MINI PROJECT (4-1)
MID-TERM PRESENTATION

Real Estate based Recommendation System (Machine Learning)

Guide name: Prof. Bhakti Paranjape
Batch Coordinator: Prof. Shamla Mantri

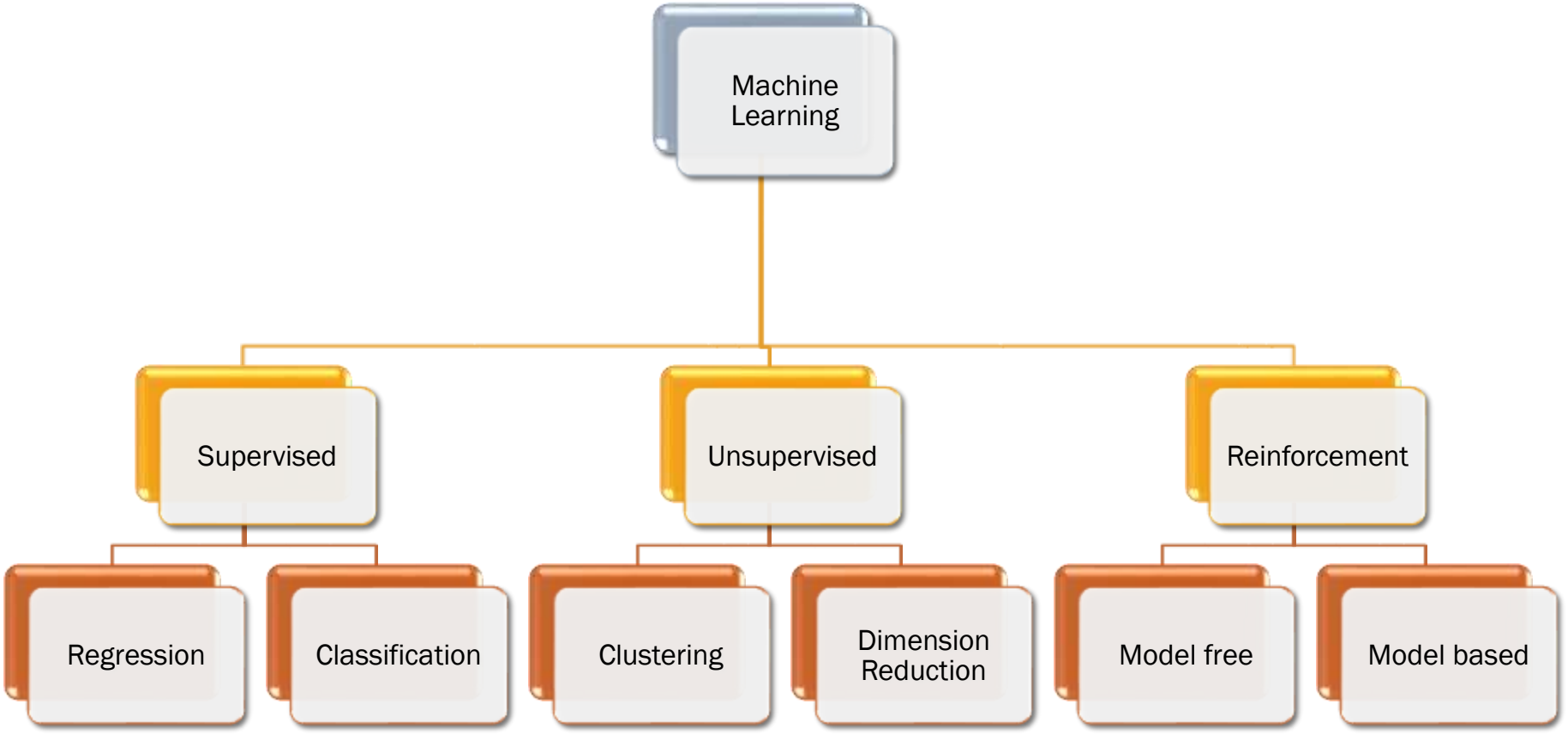
MACHINE LEARNING:

- Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.
- Tom Mitchell (1998) Well-posed Learning Problem: A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.
- Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T in this setting?

○ Classifying emails as spam or not spam.

○ Watching you label emails as spam or not spam.

○ The number (or fraction) of emails correctly classified as spam/not spam.

