EXPERIMENT 6(a)

IMPLEMENTATION OF SVM LINEAR ALGORITHM

NAME : SREENIDHI GANACHARI

REGISTRATION NUMBER : 19BCE7230

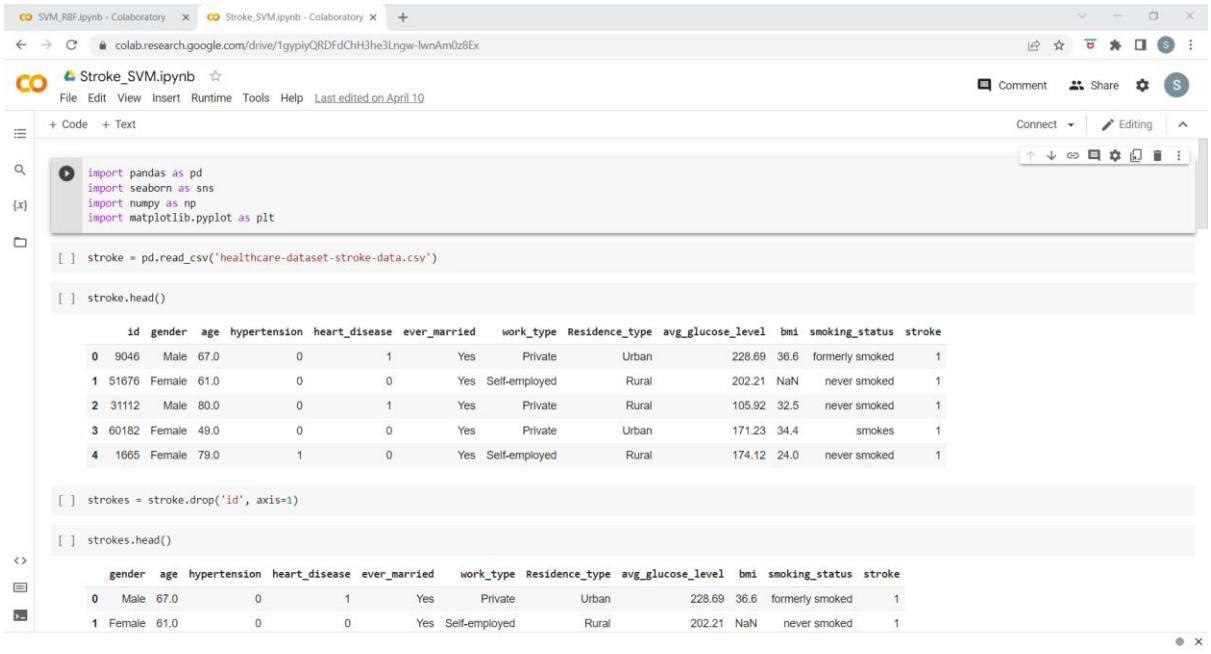
SLOT : L-23 &24

CODE –

```
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
stroke = pd.read_csv('healthcare-dataset-stroke-data.csv')
stroke.head()
strokes = stroke.drop('id', axis=1)
strokes.head()
strokes['gender'].unique()
strokes['ever_married'].unique()
strokes['work_type'].unique()
strokes['Residence_type'].unique()
strokes['smoking_status'].unique()
strokes['bmi'].fillna(strokes['bmi'].mean(), inplace = True)
strokes.info()
x = strokes.iloc[:,:-1]
y = strokes.iloc[:,-1]
from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OneHotEncoder
ct = ColumnTransformer(transformers = [('encoder', OneHotEncoder(), [0,
4,5,6,9])], remainder='passthrough')
x = np.array(ct.fit_transform(x))
print(x)
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y,random_state=1)
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train_1 = sc.fit_transform(x_train)
x_test_1 = sc.fit_transform(x_test)
from sklearn.svm import SVC
classifier = SVC(random_state = 0, kernel = 'linear')
classifier.fit(x_train_1, y_train)
y_pred = classifier.predict(x_test_1)
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
print(cm)
```

```
from sklearn.metrics import accuracy_score
acc = accuracy_score(y_test, y_pred)
acc
```

OUTPUT -



```
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
```

```
stroke = pd.read_csv('healthcare-dataset-stroke-data.csv')
```

