Random Forest Predictive Multi-Tool
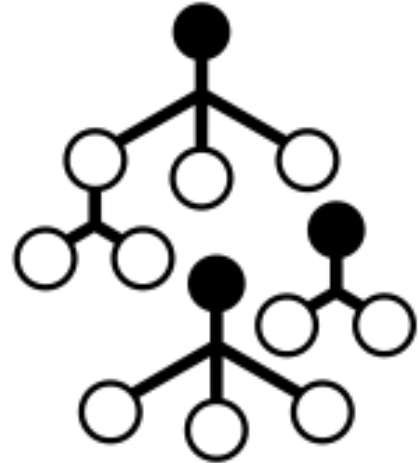
By: Pranav, Dipendra, Mustafa, Jack, Anthony

Platform for Quick Random Forest Analysis

Overall, our idea is to create a easy to use tool for anyone that wants to quickly analyze their data using a random forest model.
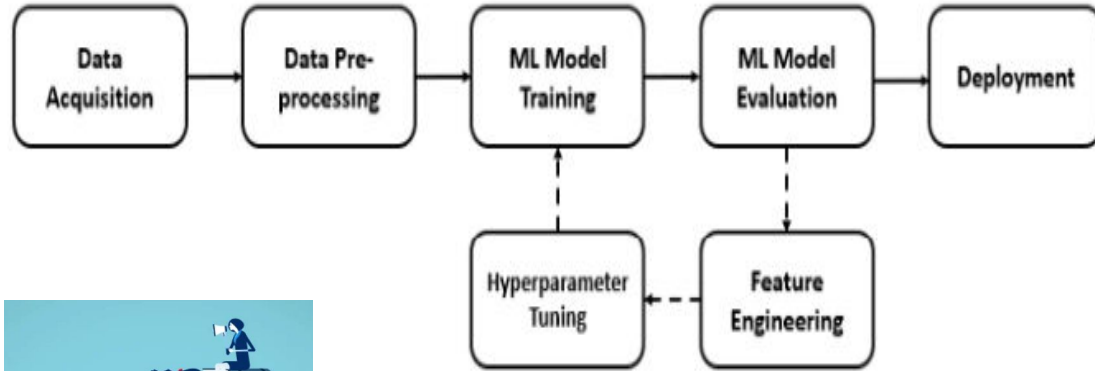
Motivation:

- Prediction model can be important for both "Seller" and "Buyer"

    Informed Decisions

    Determining average pricing

    Market volatility

Abstract-Flow

1. Data Acquisition

2. Data Pre-Processing and Exploratory Data Analysis

3. Creating a Base Model

4. Feature Engineering

5. Hyperparameter Tuning

6. Final Model Training and Evaluation

RFMT

Step 1: Input Training Data

Step 2: Input Test Data

Step 3: Quickly Look at Predicted Results

Libraries used: pandas, numpy, sklearn, ipython, ipywidgets

1. Input Training Data

Users are given a format in which they can plug-in their data to be quickly analyzed.

Expectations:

➔ **Data**

-Data is clean and has relevant features

➔ **Format**

-Data is saved in a .CSV format

➔ **Look at model accuracy score before proceeding**

Type Name of Training Dataset - CSV format

Depth 0

Estimators 0

Generate Model

2. Input Test Data

Users are given a format in which they can plug-in their test data

Expectations:

➜ **Data is formatted to features file format provided**
   -Data is clean and has relevant features

➜ **Pick desired model to be used for predictions**

Type Name of Features File - Should be CSV f

Type Name Model to use for Prediction

Generate Prediction

Data Cleaning Service

Data cleaning lays the groundwork for efficient, accurate and effective data analysis. Without cleaning data beforehand, the analysis process won't be clear or as accurate because the information in the dataset will be unorganized and scattered. Good analysis rests on clean data–it's as simple as that.

RFMT Token

Users wanting to use the Data Cleaning Service will pay for access by purchasing our RFMT token.

Service:

- -Discovering

-   -Structuring

- -Data scrubbing

- -Enriching to augment the data

- -Validating

- -Publishing and allowing for downstream analysis

Data Cleansing

Bad Data

Data Cleansing made simple. Quickly and easily remove data that may distort your analysis.

Thank You For
Listening .