Міністерство освіти і науки України

Львівський національний університет імені Івана Франка

Факультет електроніки та комп’ютерних технологій

Кафедра опто-електроніки

Звіт

Про виконання лабораторної роботи №3

З курсу «Системи машинного навчання»

Регуляризація

Виконав:

Студент групи ФеС-32

Бойко Кирило

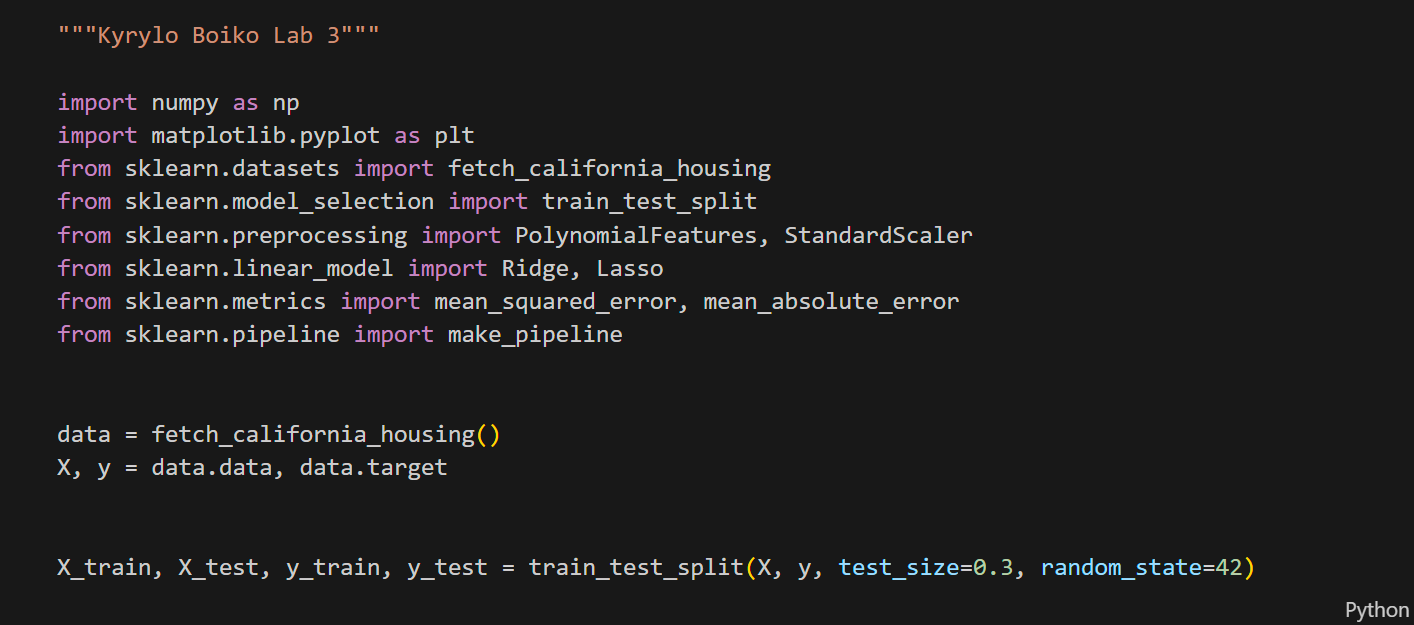
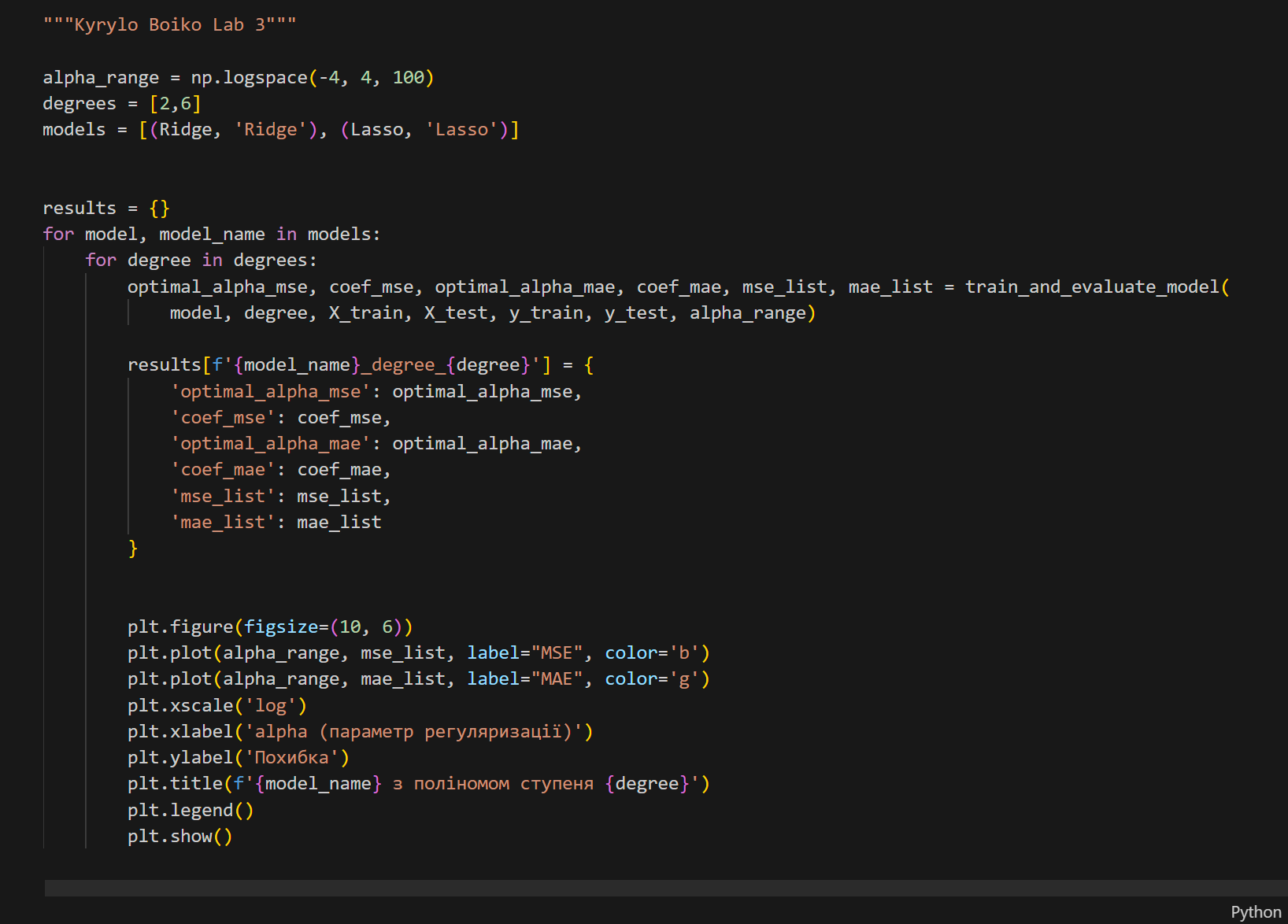
