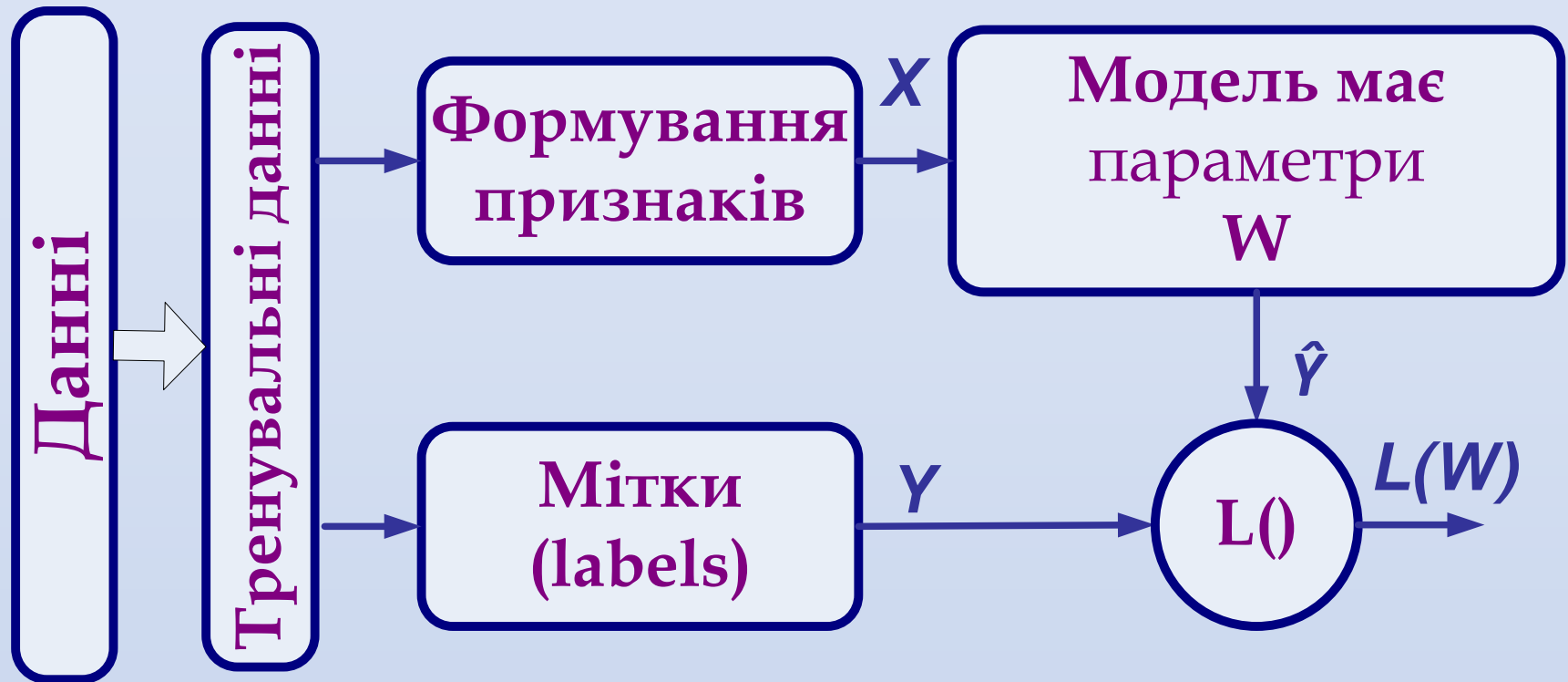


ОСНОВИ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ, НЕЙРОННИХ МЕРЕЖ та ГЛИБОКОГО НАВЧАННЯ

