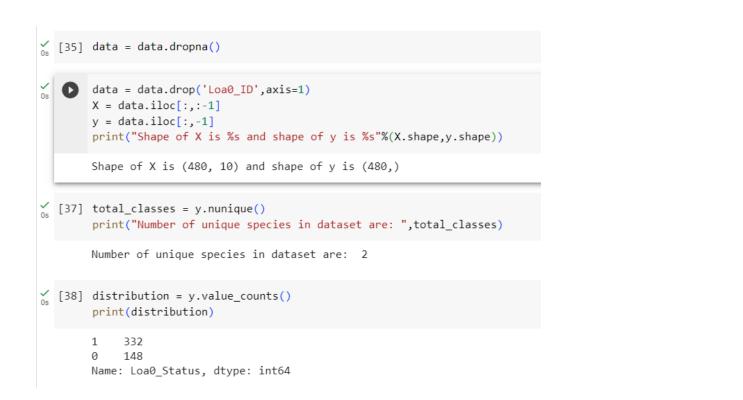# EXP 8 ADABOOST ALGORITHM

```python
[33] import pandas as pd
     import numpy as np
     from sklearn.model_selection import train_test_split
     from sklearn.ensemble import AdaBoostClassifier
     import warnings
     warnings.filterwarnings("ignore")
```

```python
[34] # Reading the dataset from the csv file
     # separator is a vertical line, as seen in the dataset
     data = pd.read_csv("loan_data_set2.csv")

     # Printing the shape of the dataset
     print(data.shape)
     data.head()
```

(614, 12)

	Loa0_ID	Ge0der	Married	Depe0de0ts	Self_Emplo1ed	Applica0tI0come	Coapplica0tI0come	Loa0Amou0t	Loa0_Amou0t_Term	Credit_Histor1	Propert1_Area	Loa0_Status
0	LP001002	1.0	0.0	0.0	0.0	5849	0.0	NaN	360.0	1.0	1	1
1	LP001003	1.0	1.0	1.0	0.0	4583	1508.0	128.0	360.0	1.0	3	0
2	LP001005	1.0	1.0	0.0	1.0	3000	0.0	66.0	360.0	1.0	1	1
3	LP001006	1.0	1.0	0.0	0.0	2583	2358.0	120.0	360.0	1.0	1	1
4	LP001008	1.0	0.0	0.0	0.0	6000	0.0	141.0	360.0	1.0	1	1

```python
[35] data = data.dropna()
```

```python
     data = data.drop('Loa0_ID',axis=1)
     X = data.iloc[:,:-1]
     y = data.iloc[:,-1]
     print("Shape of X is %s and shape of y is %s"%(X.shape,y.shape))
```

Shape of X is (480, 10) and shape of y is (480,)

```python
[37] total_classes = y.nunique()
     print("Number of unique species in dataset are: ",total_classes)
```

Number of unique species in dataset are: 2

```python
[38] distribution = y.value_counts()
     print(distribution)
```

```
1    332
0    148
Name: Loa0_Status, dtype: int64
```

```
[39] X_train, X_val, Y_train, Y_val = train_test_split(X, y, test_size=0.25, random_state=28)
```

```
[40] from sklearn.ensemble import AdaBoostClassifier
     # Creating adaboost classifier model
     adb = AdaBoostClassifier()
     adb_model = adb.fit(X_train,Y_train)
```

```
[41] print("The accuracy of the model on validation set is", adb_model.score(X_val,Y_val))
```

The accuracy of the model on validation set is 0.7916666666666666