

Binary Text Classification

ML Classifiers Result on Tf-Idf Vectorizer:

Logistic Regression Classifier Results:

Accuracy: 73.75%

Confusion Matrix:

[[371 144]

[114 354]]

Classification Report:

	precision	recall	f1-score	support
NOT	0.76	0.72	0.74	515
HOF	0.71	0.76	0.73	468
accuracy			0.74	983
macro avg	0.74	0.74	0.74	983
weighted avg	0.74	0.74	0.74	983

SVM Classifier Results:

Accuracy: 73.96%

Confusion Matrix:

[[361 154]

[102 366]]

Classification Report:

	precision	recall	f1-score	support
NOT	0.78	0.70	0.74	515
HOF	0.70	0.78	0.74	468
accuracy			0.74	983
macro avg	0.74	0.74	0.74	983
weighted avg	0.74	0.74	0.74	983

Naive Bayes Classifier Results:

Accuracy: 68.16%

Confusion Matrix:

[[296 219]

[94 374]]

Classification Report:

	precision	recall	f1-score	support
NOT	0.76	0.57	0.65	515
HOF	0.63	0.80	0.70	468
accuracy			0.74	983
macro avg	0.69	0.69	0.74	983
weighted avg	0.70	0.68	0.74	983

K-Nearest Neighbors Classifier Results:

Accuracy: 67.75%

Confusion Matrix:

[[359 156]

[161 307]]

Classification Report:

	precision	recall	f1-score	support
NOT	0.69	0.70	0.69	515
HOF	0.66	0.66	0.66	468
accuracy			0.68	983
macro avg	0.68	0.68	0.68	983
weighted avg	0.68	0.68	0.68	983

Decision Tree Classifier Results:

Accuracy: 70.19%

Confusion Matrix:

[[375 140]

[153 315]]

Classification Report:

	precision	recall	f1-score	support
NOT	0.71	0.73	0.72	515
HOF	0.69	0.67	0.68	468
accuracy			0.70	983
macro avg	0.70	0.70	0.70	983
weighted avg	0.70	0.70	0.70	983

Gradient Boosting Classifier Results:

Accuracy: 69.68%

Confusion Matrix:

[[330 185]

[113 355]]

Classification Report:

	precision	recall	f1-score	support
NOT	0.74	0.64	0.69	515
HOF	0.66	0.76	0.70	468
accuracy			0.70	983
macro avg	0.70	0.70	0.70	983
weighted avg	0.70	0.70	0.70	983

Random Forest Classifier Results:

Accuracy: 72.74%

Confusion Matrix:

[[365 150]

[118 350]]

Classification Report:

	precision	recall	f1-score	support
NOT	0.76	0.71	0.73	515
HOF	0.70	0.75	0.72	468
accuracy			0.73	983
macro avg	0.73	0.73	0.73	983
weighted avg	0.73	0.73	0.73	983

ML Classifiers Result on Custom GloVe Embeddings (50 Dimensions):

Evaluating Logistic Regression...

Accuracy: 66.73%

Classification Report:

	precision	recall	f1-score	support
NOT	0.64	0.69	0.66	468
HOF	0.70	0.65	0.67	515
accuracy			0.67	983
macro avg	0.67	0.67	0.67	983
weighted avg	0.67	0.67	0.67	983

Confusion Matrix:

[[322 146]

[181 334]]

Evaluating Support Vector Classifier...

Accuracy: 70.30%

Classification Report:

	precision	recall	f1-score	support
NOT	0.67	0.75	0.71	468
HOF	0.75	0.65	0.70	515
accuracy			0.70	983
macro avg	0.71	0.71	0.70	983
weighted avg	0.71	0.70	0.70	983

Confusion Matrix:

[[354 114]

[178 337]]

Evaluating Gaussian Naive Bayes...

Accuracy: 66.53%

Classification Report:

	precision	recall	f1-score	support
NOT	0.63	0.71	0.67	468
HOF	0.70	0.63	0.66	515
accuracy			0.67	983
macro avg	0.67	0.67	0.67	983
weighted avg	0.67	0.67	0.67	983

Confusion Matrix:

[[330 138]

[191 324]]

Evaluating K-Nearest Neighbors...

Accuracy: 69.58%

Classification Report:

	precision	recall	f1-score	support
NOT	0.68	0.68	0.68	468
HOF	0.71	0.71	0.71	515
accuracy				0.70 983
macro avg	0.70	0.70	0.70	983
weighted avg	0.70	0.70	0.70	983

Confusion Matrix:

[[318 150]

[149 366]]

Evaluating Decision Tree Classifier...

Accuracy: 65.51%

Classification Report:

	precision	recall	f1-score	support
NOT	0.63	0.65	0.64	468
HOF	0.68	0.66	0.67	515
accuracy				0.66 983
macro avg	0.65	0.65	0.65	983
weighted avg	0.65	0.65	0.66	983

Confusion Matrix:

[[305 163]

[176 339]]

Evaluating Random Forest Classifier...

Accuracy: 72.43%

Classification Report:

	precision	recall	f1-score	support
NOT	0.69	0.76	0.72	468
HOF	0.76	0.69	0.72	515
accuracy				0.72 983
macro avg	0.73	0.73	0.72	983
weighted avg	0.73	0.73	0.72	983

Confusion Matrix:

[[357 111]

[160 355]]

Evaluating Gradient Boosting Classifier...

