

# **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

```
graph TD; A[Машинне навчання] --> B[Навчання з вчителем]; A --> C[Навчання без вчителя]; A --> D[Навчання з підкріпленням]; B --> B1[Регресія]; B --> B2[Класифікація]; C --> C1[Кластеризація]; C --> C2[Зменшення розмірності]; D --> D1[Вживання у незнайомій обстановці];
```

Навчання з  
вчителем

Регресія

Класифікація

Навчання без  
вчителя

Кластеризація

Зменшення  
розмірності

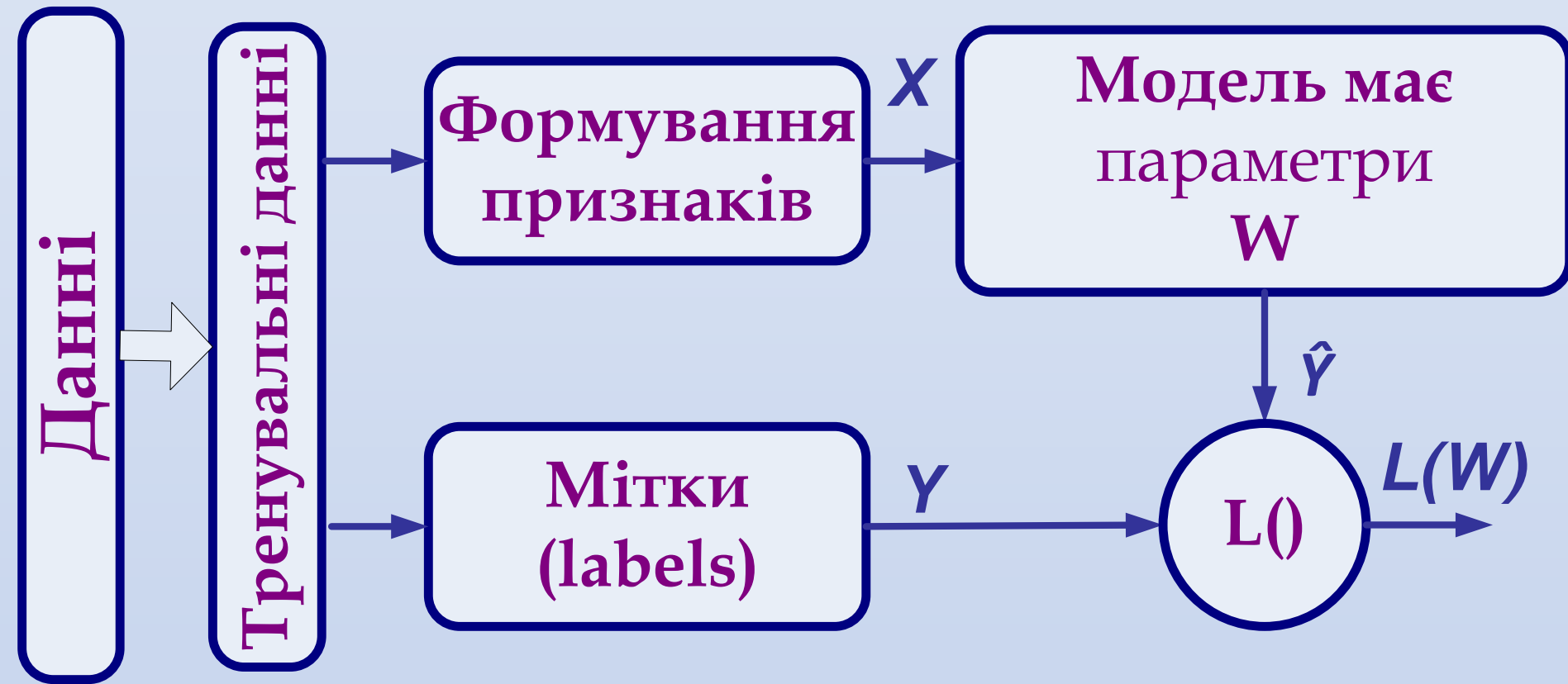
Навчання з  
підкріпленням

Вживання 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

```
graph TD; A[Машинне навчання] --> B[Навчання з вчителем]; A --> C[Навчання без вчителя]; A --> D[Навчання з підкріпленням]; B --> E[Регресія]; B --> F[Класифікація]; C --> G[Кластеризація]; C --> H[Зменшення розмірності]; D --> I[Вживання у незнайомій обстановці];
```

