I081 Aniruddh Kulkarni NLP Exp6

May 28, 2023

1 Name: Aniruddh Kulkarni

2 Roll no: I081

3 Stream: CS (AI)

4 Division: I

5 Semester: 5th Semester

6 Batch: I-3

7 Subject: NLP

8 Assignment-6

[1]:
```
import warnings
warnings.filterwarnings('ignore')

# Generate and plot a synthetic imbalanced classification dataset
from collections import Counter
import numpy as np
import pandas as pd # to work with csv files
from scipy import sparse

# matplotlib imports are used to plot confusion matrices for the classifiers
import matplotlib as mpl
import matplotlib.cm as cm
import matplotlib.pyplot as plt

# import feature extraction methods from sklearn
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction import _stop_words
from sklearn.feature_extraction.text import TfidfVectorizer

# pre-processing of text
```

```
import string
import re

# import classifiers from sklearn
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB
from sklearn.linear_model import LogisticRegression
from sklearn.svm import LinearSVC

# import different metrics to evaluate the classifiers
from sklearn.metrics import accuracy_score
from sklearn.metrics import roc_auc_score

# from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix
from sklearn import metrics

# import time function from time module to track the training duration
from time import time

# importing required ml model libraries
from sklearn.metrics import precision_recall_fscore_support as score
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC, LinearSVC
from sklearn.ensemble import RandomForestClassifier
from sklearn.naive_bayes import MultinomialNB
from keras.layers import Dense, Input, Flatten
from keras.layers import GlobalAveragePooling1D, Embedding
from keras.models import Sequential
from keras.preprocessing.text import Tokenizer

from sklearn import metrics

!pip install imbalanced-learn
from imblearn.under_sampling import RandomUnderSampler
from imblearn.over_sampling import RandomOverSampler
from imblearn.over_sampling import SMOTE
```

```
Requirement already satisfied: imbalanced-learn in
/Users/pushpakulkarni/miniconda3/lib/python3.10/site-packages (0.10.1)
Requirement already satisfied: numpy>=1.17.3 in
/Users/pushpakulkarni/miniconda3/lib/python3.10/site-packages (from imbalanced-
learn) (1.22.3)
Requirement already satisfied: joblib>=1.1.1 in
/Users/pushpakulkarni/miniconda3/lib/python3.10/site-packages (from imbalanced-
learn) (1.2.0)
```

```
Requirement already satisfied: scikit-learn>=1.0.2 in
/Users/pushpakulkarni/miniconda3/lib/python3.10/site-packages (from imbalanced-
learn) (1.2.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/Users/pushpakulkarni/miniconda3/lib/python3.10/site-packages (from imbalanced-
learn) (3.1.0)
Requirement already satisfied: scipy>=1.3.2 in
/Users/pushpakulkarni/miniconda3/lib/python3.10/site-packages (from imbalanced-
learn) (1.10.1)
```

[2]:
```
our_data = pd.read_csv("Full-Economic-News-DFE-839861.csv" , encoding =␣
↪"ISO-8859-1" )
```

[3]:
```
our_data.head()
```

[3]:

	_unit_id	_golden	_unit_state	_trusted_judgments	_last_judgment_at	positivity	positivity:confidence	relevance	relevance:confidence	articleid	date	headline	positivity_gold	relevance_gold	text
0	842613455	False	finalized	3	12/5/2015 17:48:27	3.0	0.6400	yes	0.640	wsj_398217788	1991-08-14	Yields on CDs Fell in the Latest Week	NaN	NaN	NEW YORK -- Yields on most certificates of dep…
1	842613456	False	finalized	3	12/5/2015 16:54:25	NaN	NaN	no	1.000	wsj_399019502	2007-08-21	The Morning Brief: White House Seeks to Limit …	NaN	NaN	The Wall Street Journal Online</br></br>The Mo…
2	842613457	False	finalized	3	12/5/2015 01:59:03	NaN	NaN	no	1.000	wsj_398284048	1991-11-14	Banking Bill Negotiators Set Compromise --- Pl…	NaN	NaN	WASHINGTON -- In an effort to achieve banking …
3	842613458	False	finalized	3	12/5/2015 02:19:39	NaN	0.0000	no	0.675	wsj_397959018	1986-06-16	Manager's Journal: Sniffing Out Drug Abusers I…	NaN		
4	842613459	False	finalized	3	12/5/2015 17:48:27	3.0	0.3257	yes	0.640	wsj_398838054	2002-10-04	Currency Trading: Dollar Remains in Tight Rang…	NaN		

```
3             NaN  The statistics on the enormous costs of employ…
4             NaN  NEW YORK -- Indecision marked the dollar's ton…
```

[4]:
```
our_data.shape # Number of rows (instances) and columns in the dataset
```

[4]:
```
(8000, 15)
```

[5]:
```
our_data["relevance"].unique()
```

[5]:
```
array(['yes', 'no', 'not sure'], dtype=object)
```

[6]:
```
our_data["relevance"].value_counts()
```

[6]:
```
relevance
no          6571
yes         1420
not sure       9
Name: count, dtype: int64
```

[7]:
```
our_data["relevance"].value_counts()/our_data.shape[0] # Class distribution in␣
 ↪the dataset
```

[7]:
```
relevance
no          0.821375
yes         0.177500
not sure    0.001125
Name: count, dtype: float64
```

[8]:
```
# convert label to a numerical variable
our_data = our_data[our_data.relevance != "not sure"] # removing the data where␣
 ↪we don't want relevance="not sure".
our_data.shape
```

[8]:
```
(7991, 15)
```

[9]:
```
our_data['relevance'] = our_data.relevance.map({'yes':1, 'no':0}) # relevant is␣
 ↪1, not-relevant is 0.
```

[10]:
```
our_data = our_data[["text","relevance"]] # Let us take only the two columns we␣
 ↪need.
our_data
```

[10]:

	text	relevance
0	NEW YORK -- Yields on most certificates of dep…	1
1	The Wall Street Journal Online</br></br>The Mo…	0
2	WASHINGTON -- In an effort to achieve banking …	0
3	The statistics on the enormous costs of employ…	0

```
4     NEW YORK -- Indecision marked the dollar's ton…          1
…                                                    …        …
7995  Secretary of Commerce Charles W. Sawyer said y…          1
7996  U.S. stocks inched up last week, overcoming co…          0
7997  Ben S. Bernanke cleared a key hurdle Thursday …          0
7998  The White House's push to contract out many fe…          0
7999  NEW YORK. April 17-Automobile stocks put on th…          0

[7991 rows x 2 columns]
```

[11]:
```
our_data.shape
```

[11]:
```
(7991, 2)
```

[12]:
```
import re
import nltk
nltk.download('stopwords')
from nltk.corpus import stopwords
stopwords = stopwords.words('english')
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]     /Users/pushpakulkarni/nltk_data…
[nltk_data]   Package stopwords is already up-to-date!
```

[13]:
```
def clean(doc): # doc is a string of text
    doc = doc.replace("</br>", " ") # This text contains a lot of <br/> tags.
    doc = "".join([char for char in doc if char not in string.punctuation and␣
 ↪not char.isdigit()])
    doc = " ".join([token for token in doc.split() if token not in stopwords])
    # remove punctuation and numbers
    return doc
our_data['text'] = our_data['text'].apply(clean)
```

[14]:
```
def special_char(text):
  reviews = ''
  for x in text:
    if x.isalnum():
      reviews = reviews + x
    else:
      reviews = reviews + ' '
  return reviews
our_data['text'] = our_data['text'].apply(special_char)
```

[15]:
```
def convert_lower(text):
   return text.lower()
our_data['text'] = our_data['text'].apply(convert_lower)
our_data['text'][1]
```

[15]:
```
'the wall street journal online the morning brief look days biggest news emailed
subscribers every business day sign email on friday evening congress town summer
recess americans heading midaugust weekend bush administration sent message
states the federal government make tougher national childrens insurance program
cover offspring middleincome families the state childrens health insurance
program created help children whose families couldnt afford insurance didnt
qualify medicaid administration officials tell new york times changes aimed
returning program low income focus assuring didnt become replacement private
insurance administration point man dennis smith wrote state officials saying
would new restrictions district columbia states including california new york
extend plan extend coverage children whose families make federal poverty levels
for family three family four under new limits child family making would spend
one year uninsured qualifying state wants extend coverage would assure
washington least children eligible schip medicaid enrolled one programs but
associated press reports state currently make assurances rachel klein deputy
director health policy advocacy group families usa tells ap since many families
threshold cant afford private insurance effect policy uninsured kids ann
clemency kohler deputy commissioner human services new jersey tells times
changes cause havoc program could jeopardize coverage thousands children states
already imposing waiting periods taking steps prevent parents moving children
private insurance schip currently serves million children washington post notes
the administrations new restrictions come program expires end next month
congress doesnt reauthorize subject larger political fight pits white house
democrats republicans congress state capitals'
```

[16]:
```
x = our_data['text']
y = our_data['relevance']
our_data
```

[16]:

	text	relevance
0	new york yields certificates deposit offered m…	1
1	the wall street journal online the morning bri…	0
2	washington in effort achieve banking reform se…	0
3	the statistics enormous costs employee drug ab…	0
4	new york indecision marked dollars tone trader…	1
…	…	…
7995	secretary commerce charles w sawyer said yeste…	1
7996	us stocks inched last week overcoming concern …	0
7997	ben s bernanke cleared key hurdle thursday con…	0
7998	the white houses push contract many federal fu…	0
7999	new york april automobile stocks put best show…	0

[7991 rows x 2 columns]

[17]:
```
#BoW 1000 feat
x_train1, x_test1, y_train1, y_test1 = train_test_split(x, y, test_size = 0.3,␣
 ↪random_state = 0, shuffle = True)
```

```
print(len(x_train1))
print(len(x_test1))

cv = CountVectorizer(max_features = 1000)
x_train1 = cv.fit_transform(x_train1)

x_test1 = cv.transform(x_test1)
print(x_train1.shape, x_test1.shape)
```

```
5593
2398
(5593, 1000) (2398, 1000)
```

[18]:
```
#BOW 5000 max feat
x_train2, x_test2, y_train2, y_test2 = train_test_split(x, y, test_size = 0.3,␣
 ↪random_state = 0, shuffle = True)
print(len(x_train2))
print(len(x_test2))

cv2 = CountVectorizer(max_features = 5000)
x_train2 = cv2.fit_transform(x_train2)
x_test2 = cv2.transform(x_test2)
print(x_train2.shape, x_test2.shape)
```

```
5593
2398
(5593, 5000) (2398, 5000)
```

[19]:
```
#Bag of n gram 1000 feat [bi,tri grams]
x_train3, x_test3, y_train3, y_test3 = train_test_split(x, y, test_size = 0.3,␣
 ↪random_state = 0, shuffle = True)
print(len(x_train3))
print(len(x_test3))
count_vect = CountVectorizer(ngram_range=(2,3),max_features = 1000)

x_train3 = count_vect.fit_transform(x_train3)
x_test3 = count_vect.transform(x_test3)
print(x_train3.shape, x_test3.shape)
```

```
5593
2398
(5593, 1000) (2398, 1000)
```

[20]:
```
#Bag of n gram 5000 feat [bi,tri grams]
x_train4, x_test4, y_train4, y_test4 = train_test_split(x, y, test_size = 0.3,␣
 ↪random_state = 0, shuffle = True)
print(len(x_train4))
```

```
print(len(x_test4))
count_vect2 = CountVectorizer(ngram_range=(2,3),max_features = 5000)

x_train4 = count_vect2.fit_transform(x_train4)
x_test4 = count_vect2.transform(x_test4)
print(x_train4.shape, x_test4.shape)
```

```
5593
2398
(5593, 5000) (2398, 5000)
```

[21]:
```
#TF-IDF 1000 feat
x_train5, x_test5, y_train5, y_test5 = train_test_split(x, y, test_size = 0.3,␣
 ↪random_state = 0, shuffle = True)
print(len(x_train5))
print(len(x_test5))

tfidf = TfidfVectorizer(max_features = 1000)
x_train5 = tfidf.fit_transform(x_train5)
x_test5 = tfidf.fit_transform(x_test5)
print(x_train5.shape, x_test5.shape)
```

```
5593
2398
(5593, 1000) (2398, 1000)
```

[22]:
```
#TF-IDF 5000 feat
x_train6, x_test6, y_train6, y_test6 = train_test_split(x, y, test_size = 0.3,␣
 ↪random_state = 0, shuffle = True)
print(len(x_train6))
print(len(x_test6))

tfidf2 = TfidfVectorizer(max_features = 5000)
x_train6 = tfidf2.fit_transform(x_train6)
x_test6 = tfidf2.fit_transform(x_test6)
print(x_train6.shape, x_test6.shape)
```

```
5593
2398
(5593, 5000) (2398, 5000)
```

[23]:
```
#NORMAL
X_train, X_test, Y_train, Y_test = train_test_split(x, y, random_state=1)
vect = CountVectorizer(preprocessor=clean)
X_train_dtm = vect.fit_transform(X_train)# use it to extract features from␣
 ↪training data
# transform testing data (using training data's features)
```

```
X_test_dtm = vect.transform(X_test)

n_words1 = x_test1.shape[1]
n_words2 = x_test2.shape[1]
n_words3 = x_test3.shape[1]
n_words4 = x_test4.shape[1]
n_words5 = x_test5.shape[1]
n_words6 = x_test6.shape[1]
```

[24]:
```
#create list of model and accuracy dicts
perform_list1 = [ ]
perform_list2 = [ ]
perform_list3 = [ ]
perform_list4 = [ ]
perform_list5 = [ ]
```

[25]:
```
def run_models(x_train, x_test, y_train, y_test, n_words):

  mdl1=''
  mdl2=''
  mdl3=''
  mdl4=''
  mdl5=''

#Multinomial Naive Bayes
  mdl1 = MultinomialNB(alpha=1.0,fit_prior=True)

#Logistic Regression
  mdl2 = LogisticRegression()

#Support Vector Classifer
  mdl3 = SVC()

#Random Forest
  mdl4 = RandomForestClassifier(n_estimators=100 ,criterion='entropy' ,␣
 ↪random_state=0)

#ANN
  mdl5 = Sequential()
  mdl5.add(Dense(50, input_shape=(n_words,), activation='relu'))
  #mdl5.add(Dense(50, input_shape=(n_words,), activation='relu'))  #NEWW
  mdl5.add(Dense(1, activation='sigmoid'))
  mdl5.compile(loss='binary_crossentropy', optimizer='adam',␣
 ↪metrics=['accuracy'])

#-------------------------------------------------------------------------------------------
  print()
```

```
print("FOR NAIVE BAYES: ")
print()
mdl1.fit(x_train, y_train)
y_pred = mdl1.predict(x_test)
# Performance metrics

accuracy = round(accuracy_score(y_test, y_pred) * 100, 2)

# Get precision, recall, f1 scores

precision, recall, f1score, support = score(y_test, y_pred, average='micro')

print('Test Accuracy Score of Basic Naive Bayes Model:',accuracy)

print('Precision :',precision)

print('Recall :',recall)

print('F1-score :',f1score)

#calculate AUC of model
y_pred_prob = mdl1.predict_proba(x_test)[:, 1]
auc1 = metrics.roc_auc_score(y_test, y_pred_prob)
print("ROC_AOC_Score for Naive Bayes: ", auc1)


# Add performance parameters to list

perform_list1.append(dict([('Model', 'Naive Bayes'),
                           ('Test Accuracy', round(accuracy, 2)),('Precision',␣
↪round(precision, 2)),('Recall', round(recall, 2)),('F1', round(f1score,␣
↪2)),('ROC-AUC', round(auc1, 2))]))

#-------------------------------------------------------------------------------------
print()
print("FOR LOGISTIC REGRESSION: ")
print()
mdl2.fit(x_train, y_train)
y_pred2 = mdl2.predict(x_test)
# Performance metrics

accuracy2 = round(accuracy_score(y_test, y_pred2) * 100, 2)

# Get precision, recall, f1 scores

precision2, recall2, f1score2, support2 = score(y_test, y_pred2,␣
↪average='micro')
```

```
print('Test Accuracy Score of Basic Logistic Regression Model:',accuracy2)

print('Precision :',precision2)

print('Recall :',recall2)

print('F1-score :',f1score2)

#calculate AUC of model
y_pred_prob = mdl2.predict_proba(x_test)[:, 1]
auc2 = metrics.roc_auc_score(y_test, y_pred_prob)
print("ROC_AOC_Score for Logistic Regression: ", auc2)

# Add performance parameters to list

perform_list2.append(dict([('Model', 'Logistic Regression'),
                           ('Test Accuracy', round(accuracy2,␣
↪2)),('Precision', round(precision2, 2)),('Recall', round(recall2, 2)),('F1',␣
↪round(f1score2, 2)),('ROC-AUC', round(auc2, 2))]))


#-------------------------------------------------------------------------------------

print()
print("FOR LINEAR SVC: ")
print()


mdl3.fit(x_train, y_train)
y_pred3 = mdl3.predict(x_test)
# Performance metrics

accuracy3 = round(accuracy_score(y_test, y_pred3) * 100, 2)

# Get precision, recall, f1 scores

precision3, recall3, f1score3, support3 = score(y_test, y_pred3,␣
↪average='micro')

print('Test Accuracy Score of Basic Linear SVC Model:',accuracy3)

print('Precision :',precision3)

print('Recall :',recall3)

print('F1-score :',f1score3)
```

```
 #calculate AUC of model
 #y_pred_prob = mdl3.predict_proba(x_test)[:, 1]
 #auc3 = metrics.roc_auc_score(y_test, y_pred_prob)
 #print("ROC_AOC_Score for Linear SVC: ", auc3)

 # Add performance parameters to list

 perform_list3.append(dict([('Model', 'Linear SVC'),
                            ('Test Accuracy', round(accuracy3,␣
↪2)),('Precision', round(precision3, 2)),('Recall', round(recall3, 2)),('F1',␣
↪round(f1score3, 2))]))


#-------------------------------------------------------------------------------------------------


 print()
 print("FOR RANDOM FOREST: ")
 print()
 mdl4.fit(x_train, y_train)
 y_pred4 = mdl4.predict(x_test)
 # Performance metrics

 accuracy4 = round(accuracy_score(y_test, y_pred4) * 100, 2)



 # Get precision, recall, f1 scores

 precision4, recall4, f1score4, support4 = score(y_test, y_pred4,␣
↪average='micro')

 print('Test Accuracy Score of Basic Random Forest Model:',accuracy4)

 print('Precision :',precision4)

 print('Recall :',recall4)

 print('F1-score :',f1score4)

 #calculate AUC of model
 y_pred_prob = mdl4.predict_proba(x_test)[:, 1]
 auc4 = metrics.roc_auc_score(y_test, y_pred_prob)
 print("ROC_AOC_Score for Random Forest: ", auc4)
```

```
  # Add performance parameters to list

  perform_list4.append(dict([('Model', 'Random Forest'),
                             ('Test Accuracy', round(accuracy4,␣
↪2)),('Precision', round(precision4, 2)),('Recall', round(recall4, 2)),('F1',␣
↪round(f1score4, 2)),('ROC-AUC', round(auc4, 2))]))



#-------------------------------------------------------------------------------

  print()
  print("FOR ANN: ")
  print()
  mdl5.summary()
  x_train = x_train.toarray()
  y_train = np.array(y_train)
  x_test = x_test.toarray()
  y_test = np.array(y_test)
  mdl5.fit(x_train, y_train, epochs=69, verbose=2)  #SAME
  loss, acc = mdl5.evaluate(x_test, y_test, verbose=0)
  #calculate AUC of model
  #y_pred_prob = mdl5.predict_proba(x_test)[:, 1]
  #auc5 = metrics.roc_auc_score(y_test, y_pred_prob)
  #print("ROC_AOC_Score for ANN: ", auc5)
  print('Test Accuracy:',acc)
  perform_list5.append(dict([('Model', 'ANN'),('Test Accuracy', round(acc,␣
↪2)),('Loss',round(loss,2))]))
```

[26]:
```
#BoW 1000 feat
run_models(x_train1, x_test1, y_train1, y_test1, n_words1)
```

```
Metal device set to: Apple M1

FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 67.64
Precision : 0.6763969974979149
Recall : 0.6763969974979149
F1-score : 0.6763969974979149
ROC_AOC_Score for Naive Bayes:  0.7230764847947851

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 77.31
```

```
Precision : 0.7731442869057548
Recall : 0.7731442869057548
F1-score : 0.7731442869057548
ROC_AOC_Score for Logistic Regression:  0.6721538695885237

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.61
Precision : 0.8160967472894078
Recall : 0.8160967472894078
F1-score : 0.8160967472894078

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.44
Precision : 0.8144286905754796
Recall : 0.8144286905754796
F1-score : 0.8144286905754796
ROC_AOC_Score for Random Forest:  0.7156962608183014

FOR ANN:

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 50)                50050

 dense_1 (Dense)             (None, 1)                 51

=================================================================
Total params: 50,101
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69

2023-05-28 00:56:58.047999: W
tensorflow/tsl/platform/profile_utils/cpu_utils.cc:128] Failed to get CPU
frequency: 0 Hz

175/175 - 4s - loss: 0.4473 - accuracy: 0.8196 - 4s/epoch - 21ms/step
Epoch 2/69
175/175 - 1s - loss: 0.3668 - accuracy: 0.8453 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.3106 - accuracy: 0.8714 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 1s - loss: 0.2504 - accuracy: 0.9054 - 1s/epoch - 7ms/step
Epoch 5/69
```

```
175/175 - 1s - loss: 0.1902 - accuracy: 0.9374 - 1s/epoch - 8ms/step
Epoch 6/69
175/175 - 1s - loss: 0.1391 - accuracy: 0.9623 - 1s/epoch - 7ms/step
Epoch 7/69
175/175 - 1s - loss: 0.0960 - accuracy: 0.9832 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 2s - loss: 0.0660 - accuracy: 0.9903 - 2s/epoch - 9ms/step
Epoch 9/69
175/175 - 1s - loss: 0.0456 - accuracy: 0.9959 - 1s/epoch - 8ms/step
Epoch 10/69
175/175 - 1s - loss: 0.0321 - accuracy: 0.9980 - 1s/epoch - 8ms/step
Epoch 11/69
175/175 - 2s - loss: 0.0240 - accuracy: 0.9979 - 2s/epoch - 9ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0184 - accuracy: 0.9982 - 1s/epoch - 9ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0142 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0130 - accuracy: 0.9980 - 1s/epoch - 9ms/step
Epoch 15/69
175/175 - 1s - loss: 0.0094 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0087 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0065 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0067 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 2s - loss: 0.0060 - accuracy: 0.9989 - 2s/epoch - 9ms/step
Epoch 20/69
175/175 - 1s - loss: 0.0046 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 1s - loss: 0.0048 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 22/69
175/175 - 1s - loss: 0.0034 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 23/69
175/175 - 1s - loss: 0.0036 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 0.0028 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 26/69
175/175 - 1s - loss: 0.0031 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 27/69
175/175 - 1s - loss: 0.0024 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 0.0013 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 29/69
```

```
175/175 - 1s - loss: 0.0026 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 0.0012 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 1s - loss: 0.0014 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 32/69
175/175 - 2s - loss: 0.0016 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 33/69
175/175 - 2s - loss: 0.0015 - accuracy: 0.9996 - 2s/epoch - 10ms/step
Epoch 34/69
175/175 - 2s - loss: 9.3897e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 35/69
175/175 - 2s - loss: 9.2603e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 36/69
175/175 - 2s - loss: 0.0010 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 37/69
175/175 - 1s - loss: 7.6178e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 2s - loss: 8.2363e-04 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 39/69
175/175 - 2s - loss: 9.8051e-04 - accuracy: 0.9996 - 2s/epoch - 10ms/step
Epoch 40/69
175/175 - 1s - loss: 3.1867e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 41/69
175/175 - 1s - loss: 2.7411e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 42/69
175/175 - 1s - loss: 2.4430e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 43/69
175/175 - 1s - loss: 2.2024e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 44/69
175/175 - 1s - loss: 1.9776e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 45/69
175/175 - 1s - loss: 1.7945e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 46/69
175/175 - 1s - loss: 1.6248e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 47/69
175/175 - 1s - loss: 1.4970e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 48/69
175/175 - 1s - loss: 1.3648e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 49/69
175/175 - 1s - loss: 1.2237e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 50/69
175/175 - 1s - loss: 1.1291e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 51/69
175/175 - 1s - loss: 9.9030e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 52/69
175/175 - 1s - loss: 9.3667e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 53/69
```

```
175/175 - 1s - loss: 8.8640e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 54/69
175/175 - 1s - loss: 1.0212e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 55/69
175/175 - 1s - loss: 0.0015 - accuracy: 0.9998 - 1s/epoch - 7ms/step
Epoch 56/69
175/175 - 1s - loss: 0.0082 - accuracy: 0.9979 - 1s/epoch - 7ms/step
Epoch 57/69
175/175 - 1s - loss: 0.0043 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 58/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 59/69
175/175 - 1s - loss: 0.0014 - accuracy: 0.9996 - 1s/epoch - 7ms/step
Epoch 60/69
175/175 - 1s - loss: 1.9795e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 61/69
175/175 - 1s - loss: 1.5518e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 62/69
175/175 - 1s - loss: 1.3155e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 63/69
175/175 - 1s - loss: 1.1574e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 64/69
175/175 - 1s - loss: 1.0261e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 65/69
175/175 - 1s - loss: 9.2636e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 66/69
175/175 - 1s - loss: 8.4163e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 67/69
175/175 - 1s - loss: 7.6232e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 68/69
175/175 - 1s - loss: 7.0284e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 69/69
175/175 - 1s - loss: 6.4961e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Test Accuracy: 0.7935779690742493
```

