

```
!pip install autoviml
```

```
from autoviml.Auto_ViML import Auto_ViML
```

```
Imported Auto_NLP version: 0.0.45.. Call using:
```

```
    train_nlp, test_nlp, nlp_pipeline, predictions = Auto_NLP(  
        nlp_column, train, test, target, score_type='balanced_accuracy',  
        modeltype='Classification',top_num_features=200, verbose=0,  
        build_model=True)
```

```
Imported Auto_ViML version: 0.1.677. Call using:
```

```
    m, feats, trainm, testm = Auto_ViML(train, target, test,  
        sample_submission='',  
        scoring_parameter='', KMeans_Featurizer=False,  
        hyper_param='RS',feature_reduction=True,  
        Boosting_Flag='CatBoost', Binning_Flag=False,  
        Add_Poly=0, Stacking_Flag=False,Imbalanced_Flag=False,  
        verbose=1)
```

The new Auto_ViML can solve multi-label, multi-output problems. Check if your version is $\geq 0.1.669$
To get the latest version, perform "pip install autoviml --no-cache-dir --ignore-installed"

```
datapath = 'https://web.stanford.edu/class/archive/cs/cs109/cs109.1166/stuff/'  
df = pd.read_csv(datapath+'titanic.csv')  
train=df  
print(train.shape)  
#print(test.shape)  
print(train.head())  
target = 'Survived'
```

```
(887, 8)
```

| | Survived | Pclass | Name \ |
|---|----------|--------|--|
| 0 | 0 | 3 | Mr. Owen Harris Braund |
| 1 | 1 | 1 | Mrs. John Bradley (Florence Briggs Thayer) Cumings |
| 2 | 1 | 3 | Miss. Laina Heikkinen |
| 3 | 1 | 1 | Mrs. Jacques Heath (Lily May Peel) Futrelle |
| 4 | 0 | 3 | Mr. William Henry Allen |

| | Sex | Age | Siblings/Spouses Aboard | Parents/Children Aboard | Fare |
|---|------|-------|-------------------------|-------------------------|------|
| 0 | male | 22.00 | 1 | 0 | 7.25 |