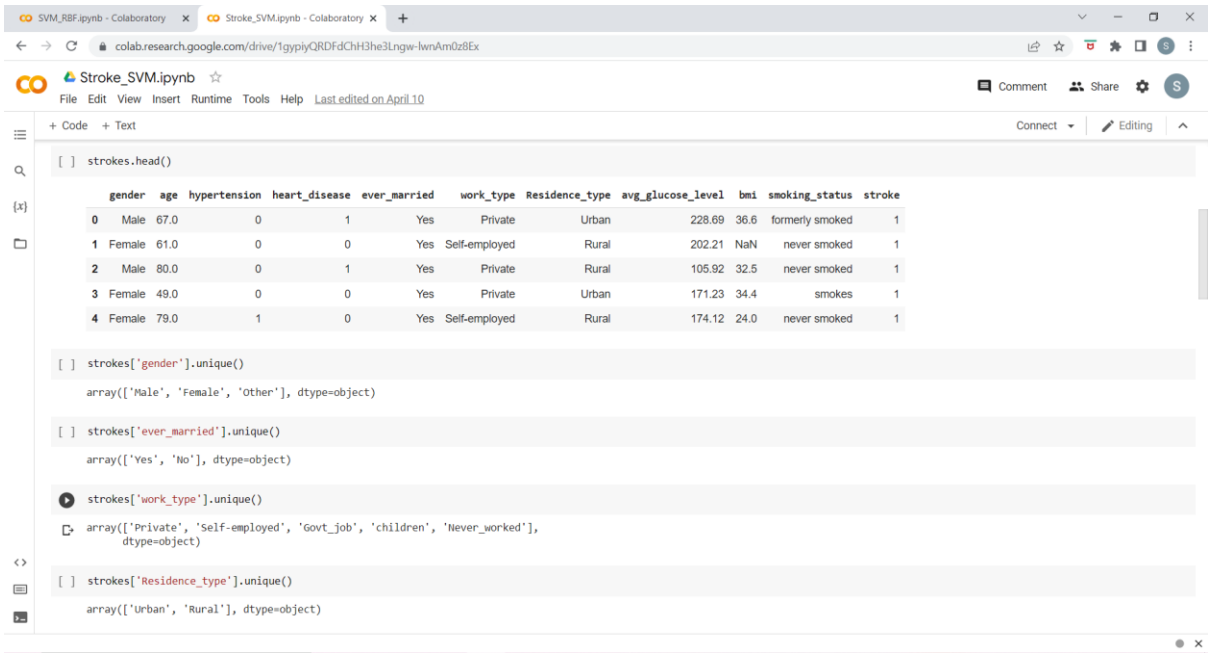
```
stroke.head()
```

	id	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	9046	Male	67.0	0	1	Yes	Private	Urban	228.69	36.6	formerly smoked	1
1	51676	Female	61.0	0	0	Yes	Self-employed	Rural	202.21	NaN	never smoked	1
2	31112	Male	80.0	0	1	Yes	Private	Rural	105.92	32.5	never smoked	1
3	60182	Female	49.0	0	0	Yes	Private	Urban	171.23	34.4	smokes	1
4	1665	Female	79.0	1	0	Yes	Self-employed	Rural	174.12	24.0	never smoked	1

```
strokes = stroke.drop('id', axis=1)
```

```
strokes.head()
```

	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	Male	67.0	0	1	Yes	Private	Urban	228.69	36.6	formerly smoked	1
1	Female	61.0	0	0	Yes	Self-employed	Rural	202.21	NaN	never smoked	1



```
strokes.head()
```

	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	Male	67.0	0	1	Yes	Private	Urban	228.69	36.6	formerly smoked	1
1	Female	61.0	0	0	Yes	Self-employed	Rural	202.21	NaN	never smoked	1
2	Male	80.0	0	1	Yes	Private	Rural	105.92	32.5	never smoked	1
3	Female	49.0	0	0	Yes	Private	Urban	171.23	34.4	smokes	1
4	Female	79.0	1	0	Yes	Self-employed	Rural	174.12	24.0	never smoked	1

```
strokes['gender'].unique()
```

```
array(['Male', 'Female', 'Other'], dtype=object)
```

```
strokes['ever_married'].unique()
```

```
array(['Yes', 'No'], dtype=object)
```

```
strokes['work_type'].unique()
```

```
array(['Private', 'Self-employed', 'Govt_job', 'children', 'Never_worked'],
      dtype=object)
```

```
strokes['Residence_type'].unique()
```

```
array(['Urban', 'Rural'], dtype=object)
```