[27]:
```
#BoW 5000 feat
run_models(x_train2, x_test2, y_train2, y_test2, n_words2)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 67.18
Precision : 0.6718098415346122
Recall : 0.6718098415346122
F1-score : 0.6718098415346122
ROC_AOC_Score for Naive Bayes:  0.7333456589003966

FOR LOGISTIC REGRESSION:
```

```
Test Accuracy Score of Basic Logistic Regression Model: 76.27
Precision : 0.762718932443703
Recall : 0.762718932443703
F1-score : 0.7627189324437029
ROC_AOC_Score for Logistic Regression:  0.6668678292234544

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.65
Precision : 0.8165137614678899
Recall : 0.8165137614678899
F1-score : 0.81651376146789

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.61
Precision : 0.8160967472894078
Recall : 0.8160967472894078
F1-score : 0.8160967472894078
ROC_AOC_Score for Random Forest:  0.7065078994094194

FOR ANN:

Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_2 (Dense)             (None, 50)                250050

 dense_3 (Dense)             (None, 1)                 51

=================================================================
Total params: 250,101
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4324 - accuracy: 0.8185 - 2s/epoch - 11ms/step
Epoch 2/69
175/175 - 1s - loss: 0.3051 - accuracy: 0.8591 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.2013 - accuracy: 0.9192 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 1s - loss: 0.1080 - accuracy: 0.9719 - 1s/epoch - 8ms/step
Epoch 5/69
175/175 - 1s - loss: 0.0512 - accuracy: 0.9925 - 1s/epoch - 9ms/step
Epoch 6/69
```

```
175/175 - 1s - loss: 0.0249 - accuracy: 0.9964 - 1s/epoch - 8ms/step
Epoch 7/69
175/175 - 1s - loss: 0.0141 - accuracy: 0.9977 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 1s - loss: 0.0102 - accuracy: 0.9980 - 1s/epoch - 8ms/step
Epoch 9/69
175/175 - 1s - loss: 0.0075 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 10/69
175/175 - 1s - loss: 0.0069 - accuracy: 0.9986 - 1s/epoch - 8ms/step
Epoch 11/69
175/175 - 1s - loss: 0.0053 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0049 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0043 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 15/69
175/175 - 1s - loss: 0.0030 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0027 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0019 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0013 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 0.0011 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 1s - loss: 0.0013 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 1s - loss: 7.4471e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 22/69
175/175 - 1s - loss: 6.1543e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 23/69
175/175 - 1s - loss: 5.6271e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 5.1452e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 4.7678e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 26/69
175/175 - 1s - loss: 4.4923e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 27/69
175/175 - 1s - loss: 4.2657e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 4.1964e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 1s - loss: 3.9623e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 30/69
```

```
175/175 - 1s - loss: 2.8726e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 1s - loss: 4.0820e-04 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 32/69
175/175 - 2s - loss: 0.0018 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 33/69
175/175 - 2s - loss: 0.0011 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 34/69
175/175 - 2s - loss: 0.0012 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 35/69
175/175 - 2s - loss: 3.5449e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 36/69
175/175 - 1s - loss: 1.9917e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 37/69
175/175 - 1s - loss: 1.8040e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 1s - loss: 1.7148e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 39/69
175/175 - 1s - loss: 1.6442e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 40/69
175/175 - 1s - loss: 1.6173e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 41/69
175/175 - 1s - loss: 1.5755e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 1s - loss: 1.5480e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 43/69
175/175 - 1s - loss: 1.4474e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 44/69
175/175 - 1s - loss: 1.4035e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 45/69
175/175 - 1s - loss: 1.3726e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 1.3381e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 47/69
175/175 - 1s - loss: 1.3056e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 48/69
175/175 - 1s - loss: 1.2773e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 49/69
175/175 - 2s - loss: 1.2524e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 50/69
175/175 - 1s - loss: 1.2302e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 51/69
175/175 - 1s - loss: 1.2075e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 52/69
175/175 - 1s - loss: 1.1964e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 1.1677e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 54/69
```

```
175/175 - 1s - loss: 1.1485e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 55/69
175/175 - 1s - loss: 1.1307e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 56/69
175/175 - 1s - loss: 1.1171e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 57/69
175/175 - 1s - loss: 1.0980e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 58/69
175/175 - 1s - loss: 1.0828e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 59/69
175/175 - 1s - loss: 1.0626e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 60/69
175/175 - 1s - loss: 1.0487e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 61/69
175/175 - 1s - loss: 1.0312e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 62/69
175/175 - 1s - loss: 1.0176e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 63/69
175/175 - 1s - loss: 1.0053e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 64/69
175/175 - 1s - loss: 9.8916e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 65/69
175/175 - 1s - loss: 9.7720e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 66/69
175/175 - 1s - loss: 9.6707e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 67/69
175/175 - 1s - loss: 9.5295e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 68/69
175/175 - 1s - loss: 9.4253e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 69/69
175/175 - 1s - loss: 9.3287e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Test Accuracy: 0.7739782929420471
```

[28]:
```
#Bag of n grams 1000 feat [bi,tri grams]
run_models(x_train3, x_test3, y_train3, y_test3, n_words3)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 76.36
Precision : 0.7635529608006673
Recall : 0.7635529608006673
F1-score : 0.7635529608006673
ROC_AOC_Score for Naive Bayes:  0.6754207262549694

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 79.73
```

```
Precision : 0.7973311092577148
Recall : 0.7973311092577148
F1-score : 0.7973311092577148
ROC_AOC_Score for Logistic Regression:  0.6465835504842437

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.86
Precision : 0.8185988323603003
Recall : 0.8185988323603003
F1-score : 0.8185988323603003

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372
ROC_AOC_Score for Random Forest:  0.7150526569155151

FOR ANN:

Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_4 (Dense)             (None, 50)                50050

 dense_5 (Dense)             (None, 1)                 51

=================================================================
Total params: 50,101
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.5077 - accuracy: 0.8178 - 2s/epoch - 10ms/step
Epoch 2/69
175/175 - 1s - loss: 0.4038 - accuracy: 0.8328 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.3577 - accuracy: 0.8507 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 1s - loss: 0.3197 - accuracy: 0.8714 - 1s/epoch - 8ms/step
Epoch 5/69
175/175 - 1s - loss: 0.2809 - accuracy: 0.8917 - 1s/epoch - 8ms/step
Epoch 6/69
175/175 - 1s - loss: 0.2453 - accuracy: 0.9126 - 1s/epoch - 8ms/step
Epoch 7/69
```

```
175/175 - 1s - loss: 0.2115 - accuracy: 0.9324 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 1s - loss: 0.1806 - accuracy: 0.9433 - 1s/epoch - 8ms/step
Epoch 9/69
175/175 - 1s - loss: 0.1527 - accuracy: 0.9564 - 1s/epoch - 8ms/step
Epoch 10/69
175/175 - 1s - loss: 0.1285 - accuracy: 0.9682 - 1s/epoch - 8ms/step
Epoch 11/69
175/175 - 1s - loss: 0.1086 - accuracy: 0.9737 - 1s/epoch - 8ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0913 - accuracy: 0.9809 - 1s/epoch - 8ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0771 - accuracy: 0.9846 - 1s/epoch - 7ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0652 - accuracy: 0.9882 - 1s/epoch - 8ms/step
Epoch 15/69
175/175 - 1s - loss: 0.0557 - accuracy: 0.9903 - 1s/epoch - 8ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0477 - accuracy: 0.9921 - 1s/epoch - 8ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0411 - accuracy: 0.9928 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0357 - accuracy: 0.9945 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 2s - loss: 0.0309 - accuracy: 0.9954 - 2s/epoch - 9ms/step
Epoch 20/69
175/175 - 2s - loss: 0.0273 - accuracy: 0.9959 - 2s/epoch - 9ms/step
Epoch 21/69
175/175 - 2s - loss: 0.0241 - accuracy: 0.9964 - 2s/epoch - 10ms/step
Epoch 22/69
175/175 - 2s - loss: 0.0213 - accuracy: 0.9970 - 2s/epoch - 9ms/step
Epoch 23/69
175/175 - 1s - loss: 0.0192 - accuracy: 0.9971 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 2s - loss: 0.0173 - accuracy: 0.9973 - 2s/epoch - 9ms/step
Epoch 25/69
175/175 - 2s - loss: 0.0156 - accuracy: 0.9975 - 2s/epoch - 10ms/step
Epoch 26/69
175/175 - 2s - loss: 0.0141 - accuracy: 0.9982 - 2s/epoch - 10ms/step
Epoch 27/69
175/175 - 2s - loss: 0.0128 - accuracy: 0.9984 - 2s/epoch - 10ms/step
Epoch 28/69
175/175 - 2s - loss: 0.0117 - accuracy: 0.9986 - 2s/epoch - 9ms/step
Epoch 29/69
175/175 - 1s - loss: 0.0108 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 0.0100 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 31/69
```

```
175/175 - 1s - loss: 0.0092 - accuracy: 0.9987 - 1s/epoch - 7ms/step
Epoch 32/69
175/175 - 1s - loss: 0.0087 - accuracy: 0.9986 - 1s/epoch - 8ms/step
Epoch 33/69
175/175 - 2s - loss: 0.0085 - accuracy: 0.9984 - 2s/epoch - 10ms/step
Epoch 34/69
175/175 - 2s - loss: 0.0075 - accuracy: 0.9989 - 2s/epoch - 9ms/step
Epoch 35/69
175/175 - 2s - loss: 0.0070 - accuracy: 0.9989 - 2s/epoch - 10ms/step
Epoch 36/69
175/175 - 1s - loss: 0.0066 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 37/69
175/175 - 1s - loss: 0.0063 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 1s - loss: 0.0059 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 39/69
175/175 - 1s - loss: 0.0056 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 40/69
175/175 - 1s - loss: 0.0054 - accuracy: 0.9989 - 1s/epoch - 7ms/step
Epoch 41/69
175/175 - 1s - loss: 0.0051 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 1s - loss: 0.0049 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 43/69
175/175 - 1s - loss: 0.0047 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 44/69
175/175 - 1s - loss: 0.0045 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 45/69
175/175 - 1s - loss: 0.0043 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 0.0042 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 47/69
175/175 - 1s - loss: 0.0041 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 48/69
175/175 - 1s - loss: 0.0039 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 49/69
175/175 - 1s - loss: 0.0038 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 50/69
175/175 - 1s - loss: 0.0037 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 51/69
175/175 - 1s - loss: 0.0036 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 52/69
175/175 - 1s - loss: 0.0035 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 0.0035 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 1s - loss: 0.0033 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 55/69
```

```
175/175 - 1s - loss: 0.0034 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 56/69
175/175 - 1s - loss: 0.0041 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 57/69
175/175 - 1s - loss: 0.0033 - accuracy: 0.9991 - 1s/epoch - 7ms/step
Epoch 58/69
175/175 - 1s - loss: 0.0031 - accuracy: 0.9991 - 1s/epoch - 7ms/step
Epoch 59/69
175/175 - 1s - loss: 0.0031 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 60/69
175/175 - 1s - loss: 0.0031 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 61/69
175/175 - 1s - loss: 0.0030 - accuracy: 0.9991 - 1s/epoch - 7ms/step
Epoch 62/69
175/175 - 1s - loss: 0.0030 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 63/69
175/175 - 1s - loss: 0.0029 - accuracy: 0.9989 - 1s/epoch - 7ms/step
Epoch 64/69
175/175 - 1s - loss: 0.0029 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 65/69
175/175 - 1s - loss: 0.0029 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 66/69
175/175 - 1s - loss: 0.0028 - accuracy: 0.9989 - 1s/epoch - 7ms/step
Epoch 67/69
175/175 - 1s - loss: 0.0028 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 68/69
175/175 - 1s - loss: 0.0028 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 69/69
175/175 - 1s - loss: 0.0027 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Test Accuracy: 0.7760633826255798
```

[29]:
```
#Bag of n grams 5000 feat [bi,tri grams]
run_models(x_train4, x_test4, y_train4, y_test4, n_words4)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 75.73
Precision : 0.7572977481234362
Recall : 0.7572977481234362
F1-score : 0.7572977481234362
ROC_AOC_Score for Naive Bayes:  0.6868817594398147

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 78.27
Precision : 0.7827356130108424
Recall : 0.7827356130108424
```

```
F1-score : 0.7827356130108424
ROC_AOC_Score for Logistic Regression:  0.6406934410541383

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.65
Precision : 0.8165137614678899
Recall : 0.8165137614678899
F1-score : 0.81651376146789

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.61
Precision : 0.8160967472894078
Recall : 0.8160967472894078
F1-score : 0.8160967472894078
ROC_AOC_Score for Random Forest:  0.7019695325935668

FOR ANN:

Model: "sequential_3"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_6 (Dense)             (None, 50)                250050

 dense_7 (Dense)             (None, 1)                 51

=================================================================
Total params: 250,101
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4938 - accuracy: 0.8210 - 2s/epoch - 14ms/step
Epoch 2/69
175/175 - 1s - loss: 0.3420 - accuracy: 0.8461 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.2400 - accuracy: 0.9054 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 1s - loss: 0.1452 - accuracy: 0.9539 - 1s/epoch - 8ms/step
Epoch 5/69
175/175 - 1s - loss: 0.0808 - accuracy: 0.9798 - 1s/epoch - 9ms/step
Epoch 6/69
175/175 - 1s - loss: 0.0463 - accuracy: 0.9932 - 1s/epoch - 8ms/step
Epoch 7/69
175/175 - 2s - loss: 0.0292 - accuracy: 0.9961 - 2s/epoch - 10ms/step
Epoch 8/69
```

```
175/175 - 2s - loss: 0.0198 - accuracy: 0.9975 - 2s/epoch - 9ms/step
Epoch 9/69
175/175 - 2s - loss: 0.0147 - accuracy: 0.9986 - 2s/epoch - 9ms/step
Epoch 10/69
175/175 - 2s - loss: 0.0116 - accuracy: 0.9982 - 2s/epoch - 10ms/step
Epoch 11/69
175/175 - 2s - loss: 0.0090 - accuracy: 0.9991 - 2s/epoch - 11ms/step
Epoch 12/69
175/175 - 2s - loss: 0.0075 - accuracy: 0.9991 - 2s/epoch - 10ms/step
Epoch 13/69
175/175 - 2s - loss: 0.0062 - accuracy: 0.9993 - 2s/epoch - 9ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0056 - accuracy: 0.9989 - 1s/epoch - 9ms/step
Epoch 15/69
175/175 - 1s - loss: 0.0047 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0043 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0036 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0035 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 0.0029 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 1s - loss: 0.0026 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 2s - loss: 0.0022 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 22/69
175/175 - 1s - loss: 0.0020 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 23/69
175/175 - 1s - loss: 0.0019 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0018 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 2s - loss: 0.0017 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 26/69
175/175 - 2s - loss: 0.0021 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 27/69
175/175 - 1s - loss: 0.0019 - accuracy: 0.9996 - 1s/epoch - 9ms/step
Epoch 28/69
175/175 - 2s - loss: 0.0019 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 29/69
175/175 - 2s - loss: 0.0014 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 30/69
175/175 - 1s - loss: 0.0012 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 2s - loss: 0.0012 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 32/69
```

```
175/175 - 2s - loss: 0.0012 - accuracy: 0.9998 - 2s/epoch - 11ms/step
Epoch 33/69
175/175 - 2s - loss: 0.0012 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 34/69
175/175 - 2s - loss: 0.0011 - accuracy: 0.9996 - 2s/epoch - 10ms/step
Epoch 35/69
175/175 - 2s - loss: 0.0011 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 36/69
175/175 - 2s - loss: 0.0010 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 37/69
175/175 - 1s - loss: 0.0010 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 2s - loss: 9.9393e-04 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 39/69
175/175 - 2s - loss: 9.5683e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 40/69
175/175 - 2s - loss: 9.6194e-04 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 41/69
175/175 - 2s - loss: 9.2750e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 42/69
175/175 - 2s - loss: 9.3145e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 43/69
175/175 - 1s - loss: 9.1933e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 44/69
175/175 - 1s - loss: 9.0936e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 45/69
175/175 - 2s - loss: 8.5575e-04 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 46/69
175/175 - 2s - loss: 8.8397e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 47/69
175/175 - 2s - loss: 8.5392e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 48/69
175/175 - 2s - loss: 8.5754e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 49/69
175/175 - 1s - loss: 8.2963e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 50/69
175/175 - 2s - loss: 8.7182e-04 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 51/69
175/175 - 2s - loss: 0.0010 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 52/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 8.5846e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 2s - loss: 8.5684e-04 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 55/69
175/175 - 2s - loss: 8.4558e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 56/69
```

```
175/175 - 1s - loss: 8.1162e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 57/69
175/175 - 1s - loss: 8.3182e-04 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 58/69
175/175 - 1s - loss: 7.9206e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 59/69
175/175 - 1s - loss: 8.0403e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 60/69
175/175 - 2s - loss: 7.7987e-04 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 61/69
175/175 - 2s - loss: 7.6670e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 62/69
175/175 - 2s - loss: 7.7666e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 63/69
175/175 - 2s - loss: 7.5677e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 64/69
175/175 - 2s - loss: 7.6299e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 65/69
175/175 - 1s - loss: 7.6276e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 66/69
175/175 - 2s - loss: 7.0189e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 67/69
175/175 - 2s - loss: 7.4495e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 68/69
175/175 - 2s - loss: 7.2966e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 69/69
175/175 - 1s - loss: 7.3996e-04 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Test Accuracy: 0.7527105808258057
```

[30]:
```
#TF-IDF 1000 feat
run_models(x_train5, x_test5, y_train5, y_test5, n_words5)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372
ROC_AOC_Score for Naive Bayes:  0.5571260963650043

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372
ROC_AOC_Score for Logistic Regression:  0.5745935179145141
```

```
FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372
ROC_AOC_Score for Random Forest:  0.5970522127299852

FOR ANN:

Model: "sequential_4"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_8 (Dense)             (None, 50)                50050

 dense_9 (Dense)             (None, 1)                 51

=================================================================
Total params: 50,101
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4681 - accuracy: 0.8235 - 2s/epoch - 13ms/step
Epoch 2/69
175/175 - 2s - loss: 0.4023 - accuracy: 0.8255 - 2s/epoch - 9ms/step
Epoch 3/69
175/175 - 2s - loss: 0.3803 - accuracy: 0.8312 - 2s/epoch - 10ms/step
Epoch 4/69
175/175 - 2s - loss: 0.3628 - accuracy: 0.8412 - 2s/epoch - 10ms/step
Epoch 5/69
175/175 - 2s - loss: 0.3459 - accuracy: 0.8500 - 2s/epoch - 10ms/step
Epoch 6/69
175/175 - 2s - loss: 0.3287 - accuracy: 0.8638 - 2s/epoch - 10ms/step
Epoch 7/69
175/175 - 2s - loss: 0.3113 - accuracy: 0.8691 - 2s/epoch - 10ms/step
Epoch 8/69
175/175 - 2s - loss: 0.2935 - accuracy: 0.8781 - 2s/epoch - 9ms/step
Epoch 9/69
```

```
175/175 - 2s - loss: 0.2769 - accuracy: 0.8900 - 2s/epoch - 9ms/step
Epoch 10/69
175/175 - 1s - loss: 0.2587 - accuracy: 0.8997 - 1s/epoch - 9ms/step
Epoch 11/69
175/175 - 2s - loss: 0.2434 - accuracy: 0.9102 - 2s/epoch - 10ms/step
Epoch 12/69
175/175 - 2s - loss: 0.2256 - accuracy: 0.9222 - 2s/epoch - 10ms/step
Epoch 13/69
175/175 - 2s - loss: 0.2097 - accuracy: 0.9313 - 2s/epoch - 10ms/step
Epoch 14/69
175/175 - 2s - loss: 0.1939 - accuracy: 0.9394 - 2s/epoch - 10ms/step
Epoch 15/69
175/175 - 2s - loss: 0.1794 - accuracy: 0.9481 - 2s/epoch - 10ms/step
Epoch 16/69
175/175 - 2s - loss: 0.1642 - accuracy: 0.9553 - 2s/epoch - 10ms/step
Epoch 17/69
175/175 - 2s - loss: 0.1505 - accuracy: 0.9625 - 2s/epoch - 10ms/step
Epoch 18/69
175/175 - 1s - loss: 0.1377 - accuracy: 0.9685 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 0.1256 - accuracy: 0.9718 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 1s - loss: 0.1141 - accuracy: 0.9769 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 1s - loss: 0.1027 - accuracy: 0.9802 - 1s/epoch - 8ms/step
Epoch 22/69
175/175 - 1s - loss: 0.0929 - accuracy: 0.9830 - 1s/epoch - 8ms/step
Epoch 23/69
175/175 - 1s - loss: 0.0833 - accuracy: 0.9859 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0749 - accuracy: 0.9877 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 0.0667 - accuracy: 0.9911 - 1s/epoch - 7ms/step
Epoch 26/69
175/175 - 1s - loss: 0.0597 - accuracy: 0.9932 - 1s/epoch - 8ms/step
Epoch 27/69
175/175 - 1s - loss: 0.0533 - accuracy: 0.9955 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 0.0479 - accuracy: 0.9966 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 1s - loss: 0.0426 - accuracy: 0.9977 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 0.0376 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 1s - loss: 0.0336 - accuracy: 0.9989 - 1s/epoch - 7ms/step
Epoch 32/69
175/175 - 1s - loss: 0.0302 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 33/69
```

```
175/175 - 1s - loss: 0.0268 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 34/69
175/175 - 1s - loss: 0.0239 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 35/69
175/175 - 1s - loss: 0.0215 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 36/69
175/175 - 1s - loss: 0.0191 - accuracy: 0.9991 - 1s/epoch - 7ms/step
Epoch 37/69
175/175 - 1s - loss: 0.0171 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 38/69
175/175 - 1s - loss: 0.0155 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 39/69
175/175 - 1s - loss: 0.0141 - accuracy: 0.9991 - 1s/epoch - 7ms/step
Epoch 40/69
175/175 - 1s - loss: 0.0125 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 41/69
175/175 - 1s - loss: 0.0114 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 1s - loss: 0.0101 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 43/69
175/175 - 1s - loss: 0.0092 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 44/69
175/175 - 1s - loss: 0.0083 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 45/69
175/175 - 1s - loss: 0.0077 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 0.0069 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 47/69
175/175 - 1s - loss: 0.0063 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 48/69
175/175 - 1s - loss: 0.0057 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 49/69
175/175 - 1s - loss: 0.0052 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 50/69
175/175 - 1s - loss: 0.0049 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 51/69
175/175 - 1s - loss: 0.0047 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 52/69
175/175 - 1s - loss: 0.0044 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 0.0040 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 1s - loss: 0.0035 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 55/69
175/175 - 1s - loss: 0.0029 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 56/69
175/175 - 1s - loss: 0.0027 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 57/69
```

```
175/175 - 1s - loss: 0.0025 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 58/69
175/175 - 1s - loss: 0.0023 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 59/69
175/175 - 1s - loss: 0.0022 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 60/69
175/175 - 1s - loss: 0.0020 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 61/69
175/175 - 1s - loss: 0.0024 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 62/69
175/175 - 1s - loss: 0.0024 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 63/69
175/175 - 1s - loss: 0.0016 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 64/69
175/175 - 1s - loss: 0.0015 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 65/69
175/175 - 1s - loss: 0.0014 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 66/69
175/175 - 1s - loss: 0.0013 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 67/69
175/175 - 1s - loss: 0.0011 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 68/69
175/175 - 1s - loss: 0.0011 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 69/69
175/175 - 1s - loss: 0.0010 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Test Accuracy: 0.7727272510528564
```

[31]:
```
#TF-IDF 5000 feat
run_models(x_train6, x_test6, y_train6, y_test6, n_words6)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 80.15
Precision : 0.8015012510425354
Recall : 0.8015012510425354
F1-score : 0.8015012510425356
ROC_AOC_Score for Naive Bayes:  0.5291551986567458

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372
ROC_AOC_Score for Logistic Regression:  0.547886572224916

FOR LINEAR SVC:
```

```
Test Accuracy Score of Basic Linear SVC Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372
ROC_AOC_Score for Random Forest:  0.5820696720120093

FOR ANN:

Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_10 (Dense)            (None, 50)                250050

 dense_11 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4696 - accuracy: 0.8242 - 2s/epoch - 11ms/step
Epoch 2/69
175/175 - 1s - loss: 0.3772 - accuracy: 0.8246 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.3329 - accuracy: 0.8439 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 1s - loss: 0.2840 - accuracy: 0.8736 - 1s/epoch - 9ms/step
Epoch 5/69
175/175 - 1s - loss: 0.2335 - accuracy: 0.9047 - 1s/epoch - 8ms/step
Epoch 6/69
175/175 - 1s - loss: 0.1883 - accuracy: 0.9356 - 1s/epoch - 8ms/step
Epoch 7/69
175/175 - 1s - loss: 0.1476 - accuracy: 0.9573 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 1s - loss: 0.1148 - accuracy: 0.9723 - 1s/epoch - 8ms/step
Epoch 9/69
175/175 - 1s - loss: 0.0874 - accuracy: 0.9852 - 1s/epoch - 8ms/step
Epoch 10/69
```

```
175/175 - 1s - loss: 0.0662 - accuracy: 0.9912 - 1s/epoch - 8ms/step
Epoch 11/69
175/175 - 1s - loss: 0.0508 - accuracy: 0.9943 - 1s/epoch - 8ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0389 - accuracy: 0.9966 - 1s/epoch - 8ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0302 - accuracy: 0.9971 - 1s/epoch - 8ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0242 - accuracy: 0.9973 - 1s/epoch - 8ms/step
Epoch 15/69
175/175 - 1s - loss: 0.0194 - accuracy: 0.9982 - 1s/epoch - 8ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0160 - accuracy: 0.9984 - 1s/epoch - 8ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0136 - accuracy: 0.9979 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0112 - accuracy: 0.9982 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 0.0096 - accuracy: 0.9982 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 1s - loss: 0.0082 - accuracy: 0.9984 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 1s - loss: 0.0071 - accuracy: 0.9986 - 1s/epoch - 8ms/step
Epoch 22/69
175/175 - 1s - loss: 0.0061 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 23/69
175/175 - 1s - loss: 0.0053 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0047 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 0.0042 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 26/69
175/175 - 1s - loss: 0.0038 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 27/69
175/175 - 1s - loss: 0.0033 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 0.0030 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 2s - loss: 0.0028 - accuracy: 0.9991 - 2s/epoch - 10ms/step
Epoch 30/69
175/175 - 2s - loss: 0.0027 - accuracy: 0.9989 - 2s/epoch - 9ms/step
Epoch 31/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 32/69
175/175 - 1s - loss: 0.0024 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 33/69
175/175 - 1s - loss: 0.0021 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 34/69
```

```
175/175 - 88s - loss: 0.0018 - accuracy: 0.9995 - 88s/epoch - 506ms/step
Epoch 35/69
175/175 - 2s - loss: 0.0016 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 36/69
175/175 - 1s - loss: 0.0015 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 37/69
175/175 - 2s - loss: 0.0014 - accuracy: 1.0000 - 2s/epoch - 11ms/step
Epoch 38/69
175/175 - 2s - loss: 0.0013 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 39/69
175/175 - 1s - loss: 0.0013 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 40/69
175/175 - 1s - loss: 0.0012 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 41/69
175/175 - 1s - loss: 0.0011 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 1s - loss: 0.0010 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 43/69
175/175 - 1s - loss: 9.6862e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 44/69
175/175 - 1s - loss: 8.9303e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 45/69
175/175 - 1s - loss: 8.5007e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 8.1221e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 47/69
175/175 - 1s - loss: 7.6160e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 48/69
175/175 - 1s - loss: 7.1402e-04 - accuracy: 1.0000 - 1s/epoch - 9ms/step
Epoch 49/69
175/175 - 1s - loss: 6.7834e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 50/69
175/175 - 1s - loss: 6.4980e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 51/69
175/175 - 1s - loss: 6.2150e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 52/69
175/175 - 1s - loss: 6.0272e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 5.6900e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 1s - loss: 5.4333e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 55/69
175/175 - 1s - loss: 5.2646e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 56/69
175/175 - 1s - loss: 5.0107e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 57/69
175/175 - 1s - loss: 4.7918e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 58/69
```

```
175/175 - 1s - loss: 4.6813e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 59/69
175/175 - 1s - loss: 4.5587e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 60/69
175/175 - 1s - loss: 4.3831e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 61/69
175/175 - 1s - loss: 4.2270e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 62/69
175/175 - 1s - loss: 4.0843e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 63/69
175/175 - 1s - loss: 3.9510e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 64/69
175/175 - 1s - loss: 3.8524e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 65/69
175/175 - 1s - loss: 3.7416e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 66/69
175/175 - 1s - loss: 3.5931e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 67/69
175/175 - 1s - loss: 3.4967e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 68/69
175/175 - 1s - loss: 3.1134e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 69/69
175/175 - 1s - loss: 2.9328e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Test Accuracy: 0.7931609749794006
```

