

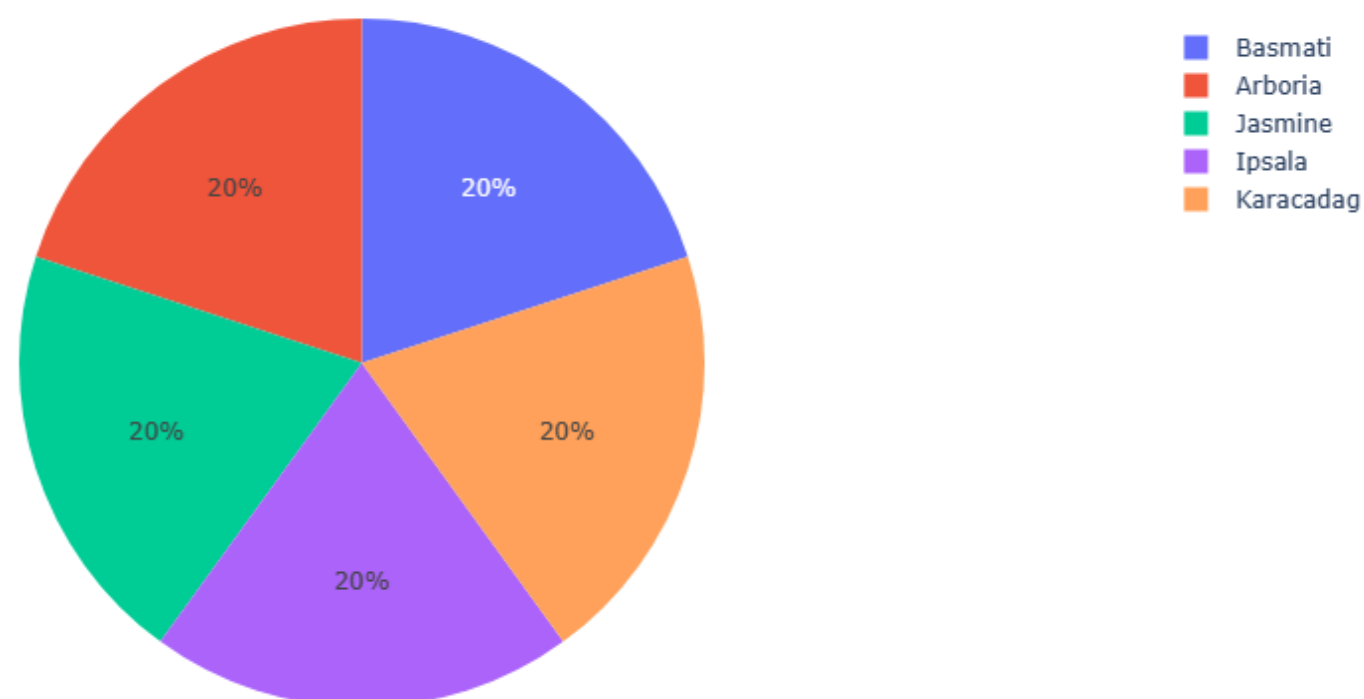
# Agriculture AI-ML Case Study

In the agriculture industry, the classification of rice varieties is a critical task that helps farmers and distributors identify the best types of rice for various uses. Traditionally, this task is performed by expert agronomists who evaluate various characteristics of the rice grain, such as shape, size, and texture. However, this process can be time-consuming, costly, and prone to errors.

Auto-ML techniques offer a promising solution to automate the classification of rice. By analyzing large datasets of rice grain images and their corresponding characteristics, Auto-ML models can learn to distinguish between different varieties of rice with high accuracy. This can significantly reduce the time and cost required for rice classification, while also improving the consistency and reliability of the results.