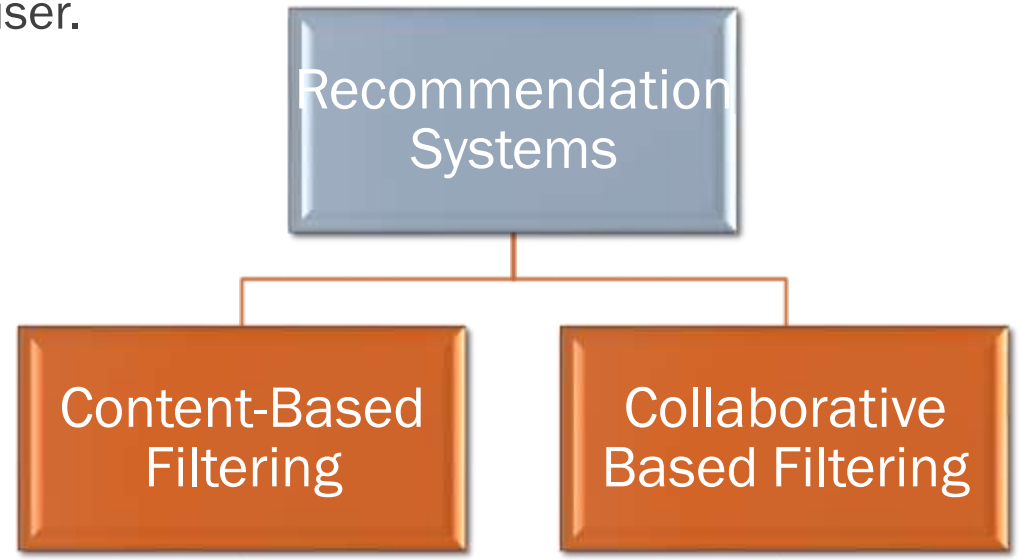

Machine Learning
Supervised
Unsupervised
Reinforcement
Regression
Classification
Clustering
Dimension Reduction
Model free
Model based

DATA SCIENCE

- Data science is a broad term for multiple disciplines, machine learning fits within data science.
- The main difference between the two is that data science as a broader term not only focuses on algorithms and statistics but also takes care of the entire data processing methodology.
- Data science, data analytics, and machine learning are some of the most in-demand domains in the industry right now.

RECOMMENDATION SYSTEMS:

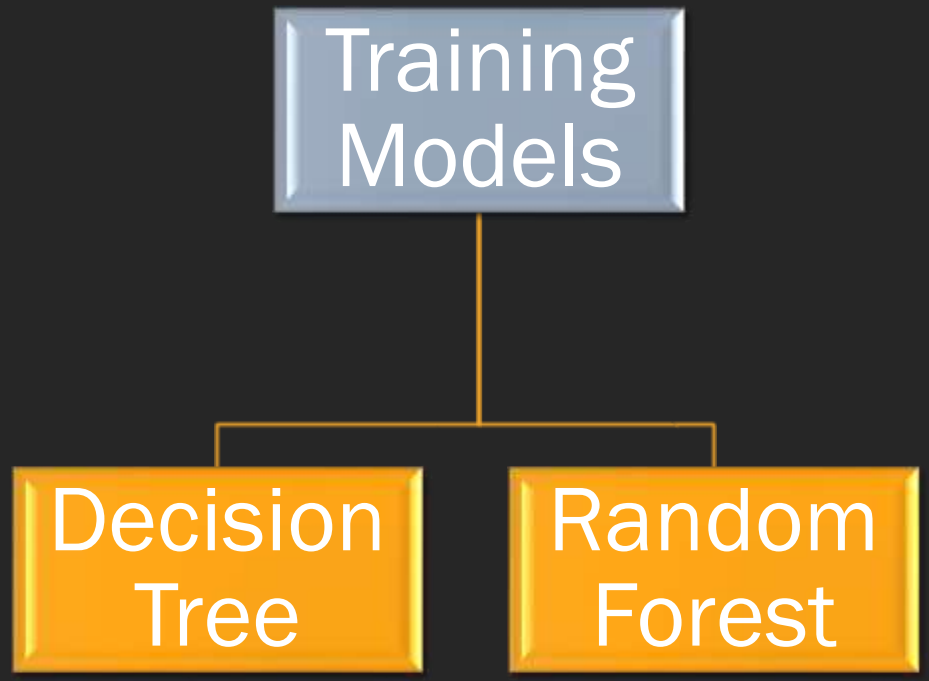
- Recommendation Systems are composed by algorithms that generate recommendations of a given type of item for users based on the information they provide
- Recommender engines (REs) also known as recommender systems are software tools and techniques providing suggestions to a user.



INTRODUCTION:

- This project outlines how the Recommendation System for Real Estate Project will address current real estate platform which has become a major priority for buyers & sellers.
- The project will compare and evaluate the results from two training models i.e. Decision Tree and Random Forest to predict the price.
- The motivation behind is to show how a recommender system can be applied in Real estate scenario.
- The main contributions of this research can be summarized as follow:
  - ❖ To analyze data of the real estate from the dataset
  - ❖ To evaluate and predict the results by using training models (Decision Tree and Random Forest)
  - ❖ To recommend the sectors of real estate based on data and compare the results.

SOFTWARE USED:



DECISION TREE

- Decision trees are constructed via an algorithmic approach that identifies ways to split a data set based on different conditions. It is one of the most widely used and practical methods for supervised learning.
- Decision Trees are a non-parametric **supervised learning** method used for both **classification** and **regression** tasks.