Accuracy: 70.80%

Classification Report:

	precision	recall	f1-score	support
NOT	0.68	0.72	0.70	468
HOF	0.73	0.70	0.71	515
accuracy				0.71 983
macro avg	0.71	0.71	0.71	983
weighted avg	0.71	0.71	0.71	983

Confusion Matrix:

[[337 131]

[156 359]]

Deep Learning Algorithms Results on Custom GloVe Embeddings(50 Dimensions):

Conv1D Accuracy: 66.43%

Conv1D Classification Report:

	precision	recall	f1-score	support
NOT	0.64	0.69	0.66	468
HOF	0.69	0.64	0.67	515
accuracy			0.66	983
macro avg	0.67	0.67	0.66	983
weighted avg	0.67	0.66	0.66	983

Conv1D Confusion Matrix:

[[321 147]

[183 332]]

LSTM Accuracy: 67.14%

LSTM Classification Report:

	precision	recall	f1-score	support
NOT	0.65	0.66	0.66	468
HOF	0.69	0.68	0.68	515
accuracy			0.67	983
macro avg	0.67	0.67	0.67	983
weighted avg	0.67	0.67	0.67	983

LSTM Confusion Matrix:

[[309 159]

[164 351]]

Multiclass Text Classification:

ML Classifiers Result on Custom GloVe Embeddings (50 Dimensions):

Logistic Regression Classifier Results:

Accuracy: 55.75%

Classification Report:

	precision	recall	f1-score	support
NONE	0.57	0.82	0.67	468
CHOF	0.64	0.43	0.51	189
SHOF	0.47	0.26	0.33	326
accuracy			0.56	983
macro avg	0.56	0.50	0.50	983
weighted avg	0.55	0.56	0.53	983

SVM Classifier Results:

Accuracy: 60.63%

Confusion Matrix:

[[405 12 51]

[63 90 36]

[207 18 101]]

Classification Report:

	precision	recall	f1-score	support
NONE	0.60	0.87	0.71	468
CHOF	0.75	0.48	0.58	189
SHOF	0.54	0.31	0.39	326
accuracy			0.61	983
macro avg	0.63	0.55	0.56	983
weighted avg	0.61	0.61	0.58	983

Naive Bayes Classifier Results:

Accuracy: 55.95%

Confusion Matrix:

[[271 32 165]

[31 90 68]

[100 37 189]]

Classification Report:

	precision	recall	f1-score	support
NONE	0.67	0.58	0.62	468
CHOF	0.57	0.48	0.52	189
SHOF	0.45	0.58	0.51	326
accuracy			0.56	983
macro avg	0.56	0.55	0.55	983
weighted avg	0.58	0.56	0.56	983

K-Nearest Neighbors Classifier Results:

Accuracy: 63.17%

Confusion Matrix:

[[377 28 63]

[39 120 30]

[163 39 124]]

Classification Report:

	precision	recall	f1-score	support
NONE	0.65	0.81	0.72	468
CHOF	0.64	0.63	0.64	189
SHOF	0.57	0.38	0.46	326
accuracy			0.63	983
macro avg	0.62	0.61	0.61	983
weighted avg	0.62	0.63	0.62	983

Decision Tree Classifier Results:

Accuracy: 52.09%

Confusion Matrix:

[[285 46 137]

[39 90 60]

[137 52 137]]

Classification Report:

	precision	recall	f1-score	support
NONE	0.62	0.61	0.61	468
CHOF	0.48	0.48	0.48	189
SHOF	0.41	0.42	0.42	326
accuracy			0.52	983
macro avg	0.50	0.50	0.50	983
weighted avg	0.52	0.52	0.52	983

Gradient Boosting Classifier Results:

Accuracy: 62.05%

Confusion Matrix:

[[380 13 75]

[51 93 45]

[166 23 137]]

Classification Report:

	precision	recall	f1-score	support
NONE	0.64	0.81	0.71	468
CHOF	0.72	0.49	0.58	189
SHOF	0.53	0.42	0.47	326
accuracy			0.62	983
macro avg	0.63	0.57	0.59	983
weighted avg	0.62	0.62	0.61	983

Random Forest Classifier Results:

Accuracy: 64.09%

Confusion Matrix:

```
[[372 14 82]
 [ 37 109 43]
 [148 29 149]]
```

Classification Report:

	precision	recall	f1-score	support
NONE	0.67	0.79	0.73	468
CHOF	0.72	0.58	0.64	189
SHOF	0.54	0.46	0.50	326
accuracy			0.64	983
macro avg	0.64	0.61	0.62	983
weighted avg	0.64	0.64	0.63	983

Deep Learning Algorithms Results on Custom GloVe Embeddings (50 Dimensions):

Conv1D Accuracy: 56.77%

Conv1D Classification Report:

	precision	recall	f1-score	support
NONE	0.47	0.49	0.48	189
CHOF	0.64	0.73	0.68	468
SHOF	0.49	0.38	0.43	326
accuracy			0.57	983
macro avg	0.53	0.53	0.53	983
weighted avg	0.56	0.57	0.56	983

Conv1D Confusion Matrix:

```
[[ 93 49 47]
 [ 45 340 83]
 [ 58 143 125]]
```

LSTM Accuracy: 63.99%

LSTM Classification Report:

	precision	recall	f1-score	support
NONE	0.85	0.46	0.60	189
CHOF	0.67	0.79	0.72	468
SHOF	0.52	0.53	0.53	326
accuracy			0.64	983
macro avg	0.68	0.59	0.62	983
weighted avg	0.66	0.64	0.53	983

Conv1D Confusion Matrix:

LSTM Confusion Matrix:

```
[[ 87 39 63]
```

```
 [ 5 368 95]
```

```
 [10 142 174]]
```