



**Online Student Training for “Artificial Intelligence & Machine Learning”**  
(4<sup>th</sup> Feb, 2021 – 17<sup>th</sup> Mar, 2021)

# Random Forest Classifier

**Faculty Trainer**

Naw Varsha Pipada

Department of Computer Science & Engineering

Engineering College Bikaner



# Contents

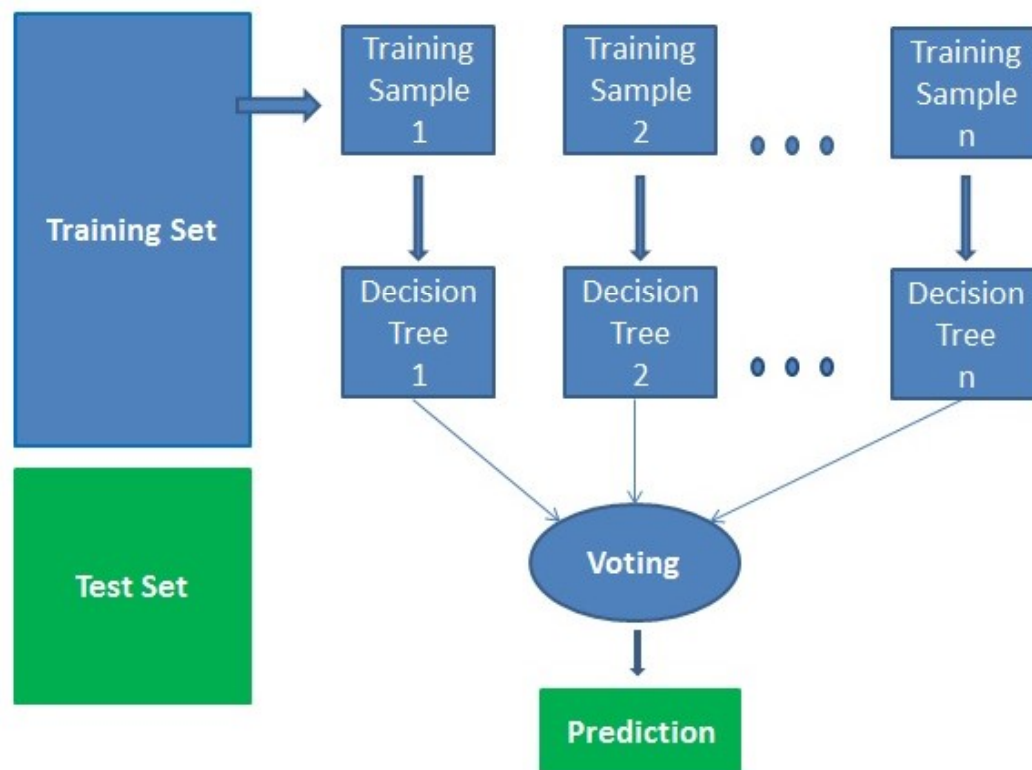
- Introduction
- How it works
- Pros & Cons
- Random Forest Vs Decision Trees
- Applications →



# Introduction

- Random forests is a supervised learning algorithm. It can be used both for classification and regression.
- It is also the most flexible and easy to use algorithm.
- A forest is comprised of trees. It is said that the more trees it has, the more robust a forest is.
- Random forests creates decision trees on randomly selected data samples, gets prediction from each tree and selects the best solution by means of voting.
- It also provides a pretty good indicator of the feature importance.

# How it Works



## How it Works

1. Select random samples from a given dataset.
2. Construct a decision tree for each sample and get a prediction result from each decision tree.
3. Perform a vote for each predicted result.
4. Select the prediction result with the most votes as the final prediction.

# Pros

- Highly accurate and robust method because of the number of decision trees participating in the process.
- Does not suffer from the overfitting problem.
- The algorithm can be used in both classification and regression problems.
- Random forests can also handle missing values. There are two ways to handle these: using median values to replace continuous variables, and computing the proximity-weighted average of missing values.
- You can get the relative feature importance, which helps in selecting the most contributing features for the classifier.

## Cons

- Random forests is slow in generating predictions because it has multiple decision trees.
- The model is difficult to interpret compared to a decision tree, where you can easily make a decision by following the path in the tree.

# Random Forests vs Decision Trees

- Random forests is a set of multiple decision trees.
- Deep decision trees may suffer from overfitting, but random forests prevents overfitting by creating trees on random subsets.
- Decision trees are computationally faster.
- Random forests is difficult to interpret, while a decision tree is easily interpretable and can be converted to rules.



# Applications

- Recommendation engines
- Image classification
- Feature selection
- Classify loyal loan applicants
- Identify fraudulent activity
- Predict diseases

See You

End of slides



Thank You for Your  
Attention