MultinomialNB

Accuracy: 0.4373985239852399

:Classification Report
C:\Users\97252\anaconda3\lib\site-packages\sklearn\metrics_classification.py:1471: 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_
C:\Users\97252\anaconda3\lib\site-packages\sklearn\metrics_classification.py:1471: 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_

	precision	recall	f1-score	support
action	0.00	0.00	0.00	1314
adult	0.00	0.00	0.00	590
adventure	0.00	0.00	0.00	775
animation	0.00	0.00	0.00	498
biography	0.00	0.00	0.00	264
comedy	0.73	0.03	0.05	7446
crime	0.00	0.00	0.00	505
documentary	0.52	0.89	0.66	13096
drama	0.38	0.87	0.52	13612
family	0.00	0.00	0.00	783
fantasy	0.00	0.00	0.00	322
game-show	0.00	0.00	0.00	193
history	0.00	0.00	0.00	243
horror	0.00	0.00	0.00	2204
music	0.00	0.00	0.00	731
musical	0.00	0.00	0.00	276
mystery	0.00	0.00	0.00	318
news	0.00	0.00	0.00	181
reality-tv	0.00	0.00	0.00	883
romance	0.00	0.00	0.00	672
sci-fi	0.00	0.00	0.00	646
short	1.00	0.00	0.00	5072
sport	0.00	0.00	0.00	431
talk-show	0.00	0.00	0.00	391
thriller	0.00	0.00	0.00	1590
war	0.00	0.00	0.00	132
western	0.00	0.00	0.00	1032
accuracy			0.44	54200
macro avg	0.10	0.07	0.05	54200
weighted avg	0.41	0.44	0.30	54200

C:\Users\97252\anaconda3\lib\site-packages\sklearn\metrics_classification.py:1471: 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_

DecisionTreeClassifier

Accuracy (Decision Tree): 0.35297047970479706

:(Classification Report (Decision Tree

	precision	recall	f1-score	support
action	0.09	0.07	0.08	1314
adult	0.16	0.11	0.13	590
adventure	0.11	0.08	0.10	775
animation	0.04	0.03	0.03	498
biography	0.01	0.01	0.01	264
comedy	0.31	0.32	0.31	7446
crime	0.04	0.03	0.04	505
documentary	0.54	0.59	0.56	13096
drama	0.40	0.46	0.42	13612
family	0.07	0.05	0.06	783
fantasy	0.00	0.00	0.00	322
game-show	0.39	0.38	0.38	193
history	0.04	0.02	0.03	243
horror	0.23	0.20	0.22	2204
music	0.29	0.26	0.28	731
musical	0.05	0.03	0.04	276
mystery	0.02	0.02	0.02	318
news	0.04	0.02	0.03	181
reality-tv	0.08	0.06	0.07	883
romance	0.06	0.04	0.05	672
sci-fi	0.11	0.07	0.09	646
short	0.21	0.20	0.20	5072
sport	0.26	0.18	0.21	431
talk-show	0.20	0.15	0.17	391
thriller	0.09	0.07	0.08	1590
war	0.04	0.03	0.04	132
western	0.47	0.43	0.45	1032
accuracy			0.35	54200
macro avg	0.16	0.15	0.15	54200
weighted avg	0.33	0.35	0.34	54200

כאשר ניסנו Word 2 vec לא עבד כי זה ערך שלילי

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)

Cell In[37], line 2
      1 nb_classifier_word2vec = MultinomialNB()
----> 2 nb_classifier_word2vec.fit(word2vec_train, train_data['Genre'])
      4 predictions_word2vec = nb_classifier_word2vec.predict(word2vec_test)
      6 accuracy_word2vec = accuracy_score(test_sol_data['Genre'], predictions_word2vec)

File ~\PycharmProjects\Movie Genre Classification\Script\lib\site-packages\sklearn\base.py:1474, in _fit_context.<locals>.decorator.<locals>
   1467     estimator._validate_params()
   1469 with config_context(
   1470     skip_parameter_validation=(
   1471         prefer_skip_nested_validation or global_skip_validation
   1472     )
   1473 ):
-> 1474     return fit_method(estimator, *args, **kwargs)

File ~\PycharmProjects\Movie Genre Classification\Script\lib\site-packages\sklearn\naive_bayes.py:759, in _BaseDiscreteNB.fit(self, X, y, sa
    757 n_classes = Y.shape[1]
    758 self._init_counters(n_classes, n_features)
--> 759 self._count(X, Y)
    760 alpha = self._check_alpha()
    761 self._update_feature_log_prob(alpha)

File ~\PycharmProjects\Movie Genre Classification\Script\lib\site-packages\sklearn\naive_bayes.py:881, in MultinomialNB._count(self, X, Y)
    879 def _count(self, X, Y):
    880     """Count and smooth feature occurrences."""
--> 881     check_non_negative(X, "MultinomialNB (input X)")
    882     self.feature_count_ += safe_sparse_dot(Y.T, X)
    883     self.class_count_ += Y.sum(axis=0)

File ~\PycharmProjects\Movie Genre Classification\Script\lib\site-packages\sklearn\utils\validation.py:1650, in check_non_negative(X, whom)
   1647     X_min = xp.min(X)
   1649 if X_min < 0:
-> 1650     raise ValueError("Negative values in data passed to %s" % whom)

ValueError: Negative values in data passed to MultinomialNB (input X)
```