміна параметру prior_scale не мала істотного впливу на прогнозування певного квантиля;
- 4) денна сезонність не має великого впливу;
- 5) точність передбачення збільшується в межах використання квантиля P70-P85, але при подальшому збільшенні точність починає погіршуватися.

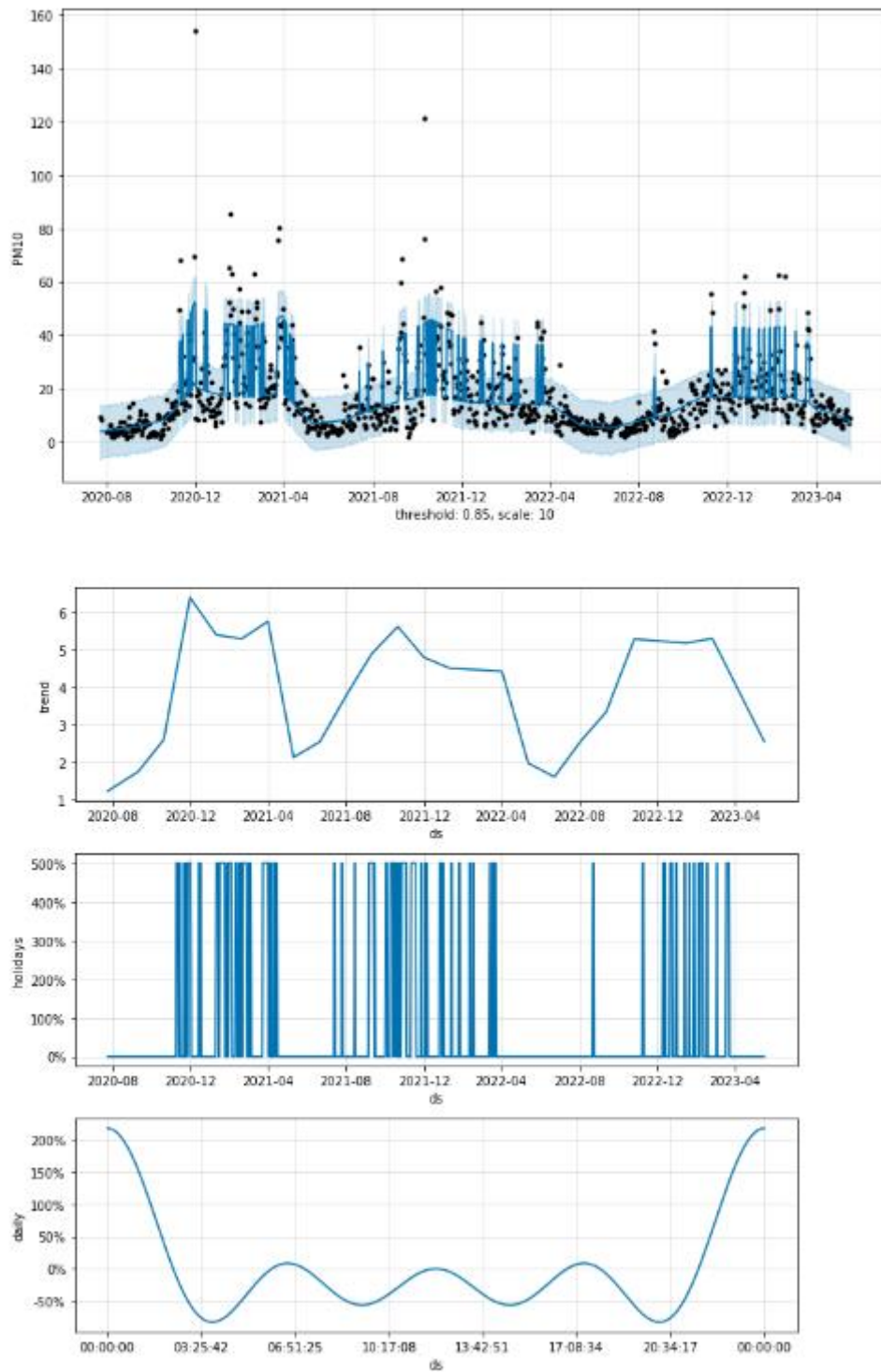


Рис. 3 – Передбачення PM10 з квантилем P85 та значенням prior_scale=10

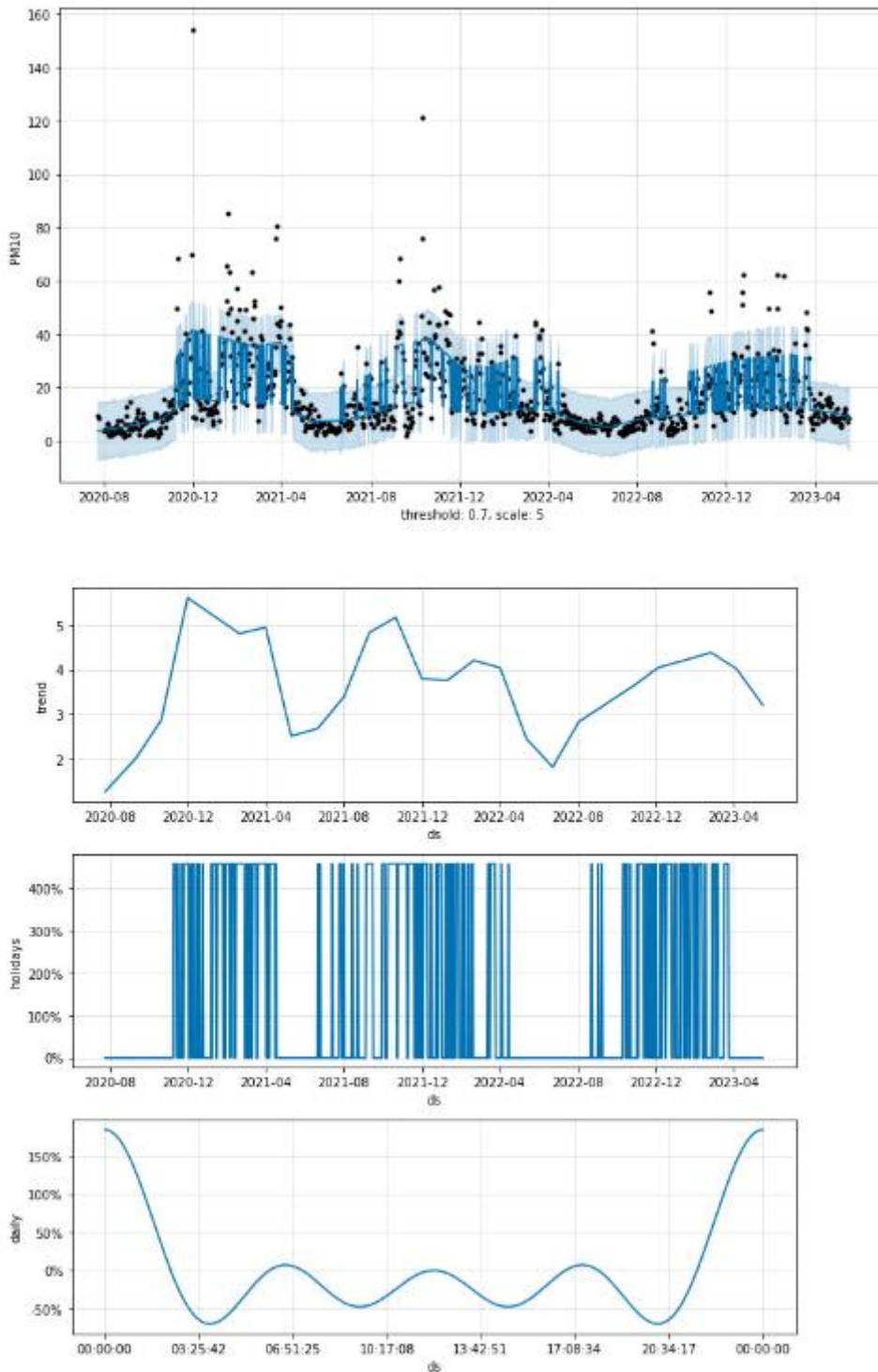


Рис. 4 – Передбачення PM10 з квантилем P70 та значенням prior_scale=5

Висновки

Розглянуто вплив параметрів моделі Prophet на точність прогнозування якості атмосферного повітря за наявними даними моніторингу атмосферного повітря EcoCity. Результати дослідження дозволяють краще налаштувати модель Prophet для подальшого аналізу та прогнозування стану атмосферного повітря.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Eco-City Громадський моніторинг стану якості повітря [Електронний ресурс] – Режим доступу до ресурсу: <https://eco-city.org.ua/>.
2. Eco-City Кабінет дослідника [Електронний ресурс] – Режим доступу до ресурсу: <https://archive.eco-city.org.ua/>.
