REPORT

JAINAM NIMISH SHROFF

MACHINE LEARNING FINAL PROJECT

Semi supervised learning is a branch of machine learning that combines small number of labeled data with large number of unlabeled data during the model training. It comes in between the supervised and unsupervised learning approaches.

1. **a self-training algorithm**

   Self-training algorithm's working is like that of supervised learning, in this the classifier is trained on the labelled data and the trained classifier is used to predict the outcome of the unlabelled data.

   STEPS:

   1. The classifier is trained on the labelled data, the trained classifier is used to predict the outcome of the unlabeled data, the most confident predictions are included in the training data. (Either based on a particular confidence threshold or taking the top n confident results that is also known as k_best results)
   2. The above step is continued until all the data becomes labelled data.
   3. The model trained using the final trained data will be the resultant model.

   The above experiment is done with different amount of unlabeled data (0%, 10%, 20%, 50%, 90%, 95%) and we analyze the interplay between the amount of unlabeled data and the performance of the resultant model. Also, the same experiment is done on the three datasets.

   Experiment Setup

   The best model and hyperparameters pair for a particular preprocessed dataset is being used as the base classifier of the self-training classifier. Three datasets are used namely heart disease dataset, marketing dataset, online shoppers' intention dataset. The accuracy, F1 score, ROC curve and runtime is calculated for each experiment.

   For heart disease dataset and online shoppers' intention dataset the classifier of choice is

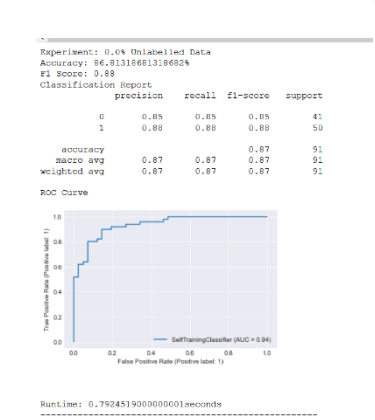
   MLP classifier along with hyperparameters as: activation = 'tanh', alpha = 0.0001, hidden_layer_sizes = (50, 100, 50), learning_rate = 'constant', solver = 'sgd',

   For Marketing dataset, the classifier of choice is: RandomForestClassifier

   With hyperparameters as: max_depth = 15, max_features = 'sqrt', min_samples_leaf = 2, min_samples_split = 2, n_estimators = 100, random_state = 42)

Results Obtained.

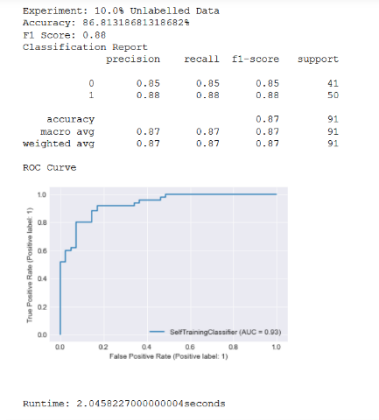
Heart Disease Dataset



```
Experiment: 0.0% Unlabelled Data
Accuracy: 86.81318681318682%
F1 Score: 0.88
Classification Report
              precision    recall  f1-score   support

           0       0.85      0.85      0.85        41
           1       0.88      0.88      0.88        50

    accuracy                           0.87        91
   macro avg       0.87      0.87      0.87        91
weighted avg       0.87      0.87      0.87        91

ROC Curve
```



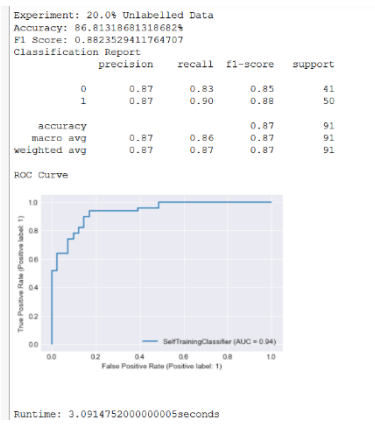
```
Runtime: 0.7924519000000001seconds
```



```
Experiment: 10.0% Unlabelled Data
Accuracy: 86.81318681318682%
F1 Score: 0.88
Classification Report
              precision    recall  f1-score   support

           0       0.85      0.85      0.85        41
           1       0.88      0.88      0.88        50

    accuracy                           0.87        91
   macro avg       0.87      0.87      0.87        91
weighted avg       0.87      0.87      0.87        91

ROC Curve
```



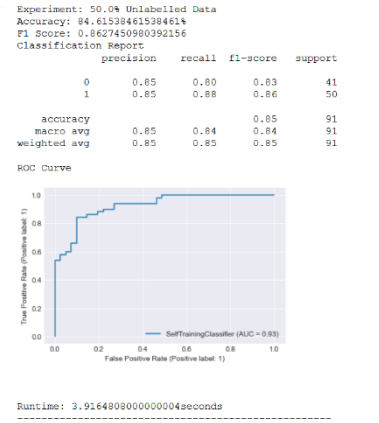
```
Runtime: 2.0458227000000004seconds
```



```
Experiment: 20.0% Unlabelled Data
Accuracy: 86.81318681318682%
F1 Score: 0.8823529411764707
Classification Report
              precision    recall  f1-score   support

           0       0.87      0.83      0.85        41
           1       0.87      0.90      0.88        50

    accuracy                           0.87        91
   macro avg       0.87      0.86      0.87        91
weighted avg       0.87      0.87      0.87        91

ROC Curve
```



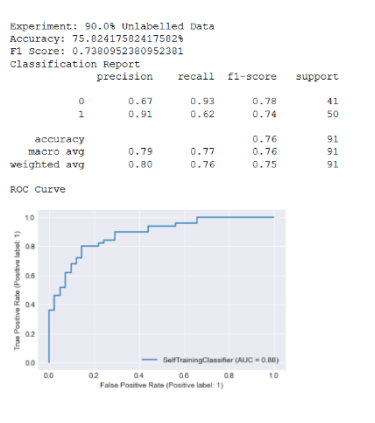
```
Runtime: 3.0914752000000005seconds
```



```
Experiment: 50.0% Unlabelled Data
Accuracy: 84.61538461538461%
F1 Score: 0.8627450980392156
Classification Report
              precision    recall  f1-score   support

           0       0.85      0.80      0.83        41
           1       0.85      0.88      0.86        50

    accuracy                           0.85        91
   macro avg       0.85      0.84      0.84        91
weighted avg       0.85      0.85      0.85        91

ROC Curve
```



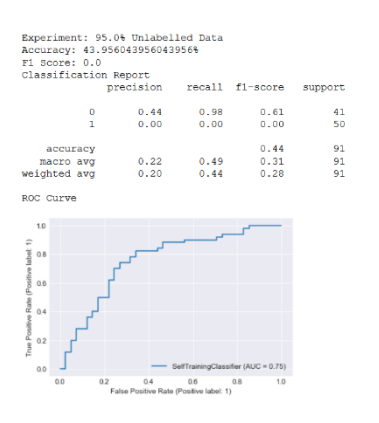
```
Runtime: 3.9164080000000004seconds
```



```
Experiment: 90.0% Unlabelled Data
Accuracy: 75.82417582417582%
F1 Score: 0.7380952380952381
Classification Report
              precision    recall  f1-score   support

           0       0.67      0.93      0.78        41
           1       0.91      0.62      0.74        50

    accuracy                           0.76        91
   macro avg       0.79      0.77      0.76        91
weighted avg       0.80      0.76      0.75        91

ROC Curve
```



```
Runtime: 2.6321207000000015seconds
```



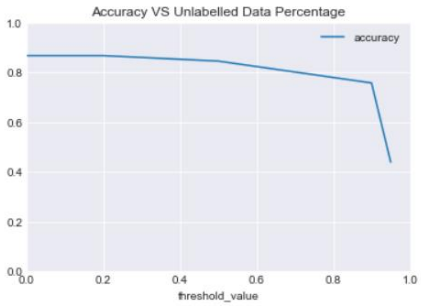
```
Experiment: 95.0% Unlabelled Data
Accuracy: 43.956043956043956%
F1 Score: 0.0
Classification Report
              precision    recall  f1-score   support

           0       0.44      0.98      0.61        41
           1       0.00      0.00      0.00        50

    accuracy                           0.44        91
   macro avg       0.22      0.49      0.31        91
weighted avg       0.20      0.44      0.28        91

ROC Curve
```

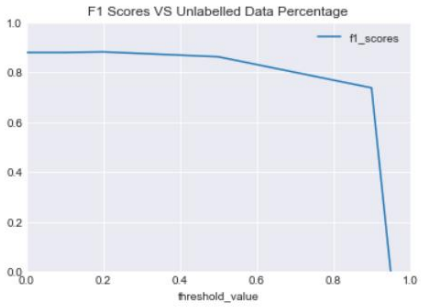


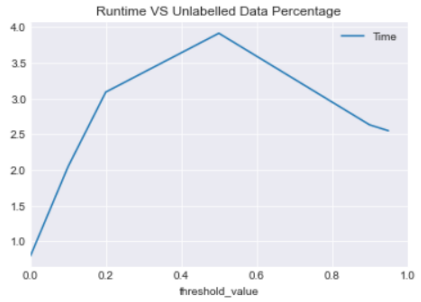
```
Runtime: 2.5493887000000015seconds
```

Here we can see the performance evaluation for different experiments done.

Above, the evaluation metrics for different levels of unlabeled data is being shown.