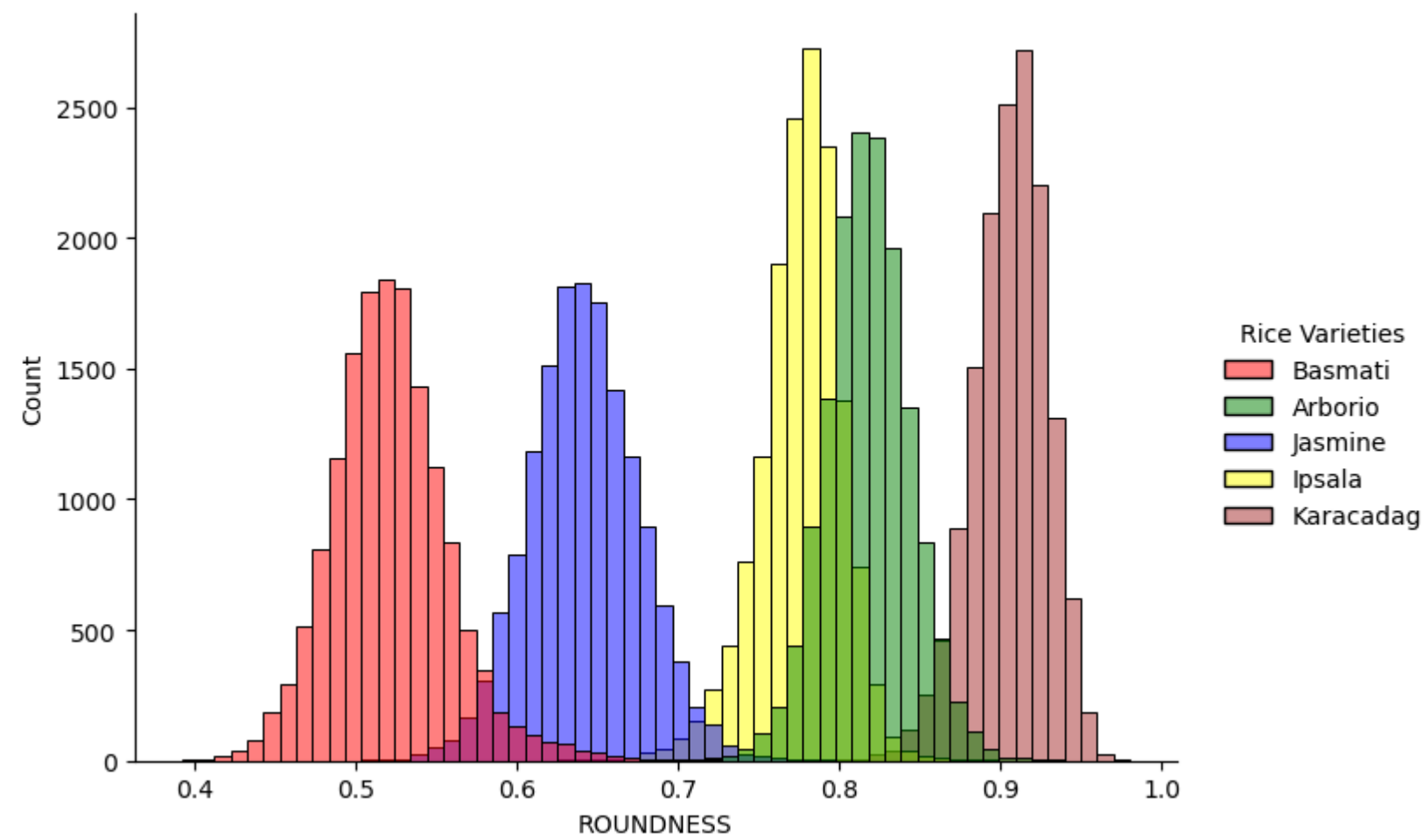
In this study, we aim to develop an Auto-ML model that can accurately classify rice varieties based on their shape, size, texture, and other relevant characteristics. The model will be trained on a large dataset of information containing rice grain varieties and their corresponding labels. Once trained, the model can be used by farmers, distributors, and other stakeholders in the agriculture industry to quickly and accurately identify the best rice varieties for their needs.

