ОСНОВИ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ, НЕЙРОННИХ МЕРЕЖ та ГЛИБОКОГО НАВЧАННЯ

Модуль 2. Навчання з вчителем

Лекція 2.1.
Регресія. Загальні визначення.

Класичний AI / Класичний ML

Машинне навчання

Навчання з вчителем
- **Регрессия**
- **Класифікація**

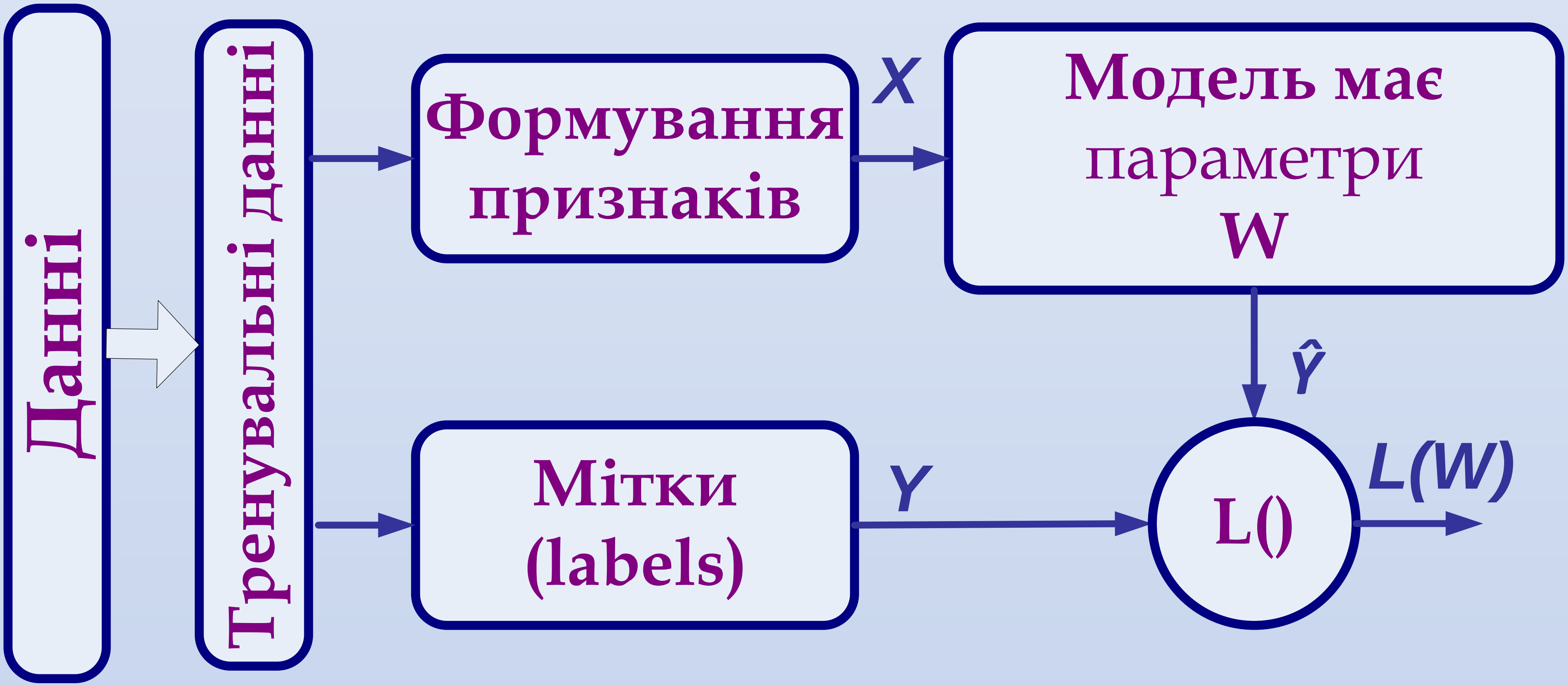
Навчання без вчителя
- **Кластерізація**
- **Зменшення розмірності**

Навчання з підкріпленням
- **Виживання у незнайомій обстановці**

Навчання з вчителем: є набір прикладів, до кожного прикладу є правильна відповідь.

Задача – навчитися по прикладах надавати правильну відповідь на питання, задане вчителем.