SVM_RBF.ipynb - Colaboratory | Stroke_SVM.ipynb - Colaboratory

colab.research.google.com/drive/1gypiyQRDFdChH3he3Lngw-lwnAm0z8Ex

Stroke_SVM.ipynb

File Edit View Insert Runtime Tools Help Last edited on April 10

Comment Share

+ Code + Text

Connect Editing

```
[ ] strokes['smoking_status'].unique()
```

```
array(['formerly smoked', 'never smoked', 'smokes', 'Unknown'],
      dtype=object)
```

```
[ ] strokes['bmi'].fillna(strokes['bmi'].mean(), inplace = True)
```

```
[ ] strokes.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5110 entries, 0 to 5109
Data columns (total 11 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   gender             5110 non-null   object 
 1   age                5110 non-null   float64
 2   hypertension       5110 non-null   int64  
 3   heart_disease      5110 non-null   int64  
 4   ever_married       5110 non-null   object 
 5   work_type          5110 non-null   object 
 6   Residence_type     5110 non-null   object 
 7   avg_glucose_level  5110 non-null   float64
 8   bmi                5110 non-null   float64
 9   smoking_status     5110 non-null   object 
 10  stroke             5110 non-null   int64  
dtypes: float64(3), int64(3), object(5)
memory usage: 439.3+ KB
```

```
x = strokes.iloc[:,:-1]
y = strokes.iloc[:,-1]
```

SVM_RBF.ipynb - Colaboratory | Stroke_SVM.ipynb - Colaboratory

colab.research.google.com/drive/1gypiyQRDFdChH3he3Lngw-lwnAm0z8Ex

Stroke_SVM.ipynb

File Edit View Insert Runtime Tools Help Last edited on April 10

Comment Share

+ Code + Text

Connect Editing

```
[ ] from sklearn.compose import ColumnTransformer
    from sklearn.preprocessing import OneHotEncoder
```

```
[ ] ct = ColumnTransformer(transformers = [('encoder', OneHotEncoder(), [0,4,5,6,9])], remainder='passthrough')
    x = np.array(ct.fit_transform(x))
    print(x)
```

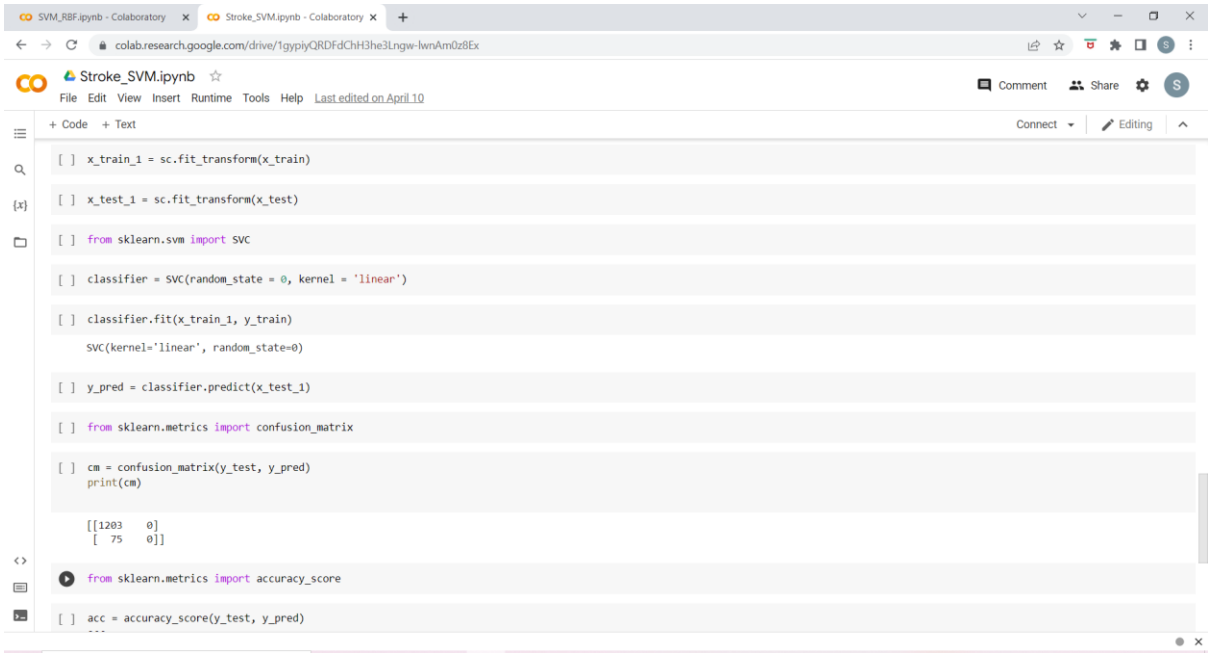
```
[[  0.           1.           0.         ...   1.         228.69
   36.6       ]
 [  1.           0.           0.         ...   0.         202.21
   28.89323691]
 [  0.           1.           0.         ...   1.         105.92
   32.5       ]
 ...
 [  1.           0.           0.         ...   0.          82.99
   30.6       ]
 [  0.           1.           0.         ...   0.         166.29
   25.6       ]
 [  1.           0.           0.         ...   0.          85.28
   26.2       ]]
```

