GENDER RECOGNITION BY VOICE

Team Skepsis

Sagnik Mitra
Spandan Pal
Sneharup Mukherjee

OBJECTIVE

In this Project we will see how to classify speech on gender basis, by using 7 different algorithms and to determine which algorithm gives the highest accuracy score.

INTRODUCTION

Overview of the Project and summary of the Libraries used.

LIBRARIES USED IN THE MODEL

Pandas

1

Open-source Python library providing high-performance, easy-to-use data structures and data analysis tools used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc

Numpy

2

Library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

Matplotlib

3

Cross-platform library for making 2D plots from data in arrays. It provides an object-oriented API that helps in embedding plots in applications using Python GUI toolkits such as PyQt, WxPython Tkinter.

Sklearn

4

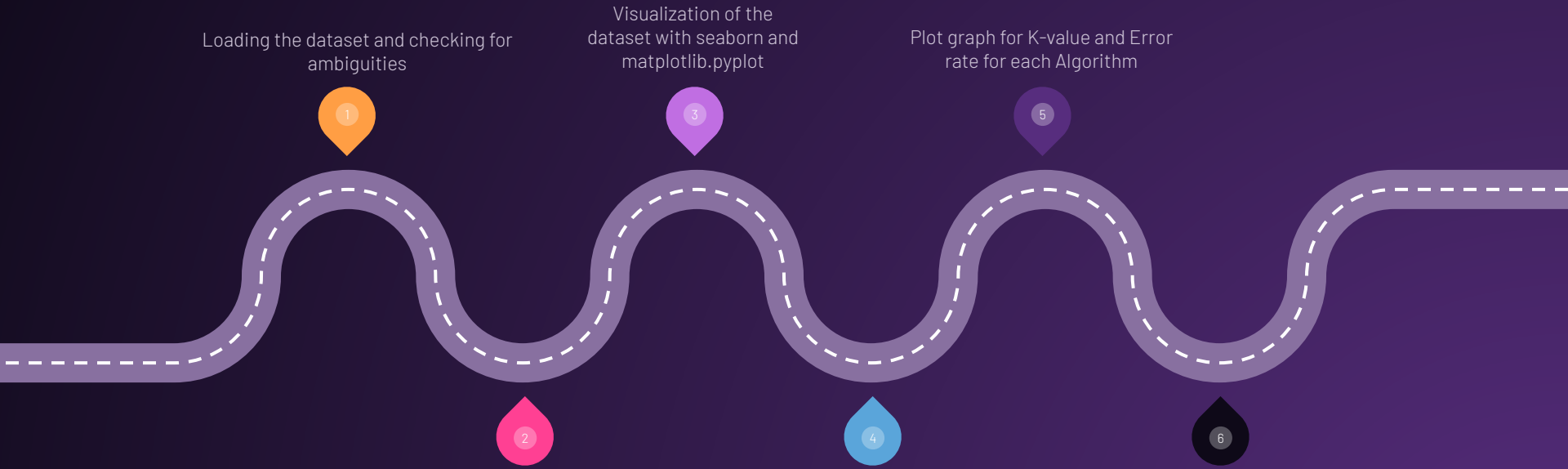
Provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistent interface in Python.

Seaborn

5

Data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

ROADMAP


1
Loading the dataset and checking for ambiguities
2
Preprocessing the data by label encoding and normalization
3
Visualization of the dataset with seaborn and matplotlib.pyplot
4
Training the model with 2 different datasets, one with all attributes and another with only 3 main attributes applying 7 different algorithms
5
Plot graph for K-value and Error rate for each Algorithm
6
Finding the accuracy score for each Algorithm and determining the best of them.

Application

A gender classification system uses face or voice of a person from a given image or audio to tell the gender (male/female) of the given person. A successful gender classification approach can boost the performance of many other applications including face recognition and smart human-computer interface.

About the Dataset

Voice Gender

This database was created to identify a voice as male or female, based upon acoustic properties of the voice and speech. The dataset consists of 3,168 recorded voice samples, collected from male and female speakers.

The voice samples are pre-processed by acoustic analysis in R using the seewave and tuneR packages, with an analyzed frequency range of 0hz-280hz (human vocal range).

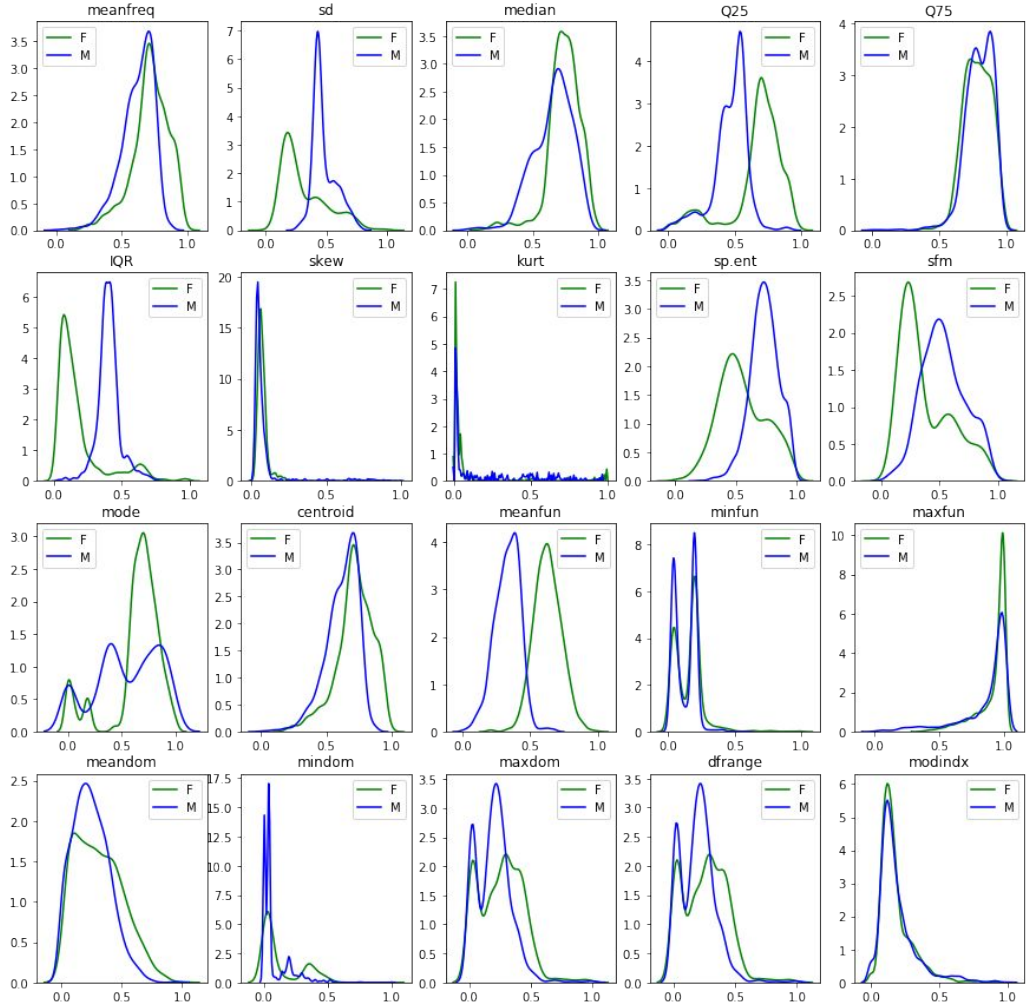
The Dataset

Size of the Dataset: 1.2 GB

The following acoustic properties of each voice are measured and included within the CSV

- meanfreq: mean frequency (in kHz)
- sd: standard deviation of frequency
- median: median frequency (in kHz)
- Q25: first quantile (in kHz)
- Q75: third quantile (in kHz)
- IQR: interquartile range (in kHz)
- skew: skewness (see note in specprop description)
- kurt: kurtosis (see note in specprop description)
- sp.ent: spectral entropy
- sfm: spectral flatness
- mode: mode frequency
- centroid: frequency centroid (see specprop)
- peakf: peak frequency (frequency with highest energy)
- meanfun: average of fundamental frequency measured across acoustic signal
- minfin: minimum fundamental frequency measured across acoustic signal
- maxfun: maximum fundamental frequency measured across acoustic signal
- meandom: average of dominant frequency measured across acoustic signal
- mindom: minimum of dominant frequency measured across acoustic signal
- maxdom: maximum of dominant frequency measured across acoustic signal
- dfrange: range of dominant frequency measured across acoustic signal
- modindex: modulation index. Calculated as the accumulated absolute difference between adjacent measurements of fundamental frequencies divided by the frequency range

