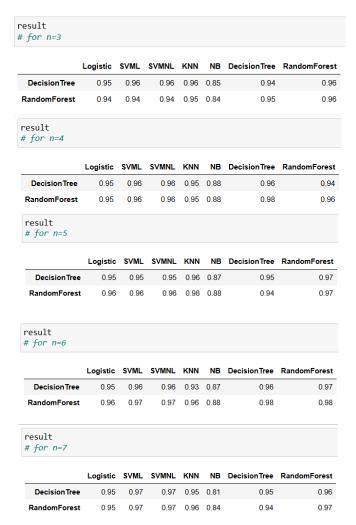
FEATURE IMPORTANCE

Feature Importance is an attribute, present in the tree based models like, **Decision Tree, Random Forest, and Gradient Boosting Classifier**. Since the feature importance from these models are impurity based importance, it can be biased towards highly unique value features. Hence we use **PERMUTATION FEATURE IMPORTANCE**.

Permutation Feature Importance is provided by scikit-learn library that uses agnostic technique and fits any estimator. This PFI works by randomly shuffling the values present in a single feature and checks the score. The larger the score decreases, higher the importance. This takes place for every single features present in the dataset.

The accuracy result for n number of important features fetched from Permutation Feature Importance is given below.



Conclusion:

From the above displayed tables, the difference in accuracy score is not positively huge as number of features is increased from n = 3 to 7. So we can finalize the number of features required for CKD prediction obtained from PFI is n=4