[88]:
```
perform_list = perform_list1 + perform_list2 + perform_list3 + perform_list4 +␣
 ↪perform_list5
model_performance_I = pd.DataFrame(data=perform_list)
model_performance_I = model_performance_I[['Model', 'Test Accuracy', 'ROC-AUC',␣
 ↪'Loss', 'Precision', 'Recall', 'F1']]
model_performance_I
```

[88]:

	Model	Test Accuracy	ROC-AUC	Loss	Precision	Recall	
0	Naive Bayes	67.76	0.72	NaN	0.68	0.68	\
1	Naive Bayes	67.26	0.73	NaN	0.67	0.67	
2	Naive Bayes	76.65	0.68	NaN	0.77	0.77	
3	Naive Bayes	75.65	0.69	NaN	0.76	0.76	
4	Naive Bayes	82.19	0.73	NaN	0.82	0.82	
5	Naive Bayes	81.78	0.72	NaN	0.82	0.82	
6	Naive Bayes	66.96	0.73	NaN	0.67	0.67	
7	Naive Bayes	68.50	0.74	NaN	0.69	0.69	
8	Naive Bayes	66.86	0.68	NaN	0.67	0.67	
9	Naive Bayes	66.57	0.70	NaN	0.67	0.67	
10	Naive Bayes	65.99	0.73	NaN	0.66	0.66	
11	Naive Bayes	68.21	0.74	NaN	0.68	0.68	
12	Naive Bayes	67.92	0.74	NaN	0.68	0.68	
13	Naive Bayes	70.75	0.79	NaN	0.71	0.71	

14	Naive Bayes	70.13	0.73	NaN	0.70	0.70
15	Naive Bayes	73.71	0.79	NaN	0.74	0.74
16	Naive Bayes	68.64	0.74	NaN	0.69	0.69
17	Naive Bayes	70.93	0.79	NaN	0.71	0.71
18	Naive Bayes	71.44	0.77	NaN	0.71	0.71
19	Naive Bayes	75.26	0.78	NaN	0.75	0.75
20	Naive Bayes	66.07	0.77	NaN	0.66	0.66
21	Naive Bayes	63.83	0.77	NaN	0.64	0.64
22	Naive Bayes	70.37	0.77	NaN	0.70	0.70
23	Naive Bayes	72.37	0.80	NaN	0.72	0.72
24	Logistic Regression	77.36	0.67	NaN	0.77	0.77
25	Logistic Regression	76.36	0.66	NaN	0.76	0.76
26	Logistic Regression	79.36	0.65	NaN	0.79	0.79
27	Logistic Regression	78.15	0.64	NaN	0.78	0.78
28	Logistic Regression	81.65	0.73	NaN	0.82	0.82
29	Logistic Regression	81.69	0.74	NaN	0.82	0.82
30	Logistic Regression	61.93	0.64	NaN	0.62	0.62
31	Logistic Regression	63.96	0.67	NaN	0.64	0.64
32	Logistic Regression	65.31	0.67	NaN	0.65	0.65
33	Logistic Regression	64.54	0.66	NaN	0.65	0.65
34	Logistic Regression	68.12	0.73	NaN	0.68	0.68
35	Logistic Regression	68.79	0.75	NaN	0.69	0.69
36	Logistic Regression	72.25	0.78	NaN	0.72	0.72
37	Logistic Regression	85.91	0.89	NaN	0.86	0.86
38	Logistic Regression	73.92	0.80	NaN	0.74	0.74
39	Logistic Regression	84.99	0.90	NaN	0.85	0.85
40	Logistic Regression	72.55	0.80	NaN	0.73	0.73
41	Logistic Regression	76.78	0.86	NaN	0.77	0.77
42	Logistic Regression	81.17	0.88	NaN	0.81	0.81
43	Logistic Regression	81.77	0.88	NaN	0.82	0.82
44	Logistic Regression	77.68	0.82	NaN	0.78	0.78
45	Logistic Regression	80.27	0.85	NaN	0.80	0.80
46	Logistic Regression	74.81	0.82	NaN	0.75	0.75
47	Logistic Regression	78.66	0.87	NaN	0.79	0.79
48	Linear SVC	81.61	NaN	NaN	0.82	0.82
49	Linear SVC	81.65	NaN	NaN	0.82	0.82
50	Linear SVC	81.90	NaN	NaN	0.82	0.82
51	Linear SVC	81.65	NaN	NaN	0.82	0.82
52	Linear SVC	81.65	NaN	NaN	0.82	0.82
53	Linear SVC	81.69	NaN	NaN	0.82	0.82
54	Linear SVC	68.02	NaN	NaN	0.68	0.68
55	Linear SVC	69.47	NaN	NaN	0.69	0.69
56	Linear SVC	68.50	NaN	NaN	0.69	0.69
57	Linear SVC	67.54	NaN	NaN	0.68	0.68
58	Linear SVC	68.50	NaN	NaN	0.69	0.69
59	Linear SVC	69.28	NaN	NaN	0.69	0.69
60	Linear SVC	84.33	NaN	NaN	0.84	0.84

61	Linear SVC	87.35	NaN	NaN	0.87	0.87
62	Linear SVC	84.00	NaN	NaN	0.84	0.84
63	Linear SVC	87.71	NaN	NaN	0.88	0.88
64	Linear SVC	88.21	NaN	NaN	0.88	0.88
65	Linear SVC	90.33	NaN	NaN	0.90	0.90
66	Linear SVC	86.66	NaN	NaN	0.87	0.87
67	Linear SVC	86.93	NaN	NaN	0.87	0.87
68	Linear SVC	79.80	NaN	NaN	0.80	0.80
69	Linear SVC	83.65	NaN	NaN	0.84	0.84
70	Linear SVC	89.35	NaN	NaN	0.89	0.89
71	Linear SVC	89.67	NaN	NaN	0.90	0.90
72	Random Forest	81.53	0.71	NaN	0.82	0.82
73	Random Forest	81.61	0.71	NaN	0.82	0.82
74	Random Forest	81.69	0.71	NaN	0.82	0.82
75	Random Forest	81.44	0.69	NaN	0.81	0.81
76	Random Forest	81.78	0.72	NaN	0.82	0.82
77	Random Forest	81.48	0.73	NaN	0.81	0.81
78	Random Forest	67.54	0.73	NaN	0.68	0.68
79	Random Forest	68.89	0.74	NaN	0.69	0.69
80	Random Forest	66.96	0.71	NaN	0.67	0.67
81	Random Forest	67.25	0.71	NaN	0.67	0.67
82	Random Forest	67.05	0.72	NaN	0.67	0.67
83	Random Forest	70.53	0.76	NaN	0.71	0.71
84	Random Forest	93.02	0.97	NaN	0.93	0.93
85	Random Forest	95.11	0.98	NaN	0.95	0.95
86	Random Forest	90.81	0.97	NaN	0.91	0.91
87	Random Forest	91.73	0.97	NaN	0.92	0.92
88	Random Forest	93.82	0.97	NaN	0.94	0.94
89	Random Forest	93.26	0.97	NaN	0.93	0.93
90	Random Forest	86.21	0.91	NaN	0.86	0.86
91	Random Forest	86.63	0.91	NaN	0.87	0.87
92	Random Forest	81.11	0.87	NaN	0.81	0.81
93	Random Forest	80.51	0.88	NaN	0.81	0.81
94	Random Forest	89.29	0.95	NaN	0.89	0.89
95	Random Forest	88.87	0.95	NaN	0.89	0.89
96	ANN	0.79	NaN	1.71	NaN	NaN
97	ANN	0.79	NaN	1.96	NaN	NaN
98	ANN	0.77	NaN	1.83	NaN	NaN
99	ANN	0.77	NaN	2.64	NaN	NaN
100	ANN	0.78	NaN	1.23	NaN	NaN
101	ANN	0.77	NaN	2.02	NaN	NaN
102	ANN	0.64	NaN	2.08	NaN	NaN
103	ANN	0.66	NaN	2.18	NaN	NaN
104	ANN	0.63	NaN	2.23	NaN	NaN
105	ANN	0.64	NaN	2.45	NaN	NaN
106	ANN	0.62	NaN	1.91	NaN	NaN
107	ANN	0.64	NaN	2.03	NaN	NaN

108	ANN	0.88	NaN	0.73	NaN	NaN
109	ANN	0.89	NaN	0.79	NaN	NaN
110	ANN	0.89	NaN	0.75	NaN	NaN
111	ANN	0.88	NaN	1.12	NaN	NaN
112	ANN	0.88	NaN	0.62	NaN	NaN
113	ANN	0.87	NaN	1.38	NaN	NaN
114	ANN	0.84	NaN	1.21	NaN	NaN
115	ANN	0.84	NaN	1.62	NaN	NaN
116	ANN	0.80	NaN	1.62	NaN	NaN
117	ANN	0.82	NaN	2.24	NaN	NaN
118	ANN	0.87	NaN	0.61	NaN	NaN
119	ANN	0.87	NaN	1.39	NaN	NaN

	F1
0	0.68
1	0.67
2	0.77
3	0.76
4	0.82
5	0.82
6	0.67
7	0.69
8	0.67
9	0.67
10	0.66
11	0.68
12	0.68
13	0.71
14	0.70
15	0.74
16	0.69
17	0.71
18	0.71
19	0.75
20	0.66
21	0.64
22	0.70
23	0.72
24	0.77
25	0.76
26	0.79
27	0.78
28	0.82
29	0.82
30	0.62
31	0.64
32	0.65

33	0.65
34	0.68
35	0.69
36	0.72
37	0.86
38	0.74
39	0.85
40	0.73
41	0.77
42	0.81
43	0.82
44	0.78
45	0.80
46	0.75
47	0.79
48	0.82
49	0.82
50	0.82
51	0.82
52	0.82
53	0.82
54	0.68
55	0.69
56	0.69
57	0.68
58	0.69
59	0.69
60	0.84
61	0.87
62	0.84
63	0.88
64	0.88
65	0.90
66	0.87
67	0.87
68	0.80
69	0.84
70	0.89
71	0.90
72	0.82
73	0.82
74	0.82
75	0.81
76	0.82
77	0.81
78	0.68
79	0.69

```
80   0.67
81   0.67
82   0.67
83   0.71
84   0.93
85   0.95
86   0.91
87   0.92
88   0.94
89   0.93
90   0.86
91   0.87
92   0.81
93   0.81
94   0.89
95   0.89
96    NaN
97    NaN
98    NaN
99    NaN
100   NaN
101   NaN
102   NaN
103   NaN
104   NaN
105   NaN
106   NaN
107   NaN
108   NaN
109   NaN
110   NaN
111   NaN
112   NaN
113   NaN
114   NaN
115   NaN
116   NaN
117   NaN
118   NaN
119   NaN
```

9 CLASS BALANCING

[33]:
```
our_data["relevance"].value_counts()
```

[33]:
```
relevance
0    6571
1    1420
Name: count, dtype: int64
```

[34]:
```
our_data.isnull().sum()
```

[34]:
```
text         0
relevance    0
dtype: int64
```

[35]:
```
#NO SAMPLING_BoW 1000 feat

x8 = np.array(our_data.iloc[:,0].values)
y8 = np.array(our_data.relevance.values)

counter = Counter(y8)
print('Count: ',counter)

#BoW 1000 feat
x8 = cv.fit_transform(our_data.text).toarray()

x_train8, x_test8, y_train8, y_test8 = train_test_split(x8, y8, test_size = 0.
 ↪3, random_state = 0, shuffle = True)
print(len(x_train8))
print(len(x_test8))

x_train8=sparse.csr_matrix(x_train8)
x_test8=sparse.csr_matrix(x_test8)
print(x_train8.shape, x_test8.shape)
```

```
Count:  Counter({0: 6571, 1: 1420})
5593
2398
(5593, 1000) (2398, 1000)
```

[36]:
```
#NO SAMPLING_BoW 5000 feat

x9 = np.array(our_data.iloc[:,0].values)
y9 = np.array(our_data.relevance.values)

counter = Counter(y9)
print('Count: ',counter)

x9 = cv2.fit_transform(our_data.text).toarray()
```

```
x_train9, x_test9, y_train9, y_test9 = train_test_split(x9, y9, test_size = 0.
↪3, random_state = 0, shuffle = True)
print(len(x_train9))
print(len(x_test9))

x_train9=sparse.csr_matrix(x_train9)
x_test9=sparse.csr_matrix(x_test9)
print(x_train9.shape, x_test9.shape)
```

```
Count:  Counter({0: 6571, 1: 1420})
5593
2398
(5593, 5000) (2398, 5000)
```

[37]:
```
#NO SAMPLING_Bag of n grams 1000 feat [bi,tri grams]

x10 = np.array(our_data.iloc[:,0].values)
y10 = np.array(our_data.relevance.values)

counter = Counter(y10)
print('Count: ',counter)


x10 = count_vect.fit_transform(our_data.text).toarray()

x_train10, x_test10, y_train10, y_test10 = train_test_split(x10, y10, test_size␣
↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train10))
print(len(x_test10))

x_train10=sparse.csr_matrix(x_train10)
x_test10=sparse.csr_matrix(x_test10)
print(x_train10.shape, x_test10.shape)
```

```
Count:  Counter({0: 6571, 1: 1420})
5593
2398
(5593, 1000) (2398, 1000)
```

[38]:
```
#NO SAMPLING_Bag of n grams 5000 feat [bi,tri grams]

x11 = np.array(our_data.iloc[:,0].values)
y11 = np.array(our_data.relevance.values)

counter = Counter(y11)
print('Count: ',counter)
```

```
x11 = count_vect2.fit_transform(our_data.text).toarray()

x_train11, x_test11, y_train11, y_test11 = train_test_split(x11, y11, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train11))
print(len(x_test11))

x_train11=sparse.csr_matrix(x_train11)
x_test11=sparse.csr_matrix(x_test11)
print(x_train11.shape, x_test11.shape)
```

```
Count:  Counter({0: 6571, 1: 1420})
5593
2398
(5593, 5000) (2398, 5000)
```

[39]:
```
#NO SAMPLING_Tf-Idf 1000 feat

x12 = np.array(our_data.iloc[:,0].values)
y12 = np.array(our_data.relevance.values)

counter = Counter(y12)
print('Count: ',counter)


x12 = tfidf.fit_transform(our_data.text).toarray()

x_train12, x_test12, y_train12, y_test12 = train_test_split(x12, y12, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train12))
print(len(x_test12))

x_train12=sparse.csr_matrix(x_train12)
x_test12=sparse.csr_matrix(x_test12)
print(x_train12.shape, x_test12.shape)
```

```
Count:  Counter({0: 6571, 1: 1420})
5593
2398
(5593, 1000) (2398, 1000)
```

[40]:
```
#NO SAMPLING_Tf-Idf 5000 feat

x13 = np.array(our_data.iloc[:,0].values)
y13 = np.array(our_data.relevance.values)
```

```
counter = Counter(y13)
print('Count: ',counter)

x13 = tfidf2.fit_transform(our_data.text).toarray()

x_train13, x_test13, y_train13, y_test13 = train_test_split(x13, y13, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train13))
print(len(x_test13))

x_train13=sparse.csr_matrix(x_train13)
x_test13=sparse.csr_matrix(x_test13)
print(x_train13.shape, x_test13.shape)
```

```
Count:  Counter({0: 6571, 1: 1420})
5593
2398
(5593, 5000) (2398, 5000)
```

[41]:
```
#Under Sampling_BoW 1000 feat

x14 = np.array(our_data.iloc[:,0].values)
y14 = np.array(our_data.relevance.values)

counter = Counter(y14)
print('Before Under Sampling: ',counter)

x14 = cv.fit_transform(our_data.text).toarray()
# Undersampling
undersample = RandomUnderSampler(sampling_strategy=0.7)  #current sample in␣
 ↪minority/0.5 = updated samples in majority class NEW
x14, y14 = undersample.fit_resample(x14, y14)

x_train14, x_test14, y_train14, y_test14 = train_test_split(x14, y14, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train14))
print(len(x_test14))

counter = Counter(y14)
print('After Under Sampling: ',counter)
x_train14=sparse.csr_matrix(x_train14)
x_test14=sparse.csr_matrix(x_test14)
print(x_train14.shape, x_test14.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
2413
1035
```

```
After Under Sampling:  Counter({0: 2028, 1: 1420})
(2413, 1000) (1035, 1000)
```

[42]:
```
#Under Sampling_BoW 5000 feat

x15 = np.array(our_data.iloc[:,0].values)
y15 = np.array(our_data.relevance.values)

counter = Counter(y15)
print('Before Under Sampling: ',counter)

x15 = cv2.fit_transform(our_data.text).toarray()
# Undersampling
undersample = RandomUnderSampler(sampling_strategy=0.7)  #current sample in␣
 ↪minority/0.5 = updated samples in majority class NEW
x15, y15 = undersample.fit_resample(x15, y15)

x_train15, x_test15, y_train15, y_test15 = train_test_split(x15, y15, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train15))
print(len(x_test15))

counter = Counter(y15)
print('After Under Sampling: ',counter)
x_train15=sparse.csr_matrix(x_train15)
x_test15=sparse.csr_matrix(x_test15)
print(x_train15.shape, x_test15.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
2413
1035
After Under Sampling:  Counter({0: 2028, 1: 1420})
(2413, 5000) (1035, 5000)
```

[43]:
```
#Under Sampling_Bag of n grams 1000 feat

x16 = np.array(our_data.iloc[:,0].values)
y16 = np.array(our_data.relevance.values)

counter = Counter(y16)
print('Before Under Sampling: ',counter)

x16 = count_vect.fit_transform(our_data.text).toarray()
# Undersampling
undersample = RandomUnderSampler(sampling_strategy=0.7)  #current sample in␣
 ↪minority/0.5 = updated samples in majority class NEW
x16, y16 = undersample.fit_resample(x16, y16)
```

```
x_train16, x_test16, y_train16, y_test16 = train_test_split(x16, y16, test_size
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train16))
print(len(x_test16))

counter = Counter(y16)
print('After Under Sampling: ',counter)
x_train16=sparse.csr_matrix(x_train16)
x_test16=sparse.csr_matrix(x_test16)
print(x_train16.shape, x_test16.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
2413
1035
After Under Sampling:  Counter({0: 2028, 1: 1420})
(2413, 1000) (1035, 1000)
```

[44]:
```
#Under Sampling_Bag of n grams 5000 feat

x17 = np.array(our_data.iloc[:,0].values)
y17 = np.array(our_data.relevance.values)

counter = Counter(y17)
print('Before Under Sampling: ',counter)

x17 = count_vect2.fit_transform(our_data.text).toarray()
# Undersampling
undersample = RandomUnderSampler(sampling_strategy=0.7)  #current sample in
 ↪minority/0.5 = updated samples in majority class NEW
x17, y17 = undersample.fit_resample(x17, y17)

x_train17, x_test17, y_train17, y_test17 = train_test_split(x17, y17, test_size
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train17))
print(len(x_test17))

counter = Counter(y17)
print('After Under Sampling: ',counter)
x_train17=sparse.csr_matrix(x_train17)
x_test17=sparse.csr_matrix(x_test17)
print(x_train17.shape, x_test17.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
2413
1035
After Under Sampling:  Counter({0: 2028, 1: 1420})
```

```
(2413, 5000) (1035, 5000)
```

[45]:
```
#Under Sampling_TF-IDF 1000 feat

x18 = np.array(our_data.iloc[:,0].values)
y18 = np.array(our_data.relevance.values)

counter = Counter(y18)
print('Before Under Sampling: ',counter)

x18 = tfidf.fit_transform(our_data.text).toarray()
# Undersampling
undersample = RandomUnderSampler(sampling_strategy=0.7)  #current sample in␣
 ↪minority/0.5 = updated samples in majority class NEW
x18, y18 = undersample.fit_resample(x18, y18)

x_train18, x_test18, y_train18, y_test18 = train_test_split(x18, y18, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train18))
print(len(x_test18))

counter = Counter(y18)
print('After Under Sampling: ',counter)
x_train18=sparse.csr_matrix(x_train18)
x_test18=sparse.csr_matrix(x_test18)
print(x_train18.shape, x_test18.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
2413
1035
After Under Sampling:  Counter({0: 2028, 1: 1420})
(2413, 1000) (1035, 1000)
```

[46]:
```
#Under Sampling_TF-IDF 5000 feat

x19 = np.array(our_data.iloc[:,0].values)
y19 = np.array(our_data.relevance.values)

counter = Counter(y19)
print('Before Under Sampling: ',counter)

x19 = tfidf2.fit_transform(our_data.text).toarray()
# Undersampling
undersample = RandomUnderSampler(sampling_strategy=0.7)  #current sample in␣
 ↪minority/0.5 = updated samples in majority class NEW
x19, y19 = undersample.fit_resample(x19, y19)
```

```
x_train19, x_test19, y_train19, y_test19 = train_test_split(x19, y19, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train19))
print(len(x_test19))

counter = Counter(y19)
print('After Under Sampling: ',counter)
x_train19=sparse.csr_matrix(x_train19)
x_test19=sparse.csr_matrix(x_test19)
print(x_train19.shape, x_test19.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
2413
1035
After Under Sampling:  Counter({0: 2028, 1: 1420})
(2413, 5000) (1035, 5000)
```

[47]:
```
#Under Sampling_I (text)
'''
x9 = x
y9 = y

counter = Counter(y9)
print('Before Under Sampling: ',counter)

# transform the dataset
undersample = RandomUnderSampler(sampling_strategy=0.7)  #current sample in␣
 ↪minority/0.5 = updated samples in majority class NEW
x9=np.array(x9)
x9=x9.reshape(-1, 1)
x9, y9 = undersample.fit_resample(x9, y9)
x9=x9.reshape(-1)
x9=pd.Series(x9)
counter = Counter(y9)
print('After Under Sampling: ',counter)



#BoW 1000 feat
x_train9, x_test9, y_train9, y_test9 = train_test_split(x9, y9, test_size = 0.
 ↪3, random_state = 0, shuffle = True)
print(len(x_train9))
print(len(x_test9))

cv = CountVectorizer(max_features = 1000)
x_train9 = cv.fit_transform(x_train9)
```

```
x_test9 = cv.transform(x_test9)
print(x_train9.shape, x_test9.shape)   '''
```

[47]:
```
"\nx9 = x\ny9 = y\n\ncounter = Counter(y9)\nprint('Before Under Sampling:
',counter)\n\n# transform the dataset\nundersample =
RandomUnderSampler(sampling_strategy=0.7)  #current sample in minority/0.5 =
updated samples in majority class NEW\nx9=np.array(x9)\nx9=x9.reshape(-1,
1)\nx9, y9 = undersample.fit_resample(x9,
y9)\nx9=x9.reshape(-1)\nx9=pd.Series(x9)\ncounter = Counter(y9)\nprint('After
Under Sampling: ',counter)\n\n\n\n#BoW 1000 feat\nx_train9, x_test9, y_train9,
y_test9 = train_test_split(x9, y9, test_size = 0.3, random_state = 0, shuffle =
True)\nprint(len(x_train9))\nprint(len(x_test9))\n\ncv =
CountVectorizer(max_features = 1000)\nx_train9 =
cv.fit_transform(x_train9)\n\nx_test9 =
cv.transform(x_test9)\nprint(x_train9.shape, x_test9.shape)   "
```

[48]:
```
#Over Sampling_BoW 1000 feat

x20 = np.array(our_data.iloc[:,0].values)
y20 = np.array(our_data.relevance.values)

counter = Counter(y20)
print('Before Under Sampling: ',counter)

x20 = cv.fit_transform(our_data.text).toarray()
# Oversampling
oversample = RandomOverSampler(sampling_strategy=0.7)   #0.5 * majority class =␣
 ↪updates sample count in minority  NEW
x20, y20 = oversample.fit_resample(x20, y20)

x_train20, x_test20, y_train20, y_test20 = train_test_split(x20, y20, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train20))
print(len(x_test20))

counter = Counter(y20)
print('After Under Sampling: ',counter)
x_train20=sparse.csr_matrix(x_train20)
x_test20=sparse.csr_matrix(x_test20)
print(x_train20.shape, x_test20.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 1000) (3351, 1000)
```

[49]:
```
#Over Sampling_BoW 5000 feat

x21 = np.array(our_data.iloc[:,0].values)
y21 = np.array(our_data.relevance.values)

counter = Counter(y21)
print('Before Under Sampling: ',counter)

x21 = cv2.fit_transform(our_data.text).toarray()
# Oversampling
oversample = RandomOverSampler(sampling_strategy=0.7)   #0.5 * majority class =␣
 ↪updates sample count in minority  NEW
x21, y21 = oversample.fit_resample(x21, y21)

x_train21, x_test21, y_train21, y_test21 = train_test_split(x21, y21, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train21))
print(len(x_test21))

counter = Counter(y21)
print('After Under Sampling: ',counter)
x_train21=sparse.csr_matrix(x_train21)
x_test21=sparse.csr_matrix(x_test21)
print(x_train21.shape, x_test21.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 5000) (3351, 5000)
```

[50]:
```
#Over Sampling_Bag of n grams 1000 feat [bi,tri grams]

x22 = np.array(our_data.iloc[:,0].values)
y22 = np.array(our_data.relevance.values)

counter = Counter(y22)
print('Before Under Sampling: ',counter)

x22 = count_vect.fit_transform(our_data.text).toarray()
# Oversampling
oversample = RandomOverSampler(sampling_strategy=0.7)   #0.5 * majority class =␣
 ↪updates sample count in minority  NEW
x22, y22 = oversample.fit_resample(x22, y22)

x_train22, x_test22, y_train22, y_test22 = train_test_split(x22, y22, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
```

```
print(len(x_train22))
print(len(x_test22))

counter = Counter(y22)
print('After Under Sampling: ',counter)
x_train22=sparse.csr_matrix(x_train22)
x_test22=sparse.csr_matrix(x_test22)
print(x_train22.shape, x_test22.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 1000) (3351, 1000)
```

