

ОСНОВИ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ, НЕЙРОННИХ МЕРЕЖ та ГЛИБОКОГО НАВЧАННЯ

Модуль 2. Навчання з вчителем

Лекція 2.3.

**Приклад вирішення задачі лінійної регресії за
допомогою Scikit-learn**

Регресія

Формально. Маємо:

Незалежні змінні $\Leftrightarrow X$

Залежна змінна $\Leftrightarrow Y$

Відомі пари $\Leftrightarrow (x_i, \hat{y}_i), i = 1, 2, \dots, m$

Регресійна функція (модель) $\Leftrightarrow y = f(W, X)$, де W – множина невідомих параметрів w_0, w_1, \dots, w_K .

Необхідно \Leftrightarrow оцінити унікальні значення W , які в деякому розумінні підходять **найкраще**, а регресійна модель у застосування до даних може розглядатися як перевизначена система для W . Інакше, знайти такі значення w_0, w_1, \dots, w_K , які мінімізують

$$L = \sum_1^m \varepsilon_i^2, \text{ де } \varepsilon_i = \hat{y}_i - y_i = \hat{y}_i - f(W, x_i)$$

Лінійна регресія

Регресія

Лінійна регресія

Поліноміальна регресія

....

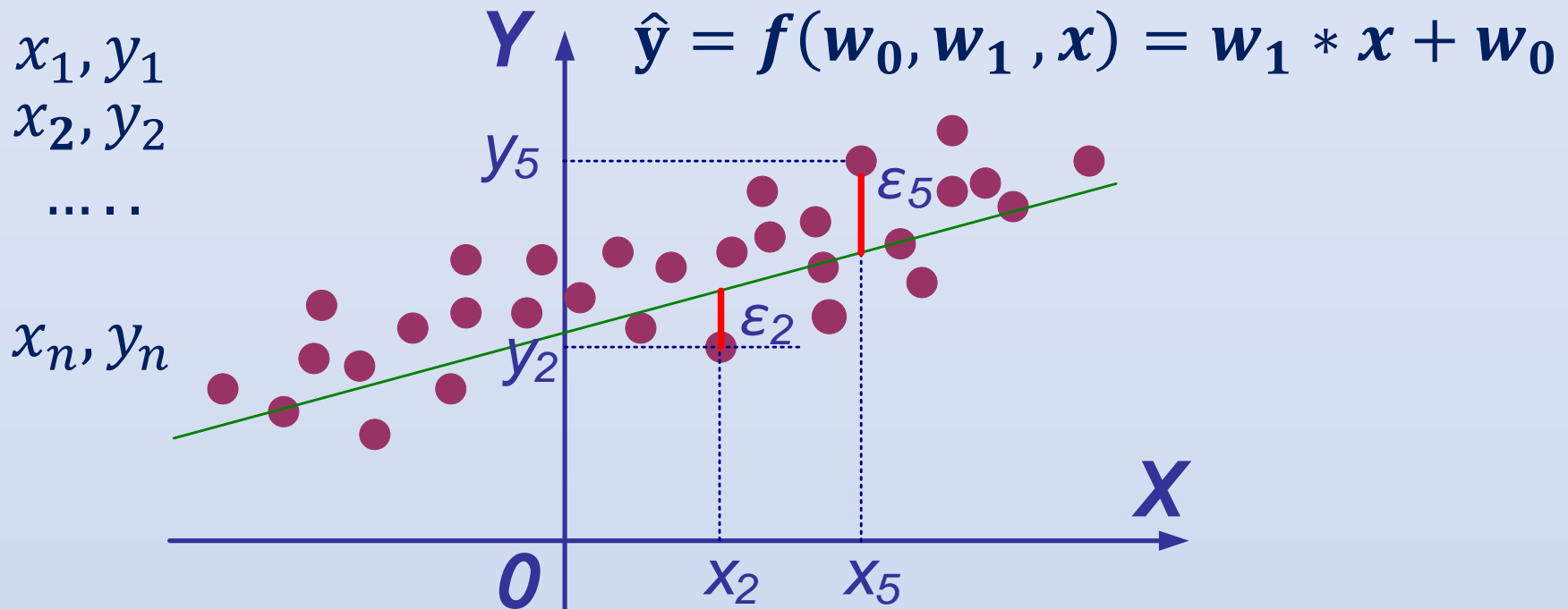
Нелінійна регресія

$$f(W, X) = w_0 + w_1 * X_1 + w_2 * X_2 + \dots$$

Оцінка якості за допомогою квадратичної помилки

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2$$

Парна лінійна регресія (SLR)



$$L(w) = \frac{1}{2n} \sum_{i=1}^n \epsilon_i^2 = \frac{1}{2n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 =$$
$$\frac{1}{2n} \sum_{i=1}^n (\hat{y}_i - w_0 - w_1 * x_i)^2$$

Знайти такі (\bar{w}) , що $L(\bar{w}) = \min_w L(w)$

Множинна лінійна регресія (MLR)

$$X = \begin{bmatrix} x_1 \\ \dots \\ x_m \end{bmatrix}, W = \begin{bmatrix} w_1 \\ \dots \\ w_m \end{bmatrix}$$

