Module 2 ALGORITHMS

Introduction

Machine Learning is a subfield of a artificial intelligence that involves the development of algorithms and statistical mode that enable computers to improve their performance in tasks through experience. Machine learning starts with data numbers, photos or text. It is a growing technology which enables computers to learn automatically from past data. Machine Learning uses various algorithms for building mathematical models and making predication using historical data or information.

Algorithms:

***1. Altman Z-score method***

Professor Altman created a model for predicting bankruptcy in 1968. The Z-score is a method that combines five financial ratios to estimate the likelihood that a business would fail within two years. Each of these ratios is assigned an important element by the Z-score algorithm, which shows the proportionate weight each ratio has in the prediction of bankruptcy. The equation reads as

Z-Score = 1.2A + 1.4B + 3.3C + 0.6D + 1.0E

where,

1. = Working Capital / Total Assets
2. = Retained Earnings / Total Assets
3. = EBIT / Total Assets
4. = Market Value of Equity / Book Value of Total

Liabilities

1. = Sales / Total Assets

**Zeta**

Zeta, which is the Z score, is the outcome of the model that determines whether the company being analyzed is under financial distress or not and if it's likely to go bankrupt.

Zeta values lie between O-4, with the upper value being flexible.

***X1= W*orking Capital*/ Total asset Ratio***

Working capital is a financial metric representing operating liquidity available to a business and is measured as the difference between current assets and current liabilities.

Total assets are taken from the company's annual financial report filings. The higher the ratio, the better.

***X2= Retained Earning / Total asset Ratio***

Retained earnings (RE) is the amount of net income left over for the business after it has paid out dividends to its shareholder. The decision to provide retained earnings is with management and is vital to the company's value.

***X3=EBIT / Total asset Ratio***

EBITi.e., earnings before interests and taxes, is calculated by subtracting interest and tax expenses from revenue generated by a company during the fiscal year.

It can also be referred to as operating profit or profit before interests and taxes. Calculating this ratio can infer how well a company uses its assets to generate profits. Hence, determining the profitability of the company.

***X4 =Market value of equity / Total asset ratio***

It is not the same as book value, but it is also known as market capitalization. It is calculated by multiplying the total number of outstanding shares by one share's market value. It can also be referred to as the total value given by the investment community to a business.

***X5 = Total sales / Total Liabilities Ratio***

Total sales refer to the total income generated by a company by sales of all goods and services. At the same time, total liabilities comprise current liabilities and non-current liabilities.

Liabilities that are meant to be paid within a year are current liabilities, and those which are payable after a year are non-current liabilities.

**Range of Z score and observations** :

* According to the model, a company with a score greater than 2.99 means it's in a safe zone and under no financial distress.
* A company with a score of 1.8 < Z > 3, i.e., greater than 1.8 but less than 3, puts the company into a gray zone, meaning the company is under financial distress and has a high chance of going bankrupt shortly.
* A company with a score of 0 < Z > 1.8, i.e., greater than zero but less than 8, puts into a distress zone, making it most vulnerable to bankruptcy because of financial distress

***2. Springate Method***

This model, created by Gordon Springate in 1978, chooses four of the nineteen most typical financial ratios to assess the chance of an entity failing. Companies are categorised as "failed companies" at the time of computation if their Springate scores are less than 0.862 and as "healthy" if their scores are greater than 0.862. The following formula determines the Springate score:

**S = 1,03X1 + 3,07X2 + 0,66X3 + 0,4X4** where,

X1 = Working Capital / Total Assets

X2 = Net Profit Before Interest and Taxes / Total

Assets

X3 = Net Profit Before Taxes / Current Liability

X4 = Sales / Total Assets

***3. Zmijewski Method***

Another model for predicting company failure based on factors like performance, leverage, and financial liquidity is the Zmijewski score. The following formula serves as a summary of the contained ratios in this model:

**X = -4,3 – 4,5X1 + 5,7X2 – 0,004X3** where,

X1 = Earnings after Tax / Total Assets

X2 = Total Debt / Total Assets

X3 = Current Assets / Current Liabilities

**Random Forest Algorithm**

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. in classification problems , Random Forest builds decision tree from various samples and uses their majority vote for classification. In Regression problems, it builds decision tree from various samples and uses their average for regression. It is based on the concept of **ensemble learning,** which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

**Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.** it is applied in various industries such as banking and e-commerce to predict behaviour and outcomes.

**Decision trees Algorithm:**

Decision trees algorithm is a popular supervised learning algorithm that is used for classification and Regression Problem. It generate the outcomes as optimized result based upon the tree structure with the condition or rules. Decision trees are versatile machine learning algorithms capable of performing both regression and classification tasks and even work in case of tasks which have multiple outputs.

The goal of using a Decision Tree is to create a training model that can use to predict the class or value of the target variable by **learning simple decision rules**inferred from prior training data.

It is considered to be the most understandable Machine Learning algorithm, and it can be easily interpreted. Unlike most Machine Learning algorithms, it works effectively with non-linear data.

Advantage of Decision tree :

* Easy to understand
* Requires minimum data cleaning
* No constraint on the data type

Disadvantage of Decision tree :

* Possibility of overfitting

**DecisionTreeRegressio ():**It is the decision tree regressor function used to build a decision tree model in Machine Learning using python. The DecisionTreeRegressor () function looks like this:

**DecisionTreeRegressor (criterion = ‘mse’, random\_state =None , max\_depth=None, min\_samples\_leaf=1,)**

* **criterion:**This function is used to measure the quality of a split in the decision tree regression. By default, it is ‘mse’ (the mean squared error), and it also supports ‘mae’ (the mean absolute error).
* **max\_depth:**This is used to add maximum depth to the decision tree after the tree is expanded.
* **min\_samples\_leaf:** This function is used to add the minimum number of samples required to be present at a leaf node.

**XGBoost Algorithm :**

XGBoost or Extreme Gradient Boosting is a machine learning algorithm that is used for the implementation of gradient boosting decision trees. It is an optimized distributed gradient boosting library designed for efficient and scalable training of machine learning models. XGBoost is an implementation of gradient-boosting decision trees. It has been used by data scientists and researchers worldwide to optimize their machine-learning models. XGBoost is known for its accuracy and has been shown to outperform other machine learning algorithms in many predictive modeling tasks