[51]:
```
#Over Sampling_Bag of n grams 5000 feat [bi,tri grams]

x23 = np.array(our_data.iloc[:,0].values)
y23 = np.array(our_data.relevance.values)

counter = Counter(y23)
print('Before Under Sampling: ',counter)

x23 = count_vect2.fit_transform(our_data.text).toarray()
# Oversampling
oversample = RandomOverSampler(sampling_strategy=0.7)   #0.5 * majority class =␣
 ↪updates sample count in minority  NEW
x23, y23 = oversample.fit_resample(x23, y23)

x_train23, x_test23, y_train23, y_test23 = train_test_split(x23, y23, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train23))
print(len(x_test23))

counter = Counter(y23)
print('After Under Sampling: ',counter)
x_train23=sparse.csr_matrix(x_train23)
x_test23=sparse.csr_matrix(x_test23)
print(x_train23.shape, x_test23.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 5000) (3351, 5000)
```

[52]:
```
#Over Sampling_TF-IDF 1000 feat

x24 = np.array(our_data.iloc[:,0].values)
y24 = np.array(our_data.relevance.values)

counter = Counter(y24)
print('Before Under Sampling: ',counter)

x24 = tfidf.fit_transform(our_data.text).toarray()
# Oversampling
oversample = RandomOverSampler(sampling_strategy=0.7)   #0.5 * majority class =
 ↪updates sample count in minority  NEW
x24, y24 = oversample.fit_resample(x24, y24)

x_train24, x_test24, y_train24, y_test24 = train_test_split(x24, y24, test_size
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train24))
print(len(x_test24))

counter = Counter(y24)
print('After Under Sampling: ',counter)
x_train24=sparse.csr_matrix(x_train24)
x_test24=sparse.csr_matrix(x_test24)
print(x_train24.shape, x_test24.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 1000) (3351, 1000)
```

[53]:
```
#Over Sampling_TF-IDF 5000 feat

x25 = np.array(our_data.iloc[:,0].values)
y25 = np.array(our_data.relevance.values)

counter = Counter(y25)
print('Before Under Sampling: ',counter)

x25 = tfidf2.fit_transform(our_data.text).toarray()
# Oversampling
oversample = RandomOverSampler(sampling_strategy=0.7)   #0.5 * majority class =
 ↪updates sample count in minority  NEW
x25, y25 = oversample.fit_resample(x25, y25)

x_train25, x_test25, y_train25, y_test25 = train_test_split(x25, y25, test_size
 ↪= 0.3, random_state = 0, shuffle = True)
```

```
print(len(x_train25))
print(len(x_test25))

counter = Counter(y25)
print('After Under Sampling: ',counter)
x_train25=sparse.csr_matrix(x_train25)
x_test25=sparse.csr_matrix(x_test25)
print(x_train25.shape, x_test25.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 5000) (3351, 5000)
```

[54]:
```
#SMOTE BoW 1000 feat

x26 = np.array(our_data.iloc[:,0].values)
y26 = np.array(our_data.relevance.values)

counter = Counter(y26)
print('Before Under Sampling: ',counter)

x26 = cv.fit_transform(our_data.text).toarray()
# SMOTE
smote = SMOTE(sampling_strategy=0.7)          # 0.7*samples in majority =␣
 ↪updated sample count in minority class
x26, y26 = smote.fit_resample(x26, y26)

x_train26, x_test26, y_train26, y_test26 = train_test_split(x26, y26, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train26))
print(len(x_test26))

counter = Counter(y26)
print('After Under Sampling: ',counter)
x_train26=sparse.csr_matrix(x_train26)
x_test26=sparse.csr_matrix(x_test26)
print(x_train26.shape, x_test26.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 1000) (3351, 1000)
```

[55]:
```
#SMOTE BoW 5000 feat

x27 = np.array(our_data.iloc[:,0].values)
y27 = np.array(our_data.relevance.values)

counter = Counter(y27)
print('Before Under Sampling: ',counter)

x27 = cv2.fit_transform(our_data.text).toarray()
# SMOTE
smote = SMOTE(sampling_strategy=0.7)         # 0.7*samples in majority =␣
 ↪updated sample count in minority class
x27, y27 = smote.fit_resample(x27, y27)

x_train27, x_test27, y_train27, y_test27 = train_test_split(x27, y27, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train27))
print(len(x_test27))

counter = Counter(y27)
print('After Under Sampling: ',counter)
x_train27=sparse.csr_matrix(x_train27)
x_test27=sparse.csr_matrix(x_test27)
print(x_train27.shape, x_test27.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 5000) (3351, 5000)
```

[56]:
```
#SMOTE Bag of n grams 1000 feat [bi,tri grams]

x28 = np.array(our_data.iloc[:,0].values)
y28 = np.array(our_data.relevance.values)

counter = Counter(y28)
print('Before Under Sampling: ',counter)

x28 = count_vect.fit_transform(our_data.text).toarray()
# SMOTE
smote = SMOTE(sampling_strategy=0.7)         # 0.7*samples in majority =␣
 ↪updated sample count in minority class
x28, y28 = smote.fit_resample(x28, y28)

x_train28, x_test28, y_train28, y_test28 = train_test_split(x28, y28, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
```

```
print(len(x_train28))
print(len(x_test28))

counter = Counter(y28)
print('After Under Sampling: ',counter)
x_train28=sparse.csr_matrix(x_train28)
x_test28=sparse.csr_matrix(x_test28)
print(x_train28.shape, x_test28.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 1000) (3351, 1000)
```

[57]:
```
#SMOTE Bag of n grams 5000 feat [bi,tri grams]

x29 = np.array(our_data.iloc[:,0].values)
y29 = np.array(our_data.relevance.values)

counter = Counter(y29)
print('Before Under Sampling: ',counter)

x29 = count_vect2.fit_transform(our_data.text).toarray()
# SMOTE
smote = SMOTE(sampling_strategy=0.7)          # 0.7*samples in majority =␣
 ↪updated sample count in minority class
x29, y29 = smote.fit_resample(x29, y29)

x_train29, x_test29, y_train29, y_test29 = train_test_split(x29, y29, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train29))
print(len(x_test29))

counter = Counter(y29)
print('After Under Sampling: ',counter)
x_train29=sparse.csr_matrix(x_train29)
x_test29=sparse.csr_matrix(x_test29)
print(x_train29.shape, x_test29.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 5000) (3351, 5000)
```

[58]:
```
#SMOTE TF-IDF 1000 feat

x30 = np.array(our_data.iloc[:,0].values)
y30 = np.array(our_data.relevance.values)

counter = Counter(y30)
print('Before Under Sampling: ',counter)

x30 = tfidf.fit_transform(our_data.text).toarray()
# SMOTE
smote = SMOTE(sampling_strategy=0.7)         # 0.7*samples in majority =␣
 ↪updated sample count in minority class
x30, y30 = smote.fit_resample(x30, y30)

x_train30, x_test30, y_train30, y_test30 = train_test_split(x30, y30, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
print(len(x_train30))
print(len(x_test30))

counter = Counter(y30)
print('After Under Sampling: ',counter)
x_train30=sparse.csr_matrix(x_train30)
x_test30=sparse.csr_matrix(x_test30)
print(x_train30.shape, x_test30.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 1000) (3351, 1000)
```

[59]:
```
#SMOTE TF-IDF 5000 feat

x31 = np.array(our_data.iloc[:,0].values)
y31 = np.array(our_data.relevance.values)

counter = Counter(y31)
print('Before Under Sampling: ',counter)

x31 = tfidf2.fit_transform(our_data.text).toarray()
# SMOTE
smote = SMOTE(sampling_strategy=0.7)         # 0.7*samples in majority =␣
 ↪updated sample count in minority class
x31, y31 = smote.fit_resample(x31, y31)

x_train31, x_test31, y_train31, y_test31 = train_test_split(x31, y31, test_size␣
 ↪= 0.3, random_state = 0, shuffle = True)
```

```
print(len(x_train31))
print(len(x_test31))

counter = Counter(y31)
print('After Under Sampling: ',counter)
x_train31=sparse.csr_matrix(x_train31)
x_test31=sparse.csr_matrix(x_test31)
print(x_train31.shape, x_test31.shape)
```

```
Before Under Sampling:  Counter({0: 6571, 1: 1420})
7819
3351
After Under Sampling:  Counter({0: 6571, 1: 4599})
(7819, 5000) (3351, 5000)
```

[60]:
```
#Normal
n_words8 = x_test8.shape[1]
n_words9 = x_test9.shape[1]
n_words10 = x_test10.shape[1]
n_words11 = x_test11.shape[1]
n_words12 = x_test12.shape[1]
n_words13 = x_test13.shape[1]
n_words14 = x_test14.shape[1]
n_words15 = x_test15.shape[1]
n_words16 = x_test16.shape[1]
n_words17 = x_test17.shape[1]
n_words18 = x_test18.shape[1]
n_words19 = x_test19.shape[1]
n_words20 = x_test20.shape[1]
n_words21 = x_test21.shape[1]
n_words22 = x_test22.shape[1]
n_words23 = x_test23.shape[1]
n_words24 = x_test24.shape[1]
n_words25 = x_test25.shape[1]
n_words26 = x_test26.shape[1]
n_words27 = x_test27.shape[1]
n_words28 = x_test28.shape[1]
n_words29 = x_test29.shape[1]
n_words30 = x_test30.shape[1]
n_words31 = x_test31.shape[1]
```

[61]:
```
#create list of model and accuracy dicts
perform_list1 = [ ]
perform_list2 = [ ]
perform_list3 = [ ]
perform_list4 = [ ]
perform_list5 = [ ]
```

[62]:
```
#1 NO SAMPLING_BoW 1000 feat
run_models(x_train8, x_test8, y_train8, y_test8, n_words8)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 67.76
Precision : 0.6776480400333611
Recall : 0.6776480400333611
F1-score : 0.6776480400333611
ROC_AOC_Score for Naive Bayes:  0.7237625305086854

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 77.36
Precision : 0.7735613010842368
Recall : 0.7735613010842368
F1-score : 0.7735613010842367
ROC_AOC_Score for Logistic Regression:  0.6703201507905223

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.61
Precision : 0.8160967472894078
Recall : 0.8160967472894078
F1-score : 0.8160967472894078

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.53
Precision : 0.8152627189324437
Recall : 0.8152627189324437
F1-score : 0.8152627189324437
ROC_AOC_Score for Random Forest:  0.7120753347961224

FOR ANN:

Model: "sequential_6"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_12 (Dense)            (None, 50)                50050

 dense_13 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4442 - accuracy: 0.8205 - 2s/epoch - 11ms/step
Epoch 2/69
175/175 - 1s - loss: 0.3588 - accuracy: 0.8478 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.3018 - accuracy: 0.8741 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 1s - loss: 0.2423 - accuracy: 0.9074 - 1s/epoch - 8ms/step
Epoch 5/69
175/175 - 1s - loss: 0.1829 - accuracy: 0.9421 - 1s/epoch - 8ms/step
Epoch 6/69
175/175 - 1s - loss: 0.1313 - accuracy: 0.9651 - 1s/epoch - 8ms/step
Epoch 7/69
175/175 - 1s - loss: 0.0910 - accuracy: 0.9848 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 1s - loss: 0.0607 - accuracy: 0.9927 - 1s/epoch - 7ms/step
Epoch 9/69
175/175 - 1s - loss: 0.0410 - accuracy: 0.9970 - 1s/epoch - 7ms/step
Epoch 10/69
175/175 - 1s - loss: 0.0301 - accuracy: 0.9980 - 1s/epoch - 8ms/step
Epoch 11/69
175/175 - 1s - loss: 0.0218 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0172 - accuracy: 0.9989 - 1s/epoch - 7ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0137 - accuracy: 0.9987 - 1s/epoch - 7ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0107 - accuracy: 0.9989 - 1s/epoch - 7ms/step
Epoch 15/69
175/175 - 1s - loss: 0.0095 - accuracy: 0.9991 - 1s/epoch - 7ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0091 - accuracy: 0.9987 - 1s/epoch - 7ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0075 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0058 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 0.0047 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 1s - loss: 0.0054 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 21/69
175/175 - 2s - loss: 0.0035 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 22/69
175/175 - 2s - loss: 0.0034 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 23/69
```

```
175/175 - 1s - loss: 0.0034 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0036 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 25/69
175/175 - 1s - loss: 0.0023 - accuracy: 0.9998 - 1s/epoch - 7ms/step
Epoch 26/69
175/175 - 337s - loss: 0.0029 - accuracy: 0.9995 - 337s/epoch - 2s/step
Epoch 27/69
175/175 - 2s - loss: 0.0033 - accuracy: 0.9993 - 2s/epoch - 10ms/step
Epoch 28/69
175/175 - 1s - loss: 0.0038 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 1s - loss: 0.0028 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 0.0016 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 1s - loss: 0.0013 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 32/69
175/175 - 1s - loss: 7.2467e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 33/69
175/175 - 1s - loss: 6.0703e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 34/69
175/175 - 1s - loss: 5.4020e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 35/69
175/175 - 1s - loss: 4.7787e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 36/69
175/175 - 1s - loss: 4.3032e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 37/69
175/175 - 1s - loss: 3.8569e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 2s - loss: 3.4639e-04 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 39/69
175/175 - 1s - loss: 3.1329e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 40/69
175/175 - 2s - loss: 2.8160e-04 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 41/69
175/175 - 1s - loss: 2.5580e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 1s - loss: 2.3240e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 43/69
175/175 - 1s - loss: 2.0990e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 44/69
175/175 - 1s - loss: 1.9032e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 45/69
175/175 - 1s - loss: 1.7192e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 1.5886e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 47/69
```

```
175/175 - 1s - loss: 1.4208e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 48/69
175/175 - 1s - loss: 1.2795e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 49/69
175/175 - 1s - loss: 1.1604e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 50/69
175/175 - 1s - loss: 1.0755e-04 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 51/69
175/175 - 1s - loss: 1.2001e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 52/69
175/175 - 1s - loss: 9.7474e-04 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 2.2436e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 2s - loss: 1.0511e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 55/69
175/175 - 2s - loss: 8.9788e-04 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 56/69
175/175 - 2s - loss: 7.7815e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 57/69
175/175 - 2s - loss: 5.8087e-05 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 58/69
175/175 - 2s - loss: 5.0545e-05 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 59/69
175/175 - 1s - loss: 4.4573e-05 - accuracy: 1.0000 - 1s/epoch - 9ms/step
Epoch 60/69
175/175 - 2s - loss: 4.0124e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 61/69
175/175 - 2s - loss: 3.6177e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 62/69
175/175 - 2s - loss: 3.2696e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 63/69
175/175 - 1s - loss: 2.9742e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 64/69
175/175 - 1s - loss: 2.7080e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 65/69
175/175 - 1s - loss: 2.4669e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 66/69
175/175 - 1s - loss: 2.2508e-05 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 67/69
175/175 - 2s - loss: 2.0535e-05 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 68/69
175/175 - 1s - loss: 1.8710e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 69/69
175/175 - 1s - loss: 1.7154e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Test Accuracy: 0.7873227596282959
```

[63]:
```
#2 NO SAMPLING_BoW 5000 feat
run_models(x_train9, x_test9, y_train9, y_test9, n_words9)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 67.26
Precision : 0.6726438698915763
Recall : 0.6726438698915763
F1-score : 0.6726438698915763
ROC_AOC_Score for Naive Bayes:  0.732863682716648

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 76.36
Precision : 0.7635529608006673
Recall : 0.7635529608006673
F1-score : 0.7635529608006673
ROC_AOC_Score for Logistic Regression:  0.6642271346196109

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.65
Precision : 0.8165137614678899
Recall : 0.8165137614678899
F1-score : 0.81651376146789

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.61
Precision : 0.8160967472894078
Recall : 0.8160967472894078
F1-score : 0.8160967472894078
ROC_AOC_Score for Random Forest:  0.7122764973529101

FOR ANN:

Model: "sequential_7"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_14 (Dense)            (None, 50)                250050

 dense_15 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4320 - accuracy: 0.8210 - 2s/epoch - 12ms/step
Epoch 2/69
175/175 - 1s - loss: 0.3099 - accuracy: 0.8579 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.2033 - accuracy: 0.9219 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 2s - loss: 0.1055 - accuracy: 0.9728 - 2s/epoch - 10ms/step
Epoch 5/69
175/175 - 2s - loss: 0.0466 - accuracy: 0.9948 - 2s/epoch - 9ms/step
Epoch 6/69
175/175 - 2s - loss: 0.0224 - accuracy: 0.9977 - 2s/epoch - 9ms/step
Epoch 7/69
175/175 - 1s - loss: 0.0129 - accuracy: 0.9982 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 2s - loss: 0.0090 - accuracy: 0.9980 - 2s/epoch - 9ms/step
Epoch 9/69
175/175 - 2s - loss: 0.0076 - accuracy: 0.9987 - 2s/epoch - 10ms/step
Epoch 10/69
175/175 - 2s - loss: 0.0062 - accuracy: 0.9989 - 2s/epoch - 10ms/step
Epoch 11/69
175/175 - 2s - loss: 0.0059 - accuracy: 0.9991 - 2s/epoch - 10ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0041 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0040 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0035 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 15/69
175/175 - 2s - loss: 0.0028 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0031 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0015 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0015 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 9.7294e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 1s - loss: 0.0011 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 1s - loss: 0.0014 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 22/69
175/175 - 2s - loss: 8.2349e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 23/69
```

```
175/175 - 1s - loss: 6.5355e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 5.5235e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 5.1851e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 26/69
175/175 - 1s - loss: 4.8070e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 27/69
175/175 - 1s - loss: 4.5452e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 4.3098e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 1s - loss: 4.0052e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 3.8235e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 2s - loss: 3.6749e-04 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 32/69
175/175 - 2s - loss: 3.5220e-04 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 33/69
175/175 - 2s - loss: 3.3779e-04 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 34/69
175/175 - 2s - loss: 3.2894e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 35/69
175/175 - 1s - loss: 3.1776e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 36/69
175/175 - 1s - loss: 3.0806e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 37/69
175/175 - 1s - loss: 2.9814e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 1s - loss: 2.8885e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 39/69
175/175 - 2s - loss: 2.8063e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 40/69
175/175 - 2s - loss: 2.7355e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 41/69
175/175 - 1s - loss: 2.6744e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 2s - loss: 2.6303e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 43/69
175/175 - 2s - loss: 2.5673e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 44/69
175/175 - 1s - loss: 2.5127e-04 - accuracy: 1.0000 - 1s/epoch - 9ms/step
Epoch 45/69
175/175 - 2s - loss: 2.5601e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 46/69
175/175 - 2s - loss: 2.4359e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 47/69
```

```
175/175 - 1s - loss: 2.4730e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 48/69
175/175 - 1s - loss: 2.3478e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 49/69
175/175 - 1s - loss: 2.3262e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 50/69
175/175 - 2s - loss: 2.3115e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 51/69
175/175 - 2s - loss: 2.1917e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 52/69
175/175 - 1s - loss: 2.1367e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 2s - loss: 2.1108e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 54/69
175/175 - 2s - loss: 2.0781e-04 - accuracy: 1.0000 - 2s/epoch - 11ms/step
Epoch 55/69
175/175 - 2s - loss: 2.0220e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 56/69
175/175 - 1s - loss: 1.9788e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 57/69
175/175 - 2s - loss: 1.9444e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 58/69
175/175 - 2s - loss: 1.9118e-04 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 59/69
175/175 - 2s - loss: 1.8769e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 60/69
175/175 - 1s - loss: 1.8661e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 61/69
175/175 - 1s - loss: 1.8202e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 62/69
175/175 - 1s - loss: 1.7920e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 63/69
175/175 - 1s - loss: 1.7751e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 64/69
175/175 - 2s - loss: 1.7465e-04 - accuracy: 1.0000 - 2s/epoch - 12ms/step
Epoch 65/69
175/175 - 2s - loss: 2.7554e-04 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 66/69
175/175 - 1s - loss: 0.0096 - accuracy: 0.9975 - 1s/epoch - 8ms/step
Epoch 67/69
175/175 - 1s - loss: 0.0088 - accuracy: 0.9971 - 1s/epoch - 8ms/step
Epoch 68/69
175/175 - 1s - loss: 0.0086 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 69/69
175/175 - 1s - loss: 0.0019 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Test Accuracy: 0.7869057655334473
```

[64]:
```
#3 NO SAMPLING_Bag of n grams 1000 feat [bi,tri grams]
run_models(x_train10, x_test10, y_train10, y_test10, n_words10)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 76.65
Precision : 0.7664720600500416
Recall : 0.7664720600500416
F1-score : 0.7664720600500416
ROC_AOC_Score for Naive Bayes:  0.6779253745053785

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 79.36
Precision : 0.7935779816513762
Recall : 0.7935779816513762
F1-score : 0.7935779816513762
ROC_AOC_Score for Logistic Regression:  0.652991682567811

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.9
Precision : 0.8190158465387823
Recall : 0.8190158465387823
F1-score : 0.8190158465387823

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372
ROC_AOC_Score for Random Forest:  0.7088642920182651

FOR ANN:

Model: "sequential_8"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_16 (Dense)            (None, 50)                50050

 dense_17 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.5131 - accuracy: 0.8090 - 2s/epoch - 11ms/step
Epoch 2/69
175/175 - 1s - loss: 0.4030 - accuracy: 0.8339 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.3581 - accuracy: 0.8493 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 1s - loss: 0.3183 - accuracy: 0.8693 - 1s/epoch - 8ms/step
Epoch 5/69
175/175 - 1s - loss: 0.2802 - accuracy: 0.8949 - 1s/epoch - 7ms/step
Epoch 6/69
175/175 - 1s - loss: 0.2454 - accuracy: 0.9161 - 1s/epoch - 7ms/step
Epoch 7/69
175/175 - 1s - loss: 0.2112 - accuracy: 0.9315 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 1s - loss: 0.1789 - accuracy: 0.9476 - 1s/epoch - 8ms/step
Epoch 9/69
175/175 - 1s - loss: 0.1533 - accuracy: 0.9574 - 1s/epoch - 8ms/step
Epoch 10/69
175/175 - 1s - loss: 0.1284 - accuracy: 0.9682 - 1s/epoch - 8ms/step
Epoch 11/69
175/175 - 1s - loss: 0.1079 - accuracy: 0.9753 - 1s/epoch - 7ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0906 - accuracy: 0.9812 - 1s/epoch - 7ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0776 - accuracy: 0.9852 - 1s/epoch - 8ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0657 - accuracy: 0.9873 - 1s/epoch - 8ms/step
Epoch 15/69
175/175 - 1s - loss: 0.0560 - accuracy: 0.9900 - 1s/epoch - 8ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0485 - accuracy: 0.9912 - 1s/epoch - 7ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0414 - accuracy: 0.9930 - 1s/epoch - 7ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0362 - accuracy: 0.9943 - 1s/epoch - 7ms/step
Epoch 19/69
175/175 - 1s - loss: 0.0318 - accuracy: 0.9957 - 1s/epoch - 7ms/step
Epoch 20/69
175/175 - 1s - loss: 0.0279 - accuracy: 0.9966 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 1s - loss: 0.0250 - accuracy: 0.9966 - 1s/epoch - 8ms/step
Epoch 22/69
175/175 - 1s - loss: 0.0219 - accuracy: 0.9970 - 1s/epoch - 8ms/step
Epoch 23/69
```

```
175/175 - 1s - loss: 0.0197 - accuracy: 0.9973 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0176 - accuracy: 0.9973 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 0.0158 - accuracy: 0.9979 - 1s/epoch - 8ms/step
Epoch 26/69
175/175 - 2s - loss: 0.0143 - accuracy: 0.9980 - 2s/epoch - 11ms/step
Epoch 27/69
175/175 - 1s - loss: 0.0130 - accuracy: 0.9982 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 0.0118 - accuracy: 0.9982 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 1s - loss: 0.0108 - accuracy: 0.9982 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 0.0099 - accuracy: 0.9982 - 1s/epoch - 7ms/step
Epoch 31/69
175/175 - 1s - loss: 0.0093 - accuracy: 0.9986 - 1s/epoch - 7ms/step
Epoch 32/69
175/175 - 1s - loss: 0.0086 - accuracy: 0.9984 - 1s/epoch - 8ms/step
Epoch 33/69
175/175 - 1s - loss: 0.0081 - accuracy: 0.9984 - 1s/epoch - 8ms/step
Epoch 34/69
175/175 - 1s - loss: 0.0075 - accuracy: 0.9986 - 1s/epoch - 8ms/step
Epoch 35/69
175/175 - 1s - loss: 0.0068 - accuracy: 0.9986 - 1s/epoch - 8ms/step
Epoch 36/69
175/175 - 1s - loss: 0.0062 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 37/69
175/175 - 1s - loss: 0.0057 - accuracy: 0.9986 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 1s - loss: 0.0053 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 39/69
175/175 - 1s - loss: 0.0050 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 40/69
175/175 - 1s - loss: 0.0047 - accuracy: 0.9989 - 1s/epoch - 8ms/step
Epoch 41/69
175/175 - 1s - loss: 0.0045 - accuracy: 0.9991 - 1s/epoch - 7ms/step
Epoch 42/69
175/175 - 1s - loss: 0.0042 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 43/69
175/175 - 1s - loss: 0.0040 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 44/69
175/175 - 1s - loss: 0.0038 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 45/69
175/175 - 1s - loss: 0.0036 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 0.0034 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 47/69
```

```
175/175 - 1s - loss: 0.0033 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 48/69
175/175 - 1s - loss: 0.0032 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 49/69
175/175 - 1s - loss: 0.0031 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 50/69
175/175 - 1s - loss: 0.0030 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 51/69
175/175 - 1s - loss: 0.0029 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 52/69
175/175 - 1s - loss: 0.0028 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 0.0027 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 1s - loss: 0.0026 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 55/69
175/175 - 1s - loss: 0.0026 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 56/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 57/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 58/69
175/175 - 1s - loss: 0.0024 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 59/69
175/175 - 1s - loss: 0.0024 - accuracy: 0.9995 - 1s/epoch - 7ms/step
Epoch 60/69
175/175 - 1s - loss: 0.0024 - accuracy: 0.9993 - 1s/epoch - 7ms/step
Epoch 61/69
175/175 - 2s - loss: 0.0023 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 62/69
175/175 - 2s - loss: 0.0023 - accuracy: 0.9995 - 2s/epoch - 11ms/step
Epoch 63/69
175/175 - 1s - loss: 0.0022 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 64/69
175/175 - 1s - loss: 0.0022 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 65/69
175/175 - 1s - loss: 0.0022 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Epoch 66/69
175/175 - 2s - loss: 0.0021 - accuracy: 0.9993 - 2s/epoch - 9ms/step
Epoch 67/69
175/175 - 2s - loss: 0.0021 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 68/69
175/175 - 1s - loss: 0.0021 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 69/69
175/175 - 1s - loss: 0.0020 - accuracy: 0.9993 - 1s/epoch - 8ms/step
Test Accuracy: 0.7710592150688171
```

