

What have I learnt?

Machine Learning

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial **intelligence** based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.

Supervised Learning

It is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately.

There are two parts:

Classification

Gives label as true or false.

Common classification algorithms are linear classifiers, support vector machines (SVM), decision trees, k-nearest neighbor, and random forest.

Regression

Predicts label in between a range.

Overfitting

Overfitting happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. To avoid Overfitting we use **Normalization** process.

Underfitting

It usually happens when we have less data to build an accurate model.

Outlier Vs Anomaly

Outlier = legitimate data point that's far away from the mean or median in a distribution.

Anomaly = illegitimate data point that's generated by a different process than whatever generated the rest of the data

Feature Engineering

Combining features or making new feature for better result, avoiding less important feature(as more feature need lot of time and storage for processing), avoiding anomaly and outliers can be part of feature engineering.

Unsupervised Learning

Unsupervised learning is a type of algorithm that learns patterns from untagged data.

Reinforcement Learning

Learning through trial and error.

What will be my next step?

Enhancing Python Coding Skills, Testing AWS machine learning services, Practice Amazon Sagemaker.