

*Білич А. О., студент 2 курсу спеціальності
122 «Комп'ютерні науки» СО Магістр
Нескородєва Т. В., к.т.н, доцент, завідувач
кафедри комп'ютерних наук та
інформаційних технологій*

ДОСЛІДЖЕННЯ ТА НАЛАШТУВАННЯ ДАНИХ ДЛЯ ПРОЕКТУВАННЯ ТА РЕАЛІЗАЦІЇ СИСТЕМИ ПРОГНОЗУВАННЯ ПОГОДИ

Донецький національний університет імені Василя Стуса, м. Вінниця

Системи для прогнозування погоди є досить важливим елементом у повсякденному житті людей. Прогнозування погоди є досить важливим чинником у будь-якій діяльності людей, наприклад: аграрна, авіа перевізна, метео-дослідження тощо. Зазвичай побудова систем з прогнозування погоди є досить складним процесом, який включає в себе побудову моделі та її практичну реалізацію. Основним та важливим елементом у побудові система такого роду є вибір методів та моделей за яким буде побудована система, а також вибір початкових даних – дата-сетів для дослідження початкових даних та налаштування даних у тестові та тренувальні набори даних для подальшого аналізу та даних та створення прогнозу погоди на основі цих даних.

В основі мого дослідження були вибрані певні регресійні моделі на основі статистичного аналізу даних для створення системи такого роду, такі як: проста лінійна регресія, множино-лінійна регресія та метод часових рядів. В результаті цього були спроектовані первинні моделі прогнозу погоди, які були представлені у вигляді таких простих моделей, а саме : ранкова модель, добова модель, вечірня модель, вечірня модель другого порядку, показані на рисунку 1:

```
mod9am_c1_fit: RainTomorrow ~ Cloud9am + Humidity9am + Pressure9am + Temp9am
mod3pm_c1_fit: RainTomorrow ~ Cloud3pm + Humidity3pm + Pressure3pm + Temp3pm
mod_ev_c2_fit: RainTomorrow ~ Cloud3pm + Humidity3pm + Pressure3pm + Temp3pm
mod_ev_c3_fit: RainTomorrow ~ Pressure3pm + Sunshine
```

Рисунок 1 – Первинні моделі прогнозу погоди.

Було прийнято рішення, щодо наповнення дата-сету за допомогою відкритого погодного API. Завданням програми є наповнення даних і в подальшому створення аналізу даних на основі якого буде створений прогноз. Отже після отримання даних у дата-сет, спроектований додаток проводить певний аналіз даних для створення тестувальних та тренувальних наборів даних для того, щоб навчити загальну модель та перевірити як вона працює. В процесі аналізу даних програма проводить Turing та ROC аналіз на основі якого проводиться створення наборів даних та створення підмножини змін.

Результати ROC аналізу програма має здатність вивести в звітах, вони мають такий вигляд, який показано на рисунках 2,3 та 4:

