Fundamentals of Machine learning

1. Write a short note on semi-supervised learning algorithm

→ Semi-supervised learning is a type of machine learning that falls in between supervised and unsupervised learning. it is a method that uses a small amount of labeled data and a large amount of unlabeled data to train a model.

The goal of semi-supervised learning is to learn a function that can accurately predict the output variable based on the input variables, similar to supervised learning.

Semi-supervised learning is particularly useful when there is a large amount of unlabeled data available, but it's too expensive or difficult to label all of it.

Examples of semi-supervised learning :- Text classification, image classification and Anomaly detection.

2. Give Examples of semi-supervised learning algorithm

→ Text classification :- In text classification, the goal is to classify a given text into one or more predefined categories. Semi-supervised learning can be used to train a text classification model using a small amount of labeled data & a large amount of unlabeled text data.

* Label Propogation :- A social network analysis when users are connected based on their interactions known user are spread to predict

* Speech recognition :- Labelling audio is a very resource & time-intensive task so semi-supervised used with a small set of labeled transcribed audio & large of unlabeled audio recordings.

* Anamoly Dectetion :- Identifying outliers or anomalies in network traffic or sensor data where only a small number of anomalies are labeled & majority of the data is unlabled.

* Natural language processing :- using semi-supervised learning argument a small labeled dataset of product reviews with a larger set of unlabled reviews.

3. write a a few applications of semi-supervised learning algorithm.

→ Self-Training :- This algorithm involves a training a model on the labeled data, then using this model to label the unlabeled data.

* **co-Training:** - Two seperate models are trained on the different views or features of the data. Each model labels the unlabeled data & they are then used to train the other model.

* **Label Propogation:** - This method involves propogating labels from labeled to unlabeled data based on the Similarity blw data points. It uses a graph based approach where nodes represent datapoints

* **Generative Adversrial networks (GANs):** The descriminator is trained to distinguish blw real & generated data & to Classify data in different categories.
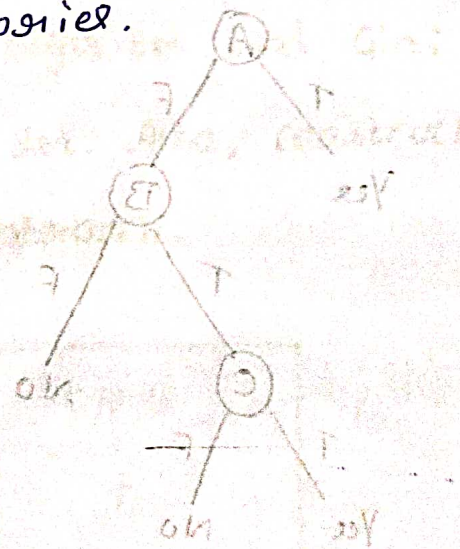
FML

Assignment -1 = Semi-Supervised learning algorithm

Submitted by :- Karthik. C

Reg. No :- 241058016