Insurance Prediction Analysis Using Machine Learning in Python

Methodology

Types of Machine Learning:

Machine Learning can be classified into 3 types as follows –

Supervised Learning algorithms are used when we have labelled data and are trying to predict a label (target) based off of known features (input variables). This is commonly used in applications where historical data predicts likely future events.

For example, it can attempt to predict the price of a product/car/house based on different features for products for which we have historical price data.

Unsupervised Learning algorithms are used when we have unlabelled data and are trying to group together similar data points based off of features. This is mainly used to explore the data and find some structure within.

For example, it can identify the image (cat or dog) based on different inputs which groups together similar segments and then attempts to recognize the image correctly. This is unsupervised learning, where a machine is not taught but learns from the data (in this case data about a dog or cat)

Reinforcement Learning occurs when a computer system receives data in a specific environment and then learns how to maximize its outcomes. That means this model keeps continues to learn until best possible behaviour is met. Reinforcement learning is frequently used for robotics, gaming, and navigation.
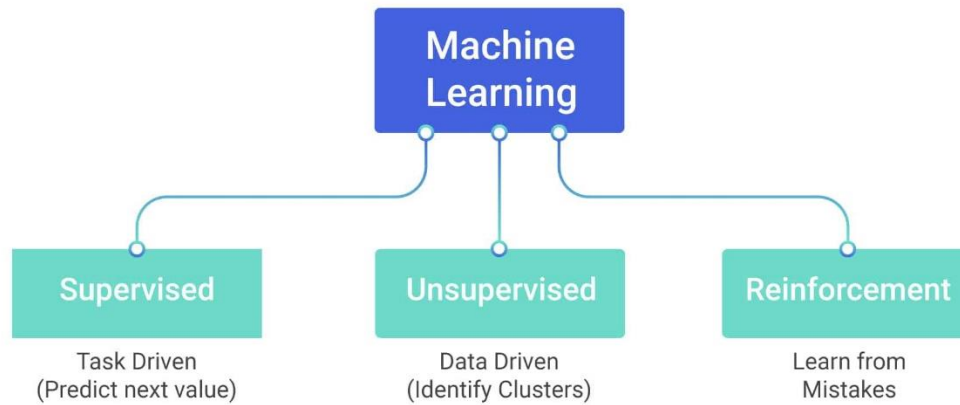
Linear Regression Algorithm:

Linear Regression is the first machine learning algorithm based on '**Supervised Learning'**. Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x).

When there is a single input variable (x), the method is referred to as '**Simple Linear Regression'**. When there are multiple input variables, the method is referred to as **'Multiple Linear Regression'.**

Software and Hardware used for evaluation:

CPU: Intel® Core™ i3-8130U CPU @ 2.20GHz × 4

Software: Anaconda3- Jupyter Notebook

Language: Python

OS: Windows, Ubuntu

This methodology is helpful to train data set to predict the precision and recall from the data file that is provided to predict the output. The algorithm divides the data set into two parts for training and testing the data purpose. After the values get predicted we will make changes into the parameters according to our convenience. The more data our model get, the more accurate values it predicts.