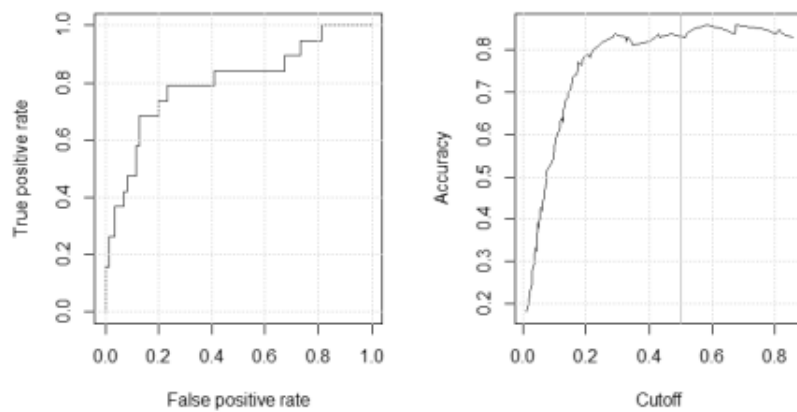


Рисунок 2 – ROC крива для ранкової моделі прогнозу погоди

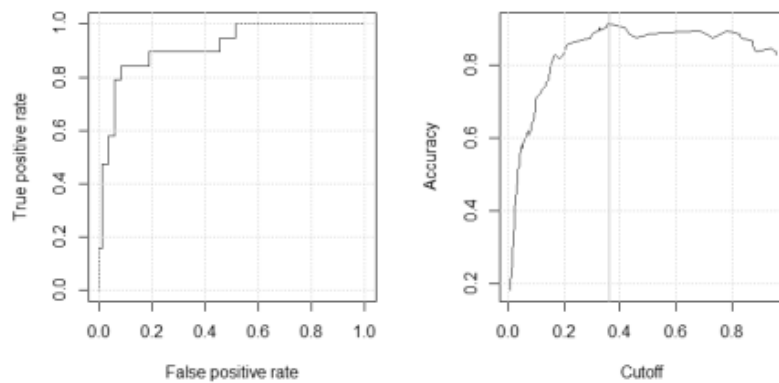


Рисунок 3 – ROC крива для добової моделі прогнозу погоди

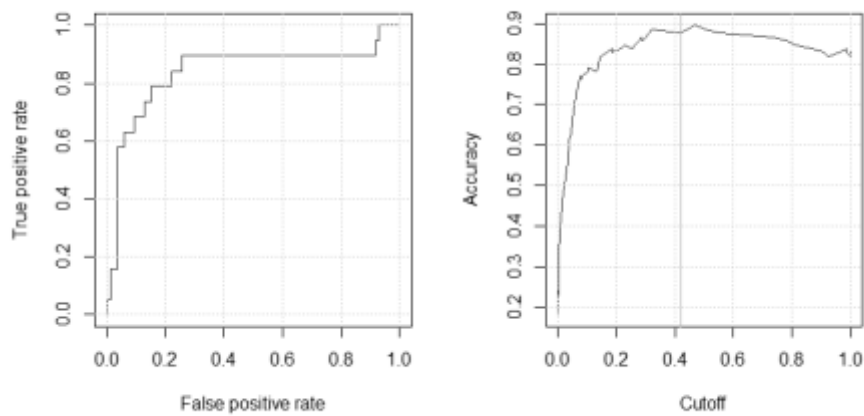


Рисунок 4 - ROC крива для вечірньої моделі прогнозу погоди

В результаті дослідження даних та створення аналізу даних цих типів ми отримаємо тренувальні та тестові набори даних, які в свою чергу дають можливість навчити модель прогнозу та можливість її тестування. Приклад тестових та тренувальних наборів даних представлені на рисунку 5 та 6:

	cutoff	accuracy	specificity
1	0.20	0.7822581	0.73333333
2	0.22	0.7862903	0.71111111
3	0.24	0.7943548	0.68888889
4	0.26	0.8064516	0.64444444
5	0.28	0.8104839	0.64444444
6	0.30	0.8145161	0.60000000
7	0.32	0.8185484	0.57777778
8	0.34	0.8185484	0.51111111
9	0.36	0.8104839	0.46666667
10	0.38	0.8266129	0.46666667
11	0.40	0.8266129	0.40000000
12	0.42	0.8306452	0.40000000
13	0.44	0.8266129	0.37777778
14	0.46	0.8346774	0.35555556
15	0.48	0.8467742	0.33333333
16	0.50	0.8508065	0.31111111
17	0.52	0.8427419	0.26666667
18	0.54	0.8467742	0.26666667
19	0.56	0.8346774	0.20000000
20	0.58	0.8387097	0.20000000
21	0.60	0.8387097	0.20000000
22	0.62	0.8346774	0.17777778
23	0.64	0.8387097	0.17777778
24	0.66	0.8346774	0.15555556
25	0.68	0.8387097	0.15555556
26	0.70	0.8387097	0.13333333
27	0.72	0.8387097	0.13333333
28	0.74	0.8346774	0.11111111
29	0.76	0.8266129	0.06666667
30	0.78	0.8266129	0.06666667
31	0.80	0.8306452	0.06666667

Рисунок 5 – Тренувальний набір даних для моделі

	cutoff	accuracy	specificity
1	0.20	0.8064516	0.7777778
2	0.22	0.8185484	0.7555556
3	0.24	0.8225806	0.7333333
4	0.26	0.8306452	0.6888889
5	0.28	0.8467742	0.6666667
6	0.30	0.8467742	0.6444444
7	0.32	0.8427419	0.6222222
8	0.34	0.8669355	0.6222222
9	0.36	0.8709677	0.6222222
10	0.38	0.8629032	0.5777778
11	0.40	0.8669355	0.5777778
12	0.42	0.8669355	0.5555556
13	0.44	0.8548387	0.4666667
14	0.46	0.8548387	0.4444444
15	0.48	0.8588710	0.4444444
16	0.50	0.8669355	0.4444444
17	0.52	0.8629032	0.4222222
18	0.54	0.8669355	0.4222222
19	0.56	0.8669355	0.3777778
20	0.58	0.8669355	0.3777778
21	0.60	0.8588710	0.3333333
22	0.62	0.8548387	0.3111111
23	0.64	0.8508065	0.2888889
24	0.66	0.8467742	0.2666667
25	0.68	0.8387097	0.2222222
26	0.70	0.8387097	0.2222222
27	0.72	0.8346774	0.2000000
28	0.74	0.8387097	0.1777778
29	0.76	0.8346774	0.1555556
30	0.78	0.8346774	0.1555556
31	0.80	0.8306452	0.1111111

Рисунок 6 – Тестовий набір даних для моделі

В результаті модель системи для прогнозування має такий вигляд, показаний на рисунку 7.