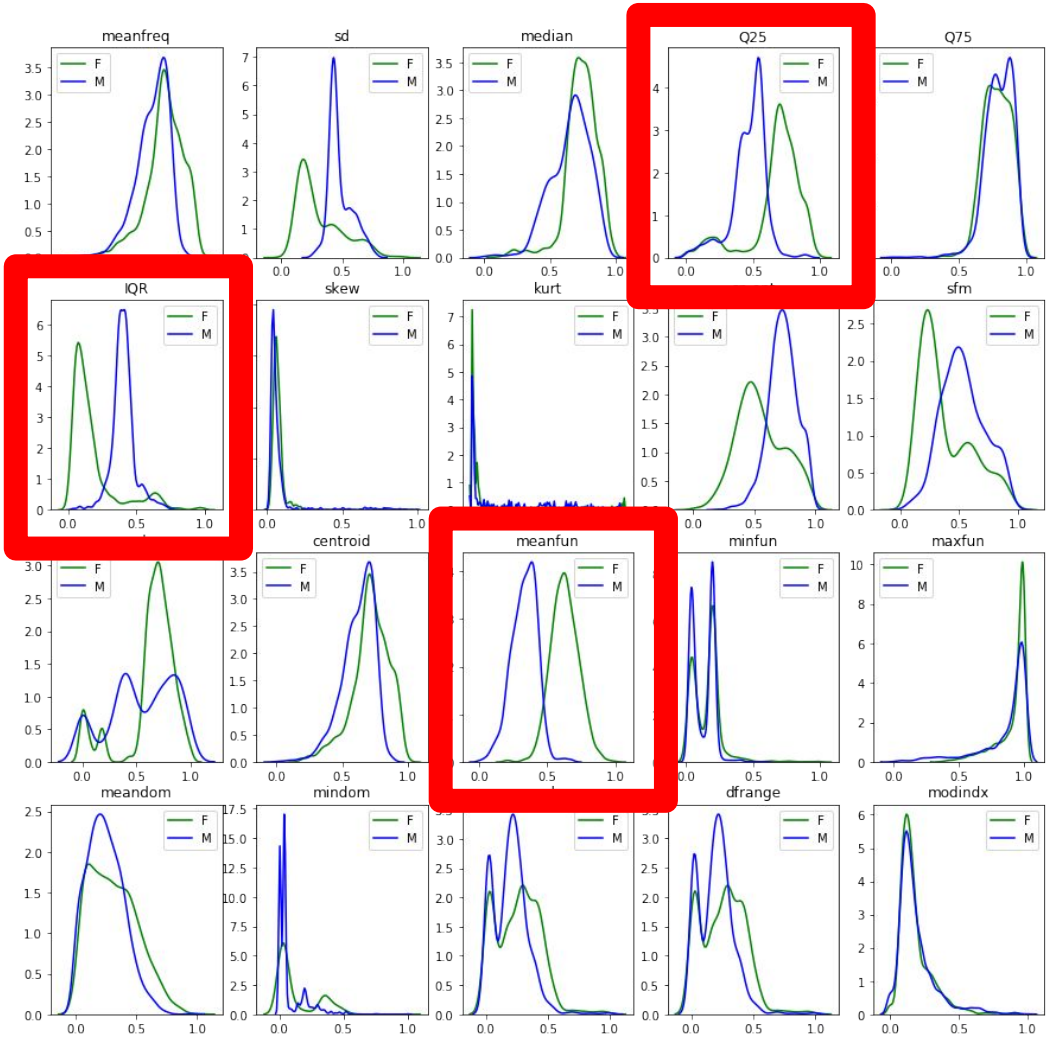
VISUALIZATION OF DIFFERENT ACOUSTIC PROPERTIES



LINK TO THE NOTEBOOK

ACOUSTIC PROPERTIES CHOSEN WITH RESPECT TO THEIR PEAK VALUE AND POSITION IN THE GRAPH:

IQR, Q25, MEANFUN

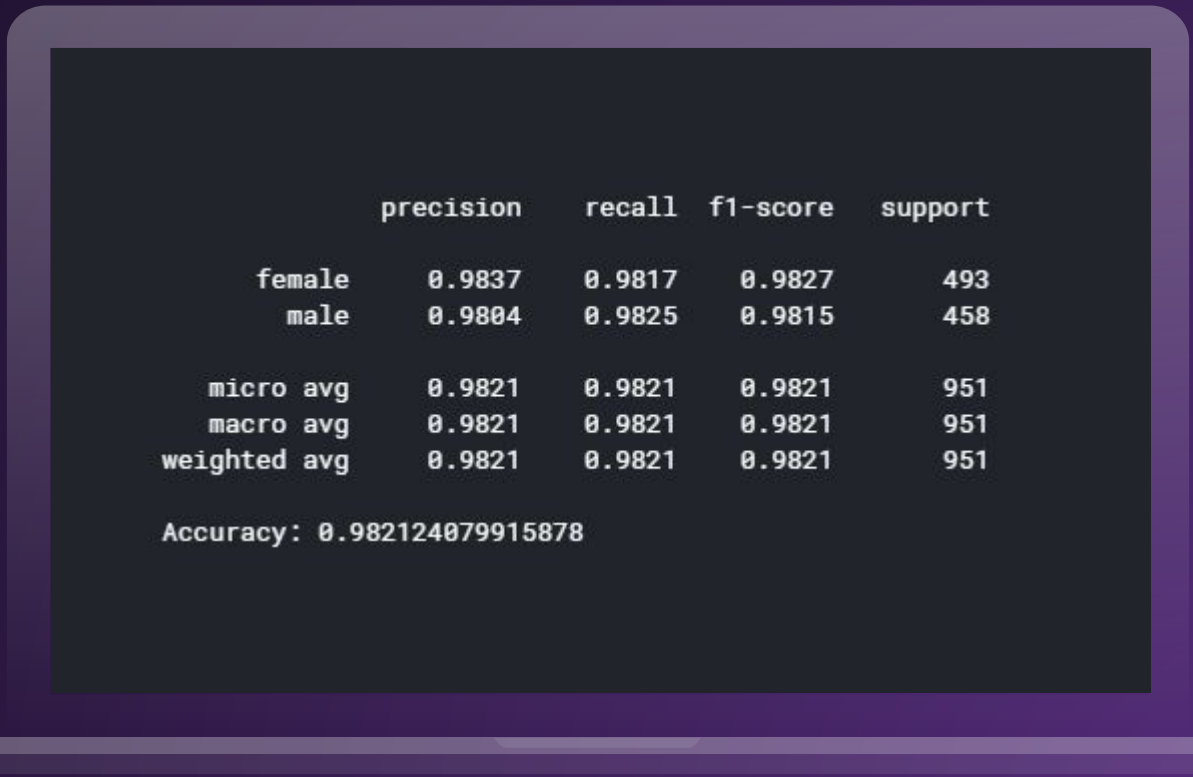


Results and Error-Rate Plot of Algorithms Used in the Model

K Nearest Neighbours

K-Nearest Neighbors Precision, F1 Score, Recall, Accuracy, Support

All 20 Attributes



	precision	recall	f1-score	support
female	0.9837	0.9817	0.9827	493
male	0.9804	0.9825	0.9815	458
micro avg	0.9821	0.9821	0.9821	951
macro avg	0.9821	0.9821	0.9821	951
weighted avg	0.9821	0.9821	0.9821	951