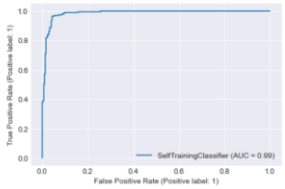
Here we can see the trend of accuracy, f1-score, and runtime with different amount of unlabeled data and can see that the accuracy remains like the supervised experiment until the point where the 90% of data is unlabeled and it drops sharply at 95% amount of unlabeled data.

Marketing Dataset.

Experiment: 0.0% Unlabelled Data
Accuracy: 95.684523809523815%
F1 Score: 0.9567809239940388
Classification Report

	precision	recall	f1-score	support
0	0.96	0.96	0.96	337
1	0.96	0.96	0.96	335
accuracy			0.96	672
macro avg	0.96	0.96	0.96	672
weighted avg	0.96	0.96	0.96	672

ROC Curve

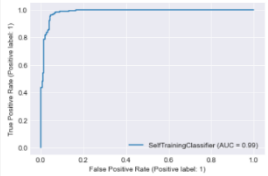


Runtime: 2.6968980000000015seconds

Experiment: 10.0% Unlabelled Data
Accuracy: 95.38690476190477%
F1 Score: 0.9538002980625933
Classification Report

	precision	recall	f1-score	support
0	0.96	0.95	0.95	337
1	0.95	0.96	0.95	335
accuracy			0.95	672
macro avg	0.95	0.95	0.95	672
weighted avg	0.95	0.95	0.95	672

ROC Curve

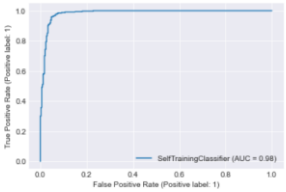


Runtime: 4.737451799999999seconds

Experiment: 20.0% Unlabelled Data
Accuracy: 95.386904761904774%
F1 Score: 0.9539375928677564
Classification Report

	precision	recall	f1-score	support
0	0.96	0.95	0.95	337
1	0.95	0.96	0.95	335
accuracy			0.95	672
macro avg	0.95	0.95	0.95	672
weighted avg	0.95	0.95	0.95	672

ROC Curve

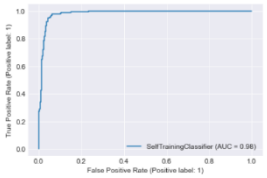


Runtime: 4.593247999999999seconds

Experiment: 50.0% Unlabelled Data
Accuracy: 95.089285714285715%
F1 Score: 0.9509658246656759
Classification Report

	precision	recall	f1-score	support
0	0.96	0.95	0.95	337
1	0.95	0.96	0.95	335
accuracy			0.95	672
macro avg	0.95	0.95	0.95	672
weighted avg	0.95	0.95	0.95	672

ROC Curve

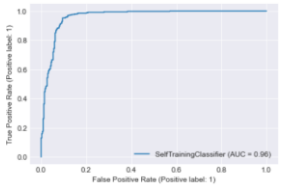


Runtime: 4.155937399999999seconds

Experiment: 90.0% Unlabelled Data
Accuracy: 91.666666666666666%
F1 Score: 0.9174041297935103
Classification Report

	precision	recall	f1-score	support
0	0.93	0.91	0.92	337
1	0.91	0.93	0.92	335
accuracy			0.92	672
macro avg	0.92	0.92	0.92	672
weighted avg	0.92	0.92	0.92	672

ROC Curve

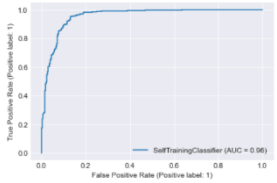


Runtime: 3.5231952999999976seconds

Experiment: 95.0% Unlabelled Data
Accuracy: 89.69047619047619%
F1 Score: 0.8823529411764706
Classification Report

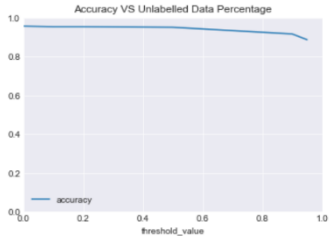
	precision	recall	f1-score	support
0	0.86	0.92	0.89	337
1	0.92	0.85	0.88	335
accuracy			0.89	672
macro avg	0.89	0.89	0.89	672
weighted avg	0.89	0.89	0.89	672

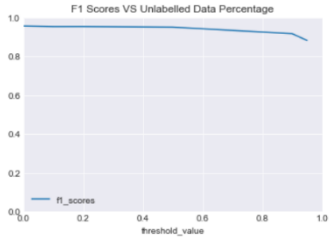
ROC Curve

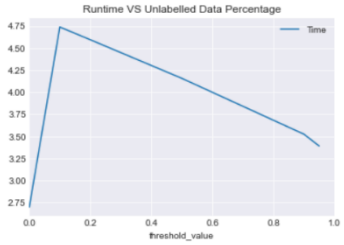


Runtime: 3.389029200000003seconds

Performance evaluation Is shown for different experiments







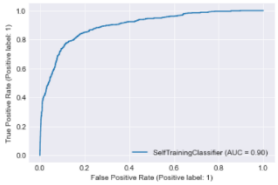
Here we can see the trend of accuracy, f1-score, and runtime with different amount of unlabeled data and can see that the accuracy remains like the supervised experiment until the point where the 90% of data is unlabeled and it drops a little at 95% amount of unlabeled data.

Online Shoppers intention dataset(imbalanced).

Experiment: 0.0% Unlabelled Data
Accuracy: 88.48337388483374%
F1 Score: 0.6077348066298343
Classification Report

	precision	recall	f1-score	support
0	0.92	0.95	0.93	3114
1	0.66	0.56	0.61	585
accuracy			0.88	3699
macro avg	0.79	0.75	0.77	3699
weighted avg	0.88	0.88	0.88	3699

ROC Curve

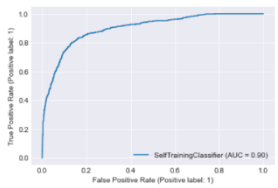


Runtime: 20.0150719999999965seconds

Experiment: 10.0% Unlabelled Data
Accuracy: 88.7537172207051%
F1 Score: 0.5961165048543688
Classification Report

	precision	recall	f1-score	support
0	0.91	0.96	0.93	3114
1	0.69	0.52	0.60	585
accuracy			0.89	3699
macro avg	0.80	0.74	0.77	3699
weighted avg	0.88	0.89	0.88	3699

ROC Curve

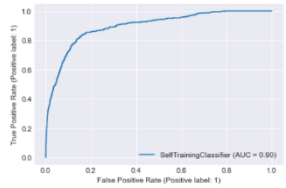


Runtime: 195.0724231seconds

Experiment: 20.0% Unlabelled Data
Accuracy: 88.48337388483374%
F1 Score: 0.5773809523809524
Classification Report

	precision	recall	f1-score	support
0	0.91	0.96	0.93	3114
1	0.69	0.50	0.58	585
accuracy			0.88	3699
macro avg	0.80	0.73	0.76	3699
weighted avg	0.88	0.88	0.88	3699

ROC Curve

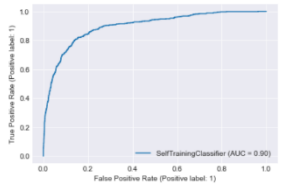


Runtime: 152.553824400000005seconds

Experiment: 50.0% Unlabelled Data
Accuracy: 89.10516355771834%
F1 Score: 0.6013847675568744
Classification Report

	precision	recall	f1-score	support
0	0.91	0.96	0.94	3114
1	0.71	0.52	0.60	585
accuracy			0.89	3699
macro avg	0.81	0.74	0.77	3699
weighted avg	0.88	0.89	0.88	3699

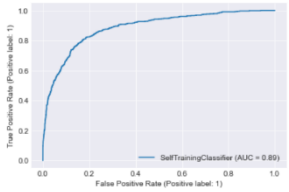
ROC Curve



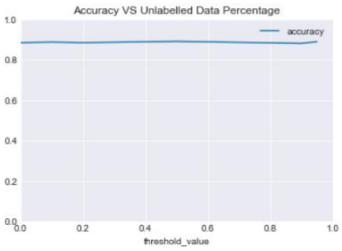
Runtime: 85.98173889999998seconds

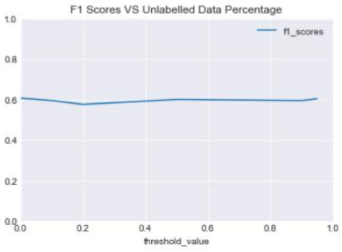
```
Experiment: 90.0% Unlabelled Data
Accuracy: 88.15896188158962%
F1 Score: 0.5959409594095941
Classification Report
              precision    recall  f1-score   support

           0       0.92      0.94      0.93      3114
           1       0.65      0.55      0.60       585

    accuracy                           0.88      3699
   macro avg       0.78      0.75      0.76      3699
weighted avg       0.88      0.88      0.88      3699
```

ROC Curve



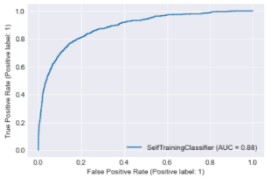
Runtime: 21.365833799999996seconds



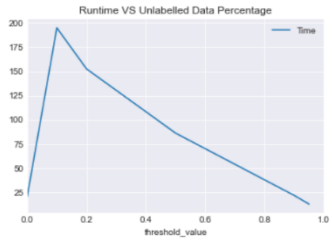



```
Experiment: 95.0% Unlabelled Data
Accuracy: 88.91592322249257%
F1 Score: 0.6042471042471043
Classification Report
              precision    recall  f1-score   support

           0       0.92      0.96      0.94      3114
           1       0.69      0.54      0.60       585

    accuracy                           0.89      3699
   macro avg       0.81      0.75      0.77      3699
weighted avg       0.88      0.89      0.88      3699
```

ROC Curve



Runtime: 12.341788000000065seconds



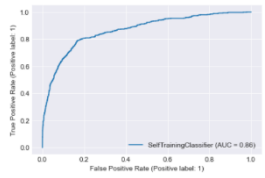
Here we can see that the accuracy and the f1 scores are pretty much the same for all the experiments.

