

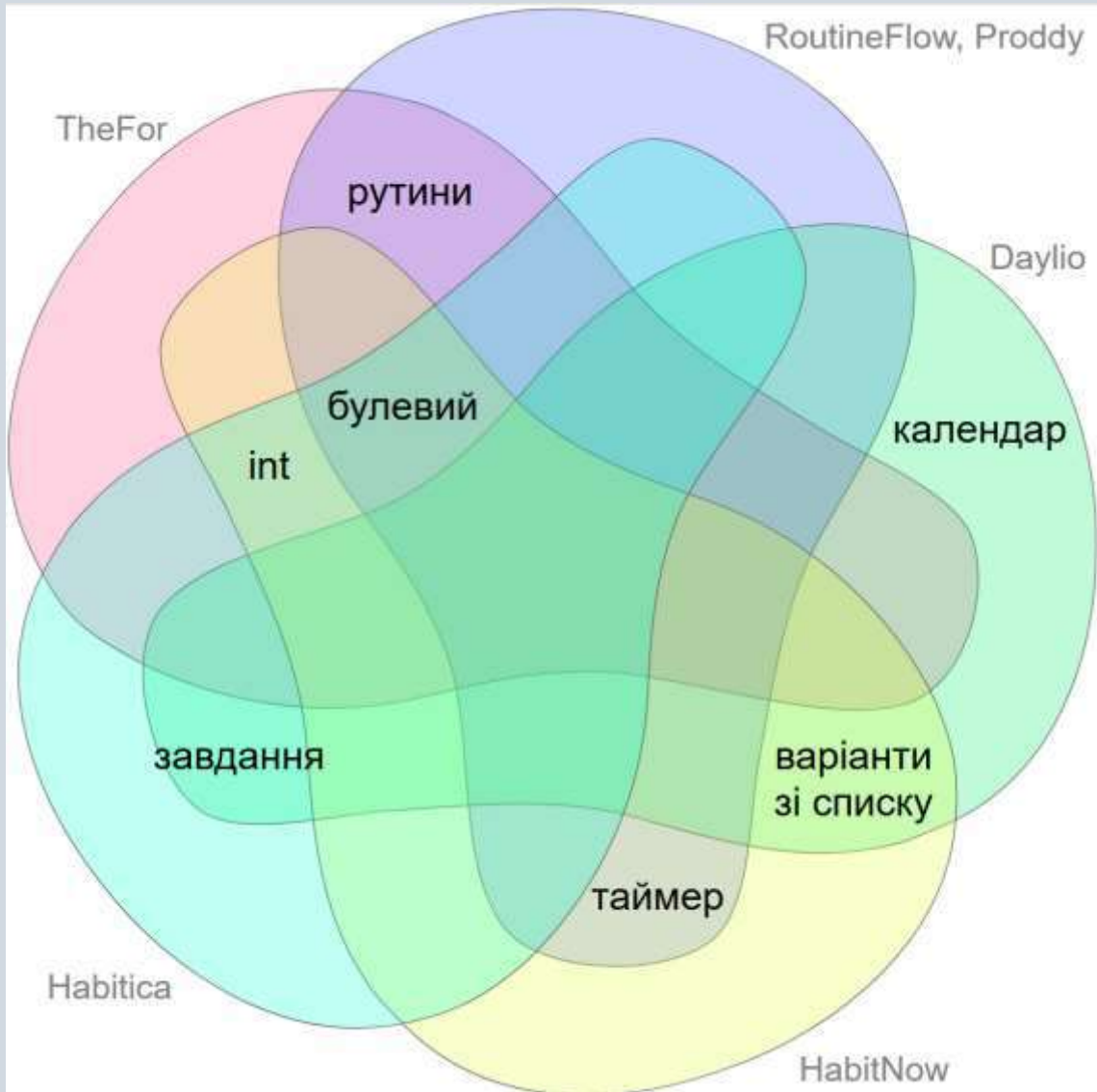
Розробка відстежувача звичок (habit tracker) з елементами прогнозування

ВИКОНАВ: КОЛОМІЙЧУК А.К.