[65]:
```
#4 NO SAMPLING_Bag of n grams 5000 feat [bi,tri grams]
run_models(x_train11, x_test11, y_train11, y_test11, n_words11)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 75.65
Precision : 0.7564637197664721
Recall : 0.7564637197664721
F1-score : 0.7564637197664721
ROC_AOC_Score for Naive Bayes:  0.6900736161934696

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 78.15
Precision : 0.7814845704753962
Recall : 0.7814845704753962
F1-score : 0.7814845704753962
ROC_AOC_Score for Logistic Regression:  0.641203905576854

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.65
Precision : 0.8165137614678899
Recall : 0.8165137614678899
F1-score : 0.81651376146789

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.44
Precision : 0.8144286905754796
Recall : 0.8144286905754796
F1-score : 0.8144286905754796
ROC_AOC_Score for Random Forest:  0.6943974483750601

FOR ANN:

Model: "sequential_9"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_18 (Dense)            (None, 50)                250050

 dense_19 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4934 - accuracy: 0.8210 - 2s/epoch - 12ms/step
Epoch 2/69
175/175 - 1s - loss: 0.3353 - accuracy: 0.8579 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.2299 - accuracy: 0.9161 - 1s/epoch - 8ms/step
Epoch 4/69
175/175 - 1s - loss: 0.1372 - accuracy: 0.9591 - 1s/epoch - 8ms/step
Epoch 5/69
175/175 - 1s - loss: 0.0769 - accuracy: 0.9828 - 1s/epoch - 8ms/step
Epoch 6/69
175/175 - 1s - loss: 0.0439 - accuracy: 0.9934 - 1s/epoch - 8ms/step
Epoch 7/69
175/175 - 1s - loss: 0.0269 - accuracy: 0.9962 - 1s/epoch - 9ms/step
Epoch 8/69
175/175 - 1s - loss: 0.0188 - accuracy: 0.9968 - 1s/epoch - 8ms/step
Epoch 9/69
175/175 - 1s - loss: 0.0143 - accuracy: 0.9973 - 1s/epoch - 8ms/step
Epoch 10/69
175/175 - 1s - loss: 0.0112 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 11/69
175/175 - 1s - loss: 0.0090 - accuracy: 0.9986 - 1s/epoch - 8ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0076 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0062 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 14/69
175/175 - 2s - loss: 0.0052 - accuracy: 0.9993 - 2s/epoch - 9ms/step
Epoch 15/69
175/175 - 2s - loss: 0.0044 - accuracy: 0.9991 - 2s/epoch - 9ms/step
Epoch 16/69
175/175 - 1s - loss: 0.0037 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0032 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0028 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 1s - loss: 0.0028 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 2s - loss: 0.0022 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 22/69
175/175 - 1s - loss: 0.0017 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 23/69
```

```
175/175 - 1s - loss: 0.0015 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0014 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 0.0012 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 26/69
175/175 - 1s - loss: 0.0011 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 27/69
175/175 - 1s - loss: 0.0011 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 9.8977e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 1s - loss: 9.2553e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 8.7507e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 1s - loss: 8.3560e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 32/69
175/175 - 1s - loss: 7.8747e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 33/69
175/175 - 1s - loss: 7.9064e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 34/69
175/175 - 1s - loss: 7.9772e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 35/69
175/175 - 1s - loss: 0.0016 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 36/69
175/175 - 1s - loss: 6.6038e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 37/69
175/175 - 1s - loss: 6.1409e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 1s - loss: 5.8776e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 39/69
175/175 - 1s - loss: 5.6425e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 40/69
175/175 - 1s - loss: 5.4287e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 41/69
175/175 - 1s - loss: 5.2674e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 1s - loss: 5.0745e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 43/69
175/175 - 1s - loss: 4.8893e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 44/69
175/175 - 1s - loss: 4.7225e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 45/69
175/175 - 1s - loss: 4.5628e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 4.4068e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 47/69
```

```
175/175 - 1s - loss: 4.2870e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 48/69
175/175 - 1s - loss: 4.1558e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 49/69
175/175 - 1s - loss: 4.0291e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 50/69
175/175 - 1s - loss: 3.9029e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 51/69
175/175 - 1s - loss: 3.7873e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 52/69
175/175 - 1s - loss: 3.6752e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 3.5816e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 1s - loss: 3.4766e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 55/69
175/175 - 2s - loss: 3.3899e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 56/69
175/175 - 1s - loss: 3.2887e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 57/69
175/175 - 1s - loss: 3.2112e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 58/69
175/175 - 2s - loss: 3.1393e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 59/69
175/175 - 2s - loss: 4.7419e-04 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 60/69
175/175 - 2s - loss: 0.0025 - accuracy: 0.9998 - 2s/epoch - 10ms/step
Epoch 61/69
175/175 - 1s - loss: 3.0982e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 62/69
175/175 - 1s - loss: 2.8659e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 63/69
175/175 - 2s - loss: 2.8048e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 64/69
175/175 - 1s - loss: 2.7374e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 65/69
175/175 - 2s - loss: 2.6769e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 66/69
175/175 - 1s - loss: 2.6176e-04 - accuracy: 1.0000 - 1s/epoch - 9ms/step
Epoch 67/69
175/175 - 1s - loss: 2.5562e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 68/69
175/175 - 1s - loss: 2.4972e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 69/69
175/175 - 1s - loss: 2.4391e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Test Accuracy: 0.7673060894012451
```

[66]:
```
#5 NO SAMPLING_Tf-Idf 1000 feat
run_models(x_train12, x_test12, y_train12, y_test12, n_words12)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 82.19
Precision : 0.8219349457881568
Recall : 0.8219349457881568
F1-score : 0.8219349457881568
ROC_AOC_Score for Naive Bayes:  0.7280538045885993

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 81.65
Precision : 0.8165137614678899
Recall : 0.8165137614678899
F1-score : 0.81651376146789
ROC_AOC_Score for Logistic Regression:  0.7340770533987752

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.65
Precision : 0.8165137614678899
Recall : 0.8165137614678899
F1-score : 0.81651376146789

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.78
Precision : 0.8177648040033361
Recall : 0.8177648040033361
F1-score : 0.8177648040033361
ROC_AOC_Score for Random Forest:  0.7247009015105796

FOR ANN:

Model: "sequential_10"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_20 (Dense)            (None, 50)                50050

 dense_21 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4819 - accuracy: 0.8169 - 2s/epoch - 12ms/step
Epoch 2/69
175/175 - 1s - loss: 0.4070 - accuracy: 0.8260 - 1s/epoch - 8ms/step
Epoch 3/69
175/175 - 1s - loss: 0.3822 - accuracy: 0.8316 - 1s/epoch - 7ms/step
Epoch 4/69
175/175 - 1s - loss: 0.3662 - accuracy: 0.8393 - 1s/epoch - 7ms/step
Epoch 5/69
175/175 - 1s - loss: 0.3498 - accuracy: 0.8491 - 1s/epoch - 7ms/step
Epoch 6/69
175/175 - 1s - loss: 0.3334 - accuracy: 0.8575 - 1s/epoch - 8ms/step
Epoch 7/69
175/175 - 1s - loss: 0.3157 - accuracy: 0.8697 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 1s - loss: 0.2975 - accuracy: 0.8777 - 1s/epoch - 8ms/step
Epoch 9/69
175/175 - 1s - loss: 0.2795 - accuracy: 0.8881 - 1s/epoch - 8ms/step
Epoch 10/69
175/175 - 1s - loss: 0.2613 - accuracy: 0.8993 - 1s/epoch - 7ms/step
Epoch 11/69
175/175 - 1s - loss: 0.2432 - accuracy: 0.9094 - 1s/epoch - 7ms/step
Epoch 12/69
175/175 - 1s - loss: 0.2262 - accuracy: 0.9217 - 1s/epoch - 8ms/step
Epoch 13/69
175/175 - 1s - loss: 0.2094 - accuracy: 0.9301 - 1s/epoch - 8ms/step
Epoch 14/69
175/175 - 1s - loss: 0.1931 - accuracy: 0.9378 - 1s/epoch - 8ms/step
Epoch 15/69
175/175 - 1s - loss: 0.1777 - accuracy: 0.9508 - 1s/epoch - 8ms/step
Epoch 16/69
175/175 - 1s - loss: 0.1622 - accuracy: 0.9566 - 1s/epoch - 8ms/step
Epoch 17/69
175/175 - 1s - loss: 0.1488 - accuracy: 0.9625 - 1s/epoch - 7ms/step
Epoch 18/69
175/175 - 1s - loss: 0.1359 - accuracy: 0.9676 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 0.1233 - accuracy: 0.9726 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 1s - loss: 0.1119 - accuracy: 0.9769 - 1s/epoch - 8ms/step
Epoch 21/69
175/175 - 1s - loss: 0.1014 - accuracy: 0.9805 - 1s/epoch - 8ms/step
Epoch 22/69
175/175 - 1s - loss: 0.0915 - accuracy: 0.9839 - 1s/epoch - 7ms/step
Epoch 23/69
```

```
175/175 - 1s - loss: 0.0828 - accuracy: 0.9862 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0747 - accuracy: 0.9880 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 0.0671 - accuracy: 0.9907 - 1s/epoch - 8ms/step
Epoch 26/69
175/175 - 1s - loss: 0.0604 - accuracy: 0.9921 - 1s/epoch - 7ms/step
Epoch 27/69
175/175 - 1s - loss: 0.0541 - accuracy: 0.9939 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 0.0485 - accuracy: 0.9950 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 1s - loss: 0.0436 - accuracy: 0.9966 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 0.0390 - accuracy: 0.9970 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 1s - loss: 0.0351 - accuracy: 0.9979 - 1s/epoch - 8ms/step
Epoch 32/69
175/175 - 1s - loss: 0.0314 - accuracy: 0.9980 - 1s/epoch - 8ms/step
Epoch 33/69
175/175 - 2s - loss: 0.0284 - accuracy: 0.9989 - 2s/epoch - 9ms/step
Epoch 34/69
175/175 - 1s - loss: 0.0253 - accuracy: 0.9987 - 1s/epoch - 8ms/step
Epoch 35/69
175/175 - 2s - loss: 0.0228 - accuracy: 0.9987 - 2s/epoch - 9ms/step
Epoch 36/69
175/175 - 1s - loss: 0.0204 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 37/69
175/175 - 1s - loss: 0.0184 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 2s - loss: 0.0166 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 39/69
175/175 - 2s - loss: 0.0150 - accuracy: 0.9995 - 2s/epoch - 10ms/step
Epoch 40/69
175/175 - 2s - loss: 0.0137 - accuracy: 0.9991 - 2s/epoch - 10ms/step
Epoch 41/69
175/175 - 1s - loss: 0.0122 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 1s - loss: 0.0112 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 43/69
175/175 - 2s - loss: 0.0100 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 44/69
175/175 - 1s - loss: 0.0091 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 45/69
175/175 - 1s - loss: 0.0083 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 0.0074 - accuracy: 0.9996 - 1s/epoch - 9ms/step
Epoch 47/69
```

```
175/175 - 1s - loss: 0.0068 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 48/69
175/175 - 1s - loss: 0.0063 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 49/69
175/175 - 1s - loss: 0.0057 - accuracy: 0.9998 - 1s/epoch - 7ms/step
Epoch 50/69
175/175 - 1s - loss: 0.0052 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 51/69
175/175 - 1s - loss: 0.0052 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 52/69
175/175 - 2s - loss: 0.0043 - accuracy: 0.9998 - 2s/epoch - 9ms/step
Epoch 53/69
175/175 - 1s - loss: 0.0039 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 2s - loss: 0.0035 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 55/69
175/175 - 1s - loss: 0.0032 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 56/69
175/175 - 2s - loss: 0.0030 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 57/69
175/175 - 1s - loss: 0.0028 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 58/69
175/175 - 1s - loss: 0.0025 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 59/69
175/175 - 2s - loss: 0.0023 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 60/69
175/175 - 2s - loss: 0.0022 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 61/69
175/175 - 1s - loss: 0.0020 - accuracy: 1.0000 - 1s/epoch - 9ms/step
Epoch 62/69
175/175 - 2s - loss: 0.0018 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 63/69
175/175 - 1s - loss: 0.0016 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 64/69
175/175 - 1s - loss: 0.0015 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 65/69
175/175 - 1s - loss: 0.0014 - accuracy: 1.0000 - 1s/epoch - 9ms/step
Epoch 66/69
175/175 - 1s - loss: 0.0014 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 67/69
175/175 - 1s - loss: 0.0011 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 68/69
175/175 - 1s - loss: 0.0011 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Epoch 69/69
175/175 - 1s - loss: 0.0010 - accuracy: 1.0000 - 1s/epoch - 7ms/step
Test Accuracy: 0.7827355861663818
```

[67]:
```
#6 NO SAMPLING_Tf-Idf 5000 feat
run_models(x_train13, x_test13, y_train13, y_test13, n_words13)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 81.78
Precision : 0.8177648040033361
Recall : 0.8177648040033361
F1-score : 0.8177648040033361
ROC_AOC_Score for Naive Bayes:  0.7198468373873985

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372
ROC_AOC_Score for Logistic Regression:  0.7449502965694226

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 81.69
Precision : 0.816930775646372
Recall : 0.816930775646372
F1-score : 0.816930775646372

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.48
Precision : 0.8148457047539617
Recall : 0.8148457047539617
F1-score : 0.8148457047539617
ROC_AOC_Score for Random Forest:  0.7292817101375464

FOR ANN:

Model: "sequential_11"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_22 (Dense)            (None, 50)                250050

 dense_23 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
175/175 - 2s - loss: 0.4801 - accuracy: 0.8234 - 2s/epoch - 12ms/step
Epoch 2/69
175/175 - 2s - loss: 0.3901 - accuracy: 0.8255 - 2s/epoch - 9ms/step
Epoch 3/69
175/175 - 2s - loss: 0.3457 - accuracy: 0.8452 - 2s/epoch - 9ms/step
Epoch 4/69
175/175 - 2s - loss: 0.2974 - accuracy: 0.8686 - 2s/epoch - 10ms/step
Epoch 5/69
175/175 - 2s - loss: 0.2440 - accuracy: 0.9020 - 2s/epoch - 10ms/step
Epoch 6/69
175/175 - 2s - loss: 0.1909 - accuracy: 0.9347 - 2s/epoch - 9ms/step
Epoch 7/69
175/175 - 1s - loss: 0.1465 - accuracy: 0.9582 - 1s/epoch - 8ms/step
Epoch 8/69
175/175 - 1s - loss: 0.1088 - accuracy: 0.9751 - 1s/epoch - 8ms/step
Epoch 9/69
175/175 - 1s - loss: 0.0814 - accuracy: 0.9855 - 1s/epoch - 8ms/step
Epoch 10/69
175/175 - 1s - loss: 0.0602 - accuracy: 0.9921 - 1s/epoch - 8ms/step
Epoch 11/69
175/175 - 1s - loss: 0.0454 - accuracy: 0.9962 - 1s/epoch - 8ms/step
Epoch 12/69
175/175 - 1s - loss: 0.0346 - accuracy: 0.9970 - 1s/epoch - 8ms/step
Epoch 13/69
175/175 - 1s - loss: 0.0269 - accuracy: 0.9977 - 1s/epoch - 8ms/step
Epoch 14/69
175/175 - 1s - loss: 0.0213 - accuracy: 0.9980 - 1s/epoch - 8ms/step
Epoch 15/69
175/175 - 2s - loss: 0.0174 - accuracy: 0.9982 - 2s/epoch - 9ms/step
Epoch 16/69
175/175 - 2s - loss: 0.0143 - accuracy: 0.9987 - 2s/epoch - 9ms/step
Epoch 17/69
175/175 - 1s - loss: 0.0120 - accuracy: 0.9991 - 1s/epoch - 8ms/step
Epoch 18/69
175/175 - 1s - loss: 0.0100 - accuracy: 0.9995 - 1s/epoch - 8ms/step
Epoch 19/69
175/175 - 1s - loss: 0.0086 - accuracy: 0.9996 - 1s/epoch - 8ms/step
Epoch 20/69
175/175 - 2s - loss: 0.0073 - accuracy: 0.9993 - 2s/epoch - 9ms/step
Epoch 21/69
175/175 - 1s - loss: 0.0064 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 22/69
175/175 - 1s - loss: 0.0056 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 23/69
```

```
175/175 - 1s - loss: 0.0050 - accuracy: 0.9998 - 1s/epoch - 8ms/step
Epoch 24/69
175/175 - 1s - loss: 0.0044 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 25/69
175/175 - 1s - loss: 0.0038 - accuracy: 0.9998 - 1s/epoch - 9ms/step
Epoch 26/69
175/175 - 1s - loss: 0.0034 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 27/69
175/175 - 1s - loss: 0.0030 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 28/69
175/175 - 1s - loss: 0.0028 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 29/69
175/175 - 1s - loss: 0.0024 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 30/69
175/175 - 1s - loss: 0.0021 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 31/69
175/175 - 1s - loss: 0.0018 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 32/69
175/175 - 2s - loss: 0.0016 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 33/69
175/175 - 2s - loss: 0.0014 - accuracy: 1.0000 - 2s/epoch - 10ms/step
Epoch 34/69
175/175 - 2s - loss: 0.0013 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 35/69
175/175 - 1s - loss: 0.0011 - accuracy: 1.0000 - 1s/epoch - 9ms/step
Epoch 36/69
175/175 - 1s - loss: 9.9842e-04 - accuracy: 1.0000 - 1s/epoch - 9ms/step
Epoch 37/69
175/175 - 1s - loss: 9.0775e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 38/69
175/175 - 1s - loss: 7.8644e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 39/69
175/175 - 1s - loss: 7.1508e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 40/69
175/175 - 1s - loss: 6.5237e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 41/69
175/175 - 1s - loss: 5.7277e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 42/69
175/175 - 1s - loss: 4.9747e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 43/69
175/175 - 1s - loss: 4.5272e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 44/69
175/175 - 1s - loss: 4.1702e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 45/69
175/175 - 1s - loss: 3.6329e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 46/69
175/175 - 1s - loss: 3.2632e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 47/69
```

```
175/175 - 1s - loss: 3.0125e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 48/69
175/175 - 1s - loss: 2.6221e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 49/69
175/175 - 1s - loss: 2.3890e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 50/69
175/175 - 1s - loss: 2.2112e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 51/69
175/175 - 1s - loss: 1.8917e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 52/69
175/175 - 1s - loss: 1.7545e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 53/69
175/175 - 1s - loss: 1.5991e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 54/69
175/175 - 1s - loss: 1.4589e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 55/69
175/175 - 1s - loss: 1.3016e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 56/69
175/175 - 1s - loss: 1.1584e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 57/69
175/175 - 1s - loss: 1.0648e-04 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 58/69
175/175 - 1s - loss: 9.6475e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 59/69
175/175 - 1s - loss: 8.6615e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 60/69
175/175 - 1s - loss: 7.8379e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 61/69
175/175 - 1s - loss: 6.9738e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 62/69
175/175 - 1s - loss: 6.4302e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 63/69
175/175 - 1s - loss: 6.1945e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 64/69
175/175 - 2s - loss: 5.3740e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 65/69
175/175 - 2s - loss: 4.7847e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 66/69
175/175 - 2s - loss: 4.3511e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 67/69
175/175 - 1s - loss: 3.9259e-05 - accuracy: 1.0000 - 1s/epoch - 8ms/step
Epoch 68/69
175/175 - 2s - loss: 3.5712e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 69/69
175/175 - 2s - loss: 3.2046e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Test Accuracy: 0.7693911790847778
```

[68]:
```
#7 Under Sampling_BoW 1000 feat
run_models(x_train14, x_test14, y_train14, y_test14, n_words14)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 66.96
Precision : 0.6695652173913044
Recall : 0.6695652173913044
F1-score : 0.6695652173913044
ROC_AOC_Score for Naive Bayes:  0.7299733629127382

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 61.93
Precision : 0.6193236714975845
Recall : 0.6193236714975845
F1-score : 0.6193236714975845
ROC_AOC_Score for Logistic Regression:  0.6444300366357593

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 68.02
Precision : 0.6801932367149759
Recall : 0.6801932367149759
F1-score : 0.6801932367149759

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 67.54
Precision : 0.6753623188405797
Recall : 0.6753623188405797
F1-score : 0.6753623188405797
ROC_AOC_Score for Random Forest:  0.731695790402837

FOR ANN:

Model: "sequential_12"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_24 (Dense)            (None, 50)                50050

 dense_25 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
76/76 - 2s - loss: 0.6295 - accuracy: 0.6486 - 2s/epoch - 21ms/step
Epoch 2/69
76/76 - 1s - loss: 0.4986 - accuracy: 0.7671 - 661ms/epoch - 9ms/step
Epoch 3/69
76/76 - 1s - loss: 0.4154 - accuracy: 0.8193 - 606ms/epoch - 8ms/step
Epoch 4/69
76/76 - 1s - loss: 0.3423 - accuracy: 0.8753 - 596ms/epoch - 8ms/step
Epoch 5/69
76/76 - 1s - loss: 0.2759 - accuracy: 0.9159 - 598ms/epoch - 8ms/step
Epoch 6/69
76/76 - 1s - loss: 0.2174 - accuracy: 0.9507 - 587ms/epoch - 8ms/step
Epoch 7/69
76/76 - 1s - loss: 0.1686 - accuracy: 0.9702 - 590ms/epoch - 8ms/step
Epoch 8/69
76/76 - 1s - loss: 0.1277 - accuracy: 0.9863 - 594ms/epoch - 8ms/step
Epoch 9/69
76/76 - 1s - loss: 0.0988 - accuracy: 0.9925 - 595ms/epoch - 8ms/step
Epoch 10/69
76/76 - 1s - loss: 0.0749 - accuracy: 0.9946 - 584ms/epoch - 8ms/step
Epoch 11/69
76/76 - 1s - loss: 0.0574 - accuracy: 0.9963 - 568ms/epoch - 7ms/step
Epoch 12/69
76/76 - 1s - loss: 0.0463 - accuracy: 0.9971 - 564ms/epoch - 7ms/step
Epoch 13/69
76/76 - 1s - loss: 0.0388 - accuracy: 0.9959 - 568ms/epoch - 7ms/step
Epoch 14/69
76/76 - 1s - loss: 0.0309 - accuracy: 0.9975 - 568ms/epoch - 7ms/step
Epoch 15/69
76/76 - 1s - loss: 0.0256 - accuracy: 0.9983 - 573ms/epoch - 8ms/step
Epoch 16/69
76/76 - 1s - loss: 0.0221 - accuracy: 0.9979 - 580ms/epoch - 8ms/step
Epoch 17/69
76/76 - 1s - loss: 0.0203 - accuracy: 0.9983 - 569ms/epoch - 7ms/step
Epoch 18/69
76/76 - 1s - loss: 0.0168 - accuracy: 0.9983 - 577ms/epoch - 8ms/step
Epoch 19/69
76/76 - 1s - loss: 0.0159 - accuracy: 0.9988 - 647ms/epoch - 9ms/step
Epoch 20/69
76/76 - 1s - loss: 0.0133 - accuracy: 0.9996 - 650ms/epoch - 9ms/step
Epoch 21/69
76/76 - 1s - loss: 0.0117 - accuracy: 0.9996 - 582ms/epoch - 8ms/step
Epoch 22/69
76/76 - 1s - loss: 0.0112 - accuracy: 0.9983 - 577ms/epoch - 8ms/step
Epoch 23/69
```

```
76/76 - 1s - loss: 0.0092 - accuracy: 0.9996 - 571ms/epoch - 8ms/step
Epoch 24/69
76/76 - 1s - loss: 0.0084 - accuracy: 0.9992 - 579ms/epoch - 8ms/step
Epoch 25/69
76/76 - 1s - loss: 0.0086 - accuracy: 0.9992 - 575ms/epoch - 8ms/step
Epoch 26/69
76/76 - 1s - loss: 0.0082 - accuracy: 0.9988 - 572ms/epoch - 8ms/step
Epoch 27/69
76/76 - 1s - loss: 0.0072 - accuracy: 0.9992 - 575ms/epoch - 8ms/step
Epoch 28/69
76/76 - 1s - loss: 0.0072 - accuracy: 0.9988 - 566ms/epoch - 7ms/step
Epoch 29/69
76/76 - 1s - loss: 0.0059 - accuracy: 0.9992 - 566ms/epoch - 7ms/step
Epoch 30/69
76/76 - 1s - loss: 0.0054 - accuracy: 0.9996 - 595ms/epoch - 8ms/step
Epoch 31/69
76/76 - 1s - loss: 0.0045 - accuracy: 0.9996 - 600ms/epoch - 8ms/step
Epoch 32/69
76/76 - 1s - loss: 0.0034 - accuracy: 1.0000 - 584ms/epoch - 8ms/step
Epoch 33/69
76/76 - 1s - loss: 0.0031 - accuracy: 1.0000 - 571ms/epoch - 8ms/step
Epoch 34/69
76/76 - 1s - loss: 0.0029 - accuracy: 1.0000 - 576ms/epoch - 8ms/step
Epoch 35/69
76/76 - 1s - loss: 0.0026 - accuracy: 1.0000 - 576ms/epoch - 8ms/step
Epoch 36/69
76/76 - 1s - loss: 0.0024 - accuracy: 1.0000 - 590ms/epoch - 8ms/step
Epoch 37/69
76/76 - 1s - loss: 0.0023 - accuracy: 1.0000 - 567ms/epoch - 7ms/step
Epoch 38/69
76/76 - 1s - loss: 0.0021 - accuracy: 1.0000 - 564ms/epoch - 7ms/step
Epoch 39/69
76/76 - 1s - loss: 0.0020 - accuracy: 1.0000 - 572ms/epoch - 8ms/step
Epoch 40/69
76/76 - 1s - loss: 0.0018 - accuracy: 1.0000 - 566ms/epoch - 7ms/step
Epoch 41/69
76/76 - 1s - loss: 0.0017 - accuracy: 1.0000 - 575ms/epoch - 8ms/step
Epoch 42/69
76/76 - 1s - loss: 0.0016 - accuracy: 1.0000 - 639ms/epoch - 8ms/step
Epoch 43/69
76/76 - 1s - loss: 0.0015 - accuracy: 1.0000 - 612ms/epoch - 8ms/step
Epoch 44/69
76/76 - 1s - loss: 0.0015 - accuracy: 1.0000 - 574ms/epoch - 8ms/step
Epoch 45/69
76/76 - 1s - loss: 0.0014 - accuracy: 1.0000 - 625ms/epoch - 8ms/step
Epoch 46/69
76/76 - 1s - loss: 0.0013 - accuracy: 1.0000 - 607ms/epoch - 8ms/step
Epoch 47/69
```

```
76/76 - 1s - loss: 0.0012 - accuracy: 1.0000 - 622ms/epoch - 8ms/step
Epoch 48/69
76/76 - 1s - loss: 0.0011 - accuracy: 1.0000 - 598ms/epoch - 8ms/step
Epoch 49/69
76/76 - 1s - loss: 0.0010 - accuracy: 1.0000 - 572ms/epoch - 8ms/step
Epoch 50/69
76/76 - 1s - loss: 9.9290e-04 - accuracy: 1.0000 - 609ms/epoch - 8ms/step
Epoch 51/69
76/76 - 1s - loss: 9.3007e-04 - accuracy: 1.0000 - 576ms/epoch - 8ms/step
Epoch 52/69
76/76 - 1s - loss: 8.8333e-04 - accuracy: 1.0000 - 573ms/epoch - 8ms/step
Epoch 53/69
76/76 - 1s - loss: 8.3152e-04 - accuracy: 1.0000 - 595ms/epoch - 8ms/step
Epoch 54/69
76/76 - 1s - loss: 7.8426e-04 - accuracy: 1.0000 - 578ms/epoch - 8ms/step
Epoch 55/69
76/76 - 1s - loss: 7.4276e-04 - accuracy: 1.0000 - 569ms/epoch - 7ms/step
Epoch 56/69
76/76 - 1s - loss: 7.0130e-04 - accuracy: 1.0000 - 570ms/epoch - 7ms/step
Epoch 57/69
76/76 - 1s - loss: 6.6234e-04 - accuracy: 1.0000 - 568ms/epoch - 7ms/step
Epoch 58/69
76/76 - 1s - loss: 6.3680e-04 - accuracy: 1.0000 - 567ms/epoch - 7ms/step
Epoch 59/69
76/76 - 1s - loss: 6.0714e-04 - accuracy: 1.0000 - 568ms/epoch - 7ms/step
Epoch 60/69
76/76 - 1s - loss: 5.9922e-04 - accuracy: 1.0000 - 570ms/epoch - 8ms/step
Epoch 61/69
76/76 - 1s - loss: 5.3798e-04 - accuracy: 1.0000 - 568ms/epoch - 7ms/step
Epoch 62/69
76/76 - 1s - loss: 5.0990e-04 - accuracy: 1.0000 - 562ms/epoch - 7ms/step
Epoch 63/69
76/76 - 1s - loss: 4.8811e-04 - accuracy: 1.0000 - 575ms/epoch - 8ms/step
Epoch 64/69
76/76 - 1s - loss: 4.6329e-04 - accuracy: 1.0000 - 596ms/epoch - 8ms/step
Epoch 65/69
76/76 - 1s - loss: 4.3967e-04 - accuracy: 1.0000 - 603ms/epoch - 8ms/step
Epoch 66/69
76/76 - 1s - loss: 4.1936e-04 - accuracy: 1.0000 - 576ms/epoch - 8ms/step
Epoch 67/69
76/76 - 1s - loss: 3.9641e-04 - accuracy: 1.0000 - 618ms/epoch - 8ms/step
Epoch 68/69
76/76 - 1s - loss: 3.8702e-04 - accuracy: 1.0000 - 576ms/epoch - 8ms/step
Epoch 69/69
76/76 - 1s - loss: 3.5688e-04 - accuracy: 1.0000 - 581ms/epoch - 8ms/step
Test Accuracy: 0.6396135091781616
```