Rice Varieties Distribution

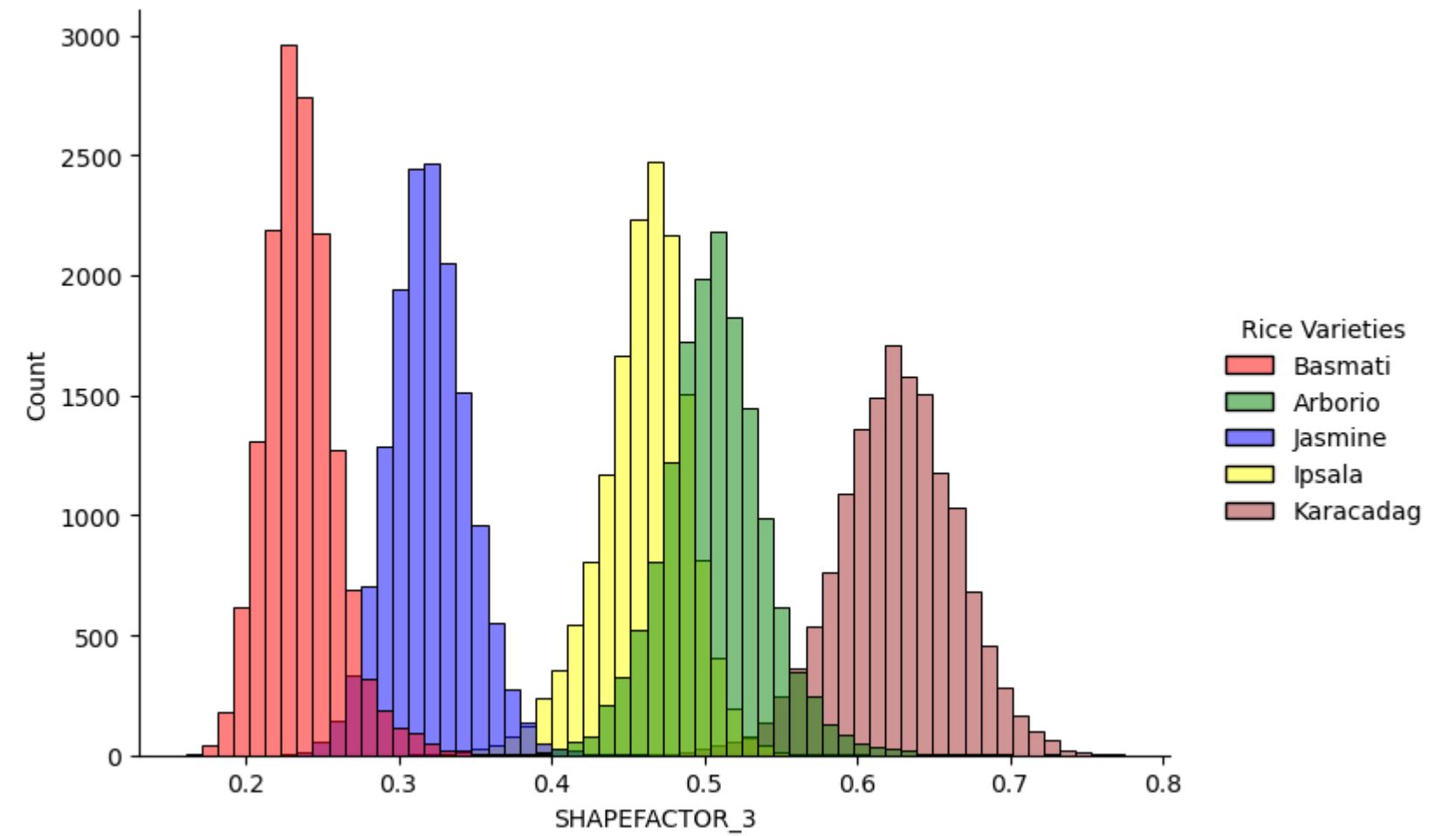


Rice Varieties	Count
Basmati	15000
Arboria	15000
Jasmine	15000
Ipasla	15000
Karacadag	15000