The diagram illustrates the classification of Machine Learning into three main categories: Supervised Learning, Unsupervised Learning, and Reinforcement Learning. Supervised Learning is further divided into Regression and Classification. Unsupervised Learning is divided into Clustering and Dimensionality Reduction. Reinforcement Learning is associated with application in unfamiliar environments. A red dashed line highlights the Supervised Learning category and its subtypes.

Навчання з  
вчителем

Регресія

Класифікація

Навчання без  
вчителя

Кластеризація

Зменшення  
розмірності

Навчання з  
підкріпленням

Вживання у  
незнайомій  
обстановці

# Регресія

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

**Регресія**  $\Leftrightarrow$  *Інтерполяція, Апроксимація, Екстраполяція*

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

1. **Інтерполяція** - це такий метод, опис вихідних даних, при якому інтерполуюча 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**  $\Leftrightarrow X$

**Залежна змінна**  $\Leftrightarrow Y$

**Відомі пари**  $\Leftrightarrow (x_i, y_i), i = 1, 2, \dots, m$

**Регресійна функція (модель)**  $\Leftrightarrow \hat{y} = f(W, X)$ , 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

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

# Регресія

Регресія

Лінійна регресія

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

Лінійна регресія

Поліноміальна регресія

Логаріфмічна регресія

...

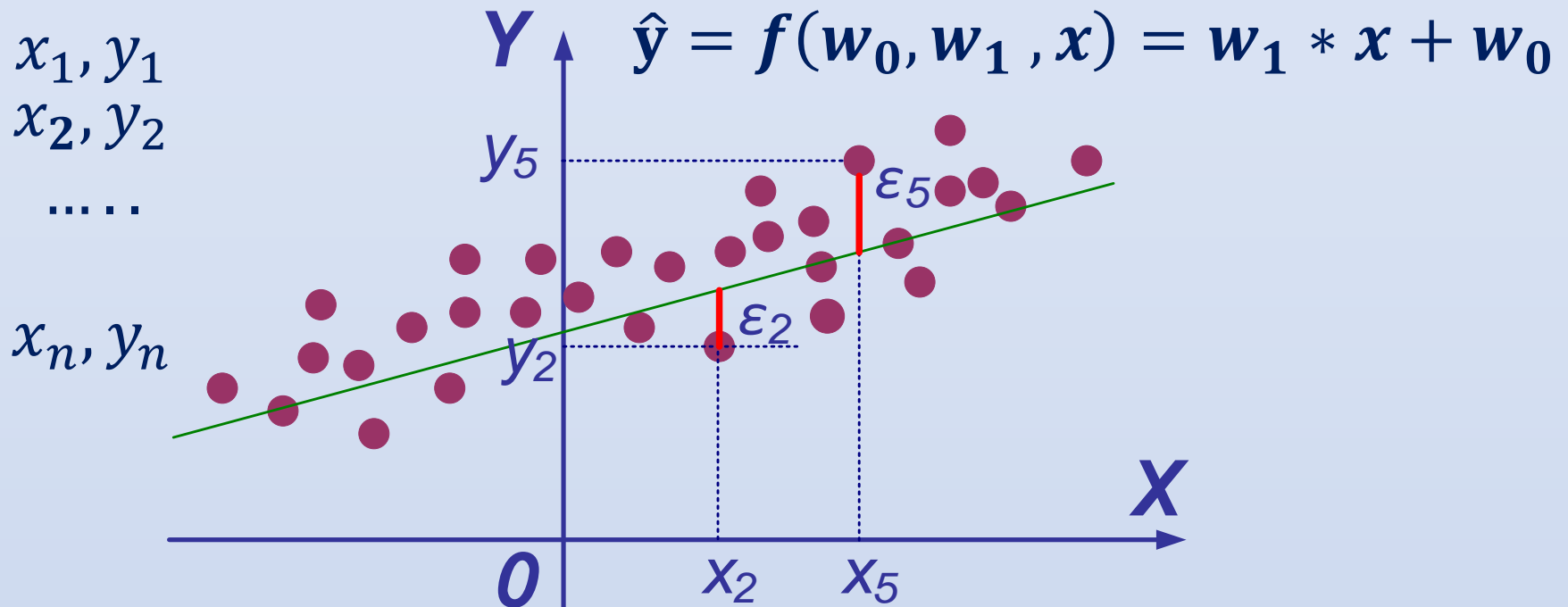
$x$  скаляр  $\Leftrightarrow$  парна лінійна регресія (найпростіша)

