

## Dataset Description: Rice-Gonen and Jasmine

- This dataset contains **morphological and geometric features of rice grains**, specifically of two rice varieties: **Gonen** and **Jasmine**. The dataset is likely used for **classification tasks** to differentiate between these two rice types based on their physical properties. **Features:** Size, shape, and geometric properties of rice grains.
- The dataset originally contained **18,185** samples. After preprocessing and filtering (handling missing values, removing outliers, and duplicates), the final dataset contains **17,490** samples.

Feature	Description
Area	Total pixel area of the rice grain.
Perimeter	The total boundary length of the grain.
MajorAxisLength	Length of the longest axis of the grain.
MinorAxisLength	Length of the shortest axis of the grain.
Eccentricity	Measures how elongated the grain is (0 to 1).
ConvexArea	Area of the smallest convex shape enclosing the grain.
EquivDiameter	Diameter of a circle with the same area as the grain.
Extent	Ratio of Area to its bounding box area.
Solidity	Ratio of Area to ConvexArea (compactness measure).
Roundness	Circularity measure: $4 * \text{Area} / (\pi * \text{MajorAxisLength}^2)$ .
AspectRatio	Ratio of MajorAxisLength to MinorAxisLength.
Compactness	Shape uniformity measurement.
Class	Target variable: 0 (Gonen) or 1 (Jasmine)