RANDOM FOREST

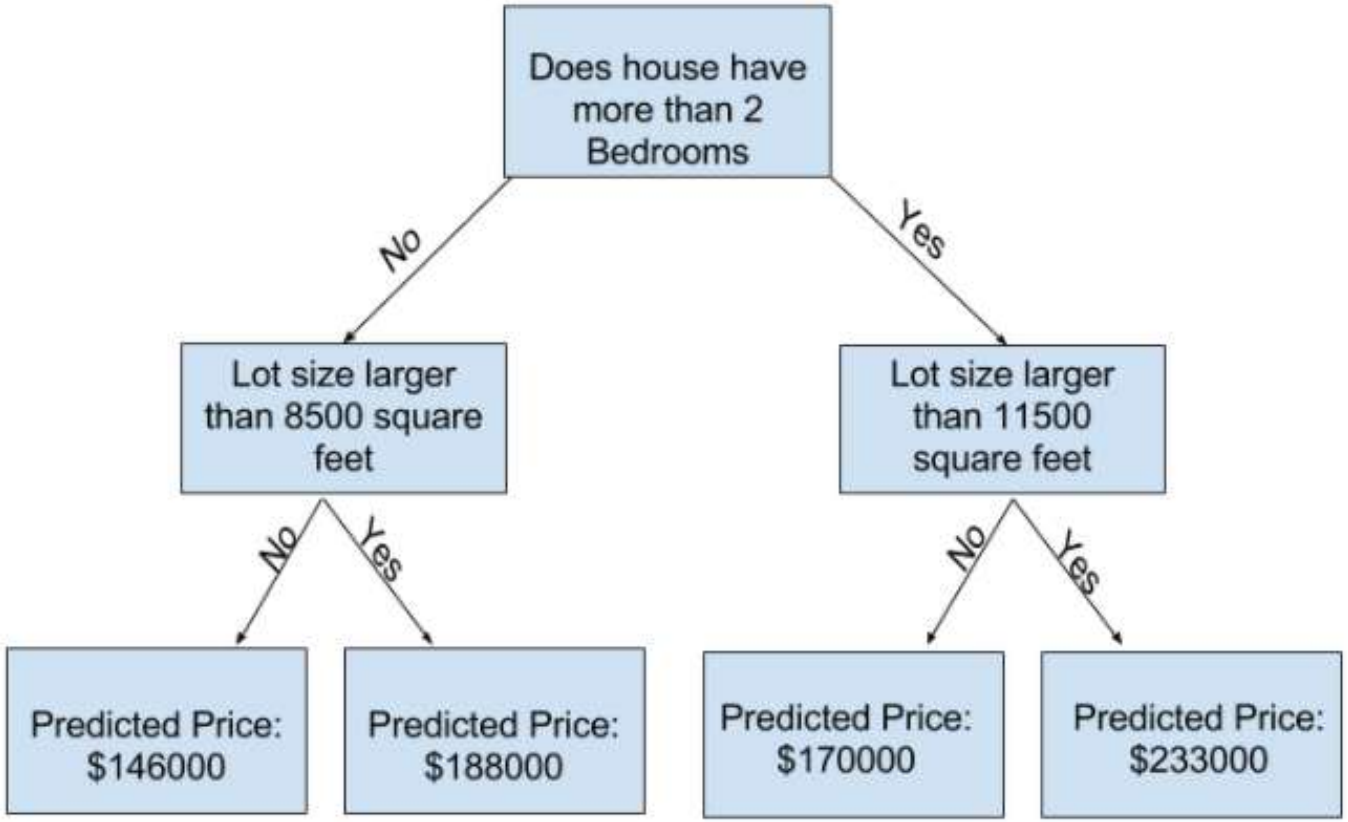
- Random forest, like its name implies, consists of a large number of individual decision trees that operate as an ensemble.
- Each individual tree in the random forest spits out a class prediction and the class with the most votes becomes our model's prediction

BLOCK SCHEMATIC:

PROCESS METHODOLOGY

1. Import data from the dataset
2. Read the data and store data in DataFrame
3. Choose variables/columns, with the columns property of the DataFrame
4. Use the dot notation to select the column we want to predict
5. Select multiple features by providing a list of column names
6. Use the scikit-learn library to create your models.
7. Define a decision tree model and a Random Forest with scikit-learn and fitting it with the features and target variable.

DECISION TREE FLOWCHART



STATUS OF THE PROJECT:

COMPLETED:

- Built a machine Learning model using Decision Tree

YET TO COMPLETE:

- A machine learning model using Decision tree and Random forest.
- Comparison and analysis of the Prediction of Prices using both the training models.

SCOPE OF IMPROVEMENT:

- Use Data Science tools to select the best possible features out of the dataset.
- Analysis of the project using Graphs

THANK YOU!

SANSKRUTI RAUT (ECE14)

FINAL YEAR ECE-C