Accuracy: 0.982124079915878

K-Nearest Neighbors Precision, F1 Score, Recall, Accuracy, Support

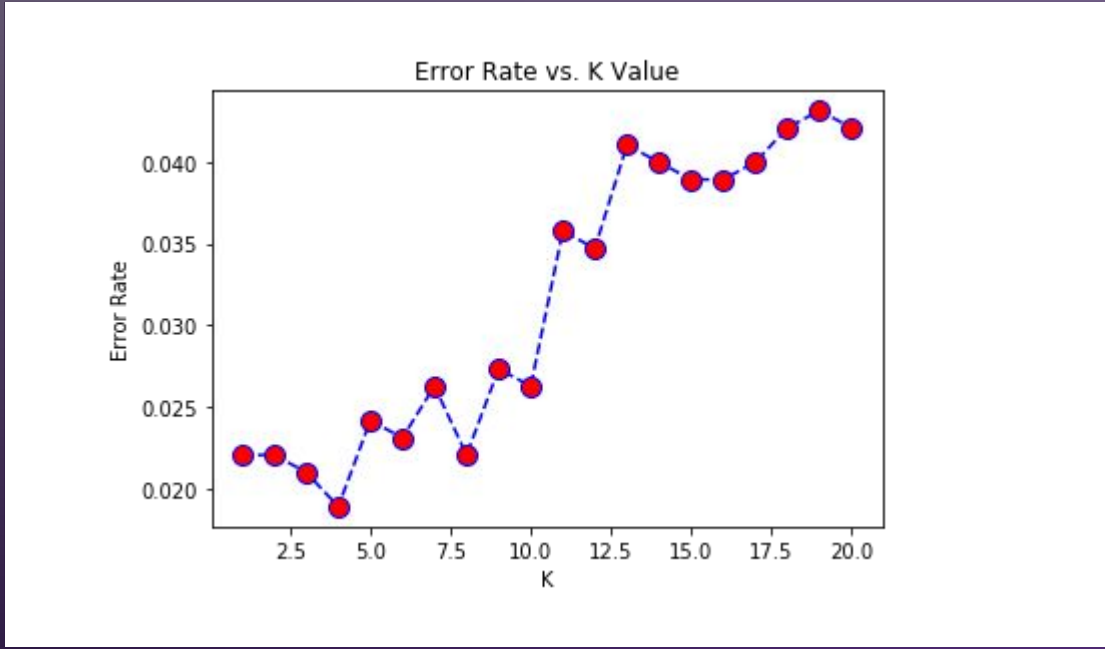
Major 3 Attributes
(IQR, MEANFUN, Q25)

	precision	recall	f1-score	support
female	0.9642	0.9838	0.9739	493
male	0.9821	0.9607	0.9713	458
micro avg	0.9727	0.9727	0.9727	951
macro avg	0.9732	0.9722	0.9726	951
weighted avg	0.9728	0.9727	0.9726	951

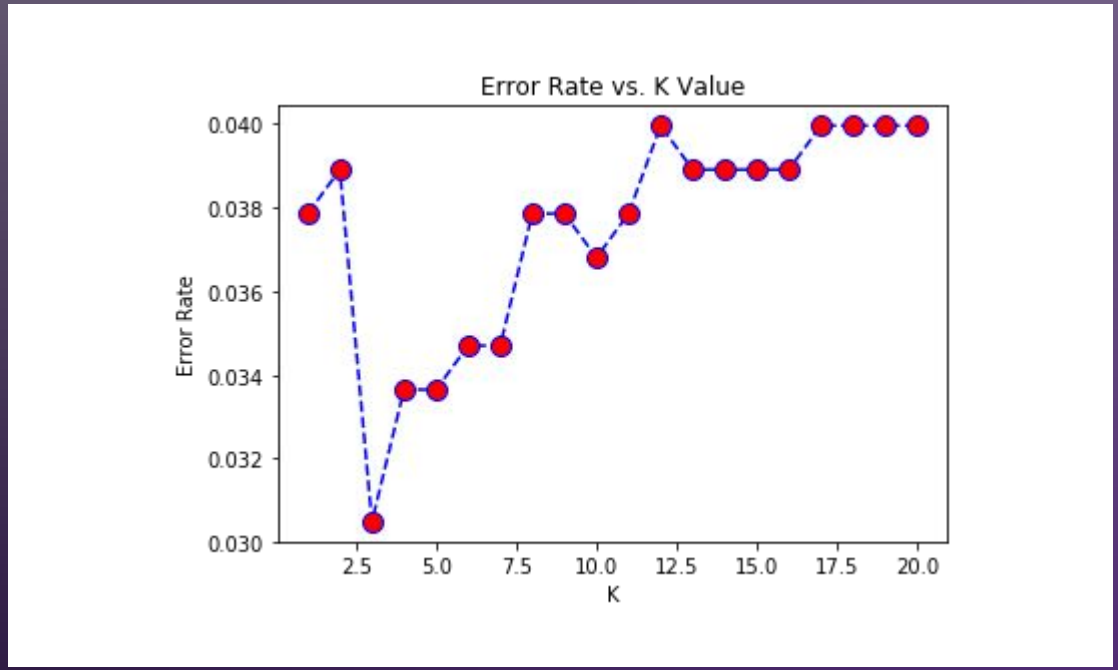
Accuracy: 0.9726603575184016

K-Nearest Neighbors Error Rate Vs K-Value Plot

All 20 Attributes



K-Nearest Neighbors Error Rate Vs K-Value Plot



Major 3 Attributes
(IQR, MEANFUN, Q25)

Naive Bayes

Naive Bayes

Precision, F1 Score, Recall, Accuracy, Support

All 20 Attributes

	precision	recall	f1-score	support
female	0.8970	0.9006	0.8988	493
male	0.8925	0.8886	0.8906	458
micro avg	0.8948	0.8948	0.8948	951
macro avg	0.8948	0.8946	0.8947	951
weighted avg	0.8948	0.8948	0.8948	951

Accuracy: 0.8948475289169295

Naive Bayes

Precision, F1 Score, Recall, Accuracy, Support

Major 3 Attributes

(IQR, MEANFUN, Q25)

	precision	recall	f1-score	support
female	0.9735	0.9696	0.9715	493
male	0.9674	0.9716	0.9695	458
micro avg	0.9706	0.9706	0.9706	951
macro avg	0.9705	0.9706	0.9705	951
weighted avg	0.9706	0.9706	0.9706	951

Accuracy: 0.9705573080967402

Random Forest

Random Forest

Precision, F1 Score, Recall, Accuracy, Support

All 20 Attributes

	precision	recall	f1-score	support
female	0.9759	0.9858	0.9808	493
male	0.9845	0.9738	0.9791	458
micro avg	0.9800	0.9800	0.9800	951
macro avg	0.9802	0.9798	0.9800	951
weighted avg	0.9801	0.9800	0.9800	951