Simple Linear Regression (SLR)

$x$  вектор  $\Leftrightarrow$  множинна лінійна регресія

Multiple Linear Regression (MLR)

# SLR

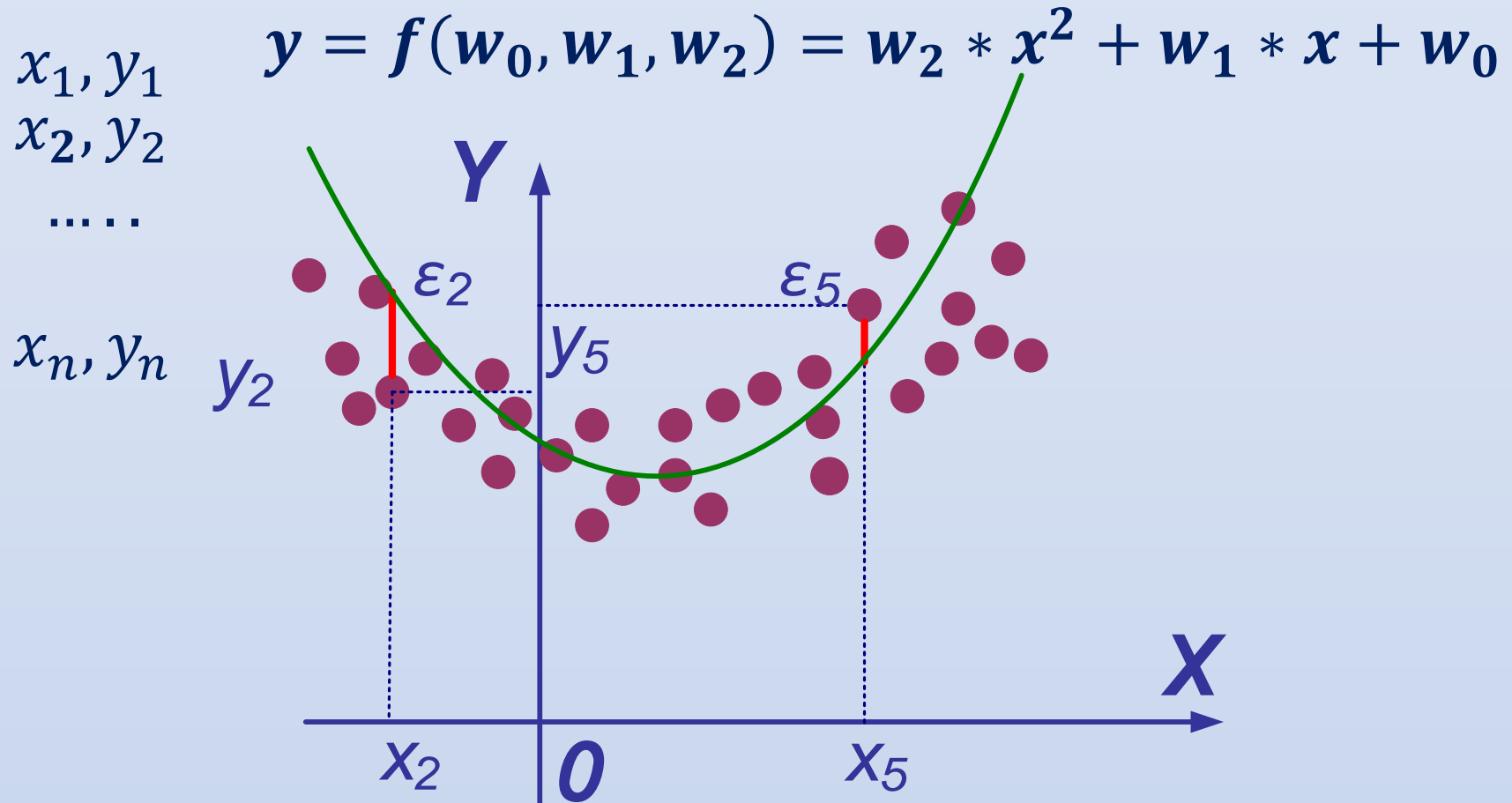


$$L(w) = \frac{1}{2n} \sum_{i=1}^n \epsilon_i^2 = \frac{1}{2n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 =$$
$$\frac{1}{2n} \sum_{i=1}^n (\hat{y}_i - w_0 - w_1 * x_i)^2$$

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



# SLR



$$L(w) = \sum_{i=1}^n \epsilon_i^2 = (\hat{y}_1 - y_1)^2 + \dots + (\hat{y}_n - y_n)^2$$

