**RANDOM FOREST**

**Business Problem** = ﻿﻿﻿﻿Create a Random Forest Model to classify 'Class Variable'.

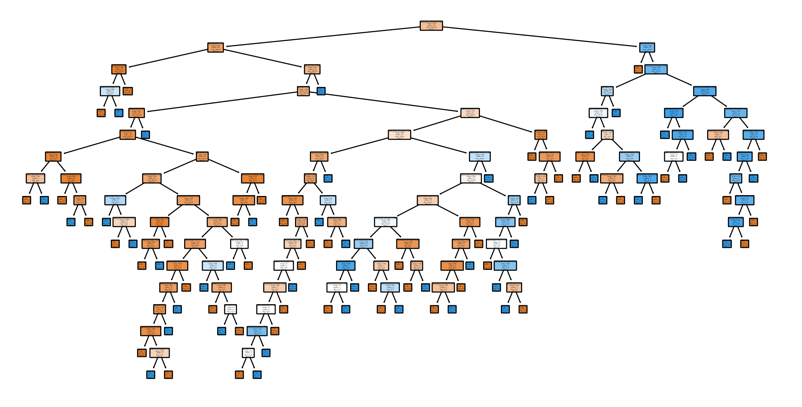
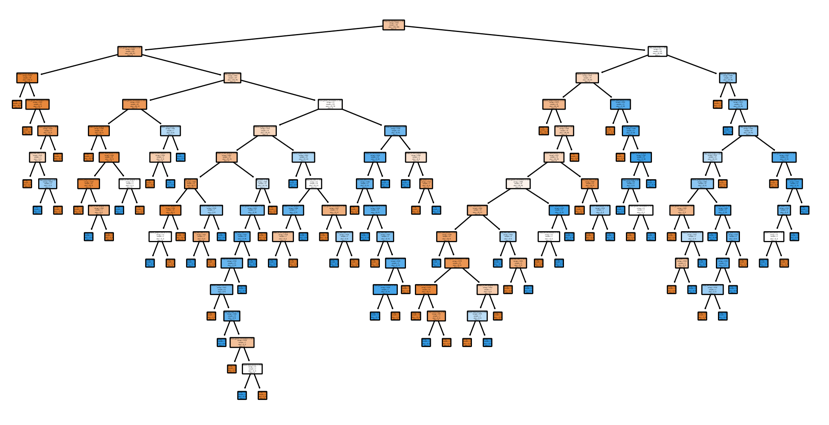
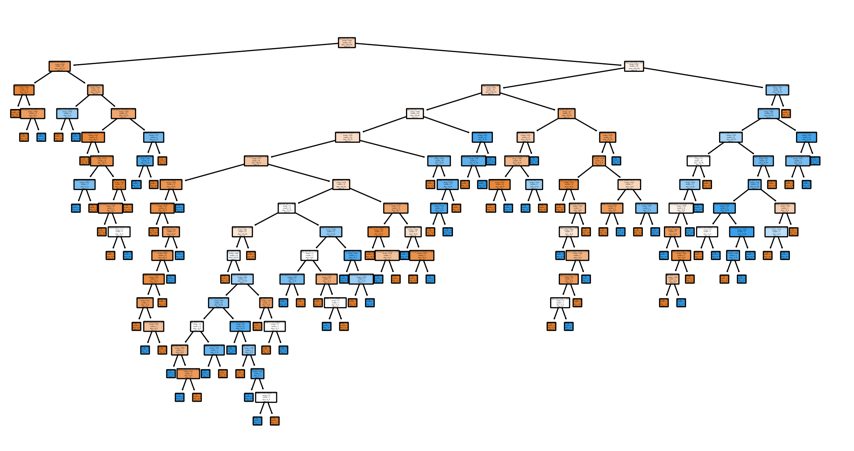
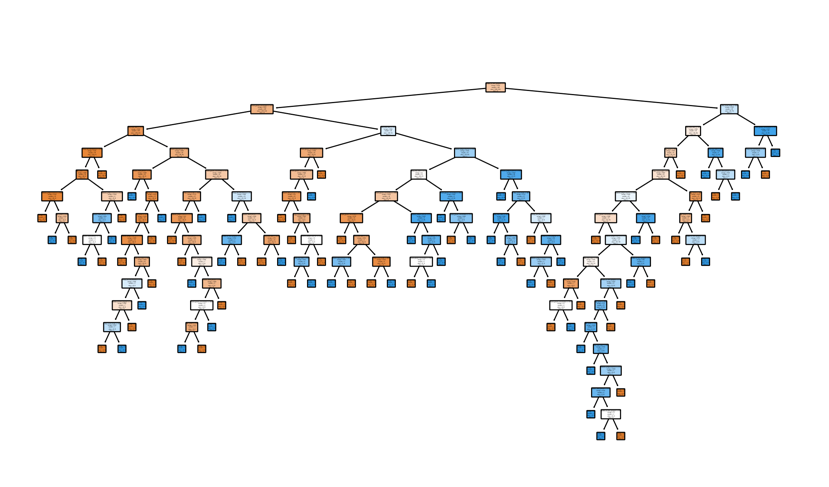
* **Name of the File: -** Diabetes.csv
* **Size of the File: -** 29 KB
* **Necessary Data : -** 768 Observations, 9 Features.

**Exploratory data Analysis** =

* **Outliers: -**  Outliers are not presents.
* **Missing Value: -** Data don’t have Missing Values
* **Output:** - Categorical
* **Sampling:**- Stratified Sampling (65% - 35%)

**Building Random Forest =**

* **﻿Criterion:-** Entropy
* **No. of ﻿Estimators: -** 35
* **Trees: -**



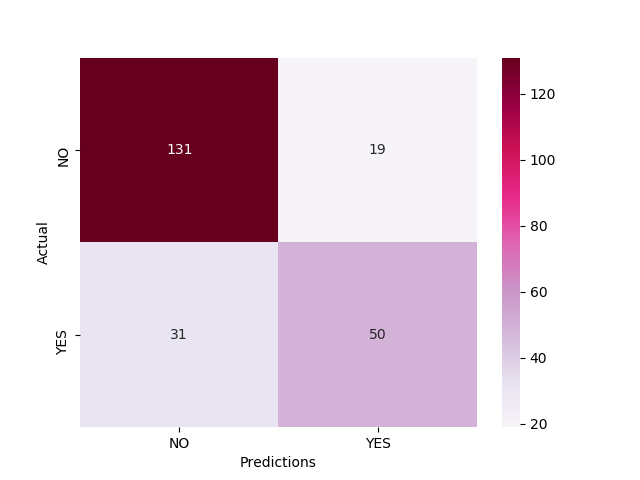
Tree 15

Tree 10

Tree 2

Tree 34

These are four Decision Trees out of 35 from Random forest. We can see any trees from a random forest by using our model.

* **Accuracy Score :-** 77 % - 80 %
* **Confusion Matrix : -**

Accuracies of model on Train and Test data are high So we can use this Random Forest model as Final model for Prediction.

**Python code file**: - [Fraud\_check Analysis.py](https://github.com/nilaydeshmukh0/Decision-Tree/blob/master/Fraud%20Checks%20Analysis/Fraud_check%20Analysis.py)