Knowledge Discovery & Data Mining Lab-08

Name: Gurvinder Kaur Matharu
PRN: 20190802077

AIM:

To implement SVM using python.

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.svm import SVC
```

```
data = pd.read_csv('Fish.csv')
```

```
data.head()
```

	Species	Weight	Length1	Length2	Length3	Height	Width
0	Bream	242.0	23.2	25.4	30.0	11.5200	4.0200
1	Bream	290.0	24.0	26.3	31.2	12.4800	4.3056
2	Bream	340.0	23.9	26.5	31.1	12.3778	4.6961
3	Bream	363.0	26.3	29.0	33.5	12.7300	4.4555
4	Bream	430.0	26.5	29.0	34.0	12.4440	5.1340

```
data.shape
```

```
(159, 7)
```

```
data.isnull().sum()
```

```
Species    0
Weight     0
Length1    0
Length2    0
Length3    0
Height     0
Width      0
dtype: int64
```

```
X = data.iloc[:,1:]
y = data.iloc[:,0]
```

```
X
```

	Weight	Length1	Length2	Length3	Height	Width
0	242.0	23.2	25.4	30.0	11.5200	4.0200
1	290.0	24.0	26.3	31.2	12.4800	4.3056
2	340.0	23.9	26.5	31.1	12.3778	4.6961
3	363.0	26.3	29.0	33.5	12.7300	4.4555
4	430.0	26.5	29.0	34.0	12.4440	5.1340
...	...	...	...	...	...	...

```
y
```

```
0      Bream
1      Bream
2      Bream
3      Bream
4      Bream
       ...
154    Smelt
155    Smelt
156    Smelt
157    Smelt
158    Smelt
Name: Species, Length: 159, dtype: object
```

```
labelencoder = LabelEncoder()
y = labelencoder.fit_transform(y)
```

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

```
model = SVC(kernel = 'linear')
```

```
model.fit(X_train, y_train)
```

```
SVC(kernel='linear')
```

```
svm_pred = model.predict(X_test)
svm_pred
```

```
array([3, 2, 5, 4, 0, 4, 2, 2, 5, 0, 5, 4, 0, 5, 3, 2, 0, 2, 2, 1, 2, 4,
       2, 0, 2, 1, 3, 5, 4, 2, 2, 0])
```

```
accuracy = model.score(X_test, y_test)
accuracy
```

```
0.9375
```

✓ 0s completed at 9:26 PM