Online Shoppers intention dataset (Balanced - Up sampled).

Experiment: 0.0% Unlabelled Data
Accuracy: 83.96864017301974%
F1 Score: 0.5896193771626297
Classification Report

	precision	recall	f1-score	support
0	0.94	0.86	0.90	3114
1	0.50	0.73	0.59	585
accuracy			0.84	3699
macro avg	0.72	0.79	0.75	3699
weighted avg	0.87	0.84	0.85	3699

ROC Curve

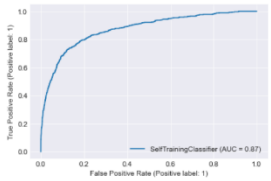


Runtime: 32.776894999999975seconds

Experiment: 10.0% Unlabelled Data
Accuracy: 83.59015950256826%
F1 Score: 0.5906945380984491
Classification Report

	precision	recall	f1-score	support
0	0.95	0.85	0.90	3114
1	0.49	0.75	0.59	585
accuracy			0.84	3699
macro avg	0.72	0.80	0.74	3699
weighted avg	0.87	0.84	0.85	3699

ROC Curve

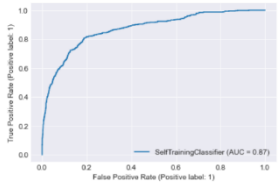


Runtime: 335.1626057seconds

Experiment: 20.0% Unlabelled Data
Accuracy: 83.48202216815356%
F1 Score: 0.583503749147921
Classification Report

	precision	recall	f1-score	support
0	0.94	0.85	0.90	3114
1	0.49	0.73	0.58	585
accuracy			0.83	3699
macro avg	0.71	0.79	0.74	3699
weighted avg	0.87	0.83	0.85	3699

ROC Curve

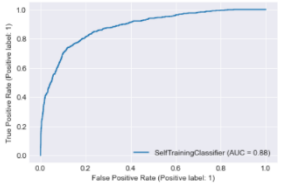


Runtime: 308.30809080000006seconds

Experiment: 50.0% Unlabelled Data
Accuracy: 83.40091916734252%
F1 Score: 0.5944517833553502
Classification Report

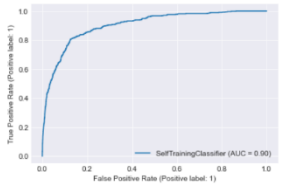
	precision	recall	f1-score	support
0	0.95	0.85	0.90	3114
1	0.48	0.77	0.59	585
accuracy			0.83	3699
macro avg	0.72	0.81	0.75	3699
weighted avg	0.88	0.83	0.85	3699

ROC Curve



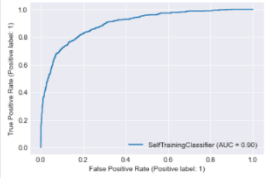
Runtime: 182.984272899999995seconds

```
Experiment: 90.0% Unlabelled Data
Accuracy: 84.86077318194106%
F1 Score: 0.6310935441370225
Classification Report
              precision    recall  f1-score   support

           0       0.96      0.85      0.90      3114
           1       0.51      0.82      0.63       585

    accuracy                           0.85      3699
   macro avg       0.74      0.84      0.77      3699
weighted avg       0.89      0.85      0.86      3699

ROC Curve
```

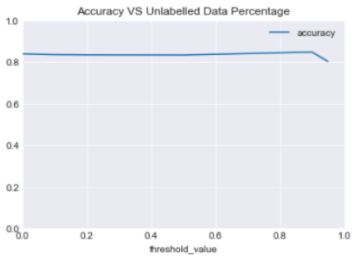


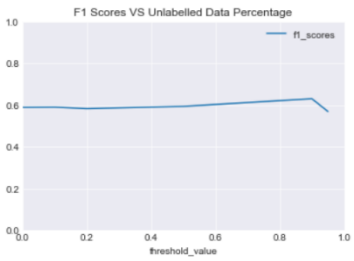

```
Runtime: 33.79106069999989seconds
```

```
Experiment: 95.0% Unlabelled Data
Accuracy: 80.31900513652339%
F1 Score: 0.5697399527186762
Classification Report
              precision    recall  f1-score   support

           0       0.96      0.80      0.87      3114
           1       0.44      0.82      0.57       585

    accuracy                           0.80      3699
   macro avg       0.70      0.81      0.72      3699
weighted avg       0.88      0.80      0.82      3699

ROC Curve
```




```
Runtime: 19.71918430000005seconds
```





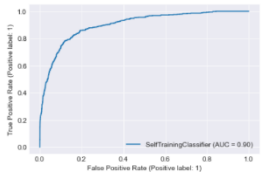
Runtime VS Unlabelled Data Percentage

Online Shoppers intention dataset (Balanced – Down sampled)

Experiment: 0.0% Unlabelled Data
Accuracy: 83.56312516896458%
F1 Score: 0.6127388535031847
Classification Report

	precision	recall	f1-score	support
0	0.96	0.84	0.90	3114
1	0.49	0.82	0.61	585
accuracy			0.84	3699
macro avg	0.73	0.83	0.75	3699
weighted avg	0.89	0.84	0.85	3699

ROC Curve

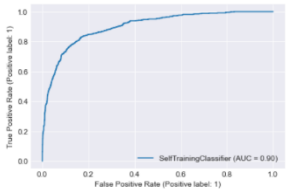


Runtime: 9.218825700000025seconds

Experiment: 10.0% Unlabelled Data
Accuracy: 83.022438496891064
F1 Score: 0.6075
Classification Report

	precision	recall	f1-score	support
0	0.96	0.83	0.89	3114
1	0.48	0.83	0.61	585
accuracy			0.83	3699
macro avg	0.72	0.83	0.75	3699
weighted avg	0.89	0.83	0.85	3699

ROC Curve

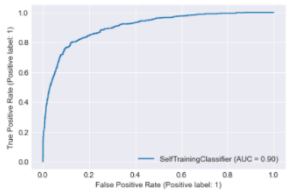


Runtime: 49.654091699999984seconds

Experiment: 20.0% Unlabelled Data
Accuracy: 82.69802649364694%
F1 Score: 0.6024844720496894
Classification Report

	precision	recall	f1-score	support
0	0.96	0.83	0.89	3114
1	0.47	0.83	0.60	585
accuracy			0.83	3699
macro avg	0.72	0.83	0.75	3699
weighted avg	0.89	0.83	0.84	3699

ROC Curve

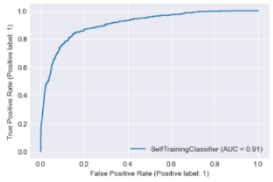


Runtime: 43.977011100000003seconds

Experiment: 50.0% Unlabelled Data
Accuracy: 85.99621519329548%
F1 Score: 0.6447187928669409
Classification Report

	precision	recall	f1-score	support
0	0.96	0.87	0.91	3114
1	0.54	0.80	0.64	585
accuracy			0.86	3699
macro avg	0.75	0.84	0.78	3699
weighted avg	0.89	0.86	0.87	3699

ROC Curve

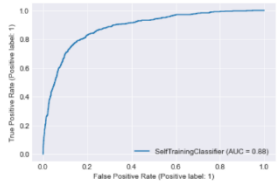


Runtime: 29.771075699999983seconds

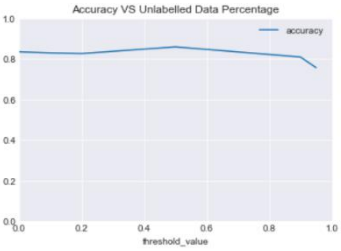
Experiment: 90.0% Unlabelled Data
Accuracy: 80.99486347661538
F1 Score: 0.5741974560872198
Classification Report

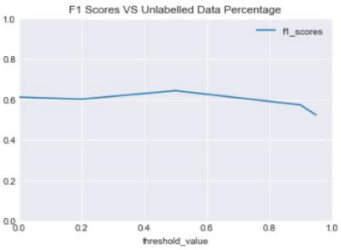
	precision	recall	f1-score	support
0	0.96	0.81	0.88	3114
1	0.44	0.81	0.57	585
accuracy			0.81	3699
macro avg	0.70	0.81	0.73	3699
weighted avg	0.88	0.81	0.83	3699

ROC Curve



Runtime: 8.1549448000000067seconds

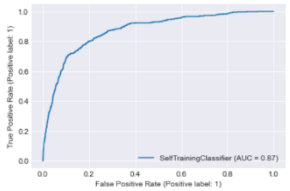




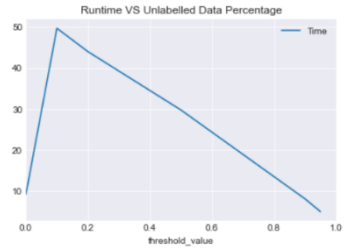
Experiment: 95.0% Unlabelled Data
Accuracy: 75.75020275750202%
F1 Score: 0.5241379310344828
Classification Report

	precision	recall	f1-score	support
0	0.96	0.74	0.84	3114
1	0.38	0.84	0.52	585
accuracy			0.76	3699
macro avg	0.67	0.79	0.68	3699
weighted avg	0.87	0.76	0.79	3699

ROC Curve



Runtime: 5.024844599999824seconds



Conclusion:

We can say that even with less amount of labeled data and more amount of unlabeled data the model's performance is remaining similar. The task of labelling is expensive in real world and thus getting higher performances from model even with large amount of unlabeled data is a boon.