Weather Forecast report block scheme

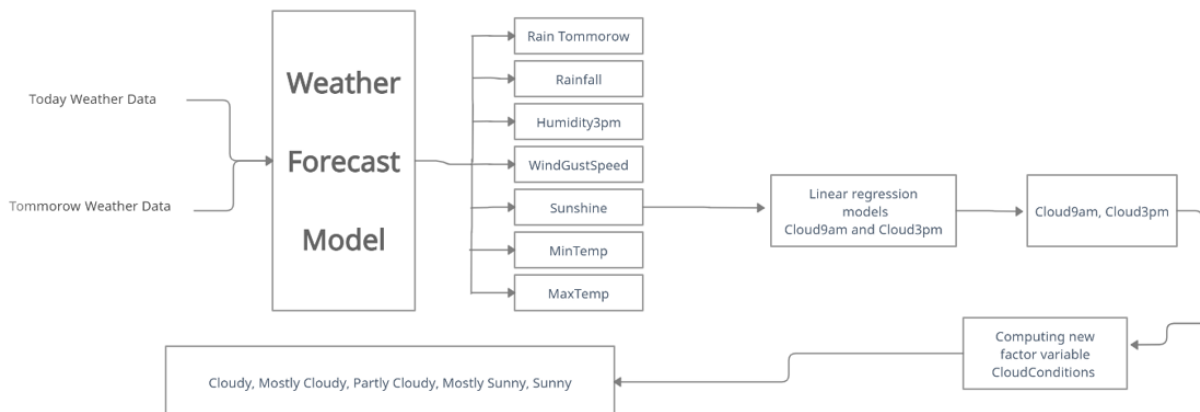


Рисунок 7 – Загальна модель прогнозу погоди

Висновок. У результаті аналізу даних, вдалось створити додаток з прогнозування погоди на основі статистичного аналізу даних. У процесі аналізування та налаштування даних були методи ROC та Turing аналізу за допомогою яких був створений тестувальний та тренувальний набір даних для кожної первинної моделі, яка використовуються у загальній системі. Дата-сет

представлений генерованими даними які наповнюються з відкритого погодного API, після чого за допомогою цього дата-сету виконується аналіз даних, що дає можливість створення прогнозу погоди на наступні дні.

Список літературних джерел.

1. Джейсон У. Осборн., Найкращі практики з логістичної регресії. - SAGE Publications, Inc, 2014 – 500 с.
2. Рігбі Р., Хелер Г., Розподіли для моделювання розташування, масштабу та форми: використання GAMLSS у R (Chapman & Hall/CRC The R Series) 1 -е видання. - Chapman and Hall/CRC; 1st edition, 2019 – 588 с.
3. Ануша Ілюккумбура., Вступ до регресійного аналізу. - Independently published, 2020 – 121 с.
4. Алан Агрешті., Основи лінійної та узагальненої лінійних моделей. - Wiley; 1st edition, 2015 – 469 с.

УДК 004.8

*Васильченко Д. Н., студент 3 курсу
спеціальності І22 «Комп'ютерні науки»
Потапова Н.А., к.е.н., доцент, доцент
кафедри інформаційних технологій*

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ

Донецький національний університет імені В. Стуса, м. Вінниця

Існує дві основних складових у визначенні поняття математичного моделювання технологічних процесів: математична модель і технологічний процес:

модель – проект, що імітує властивості певної сутності у тому чи іншому зручному для дослідження вигляді .

математична модель – сукупність математичних співвідношень, які описують певні явища та процеси.

Виробничий процес це сукупність усіх дій людей і знарядь праці, що застосовуються для виготовлення чи ремонту виробів, які випускаються. Технологічний процес є його частиною, що відповідає за сукупність дій по зміні стану предмета праці. Отже, математичне моделювання технологічних процесів – це процес створення математичної моделі, об'єктом якої є технологічний процес чи його складові.

Типи задач математичного моделювання технологічних процесів:

- прогнозування з різним рівнем достовірності кінцевих чи проміжкових характеристичних показників технологічного процесу, що неможливо або занадто недоцільно визначати в реальних умовах;