КЕРІВНИК: СТ. В. БОРОЗЕННИЙ С.О.

Мета та задачі

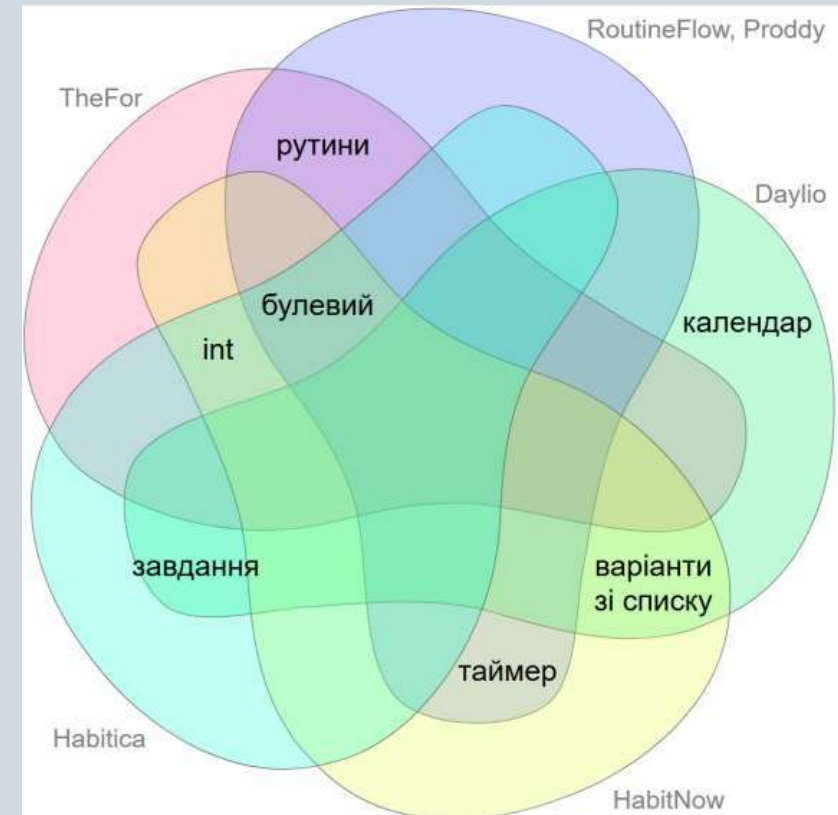
- Надати можливість відслідковувати показники різних типів
- Полегшити та пришвидшити процес внесення даних за допомогою автозаповнень
- Полегшити процес планування за рахунок прогнозів
- Розробити модуль програми для генерації автозаповнень та короткострокових прогнозів
- Розробити модуль програми для відстежування звичок, що об'єднував би функціональні можливості наявних застосунків



Звичні ДОДАТКИ

Типи та форма даних

- Булеве значення
- Число
- Діапазон
- Тривалість
- Час доби
- Варіанти з переліку
- Композитний показник
- Декілька показників на добу



Yesterday

+ACTIVITY +GROUP

<< < 21 22 23 24 25 > >>

Reading

15:37:33 21:12:55

15:36:38	01:15:26
15:36:41	01:14:50
15:37:23	01:13:03
15:37:33	01:13:39
15:37:35	01:12:49
15:36:16	

- Сонячно
- Хмарно
- Дощ
- Гроза
- Іній
- Град

Weather

21:12:55 0 NOTHING SELECTED

07:46:45	25,05	Хмарно
07:46:14	24,87	Дощ
07:46:13	25,2	
07:51:15	25,19	
07:50:10	25,18	
07:46:03	24,96	

