Classification of dry beans

Question/need: Dry beans are among the most edible legumes in the world. A uniform seed variety is one of the quality factors for dry beans. Therefore, a crop production Company (Company X) requires an automated system that classifies the harvested seeds with the same size, shape , form and structure in the same classes during storage and packaging to provide uniform varieties of seeds to the market.

This project is designed to create a classification model that classifies seeds into 7 Registered bean groups that will be used with a computer vision system in the crop production line.

Data Description: Data source : UCI machine learning repository

Data points : 13,611
Features : 16
1.) Area (A): The area of a bean zone and the number of pixels within its boundaries.
2.) Perimeter (P): Bean circumference is defined as the length of its border.
3.) Major axis length (L):Distance between the ends of the longest line that can be drawn from a bean.
4.) Minor axis length (l): The longest line that can be drawn from the bean while standing perpendicular to the main axis.
5.) Aspect ratio (K): Defines the relationship between L and l.
6.) Eccentricity (Ec): Eccentricity of the ellipse having the same moments as the region.
7.) Convex area (C): Number of pixels in the smallest convex polygon that can contain the area of a bean seed.
8.) Equivalent diameter (Ed): The diameter of a circle having the same area as a bean seed area.
9.) Extent (Ex): The ratio of the pixels in the bounding box to the bean area.
10.)Solidity (S): Also known as convexity. The ratio of the pixels in the convex shell to those found in beans.
11.)Roundness (R): Calculated with the following formula: (4piA)/(P^2)
12.)Compactness (CO): Measures the roundness of an object: Ed/L
13.)ShapeFactor1 (SF1)
14.)ShapeFactor2 (SF2)
15.)ShapeFactor3 (SF3)
16.)ShapeFactor4 (SF4)
Target : 7 Classes
Class (Seker, Barbunya, Bombay, Cali, Dermosan, Horoz and Sira)

Tools:

- Python
- Tableau

MVP Goal: A logistic regression Baseline model