| | | | | | |
|---|--------|-------|---|---|-------|
| 1 | female | 38.00 | 1 | 0 | 71.28 |
| 2 | female | 26.00 | 0 | 0 | 7.92 |
| 3 | female | 35.00 | 1 | 0 | 53.10 |
| 4 | male | 35.00 | 0 | 0 | 8.05 |

```
df['Survived'].value_counts()
```

```
0    545
1    342
Name: Survived, dtype: int64
```

```
num = int(0.9*df.shape[0])
train = df[:num]
test = df[num:]
sample_submission=''
scoring_parameter = 'balanced-accuracy'
```

```
train['Survived'].value_counts()
```

```
0    489
1    309
Name: Survived, dtype: int64
```

```
m, feats, trainm, testm = Auto_ViML(train, target, test, sample_submission,
                                     scoring_parameter=scoring_parameter,
                                     hyper_param='GS',feature_reduction=True,
                                     Boosting_Flag=True,Binning_Flag=False,
                                     Add_Poly=0, Stacking_Flag=False,
                                     Imbalanced_Flag=False,
                                     verbose=1)
```

```
#####
precision    recall  f1-score   support

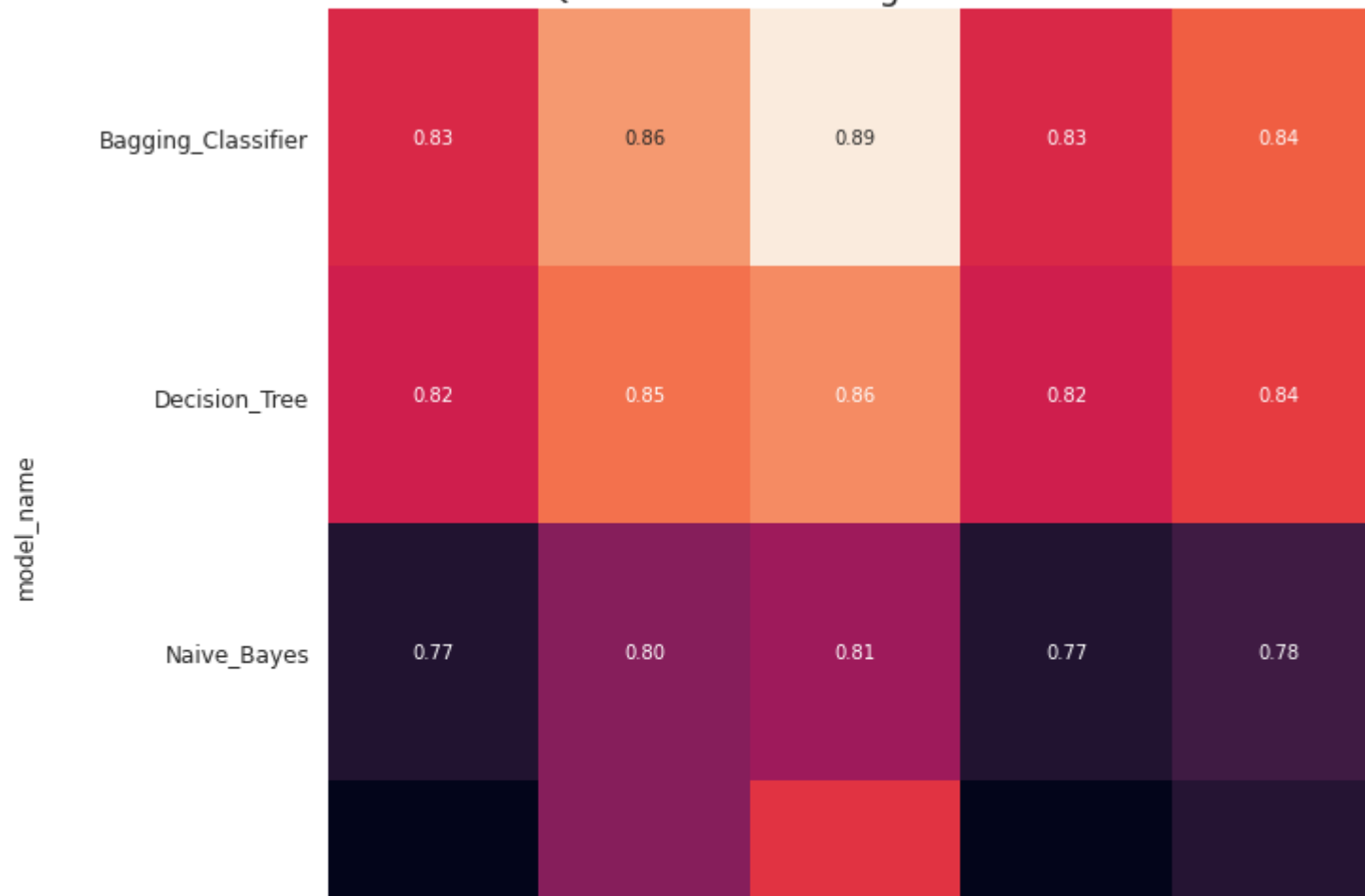
     0       0.85     0.96     0.90        49
     1       0.92     0.74     0.82        31

 accuracy          0.88        80
 macro avg       0.89     0.85     0.86        80
weighted avg       0.88     0.88     0.87        80
```

```
[[47  2]
 [ 8 23]]
```

```
#####
Single Model is better than Ensembling Models for this data set.
```

