"C:\Program Files (x86)\Python37-32\python.exe" D:/cmpt400-project/Fake\_News\_Model\_Zoo/code/NLTK/shallow\_models.py

Training set size: (11524, 2)

------------------------------------------------------------------------

logistic regression

C:\Program Files (x86)\Python37-32\lib\site-packages\sklearn\model\_selection\\_split.py:296: FutureWarning: Setting a random\_state has no effect since shuffle is False. This will raise an error in 0.24. You should leave random\_state to its default (None), or set shuffle=True.

FutureWarning

Classification Report :

Model Name: Logistic Regression-CV-TFIDF

Accuracy: 0.2478295185477506

Precision score: 0.2616014219791488

Recall score: 0.2478295185477506

F1 score: 0.2372952002615486

precision recall f1-score support

barely-true 0.27 0.18 0.22 212

false 0.27 0.33 0.30 249

half-true 0.24 0.32 0.27 265

mostly-true 0.21 0.28 0.24 241

pants-fire 0.40 0.02 0.04 92

true 0.26 0.18 0.21 208

accuracy 0.25 1267

macro avg 0.28 0.22 0.22 1267

weighted avg 0.26 0.25 0.24 1267

StratifiedKFold score for LR: [0.2516269 0.25466377 0.24815618 0.2494577 0.25607639]

StratifiedKFold mean score for LR: 0.2519961888406845

------------------------------------------------------------------------

Naive Bayes

Classification Report :

C:\Program Files (x86)\Python37-32\lib\site-packages\sklearn\metrics\\_classification.py:1272: UndefinedMetricWarning: Precision is ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

C:\Program Files (x86)\Python37-32\lib\site-packages\sklearn\metrics\\_classification.py:1272: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

Model Name: MultinomialNB-CV-TFIDF

Accuracy: 0.23362273086029992

Precision score: 0.2212539648529418

Recall score: 0.23362273086029992

F1 score: 0.18791818312806927

precision recall f1-score support

barely-true 0.26 0.05 0.09 212

false 0.26 0.29 0.27 249

half-true 0.23 0.54 0.33 265

mostly-true 0.21 0.27 0.24 241

pants-fire 0.00 0.00 0.00 92

true 0.24 0.02 0.04 208

accuracy 0.23 1267

macro avg 0.20 0.20 0.16 1267

weighted avg 0.22 0.23 0.19 1267

StratifiedKFold score for nb\_muti\_cv: [0.23947939 0.23167028 0.23123644 0.23340564 0.23871528]

StratifiedKFold mean score for nb\_muti\_cv: 0.23490140696553388

Classification Report :

Model Name: ComplementNB-CV-TFIDF

Accuracy: 0.2517758484609313

Precision score: 0.2533143111514316

Recall score: 0.2517758484609313

F1 score: 0.2461437947832765

precision recall f1-score support

barely-true 0.26 0.20 0.23 212

false 0.28 0.31 0.29 249

half-true 0.26 0.33 0.29 265

mostly-true 0.22 0.29 0.25 241

pants-fire 0.28 0.10 0.15 92

true 0.22 0.17 0.19 208

accuracy 0.25 1267

macro avg 0.26 0.23 0.23 1267

weighted avg 0.25 0.25 0.25 1267

StratifiedKFold score for NB\_comple: [0.24815618 0.24121475 0.24598698 0.24208243 0.25477431]

StratifiedKFold mean score for NB\_comple: 0.24644293052542782

------------------------------------------------------------------------

Support Vector Machine

Classification Report :

Model Name: SVM-CV-TFIDF

Accuracy: 0.2525651144435675

Precision score: 0.2512064153374705

Recall score: 0.2525651144435675

F1 score: 0.25015485176986796

precision recall f1-score support

barely-true 0.25 0.19 0.22 212

false 0.30 0.32 0.31 249

half-true 0.26 0.29 0.27 265

mostly-true 0.23 0.27 0.25 241

pants-fire 0.20 0.12 0.15 92

true 0.22 0.24 0.23 208

accuracy 0.25 1267

macro avg 0.25 0.24 0.24 1267

weighted avg 0.25 0.25 0.25 1267

StratifiedKFold score for SVM: [0.23644252 0.24295011 0.24685466 0.23383948 0.24956597]

StratifiedKFold mean score for SVM: 0.24193054802362016

------------------------------------------------------------------------

Decision Tree

Classification Report :

Model Name: DT-CV-TFIDF

Accuracy: 0.2067876874506709

Precision score: 0.2029740298298947

Recall score: 0.2067876874506709

F1 score: 0.20393717479864817

precision recall f1-score support

barely-true 0.19 0.16 0.17 212

false 0.23 0.28 0.25 249

half-true 0.19 0.16 0.17 265

mostly-true 0.21 0.22 0.21 241

pants-fire 0.11 0.09 0.10 92

true 0.24 0.26 0.25 208

accuracy 0.21 1267

macro avg 0.19 0.20 0.19 1267

weighted avg 0.20 0.21 0.20 1267

StratifiedKFold score for DT: [0.1978308 0.20390456 0.21735358 0.2 0.18706597]

StratifiedKFold mean score for DT: 0.2012309818630995

------------------------------------------------------------------------

Random Forest

Classification Report :

C:\Program Files (x86)\Python37-32\lib\site-packages\sklearn\metrics\\_classification.py:1272: UndefinedMetricWarning: Precision is ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

C:\Program Files (x86)\Python37-32\lib\site-packages\sklearn\metrics\\_classification.py:1272: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

\_warn\_prf(average, modifier, msg\_start, len(result))

Model Name: Random Forest

Accuracy: 0.2107340173638516

Precision score: 0.1279755067963741

Recall score: 0.2107340173638516

F1 score: 0.13413183827443034

precision recall f1-score support

barely-true 0.00 0.00 0.00 212

false 0.22 0.18 0.20 249

half-true 0.21 0.74 0.32 265

mostly-true 0.22 0.10 0.14 241

pants-fire 0.00 0.00 0.00 92

true 0.00 0.00 0.00 208

accuracy 0.21 1267

macro avg 0.11 0.17 0.11 1267

weighted avg 0.13 0.21 0.13 1267

StratifiedKFold score for RF: [0.21778742 0.22733189 0.22255965 0.21691974 0.22178819]

StratifiedKFold mean score for RF: 0.22127737858520127

Process finished with exit code 0