Training Day-13 Report:

<u>Random Forest Algorithm in Machine Learning</u>

-

Machine learning, a fascinating blend of computer science and statistics, has witnessed incredible progress, with one standout algorithm being the **Random Forest**. **Random forests or Random Decision Trees**

What is the Random Forest Algorithm?
Random Forest algorithm is a powerful tree learning technique in Machine Lea\ It works by creating a number of Decision Tree during the training phase. Each tree is constructed using a random subset of the data set to measure a random subset of features in each partition. This randomness introduces variability among individual trees, reducing the risk of overfitting and improving overall prediction performance. In prediction, the algorithm aggregates the results of all trees, either by voting (for classification tasks) or by averaging (for regression tasks) This collaborative decision-making process, supported by multiple trees with their insights, provides an example stable and precise results. Random forests are widely used for classification and regression functions, which are known for their ability to handle complex data, reduce overfitting, and provide reliable forecasts in different environments.

What are Ensemble Learning models?
Ensemble learning models work just like a group of diverse experts teaming up to make decisions – think of them as a bunch of friends with different strengths tackling a problem together. Picture it as a group of friends with different skills working on a project. Each friend excels in a particular area, and by combining their strengths, they create a more robust solution than any individual could achieve alone.
Similarly, in ensemble learning, different models, often of the same type or different types, team up to enhance predictive performance.

How Does Random Forest Work?
The random Forest algorithm works in several steps which are discussed below–>

- **Ensemble of Decision Trees:** Random Forest leverages the power of <u>ensemble learning </u>by constructing an army of Decision Trees. These trees are like individual experts, each specializing in a particular aspect of the data. Importantly, they operate independently, minimizing the risk of the model being overly influenced by the nuances of a single tree.

-
- **Random Feature Selection:** To ensure that each decision tree in the ensemble brings a unique perspective, Random Forest employs <u>random feature selection</u>. During the training of each tree, a random subset of features is chosen. This randomness ensures that each tree focuses on different aspects of the data, fostering a diverse set of predictors within the ensemble.
- <u>Bootstrap Aggregating or Bagging:</u> The technique of bagging is acornerstone of Random Forest's training strategy which involves creating multiple bootstrap samples from the original dataset, allowing instances to be sampled with replacement. This results in different subsets of data for each decision tree, introducing variability in the training process and making the model more robust.
- **Decision Making and Voting:** When it comes to making predictions, each decision tree in the Random Forest casts its vote. For <u>classification tasks</u>, the final prediction is

determined by the <u>mode</u> (most frequent prediction) across all the trees. In <u>regression tasks</u>, the average of the individual tree predictions is taken. This internal voting mechanism ensures a balanced and collective decision-making process.