2. A Semi Supervised Ensemble Approach

In statistics and machine learning, ensemble methods use multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone.

STEPS:

1. The base classifiers are trained on the unlabeled data using pseudo labelling approach (self-training) and the majority voting is used to make predictions
2. The above step is continued until all the data becomes labelled data.
3. The model trained using the final trained data will be the resultant model.

The above experiment is done with different amount of unlabeled data (0%, 10%, 20%, 50%, 90%, 95%) and we analyze the interplay between the amount of unlabeled data and the performance of the resultant model. Also, the same experiment is done on the three datasets.

Experiment Setup

The base models are wrapped under self-training classifier module and these base models are used to form an ensemble which is then run on the data with different percentages of unlabeled records.

The random forest classifier is used as the based classifier.

Results Obtained

Heart Disease Dataset.

```
Experiment: 0.0% Unlabelled Data
Accuracy: 81.318681318681318%
F1 Score: 0.8316831683168315
Classification Report
              precision    recall  f1-score   support

           0       0.80      0.78      0.79        41
           1       0.82      0.84      0.83        50

    accuracy                           0.81        91
   macro avg       0.81      0.81      0.81        91
weighted avg       0.81      0.81      0.81        91

ROC Curve
```

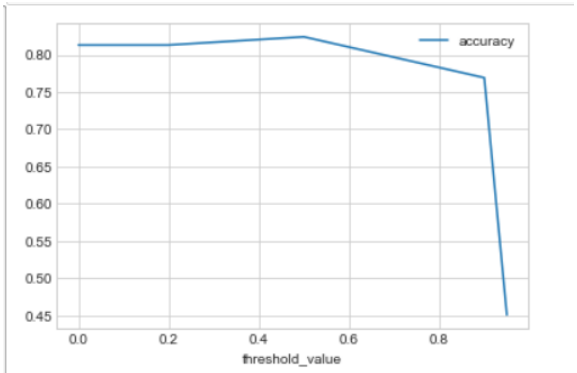
```
Experiment: 10.0% Unlabelled Data
Accuracy: 81.318681318681318%
F1 Score: 0.8316831683168315
Classification Report
              precision    recall  f1-score   support

           0       0.80      0.78      0.79        41
           1       0.82      0.84      0.83        50

    accuracy                           0.81        91
   macro avg       0.81      0.81      0.81        91
weighted avg       0.81      0.81      0.81        91

ROC Curve
```

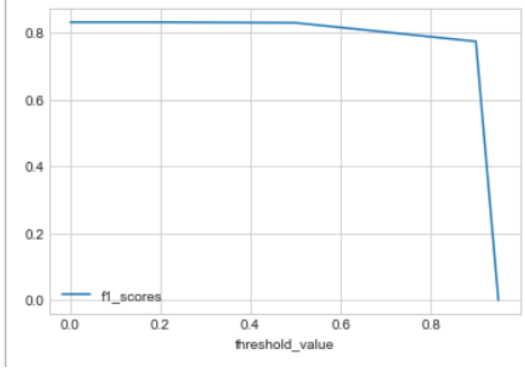
```
Experiment: 20.0% Unlabelled Data
Accuracy: 81.318681318681318%
F1 Score: 0.8316831683168315
Classification Report
              precision    recall  f1-score   support

           0       0.80      0.78      0.79        41
           1       0.82      0.84      0.83        50

    accuracy                           0.81        91
   macro avg       0.81      0.81      0.81        91
weighted avg       0.81      0.81      0.81        91

ROC Curve
```

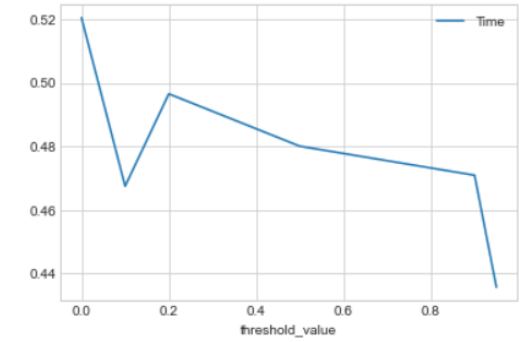
```
Experiment: 50.0% Unlabelled Data
Accuracy: 82.417582417582418%
F1 Score: 0.8297872340425532
Classification Report
              precision    recall  f1-score   support

           0       0.77      0.88      0.82        41
           1       0.89      0.78      0.83        50

    accuracy                           0.82        91
   macro avg       0.83      0.83      0.82        91
weighted avg       0.83      0.82      0.82        91

ROC Curve
```

```
Experiment: 90.0% Unlabelled Data
Accuracy: 76.923076923076923%
F1 Score: 0.7741935483870969
Classification Report
              precision    recall  f1-score   support

           0       0.71      0.83      0.76        41
           1       0.84      0.72      0.77        50

    accuracy                           0.77        91
   macro avg       0.77      0.77      0.77        91
weighted avg       0.78      0.77      0.77        91

ROC Curve
```

```
Experiment: 95.0% Unlabelled Data
Accuracy: 45.054945054945054%
F1 Score: 0.0
Classification Report
              precision    recall  f1-score   support

           0       0.45      1.00      0.62        41
           1       0.00      0.00      0.00        50

    accuracy                           0.45        91
   macro avg       0.23      0.50      0.31        91
weighted avg       0.20      0.45      0.28        91

ROC Curve
```


accuracy
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.0
0.2
0.4
0.6
0.8
threshold_value


0.8
0.6
0.4
0.2
0.0
f1_scores
0.0
0.2
0.4
0.6
0.8
threshold_value


Time
0.52
0.50
0.48
0.46
0.44
0.0
0.2
0.4
0.6
0.8
threshold_value

Marketing Dataset.

Experiment: 0.0% Unlabelled Data
Accuracy: 95.684523809523814%
F1 Score: 0.9567809239940388
Classification Report

	precision	recall	f1-score	support
0	0.96	0.96	0.96	337
1	0.96	0.96	0.96	335
accuracy			0.96	672
macro avg	0.96	0.96	0.96	672
weighted avg	0.96	0.96	0.96	672

ROC Curve

Experiment: 10.0% Unlabelled Data
Accuracy: 95.089285714285714%
F1 Score: 0.9512555391432792
Classification Report

	precision	recall	f1-score	support
0	0.96	0.94	0.95	337
1	0.94	0.96	0.95	335
accuracy			0.95	672
macro avg	0.95	0.95	0.95	672
weighted avg	0.95	0.95	0.95	672

ROC Curve

Experiment: 20.0% Unlabelled Data
Accuracy: 94.940476190476195%
F1 Score: 0.9492537313432836
Classification Report

	precision	recall	f1-score	support
0	0.95	0.95	0.95	337
1	0.95	0.95	0.95	335
accuracy			0.95	672
macro avg	0.95	0.95	0.95	672
weighted avg	0.95	0.95	0.95	672

ROC Curve

Experiment: 50.0% Unlabelled Data
Accuracy: 95.386904761904773%
F1 Score: 0.9542097488921715
Classification Report

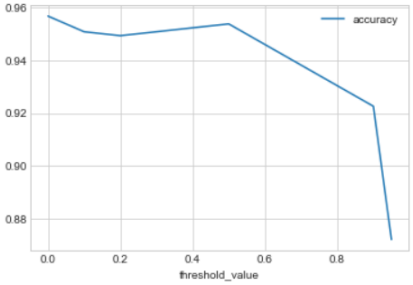
	precision	recall	f1-score	support
0	0.96	0.94	0.95	337
1	0.94	0.96	0.95	335
accuracy			0.95	672
macro avg	0.95	0.95	0.95	672
weighted avg	0.95	0.95	0.95	672

ROC Curve

Experiment: 90.0% Unlabelled Data
Accuracy: 92.261904761904773%
F1 Score: 0.9235294117647059
Classification Report

	precision	recall	f1-score	support
0	0.94	0.91	0.92	337
1	0.91	0.94	0.92	335
accuracy			0.92	672
macro avg	0.92	0.92	0.92	672
weighted avg	0.92	0.92	0.92	672

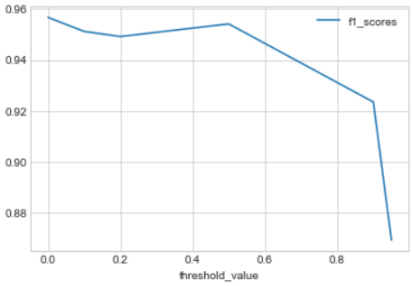
ROC Curve

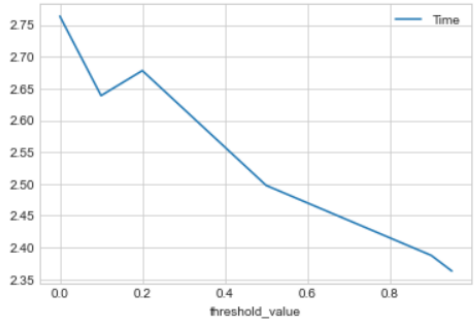


Experiment: 95.0% Unlabelled Data
Accuracy: 87.202380952380952%
F1 Score: 0.8693009118541033
Classification Report

	precision	recall	f1-score	support
0	0.86	0.89	0.87	337
1	0.89	0.85	0.87	335
accuracy			0.87	672
macro avg	0.87	0.87	0.87	672
weighted avg	0.87	0.87	0.87	672