Загальний процес ML з вчителем



Класичний AI / Класичний ML

Машинне навчання

- **Навчання з вчителем**
  - Регрессия
  - Класифікація
- **Навчання без вчителя**
  - Кластерізація
  - Зменшення розмірності
- **Навчання з підкріпленням**
  - Виживання у незнайомій обстановці

Регресія

Регресія (**лат. regressus**) – «Повернення» (Повернення до середнього) (*деградація*)

Регресія ⇔ ***Інтерполяція, Апроксимація, Екстраполяція***

Математично це наближення функцій до визначених даних (вхідних точок).

1. **Інтерполяція** - це такий метод, опис вихідних даних, при якому інтерполююча функція проходить через вихідні точки.
2. **Апроксимація** - це такий метод, при якому апроксимуюча функція наближається до вихідних даних, виконуючи якусь умову.
3. **Екстраполяція** – прогноз вихідних даних.

Функція може бути: лінійна, кубічна, поліноміальна, сплайн, і т.д.

Регресія

Формально. Маємо:

Незалежні змінні ⬄ $\boldsymbol{X}$

Залежна змінна ⬄ $\boldsymbol{Y}$

Відомі пари ⬄ $(\boldsymbol{x}_i, \boldsymbol{y}_i),\ i = 1,2, \dots, m$

Регресійна функція (модель) ⬄ $\widehat{\boldsymbol{y}} = \boldsymbol{f}(\boldsymbol{W}, \boldsymbol{X})$, де $\boldsymbol{W}$ – множина невідомих параметрів $w_0, w_1, \dots w_K$.

Необхідно ⬄ оцінити унікальні значення $\boldsymbol{W}$, які в деякому розумінні підходять **найкраще**, а регресійна модель у застосування до даних може розглядатися як перевизначена система для $\boldsymbol{W}$.

Інакше, знайти такі значення $w_0, w_1, \dots w_K$, які мінімізують

$$\boldsymbol{L} = \sum_1^m \boldsymbol{\varepsilon}_i^2\ ,\ \text{де}\ \boldsymbol{\varepsilon}_i = \boldsymbol{y}_i - \widehat{\boldsymbol{y}}_i = \boldsymbol{y}_i - \boldsymbol{f}(\boldsymbol{W}, \boldsymbol{x}_i)$$

Регресія

- Регрессія
  - Лінійна регресія
  - Поліноміальна регресія
  - Логаріфмічна регресія
  - ....

Лінійна регресія

$$\hat{y}_i = f(W, X) = w_0 * 1 + w_1 * f(x_1) + w_2 * f(x_2) + \cdots$$

x скаляр ⇔ парна лінійна регресія (найпростіша)

Simple Linear Regression (SLR)