Модуль 7. TensorFlow / KERAS

Лекція 7.5. Тренування моделі

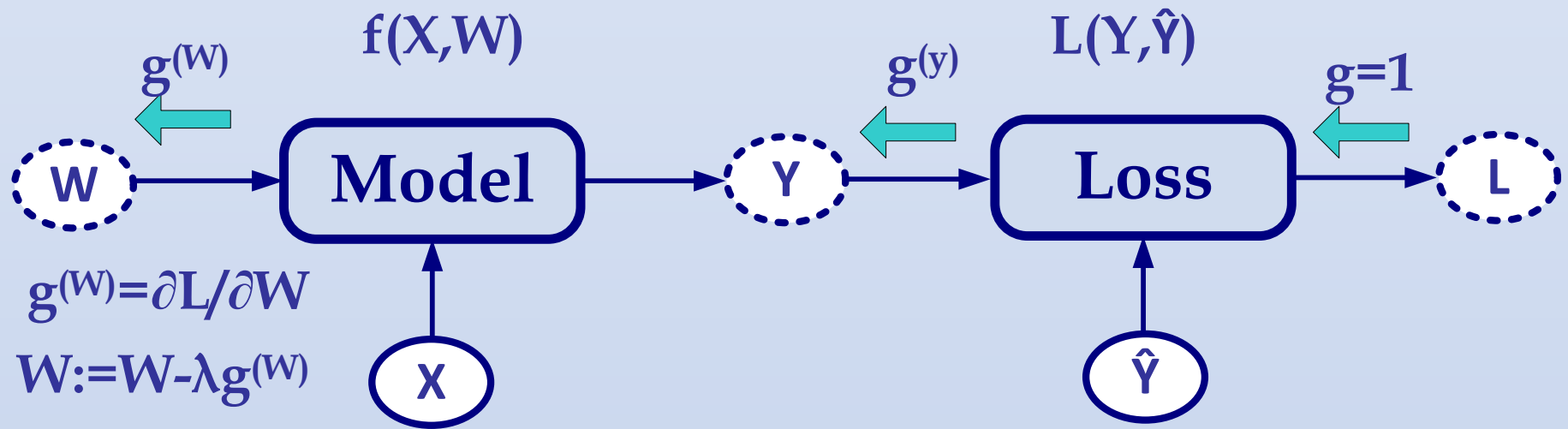
Навчання з вчителем



Навчання: знайти W , що мінімізують похибку (втрати) моделі

← Backward

Процес навчання \rightarrow пошук параметрів \bar{W} , які мінімізують втрати (Loss)



Загальний підхід \rightarrow використання методів градієнтного методу (gradient descent)

Градiєнтний спуск

Маєм тренувальний набір

$X = \{x_i | i = 0, 1, \dots < N - 1\}$ - множина векторів ознак

$Y = \{y_i | i = 0, 1, \dots < N - 1\}$ - множина міток

Деяким чином визначені початкові значення ваг моделі W

Визначена функція похибки (втрат, Loss)

$$L(W) = F(W, X, Y)$$

! Важливо: $L(W)$ залежить тільки від W

Градiєнтний спуск

Визначена функція $L(W)$

Необхідно знайти таке \bar{W} , що

$$L(\bar{W}) = \min (L(W))$$

\bar{W} - ваги, для якої функція похибки досягає свого мінімального значення.

Узагальнено ітераційний процес пошуку \bar{W} :

$$W^{(t+1)} = W^{(t)} - \Delta W^{(t)}; \Delta W^{(t)} = \lambda \nabla L(W^{(t)});$$

$$\nabla L(W^{(t)}) = \frac{\partial L(W)}{\partial W}$$

t - ітерація, $t=1,2, \dots$

$\Delta W^{(t)}$ - крок оптимізації ваг W

$\nabla L(W^{(t)})$ - градієнт функції похибки в точці $W^{(t)}$

λ - швидкість навчання (розмір кроку навчання - learning rate)

Тренування.

Тренувальний процес деякої моделі за допомогою навчання з вчителем (керованого навчання) зазвичай складається з наступних кроків:

- Отримати навчальні дані X, Y .
- Додати визначення моделі $\hat{Y} = M(W, X)$.
- Додати визначення функції втрат $L(\hat{Y}, Y)$.
- Перебрати дані навчання X , обчисливши втрати $L(M(W, X), Y)$ від ідеального значення.
- Обчислити градієнти для втрат

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

- Налаштувати W за деяким градієнтним методом.
- Оцінити результат.

Тренування.

Training dataset - тренувальний набір даних - набір прикладів, що використовують для пошуку параметрів моделі (ваг з'єднань між нейронами в NN)

Validation dataset - затверджувальний набір даних - оцінку допасованості моделі на тренувальному наборі даних при налаштуванні гіперпараметрів моделі

Test dataset - випробувальний набір даних - який використовують для забезпечення неупередженої оцінки допасованості остаточної моделі на тренувальному наборі даних.