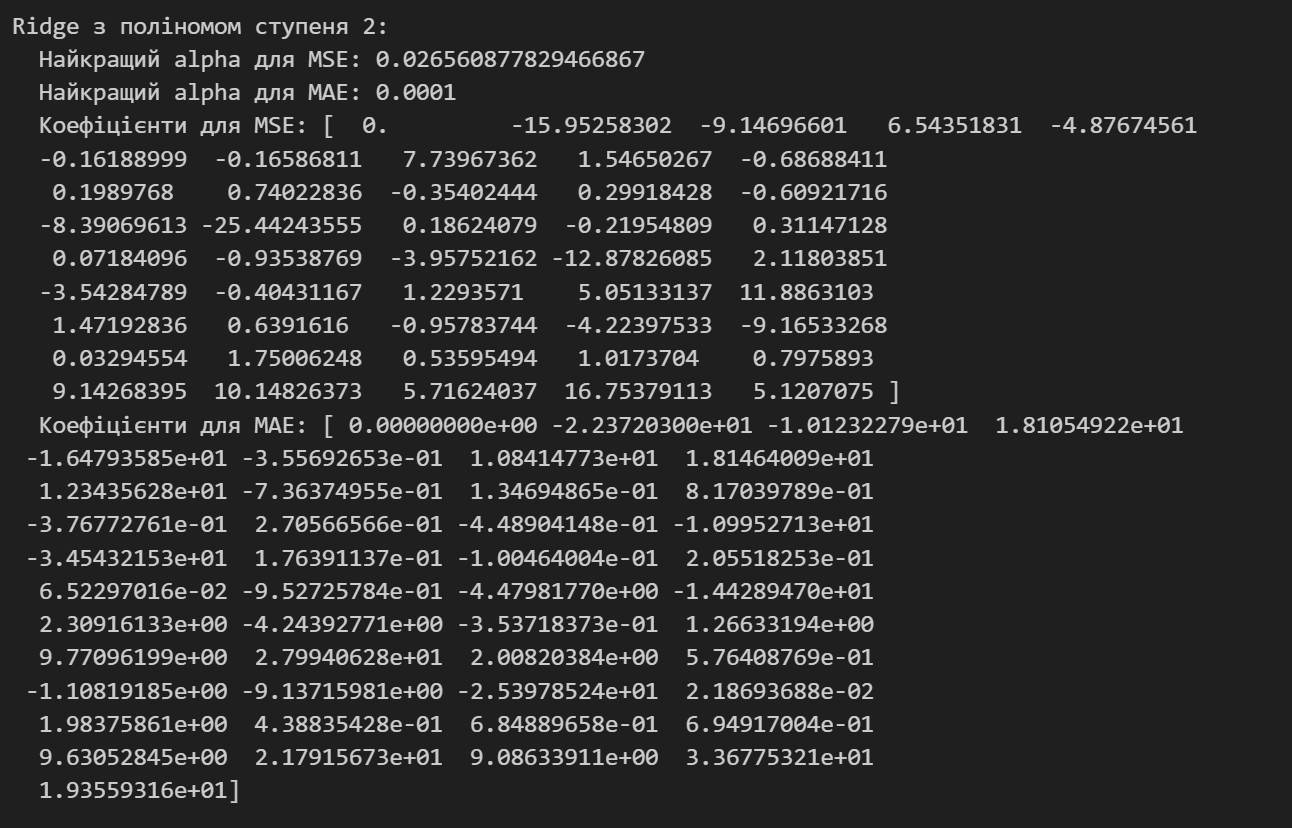
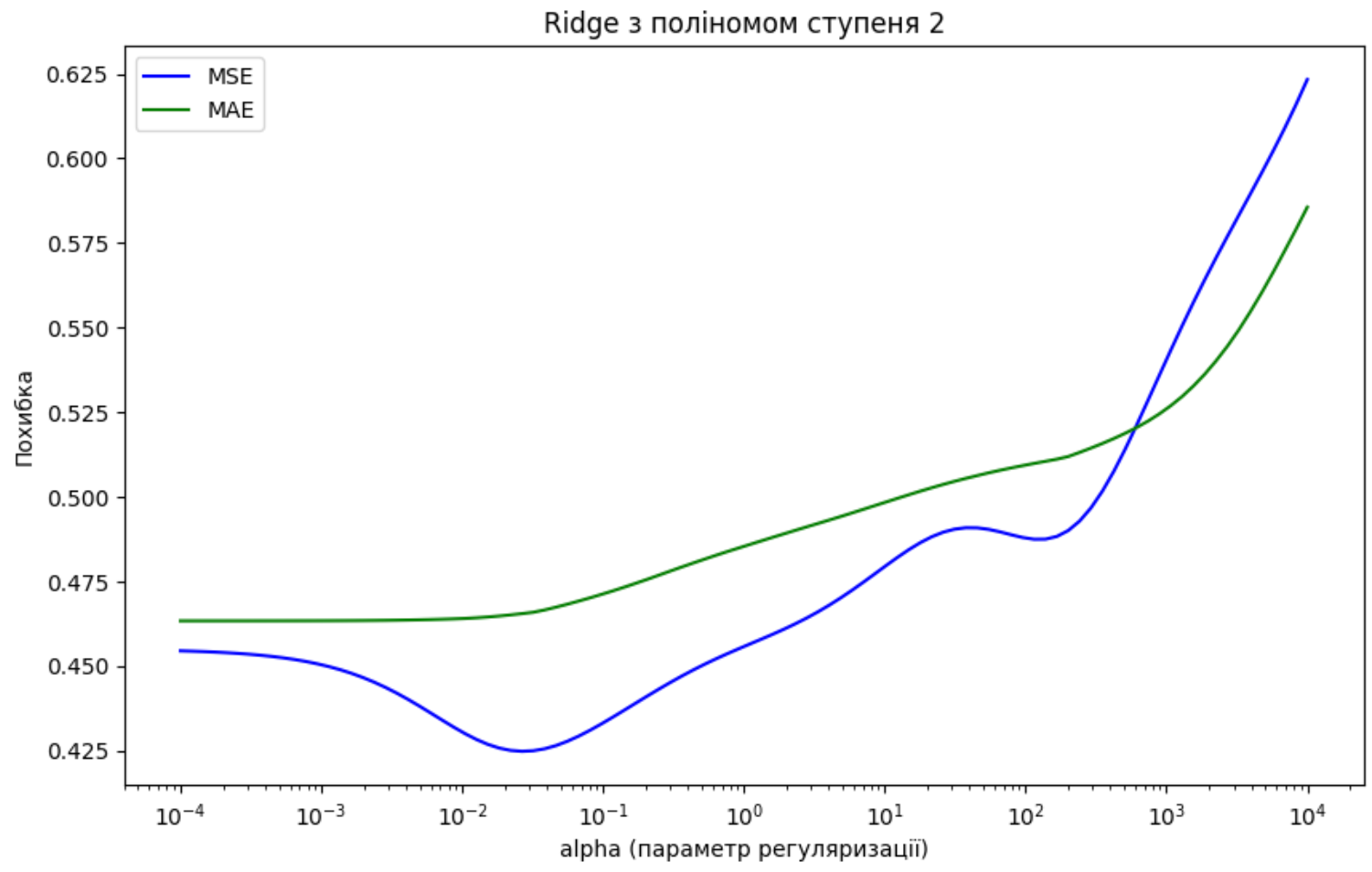
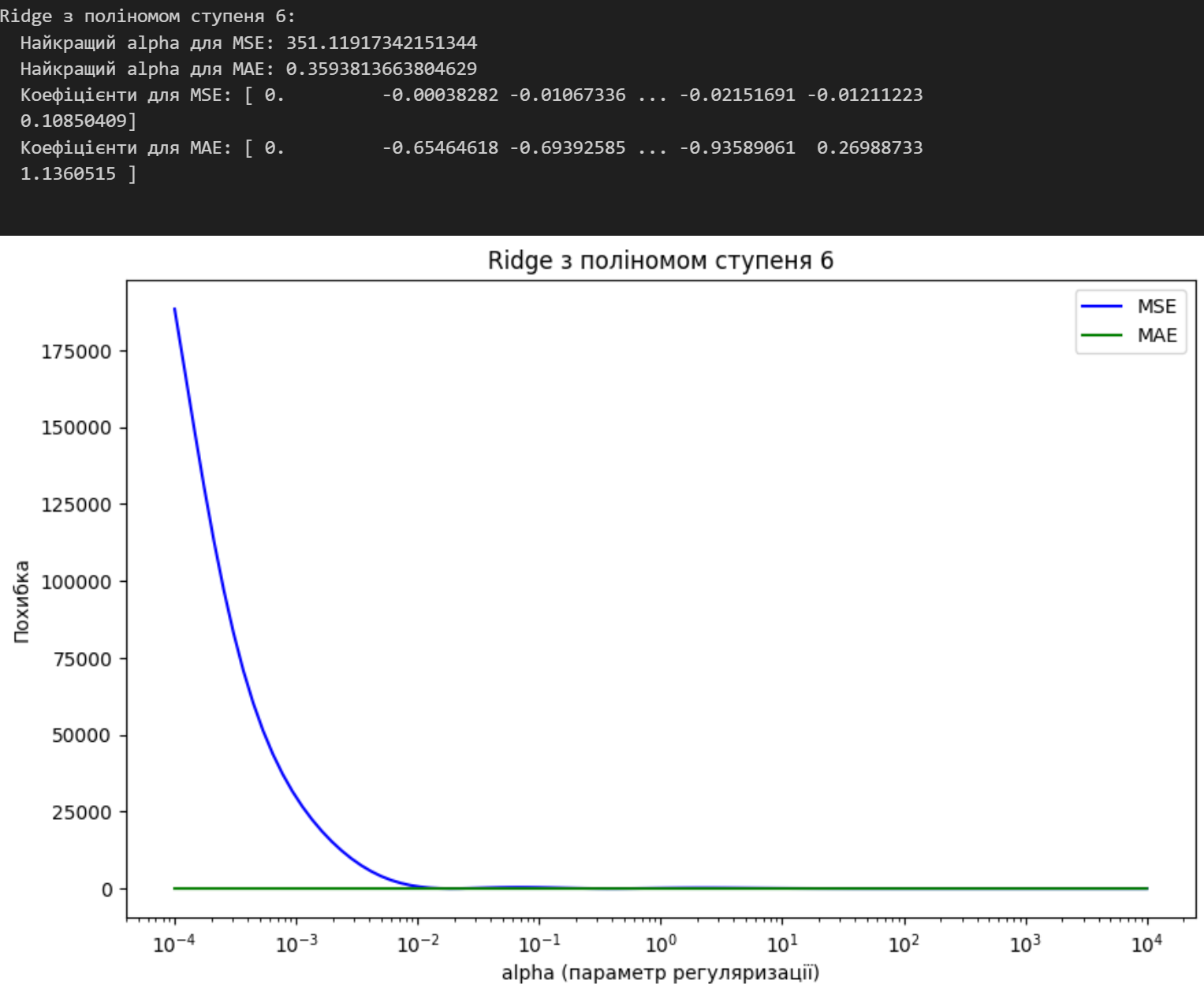