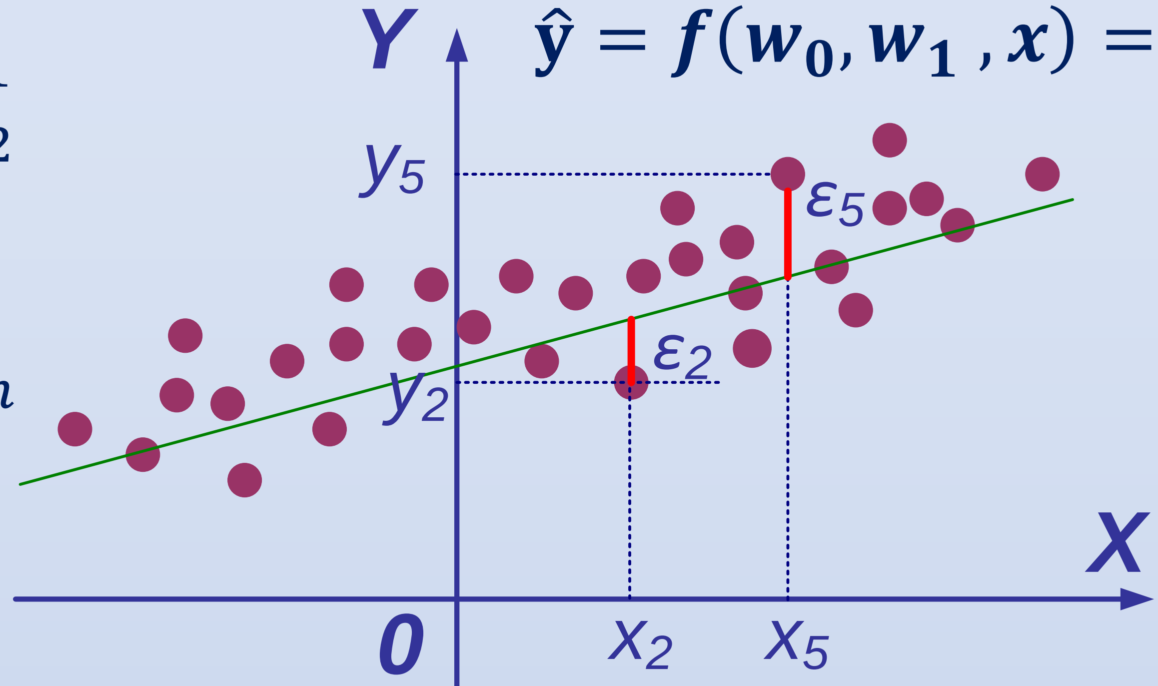
x вектор ⇔ множинна лінійна регресія

Multiple Linear Regression (MLR)

SLR

x_1, y_1
x_2, y_2
$\ldots\ldots$
x_n, y_n

$$\hat{y} = f(w_0, w_1, x) = w_1 * x + w_0$$



$$L(w) = \frac{1}{2n}\sum_{1}^{n}\varepsilon_i^2 = \frac{1}{2n}\sum_{1}^{n}(\hat{y}_i - y_i)^2 =$$

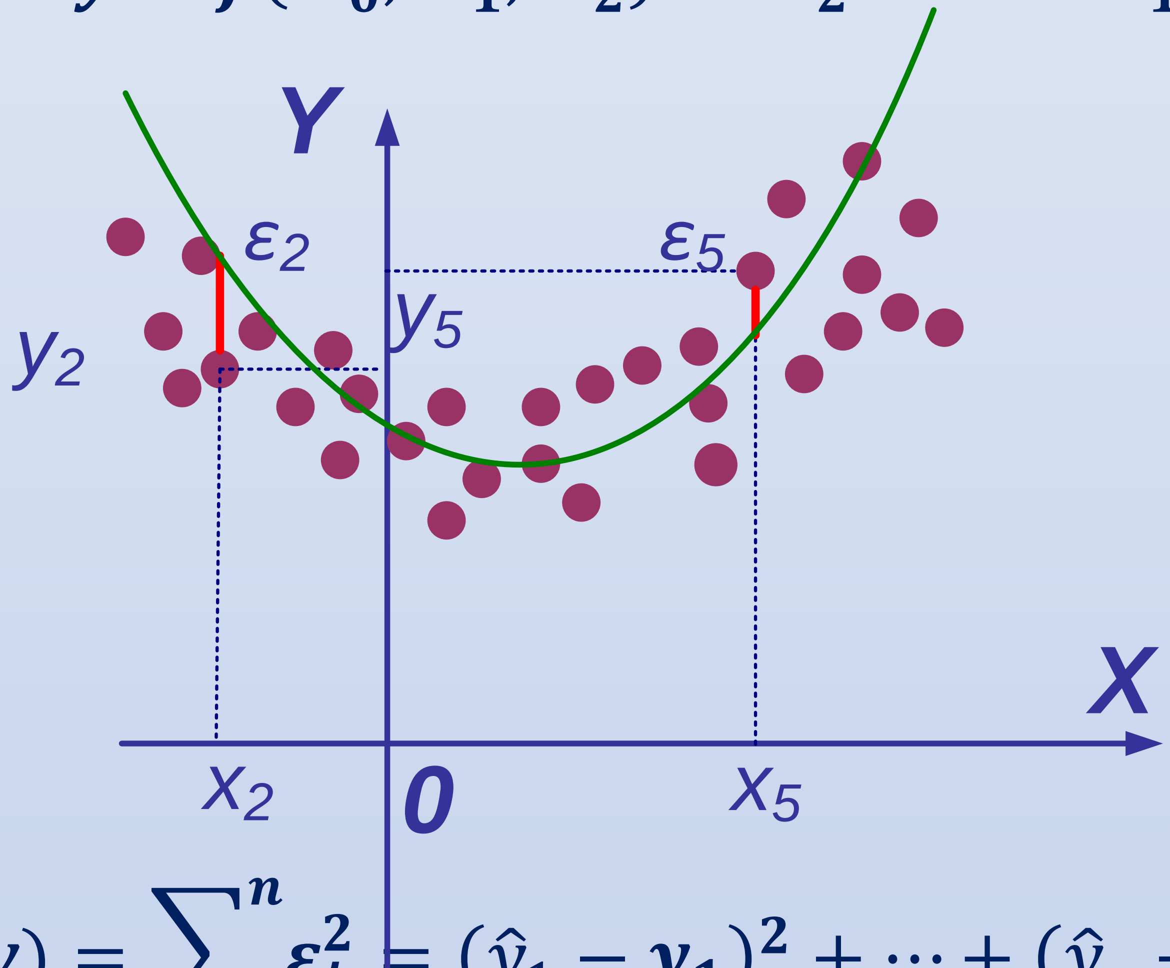
$$\frac{1}{2n}\sum_{1}^{n}(\hat{y}_i - w_0 - w_1 * x_i)^2$$