Тренування.

Розбиття даних



ML модель працює на основі навчальних даних, які надані моделі.

Тренування – процес визначення параметрів моделі.

Валідація – процес перевірки моделі.

Тестування – процес оцінки якості моделі.

Приклад. Регресія

Формально. Маємо:

Незалежні змінні $\Leftrightarrow X$

Залежна змінна $\Leftrightarrow Y$

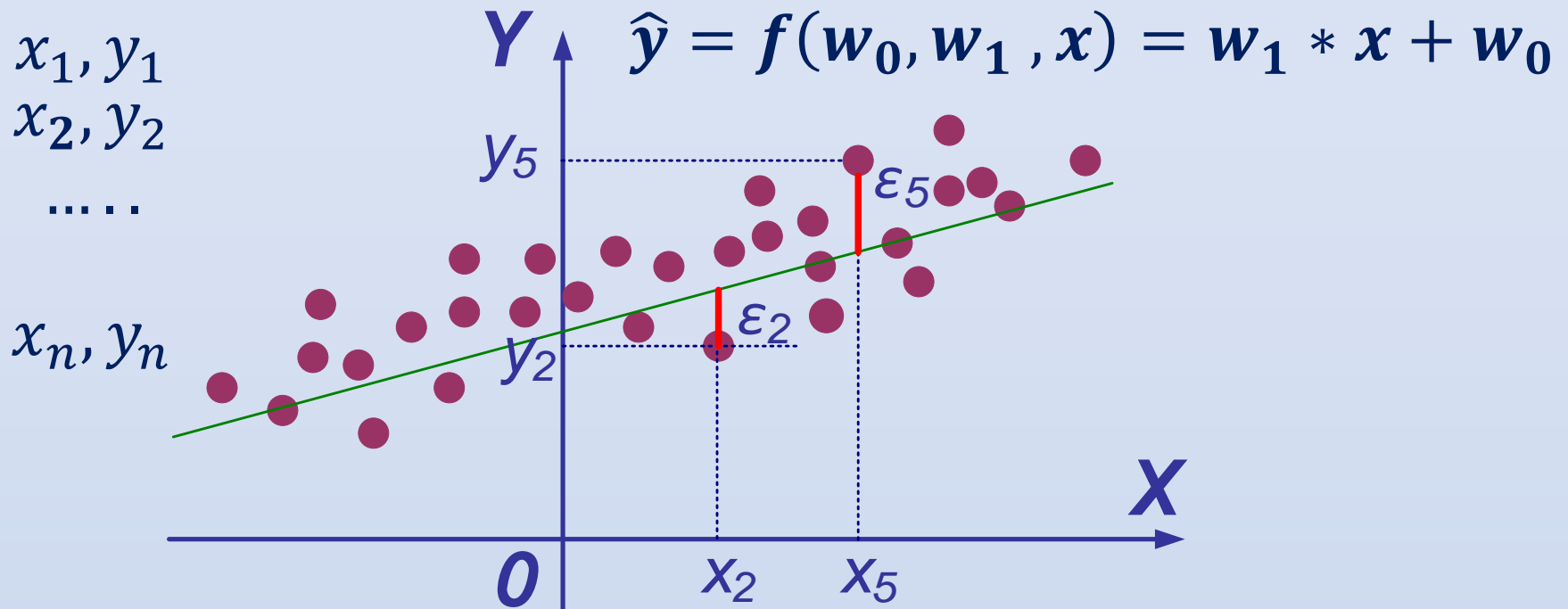
Відомі пари $\Leftrightarrow (x_i, \hat{y}_i), i = 1, 2, \dots, m$

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

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

SLR



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

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

Tensor Flow

TensorFlow Official

<https://www.tensorflow.org/>

TensorFlow API Documentation

https://www.tensorflow.org/api_docs/python/tf

TensorFlow on GitHub

<https://github.com/tensorflow/tensorflow>

Приклади дивись

Lec_07_05_Exmpl_1.md

Lec_07_05_Exmpl_2.md

Lec_07_05_Exmpl_3.md

The END

Модуль 7. Лес. 7.5.