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

# 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

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

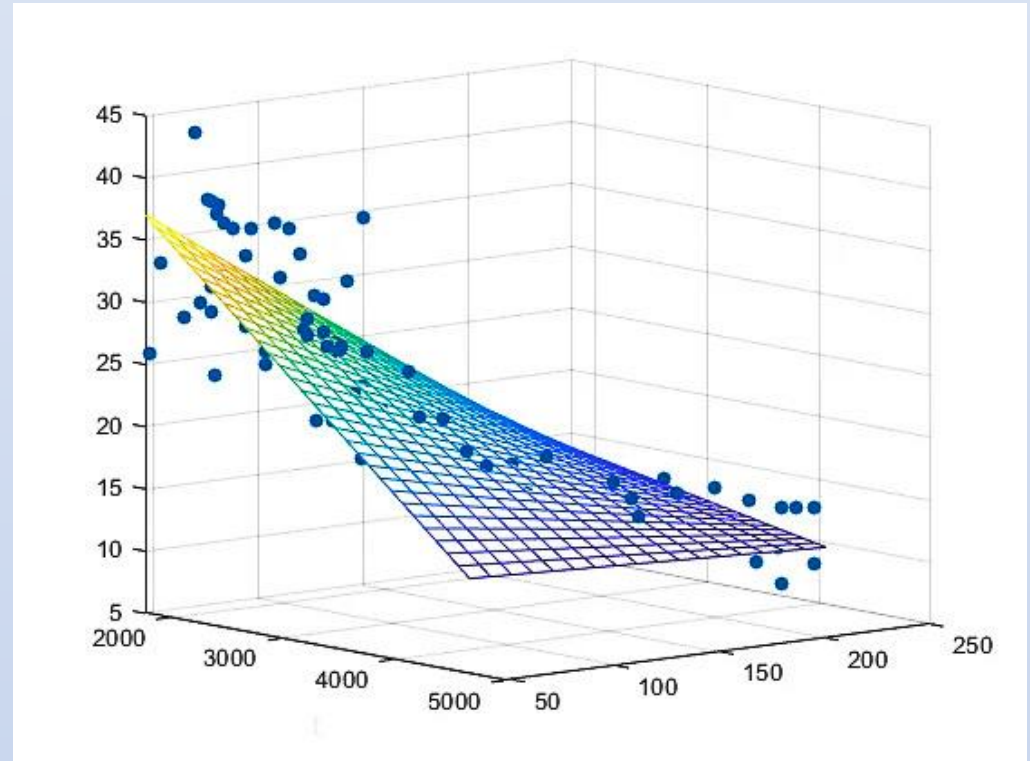
$X_1, y_1$

$X_2, y_2$

.....

$X_n, y_n$

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



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

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

# Лінійна регресія. MLR

## Типове оптимізаційне завдання. MLR

1. Знаходимо часткові похідні за параметрами  $w_i$

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