QuickML Ensembling Models Results



Logistic_Regression_CV



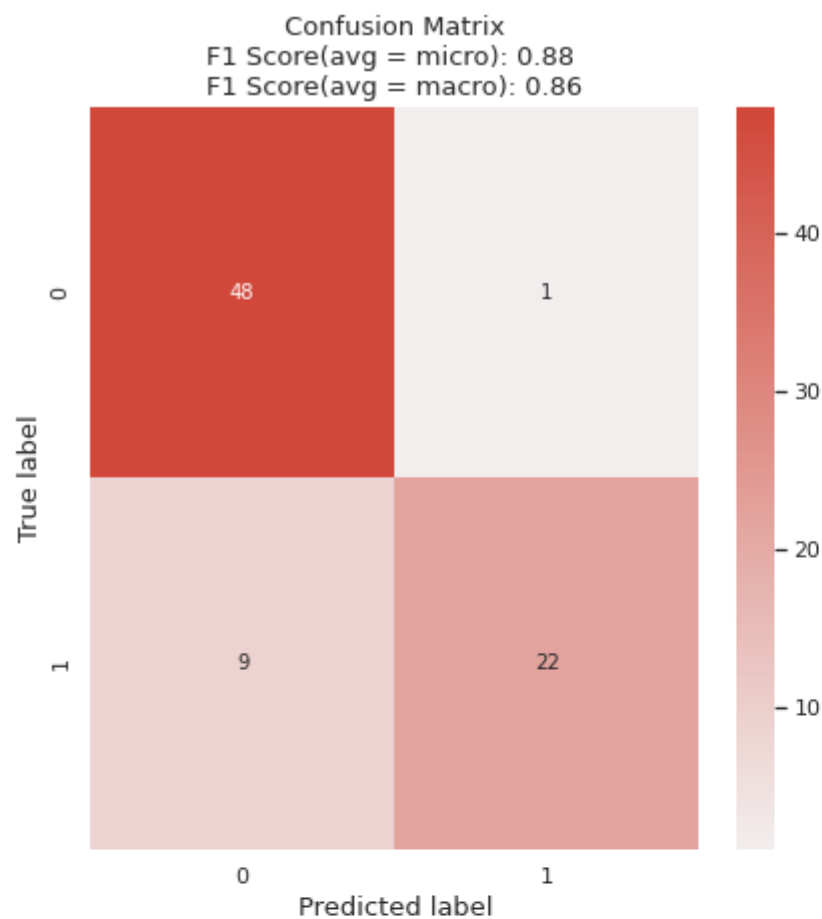
bal_accuracy_score

accuracy_score

ave_precision_score

ave_recall_score

ave_f1_score



True label

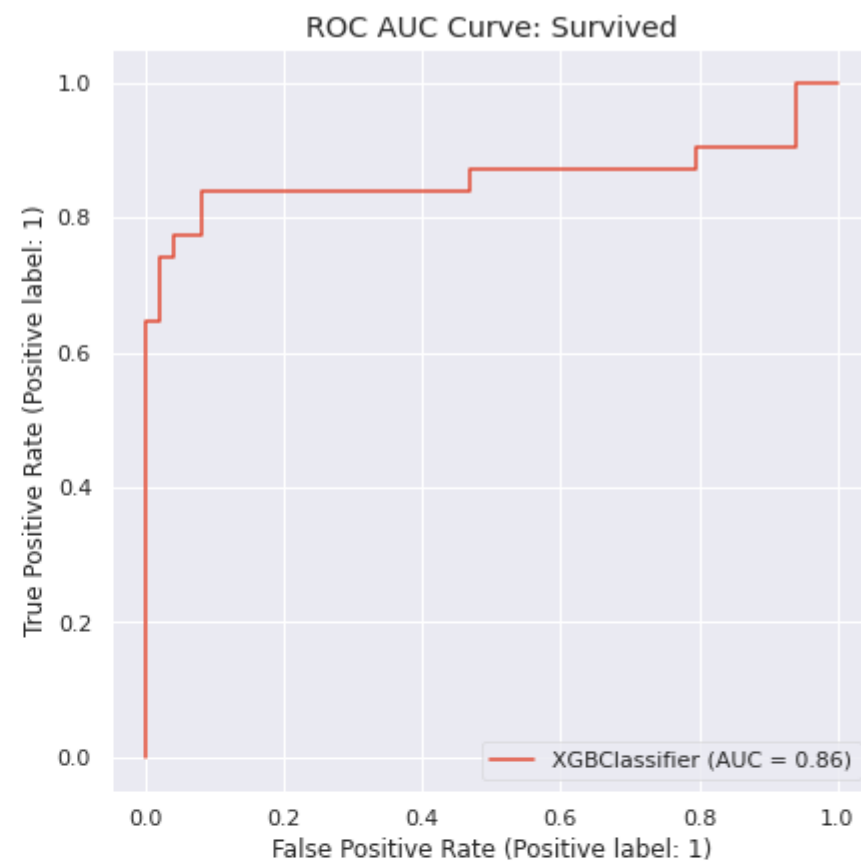
0

1

Predicted label

0

1



True Positive Rate (Positive label: 1)