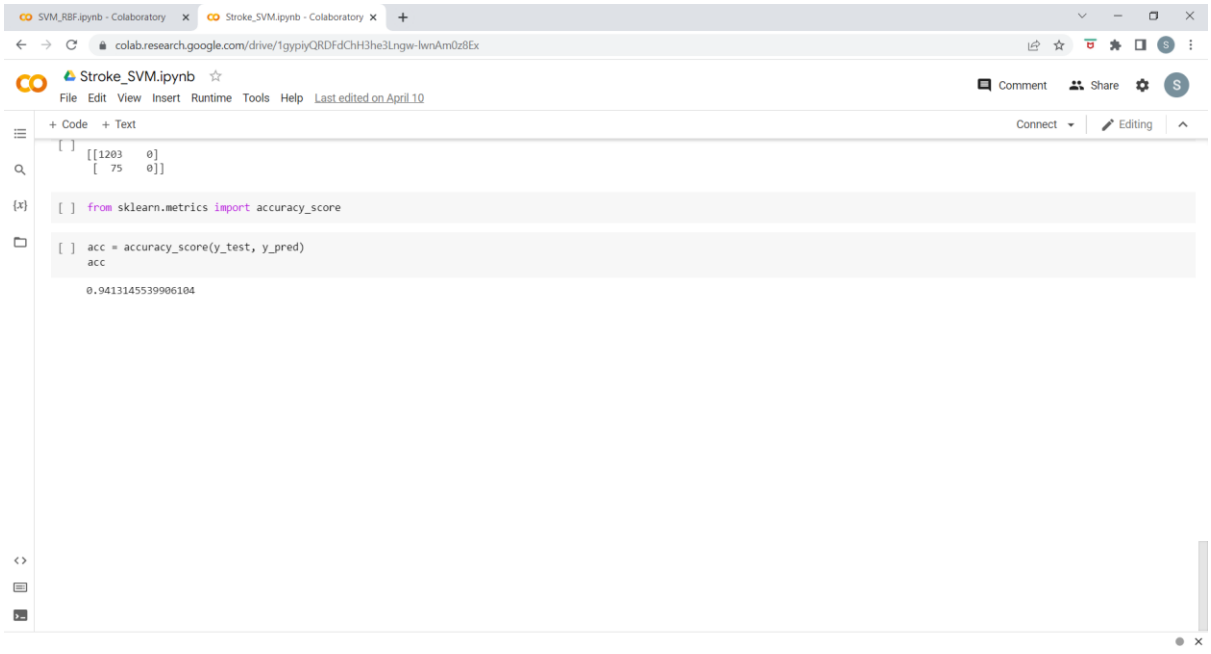
```
[ ] from sklearn.model_selection import train_test_split
```

```
[ ] x_train, x_test, y_train, y_test = train_test_split(x,y,random_state=1)
```

```
[ ] from sklearn.preprocessing import StandardScaler
```

```
sc = StandardScaler()
```


SVM_RBF.ipynb - Colaboratory
Stroke_SVM.ipynb - Colaboratory
colab.research.google.com/drive/1gypiyQRDFdChH3he3Lngw-lwnAm0z8Ex
Stroke_SVM.ipynb
File Edit View Insert Runtime Tools Help Last edited on April 10
Comment
Share
+ Code + Text
Connect
Editing
[ ] x_train_1 = sc.fit_transform(x_train)
[ ] x_test_1 = sc.fit_transform(x_test)
[ ] from sklearn.svm import SVC
[ ] classifier = SVC(random_state = 0, kernel = 'linear')
[ ] classifier.fit(x_train_1, y_train)
SVC(kernel='linear', random_state=0)
[ ] y_pred = classifier.predict(x_test_1)
[ ] from sklearn.metrics import confusion_matrix
[ ] cm = confusion_matrix(y_test, y_pred)
print(cm)
[[1203 0]
[75 0]]
from sklearn.metrics import accuracy_score
[ ] acc = accuracy_score(y_test, y_pred)


SVM_RBF.ipynb - Colaboratory
Stroke_SVM.ipynb - Colaboratory
colab.research.google.com/drive/1gypiyQRDFdChH3he3Lngw-lwnAm0z8Ex
Stroke_SVM.ipynb
File Edit View Insert Runtime Tools Help Last edited on April 10
Comment
Share
+ Code + Text
Connect
Editing
[ ]
[[1203 0]
[75 0]]
[ ] from sklearn.metrics import accuracy_score
[ ] acc = accuracy_score(y_test, y_pred)
acc
0.9413145539906104