Machine Learning

Chap2

- The gradient of a function ::= a vector of partial derivatives

$$f([x^{(1)}, x^{(2)}]) = ax^{(1)} + bx^{(2)} + c, \qquad \nabla f \text{ is given by the vector } [\tfrac{\partial f}{\partial x^{(1)}}, \tfrac{\partial f}{\partial x^{(2)}}]$$

- Random variables

  - Usually expressed by italic capital letter

  - Expectation value

$$\mathbb{E}[X] \stackrel{\text{def}}{=} \sum_{i=1}^{k} [x_i \cdot \Pr(X = x_i)]$$
$$= x_1 \cdot \Pr(X = x_1) + x_2 \cdot \Pr(X = x_2) + \cdots + x_k \cdot \Pr(X = x_k),$$

- Hyperparameters -> set by data analysts
- Parameters -> directly modified by learning algorithms
- Classification ::= assign labels to unlabeled examples
- Regression ::= predicting real valued labels (target)

- Model based & Instrumental Based Learning

  - Model-based:

    - Creates models

    - Develops parameter

    - Like w* & b* of SVM

  - Instrumental Based Learning

    - K-Nearest Neighbors (KNN)

    - Look at the close neighborhood of the input, output the most often ones.