ROC Curve





Online Shopper Intention – imbalanced

```
Experiment: 0.0% Unlabelled Data
Accuracy: 89.051094890510 95%
F1 Score: 0.6033300685602352
Classification Report
              precision    recall  f1-score   support

           0       0.92      0.96      0.94      3114
           1       0.71      0.53      0.60       585

    accuracy                           0.89      3699
   macro avg       0.81      0.74      0.77      3699
weighted avg       0.88      0.89      0.88      3699

ROC Curve
```

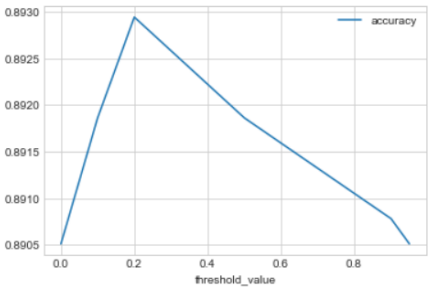
```
Experiment: 10.0% Unlabelled Data
Accuracy: 89.18626655852934%
F1 Score: 0.6146435452793834
Classification Report
              precision    recall  f1-score   support

           0       0.92      0.96      0.94      3114
           1       0.70      0.55      0.61       585

    accuracy                           0.89      3699
   macro avg       0.81      0.75      0.78      3699
weighted avg       0.88      0.89      0.89      3699

ROC Curve
```

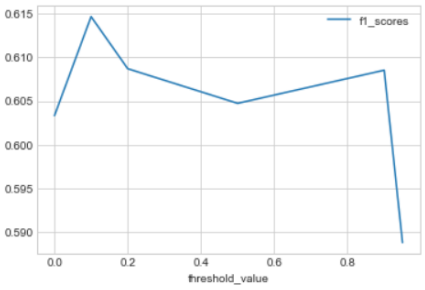
```
Experiment: 20.0% Unlabelled Data
Accuracy: 89.29440389294405%
F1 Score: 0.6086956521739131
Classification Report
              precision    recall  f1-score   support

           0       0.92      0.96      0.94      3114
           1       0.72      0.53      0.61       585

    accuracy                           0.89      3699
   macro avg       0.82      0.74      0.77      3699
weighted avg       0.88      0.89      0.89      3699

ROC Curve
```

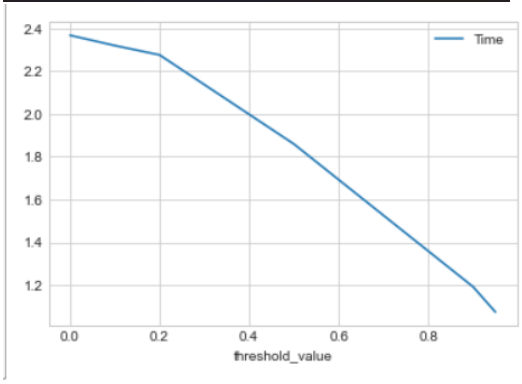
```
Experiment: 50.0% Unlabelled Data
Accuracy: 89.18626655852934%
F1 Score: 0.6047430830039526
Classification Report
              precision    recall  f1-score   support

           0       0.91      0.96      0.94      3114
           1       0.72      0.52      0.60       585

    accuracy                           0.89      3699
   macro avg       0.82      0.74      0.77      3699
weighted avg       0.88      0.89      0.88      3699

ROC Curve
```

```
Experiment: 90.0% Unlabelled Data
Accuracy: 89.07812922411462%
F1 Score: 0.6085271317829457
Classification Report
              precision    recall  f1-score   support

           0       0.92      0.96      0.94      3114
           1       0.70      0.54      0.61       585

    accuracy                           0.89      3699
   macro avg       0.81      0.75      0.77      3699
weighted avg       0.88      0.89      0.88      3699

ROC Curve
```

```
Experiment: 95.0% Unlabelled Data
Accuracy: 89.05109489051095%
F1 Score: 0.5888324873096447
Classification Report
              precision    recall  f1-score   support

           0       0.91      0.96      0.94      3114
           1       0.72      0.50      0.59       585

    accuracy                           0.89      3699
   macro avg       0.82      0.73      0.76      3699
weighted avg       0.88      0.89      0.88      3699

ROC Curve
```







Online shopper intention – balanced up sampled

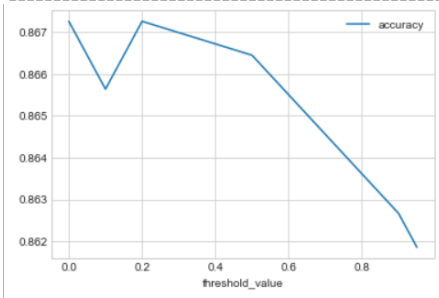
```
Experiment: 0.0% Unlabelled Data
Accuracy: 86.72614220059475%
F1 Score: 0.6449746926970354
Classification Report
              precision    recall  f1-score   support

           0       0.95      0.89      0.92      3114
           1       0.56      0.76      0.64       585

    accuracy                           0.87      3699
   macro avg       0.76      0.82      0.78      3699
weighted avg       0.89      0.87      0.88      3699

ROC Curve
```

```
Experiment: 20.0% Unlabelled Data
Accuracy: 86.72614220059475%
F1 Score: 0.6505338078291815
Classification Report
              precision    recall  f1-score   support

           0       0.96      0.88      0.92      3114
           1       0.56      0.78      0.65       585

    accuracy                           0.87      3699
   macro avg       0.76      0.83      0.78      3699
weighted avg       0.89      0.87      0.88      3699

ROC Curve
```

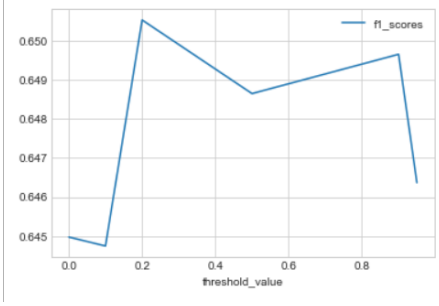
```
Experiment: 10.0% Unlabelled Data
Accuracy: 86.5639361989727%
F1 Score: 0.6447462473195139
Classification Report
              precision    recall  f1-score   support

           0       0.95      0.88      0.92      3114
           1       0.55      0.77      0.64       585

    accuracy                           0.87      3699
   macro avg       0.75      0.83      0.78      3699
weighted avg       0.89      0.87      0.87      3699

ROC Curve
```

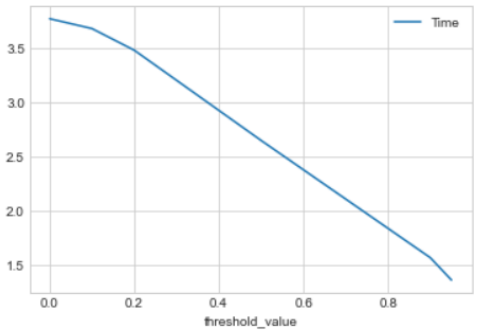
```
Experiment: 50.0% Unlabelled Data
Accuracy: 86.64503919978372%
F1 Score: 0.6486486486486488
Classification Report
              precision    recall  f1-score   support

           0       0.96      0.88      0.92      3114
           1       0.56      0.78      0.65       585

    accuracy                           0.87      3699
   macro avg       0.76      0.83      0.78      3699
weighted avg       0.89      0.87      0.88      3699

ROC Curve
```

```
Experiment: 90.0% Unlabelled Data
Accuracy: 86.26655852933224%
F1 Score: 0.6496551724137932
Classification Report
              precision    recall  f1-score   support

           0       0.96      0.87      0.91      3114
           1       0.54      0.81      0.65       585

    accuracy                           0.86      3699
   macro avg       0.75      0.84      0.78      3699
weighted avg       0.89      0.86      0.87      3699

ROC Curve
```




```
Experiment: 95.0% Unlabelled Data
Accuracy: 86.18545552852123%
F1 Score: 0.6463667820069203
Classification Report
              precision    recall  f1-score   support

           0       0.96      0.87      0.91      3114
           1       0.54      0.80      0.65       585

    accuracy                           0.86      3699
   macro avg       0.75      0.84      0.78      3699
weighted avg       0.89      0.86      0.87      3699

ROC Curve
```





Online shopper intention – balanced down sampled

Experiment: 0.0% Unlabelled Data
Accuracy: 84.18491484184915%
F1 Score: 0.6257197696737044
Classification Report

	precision	recall	f1-score	support
0	0.96	0.84	0.90	3114
1	0.50	0.84	0.63	585
accuracy			0.84	3699
macro avg	0.73	0.84	0.76	3699
weighted avg	0.89	0.84	0.86	3699

