

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ
СІКОРСЬКОГО»

Факультет інформатики та обчислювальної техніки

Кафедра інформатики та програмної інженерії

Практикум №5

з курсу «Аналіз даних в інформаційних системах»

на тему: «Регресійні моделі»

Викладач:
Ліхоузова Т.А.

Виконав:
студент 2 курсу
групи ІП-11 Сідак
Кирил з ФІОТ

Київ-2023

ЗМІСТ

1. ЗАВДАННЯ	3
2. ОСНОВНЕ ЗАВДАННЯ	4
3. ДОДАТКОВЕ ЗАВДАННЯ	5
4. ВИСНОВОК.....	6

1. ЗАВДАННЯ

Основне завдання

Завантажити [дані про якість червоного вина](#)

1. Дослідити дані, підготувати їх для побудови регресійної моделі
2. Розділити дані на навчальну та тестову вибірки
3. Побудувати декілька регресійних моделей для прогнозу якості вина (12 - quality). Використати лінійну регресію та поліноміальну регресію обраного вами виду
4. Використовуючи тестову вибірку, з'ясувати яка з моделей краща

Додаткове завдання

Завантажити дані файлу Data4.csv

1. Дослідити дані, сказати чи є мультиколінеарність, побудувати діаграми розсіювання
2. Побудувати декілька регресійних моделей (використати лінійну регресію та поліноміальну регресію обраного вами виду)
3. Використовуючи тестову вибірку з файлу Data4t.csv, з'ясувати яка з моделей краща

2. ОСНОВНЕ ЗАВДАННЯ

Звіт по основному завданню було оформлено за допомогою блокноту з кодом.

3. ДОДАТКОВЕ ЗАВДАННЯ

Звіт по додатковому завданню було оформлено за допомогою блокноту з кодом.

4. ВИСНОВОК

Отже, при виконанні даної лабораторної було досліджено два різні набори даних, побудовано моделі лінійної та поліноміальних регресій різних порядків для кожного з цих наборів, аналітично та графічно обрано оптимальні моделі (для обох наборів – це лінійна регресія). Крім того, у рамках додаткового завдання було досліджено дані на мультиколінеарність за допомогою кореляційної матриці й фактору дисперсії інфляції та побудовано діаграми розсіювання.