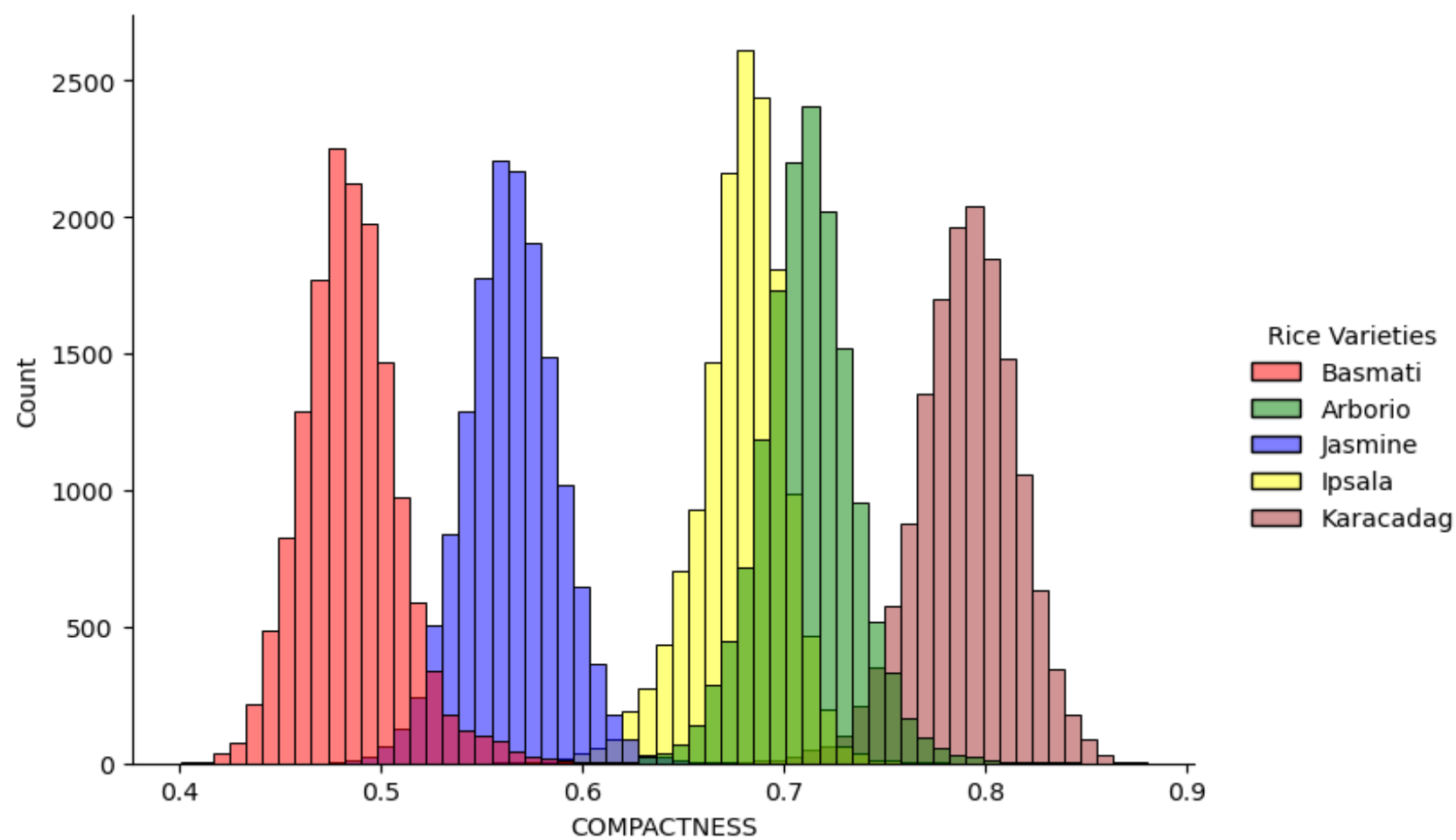
# Features Responsible



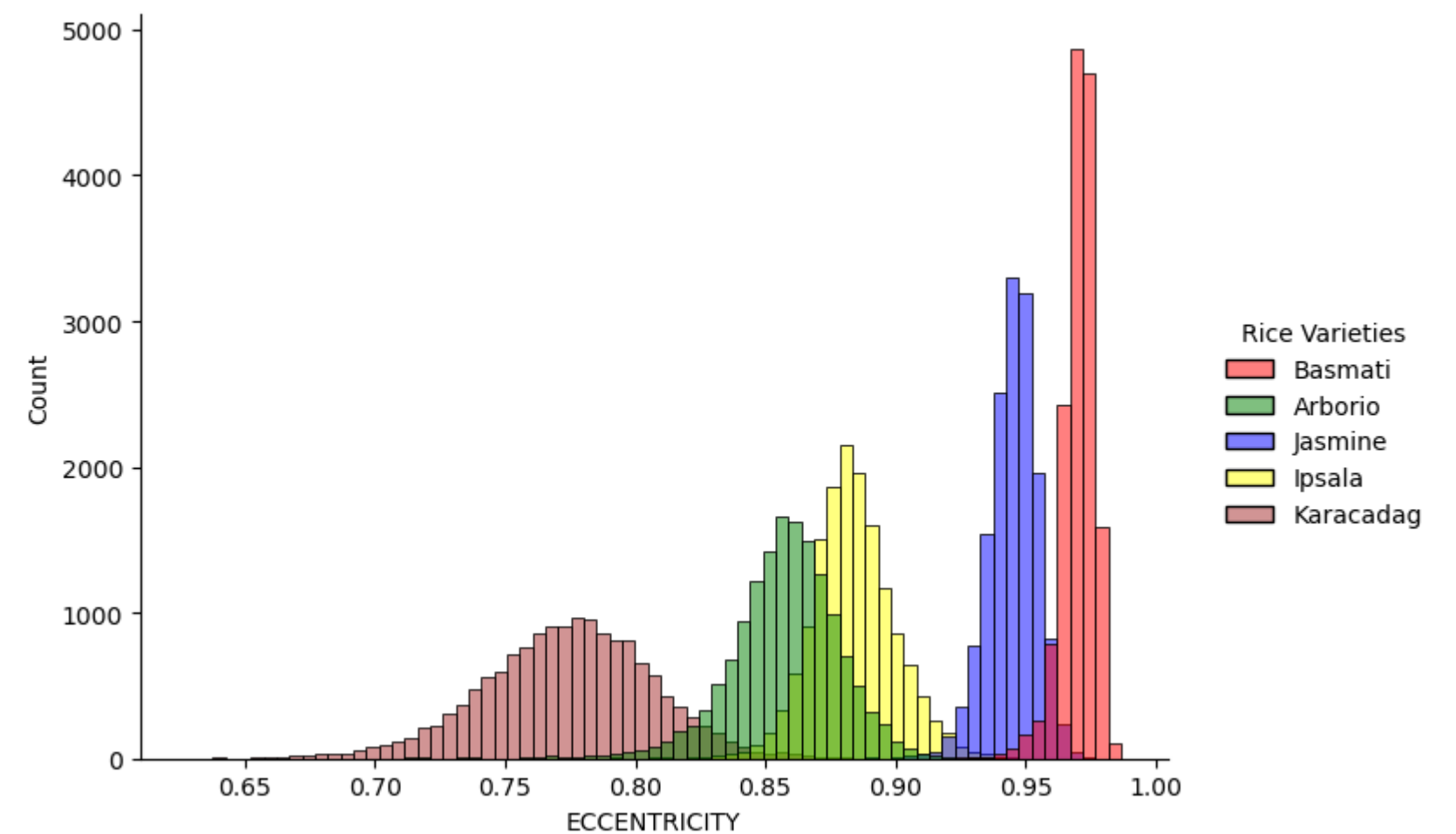
The **roundness** of rice grains is determined by their genetic makeup, and it can have a significant impact on the type of rice that is produced.



The **shape factor**, also known as the length-to-width ratio, is an important characteristic that determines the type of rice produced. The shape factor is calculated by dividing the length of the rice grain by its width.



**Compactness**, also known as the density or hardness of the rice grain, is influenced by various factors such as the genetics of the rice variety, growing conditions, and post-harvest processing methods.



**Eccentricity** refers to the degree of elongation of the rice grain, specifically the ratio of the length of the grain to the width of its central portion.

# Auto-ML Methodology Results

Algorithms	25 percentiles	50 percentiles	75 percentiles	90 percentiles
Decision Tree	99.42	99.71	99.71	99.7
Random Forest	99.95	99.98	99.99	99.99
XGBoost	99.99	99.99	99.99	99.99
MLP	79.77	83.35	82.71	84.43
Total Features	27	53	79	95
<b>Avg. Accuracy in percentile</b>	<b>94.78</b>	<b>95.76</b>	<b>95.6</b>	<b>96.03</b>

- Based on our observation , XGBoost was the best performing algorithm with 99.99% accuracy across all percentile.
- 90<sup>th</sup> percentile is the best percentile with an average accuracy of 96.03%.

# Conclusion

The rice agriculture industry is a crucial sector in many countries, contributing significantly to the economy and food security. It involves the cultivation, harvesting, and processing of rice for consumption and export. The dataset contains different feature extraction-like shapes and features for five different rice varieties of the same brand. The dataset contains 75000 records with 1 Categorical feature and 107 Numerical features. The five different varieties of rice are Arborio , Basmati, Ipsala, Jasmine and Karacadag.

For classification, models were created with algorithms using Auto-ML techniques like Decision tree, Multilayer Perceptron, Random forest and XGBoost . With these models, performance measurement values were obtained for feature sets of 27, 53, 79 and 95. The Auto-ML algorithms were able to predict the of Rice varieties based on their features with an average accuracy between 94% – 97% and helped to identify factors that determine the type of Rice. The major factors include Roundness, Compactness, Shape factor and Eccentricity. When the results are examined, it is observed that with the addition of each new feature, the success of classification increases. Based on the performance measurement values obtained, it is possible to say that the study achieved success in classifying rice varieties.