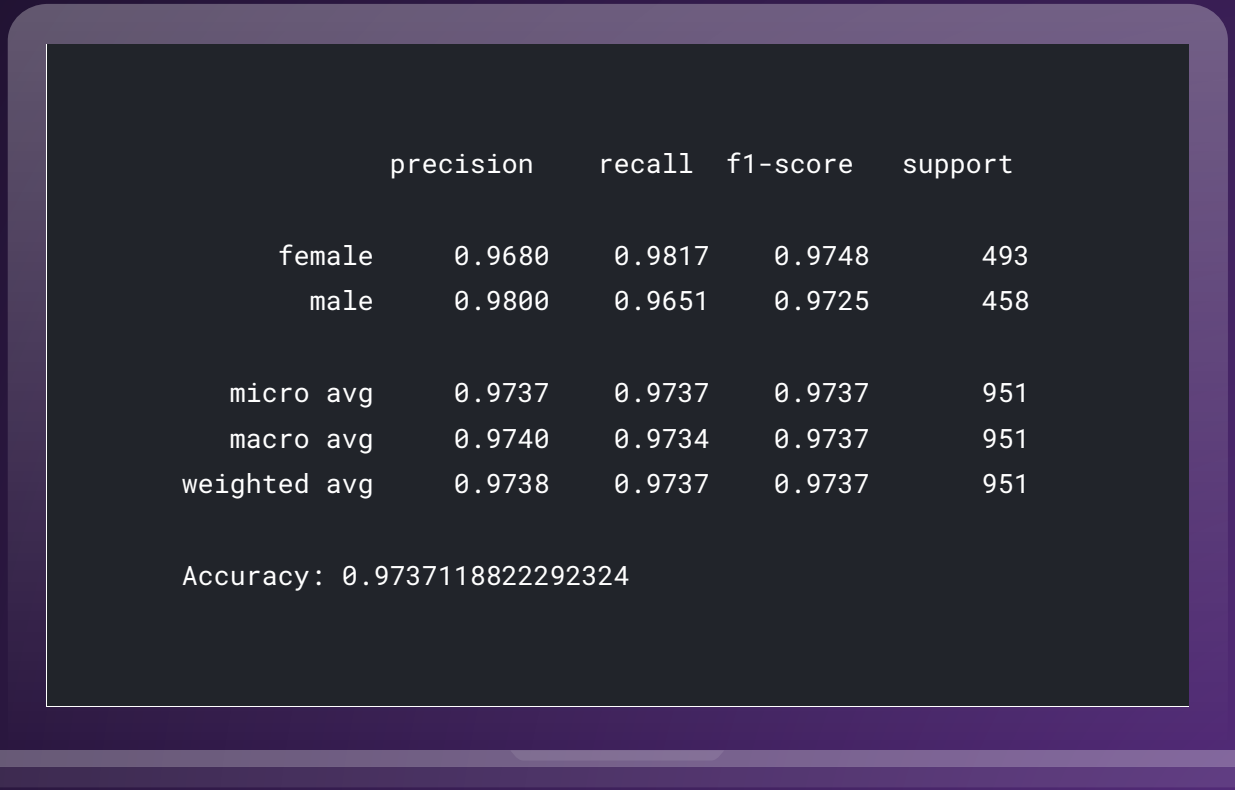
Accuracy: 0.9800210304942166

Random Forest

Precision, F1 Score, Recall, Accuracy, Support

Major 3 Attributes

(IQR, MEANFUN, Q25)



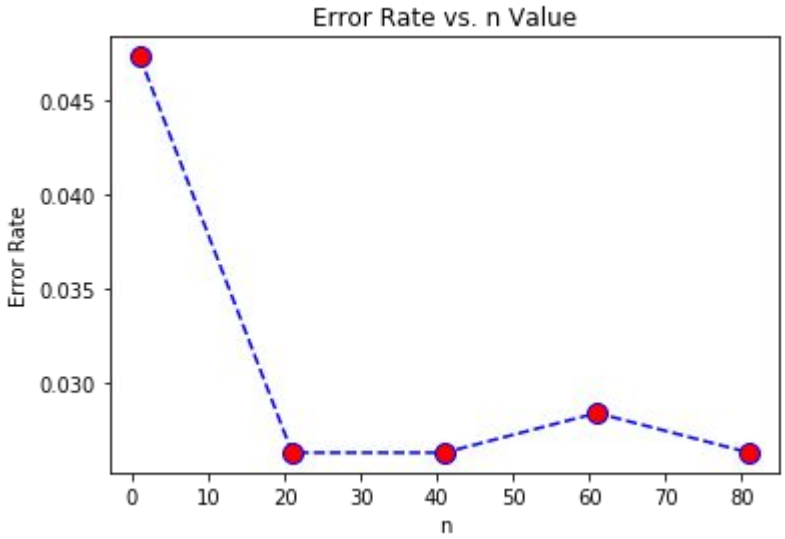
	precision	recall	f1-score	support
female	0.9680	0.9817	0.9748	493
male	0.9800	0.9651	0.9725	458
micro avg	0.9737	0.9737	0.9737	951
macro avg	0.9740	0.9734	0.9737	951
weighted avg	0.9738	0.9737	0.9737	951

Accuracy: 0.9737118822292324

Random Forest

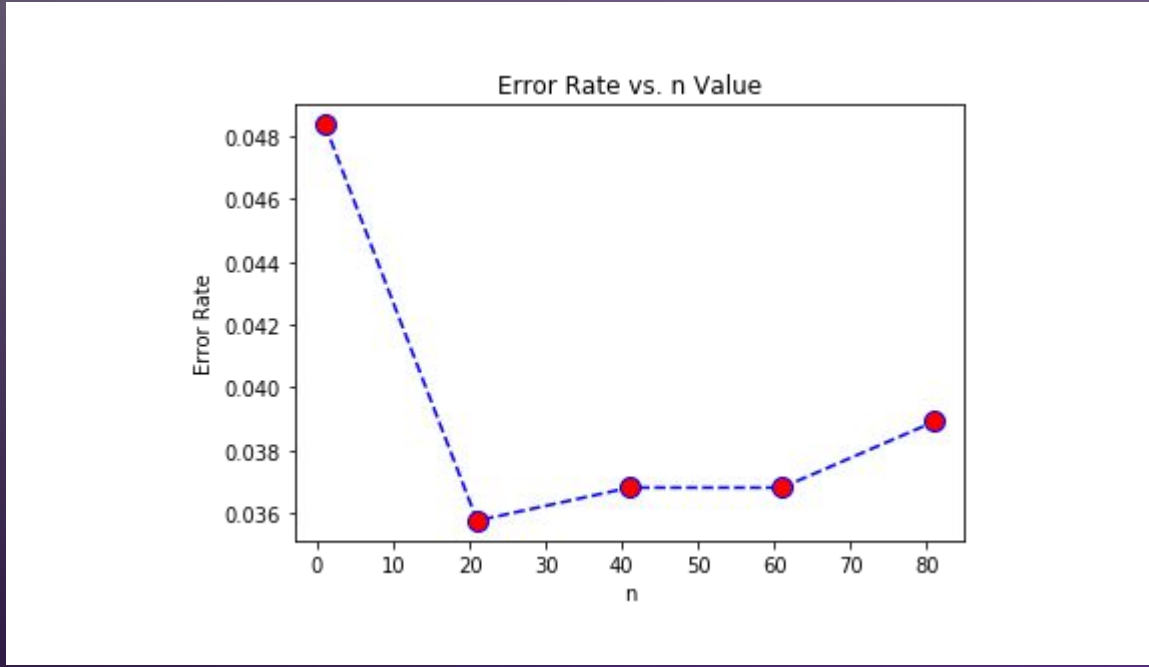
Error Rate Vs K-Value Plot

All 20 Attributes



Random Forest Error Rate Vs K-Value Plot

Major 3 Attributes
(IQR, MEANFUN, Q25)



Decision Tree

Decision Tree

Precision, F1 Score, Recall, Accuracy, Support

All 20 Attributes

	precision	recall	f1-score	support
female	0.9523	0.9716	0.9618	493
male	0.9688	0.9476	0.9581	458
micro avg	0.9600	0.9600	0.9600	951
macro avg	0.9605	0.9596	0.9600	951
weighted avg	0.9602	0.9600	0.9600	951

Accuracy: 0.9600420609884333

Decision Tree

Precision, F1 Score, Recall, Accuracy, Support

Major 3 Attributes

(IQR, MEANFUN, Q25)

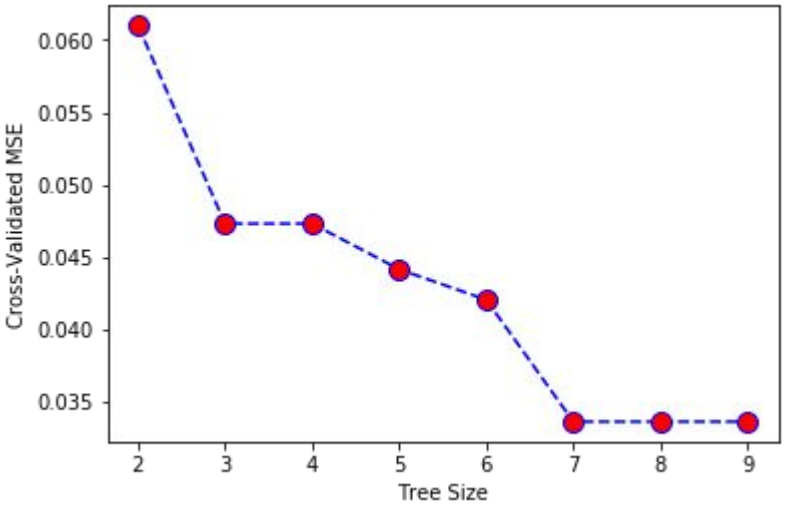
	precision	recall	f1-score	support
female	0.9621	0.9777	0.9698	493
male	0.9756	0.9585	0.9670	458
micro avg	0.9685	0.9685	0.9685	951
macro avg	0.9688	0.9681	0.9684	951
weighted avg	0.9686	0.9685	0.9684	951

