Aprententatge Automàtic per a Xarxes (ML4Net)

Seminar 4 - Neural Networks

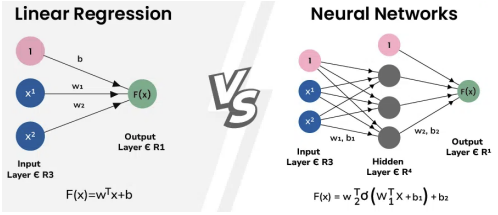
Francesc Wilhelmi & Boris Bellalta

School of Engineering
Universitat Pompeu Fabra

Supervised Learning through Neural Networks

Key concepts:

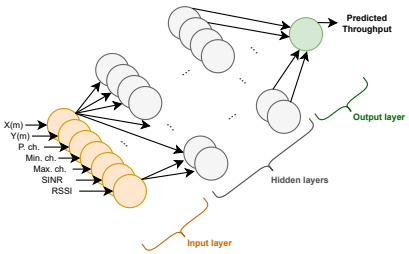
- Features (x): Available data that are fed into the neural network $\mathbf{x} = [x_1, x_2, ..., x_n]^T$, where $x_i \in \mathbb{R}^d$.
- Labels (y): The ground truth associated with the features.
  - **Regression:** Labels are continuous numerical values ($y \in \mathbb{R}$).
  - **Classification:** Labels are discrete categorical values ($y \in \{c_1, c_2, ..., c_k\}$).
- Model (h): Function $h = f(\mathbf{x}; \theta) \rightarrow \hat{y}$ that learns the mapping between input features and output labels.
- Goal: Find the optimal parameters θ that minimize a loss function $L(y, f(\mathbf{x}; \theta))$.



[Source: https://www.geeksforgeeks.org/linear-regression-vs-neural-networks-understanding-key-differences/]

Feed-forward neural networks

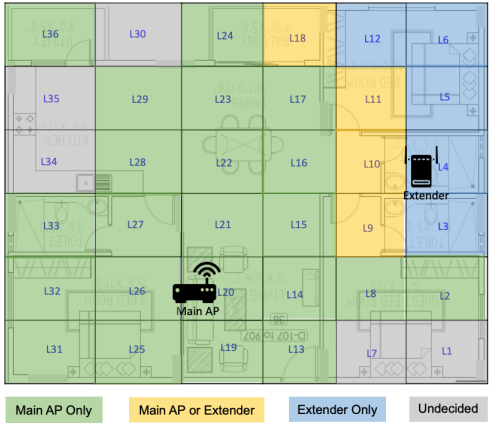
- **Input Layer:** Receives the input features **x**, so the number of neurons (often) matches the number of features.
- **Hidden Layers:** One or more intermediate layers between the input and output layers.
- **Output Layer:** Produces the final prediction of the network.
- **Weights and Biases (θ):** The learnable parameters of the network.
- **Activation Functions:** Non-linear functions applied to the weighted sum of inputs in each neuron.



Problem to solve

Wi-Fi performance prediction:

- Our goal is to predict the performance of a Wi-Fi deployment (throughput & delay).
- We will use multiple measurements taken from STA devices over one month.
- The measurements were taken at multiple locations (up to 36).
- STAs can connect to main AP or to an extender device (acting as AP).



Thangadorai, K. K., Sivalingam, K. M., Gupta, H. P., & Kanagarathinam, M. R. (2024, April). Stickyless: An intelligent method for solving sticky client problem in wi-fi networks. In 2024 IEEE Wireless Communications and Networking Conference (WCNC) (pp. 1-6). IEEE.

Dataset

The provided dataset is divided into train (`train_data.csv`) and test (`test_data.csv`) samples.

- [Feature] **LocationNumber:** Represents the zone number where data was collected (up to 35).
- [Feature] **RSSI:** The signal strength at the given location.
- [Feature] **TxLinkSpeed:** Transmission link speed (downlink, AP to STA).
- [Feature] **RxLinkSpeed:** Transmission link speed (uplink, STA to AP).
- [Feature] **MainAPConnect:** A binary feature that indicates whether the client is connected to the main AP (1) or an extender (0).
- [Feature] **TxThroughput:** The actual throughput achieved.
- [Feature] **AvgPingLatency:** The average latency experienced (measured using ping tests to a server).