Знайти такі $(\overline{w})$, що $L(\overline{w}) = \min_{w} L(w)$

SLR

$$\boldsymbol{y = f(w_0, w_1, w_2) = w_2 * x^2 + w_1 * x + w_0}$$

x_1, y_1
x_2, y_2
$\ldots\ldots$
x_n, y_n



$$\boldsymbol{L(w) = \sum_1^n \varepsilon_i^2} = (\hat{y}_1 - \boldsymbol{y_1})^2 + \cdots + (\hat{y}_n - \boldsymbol{y_n})^2$$

Знайти такі $(\overline{\boldsymbol{w}})$, що $\boldsymbol{L(\overline{w})} = \min_{\boldsymbol{w,b}} \boldsymbol{L(w)}$

Лінійна регресія. SLR

Типова оптимізаційне завдання.

Знаходимо часткові похідні за параметрами w_i

$$\frac{\partial L}{\partial w_i} = a_{i,j} * w_j + b_i$$

2 Дорівнюємо похідні 0, отримуємо систему лінійних алгебраїчних рівнянь

$$a_{i,j} * w_j + b_i = 0$$

3. Вирішуємо систему та отримуємо

$$\overline{w}_j$$

MLR

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

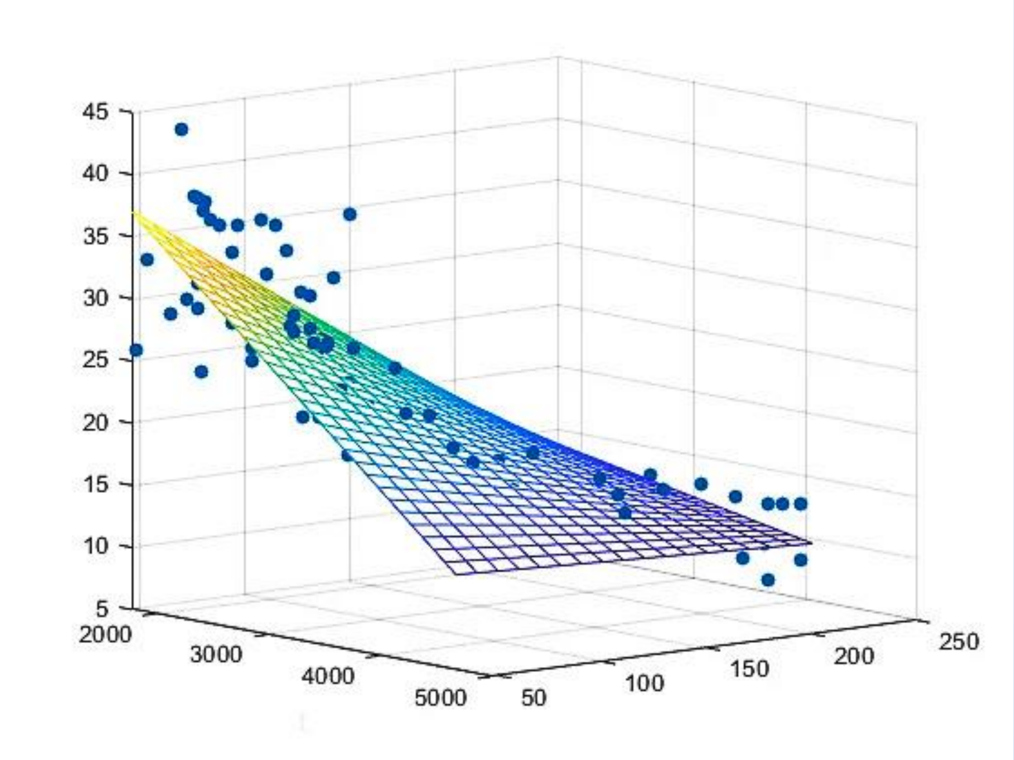
$$\widehat{\boldsymbol{y}} = \boldsymbol{f}(\boldsymbol{w_0}, \boldsymbol{W}) = \boldsymbol{W^T} * \boldsymbol{X} + \boldsymbol{w_0}$$