Accuracy: 0.9684542586750788

Decision Tree

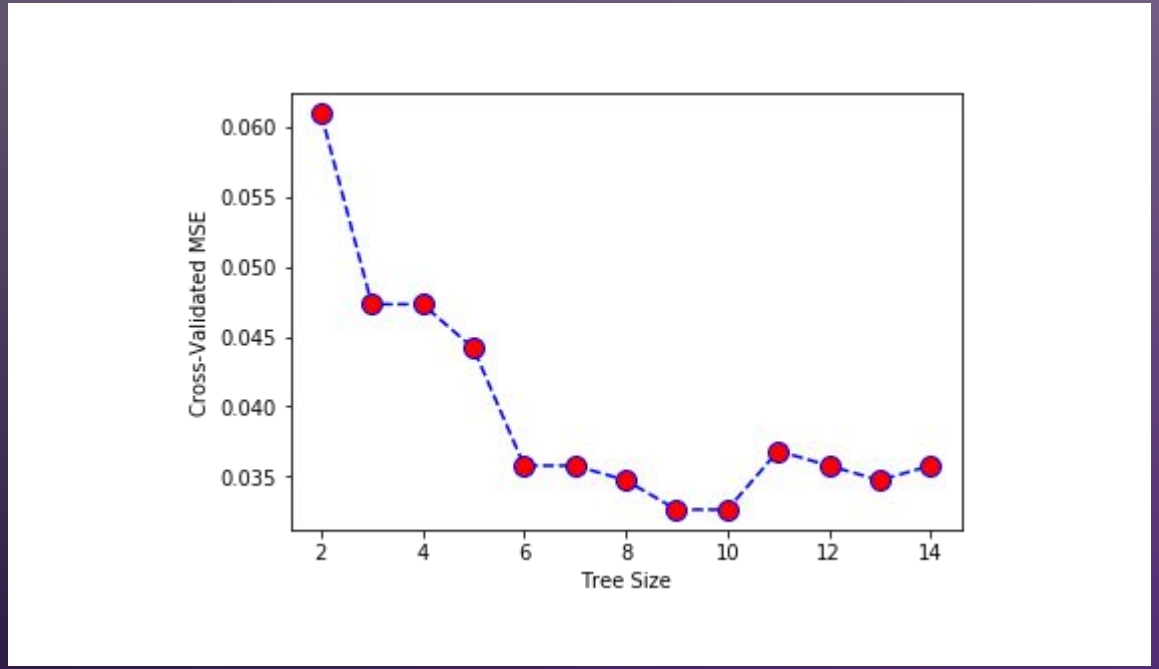
Error Rate Vs K-Value Plot

All 20 Attributes



Decision Tree Error Rate Vs K-Value Plot

Major 3 Attributes
(IQR, MEANFUN, Q25)

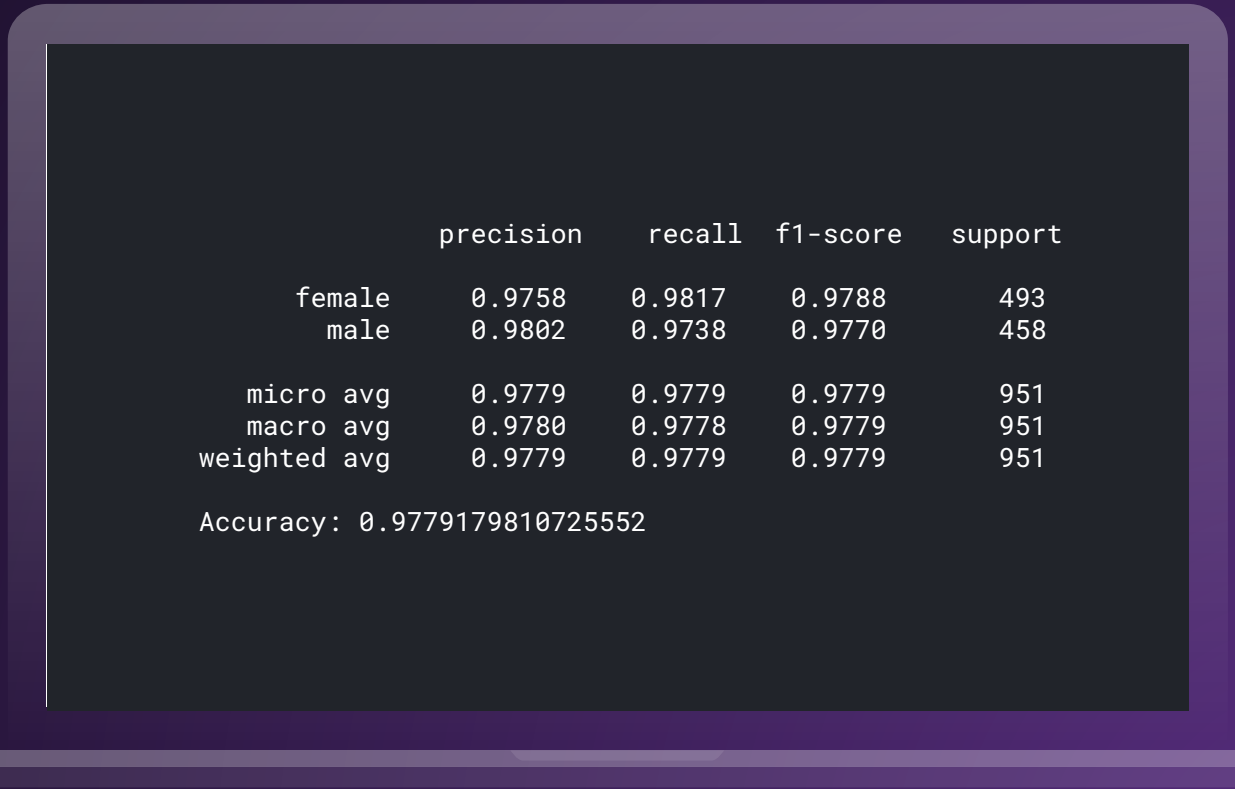


XGBoost

XGBoost

Precision, F1 Score, Recall, Accuracy, Support

All 20 Attributes



	precision	recall	f1-score	support
female	0.9758	0.9817	0.9788	493
male	0.9802	0.9738	0.9770	458
micro avg	0.9779	0.9779	0.9779	951
macro avg	0.9780	0.9778	0.9779	951
weighted avg	0.9779	0.9779	0.9779	951

Accuracy: 0.9779179810725552

XGBoost

Precision, F1 Score, Recall, Accuracy, Support

	precision	recall	f1-score	support
female	0.9583	0.9797	0.9689	493
male	0.9776	0.9541	0.9657	458
micro avg	0.9674	0.9674	0.9674	951
macro avg	0.9680	0.9669	0.9673	951
weighted avg	0.9676	0.9674	0.9674	951