ROC Curve

Experiment: 10.0% Unlabelled Data
Accuracy: 84.13084617464179%
F1 Score: 0.6244401791426745
Classification Report

	precision	recall	f1-score	support
0	0.96	0.84	0.90	3114
1	0.50	0.83	0.62	585
accuracy			0.84	3699
macro avg	0.73	0.84	0.76	3699
weighted avg	0.89	0.84	0.86	3699

ROC Curve

Experiment: 20.0% Unlabelled Data
Accuracy: 82.9683698296837%
F1 Score: 0.6091811414392059
Classification Report

	precision	recall	f1-score	support
0	0.96	0.83	0.89	3114
1	0.48	0.84	0.61	585
accuracy			0.83	3699
macro avg	0.72	0.83	0.75	3699
weighted avg	0.89	0.83	0.85	3699

ROC Curve

Experiment: 50.0% Unlabelled Data
Accuracy: 83.860502838605503%
F1 Score: 0.6199872692552515
Classification Report

	precision	recall	f1-score	support
0	0.96	0.84	0.90	3114
1	0.49	0.83	0.62	585
accuracy			0.84	3699
macro avg	0.73	0.84	0.76	3699
weighted avg	0.89	0.84	0.85	3699

ROC Curve

Experiment: 90.0% Unlabelled Data
Accuracy: 79.994593133279278%
F1 Score: 0.5756880733944955
Classification Report

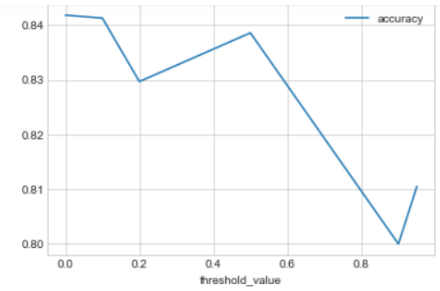
	precision	recall	f1-score	support
0	0.97	0.79	0.87	3114
1	0.43	0.86	0.58	585
accuracy			0.80	3699
macro avg	0.70	0.82	0.72	3699
weighted avg	0.88	0.80	0.82	3699

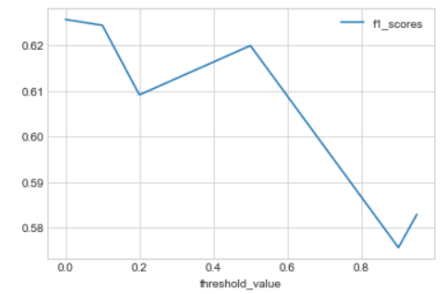
ROC Curve

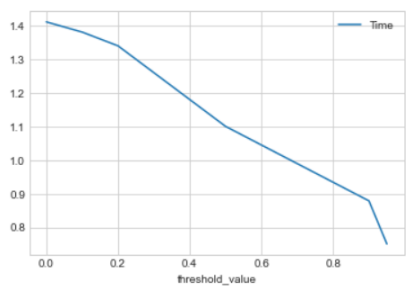
Experiment: 95.0% Unlabelled Data
Accuracy: 81.048932143822658%
F1 Score: 0.5829863176680548
Classification Report

	precision	recall	f1-score	support
0	0.96	0.81	0.88	3114
1	0.45	0.84	0.58	585
accuracy			0.81	3699
macro avg	0.71	0.82	0.73	3699
weighted avg	0.88	0.81	0.83	3699

ROC Curve







3. Intrinsically semi supervised learning method

These methods, which we call intrinsically semi-supervised, do not rely on any intermediate steps, or supervised base learners. Usually, they are extensions of existing supervised methods to include unlabeled samples in the objective function.

STEPS:

Pseudo-labelling as a form of margin maximization

when using self-training with supervised SVMs, the decision boundary is iteratively pushed away from the unlabeled samples. Even though the unlabeled data are not explicitly incorporated into the loss function, this amounts to exploiting the low-density assumption, as done in the case of S3VMs.

The above experiment is done with different amount of unlabeled data (0%, 10%, 20%, 50%, 90%, 95%) and we analyze the interplay between the amount of unlabeled data and the performance of the resultant model. Also, the same experiment is done on the three datasets.

Experiment Setup

The SVM classifier is used as the base classifier.

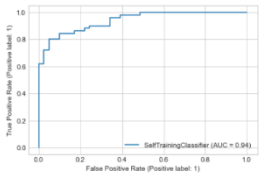
Results Obtained

Heart Disease Dataset

Experiment: 0.0% Unlabelled Data
Accuracy: 84.61538461538461%
F1 Score: 0.8599999999999999
Classification Report

	precision	recall	f1-score	support
0	0.83	0.83	0.83	41
1	0.86	0.86	0.86	50
accuracy			0.85	91
macro avg	0.84	0.84	0.84	91
weighted avg	0.85	0.85	0.85	91

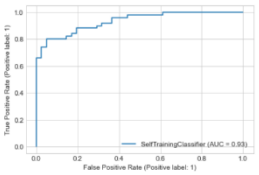
ROC Curve



Experiment: 10.0% Unlabelled Data
Accuracy: 83.51648351648352%
F1 Score: 0.8514851485148515
Classification Report

	precision	recall	f1-score	support
0	0.82	0.80	0.81	41
1	0.84	0.86	0.85	50
accuracy			0.84	91
macro avg	0.83	0.83	0.83	91
weighted avg	0.83	0.84	0.83	91

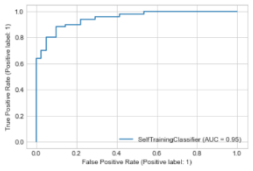
ROC Curve



Experiment: 20.0% Unlabelled Data
Accuracy: 85.71428571428571%
F1 Score: 0.8785046728971964
Classification Report

	precision	recall	f1-score	support
0	0.91	0.76	0.83	41
1	0.82	0.94	0.88	50
accuracy			0.86	91
macro avg	0.87	0.85	0.85	91
weighted avg	0.86	0.86	0.86	91

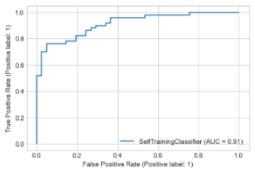
ROC Curve



Experiment: 50.0% Unlabelled Data
Accuracy: 80.21978021978022%
F1 Score: 0.8421052631578947
Classification Report

	precision	recall	f1-score	support
0	0.93	0.61	0.74	41
1	0.75	0.96	0.84	50
accuracy			0.80	91
macro avg	0.84	0.78	0.79	91
weighted avg	0.83	0.80	0.79	91

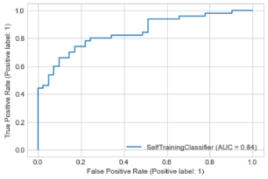
ROC Curve



Experiment: 90.0% Unlabelled Data
Accuracy: 67.03296703296702%
F1 Score: 0.7619047619047619
Classification Report

	precision	recall	f1-score	support
0	0.87	0.32	0.46	41
1	0.63	0.96	0.76	50
accuracy			0.67	91
macro avg	0.75	0.64	0.61	91
weighted avg	0.74	0.67	0.63	91

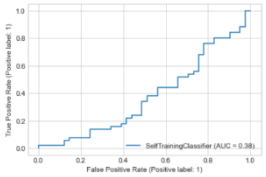
ROC Curve

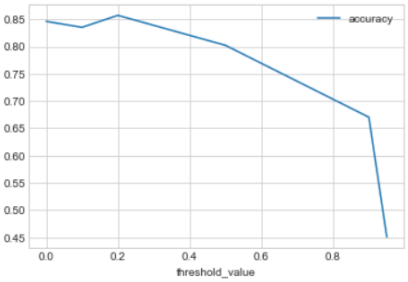


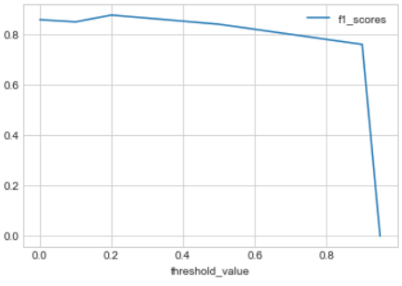
Experiment: 95.0% Unlabelled Data
Accuracy: 45.054945054945056%
F1 Score: 0.0
Classification Report

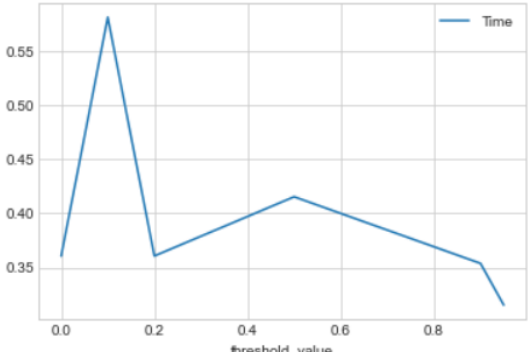
	precision	recall	f1-score	support
0	0.45	1.00	0.62	41
1	0.00	0.00	0.00	50
accuracy			0.45	91
macro avg	0.23	0.50	0.31	91
weighted avg	0.20	0.45	0.28	91

ROC Curve







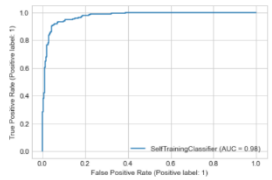


Marketing Dataset

Experiment: 0.0% Unlabelled Data
Accuracy: 92.85714285714286%
F1 Score: 0.9266055045871558
Classification Report

	precision	recall	f1-score	support
0	0.91	0.95	0.93	337
1	0.95	0.90	0.93	335
accuracy			0.93	672
macro avg	0.93	0.93	0.93	672
weighted avg	0.93	0.93	0.93	672