X_1, y_1

$X_{\mathbf{2}}, y_2$

……

X_n, y_n



$$\boldsymbol{L}(\boldsymbol{w}) = \sum_{\boldsymbol{1}}^{\boldsymbol{n}} \boldsymbol{\varepsilon}_{\boldsymbol{i}}^{\boldsymbol{2}} = (\hat{y}_1 - \boldsymbol{y_1})^2 + \cdots + (\hat{y}_i - \boldsymbol{y_i})^2 + \cdots$$

Знайти такі $(\overline{\boldsymbol{W}})$, **що** $\boldsymbol{L}(\overline{\boldsymbol{W}}) = \min_{\boldsymbol{w}} \boldsymbol{L}(\boldsymbol{W})$

Лінійна регресія. MLR

Типове оптимізаційне завдання. MLR

1. Знаходимо часткові похідні за параметрами w_i

$$\frac{\partial L}{\partial \boldsymbol{W}} = \frac{1}{2n}\left[y^T * y - 2y^T * \boldsymbol{W}^{\boldsymbol{T}} * \boldsymbol{X} \quad + \boldsymbol{W}^{\boldsymbol{T}} * \boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X} * \boldsymbol{W}\right] =$$

$$\frac{1}{2n}\left[-\boldsymbol{2}y^T * \boldsymbol{X} \quad + \boldsymbol{2} * \boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X} * \boldsymbol{W}\right]$$

2 Дорівнюємо похідні 0, отримуємо систему лінійних алгебраїчних рівнянь

$$\frac{\partial L}{\partial \boldsymbol{W}} = \boldsymbol{0} \Longrightarrow -y^T * \boldsymbol{X} \quad + \boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X} * \boldsymbol{W} = \boldsymbol{0}$$

$$\Longrightarrow y^T * \boldsymbol{X} \quad = \boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X} * \boldsymbol{W} \Longrightarrow \boldsymbol{W} = (\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}^{-\boldsymbol{1}})y^T * \boldsymbol{X}$$

Загальна регресія

Типова оптимізаційне завдання.

1. Знаходимо часткові похідні за параметрами w_i

$$\frac{\partial L}{\partial \boldsymbol{W}}$$

2 Виконуємо мінімізацію $\boldsymbol{L}(\boldsymbol{W})$ за допомогою ітераційного градієнтного спуску

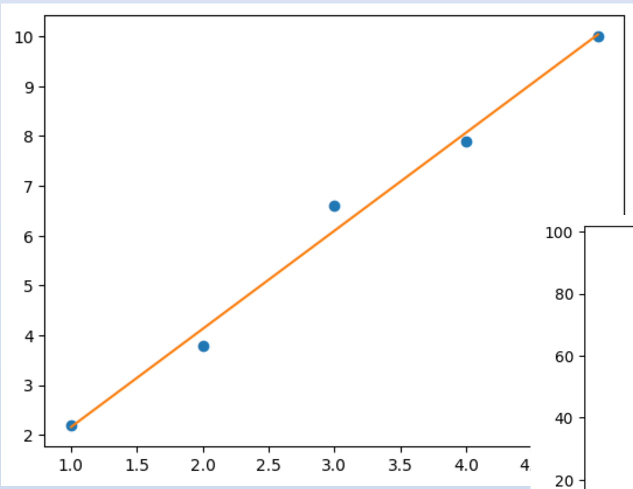
$$\boldsymbol{W}^{(t+1)} = \boldsymbol{W}^{(t)} - \boldsymbol{\lambda}\frac{\partial L}{\partial \boldsymbol{W}}$$

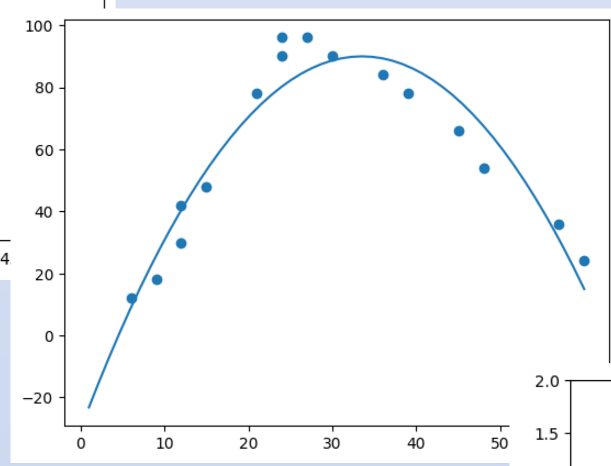
Gradient Descent Algorithm (GDA)

