Nama: Tri Hesti Wahyuningsih

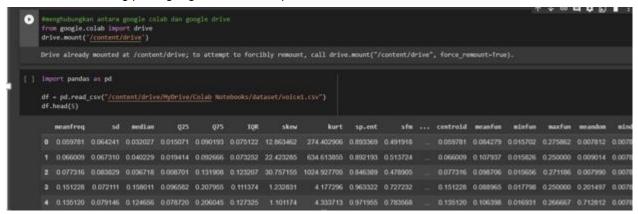
ID SIB: 3261839

## Klasifikasi voice gender

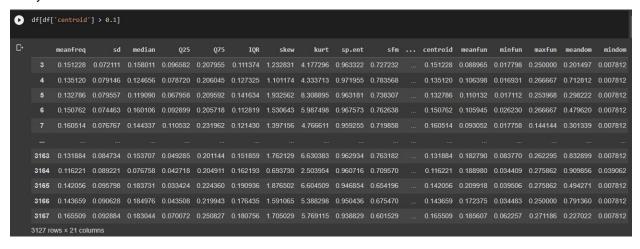
Dataset yang digunakan diambil dari situs kaggle

Link: https://www.kaggle.com/datasets/primaryobjects/voicegender

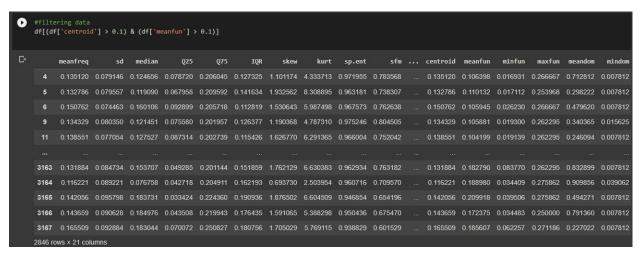
1. Melakukan mounting pada google drive dan import dataset



2. Menyeleksi data



3. Filtering data dengan minimal atribut centroid lebih dari 0.1 dan meanfun lebih dari 0.1



4. Membagi data independen dan dependen

```
[ ] #bagi data independen dan dependen variable
X = df.iloc[:,0:20] #independen variable
y = df.iloc[:,20] #dependen variable
```

5. membagi data

```
[ ] from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
X_train
```

6. Membuat model dengan svm data dan mengecek akurasinya dengan model svm didapat skor 62% dari model tersebut

```
[ ] from sklearn import svm

model = svm.SVC()
model = model.fit(X_train, y_train)

[ ] from sklearn.metrics import accuracy_score

    y_pred = model.predict(X_test)
    print(accuracy_score(y_pred, y_test))

0.6277602523659306
```

7. Membuat model dengan decision tree dan didapatkan model akurasi senilai 94%

```
[ ] from sklearn import tree

model2 = tree.DecisionTreeClassifier()
model2 = model2.fit(X_train, y_train)
y_pred2 = model2.predict(X_test)
accuracy_score(y_pred2, y_test)

0.9479495268138801
```

8. Membuat model naïve bayes

```
#Import Gaussian Naive Bayes model
from sklearn.naive_bayes import GaussianNB

#Create a Gaussian Classifier
gnb = GaussianNB()

#Train the model using the training sets
gnb.fit(X_train, y_train)

#Predict the response for test dataset
y_pred = gnb.predict[(X_test)]
```

9. Mengecek hasil akurasi model menggunakan naïve bayes Gaussian

```
[32] #Import scikit-learn metrics module for accuracy calculation
    from sklearn import metrics

# Model Accuracy, how often is the classifier correct?
    print("Accuracy:",metrics.accuracy_score(y_test, y_pred))

Accuracy: 0.886435331230284
```

 Kesimpulan dari penggunaan 3 model yaitu svm, decision tree dan naïve bayes yang mendapatkan akurasi tertinggi adalah decision tree dengan akurasi 0.95