1.0

0.8

0.6

0.4

0.2

0.0

0.0

0.2

0.4

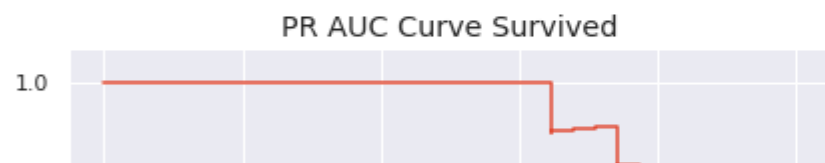
0.6

0.8

1.0

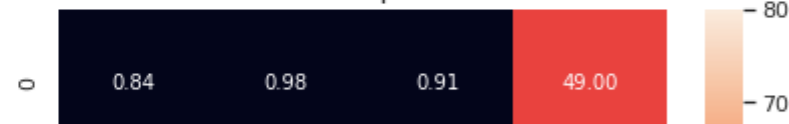
False Positive Rate (Positive label: 1)

XGBClassifier (AUC = 0.86)



1.0

Classification Report: Survived



0

0.84

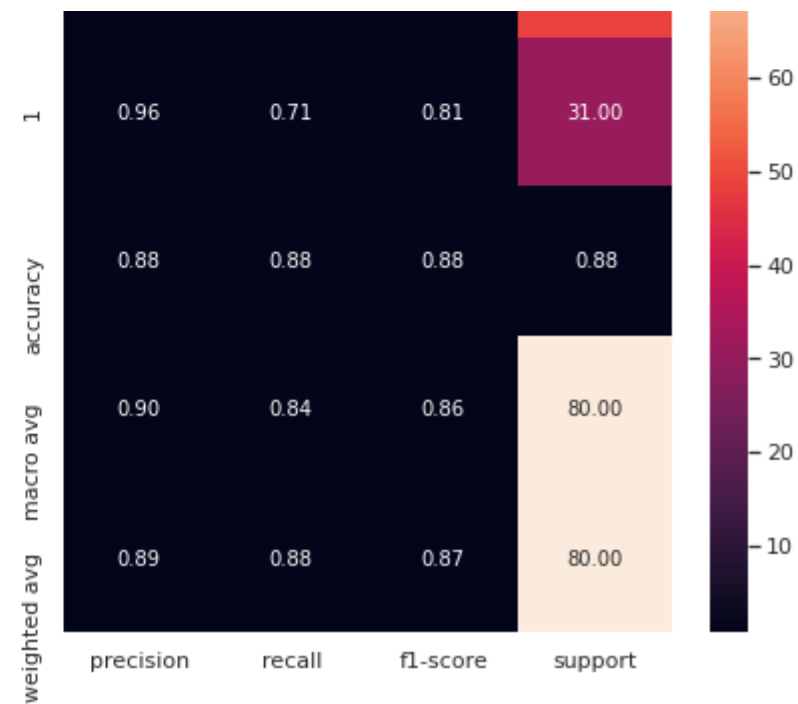
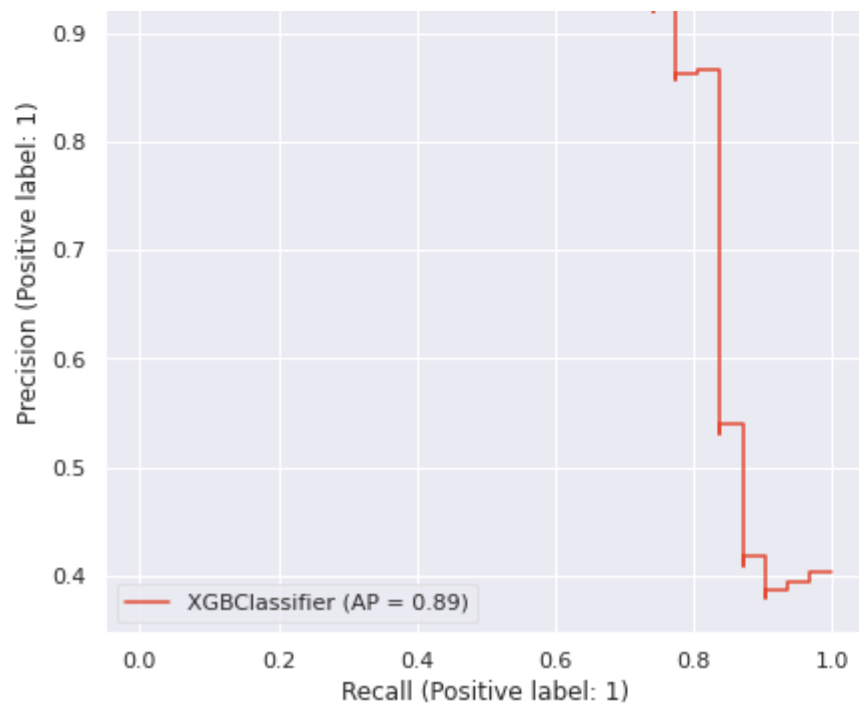
0.98