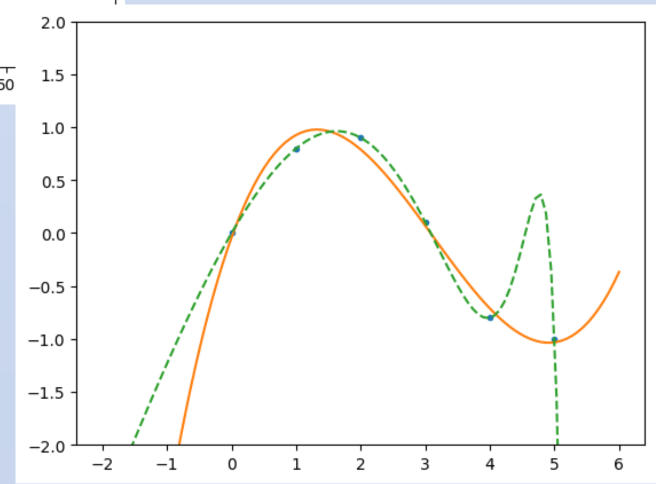
$\boldsymbol{\lambda}$ – крок, швидкість навчання (**learning rate**)

Приклади

Дивись: lec_02_01_Exmpl_1.md







Проблеми регресії

Маємо множину вхідних даних

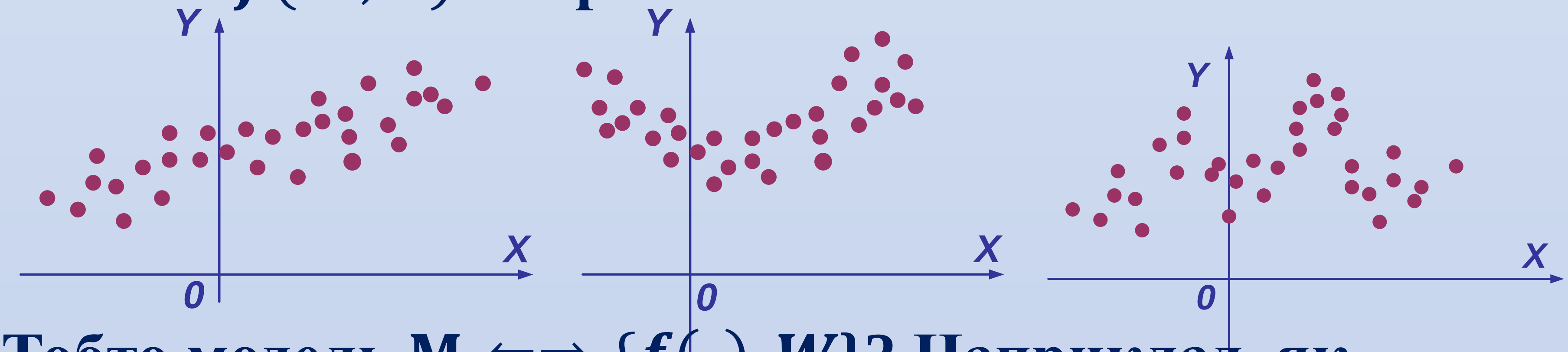
$$X_1, y_1, \ldots, X_n, y_n$$

Вирішуємо завдання регресії, тобто пошуку

$$\boldsymbol{\hat{y} = f(W, X)}$$

Можемо знайти $\boldsymbol{W}$, якщо залежність $\boldsymbol{\hat{y} = f(W, X)}$ відомо.

Але як $\boldsymbol{f(W, X)}$ обирати?



Тобто модель М $\Longleftrightarrow \{\boldsymbol{f(.), W}\}$? Наприклад, як визначити порядок поліному.

Необхідно оцінювати якість регресі!

Контрольні запитання

- Поясніть сутність навчання з вчителем.
- Надайте формальну постановку задачі регресії.
- Вкажіть особливості лінійної регресії та її різновиди (SLR, MLR).
- Надайте загальний підхід до вирішення задачі лінійної регресії за допомогою мінімізації середньоквадратичної похибки.
- Надайте підхід до вирішення загальної задачі регресії за допомогою мінімізації похибки градієнтним методом.

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.

Корисні та цікави посилання

- **Метод найменших квадратів**

https://uk.wikipedia.org/wiki/метод_найменьших_квадратів

- **MSE**

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_squared_error.html#sklearn.metrics.mean_squared_error

The END
Модуль 2. Лекція 01.