

## Final Results:

The summary of accuracies using all algorithms on the Wine dataset is presented below in a tabular form:

Algorithm	Setting	Wine - Raw	Wine PCA (PC=3)	Wine LDA (Des=3)
KNN	K=40 Weighted	98.84	88.92	99.92
Naïve Bayes	Multinomial NB	44.62	44.54	44.54
Naïve Bayes	Complement NB	48.54	46.92	0.15
Decision Tree		53.20	50.34	54.60
Random Forest		51.64	43.42	54.22
Gradient Boosting		59.92	54	55.3

The summary of accuracies for all algorithms applied on the Abalone dataset is presented below in a tabular form:

Algorithm	Setting	Abalone - Raw	Abalone PCA (PC =3)	Abalone LDA (Des=3)
KNN	K=68 Weighted	27.05	62.67	69.61
Naïve Bayes	Multinomial NB	16.99	16.99	16.99
Naïve Bayes	Complement NB	19.14	17.22	21.53
Decision Tree		26.74	25.83	25.85
Random Forest		27.79	26.23	26.93
Gradient Boosting		25.47	11.48	21.05

### Observations (Wine Dataset):

- Based on the table provided, it appears that the Wine LDA (Des=3) algorithm outperforms the other algorithms in terms of classification accuracy. The KNN algorithm also performs well, but its performance decreases significantly when using PCA. The Naive Bayes algorithms perform poorly, with the Multinomial NB algorithm having a slightly higher accuracy than the Complement NB algorithm. The Decision Tree, Random Forest, and Gradient Boosting algorithms all have similar accuracies, with the Decision Tree and Random Forest algorithms having slightly higher accuracies than the Gradient Boosting algorithm.
- Of all the algorithms, KNN has produced the best accuracy results. PCA with KNN has given less accuracy but that may have created a more generalized model, but LDA and KNN has given close to 100% accuracy.
- LDA and KNN has led to better accuracies compared to PCA as the separation between classes is handled better with LDA and there is some loss of information after using PCA.

- In the case of the Wine dataset, LDA performs better than PCA because LDA is specifically designed for classification problems and considers the class labels of the data points. The Wine dataset has three classes, and LDA can find the best linear combinations of features that maximize the separation between these classes. PCA, on the other hand, does not consider the class labels and may not be able to find the best combinations of features for classification. Therefore, LDA is a better choice for the Wine dataset.

#### Observations (Abalone Dataset):

- KNN with dimensionality reduction has provided the best results of all the other algorithms.
- Based on the table provided, it appears that the Abalone LDA (Des=3) algorithm outperforms the other algorithms in terms of classification accuracy. The KNN algorithm also performs relatively well, especially when compared to the other algorithms. The Naive Bayes algorithms perform poorly, with the Multinomial NB algorithm having the same accuracy as the other Naive Bayes algorithm, and both having a significantly lower accuracy than the other algorithms. The Decision Tree, Random Forest, and Gradient Boosting algorithms all have similar accuracies, with the Decision Tree and Random Forest algorithms having slightly higher accuracies than the Gradient Boosting algorithm.
- It is worth noting that PCA does not seem to provide much benefit in this case, as the accuracies of the algorithms do not improve much when using PCA. This may be because the Abalone dataset is already relatively low-dimensional. Therefore, using PCA to reduce the dimensionality of the dataset does not provide much additional information to the algorithms.