0.91

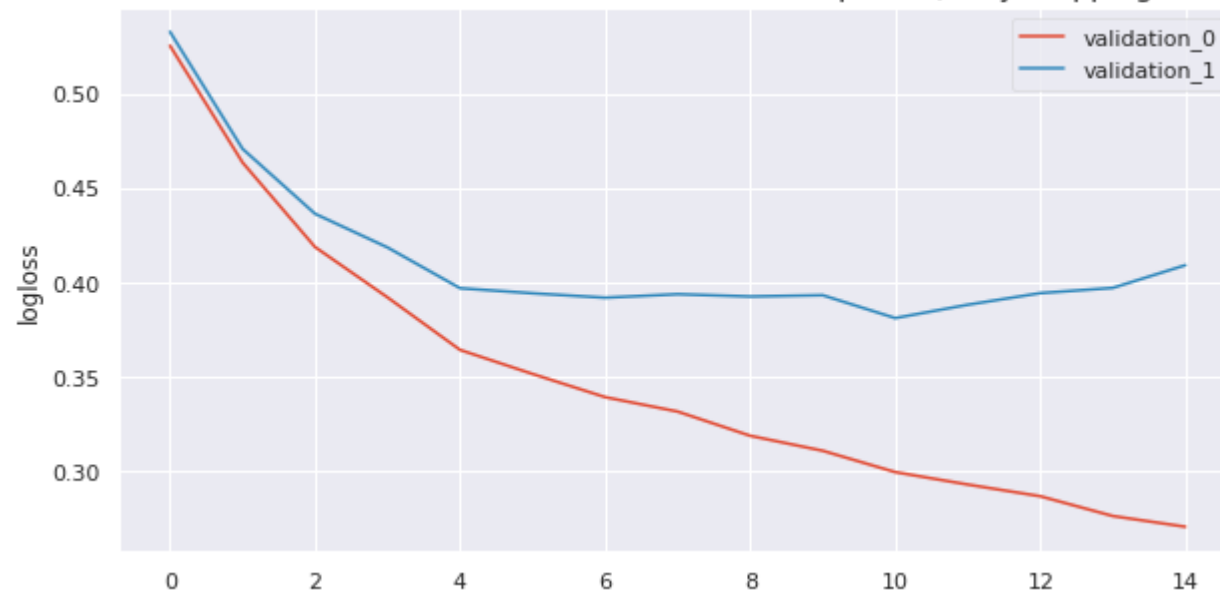
49.00

80

70



Survived Results Train and Validation Metrics across Epochs (Early Stopping in effect)



Time taken for this Target (in seconds) = 121

Binning_Flag set to False or there are no float vars in data set to be binned

Performing MinMax scaling of train and test data...

TRAINING MODEL ON FULL TRAIN DATA

```
Actual Training time taken in seconds = 0
Training of models completed. Now starting predictions on test data...
Making test Data predictions using modified Threshold = 0.301
Calculating weighted average ensemble of 5 classifiers
Plotting Feature Importances to explain the output of model
##### P R E D I C T I O N   O N   T E S T   C O M P L E T E D #####
Time taken thus far (in seconds) = 122
Writing Output files to disk...
Saving predictions to ./Survived/Survived_Binary_Classification_test_modified.csv
Saving predictions to ./Survived/Survived_Binary_Classification_submission.csv
Saving predictions to ./Survived/Survived_Binary_Classification_train_modified.csv
##### C O M P L E T E D #####
Time Taken in mins = 2.0 for the Entire Process
```