2 Дорівнюємо похідні 0, отримуємо систему лінійних алгебраїчних рівнянь

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

$$\Rightarrow y^T * X = X^T X * W \Rightarrow W = (X^T X^{-1}) y^T * X$$

# Загальна регресія

Типова оптимізаційне завдання.

1. Знаходимо часткові похідні за параметрами  $w_i$

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

2 Виконуємо мінімізацію  $L(W)$  за допомогою ітераційного градієнтного спуску

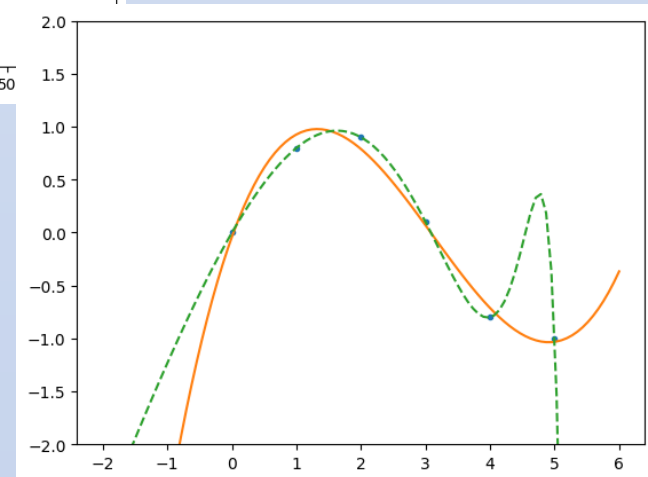
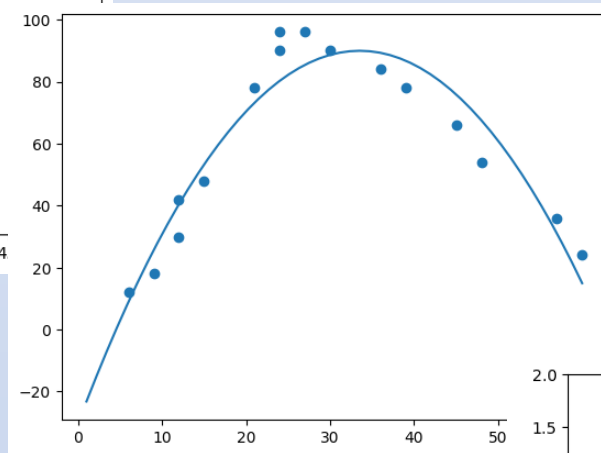
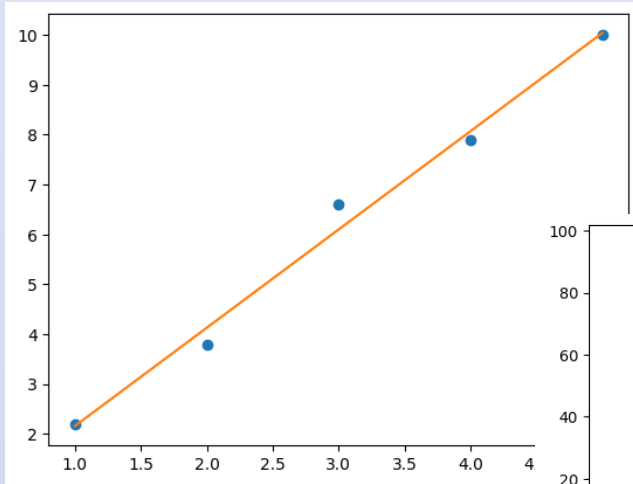
$$W^{(t+1)} = W^{(t)} - \lambda \frac{\partial L}{\partial W}$$