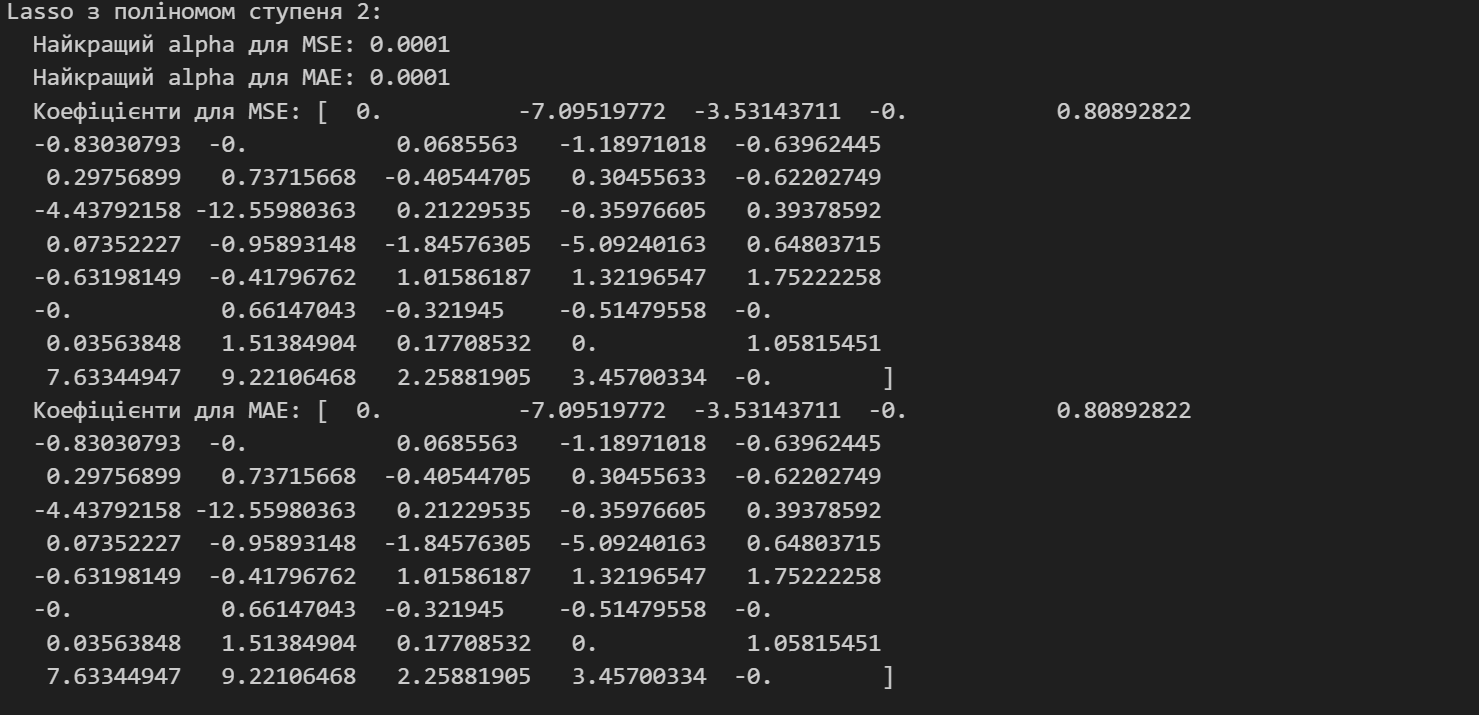
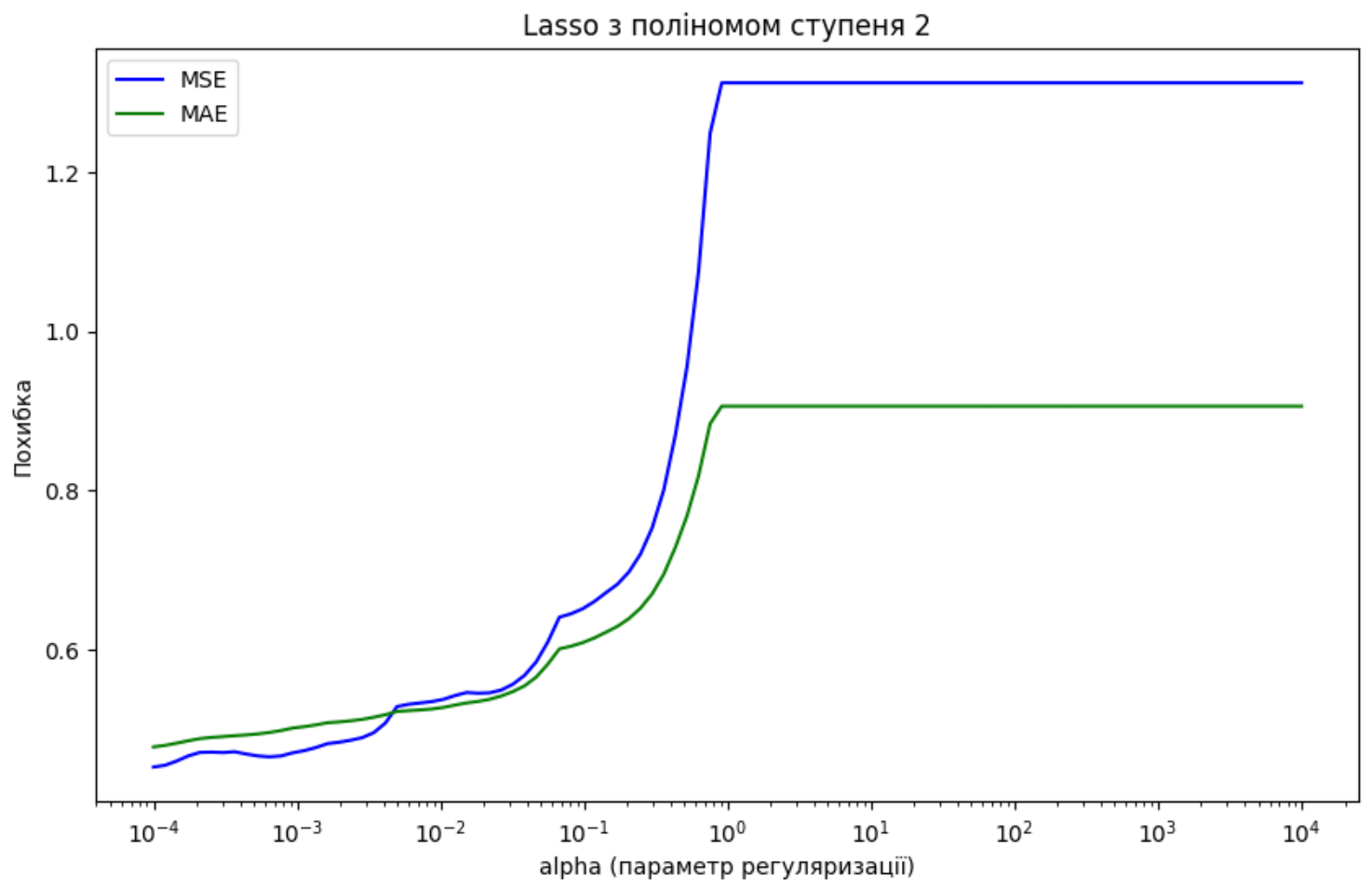
Перевірив:

Доцент Колич І.І.

Львів 2024

**Мета:** засвоїти основи регресійного аналізу з використанням регуляризації.

**Хід роботи:**

1. Завантаження готових наборів даних з Scikit-learn:
   * Завантажити набір sklearn.datasets.fetch\_california\_housing
2. Поділ даних на тренувальну та тестову вибірки:
   * Поділити дані на тренувальний 70% та тестовий 30% набори
3. Створити наступні регресійні моделі
   * Поліноміальна регресія з ступенем поліному 2 та з регуляризацією: Ridge
   * Поліноміальна регресія з ступенем поліному 20 та з регуляризацією: Ridge
   * Поліноміальна регресія з ступенем поліному 2 та з регуляризацією: Lasso
   * Поліноміальна регресія з ступенем поліному 20 та з регуляризацією: Lasso
4. Навчання та оцінка моделей:
   * Для кожної створеної моделі виконати навчання виконати наступні операції
     1. Знайти оптимальний параметр регуляризації, який забезпечує найменшу середньоквадратичну помилка (похибку)
     2. Вивести коефіцієнти моделі
     3. Знайти оптимальний параметр регуляризації, який забезпечує найменшу середню абсолютну помилка (похибку)
     4. Вивести коефіцієнти моделі та порівняти з попередніми
   * Оцінити продуктивності моделі на тестових даних.
   * Результати оцінки похибок передбачення від коефіцієнта параметра регуляризації зобразити на графіках (середньоквадратична помилка та середня абсолютна помилка)
5. Оформлення звіту:
6. Завантаження бібліотек та поділ тестових даних  
   
7. Функція для тренування та оцінки моделі підбирає значення альфа для MAE/MSE. Усі дані для оцінювання зберігаються в масивах, а наприкінці результати зберігаються у словнику result  
   
8. Додавання даних оцінювання до словника `result` та вивід графіків для моделей, які використовують регуляризацію Ridge (L2) та Lasso (L1) з поліномами степеня 2 та 6.  
   
9. Оцінка моделі з поліномом ступеня 2 за допомогою регуляризації Ridge показує графік, що ілюструє залежність похибки (MSE/MAE) від значення alpha. Високі значення alpha можуть призвести до недостатньої адаптації моделі, тоді як низькі значення можуть викликати перенасичення даними, що, в свою чергу, збільшує похибку  
     
     
   
10. Оцінка моделі поліному 6 ступеня з використанням Ridge регуляризації  
    
11. Оцінка моделі поліному 2 ступеня з використанням Lasso регуляризації  
      
    

**Висновок:** Виконавши дану лабораторну роботу, я ознайомився з методами регуляризації L1/L2 (Lasso та Ridge) і провів експерименти з їх застосуванням. Аналіз отриманих результатів дозволив мені зробити висновок про те, що значення гіперпараметра alpha суттєво впливає на якість моделі, визначаючи ступінь її складності. Низькі значення alpha призводять до перенавчання моделі, тоді як високі – до її недонавчання, що проявляється у збільшенні значень MSE та MAE.