# Comparative Analysis of Traditional Machine Learning Algorithms and Neural Networks

**Dataset Name:** pd_speech_features_mod.csv

**Testing Procedure:**

- **Data Preprocessing:**

  To make the data optimal for our machine learning algorithms to process we first did the following processes:-

  - Skewness Reduction: Skewness is undesirable for any ML algorithm as it leads to excessive amount of variance in the final estimate. We used Box-Cox, Cube Root and Log Transformation to reduce skewness in our dataset.

  - Kurtosis Reduction: Kurtosis is the measure of whether the data is heavily left or right tailed which may lead to outliers being classified as output and the output estimate to have higher distance from the mean.

  - Outlier Detection: We also checked for outliers in the dataset which we then replaced with the mean value for that attribute.

  - Principal Component Analysis: We did Principal Component Analysis to identify the strongly correlated attributes and reduce the dimension of the dataset.

- **Data Testing:**

  After Preprocessing we ran the dataset through the following ML and neural network algorithms:

  - KNN Classifier
  - SVM Classifier
  - Decision Tree Classifier
  - Random Forest Classifier
  - Naive Bayes Classifier
  - Logistic Regression Classifier
  - XGBoost Classifier
  - Gradient Boost Classifier

- ■ ANN Classifier
- ■ MLP Classifier

After Running each algorithm we accrued their respective accuracy, recall and F1-Score Values. Which are presented in the Table below:

| Name | Accuracy | Precision | | Recall | | F1-Score | |
|---|---|---|---|---|---|---|---|
| | | Class A (0) | Class B (1) | Class A (0) | Class B (1) | Class A (0) | Class B (1) |
| KNN Classifier | 96% | 1 | 0.95 | 0.86 | 1 | 0.92 | 0.97 |
| SVM Classifier | 88% | 1 | 0.86 | 0.57 | 1 | 0.73 | 0.92 |
| Decision Tree Classifier | 87% | 0.79 | 0.89 | 0.71 | 0.93 | 0.75 | 0.91 |
| Random Forest Classifier | 91% | 1 | 0.89 | 0.67 | 1 | 0.8 | 0.94 |
| Naïve Bayes Classifier | 89% | 1 | 0.87 | 0.62 | 1 | 0.76 | 0.93 |
| Logistic Regression Classifier | 88% | 1 | 0.86 | 0.57 | 1 | 0.73 | 0.92 |
| XGBoost Classifier | 89% | 0.88 | 0.9 | 0.71 | 0.96 | 0.79 | 0.93 |
| Gradient Boost Classifier | 87% | 0.82 | 0.88 | 0.67 | 0.95 | 0.74 | 0.91 |
| | | | | | | | |
| ANN Classifier | 100% | 1 | 1 | 1 | 1 | 1 | 1 |
| MLP Classifier With PCA | 96% | 0.88 | 1 | 1 | 0.94 | 0.94 | 0.97 |