Accuracy: 0.9674027339642481

Major 3 Attributes

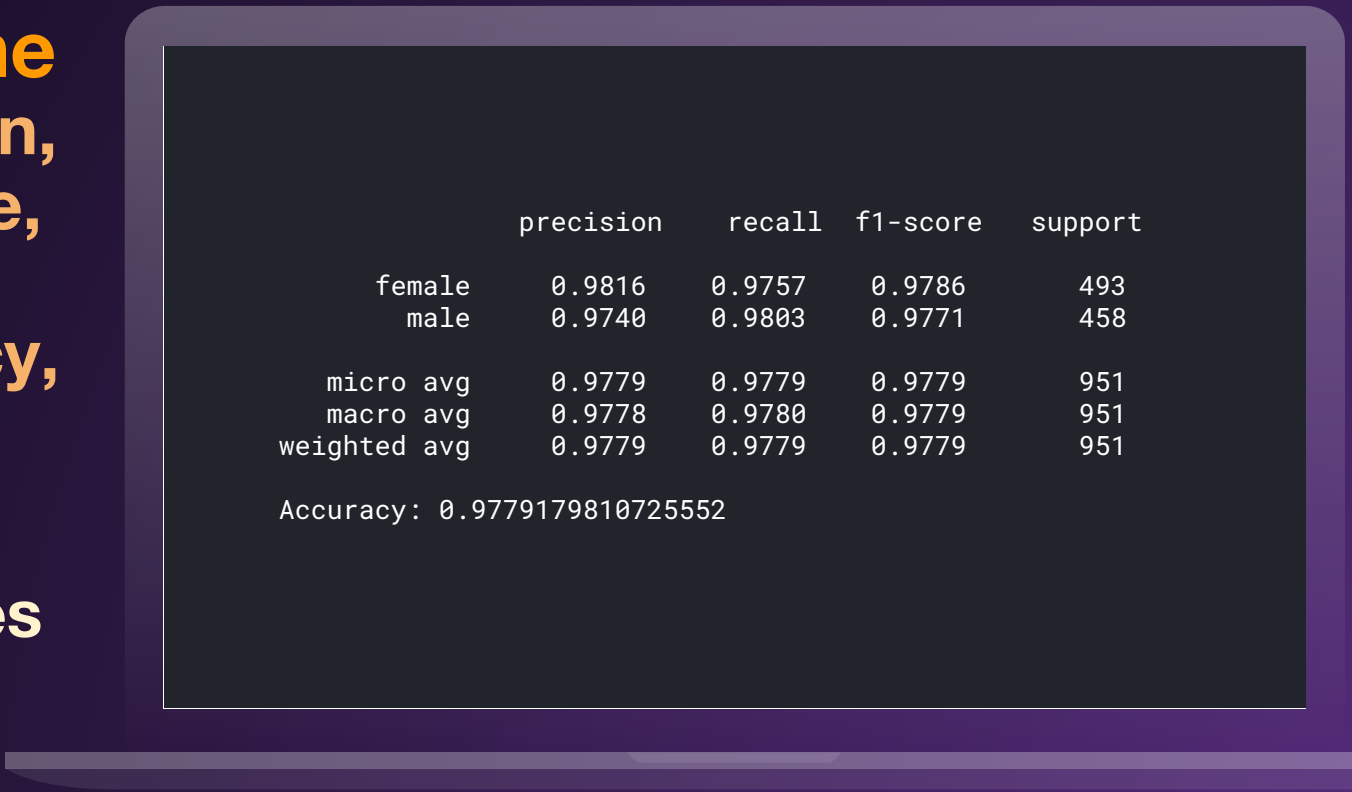
(IQR, MEANFUN, Q25)

Support Vector Machine

Support Vector Machine

Precision, F1 Score, Recall, Accuracy, Support

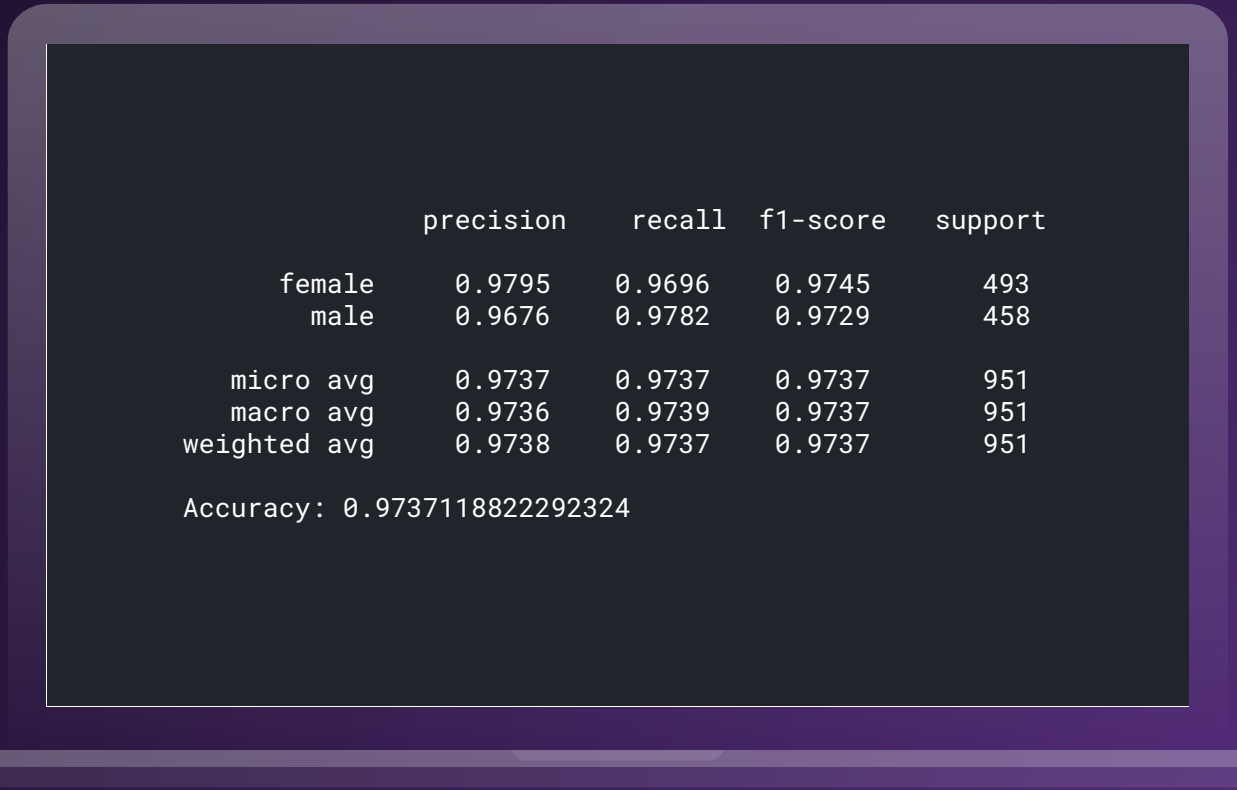
All 20 Attributes



Support Vector Machine

Precision, F1 Score, Recall, Accuracy, Support

Major 3 Attributes
(IQR, MEANFUN, Q25)



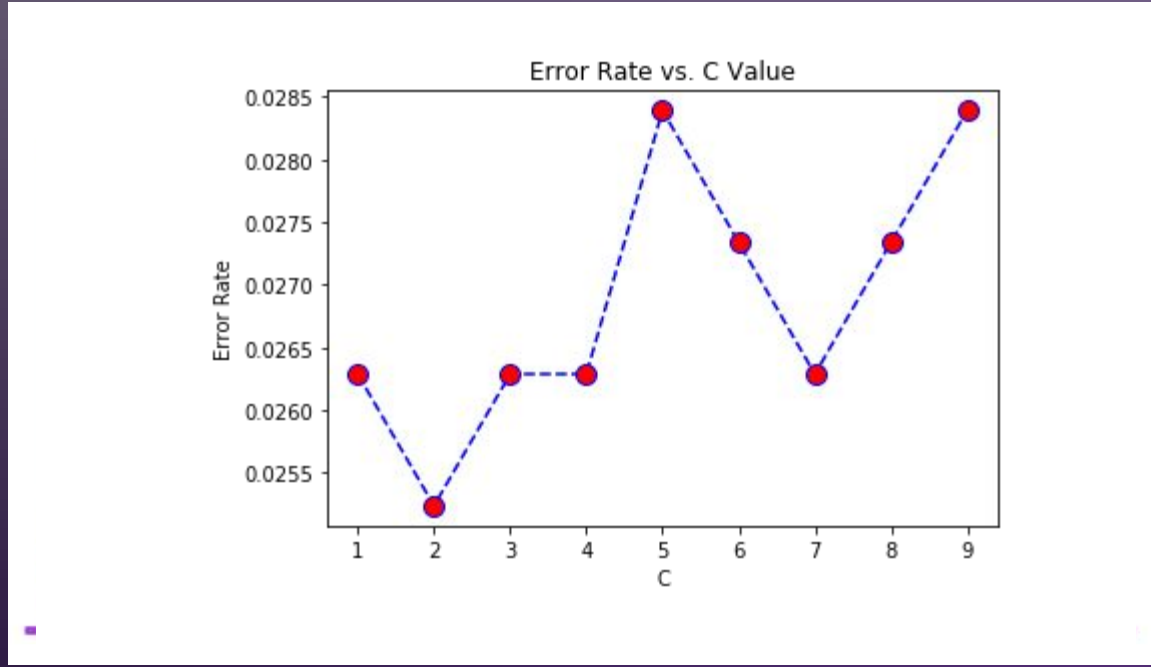
	precision	recall	f1-score	support
female	0.9795	0.9696	0.9745	493
male	0.9676	0.9782	0.9729	458
micro avg	0.9737	0.9737	0.9737	951
macro avg	0.9736	0.9739	0.9737	951
weighted avg	0.9738	0.9737	0.9737	951

