Prior Knowledge for Fertilizers Recommendation System for Disease Prediction:

Date	19 September 2022
Team ID	PNT2022TMID33700
Project Name	Fertilizers Recommendation System for Disease Prediction

Machine Learning Algorithms:

Machine Learning algorithms are the programs that can learn the hidden patterns from the data, predict the output, and improve the performance from experiences on their own. Different algorithms can be used in machine learning for different tasks, such as simple linear regression that can be used **for prediction problem**s like **stock market prediction,** and **the KNN algorithm can be used for classification problems**.

In this topic, we will see the overview of some popular and most commonly used machine learning algorithms along with their use cases and categories.

Types of Machine Learning Algorithms:

Machine Learning Algorithm can be broadly classified into three types:

1. **Supervised Learning Algorithms**
2. **Unsupervised Learning Algorithms**
3. **Reinforcement Learning algorithm**

The below diagram illustrates the different ML algorithm, along with the categories:

1)Supervised Learning Algorithm:

Supervised learning is a type of Machine learning in which the machine needs external supervision to learn. The supervised learning models are trained using the labeled dataset.

2) Unsupervised Learning Algorithm:

It is a type of machine learning in which the machine does not need any external supervision to learn from the data, hence called unsupervised learning.

3) Reinforcement Learning:

In Reinforcement learning, an agent interacts with its environment by producing actions, and learn with the help of feedback.

List of Popular Machine Learning Algorithm:

- ➢ **Linear Regression Algorithm**
- ➢ **Logistic Regression Algorithm**
- ➢ **Decision Tree**
- ➢ **SVM**
- ➢ **Naïve Bayes**
- ➢ **KNN**
- ➢ **K-Means Clustering**
- ➢ **Random Forest**
- ➢ **Apriori**
- ➢ **PCA**

1. Linear Regression:

Linear regression is one of the most popular and simple machine learning algorithms that is used for predictive analysis. Here, **predictive analysis** defines prediction of something, and linear regression makes predictions for *continuous numbers* such as **salary, age, etc.**

2. Logistic Regression:

Logistic regression is the supervised learning algorithm, which is used to **predict the categorical variables or discrete values**. It can be used for

the *classification problems in machine learning*, and the output of the logistic regression algorithm can be either Yes or NO, 0 or 1, Red or Blue,

3. Decision Tree Algorithm:

A decision tree is a supervised learning algorithm that is mainly used to solve the classification problems but can also be used for solving the regression problems.

The **internal node** is used to represent the **features of the dataset, branches show the decision rules,** and **leaf nodes represent the outcome of the problem.**

4. Support Vector Machine Algorithm:

A support vector machine or SVM is a supervised learning algorithm that can also be used for classification and regression problems. However, it is primarily used for classification problems. The goal of SVM is to create a hyperplane or decision boundary that can segregate datasets into different classes.

The data points that help to define the hyperplane are known as **support vectors**, and hence it is named as support vector machine algorithm.

Some real-life applications of SVM are **face detection, image classification, Drug discovery**
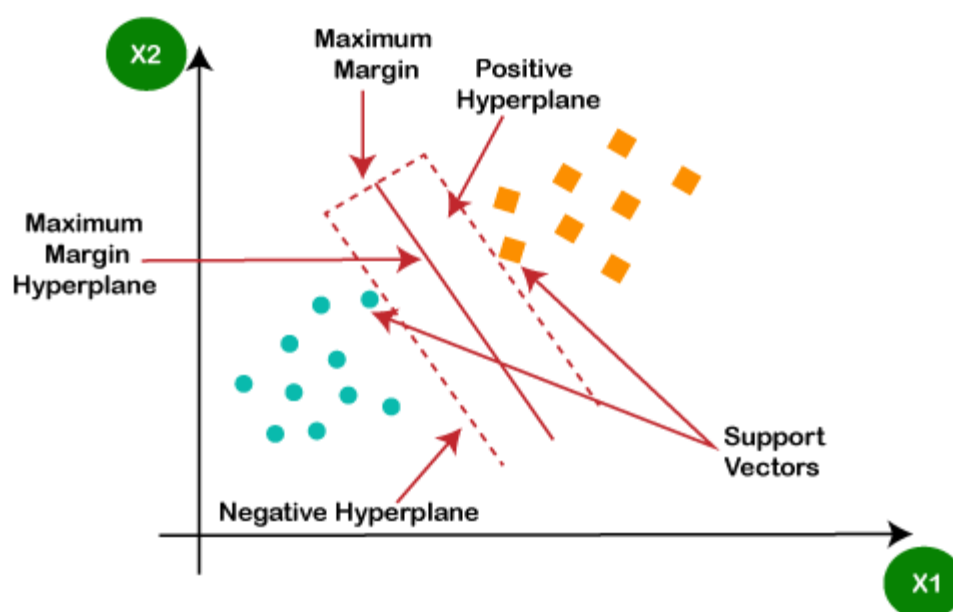
6. K-Nearest Neighbour (KNN):

K-Nearest Neighbour is a supervised learning algorithm that can be used for both classification and regression problems. This algorithm works by assuming the similarities between the new data point and available data points. Based on these similarities, the new data points are put in the most similar categories. It is also known as the lazy learner algorithm as it stores all the available datasets and classifies each new case with the help of K-neighbours. The new case is assigned to the nearest class with most similarities, and any distance function measures the distance between the data points. The distance function can be **Euclidean, Minkowski, Manhattan, or Hamming distance**, based on the requirement.

7. K-Means Clustering:

K-means clustering is one of the simplest unsupervised learning algorithms, which is used to solve the clustering problems. The datasets are grouped into K different clusters based on similarities and dissimilarities

Flask:

Flask is a web development framework. It is a framework with a built-in development server and a debugger.

Flask framework in itself is different from the other archetypes as it allows web developers to be flexible and to comfortably accommodate the frequently released changes in the software development community.

Unlike the Django framework, Flask is very Pythonic. It's easy to get started with Flask, because it doesn't have a huge learning curve.

On top of that it's very explicit, which increases readability. To create the "Hello World" app, you only need a few lines of code.

This is a boilerplate code example.

```python
from flask import Flask
app = Flask(__name__)

@app.route('/')
def hello_world():
    return 'Hello World!'
```

```
if __name__ == '__main__':
    app.run()
```

If you want to develop on your local computer, you can do so easily. Save this program as server.py and run it with python server.py.

```
$ python server.py
 * Serving Flask app "hello"
 * Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
```

It then starts a web server which is available only on your computer. In a web browser open localhost on port 5000 (the url) and you'll see "Hello World" show up.