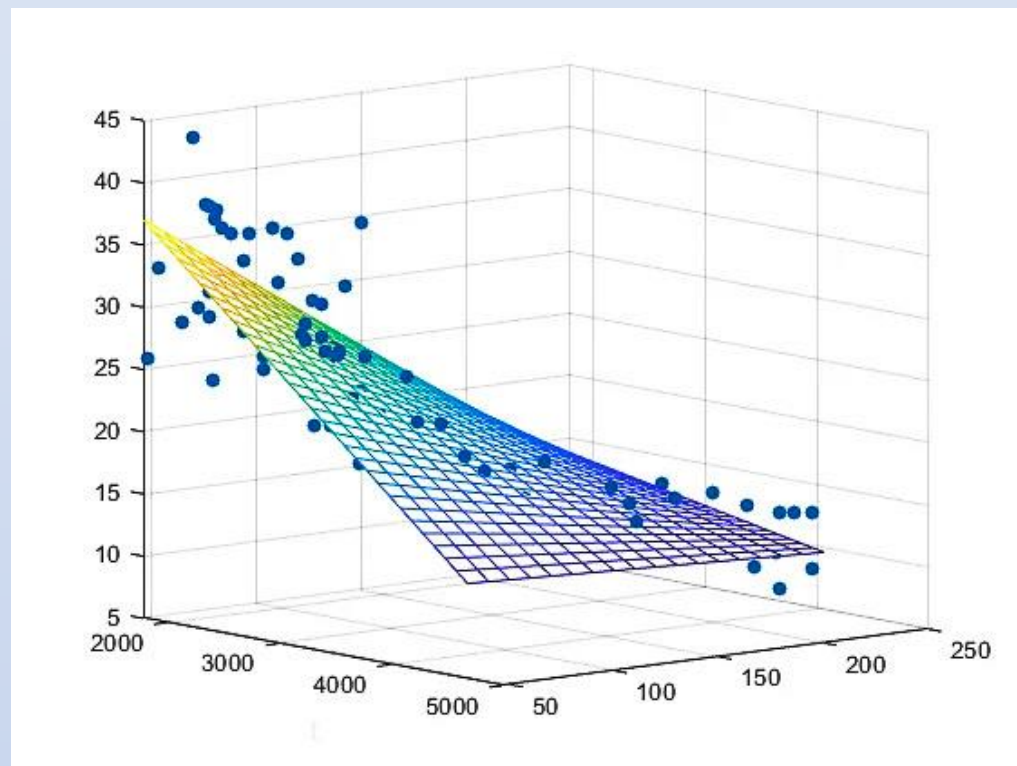
X_1, y_1

X_2, y_2

.....

X_n, y_n

$$\hat{y} = f(w_0, W) = W^T * X + w_0$$



$$L(w) = \sum_1^n \varepsilon_i^2 = (\hat{y}_1 - y_1)^2 + \dots + (\hat{y}_i - y_i)^2 + \dots$$

Знайти такі (\bar{W}) , що $L(\bar{W}) = \min_w L(W)$

Лінійна регресія з scikit-learn

Бібліотека	<i>scikit-learn</i>
Модуль	<i>linear_model</i>
Метод	<i>LinearRegression</i>

LinearRegression(*, *fit_intercept=True, copy_X=True, n_jobs=None, positive=False*)

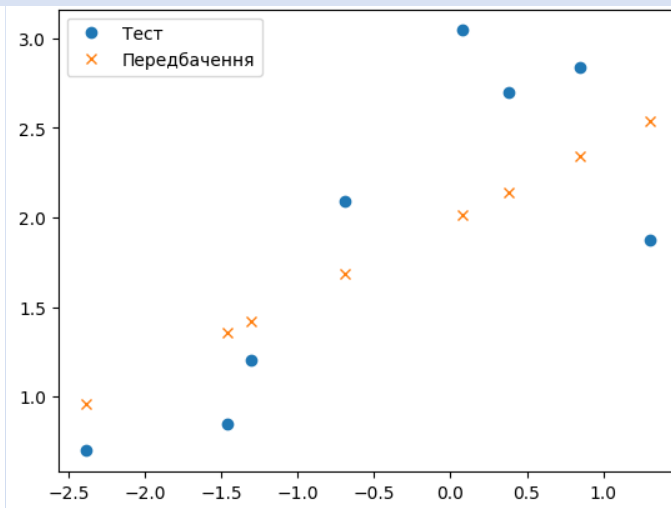
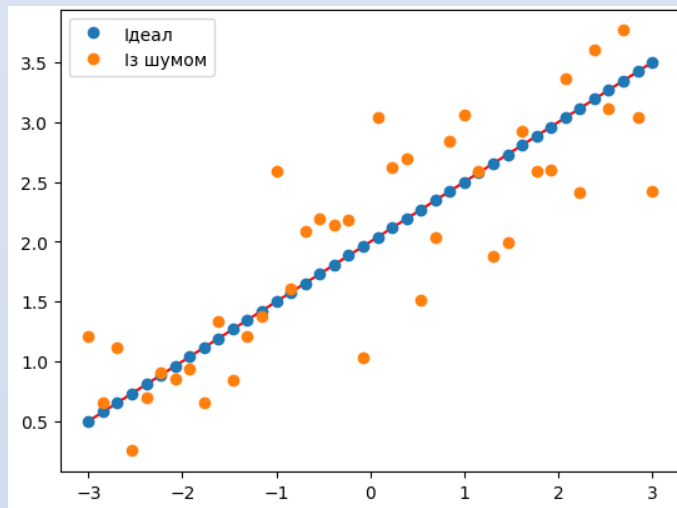
Лінійна регресія методом найменших квадратів.