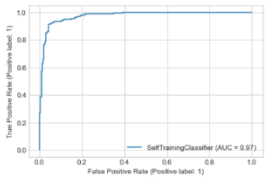
ROC Curve



Experiment: 10.0% Unlabelled Data
Accuracy: 92.70833333333334%
F1 Score: 0.9247311827956989
Classification Report

	precision	recall	f1-score	support
0	0.90	0.96	0.93	337
1	0.95	0.90	0.92	335
accuracy			0.93	672
macro avg	0.93	0.93	0.93	672
weighted avg	0.93	0.93	0.93	672

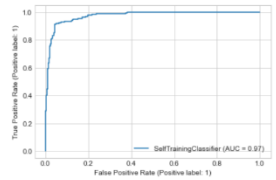
ROC Curve



Experiment: 20.0% Unlabelled Data
Accuracy: 92.70833333333334%
F1 Score: 0.9247311827956989
Classification Report

	precision	recall	f1-score	support
0	0.90	0.96	0.93	337
1	0.95	0.90	0.92	335
accuracy			0.93	672
macro avg	0.93	0.93	0.93	672
weighted avg	0.93	0.93	0.93	672

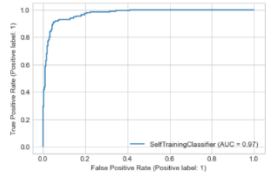
ROC Curve



Experiment: 50.0% Unlabelled Data
Accuracy: 92.11309523809523%
F1 Score: 0.9180834621329211
Classification Report

	precision	recall	f1-score	support
0	0.89	0.96	0.92	337
1	0.95	0.89	0.92	335
accuracy			0.92	672
macro avg	0.92	0.92	0.92	672
weighted avg	0.92	0.92	0.92	672

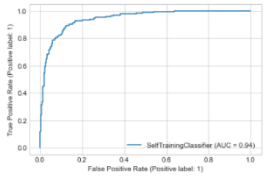
ROC Curve



Experiment: 90.0% Unlabelled Data
Accuracy: 88.39285714285714%
F1 Score: 0.8842729970326641
Classification Report

	precision	recall	f1-score	support
0	0.89	0.88	0.88	337
1	0.88	0.89	0.88	335
accuracy			0.88	672
macro avg	0.88	0.88	0.88	672
weighted avg	0.88	0.88	0.88	672

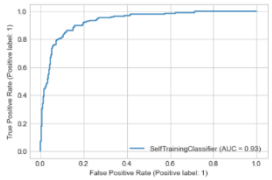
ROC Curve

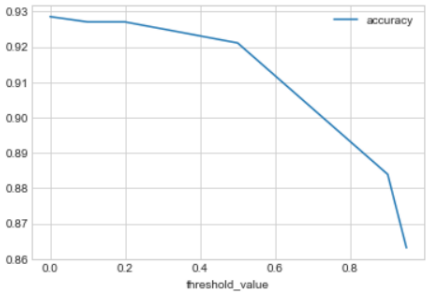


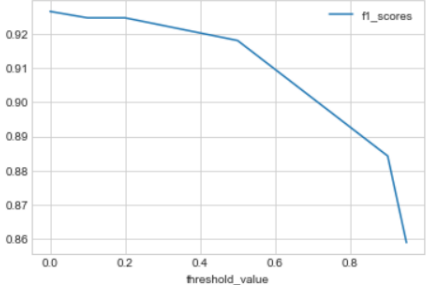
Experiment: 95.0% Unlabelled Data
Accuracy: 86.30952380952381%
F1 Score: 0.8588957055214723
Classification Report

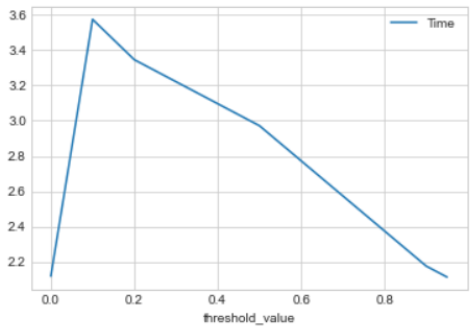
	precision	recall	f1-score	support
0	0.85	0.89	0.87	337
1	0.88	0.84	0.86	335
accuracy			0.86	672
macro avg	0.86	0.86	0.86	672
weighted avg	0.86	0.86	0.86	672

ROC Curve







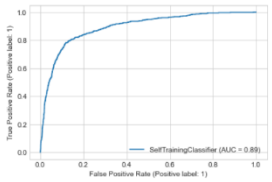


Online Shopper Intention – imbalanced

Experiment: 0.0% Unlabelled Data
Accuracy: 88.26709921600433%
F1 Score: 0.5241228070175439
Classification Report

	precision	recall	f1-score	support
0	0.90	0.97	0.93	3114
1	0.73	0.41	0.52	585
accuracy			0.88	3699
macro avg	0.81	0.69	0.73	3699
weighted avg	0.87	0.88	0.87	3699

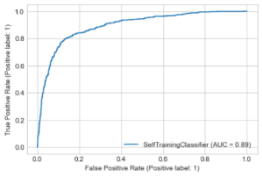
ROC Curve



Experiment: 10.0% Unlabelled Data
Accuracy: 88.24006488240065%
F1 Score: 0.5235487404162104
Classification Report

	precision	recall	f1-score	support
0	0.90	0.97	0.93	3114
1	0.73	0.41	0.52	585
accuracy			0.88	3699
macro avg	0.81	0.69	0.73	3699
weighted avg	0.87	0.88	0.87	3699

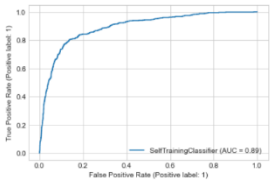
ROC Curve



Experiment: 20.0% Unlabelled Data
Accuracy: 88.29413354960 8%
F1 Score: 0.5183537263626251
Classification Report

	precision	recall	f1-score	support
0	0.90	0.97	0.93	3114
1	0.74	0.40	0.52	585
accuracy			0.88	3699
macro avg	0.82	0.69	0.73	3699
weighted avg	0.87	0.88	0.87	3699

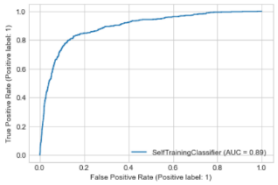
ROC Curve



Experiment: 50.0% Unlabelled Data
Accuracy: 88.15896188158962%
F1 Score: 0.5144124168514412
Classification Report

	precision	recall	f1-score	support
0	0.90	0.97	0.93	3114
1	0.73	0.40	0.51	585
accuracy			0.88	3699
macro avg	0.81	0.68	0.72	3699
weighted avg	0.87	0.88	0.87	3699

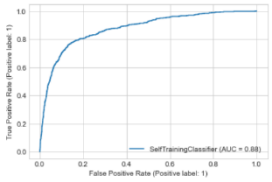
ROC Curve



Experiment: 90.0% Unlabelled Data
Accuracy: 87.26682887266828%
F1 Score: 0.4290909090909091
Classification Report

	precision	recall	f1-score	support
0	0.88	0.98	0.93	3114
1	0.74	0.30	0.43	585
accuracy			0.87	3699
macro avg	0.81	0.64	0.68	3699
weighted avg	0.86	0.87	0.85	3699

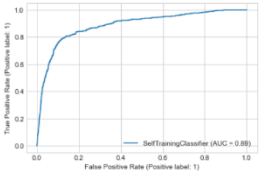
ROC Curve

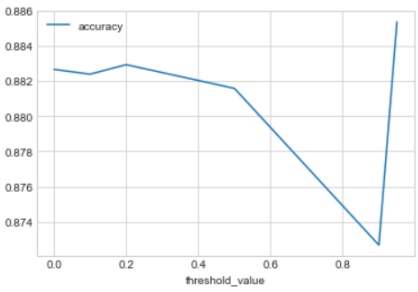


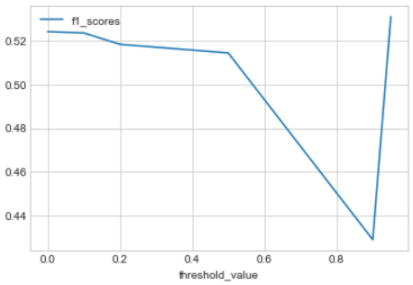
Experiment: 95.0% Unlabelled Data
Accuracy: 88.53744255204109%
F1 Score: 0.5309734513274337
Classification Report

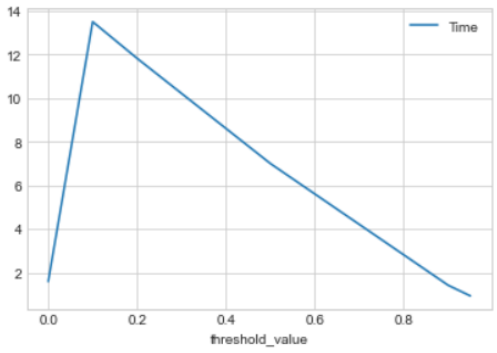
	precision	recall	f1-score	support
0	0.90	0.97	0.93	3114
1	0.75	0.41	0.53	585
accuracy			0.89	3699
macro avg	0.83	0.69	0.73	3699
weighted avg	0.87	0.89	0.87	3699