Accuracy: 0.9737118822292324

Support Vector Machine

Error Rate Vs K-Value Plot

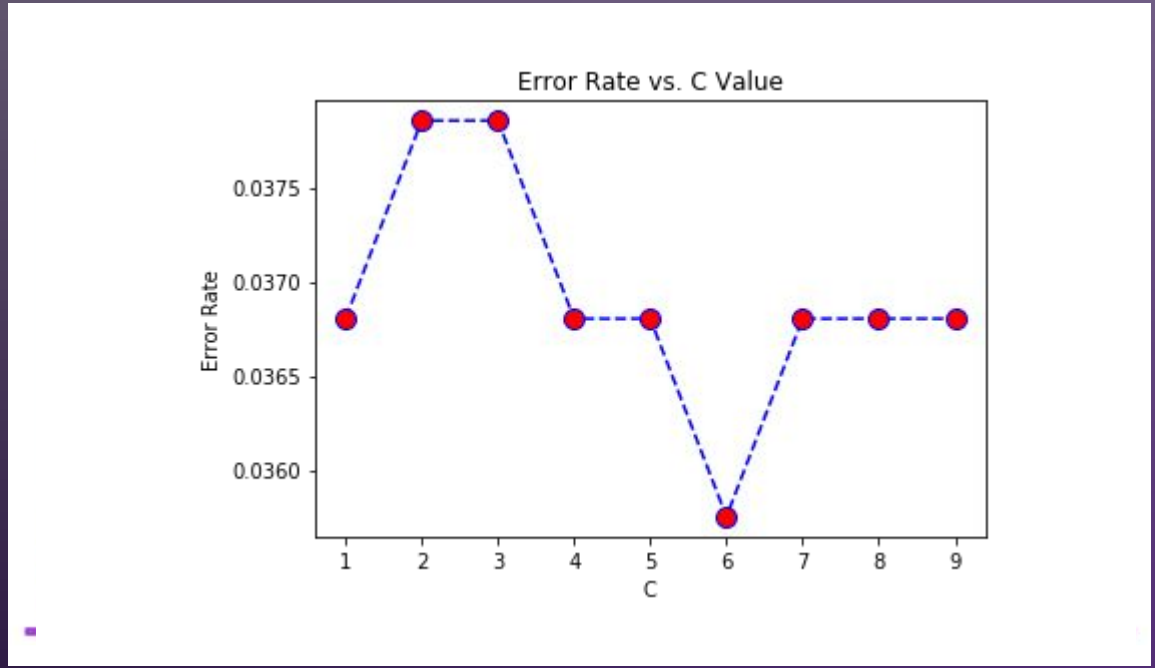
All 20 Attributes



Support Vector Machine

Error Rate Vs K-Value Plot

Major 3 Attributes (IQR, MEANFUN, Q25)

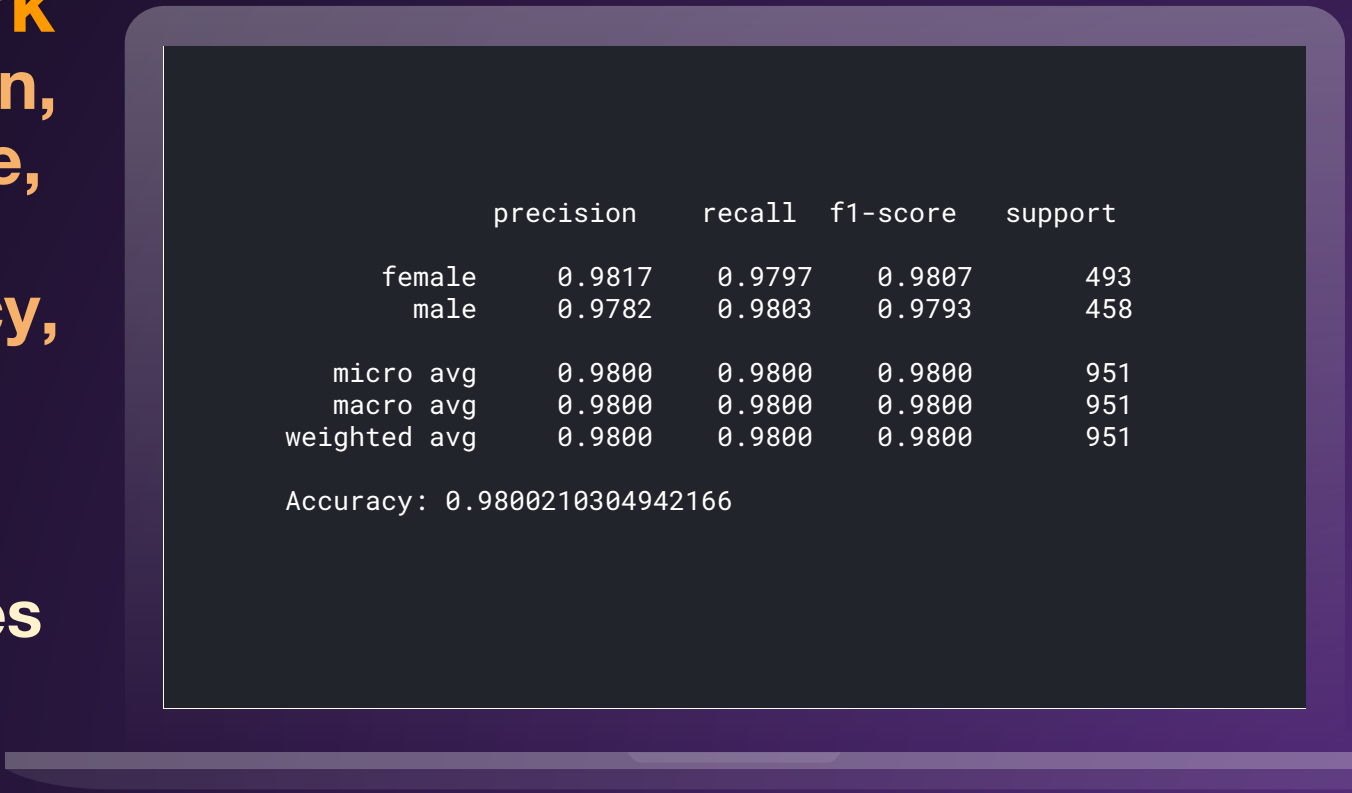


Neural Network

Neural Network

Precision, F1 Score, Recall, Accuracy, Support

All 20 Attributes



	precision	recall	f1-score	support
female	0.9817	0.9797	0.9807	493
male	0.9782	0.9803	0.9793	458
micro avg	0.9800	0.9800	0.9800	951
macro avg	0.9800	0.9800	0.9800	951
weighted avg	0.9800	0.9800	0.9800	951