Today

+ACTIVITY +GROUP

<< < 22 23 24 25 26 > >>

Reading

16:15:38 01:06:19

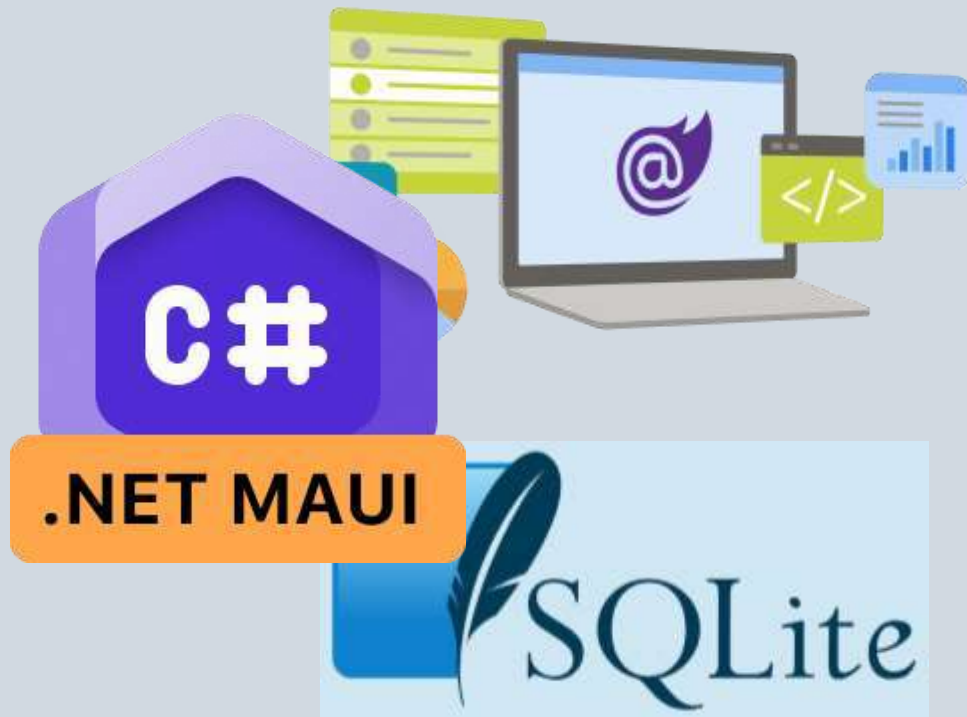
Weather

07:48:34 25,7 ХМАРНО

07:35:21	25,68	Хмарно
07:33:57	25,69	
07:34:33	25,66	
07:34:31	25,7	
07:33:34	25,58	
07:33:24	25,71	

Використані програмні рішення

РОЗРОБКА



ПОРІВНЯННЯ



Проблематика предметної області

- Різниця між активностями та метриками
- Множинність або відсутність записів

Reading

16:15:38 01:06:19

Weather

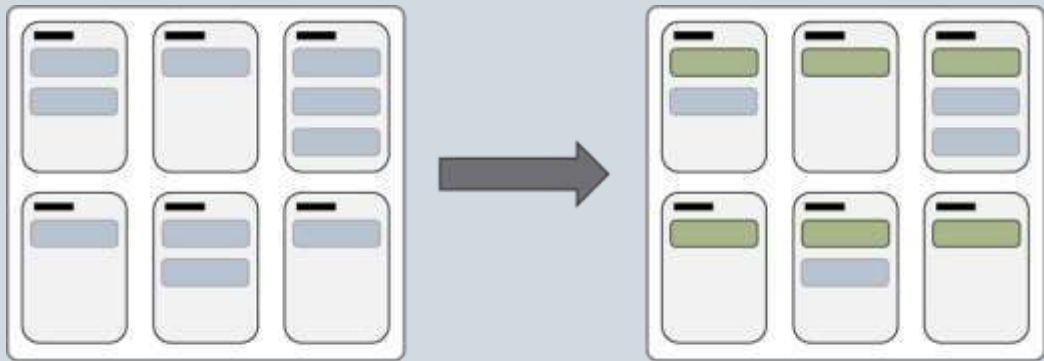
07:48:34 25.7 ХМАРНО

07:35:21	25,68	Хмарно
07:33:57	25,69	
07:34:33	25,66	
07:34:31	25,7	
07:33:34	25,58	
07:33:24	25,71	