ROC Curve







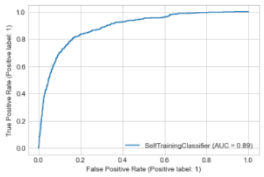


Online shopper intention – balanced up sampled

Experiment: 0.0% Unlabelled Data
Accuracy: 85.56366585563666%
F1 Score: 0.6163793103448275
Classification Report

	precision	recall	f1-score	support
0	0.95	0.88	0.91	3114
1	0.53	0.73	0.62	585
accuracy			0.86	3699
macro avg	0.74	0.81	0.76	3699
weighted avg	0.88	0.86	0.86	3699

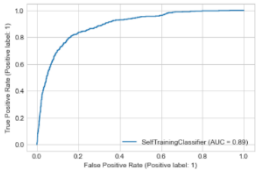
ROC Curve



Experiment: 10.0% Unlabelled Data
Accuracy: 85.64476885644768%
F1 Score: 0.6171593366979091
Classification Report

	precision	recall	f1-score	support
0	0.95	0.88	0.91	3114
1	0.53	0.73	0.62	585
accuracy			0.86	3699
macro avg	0.74	0.81	0.76	3699
weighted avg	0.88	0.86	0.87	3699

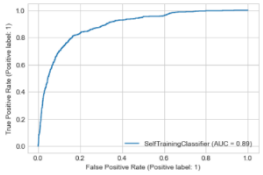
ROC Curve



Experiment: 20.0% Unlabelled Data
Accuracy: 85.77994052446607%
F1 Score: 0.6193921852387844
Classification Report

	precision	recall	f1-score	support
0	0.95	0.88	0.91	3114
1	0.54	0.73	0.62	585
accuracy			0.86	3699
macro avg	0.74	0.81	0.77	3699
weighted avg	0.88	0.86	0.87	3699

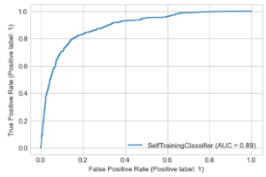
ROC Curve



Experiment: 50.0% Unlabelled Data
Accuracy: 86.45579886455799%
F1 Score: 0.6275092936802973
Classification Report

	precision	recall	f1-score	support
0	0.94	0.89	0.92	3114
1	0.56	0.72	0.63	585
accuracy			0.86	3699
macro avg	0.75	0.81	0.77	3699
weighted avg	0.88	0.86	0.87	3699

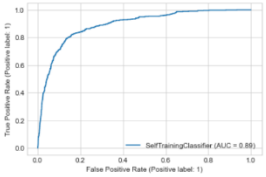
ROC Curve

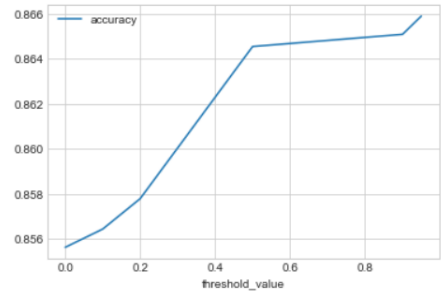


Experiment: 90.0% Unlabelled Data
Accuracy: 86.50986753176534%
F1 Score: 0.6262172284644195
Classification Report

	precision	recall	f1-score	support
0	0.94	0.89	0.92	3114
1	0.56	0.71	0.63	585
accuracy			0.87	3699
macro avg	0.75	0.80	0.77	3699
weighted avg	0.88	0.87	0.87	3699

ROC Curve

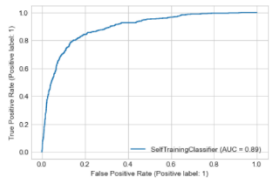


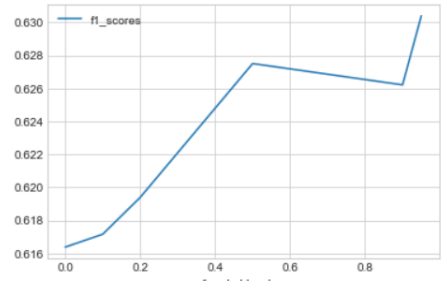


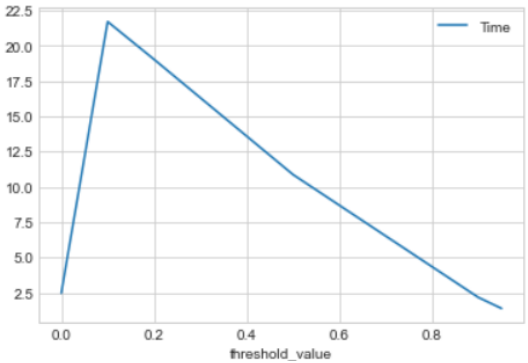
Experiment: 95.0% Unlabelled Data
Accuracy: 86.59097053257638%
F1 Score: 0.6304023845007451
Classification Report

	precision	recall	f1-score	support
0	0.94	0.89	0.92	3114
1	0.56	0.72	0.63	585
accuracy			0.87	3699
macro avg	0.75	0.81	0.77	3699
weighted avg	0.88	0.87	0.87	3699

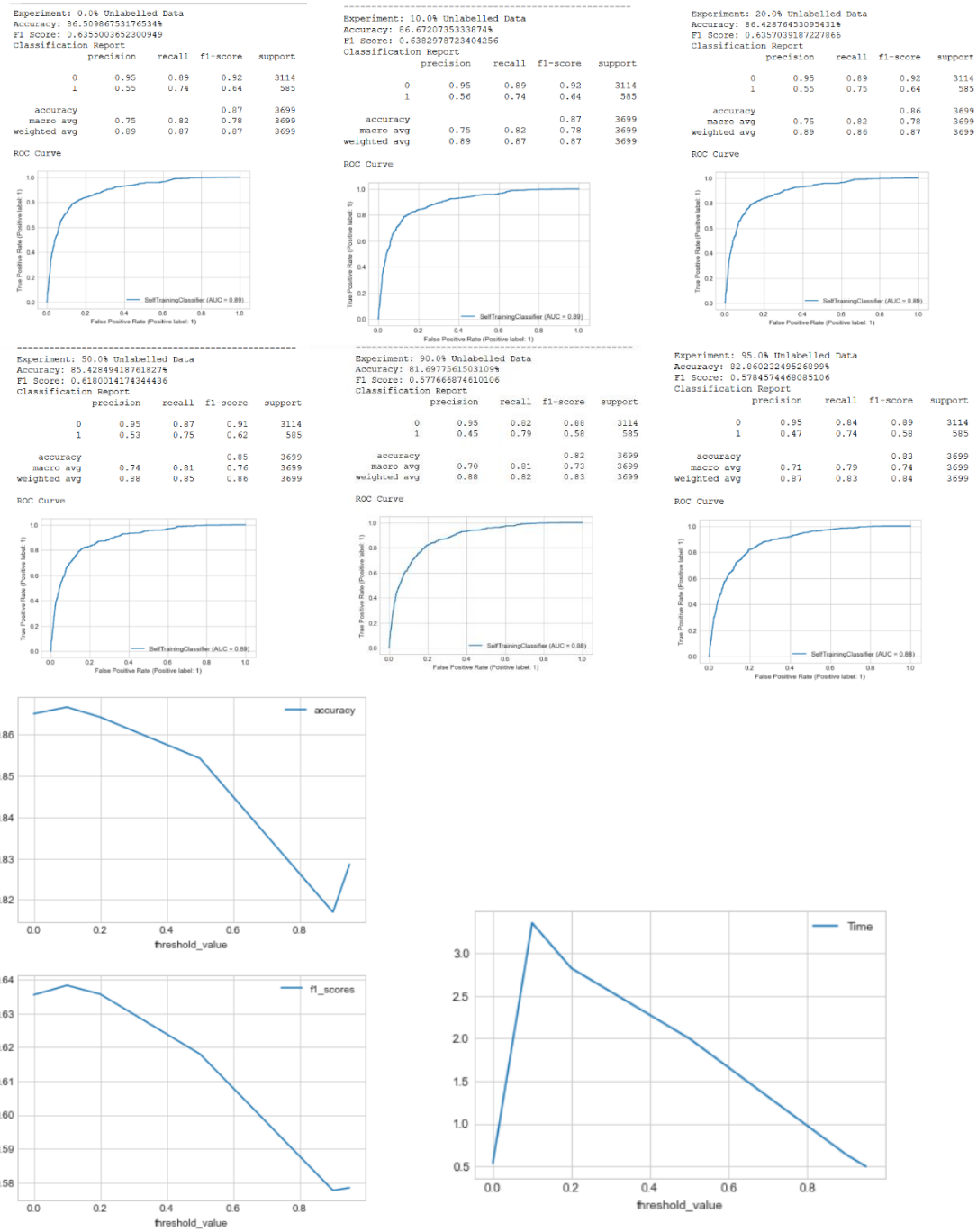
ROC Curve







Online shopper intention – balanced down sampled



Interplay between class imbalance and SSL.

When the class is balanced using up sampling/down sampling it is observed that there is a gain in F1 score overall all the experiments (with different amounts of unlabeled data) and since the classes are imbalanced the more precise measure of the performance of the model will be f1 score and not the accuracy.

Best SSL algorithm

The overall best performance is obtained using the Self Training approach, considering both the accuracy and F1 score. although we are including our ‘pseudo-labeled’ data with labeled training data, some of the ‘pseudo-labeled’ data is certainly going to be incorrect. When enough of the ‘pseudo-labels’ are incorrect, the self-training algorithm can reinforce poor classification decisions, and classifier performance can actually get worse.