3. Документація бібліотеки Prophet / Facebook [Електронний ресурс] – Режим доступу до ресурсу: https://facebook.github.io/prophet/docs/quick_start.html.
4. Мокін В. Б. Інформаційна технологія аналізу та прогнозування кількості нових випадків на коронавірус SARS-CoV-2 в Україні на основі моделі Prophet / В. Б. Мокін, А. В. Лосенко, А. Р. Ящолт // Вісник Вінницького політехнічного інституту. – 2020. – № 5. – С. 71–83.
5. Мокін В. Б. Інформаційна технологія аналізу та прогнозування багатохвильової кількості нових випадків захворювань на коронавірус COVID-19 на основі моделі Prophet / В. Б. Мокін, А. В. Лосенко, А. Р. Ящолт // Вісник Вінницького політехнічного інституту. - Вип. 6, с. 65–75, 2020.
6. Мокін В. Б. Прогнозування хвиль коронавірусу на основі відновленої когнітивної карти міжрегіонального впливу / В. Б. Мокін, М. В. Дратований, А. В. Лосенко, С. О. Жуков // Інформаційні технології та комп'ютерна інженерія, 2021, Том 52, Вип. 3, с. 86–94, Груд 2021.

Шмундяк Дмитро Олександрович — аспірант кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, dimashmund@gmail.com.

Shmundiak Dmytro O. – Postgraduate student of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, dimashmund@gmail.com.