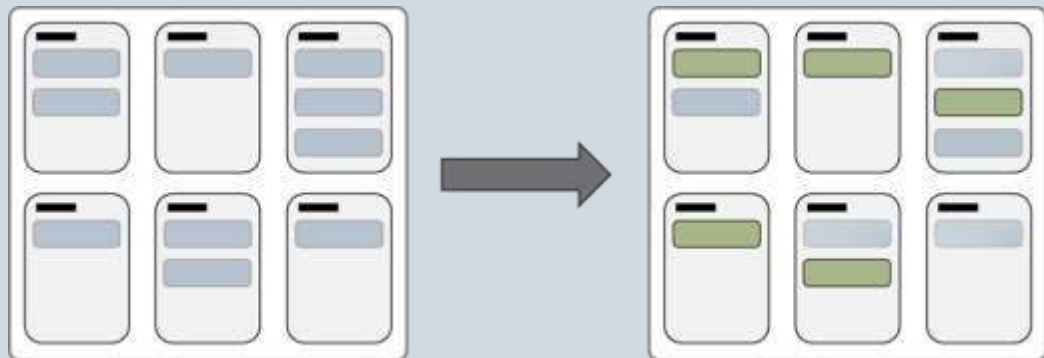
Читання – активність
Погода – метрика

Проблематика предметної області

- Групування вхідних даних
 - Кожен n-тий (загалом)
 - Кожен n-тий (на добу)
 - За часом запису
 - За бінарним значенням інших показників




Групування із вибором кожного 1-го значення у дні






Групування із вибором кожного 2-го значення


Edit options ×


 CHANGE ICON ×


Title
Погодні умови


  Random


Option #1
Сонячно 

Option #2
Хмарно 

Option #3
Дощ 

Option #4
Гроза 

Option #5
Іній 

Option #6
Град 

ADD OPTION

*Меню налаштувань варіантів переліку
(значення незалежні)* SAVE

Проблематика предметної області

Зведення різних типів до числового значення та навпаки:

- Один варіант з переліку
 - Значення незалежні
 - Значення порівнювані
 - Значення порівнювані і циклічні
- Декілька варіантів з переліку
- Бінарне значення
- Тривалість та час доби

Алгоритми прогнозування

Exponential
smoothing (ETS)
(Holt's linear)

Autoregressive
modeling (AR)

Weighted
moving average
(WMA)

Алгоритми прогнозування

Exponential
smoothing (ETS)

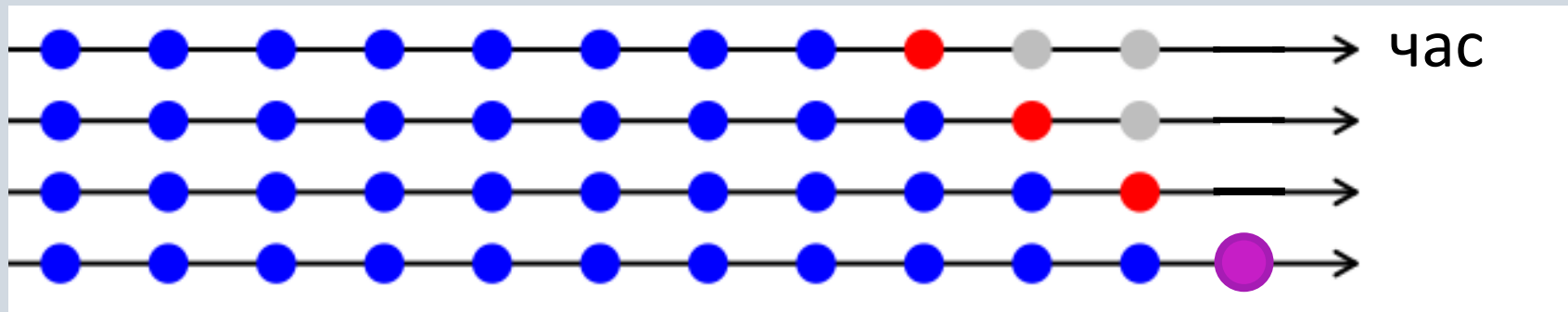
Autoregressive
Integrated Moving
average (ARIMA)