[69]:
```
#8 Under Sampling_BoW 5000 feat
run_models(x_train15, x_test15, y_train15, y_test15, n_words15)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 68.5
Precision : 0.6850241545893719
Recall : 0.6850241545893719
F1-score : 0.6850241545893719
ROC_AOC_Score for Naive Bayes:  0.7411808587921918

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 63.96
Precision : 0.6396135265700483
Recall : 0.6396135265700483
F1-score : 0.6396135265700483
ROC_AOC_Score for Logistic Regression:  0.6671184295835709

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 69.47
Precision : 0.6946859903381642
Recall : 0.6946859903381642
F1-score : 0.6946859903381642

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 68.89
Precision : 0.6888888888888889
Recall : 0.6888888888888889
F1-score : 0.6888888888888889
ROC_AOC_Score for Random Forest:  0.735788996773866

FOR ANN:

Model: "sequential_13"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_26 (Dense)            (None, 50)                250050

 dense_27 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
76/76 - 2s - loss: 0.6242 - accuracy: 0.6540 - 2s/epoch - 32ms/step
Epoch 2/69
76/76 - 1s - loss: 0.3983 - accuracy: 0.8396 - 610ms/epoch - 8ms/step
Epoch 3/69
76/76 - 1s - loss: 0.2378 - accuracy: 0.9329 - 624ms/epoch - 8ms/step
Epoch 4/69
76/76 - 1s - loss: 0.1271 - accuracy: 0.9814 - 642ms/epoch - 8ms/step
Epoch 5/69
76/76 - 1s - loss: 0.0651 - accuracy: 0.9963 - 611ms/epoch - 8ms/step
Epoch 6/69
76/76 - 1s - loss: 0.0374 - accuracy: 0.9975 - 611ms/epoch - 8ms/step
Epoch 7/69
76/76 - 1s - loss: 0.0250 - accuracy: 0.9983 - 619ms/epoch - 8ms/step
Epoch 8/69
76/76 - 1s - loss: 0.0193 - accuracy: 0.9983 - 625ms/epoch - 8ms/step
Epoch 9/69
76/76 - 1s - loss: 0.0134 - accuracy: 0.9988 - 625ms/epoch - 8ms/step
Epoch 10/69
76/76 - 1s - loss: 0.0129 - accuracy: 0.9979 - 612ms/epoch - 8ms/step
Epoch 11/69
76/76 - 1s - loss: 0.0094 - accuracy: 0.9983 - 607ms/epoch - 8ms/step
Epoch 12/69
76/76 - 1s - loss: 0.0062 - accuracy: 0.9992 - 605ms/epoch - 8ms/step
Epoch 13/69
76/76 - 1s - loss: 0.0048 - accuracy: 1.0000 - 617ms/epoch - 8ms/step
Epoch 14/69
76/76 - 1s - loss: 0.0038 - accuracy: 0.9996 - 605ms/epoch - 8ms/step
Epoch 15/69
76/76 - 1s - loss: 0.0034 - accuracy: 1.0000 - 606ms/epoch - 8ms/step
Epoch 16/69
76/76 - 1s - loss: 0.0028 - accuracy: 1.0000 - 614ms/epoch - 8ms/step
Epoch 17/69
76/76 - 1s - loss: 0.0024 - accuracy: 1.0000 - 621ms/epoch - 8ms/step
Epoch 18/69
76/76 - 1s - loss: 0.0021 - accuracy: 1.0000 - 618ms/epoch - 8ms/step
Epoch 19/69
76/76 - 1s - loss: 0.0019 - accuracy: 1.0000 - 637ms/epoch - 8ms/step
Epoch 20/69
76/76 - 1s - loss: 0.0016 - accuracy: 1.0000 - 662ms/epoch - 9ms/step
Epoch 21/69
76/76 - 1s - loss: 0.0014 - accuracy: 1.0000 - 657ms/epoch - 9ms/step
Epoch 22/69
76/76 - 1s - loss: 0.0013 - accuracy: 1.0000 - 619ms/epoch - 8ms/step
Epoch 23/69
```

```
76/76 - 1s - loss: 0.0012 - accuracy: 1.0000 - 620ms/epoch - 8ms/step
Epoch 24/69
76/76 - 1s - loss: 0.0011 - accuracy: 1.0000 - 616ms/epoch - 8ms/step
Epoch 25/69
76/76 - 1s - loss: 9.5039e-04 - accuracy: 1.0000 - 607ms/epoch - 8ms/step
Epoch 26/69
76/76 - 1s - loss: 8.7266e-04 - accuracy: 1.0000 - 612ms/epoch - 8ms/step
Epoch 27/69
76/76 - 1s - loss: 8.0823e-04 - accuracy: 1.0000 - 607ms/epoch - 8ms/step
Epoch 28/69
76/76 - 1s - loss: 7.3037e-04 - accuracy: 1.0000 - 614ms/epoch - 8ms/step
Epoch 29/69
76/76 - 1s - loss: 6.7300e-04 - accuracy: 1.0000 - 609ms/epoch - 8ms/step
Epoch 30/69
76/76 - 1s - loss: 6.0775e-04 - accuracy: 1.0000 - 619ms/epoch - 8ms/step
Epoch 31/69
76/76 - 1s - loss: 5.7133e-04 - accuracy: 1.0000 - 623ms/epoch - 8ms/step
Epoch 32/69
76/76 - 1s - loss: 5.3227e-04 - accuracy: 1.0000 - 623ms/epoch - 8ms/step
Epoch 33/69
76/76 - 1s - loss: 4.8499e-04 - accuracy: 1.0000 - 628ms/epoch - 8ms/step
Epoch 34/69
76/76 - 1s - loss: 4.5050e-04 - accuracy: 1.0000 - 633ms/epoch - 8ms/step
Epoch 35/69
76/76 - 1s - loss: 4.1788e-04 - accuracy: 1.0000 - 634ms/epoch - 8ms/step
Epoch 36/69
76/76 - 1s - loss: 3.9318e-04 - accuracy: 1.0000 - 632ms/epoch - 8ms/step
Epoch 37/69
76/76 - 1s - loss: 3.6728e-04 - accuracy: 1.0000 - 635ms/epoch - 8ms/step
Epoch 38/69
76/76 - 1s - loss: 3.4499e-04 - accuracy: 1.0000 - 630ms/epoch - 8ms/step
Epoch 39/69
76/76 - 1s - loss: 3.2014e-04 - accuracy: 1.0000 - 623ms/epoch - 8ms/step
Epoch 40/69
76/76 - 1s - loss: 3.0176e-04 - accuracy: 1.0000 - 620ms/epoch - 8ms/step
Epoch 41/69
76/76 - 1s - loss: 2.7966e-04 - accuracy: 1.0000 - 616ms/epoch - 8ms/step
Epoch 42/69
76/76 - 1s - loss: 2.6502e-04 - accuracy: 1.0000 - 614ms/epoch - 8ms/step
Epoch 43/69
76/76 - 1s - loss: 2.4553e-04 - accuracy: 1.0000 - 614ms/epoch - 8ms/step
Epoch 44/69
76/76 - 1s - loss: 2.3227e-04 - accuracy: 1.0000 - 622ms/epoch - 8ms/step
Epoch 45/69
76/76 - 1s - loss: 2.1814e-04 - accuracy: 1.0000 - 662ms/epoch - 9ms/step
Epoch 46/69
76/76 - 1s - loss: 2.0614e-04 - accuracy: 1.0000 - 694ms/epoch - 9ms/step
Epoch 47/69
```

```
76/76 - 1s - loss: 1.9599e-04 - accuracy: 1.0000 - 656ms/epoch - 9ms/step
Epoch 48/69
76/76 - 1s - loss: 1.8159e-04 - accuracy: 1.0000 - 636ms/epoch - 8ms/step
Epoch 49/69
76/76 - 1s - loss: 1.7384e-04 - accuracy: 1.0000 - 624ms/epoch - 8ms/step
Epoch 50/69
76/76 - 1s - loss: 1.6332e-04 - accuracy: 1.0000 - 621ms/epoch - 8ms/step
Epoch 51/69
76/76 - 1s - loss: 1.5551e-04 - accuracy: 1.0000 - 620ms/epoch - 8ms/step
Epoch 52/69
76/76 - 1s - loss: 1.4689e-04 - accuracy: 1.0000 - 614ms/epoch - 8ms/step
Epoch 53/69
76/76 - 1s - loss: 1.3784e-04 - accuracy: 1.0000 - 614ms/epoch - 8ms/step
Epoch 54/69
76/76 - 1s - loss: 1.2927e-04 - accuracy: 1.0000 - 614ms/epoch - 8ms/step
Epoch 55/69
76/76 - 1s - loss: 1.2709e-04 - accuracy: 1.0000 - 612ms/epoch - 8ms/step
Epoch 56/69
76/76 - 1s - loss: 1.1774e-04 - accuracy: 1.0000 - 612ms/epoch - 8ms/step
Epoch 57/69
76/76 - 1s - loss: 1.1084e-04 - accuracy: 1.0000 - 613ms/epoch - 8ms/step
Epoch 58/69
76/76 - 1s - loss: 1.0505e-04 - accuracy: 1.0000 - 608ms/epoch - 8ms/step
Epoch 59/69
76/76 - 1s - loss: 9.9577e-05 - accuracy: 1.0000 - 604ms/epoch - 8ms/step
Epoch 60/69
76/76 - 1s - loss: 9.4613e-05 - accuracy: 1.0000 - 611ms/epoch - 8ms/step
Epoch 61/69
76/76 - 1s - loss: 8.9309e-05 - accuracy: 1.0000 - 605ms/epoch - 8ms/step
Epoch 62/69
76/76 - 1s - loss: 8.5219e-05 - accuracy: 1.0000 - 613ms/epoch - 8ms/step
Epoch 63/69
76/76 - 1s - loss: 8.1059e-05 - accuracy: 1.0000 - 613ms/epoch - 8ms/step
Epoch 64/69
76/76 - 1s - loss: 7.7053e-05 - accuracy: 1.0000 - 618ms/epoch - 8ms/step
Epoch 65/69
76/76 - 1s - loss: 7.3598e-05 - accuracy: 1.0000 - 619ms/epoch - 8ms/step
Epoch 66/69
76/76 - 1s - loss: 6.9832e-05 - accuracy: 1.0000 - 616ms/epoch - 8ms/step
Epoch 67/69
76/76 - 1s - loss: 6.5422e-05 - accuracy: 1.0000 - 615ms/epoch - 8ms/step
Epoch 68/69
76/76 - 1s - loss: 6.3187e-05 - accuracy: 1.0000 - 625ms/epoch - 8ms/step
Epoch 69/69
76/76 - 1s - loss: 6.0510e-05 - accuracy: 1.0000 - 627ms/epoch - 8ms/step
Test Accuracy: 0.6618357300758362
```

[70]:
```
#9 Under Sampling_Bag of n grams 1000 feat
run_models(x_train16, x_test16, y_train16, y_test16, n_words16)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 66.86
Precision : 0.6685990338164252
Recall : 0.6685990338164252
F1-score : 0.6685990338164252
ROC_AOC_Score for Naive Bayes:  0.6815735410140841

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 65.31
Precision : 0.6531400966183575
Recall : 0.6531400966183575
F1-score : 0.6531400966183575
ROC_AOC_Score for Logistic Regression:  0.6728520430880275

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 68.5
Precision : 0.6850241545893719
Recall : 0.6850241545893719
F1-score : 0.6850241545893719

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 66.96
Precision : 0.6695652173913044
Recall : 0.6695652173913044
F1-score : 0.6695652173913044
ROC_AOC_Score for Random Forest:  0.7073826132466782

FOR ANN:

Model: "sequential_14"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_28 (Dense)            (None, 50)                50050

 dense_29 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
76/76 - 1s - loss: 0.6717 - accuracy: 0.6063 - 1s/epoch - 17ms/step
Epoch 2/69
76/76 - 1s - loss: 0.5759 - accuracy: 0.7385 - 573ms/epoch - 8ms/step
Epoch 3/69
76/76 - 1s - loss: 0.5081 - accuracy: 0.7704 - 587ms/epoch - 8ms/step
Epoch 4/69
76/76 - 1s - loss: 0.4513 - accuracy: 0.8052 - 581ms/epoch - 8ms/step
Epoch 5/69
76/76 - 1s - loss: 0.3999 - accuracy: 0.8396 - 583ms/epoch - 8ms/step
Epoch 6/69
76/76 - 1s - loss: 0.3560 - accuracy: 0.8599 - 575ms/epoch - 8ms/step
Epoch 7/69
76/76 - 1s - loss: 0.3150 - accuracy: 0.8856 - 574ms/epoch - 8ms/step
Epoch 8/69
76/76 - 1s - loss: 0.2777 - accuracy: 0.9055 - 604ms/epoch - 8ms/step
Epoch 9/69
76/76 - 1s - loss: 0.2434 - accuracy: 0.9291 - 581ms/epoch - 8ms/step
Epoch 10/69
76/76 - 1s - loss: 0.2148 - accuracy: 0.9412 - 580ms/epoch - 8ms/step
Epoch 11/69
76/76 - 1s - loss: 0.1864 - accuracy: 0.9544 - 577ms/epoch - 8ms/step
Epoch 12/69
76/76 - 1s - loss: 0.1639 - accuracy: 0.9631 - 577ms/epoch - 8ms/step
Epoch 13/69
76/76 - 1s - loss: 0.1436 - accuracy: 0.9710 - 567ms/epoch - 7ms/step
Epoch 14/69
76/76 - 1s - loss: 0.1261 - accuracy: 0.9751 - 583ms/epoch - 8ms/step
Epoch 15/69
76/76 - 1s - loss: 0.1118 - accuracy: 0.9793 - 571ms/epoch - 8ms/step
Epoch 16/69
76/76 - 1s - loss: 0.0990 - accuracy: 0.9826 - 575ms/epoch - 8ms/step
Epoch 17/69
76/76 - 1s - loss: 0.0879 - accuracy: 0.9843 - 589ms/epoch - 8ms/step
Epoch 18/69
76/76 - 1s - loss: 0.0787 - accuracy: 0.9880 - 575ms/epoch - 8ms/step
Epoch 19/69
76/76 - 1s - loss: 0.0707 - accuracy: 0.9892 - 571ms/epoch - 8ms/step
Epoch 20/69
76/76 - 1s - loss: 0.0638 - accuracy: 0.9909 - 576ms/epoch - 8ms/step
Epoch 21/69
76/76 - 1s - loss: 0.0578 - accuracy: 0.9917 - 573ms/epoch - 8ms/step
Epoch 22/69
76/76 - 1s - loss: 0.0528 - accuracy: 0.9930 - 578ms/epoch - 8ms/step
Epoch 23/69
```

```
76/76 - 1s - loss: 0.0482 - accuracy: 0.9938 - 570ms/epoch - 7ms/step
Epoch 24/69
76/76 - 1s - loss: 0.0442 - accuracy: 0.9934 - 570ms/epoch - 8ms/step
Epoch 25/69
76/76 - 1s - loss: 0.0409 - accuracy: 0.9942 - 586ms/epoch - 8ms/step
Epoch 26/69
76/76 - 1s - loss: 0.0375 - accuracy: 0.9946 - 640ms/epoch - 8ms/step
Epoch 27/69
76/76 - 1s - loss: 0.0347 - accuracy: 0.9954 - 626ms/epoch - 8ms/step
Epoch 28/69
76/76 - 1s - loss: 0.0325 - accuracy: 0.9950 - 575ms/epoch - 8ms/step
Epoch 29/69
76/76 - 1s - loss: 0.0303 - accuracy: 0.9954 - 572ms/epoch - 8ms/step
Epoch 30/69
76/76 - 1s - loss: 0.0284 - accuracy: 0.9954 - 569ms/epoch - 7ms/step
Epoch 31/69
76/76 - 1s - loss: 0.0264 - accuracy: 0.9959 - 581ms/epoch - 8ms/step
Epoch 32/69
76/76 - 1s - loss: 0.0249 - accuracy: 0.9959 - 568ms/epoch - 7ms/step
Epoch 33/69
76/76 - 1s - loss: 0.0235 - accuracy: 0.9963 - 574ms/epoch - 8ms/step
Epoch 34/69
76/76 - 1s - loss: 0.0220 - accuracy: 0.9971 - 571ms/epoch - 8ms/step
Epoch 35/69
76/76 - 1s - loss: 0.0208 - accuracy: 0.9971 - 574ms/epoch - 8ms/step
Epoch 36/69
76/76 - 1s - loss: 0.0198 - accuracy: 0.9967 - 581ms/epoch - 8ms/step
Epoch 37/69
76/76 - 1s - loss: 0.0188 - accuracy: 0.9971 - 577ms/epoch - 8ms/step
Epoch 38/69
76/76 - 1s - loss: 0.0179 - accuracy: 0.9971 - 578ms/epoch - 8ms/step
Epoch 39/69
76/76 - 1s - loss: 0.0169 - accuracy: 0.9975 - 568ms/epoch - 7ms/step
Epoch 40/69
76/76 - 1s - loss: 0.0161 - accuracy: 0.9975 - 614ms/epoch - 8ms/step
Epoch 41/69
76/76 - 1s - loss: 0.0154 - accuracy: 0.9979 - 578ms/epoch - 8ms/step
Epoch 42/69
76/76 - 1s - loss: 0.0147 - accuracy: 0.9979 - 573ms/epoch - 8ms/step
Epoch 43/69
76/76 - 1s - loss: 0.0141 - accuracy: 0.9979 - 585ms/epoch - 8ms/step
Epoch 44/69
76/76 - 1s - loss: 0.0134 - accuracy: 0.9979 - 600ms/epoch - 8ms/step
Epoch 45/69
76/76 - 1s - loss: 0.0129 - accuracy: 0.9979 - 609ms/epoch - 8ms/step
Epoch 46/69
76/76 - 1s - loss: 0.0122 - accuracy: 0.9983 - 578ms/epoch - 8ms/step
Epoch 47/69
```

```
76/76 - 1s - loss: 0.0118 - accuracy: 0.9979 - 595ms/epoch - 8ms/step
Epoch 48/69
76/76 - 1s - loss: 0.0113 - accuracy: 0.9979 - 601ms/epoch - 8ms/step
Epoch 49/69
76/76 - 1s - loss: 0.0110 - accuracy: 0.9983 - 574ms/epoch - 8ms/step
Epoch 50/69
76/76 - 1s - loss: 0.0106 - accuracy: 0.9983 - 577ms/epoch - 8ms/step
Epoch 51/69
76/76 - 1s - loss: 0.0102 - accuracy: 0.9983 - 603ms/epoch - 8ms/step
Epoch 52/69
76/76 - 1s - loss: 0.0099 - accuracy: 0.9983 - 583ms/epoch - 8ms/step
Epoch 53/69
76/76 - 1s - loss: 0.0096 - accuracy: 0.9983 - 617ms/epoch - 8ms/step
Epoch 54/69
76/76 - 1s - loss: 0.0092 - accuracy: 0.9983 - 646ms/epoch - 8ms/step
Epoch 55/69
76/76 - 1s - loss: 0.0089 - accuracy: 0.9983 - 568ms/epoch - 7ms/step
Epoch 56/69
76/76 - 1s - loss: 0.0086 - accuracy: 0.9983 - 571ms/epoch - 8ms/step
Epoch 57/69
76/76 - 1s - loss: 0.0084 - accuracy: 0.9983 - 567ms/epoch - 7ms/step
Epoch 58/69
76/76 - 1s - loss: 0.0081 - accuracy: 0.9983 - 575ms/epoch - 8ms/step
Epoch 59/69
76/76 - 1s - loss: 0.0078 - accuracy: 0.9983 - 574ms/epoch - 8ms/step
Epoch 60/69
76/76 - 1s - loss: 0.0076 - accuracy: 0.9983 - 586ms/epoch - 8ms/step
Epoch 61/69
76/76 - 1s - loss: 0.0075 - accuracy: 0.9983 - 591ms/epoch - 8ms/step
Epoch 62/69
76/76 - 1s - loss: 0.0073 - accuracy: 0.9983 - 607ms/epoch - 8ms/step
Epoch 63/69
76/76 - 1s - loss: 0.0070 - accuracy: 0.9983 - 605ms/epoch - 8ms/step
Epoch 64/69
76/76 - 1s - loss: 0.0068 - accuracy: 0.9988 - 576ms/epoch - 8ms/step
Epoch 65/69
76/76 - 1s - loss: 0.0066 - accuracy: 0.9988 - 600ms/epoch - 8ms/step
Epoch 66/69
76/76 - 1s - loss: 0.0064 - accuracy: 0.9988 - 602ms/epoch - 8ms/step
Epoch 67/69
76/76 - 1s - loss: 0.0063 - accuracy: 0.9992 - 607ms/epoch - 8ms/step
Epoch 68/69
76/76 - 1s - loss: 0.0062 - accuracy: 0.9988 - 575ms/epoch - 8ms/step
Epoch 69/69
76/76 - 1s - loss: 0.0060 - accuracy: 0.9992 - 573ms/epoch - 8ms/step
Test Accuracy: 0.634782612323761
```

