

Q 3: Write a paragraph about your experience of working with the standard ML pipeline in your own words.

Experience with the ML Pipeline in WEKA:

The process start with preparing and loading the ARFF file into WEKA's workbench. The ARFF file have the attributes and the dataset. Running the J48 classification algorithm, a decision tree implementation, provided clear insights in the dataset. The algorithm created a tree with 5 leaves and a size of 9, signifying a relatively simple model. The key decision splits were based on gender, name_length, and consonant_count, reflecting the significance of these features in classification.

Results Analysis:

The model achieved a 65% accuracy, correctly classifying 65 out of 100 instances. Additionally, the precision and recall for class 0 (female names) were higher than for class 1 (male names), with a TP (True Positive) rate of 0.776 for class 0 compared to 0.476 for class 1. This suggests that the model performed better at identifying female names than male names, possibly due to an imbalance in the dataset or overlapping feature values.