Seasonal and Trend
decomposition
using Loess (STL)

XGBoost &
LightGBM

Оцінка точності

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i^{\text{факт}} - y_i^{\text{прогн}})^2$$

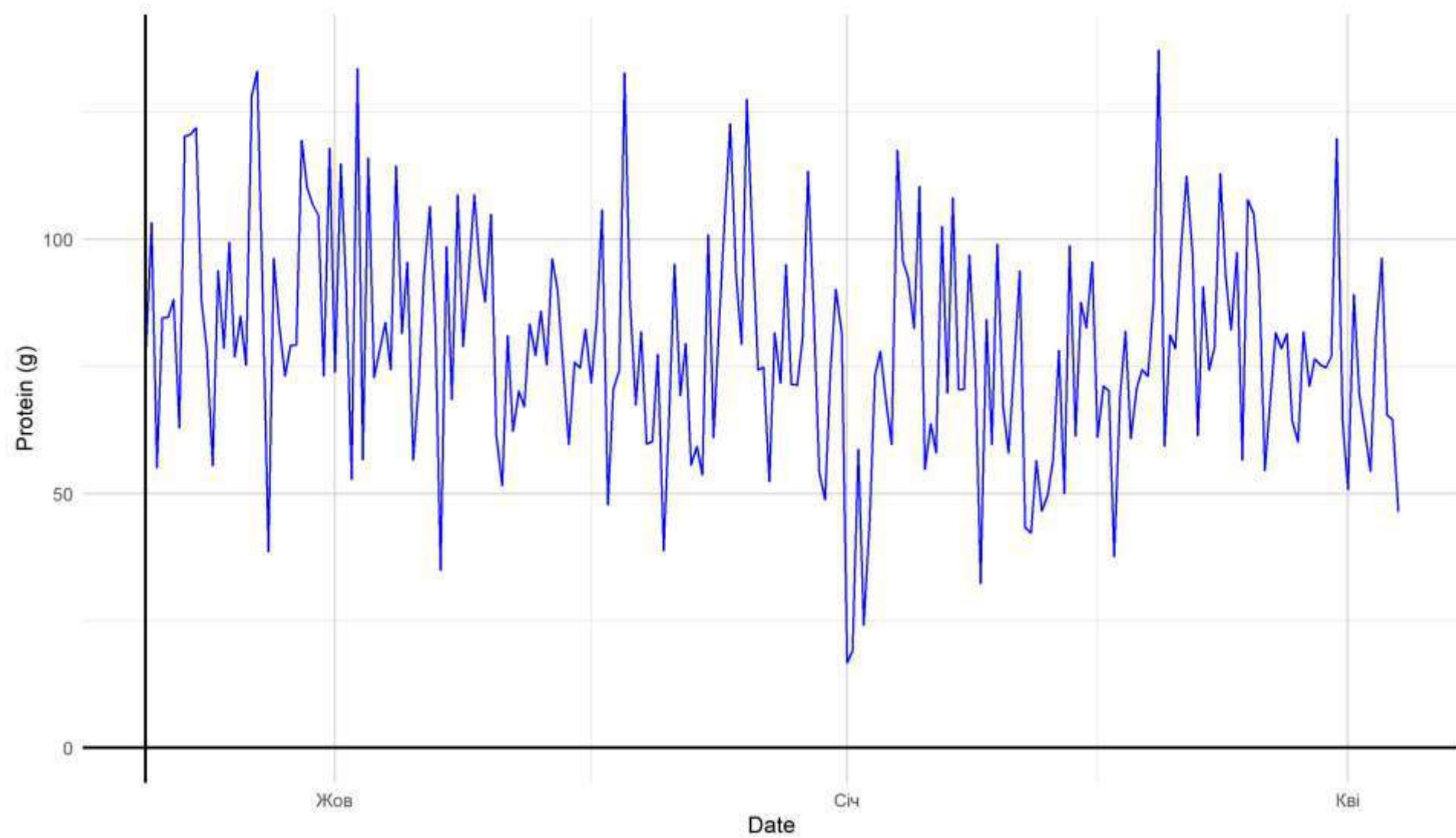