LinearRegression підходить до лінійної моделі з коефіцієнтами W , щоб мінімізувати залишкову суму квадратів між спостережуваними цілями в наборі даних і цілями, передбаченими лінійною апроксимацією.

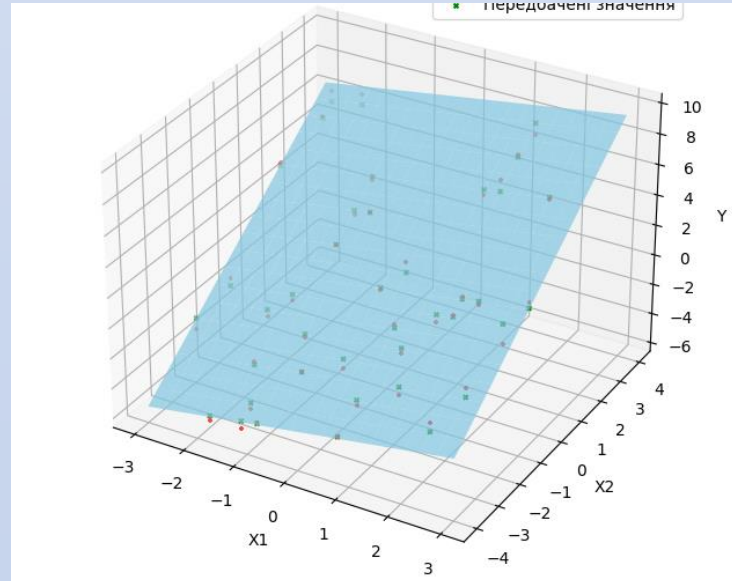
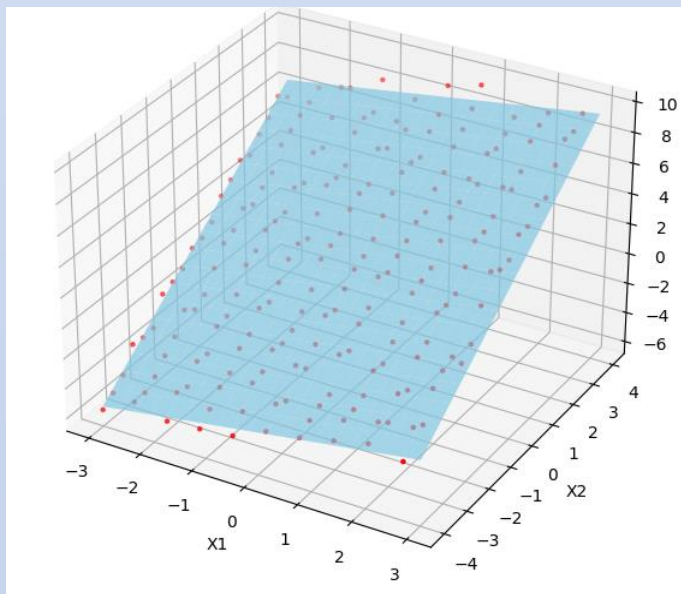
Результати

Приклад - lec_02_03_Example_1

MSE = 0.32



MSE = 0.23



Контрольні запитання

- Надайте загальну постановку задачі регресії.
- Вкажіть особливості задачі парної лінійної регресії (SLR).
- Вкажіть особливості задачі множинної лінійної регресії (MLR).
- Поясніть сутність методу вирішення задачі лінійної регресії з мінімізацією MSE.
- Наведіть приклад застосування *LinearRegression* бібліотеки *scikit-learn* для вирішення задачі лінійної регресії.

Рекомендована ЛІТЕРАТУРА

- **Глибинне навчання:** Навчальний посібник / Уклад.: В.В. Литвин, Р.М. Пелещак, В.А. Висоцька В.А. – Львів: Видавництво Львівської політехніки, 2021. – 264 с.
- Тимощук П. В., Лобур М. В. **Principles of Artificial Neural Networks and Their Applications: Принципи штучних нейронних мереж та їх застосування:** Навчальний посібник. – Львів : Видавництво Львівської політехніки, 2020. – 292 с.
- Morales M. **Grokking Deep Reinforcement Learning.** – Manning, 2020. – 907 с.
- Trask Andrew W. **Grokking Deep Learning.** – Manning, 2019. – 336 с.

The END

Модуль 2. Лекція 03.