Accuracy: 0.9800210304942166

Neural Network

Precision, F1 Score, Recall, Accuracy, Support

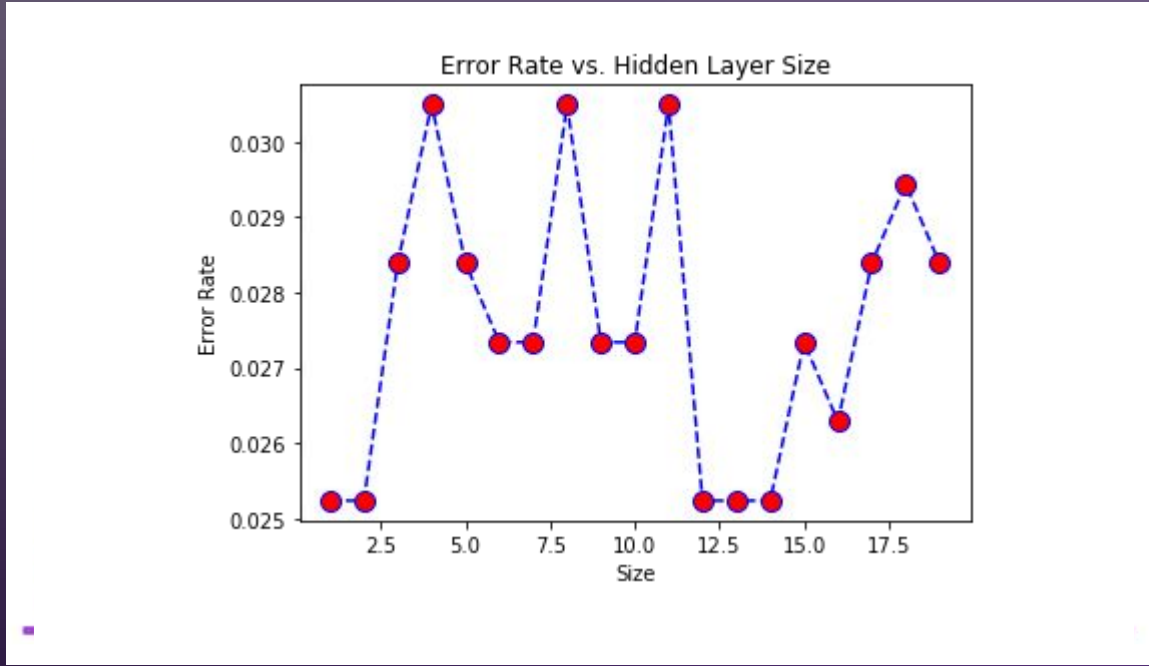
Major 3 Attributes
(IQR, MEANFUN, Q25)

	precision	recall	f1-score	support
female	0.9716	0.9716	0.9716	493
male	0.9694	0.9694	0.9694	458
micro avg	0.9706	0.9706	0.9706	951
macro avg	0.9705	0.9705	0.9705	951
weighted avg	0.9706	0.9706	0.9706	951

Accuracy: 0.9705573080967402

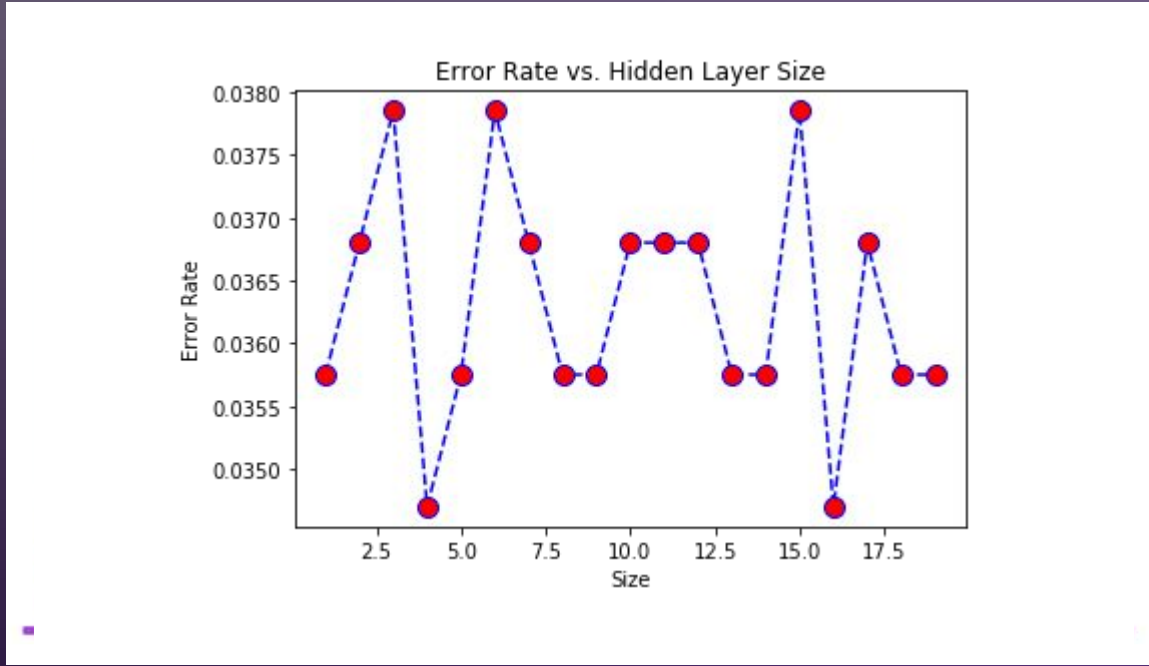
Neural Network Error Rate Vs K-Value Plot

All 20 Attributes



Neural Network Error Rate Vs K-Value Plot

Major 3 Attributes
(IQR, MEANFUN, Q25)

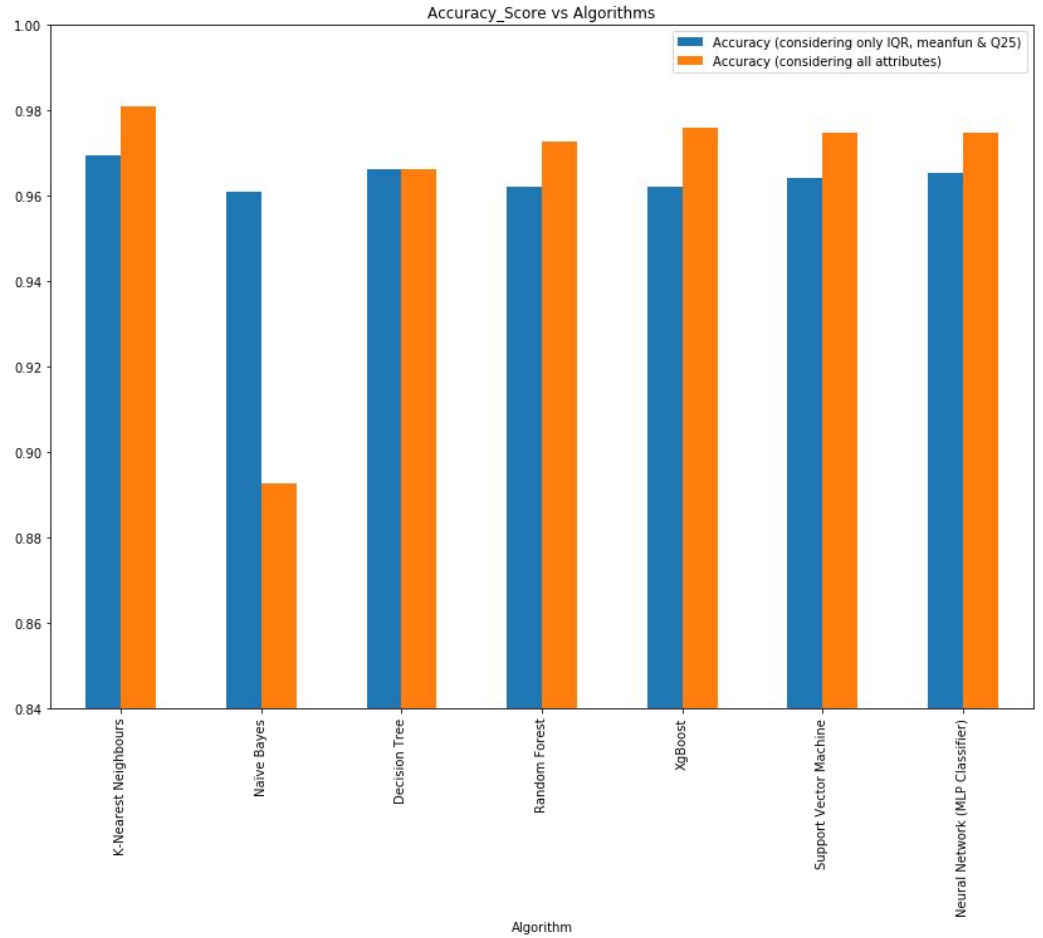


CONCLUSION OF USAGE OF ALL THE PREVIOUSLY DISCUSSED ALGORITHMS

XGBoost has the highest accuracy score amongst all the seven algorithms we used. The Graphical Comparison is presented in the next slide

LINK TO THE NOTEBOOK

GRAPHICAL COMPARISON BETWEEN ACCURACIES OF DIFFERENT ALGORITHMS



MEMBERS

Sagnik Mitra

Sneharup Mukherjee

Spandan Pal

THANK YOU