Lecture 3 (Supervised Learning)

1 Difference between ML and DL

ML is the generic term given to the technique to analyse data and find *patterns*. DL is about the **Artificial Neural Networks (ANNs)**

2 Machine Learning Settings

1. Supervisied Learning
2. Un-Supervised Learning
3. Semi-Supervisied Learning
4. (Deep) Reinforcement Learning

3 Supervisied Learning

The data is given as $\{x_i, y_i\}_{i=1}^m$, such that $\forall i \in \{1, 2, \ldots, m\}$, $x_i \in \mathbb{R}^n$. Different analysis for $m < n$ and $m \geq n$.
$\forall i \in \{1, 2, \ldots, m\}$, y_i can have different ranges such as D, $\mathbb{R}^p$ $(1 \leq p)$. For now, the range considered will be $\mathbb{R}$.
The ML model (μ) is defined by h_θ, which is also called the hypothesis and the purpose of learning is to "learn" the value of this h_θ.

3.1 Example

Classify between a monkey and chimpanzee:
We can classify this such that $y_i \in \{-1, 1\}$. The input contains the weight and height of each animal. Thus the input is given as $x_i \in \mathbb{R}^2$.
To interpret the data, we can plot all x_i on the 2-D plane and mark each vertex with the corresponding y_i. Now, to classify the data, we can make a separator on this plane using any $f(\{w, h\})$ which will be stored in h_θ.

Questions to ponder -

1. What hypothesis space should be used? (dimension of the hypothesis)
2. What is a good separator?

3. How should this separator be found out?

There is another class of problems where regression is performed (instead of classification as discussed above).

3.2 Hyperplane

Defined as:

$$\sum_{j=1}^{n} \theta_j \cdot x_j + \theta_0 = 0$$

4 Topics to be Discussed in the Course

1. Logistic Regression
2. GDA
3.
4. Decision Trees
5. Support Vector Machines (SVMs)
6. (Deep) Neural Networks (D)NNs

5 Unsupervised Learning

The data is given as $\{x_i\}_{i=1}^{m}$ (no y_i). Aim is to still find pattern in the data such as:

- Clustering
- Density estimation
- Expectation Maximisation (EM)
- Principal Component Analysis (PCA)