**Gradient Descent Algorithm (GDA)**

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

# Приклади

Дивись: `lec_02_01_Exmpl_1.md`



# Проблеми регресії

Маємо множину вхідних даних

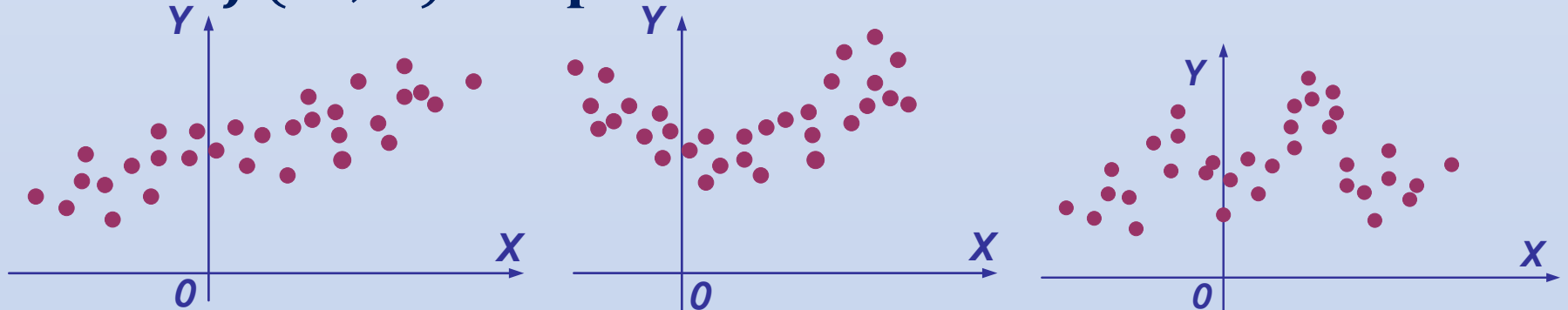
$$X_1, y_1, \dots, X_n, y_n$$

Вирішуємо завдання регресії, тобто пошуку

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

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

Але як  $f(W, X)$  обирати?



Тобто модель  $M \Leftrightarrow \{f(\cdot), W\}$ ? Наприклад, як визначити порядок поліному.

Необхідно оцінювати якість регресії!

# Контрольні запитання

- **Надайте визначення штучного інтелекту.**
- **Вкажіть особливості слабого штучного інтелекту (NAI).**
- **Вкажіть особливості сильного штучного інтелекту (AGI).**
- **Вкажіть особливості суперсильного штучного інтелекту (SI).**
- **Пояснить відмінності реактивних систем штучного інтелекту (RM).**
- **Пояснить відмінності систем штучного інтелекту з обмеженою пам'яттю (LM).**
- **Опишіть можливості гіпотетичних емоційних (TM) та самосвідомих (SA) систем ШІ.**



## 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\\_найменших\\_квадратів](https://uk.wikipedia.org/wiki/метод_найменших_квадратів)

- **MSE**

[https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean\\_squared\\_error.html#sklearn.metrics.mean\\_squared\\_error](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_squared_error.html#sklearn.metrics.mean_squared_error)

**The END**

**Модуль 2. Лекція 01.**