Сині точки – відомі значення

Сірі точки – відомі значення, не задіяні в прогнозуванні

Червоні точки – відомі значення, що прогнозуються для оцінки точності

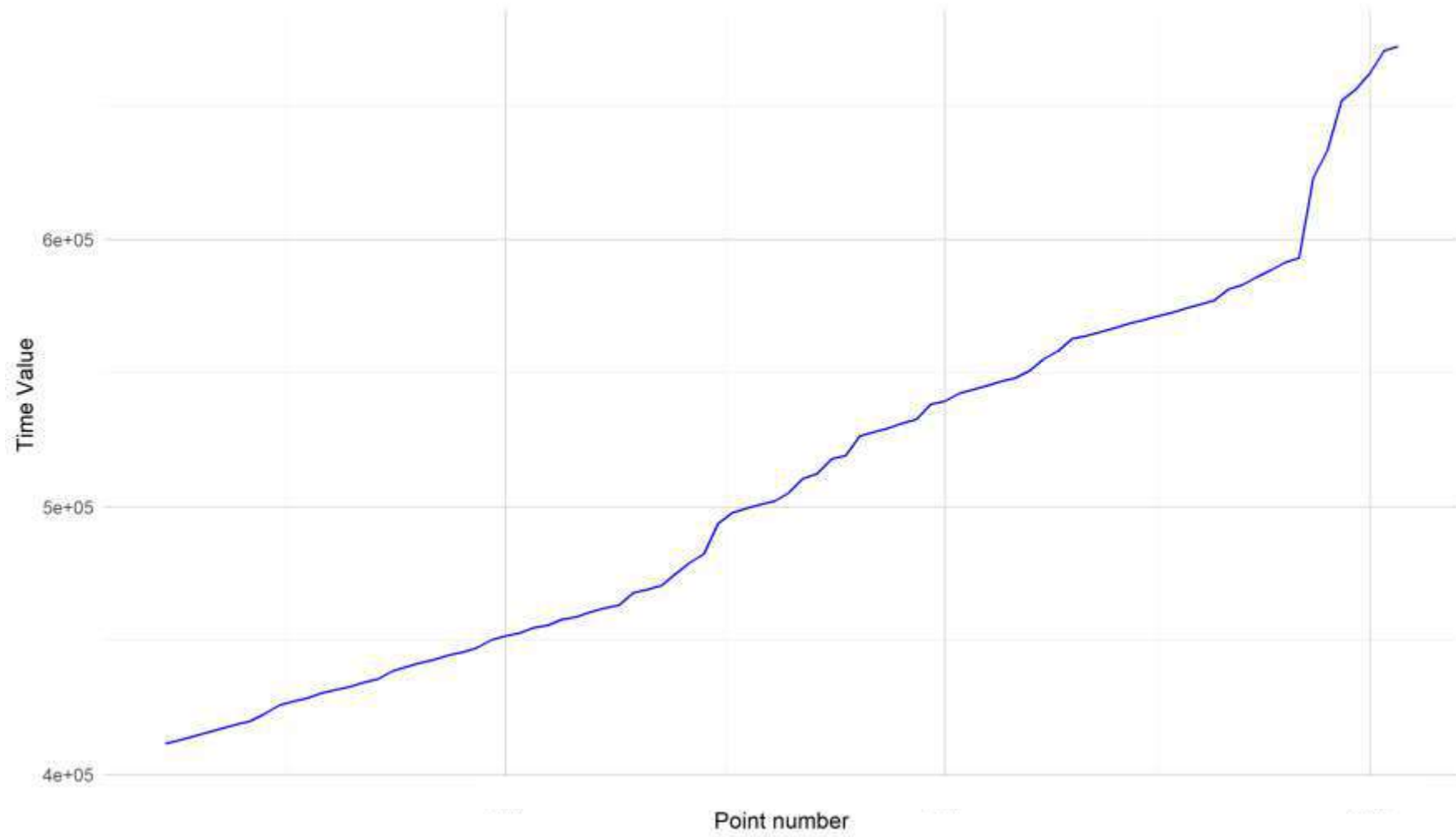
Рожева точка – невідоме (цільове) значення