[71]:
```
#10 Under Sampling_Bag of n grams 5000 feat
run_models(x_train17, x_test17, y_train17, y_test17, n_words17)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 66.57
Precision : 0.6657004830917874
Recall : 0.6657004830917874
F1-score : 0.6657004830917874
ROC_AOC_Score for Naive Bayes:  0.7043673887061874

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 64.54
Precision : 0.6454106280193237
Recall : 0.6454106280193237
F1-score : 0.6454106280193237
ROC_AOC_Score for Logistic Regression:  0.6569166595061593

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 67.54
Precision : 0.6753623188405797
Recall : 0.6753623188405797
F1-score : 0.6753623188405797

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 67.25
Precision : 0.672463768115942
Recall : 0.672463768115942
F1-score : 0.672463768115942
ROC_AOC_Score for Random Forest:  0.7066952045431466

FOR ANN:

Model: "sequential_15"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_30 (Dense)            (None, 50)                250050

 dense_31 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
76/76 - 1s - loss: 0.6568 - accuracy: 0.6490 - 1s/epoch - 18ms/step
Epoch 2/69
76/76 - 1s - loss: 0.4709 - accuracy: 0.8036 - 632ms/epoch - 8ms/step
Epoch 3/69
76/76 - 1s - loss: 0.3041 - accuracy: 0.8993 - 627ms/epoch - 8ms/step
Epoch 4/69
76/76 - 1s - loss: 0.1861 - accuracy: 0.9577 - 623ms/epoch - 8ms/step
Epoch 5/69
76/76 - 1s - loss: 0.1135 - accuracy: 0.9818 - 627ms/epoch - 8ms/step
Epoch 6/69
76/76 - 1s - loss: 0.0718 - accuracy: 0.9917 - 713ms/epoch - 9ms/step
Epoch 7/69
76/76 - 1s - loss: 0.0488 - accuracy: 0.9954 - 675ms/epoch - 9ms/step
Epoch 8/69
76/76 - 1s - loss: 0.0352 - accuracy: 0.9967 - 625ms/epoch - 8ms/step
Epoch 9/69
76/76 - 1s - loss: 0.0270 - accuracy: 0.9979 - 628ms/epoch - 8ms/step
Epoch 10/69
76/76 - 1s - loss: 0.0210 - accuracy: 0.9979 - 623ms/epoch - 8ms/step
Epoch 11/69
76/76 - 1s - loss: 0.0169 - accuracy: 0.9983 - 646ms/epoch - 8ms/step
Epoch 12/69
76/76 - 1s - loss: 0.0140 - accuracy: 0.9988 - 616ms/epoch - 8ms/step
Epoch 13/69
76/76 - 1s - loss: 0.0118 - accuracy: 0.9992 - 623ms/epoch - 8ms/step
Epoch 14/69
76/76 - 1s - loss: 0.0103 - accuracy: 0.9992 - 623ms/epoch - 8ms/step
Epoch 15/69
76/76 - 1s - loss: 0.0090 - accuracy: 0.9992 - 629ms/epoch - 8ms/step
Epoch 16/69
76/76 - 1s - loss: 0.0079 - accuracy: 0.9992 - 608ms/epoch - 8ms/step
Epoch 17/69
76/76 - 1s - loss: 0.0071 - accuracy: 0.9996 - 608ms/epoch - 8ms/step
Epoch 18/69
76/76 - 1s - loss: 0.0063 - accuracy: 0.9996 - 791ms/epoch - 10ms/step
Epoch 19/69
76/76 - 1s - loss: 0.0057 - accuracy: 0.9996 - 753ms/epoch - 10ms/step
Epoch 20/69
76/76 - 1s - loss: 0.0061 - accuracy: 0.9992 - 747ms/epoch - 10ms/step
Epoch 21/69
76/76 - 1s - loss: 0.0050 - accuracy: 0.9996 - 768ms/epoch - 10ms/step
Epoch 22/69
76/76 - 1s - loss: 0.0045 - accuracy: 0.9996 - 755ms/epoch - 10ms/step
Epoch 23/69
```

```
76/76 - 1s - loss: 0.0042 - accuracy: 0.9996 - 658ms/epoch - 9ms/step
Epoch 24/69
76/76 - 1s - loss: 0.0040 - accuracy: 0.9996 - 731ms/epoch - 10ms/step
Epoch 25/69
76/76 - 1s - loss: 0.0038 - accuracy: 0.9996 - 670ms/epoch - 9ms/step
Epoch 26/69
76/76 - 1s - loss: 0.0036 - accuracy: 0.9996 - 644ms/epoch - 8ms/step
Epoch 27/69
76/76 - 1s - loss: 0.0034 - accuracy: 0.9996 - 789ms/epoch - 10ms/step
Epoch 28/69
76/76 - 1s - loss: 0.0033 - accuracy: 0.9996 - 795ms/epoch - 10ms/step
Epoch 29/69
76/76 - 1s - loss: 0.0031 - accuracy: 0.9996 - 856ms/epoch - 11ms/step
Epoch 30/69
76/76 - 1s - loss: 0.0030 - accuracy: 0.9996 - 961ms/epoch - 13ms/step
Epoch 31/69
76/76 - 1s - loss: 0.0029 - accuracy: 0.9996 - 703ms/epoch - 9ms/step
Epoch 32/69
76/76 - 1s - loss: 0.0028 - accuracy: 0.9996 - 752ms/epoch - 10ms/step
Epoch 33/69
76/76 - 1s - loss: 0.0027 - accuracy: 0.9996 - 730ms/epoch - 10ms/step
Epoch 34/69
76/76 - 1s - loss: 0.0026 - accuracy: 0.9996 - 621ms/epoch - 8ms/step
Epoch 35/69
76/76 - 1s - loss: 0.0026 - accuracy: 0.9996 - 671ms/epoch - 9ms/step
Epoch 36/69
76/76 - 1s - loss: 0.0025 - accuracy: 0.9996 - 625ms/epoch - 8ms/step
Epoch 37/69
76/76 - 1s - loss: 0.0024 - accuracy: 0.9996 - 670ms/epoch - 9ms/step
Epoch 38/69
76/76 - 1s - loss: 0.0024 - accuracy: 0.9996 - 621ms/epoch - 8ms/step
Epoch 39/69
76/76 - 1s - loss: 0.0023 - accuracy: 0.9996 - 619ms/epoch - 8ms/step
Epoch 40/69
76/76 - 1s - loss: 0.0023 - accuracy: 0.9996 - 628ms/epoch - 8ms/step
Epoch 41/69
76/76 - 1s - loss: 0.0022 - accuracy: 0.9996 - 616ms/epoch - 8ms/step
Epoch 42/69
76/76 - 1s - loss: 0.0022 - accuracy: 0.9996 - 630ms/epoch - 8ms/step
Epoch 43/69
76/76 - 1s - loss: 0.0021 - accuracy: 0.9996 - 618ms/epoch - 8ms/step
Epoch 44/69
76/76 - 1s - loss: 0.0021 - accuracy: 0.9996 - 611ms/epoch - 8ms/step
Epoch 45/69
76/76 - 1s - loss: 0.0021 - accuracy: 0.9996 - 605ms/epoch - 8ms/step
Epoch 46/69
76/76 - 1s - loss: 0.0020 - accuracy: 0.9996 - 607ms/epoch - 8ms/step
Epoch 47/69
```

```
76/76 - 1s - loss: 0.0020 - accuracy: 0.9996 - 604ms/epoch - 8ms/step
Epoch 48/69
76/76 - 1s - loss: 0.0019 - accuracy: 0.9996 - 602ms/epoch - 8ms/step
Epoch 49/69
76/76 - 1s - loss: 0.0019 - accuracy: 0.9996 - 604ms/epoch - 8ms/step
Epoch 50/69
76/76 - 1s - loss: 0.0019 - accuracy: 0.9996 - 610ms/epoch - 8ms/step
Epoch 51/69
76/76 - 1s - loss: 0.0019 - accuracy: 0.9996 - 609ms/epoch - 8ms/step
Epoch 52/69
76/76 - 1s - loss: 0.0019 - accuracy: 0.9996 - 604ms/epoch - 8ms/step
Epoch 53/69
76/76 - 1s - loss: 0.0018 - accuracy: 0.9996 - 616ms/epoch - 8ms/step
Epoch 54/69
76/76 - 1s - loss: 0.0018 - accuracy: 0.9996 - 657ms/epoch - 9ms/step
Epoch 55/69
76/76 - 1s - loss: 0.0018 - accuracy: 0.9996 - 645ms/epoch - 8ms/step
Epoch 56/69
76/76 - 1s - loss: 0.0018 - accuracy: 0.9996 - 607ms/epoch - 8ms/step
Epoch 57/69
76/76 - 1s - loss: 0.0017 - accuracy: 0.9996 - 601ms/epoch - 8ms/step
Epoch 58/69
76/76 - 1s - loss: 0.0017 - accuracy: 0.9996 - 604ms/epoch - 8ms/step
Epoch 59/69
76/76 - 1s - loss: 0.0018 - accuracy: 0.9996 - 607ms/epoch - 8ms/step
Epoch 60/69
76/76 - 1s - loss: 0.0017 - accuracy: 0.9996 - 609ms/epoch - 8ms/step
Epoch 61/69
76/76 - 1s - loss: 0.0017 - accuracy: 0.9996 - 604ms/epoch - 8ms/step
Epoch 62/69
76/76 - 1s - loss: 0.0017 - accuracy: 0.9996 - 601ms/epoch - 8ms/step
Epoch 63/69
76/76 - 1s - loss: 0.0017 - accuracy: 0.9996 - 613ms/epoch - 8ms/step
Epoch 64/69
76/76 - 1s - loss: 0.0017 - accuracy: 0.9996 - 606ms/epoch - 8ms/step
Epoch 65/69
76/76 - 1s - loss: 0.0016 - accuracy: 0.9996 - 607ms/epoch - 8ms/step
Epoch 66/69
76/76 - 1s - loss: 0.0016 - accuracy: 0.9996 - 601ms/epoch - 8ms/step
Epoch 67/69
76/76 - 1s - loss: 0.0016 - accuracy: 0.9996 - 609ms/epoch - 8ms/step
Epoch 68/69
76/76 - 1s - loss: 0.0016 - accuracy: 0.9996 - 611ms/epoch - 8ms/step
Epoch 69/69
76/76 - 1s - loss: 0.0016 - accuracy: 0.9996 - 608ms/epoch - 8ms/step
Test Accuracy: 0.643478274345398
```

[72]:
```
#11 Under Sampling_TF-IDF 1000 feat
run_models(x_train18, x_test18, y_train18, y_test18, n_words18)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 65.99
Precision : 0.659903381642512
Recall : 0.659903381642512
F1-score : 0.659903381642512
ROC_AOC_Score for Naive Bayes:  0.7255091120710531

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 68.12
Precision : 0.6811594202898551
Recall : 0.6811594202898551
F1-score : 0.6811594202898551
ROC_AOC_Score for Logistic Regression:  0.7268253435090652

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 68.5
Precision : 0.6850241545893719
Recall : 0.6850241545893719
F1-score : 0.6850241545893719

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 67.05
Precision : 0.6705314009661836
Recall : 0.6705314009661836
F1-score : 0.6705314009661836
ROC_AOC_Score for Random Forest:  0.7249349695743533

FOR ANN:

Model: "sequential_16"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_32 (Dense)            (None, 50)                50050

 dense_33 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
76/76 - 1s - loss: 0.6600 - accuracy: 0.5972 - 1s/epoch - 20ms/step
Epoch 2/69
76/76 - 1s - loss: 0.5793 - accuracy: 0.7174 - 592ms/epoch - 8ms/step
Epoch 3/69
76/76 - 1s - loss: 0.5280 - accuracy: 0.7468 - 606ms/epoch - 8ms/step
Epoch 4/69
76/76 - 1s - loss: 0.4944 - accuracy: 0.7654 - 651ms/epoch - 9ms/step
Epoch 5/69
76/76 - 1s - loss: 0.4672 - accuracy: 0.7766 - 617ms/epoch - 8ms/step
Epoch 6/69
76/76 - 1s - loss: 0.4436 - accuracy: 0.7961 - 578ms/epoch - 8ms/step
Epoch 7/69
76/76 - 1s - loss: 0.4243 - accuracy: 0.8044 - 567ms/epoch - 7ms/step
Epoch 8/69
76/76 - 1s - loss: 0.4037 - accuracy: 0.8193 - 559ms/epoch - 7ms/step
Epoch 9/69
76/76 - 1s - loss: 0.3868 - accuracy: 0.8284 - 572ms/epoch - 8ms/step
Epoch 10/69
76/76 - 1s - loss: 0.3713 - accuracy: 0.8380 - 570ms/epoch - 8ms/step
Epoch 11/69
76/76 - 1s - loss: 0.3569 - accuracy: 0.8417 - 570ms/epoch - 8ms/step
Epoch 12/69
76/76 - 1s - loss: 0.3419 - accuracy: 0.8533 - 575ms/epoch - 8ms/step
Epoch 13/69
76/76 - 1s - loss: 0.3286 - accuracy: 0.8579 - 568ms/epoch - 7ms/step
Epoch 14/69
76/76 - 1s - loss: 0.3162 - accuracy: 0.8661 - 585ms/epoch - 8ms/step
Epoch 15/69
76/76 - 1s - loss: 0.3031 - accuracy: 0.8757 - 577ms/epoch - 8ms/step
Epoch 16/69
76/76 - 1s - loss: 0.2911 - accuracy: 0.8819 - 578ms/epoch - 8ms/step
Epoch 17/69
76/76 - 1s - loss: 0.2798 - accuracy: 0.8856 - 634ms/epoch - 8ms/step
Epoch 18/69
76/76 - 1s - loss: 0.2681 - accuracy: 0.8935 - 618ms/epoch - 8ms/step
Epoch 19/69
76/76 - 1s - loss: 0.2557 - accuracy: 0.9014 - 613ms/epoch - 8ms/step
Epoch 20/69
76/76 - 1s - loss: 0.2452 - accuracy: 0.9097 - 616ms/epoch - 8ms/step
Epoch 21/69
76/76 - 1s - loss: 0.2347 - accuracy: 0.9142 - 607ms/epoch - 8ms/step
Epoch 22/69
76/76 - 1s - loss: 0.2232 - accuracy: 0.9213 - 619ms/epoch - 8ms/step
Epoch 23/69
```

```
76/76 - 1s - loss: 0.2129 - accuracy: 0.9279 - 585ms/epoch - 8ms/step
Epoch 24/69
76/76 - 1s - loss: 0.2016 - accuracy: 0.9324 - 777ms/epoch - 10ms/step
Epoch 25/69
76/76 - 1s - loss: 0.1921 - accuracy: 0.9374 - 766ms/epoch - 10ms/step
Epoch 26/69
76/76 - 1s - loss: 0.1818 - accuracy: 0.9490 - 608ms/epoch - 8ms/step
Epoch 27/69
76/76 - 1s - loss: 0.1717 - accuracy: 0.9573 - 574ms/epoch - 8ms/step
Epoch 28/69
76/76 - 1s - loss: 0.1622 - accuracy: 0.9569 - 575ms/epoch - 8ms/step
Epoch 29/69
76/76 - 1s - loss: 0.1516 - accuracy: 0.9677 - 570ms/epoch - 8ms/step
Epoch 30/69
76/76 - 1s - loss: 0.1440 - accuracy: 0.9697 - 651ms/epoch - 9ms/step
Epoch 31/69
76/76 - 1s - loss: 0.1349 - accuracy: 0.9797 - 734ms/epoch - 10ms/step
Epoch 32/69
76/76 - 1s - loss: 0.1283 - accuracy: 0.9822 - 661ms/epoch - 9ms/step
Epoch 33/69
76/76 - 1s - loss: 0.1184 - accuracy: 0.9855 - 582ms/epoch - 8ms/step
Epoch 34/69
76/76 - 1s - loss: 0.1105 - accuracy: 0.9905 - 569ms/epoch - 7ms/step
Epoch 35/69
76/76 - 1s - loss: 0.1041 - accuracy: 0.9909 - 578ms/epoch - 8ms/step
Epoch 36/69
76/76 - 1s - loss: 0.0975 - accuracy: 0.9930 - 581ms/epoch - 8ms/step
Epoch 37/69
76/76 - 1s - loss: 0.0901 - accuracy: 0.9950 - 571ms/epoch - 8ms/step
Epoch 38/69
76/76 - 1s - loss: 0.0851 - accuracy: 0.9954 - 563ms/epoch - 7ms/step
Epoch 39/69
76/76 - 1s - loss: 0.0789 - accuracy: 0.9975 - 560ms/epoch - 7ms/step
Epoch 40/69
76/76 - 1s - loss: 0.0739 - accuracy: 0.9979 - 572ms/epoch - 8ms/step
Epoch 41/69
76/76 - 1s - loss: 0.0680 - accuracy: 0.9988 - 573ms/epoch - 8ms/step
Epoch 42/69
76/76 - 1s - loss: 0.0644 - accuracy: 0.9988 - 567ms/epoch - 7ms/step
Epoch 43/69
76/76 - 1s - loss: 0.0594 - accuracy: 0.9992 - 600ms/epoch - 8ms/step
Epoch 44/69
76/76 - 1s - loss: 0.0553 - accuracy: 0.9988 - 583ms/epoch - 8ms/step
Epoch 45/69
76/76 - 1s - loss: 0.0521 - accuracy: 0.9996 - 576ms/epoch - 8ms/step
Epoch 46/69
76/76 - 1s - loss: 0.0484 - accuracy: 0.9996 - 573ms/epoch - 8ms/step
Epoch 47/69
```

```
76/76 - 1s - loss: 0.0453 - accuracy: 0.9992 - 567ms/epoch - 7ms/step
Epoch 48/69
76/76 - 1s - loss: 0.0420 - accuracy: 0.9992 - 568ms/epoch - 7ms/step
Epoch 49/69
76/76 - 1s - loss: 0.0393 - accuracy: 0.9992 - 568ms/epoch - 7ms/step
Epoch 50/69
76/76 - 1s - loss: 0.0366 - accuracy: 0.9996 - 642ms/epoch - 8ms/step
Epoch 51/69
76/76 - 1s - loss: 0.0344 - accuracy: 0.9996 - 718ms/epoch - 9ms/step
Epoch 52/69
76/76 - 1s - loss: 0.0322 - accuracy: 0.9996 - 602ms/epoch - 8ms/step
Epoch 53/69
76/76 - 1s - loss: 0.0304 - accuracy: 0.9996 - 660ms/epoch - 9ms/step
Epoch 54/69
76/76 - 1s - loss: 0.0282 - accuracy: 0.9996 - 575ms/epoch - 8ms/step
Epoch 55/69
76/76 - 1s - loss: 0.0265 - accuracy: 0.9996 - 578ms/epoch - 8ms/step
Epoch 56/69
76/76 - 1s - loss: 0.0248 - accuracy: 0.9996 - 655ms/epoch - 9ms/step
Epoch 57/69
76/76 - 1s - loss: 0.0233 - accuracy: 0.9992 - 759ms/epoch - 10ms/step
Epoch 58/69
76/76 - 1s - loss: 0.0221 - accuracy: 0.9996 - 663ms/epoch - 9ms/step
Epoch 59/69
76/76 - 1s - loss: 0.0208 - accuracy: 0.9988 - 582ms/epoch - 8ms/step
Epoch 60/69
76/76 - 1s - loss: 0.0195 - accuracy: 0.9996 - 622ms/epoch - 8ms/step
Epoch 61/69
76/76 - 1s - loss: 0.0183 - accuracy: 0.9996 - 635ms/epoch - 8ms/step
Epoch 62/69
76/76 - 1s - loss: 0.0174 - accuracy: 0.9992 - 625ms/epoch - 8ms/step
Epoch 63/69
76/76 - 1s - loss: 0.0164 - accuracy: 0.9996 - 620ms/epoch - 8ms/step
Epoch 64/69
76/76 - 1s - loss: 0.0154 - accuracy: 0.9996 - 634ms/epoch - 8ms/step
Epoch 65/69
76/76 - 1s - loss: 0.0148 - accuracy: 0.9996 - 576ms/epoch - 8ms/step
Epoch 66/69
76/76 - 1s - loss: 0.0140 - accuracy: 0.9996 - 577ms/epoch - 8ms/step
Epoch 67/69
76/76 - 1s - loss: 0.0131 - accuracy: 0.9992 - 580ms/epoch - 8ms/step
Epoch 68/69
76/76 - 1s - loss: 0.0125 - accuracy: 0.9992 - 560ms/epoch - 7ms/step
Epoch 69/69
76/76 - 1s - loss: 0.0118 - accuracy: 0.9996 - 566ms/epoch - 7ms/step
Test Accuracy: 0.6202898621559143
```

[73]:
```
#12 Under Sampling_TF-IDF 5000 feat
run_models(x_train19, x_test19, y_train19, y_test19, n_words19)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 68.21
Precision : 0.6821256038647343
Recall : 0.6821256038647343
F1-score : 0.6821256038647343
ROC_AOC_Score for Naive Bayes:  0.7441277330354562

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 68.79
Precision : 0.6879227053140097
Recall : 0.6879227053140097
F1-score : 0.6879227053140097
ROC_AOC_Score for Logistic Regression:  0.7474163587648515

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 69.28
Precision : 0.6927536231884058
Recall : 0.6927536231884058
F1-score : 0.6927536231884058

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 70.53
Precision : 0.7053140096618358
Recall : 0.7053140096618358
F1-score : 0.7053140096618358
ROC_AOC_Score for Random Forest:  0.7555051282251576

FOR ANN:

Model: "sequential_17"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_34 (Dense)            (None, 50)                250050

 dense_35 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
76/76 - 2s - loss: 0.6545 - accuracy: 0.6121 - 2s/epoch - 21ms/step
Epoch 2/69
76/76 - 1s - loss: 0.5478 - accuracy: 0.7513 - 629ms/epoch - 8ms/step
Epoch 3/69
76/76 - 1s - loss: 0.4577 - accuracy: 0.7978 - 647ms/epoch - 9ms/step
Epoch 4/69
76/76 - 1s - loss: 0.3765 - accuracy: 0.8541 - 711ms/epoch - 9ms/step
Epoch 5/69
76/76 - 1s - loss: 0.3060 - accuracy: 0.9026 - 666ms/epoch - 9ms/step
Epoch 6/69
76/76 - 1s - loss: 0.2446 - accuracy: 0.9387 - 641ms/epoch - 8ms/step
Epoch 7/69
76/76 - 1s - loss: 0.1961 - accuracy: 0.9586 - 631ms/epoch - 8ms/step
Epoch 8/69
76/76 - 1s - loss: 0.1561 - accuracy: 0.9722 - 627ms/epoch - 8ms/step
Epoch 9/69
76/76 - 1s - loss: 0.1262 - accuracy: 0.9822 - 604ms/epoch - 8ms/step
Epoch 10/69
76/76 - 1s - loss: 0.1012 - accuracy: 0.9892 - 609ms/epoch - 8ms/step
Epoch 11/69
76/76 - 1s - loss: 0.0817 - accuracy: 0.9925 - 590ms/epoch - 8ms/step
Epoch 12/69
76/76 - 1s - loss: 0.0663 - accuracy: 0.9950 - 601ms/epoch - 8ms/step
Epoch 13/69
76/76 - 1s - loss: 0.0550 - accuracy: 0.9954 - 595ms/epoch - 8ms/step
Epoch 14/69
76/76 - 1s - loss: 0.0452 - accuracy: 0.9967 - 639ms/epoch - 8ms/step
Epoch 15/69
76/76 - 1s - loss: 0.0383 - accuracy: 0.9967 - 613ms/epoch - 8ms/step
Epoch 16/69
76/76 - 1s - loss: 0.0322 - accuracy: 0.9979 - 613ms/epoch - 8ms/step
Epoch 17/69
76/76 - 1s - loss: 0.0278 - accuracy: 0.9983 - 662ms/epoch - 9ms/step
Epoch 18/69
76/76 - 1s - loss: 0.0240 - accuracy: 0.9992 - 728ms/epoch - 10ms/step
Epoch 19/69
76/76 - 1s - loss: 0.0213 - accuracy: 0.9983 - 610ms/epoch - 8ms/step
Epoch 20/69
76/76 - 1s - loss: 0.0187 - accuracy: 0.9988 - 637ms/epoch - 8ms/step
Epoch 21/69
76/76 - 1s - loss: 0.0165 - accuracy: 0.9988 - 634ms/epoch - 8ms/step
Epoch 22/69
76/76 - 1s - loss: 0.0145 - accuracy: 0.9983 - 627ms/epoch - 8ms/step
Epoch 23/69
```

```
76/76 - 1s - loss: 0.0134 - accuracy: 0.9983 - 636ms/epoch - 8ms/step
Epoch 24/69
76/76 - 1s - loss: 0.0121 - accuracy: 0.9992 - 619ms/epoch - 8ms/step
Epoch 25/69
76/76 - 1s - loss: 0.0106 - accuracy: 0.9992 - 594ms/epoch - 8ms/step
Epoch 26/69
76/76 - 1s - loss: 0.0100 - accuracy: 0.9996 - 596ms/epoch - 8ms/step
Epoch 27/69
76/76 - 1s - loss: 0.0090 - accuracy: 0.9992 - 601ms/epoch - 8ms/step
Epoch 28/69
76/76 - 1s - loss: 0.0085 - accuracy: 0.9992 - 611ms/epoch - 8ms/step
Epoch 29/69
76/76 - 1s - loss: 0.0077 - accuracy: 0.9988 - 655ms/epoch - 9ms/step
Epoch 30/69
76/76 - 1s - loss: 0.0070 - accuracy: 0.9992 - 629ms/epoch - 8ms/step
Epoch 31/69
76/76 - 1s - loss: 0.0068 - accuracy: 0.9996 - 603ms/epoch - 8ms/step
Epoch 32/69
76/76 - 1s - loss: 0.0067 - accuracy: 0.9992 - 599ms/epoch - 8ms/step
Epoch 33/69
76/76 - 1s - loss: 0.0060 - accuracy: 0.9992 - 592ms/epoch - 8ms/step
Epoch 34/69
76/76 - 1s - loss: 0.0055 - accuracy: 0.9988 - 597ms/epoch - 8ms/step
Epoch 35/69
76/76 - 1s - loss: 0.0048 - accuracy: 0.9996 - 606ms/epoch - 8ms/step
Epoch 36/69
76/76 - 1s - loss: 0.0052 - accuracy: 0.9992 - 622ms/epoch - 8ms/step
Epoch 37/69
76/76 - 1s - loss: 0.0049 - accuracy: 0.9992 - 608ms/epoch - 8ms/step
Epoch 38/69
76/76 - 1s - loss: 0.0042 - accuracy: 0.9996 - 598ms/epoch - 8ms/step
Epoch 39/69
76/76 - 1s - loss: 0.0043 - accuracy: 0.9992 - 598ms/epoch - 8ms/step
Epoch 40/69
76/76 - 1s - loss: 0.0036 - accuracy: 0.9996 - 646ms/epoch - 8ms/step
Epoch 41/69
76/76 - 1s - loss: 0.0041 - accuracy: 0.9992 - 723ms/epoch - 10ms/step
Epoch 42/69
76/76 - 1s - loss: 0.0034 - accuracy: 0.9996 - 614ms/epoch - 8ms/step
Epoch 43/69
76/76 - 1s - loss: 0.0031 - accuracy: 0.9996 - 610ms/epoch - 8ms/step
Epoch 44/69
76/76 - 1s - loss: 0.0031 - accuracy: 0.9996 - 614ms/epoch - 8ms/step
Epoch 45/69
76/76 - 1s - loss: 0.0029 - accuracy: 0.9996 - 600ms/epoch - 8ms/step
Epoch 46/69
76/76 - 1s - loss: 0.0027 - accuracy: 0.9996 - 605ms/epoch - 8ms/step
Epoch 47/69
```

```
76/76 - 1s - loss: 0.0024 - accuracy: 0.9996 - 608ms/epoch - 8ms/step
Epoch 48/69
76/76 - 1s - loss: 0.0025 - accuracy: 0.9996 - 602ms/epoch - 8ms/step
Epoch 49/69
76/76 - 1s - loss: 0.0025 - accuracy: 0.9996 - 598ms/epoch - 8ms/step
Epoch 50/69
76/76 - 1s - loss: 0.0024 - accuracy: 0.9996 - 600ms/epoch - 8ms/step
Epoch 51/69
76/76 - 1s - loss: 0.0021 - accuracy: 0.9996 - 595ms/epoch - 8ms/step
Epoch 52/69
76/76 - 1s - loss: 0.0021 - accuracy: 0.9996 - 608ms/epoch - 8ms/step
Epoch 53/69
76/76 - 1s - loss: 0.0022 - accuracy: 0.9996 - 602ms/epoch - 8ms/step
Epoch 54/69
76/76 - 1s - loss: 0.0019 - accuracy: 0.9996 - 617ms/epoch - 8ms/step
Epoch 55/69
76/76 - 1s - loss: 0.0020 - accuracy: 0.9996 - 661ms/epoch - 9ms/step
Epoch 56/69
76/76 - 1s - loss: 0.0016 - accuracy: 0.9996 - 627ms/epoch - 8ms/step
Epoch 57/69
76/76 - 1s - loss: 0.0016 - accuracy: 0.9996 - 594ms/epoch - 8ms/step
Epoch 58/69
76/76 - 1s - loss: 0.0014 - accuracy: 0.9996 - 590ms/epoch - 8ms/step
Epoch 59/69
76/76 - 1s - loss: 0.0017 - accuracy: 0.9996 - 603ms/epoch - 8ms/step
Epoch 60/69
76/76 - 1s - loss: 0.0013 - accuracy: 0.9996 - 629ms/epoch - 8ms/step
Epoch 61/69
76/76 - 1s - loss: 0.0014 - accuracy: 0.9996 - 716ms/epoch - 9ms/step
Epoch 62/69
76/76 - 1s - loss: 0.0015 - accuracy: 0.9996 - 676ms/epoch - 9ms/step
Epoch 63/69
76/76 - 1s - loss: 0.0015 - accuracy: 0.9996 - 706ms/epoch - 9ms/step
Epoch 64/69
76/76 - 1s - loss: 0.0016 - accuracy: 0.9996 - 640ms/epoch - 8ms/step
Epoch 65/69
76/76 - 1s - loss: 0.0021 - accuracy: 0.9992 - 608ms/epoch - 8ms/step
Epoch 66/69
76/76 - 1s - loss: 0.0014 - accuracy: 0.9996 - 599ms/epoch - 8ms/step
Epoch 67/69
76/76 - 1s - loss: 0.0013 - accuracy: 0.9996 - 618ms/epoch - 8ms/step
Epoch 68/69
76/76 - 1s - loss: 0.0013 - accuracy: 0.9996 - 604ms/epoch - 8ms/step
Epoch 69/69
76/76 - 1s - loss: 0.0012 - accuracy: 0.9996 - 603ms/epoch - 8ms/step
Test Accuracy: 0.643478274345398
```

