9. Results

From the Exploratory analysis, Aftertaste of coffee is largely dependent upon these attributes Flavor, Acidity, CupperPoints, Balance, Aroma, Body, Uniformity, Clean Cup, Sweetness, Moisture, and Category Two Defects. When these attributes are visually analyzed, Moisture Category Two Defects, Uniformity, Clean Cup, and Sweetness have minimal impact in Aftertaste. Hence, Flavor, Acidity, CupperPoints, Balance, Body, and Aroma affects the Aftertaste of coffee largely.

A dataframe with these final features are selected along with target variable. This frame is divided into training and testing set. The training set is used in training six different machine learning algorithms. Each model are hyper tuned manually. These trained models are evaluated by both training set and testing set using R-square approach. Most of the model gets overfitted and this problem is minimal in Linear Regression with training accuracy 79.6% and testing accuracy 77.9%.

In second case, we take only four features dropping Aroma and Body. Again, the models are trained and evaluated similarly. The problem of overfitting is minimal in most of the models except Decision Tree. The Support Vector Machine performs better than other. It's training and testing accuracy of is 82.9% and 81.2% respectively and is saved for future purposes.

10. Future Work

The dataset has 44 columns, we've used mainly the quality features to study the factors affecting the Aftertaste. There are many other features of Bean Metadata and Farm metadata. Future enhancements of this project are listed below:

i.   These features can be used to identify the Country, Farm producing the utmost quality Coffee.
ii. The mill processing good quality coffee can identified.
iii. The quality of coffee based upon its color, and processing method can be found.

These future findings might be beneficial for coffee exporters, sellers, and even countries where people consume maximum coffee in average.

11. References

API reference. API reference - seaborn 0.11.1 documentation. (n.d.). https://seaborn.pydata.org/api.html.

User Guide. User Guide - pandas 1.3.0 documentation. (n.d.). https://pandas.pydata.org/docs/user_guide/index.html.

User Guide. scikit. (n.d.). https://scikit-learn.org/stable/user_guide.html.