Cloud Counselage IP program 2020

Data Science Project

Report

Tech Intern – Data Science

User ID: 2916

Krunal Doshi

Email: [krunaldoshi2000@gmail.com](mailto:krunaldoshi2000@gmail.com) Contact no: 8369711415

Data Science Project

# 

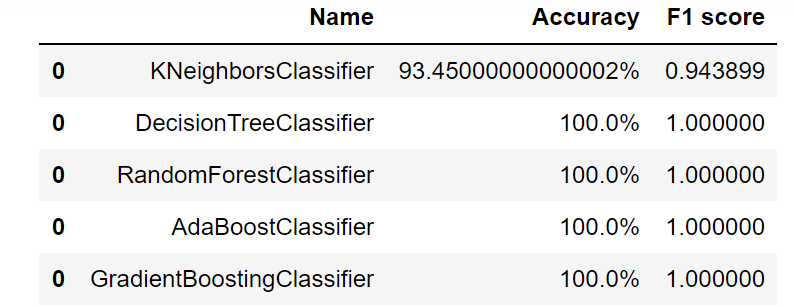
This project was provided by **Cloud Counselage** Internship Program, Students from different cities from the state of Maharashtra had applied for this Internship Program. We were provided with the dataset consisting information of students. We were expected to build a data visualization model and find the best data classification model using the provided dataset.

Under guidance of Mr. Jayanth, Mr. Nirbhey Singh Pahwa and Mr. Harish Sapte, interns were required to complete the tasks mentioned in the problem statement.

**Requirements:**

1. Classification Model
2. Visualization Model

**Conclusion:**



Out of these classifiers:

1. Decision Tree Classifier is very prone to overfits and hence cannot be accepted with 100% accuracy
2. Random Forest Classifier solves the problem with decision trees and hence is not prone to overfits, but it needs a large dataset to function reliably and hence cannot be considered ideal for a dataset with 10000 records
3. Ada Boost Classifier and Gradient Boosting Classifier are both affected by the collinearity of variables present in the dataset and hence cannot be considered ideal.
4. KNN requires high computational time, but deals well with noise in a dataset, making it immune to overfit with an appropriate K-value.

Hence, KNN provides us with the most reliable performance and a good metric score, making it the ideal classifier to use.

**Overall Experience:**

* I started with this Internship Program from May.
* The program started with an induction program.
* After completing the induction program task, I received an offer letter from Clour Counselage as a Tech Intern.
* The Bitrix24 platform was introduced and was used throughout the internship.
* The Data Science Training was provided by Edureka.
* An assignment was given to test our understanding of data science
* The live project started from 15th June 2020 and continued throughout July 2020 with the implementation of the project.
* The final project was submitted on 31st July 2020.

Special Thanks to **Cloud Counselage** for providing the internship program.

Working with Cloud Counselage under guidance of Jayanth Sir, Nirbhey Sir and Harish Sir following all the industry level project management guidelines has given me a whole new experience and has upskilled me, my personality and professionalism.