Порівняння

Метрика	№ файла	Групування		R (6)	R (4)
		Для методу з найменшою обрахованою похибкою	Для середнього значення		
Частка (%) первинних прогнозів (один на точку), з похибкою в межах 5%	1	2.15	0.7	2.84	3.55
	2	5.31	4.83	7.35	6.86
	3	11.8	0.45	18.06	21.3
	4	5.48	7.76	12.5	16.67
	5	11.42	6.85	15.74	15.28
	6	2.69	4.84	3.83	4.92
	7	7.76	10.96	15.74	17.59
	8	8.09	15.6	17.65	16.47
Середня частка (%)		14.05	12.93	24	26.38

Метрика	№ файла	Групування (6)	Групування (4)	R (6)	R (4)
Частка (%) вторинних прогнозів (x на точку), де хоча б один з похибкою в межах 5%	1	6.47	4.93	12.77	9.93
	2	18.35	16.43	22.06	18.14
	3	39.27	33.33	47.22	33.33
	4	26.02	20.09	31.94	24.07
	5	26.94	23.29	36.11	25.93
	6	16.13	15.59	16.94	10.38
	7	27.40	24.2	34.72	20.83
	8	20.81	19.07	41.76	21.18
Середня частка (%)		46.66	40.32	61.73	41.79
Загальний час виконання		<1 хв		~131 хв	>8 хв



Порівняння

Метрика	№ файла	Групування		R (6)	R (4)
		Для значення з найменшою обрахованою похибкою	Для середнього значення		
Частка (%) первинних прогнозів (один на точку), з похибкою в межах 200 хв	1	57.41	14.82	5.88	52.94
	2	56.52	0	5	0
	3	10.98	0	11.39	5.06
	4	0	0	20	10
	5	0	0	9.09	9.09
	6	25.15	1.8	17.68	33.54
	7	9.37	0	0	0
Середня частка (%)		13.77	1.55	6.62	12.89

Метрика	№ файла	Групування (6)	Групування (4)	R (6)	R (4)
Частка (%) вторинних прогнозів (x на точку), де хоча б один з похибкою в межах 200 хв	1	81.48	81.48	84.31	82.35
	2	60.87	60.87	35	30
	3	31.71	30.49	36.71	15.19
	4	0	0	50	40
	5	0	0	63.64	63.64
	6	61.68	55.1	59.15	45.12
	7	12.5	12.5	13.79	6.9
Середня частка (%)		26.93	25.23	28.41	21.81
Загальний час виконання		~6.5 с		~25 хв	<2хв

Висновки

- Розроблено прототип застосунку із підтримкою низки типів даних
- Реалізовано механізми автоматизованого прогнозування в реальному часі для автозаповнень та короткострокових прогнозів

Можливі покращення

- Довгострокове прогнозування
- Додаткові налаштування прогнозів

Дякую за увагу!