[74]:
```
#13 Over Sampling_BoW 1000 feat
run_models(x_train20, x_test20, y_train20, y_test20, n_words20)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 67.92
Precision : 0.679200238734706
Recall : 0.679200238734706
F1-score : 0.679200238734706
ROC_AOC_Score for Naive Bayes:  0.7389964582430073

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 72.25
Precision : 0.7224709042076992
Recall : 0.7224709042076992
F1-score : 0.7224709042076992
ROC_AOC_Score for Logistic Regression:  0.7794959660881894

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 84.33
Precision : 0.8433303491495077
Recall : 0.8433303491495077
F1-score : 0.8433303491495077

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 93.02
Precision : 0.9301700984780662
Recall : 0.9301700984780662
F1-score : 0.9301700984780662
ROC_AOC_Score for Random Forest:  0.97043372608069

FOR ANN:

Model: "sequential_18"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_36 (Dense)            (None, 50)                50050

 dense_37 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5511 - accuracy: 0.7252 - 3s/epoch - 14ms/step
Epoch 2/69
245/245 - 2s - loss: 0.3959 - accuracy: 0.8322 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.2841 - accuracy: 0.9009 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.1969 - accuracy: 0.9432 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.1278 - accuracy: 0.9761 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.0831 - accuracy: 0.9868 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.0551 - accuracy: 0.9944 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.0380 - accuracy: 0.9962 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.0267 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0209 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0174 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0161 - accuracy: 0.9981 - 2s/epoch - 7ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0128 - accuracy: 0.9981 - 2s/epoch - 7ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0115 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0131 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0090 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0081 - accuracy: 0.9987 - 2s/epoch - 7ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0093 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0070 - accuracy: 0.9990 - 2s/epoch - 9ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0077 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0058 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0087 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0078 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0046 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0052 - accuracy: 0.9991 - 2s/epoch - 7ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0085 - accuracy: 0.9990 - 2s/epoch - 7ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0061 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0031 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0031 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0031 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0041 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 35/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0063 - accuracy: 0.9992 - 2s/epoch - 7ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0035 - accuracy: 0.9996 - 2s/epoch - 7ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0033 - accuracy: 0.9996 - 2s/epoch - 7ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0047 - accuracy: 0.9992 - 2s/epoch - 7ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0068 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0052 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0092 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0069 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 2s - loss: 0.0036 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 0.0034 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0031 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 0.0047 - accuracy: 0.9992 - 2s/epoch - 9ms/step
Epoch 48/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 49/69
245/245 - 2s - loss: 0.0038 - accuracy: 0.9994 - 2s/epoch - 9ms/step
Epoch 50/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 0.0037 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 53/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 54/69
245/245 - 2s - loss: 0.0033 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0024 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0018 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0023 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 0.0025 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 62/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 63/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9997 - 2s/epoch - 9ms/step
Epoch 65/69
245/245 - 2s - loss: 6.7253e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 66/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 67/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 68/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 69/69
245/245 - 2s - loss: 0.0022 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Test Accuracy: 0.8821247220039368
```

[75]:
```
#14 Over Sampling_BoW 5000 feat
run_models(x_train21, x_test21, y_train21, y_test21, n_words21)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 70.75
Precision : 0.7075499850790808
Recall : 0.7075499850790808
F1-score : 0.7075499850790808
ROC_AOC_Score for Naive Bayes:  0.7874682150013476

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 85.91
Precision : 0.8591465234258431
Recall : 0.8591465234258431
F1-score : 0.8591465234258431
ROC_AOC_Score for Logistic Regression:  0.8873445171609696

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 87.35
Precision : 0.8734706057893166
Recall : 0.8734706057893166
F1-score : 0.8734706057893166

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 95.11
Precision : 0.9510593852581319
Recall : 0.9510593852581319
F1-score : 0.9510593852581319
ROC_AOC_Score for Random Forest:  0.9789164391427132

FOR ANN:

Model: "sequential_19"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_38 (Dense)            (None, 50)                250050

 dense_39 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.4991 - accuracy: 0.7630 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.2401 - accuracy: 0.9152 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.1027 - accuracy: 0.9763 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.0437 - accuracy: 0.9948 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.0214 - accuracy: 0.9974 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.0172 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.0113 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.0093 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.0093 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0080 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0098 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0089 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0110 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0082 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0065 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0087 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0059 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0053 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0049 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0052 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0062 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0040 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0028 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0033 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0049 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0048 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0035 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0047 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0049 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0051 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0031 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0031 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 35/69
245/245 - 2s - loss: 9.8424e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0029 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0023 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0029 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 2s - loss: 0.0026 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 0.0013 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0024 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 0.0021 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 48/69
245/245 - 2s - loss: 0.0054 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 49/69
245/245 - 2s - loss: 0.0042 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 0.0043 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 3s - loss: 7.9150e-04 - accuracy: 0.9997 - 3s/epoch - 11ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0013 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 53/69
245/245 - 3s - loss: 0.0012 - accuracy: 0.9996 - 3s/epoch - 11ms/step
Epoch 54/69
245/245 - 3s - loss: 0.0014 - accuracy: 0.9996 - 3s/epoch - 11ms/step
Epoch 55/69
245/245 - 3s - loss: 7.4521e-04 - accuracy: 0.9997 - 3s/epoch - 11ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 58/69
245/245 - 2s - loss: 7.8235e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 7.1117e-04 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 0.0011 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 62/69
245/245 - 2s - loss: 6.8873e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 2.9557e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 2.3058e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 4.0379e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 3.5414e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 67/69
245/245 - 2s - loss: 2.6203e-04 - accuracy: 0.9999 - 2s/epoch - 10ms/step
Epoch 68/69
245/245 - 2s - loss: 3.4616e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 69/69
245/245 - 2s - loss: 2.3008e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Test Accuracy: 0.8946583271026611
```

[76]:
```
#15 Over Sampling_Bag of n grams 1000 feat [bi,tri grams]
run_models(x_train22, x_test22, y_train22, y_test22, n_words22)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 70.13
Precision : 0.7012831990450612
Recall : 0.7012831990450612
F1-score : 0.7012831990450612
ROC_AOC_Score for Naive Bayes:  0.72574672775636

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 73.92
Precision : 0.7391823336317517
Recall : 0.7391823336317517
F1-score : 0.7391823336317517
ROC_AOC_Score for Logistic Regression:  0.7956133039993438

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 84.0
Precision : 0.8400477469412115
Recall : 0.8400477469412115
F1-score : 0.8400477469412115

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 90.81
Precision : 0.9080871381677111
Recall : 0.9080871381677111
F1-score : 0.9080871381677111
ROC_AOC_Score for Random Forest:  0.9683371245854766

FOR ANN:

Model: "sequential_20"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_40 (Dense)            (None, 50)                50050

 dense_41 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5975 - accuracy: 0.6978 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.4748 - accuracy: 0.7835 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.3997 - accuracy: 0.8342 - 2s/epoch - 7ms/step
Epoch 4/69
245/245 - 2s - loss: 0.3335 - accuracy: 0.8733 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.2759 - accuracy: 0.9077 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.2238 - accuracy: 0.9304 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.1806 - accuracy: 0.9506 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.1429 - accuracy: 0.9673 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.1136 - accuracy: 0.9770 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0908 - accuracy: 0.9835 - 2s/epoch - 8ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0727 - accuracy: 0.9880 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0596 - accuracy: 0.9900 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0487 - accuracy: 0.9922 - 2s/epoch - 8ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0408 - accuracy: 0.9930 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0344 - accuracy: 0.9942 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0297 - accuracy: 0.9950 - 2s/epoch - 7ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0259 - accuracy: 0.9950 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0225 - accuracy: 0.9957 - 2s/epoch - 8ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0200 - accuracy: 0.9958 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0183 - accuracy: 0.9963 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0152 - accuracy: 0.9971 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0138 - accuracy: 0.9977 - 2s/epoch - 7ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0125 - accuracy: 0.9976 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0116 - accuracy: 0.9978 - 2s/epoch - 10ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0108 - accuracy: 0.9978 - 2s/epoch - 9ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0100 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0094 - accuracy: 0.9981 - 2s/epoch - 9ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0086 - accuracy: 0.9980 - 2s/epoch - 10ms/step
Epoch 29/69
245/245 - 3s - loss: 0.0081 - accuracy: 0.9983 - 3s/epoch - 10ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0076 - accuracy: 0.9985 - 2s/epoch - 10ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0100 - accuracy: 0.9974 - 2s/epoch - 10ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0078 - accuracy: 0.9976 - 2s/epoch - 10ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0074 - accuracy: 0.9981 - 2s/epoch - 10ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0065 - accuracy: 0.9985 - 2s/epoch - 9ms/step
Epoch 35/69
245/245 - 2s - loss: 0.0061 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0057 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0055 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0056 - accuracy: 0.9985 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0057 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0062 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0053 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0051 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0052 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 3s - loss: 0.0051 - accuracy: 0.9983 - 3s/epoch - 11ms/step
Epoch 45/69
245/245 - 2s - loss: 0.0046 - accuracy: 0.9986 - 2s/epoch - 10ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0048 - accuracy: 0.9982 - 2s/epoch - 10ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 0.0051 - accuracy: 0.9982 - 2s/epoch - 9ms/step
Epoch 48/69
245/245 - 2s - loss: 0.0045 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 49/69
245/245 - 2s - loss: 0.0048 - accuracy: 0.9982 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 0.0044 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 0.0045 - accuracy: 0.9982 - 2s/epoch - 8ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0044 - accuracy: 0.9985 - 2s/epoch - 8ms/step
Epoch 53/69
245/245 - 2s - loss: 0.0042 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 54/69
245/245 - 2s - loss: 0.0041 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0046 - accuracy: 0.9982 - 2s/epoch - 8ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0047 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0045 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 0.0044 - accuracy: 0.9985 - 2s/epoch - 7ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0044 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 0.0059 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 0.0042 - accuracy: 0.9983 - 2s/epoch - 7ms/step
Epoch 62/69
245/245 - 2s - loss: 0.0044 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 0.0048 - accuracy: 0.9983 - 2s/epoch - 7ms/step
Epoch 64/69
245/245 - 2s - loss: 0.0042 - accuracy: 0.9985 - 2s/epoch - 7ms/step
Epoch 65/69
245/245 - 2s - loss: 0.0043 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 0.0043 - accuracy: 0.9982 - 2s/epoch - 8ms/step
Epoch 67/69
245/245 - 2s - loss: 0.0042 - accuracy: 0.9985 - 2s/epoch - 9ms/step
Epoch 68/69
245/245 - 2s - loss: 0.0043 - accuracy: 0.9986 - 2s/epoch - 9ms/step
Epoch 69/69
245/245 - 2s - loss: 0.0040 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Test Accuracy: 0.8949567079544067
```

[77]:
```
#16 Over Sampling_Bag of n grams 5000 feat [bi,tri grams]
run_models(x_train23, x_test23, y_train23, y_test23, n_words23)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 73.71
Precision : 0.7370934049537452
Recall : 0.7370934049537452
F1-score : 0.7370934049537452
ROC_AOC_Score for Naive Bayes:  0.7920367592777042

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 84.99
Precision : 0.8498955535660997
Recall : 0.8498955535660997
F1-score : 0.8498955535660998
ROC_AOC_Score for Logistic Regression:  0.9004016334852764

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 87.71
Precision : 0.877051626380185
Recall : 0.877051626380185
F1-score : 0.8770516263801849

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 91.73
Precision : 0.9173381080274545
Recall : 0.9173381080274545
F1-score : 0.9173381080274545
ROC_AOC_Score for Random Forest:  0.9693192371014425

FOR ANN:

Model: "sequential_21"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_42 (Dense)            (None, 50)                250050

 dense_43 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5506 - accuracy: 0.7376 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.2943 - accuracy: 0.8999 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.1547 - accuracy: 0.9572 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.0797 - accuracy: 0.9850 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.0430 - accuracy: 0.9933 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.0265 - accuracy: 0.9965 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.0176 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.0131 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.0108 - accuracy: 0.9982 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0082 - accuracy: 0.9985 - 2s/epoch - 8ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0065 - accuracy: 0.9990 - 2s/epoch - 10ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0063 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0048 - accuracy: 0.9991 - 2s/epoch - 9ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0035 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0032 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0024 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0012 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0011 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 9.8374e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 26/69
245/245 - 2s - loss: 9.3497e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 27/69
245/245 - 2s - loss: 8.7471e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 28/69
245/245 - 2s - loss: 8.3675e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 7.9438e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 30/69
245/245 - 2s - loss: 7.4753e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 31/69
245/245 - 2s - loss: 8.2676e-04 - accuracy: 0.9997 - 2s/epoch - 9ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0029 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0021 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 35/69
245/245 - 2s - loss: 6.4141e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 36/69
245/245 - 2s - loss: 6.2652e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 37/69
245/245 - 2s - loss: 6.0062e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 38/69
245/245 - 2s - loss: 5.9039e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 5.8075e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 40/69
245/245 - 2s - loss: 5.6467e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 41/69
245/245 - 2s - loss: 5.6159e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 5.5365e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 43/69
245/245 - 2s - loss: 5.4241e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 2s - loss: 5.4175e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 5.3278e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 46/69
245/245 - 2s - loss: 5.2428e-04 - accuracy: 0.9999 - 2s/epoch - 10ms/step
Epoch 47/69
```

```
245/245 - 3s - loss: 5.2635e-04 - accuracy: 0.9999 - 3s/epoch - 11ms/step
Epoch 48/69
245/245 - 2s - loss: 5.2336e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 49/69
245/245 - 2s - loss: 5.1763e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 50/69
245/245 - 2s - loss: 5.1516e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 51/69
245/245 - 2s - loss: 5.1846e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 52/69
245/245 - 2s - loss: 5.1353e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 53/69
245/245 - 2s - loss: 5.0537e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 54/69
245/245 - 2s - loss: 5.1463e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 55/69
245/245 - 2s - loss: 5.4203e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0021 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 5.4191e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 59/69
245/245 - 2s - loss: 4.9832e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 60/69
245/245 - 2s - loss: 4.9582e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 4.9471e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 62/69
245/245 - 2s - loss: 5.0020e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 4.9023e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 5.0047e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 4.9589e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 4.9095e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 67/69
245/245 - 2s - loss: 4.9787e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 68/69
245/245 - 2s - loss: 4.9937e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 69/69
245/245 - 2s - loss: 4.9574e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Test Accuracy: 0.8848105072975159
```

[78]:
```
#17 Over Sampling_TF-IDF 1000 feat
run_models(x_train24, x_test24, y_train24, y_test24, n_words24)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 68.64
Precision : 0.6863622799164428
Recall : 0.6863622799164428
F1-score : 0.6863622799164428
ROC_AOC_Score for Naive Bayes:  0.74403262283364

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 72.55
Precision : 0.7254550880334228
Recall : 0.7254550880334228
F1-score : 0.7254550880334228
ROC_AOC_Score for Logistic Regression:  0.8009226438088096

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 88.21
Precision : 0.8821247388839153
Recall : 0.8821247388839153
F1-score : 0.8821247388839152

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 93.82
Precision : 0.9382273948075202
Recall : 0.9382273948075202
F1-score : 0.9382273948075202
ROC_AOC_Score for Random Forest:  0.9720008217227761

FOR ANN:

Model: "sequential_22"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_44 (Dense)            (None, 50)                50050

 dense_45 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5950 - accuracy: 0.6809 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.5097 - accuracy: 0.7532 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.4711 - accuracy: 0.7793 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.4451 - accuracy: 0.7963 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.4234 - accuracy: 0.8060 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.4048 - accuracy: 0.8172 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.3867 - accuracy: 0.8235 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.3671 - accuracy: 0.8387 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.3480 - accuracy: 0.8474 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.3279 - accuracy: 0.8593 - 2s/epoch - 7ms/step
Epoch 11/69
245/245 - 2s - loss: 0.3076 - accuracy: 0.8725 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.2848 - accuracy: 0.8907 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.2633 - accuracy: 0.9018 - 2s/epoch - 8ms/step
Epoch 14/69
245/245 - 2s - loss: 0.2400 - accuracy: 0.9203 - 2s/epoch - 7ms/step
Epoch 15/69
245/245 - 2s - loss: 0.2181 - accuracy: 0.9332 - 2s/epoch - 7ms/step
Epoch 16/69
245/245 - 2s - loss: 0.1965 - accuracy: 0.9439 - 2s/epoch - 7ms/step
Epoch 17/69
245/245 - 2s - loss: 0.1742 - accuracy: 0.9538 - 2s/epoch - 7ms/step
Epoch 18/69
245/245 - 2s - loss: 0.1549 - accuracy: 0.9633 - 2s/epoch - 7ms/step
Epoch 19/69
245/245 - 2s - loss: 0.1355 - accuracy: 0.9719 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.1180 - accuracy: 0.9783 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.1020 - accuracy: 0.9841 - 2s/epoch - 7ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0881 - accuracy: 0.9887 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0756 - accuracy: 0.9919 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0650 - accuracy: 0.9941 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0558 - accuracy: 0.9954 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0478 - accuracy: 0.9964 - 2s/epoch - 7ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0410 - accuracy: 0.9972 - 2s/epoch - 7ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0353 - accuracy: 0.9976 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0304 - accuracy: 0.9982 - 2s/epoch - 7ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0269 - accuracy: 0.9983 - 2s/epoch - 7ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0241 - accuracy: 0.9985 - 2s/epoch - 8ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0201 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0175 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0155 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 35/69
245/245 - 2s - loss: 0.0137 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0136 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0113 - accuracy: 0.9987 - 2s/epoch - 9ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0103 - accuracy: 0.9987 - 2s/epoch - 7ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0093 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0082 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0073 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0068 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0062 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 2s - loss: 0.0062 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 45/69
245/245 - 2s - loss: 0.0060 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0045 - accuracy: 0.9994 - 2s/epoch - 7ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 0.0037 - accuracy: 0.9995 - 2s/epoch - 7ms/step
Epoch 48/69
245/245 - 2s - loss: 0.0048 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 49/69
245/245 - 2s - loss: 0.0043 - accuracy: 0.9991 - 2s/epoch - 7ms/step
Epoch 50/69
245/245 - 2s - loss: 0.0034 - accuracy: 0.9995 - 2s/epoch - 7ms/step
Epoch 51/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9996 - 2s/epoch - 7ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0033 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 53/69
245/245 - 2s - loss: 0.0025 - accuracy: 0.9997 - 2s/epoch - 7ms/step
Epoch 54/69
245/245 - 2s - loss: 0.0024 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0041 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9992 - 2s/epoch - 9ms/step
Epoch 58/69
245/245 - 2s - loss: 0.0023 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0026 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Epoch 61/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9997 - 2s/epoch - 7ms/step
Epoch 62/69
245/245 - 2s - loss: 0.0024 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 8.2941e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 7.3869e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 67/69
245/245 - 2s - loss: 0.0010 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 68/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 69/69
245/245 - 2s - loss: 0.0024 - accuracy: 0.9994 - 2s/epoch - 7ms/step
Test Accuracy: 0.8755595088005066
```

[79]:
```
#18 Over Sampling_TF-IDF 5000 feat
run_models(x_train25, x_test25, y_train25, y_test25, n_words25)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 70.93
Precision : 0.709340495374515
Recall : 0.709340495374515
F1-score : 0.709340495374515
ROC_AOC_Score for Naive Bayes:  0.7892094323814434

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 76.78
Precision : 0.7678304983586989
Recall : 0.7678304983586989
F1-score : 0.7678304983586989
ROC_AOC_Score for Logistic Regression:  0.8581133039993438

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 90.33
Precision : 0.9033124440465533
Recall : 0.9033124440465533
F1-score : 0.9033124440465533

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 93.26
Precision : 0.9325574455386452
Recall : 0.9325574455386452
F1-score : 0.9325574455386452
ROC_AOC_Score for Random Forest:  0.9727244064847257

FOR ANN:

Model: "sequential_23"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_46 (Dense)            (None, 50)                250050

 dense_47 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5838 - accuracy: 0.6929 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.4158 - accuracy: 0.8163 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.3020 - accuracy: 0.8894 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.2261 - accuracy: 0.9279 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.1728 - accuracy: 0.9492 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.1303 - accuracy: 0.9659 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.0972 - accuracy: 0.9795 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.0731 - accuracy: 0.9894 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.0540 - accuracy: 0.9933 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0409 - accuracy: 0.9964 - 2s/epoch - 8ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0309 - accuracy: 0.9985 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0240 - accuracy: 0.9988 - 2s/epoch - 9ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0190 - accuracy: 0.9990 - 2s/epoch - 9ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0148 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0122 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0102 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0083 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0072 - accuracy: 0.9994 - 2s/epoch - 10ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0064 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0059 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0055 - accuracy: 0.9992 - 2s/epoch - 9ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0044 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0035 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0035 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0025 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0034 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0022 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 35/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0025 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 9.5194e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0011 - accuracy: 0.9997 - 2s/epoch - 9ms/step
Epoch 41/69
245/245 - 2s - loss: 6.4017e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 7.4642e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0011 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 2s - loss: 5.7454e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 9.3167e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 4.3404e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 5.7786e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 48/69
245/245 - 2s - loss: 5.1290e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 49/69
245/245 - 2s - loss: 6.3467e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 5.0030e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 6.7115e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0010 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 53/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 54/69
245/245 - 2s - loss: 5.8183e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 56/69
245/245 - 2s - loss: 2.9162e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 5.3446e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 9.9882e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 7.2626e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 4.5744e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 5s - loss: 7.6290e-05 - accuracy: 1.0000 - 5s/epoch - 20ms/step
Epoch 62/69
245/245 - 2s - loss: 4.0056e-05 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 63/69
245/245 - 2s - loss: 3.6126e-05 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 3.4082e-05 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 3.2040e-05 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 3.3166e-05 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 67/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9997 - 2s/epoch - 9ms/step
Epoch 68/69
245/245 - 2s - loss: 2.2639e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 69/69
245/245 - 2s - loss: 7.2051e-05 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Test Accuracy: 0.8683974742889404
```

[80]:
```
#19 SMOTE BoW 1000 feat
run_models(x_train26, x_test26, y_train26, y_test26, n_words26)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 71.44
Precision : 0.7144136078782453
Recall : 0.7144136078782453
F1-score : 0.7144136078782454
ROC_AOC_Score for Naive Bayes:  0.7707796259623385

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 81.17
Precision : 0.8116980005968367
Recall : 0.8116980005968367
F1-score : 0.8116980005968367
ROC_AOC_Score for Logistic Regression:  0.8813771399946098

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 86.66
Precision : 0.8666069829901522
Recall : 0.8666069829901522
F1-score : 0.8666069829901522

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 86.21
Precision : 0.8621307072515667
Recall : 0.8621307072515667
F1-score : 0.8621307072515667
ROC_AOC_Score for Random Forest:  0.9147530364194565

FOR ANN:

Model: "sequential_24"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_48 (Dense)            (None, 50)                50050

 dense_49 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 4s - loss: 0.4584 - accuracy: 0.8185 - 4s/epoch - 17ms/step
Epoch 2/69
245/245 - 2s - loss: 0.3125 - accuracy: 0.8779 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.2419 - accuracy: 0.9054 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.1877 - accuracy: 0.9340 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.1373 - accuracy: 0.9595 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.0987 - accuracy: 0.9766 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.0680 - accuracy: 0.9875 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.0471 - accuracy: 0.9940 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.0335 - accuracy: 0.9965 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0246 - accuracy: 0.9973 - 2s/epoch - 9ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0192 - accuracy: 0.9976 - 2s/epoch - 9ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0154 - accuracy: 0.9982 - 2s/epoch - 10ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0128 - accuracy: 0.9981 - 2s/epoch - 9ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0132 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0104 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0094 - accuracy: 0.9982 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0081 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0086 - accuracy: 0.9985 - 2s/epoch - 8ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0084 - accuracy: 0.9986 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0051 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0049 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0058 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0043 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0039 - accuracy: 0.9990 - 2s/epoch - 9ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0035 - accuracy: 0.9992 - 2s/epoch - 9ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0031 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0029 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0023 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9995 - 2s/epoch - 10ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9996 - 2s/epoch - 10ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 8.3365e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 35/69
245/245 - 2s - loss: 9.6194e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0011 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0021 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9995 - 2s/epoch - 7ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0064 - accuracy: 0.9994 - 2s/epoch - 7ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0078 - accuracy: 0.9973 - 2s/epoch - 7ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9996 - 2s/epoch - 7ms/step
Epoch 44/69
245/245 - 2s - loss: 8.4205e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 4.8370e-04 - accuracy: 0.9997 - 2s/epoch - 7ms/step
Epoch 48/69
245/245 - 2s - loss: 2.6712e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 49/69
245/245 - 2s - loss: 2.4621e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 2.2986e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 2.1641e-04 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Epoch 52/69
245/245 - 2s - loss: 2.0424e-04 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Epoch 53/69
245/245 - 2s - loss: 1.9412e-04 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Epoch 54/69
245/245 - 2s - loss: 1.8568e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 55/69
245/245 - 2s - loss: 1.7709e-04 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Epoch 56/69
245/245 - 2s - loss: 1.6887e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 1.6220e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 1.6095e-04 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0018 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 1.5696e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 1.4406e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 62/69
245/245 - 2s - loss: 1.3731e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 1.3227e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 1.2745e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 1.2354e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 1.3686e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 67/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 68/69
245/245 - 2s - loss: 3.6459e-04 - accuracy: 0.9997 - 2s/epoch - 9ms/step
Epoch 69/69
245/245 - 2s - loss: 4.3847e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Test Accuracy: 0.8433303236961365
```

[81]:
```
#20 SMOTE BoW 5000 feat
run_models(x_train27, x_test27, y_train27, y_test27, n_words27)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 75.26
Precision : 0.7526111608475082
Recall : 0.7526111608475082
F1-score : 0.7526111608475082
ROC_AOC_Score for Naive Bayes:  0.778436405102005

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 81.77
Precision : 0.8176663682482841
Recall : 0.8176663682482841
F1-score : 0.8176663682482841
ROC_AOC_Score for Logistic Regression:  0.880901097388064

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 86.93
Precision : 0.8692927484333035
Recall : 0.8692927484333035
F1-score : 0.8692927484333035

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 86.63
Precision : 0.8663085646075799
Recall : 0.8663085646075799
F1-score : 0.8663085646075799
ROC_AOC_Score for Random Forest:  0.9098615155438896

FOR ANN:

Model: "sequential_25"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_50 (Dense)            (None, 50)                250050

 dense_51 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.4276 - accuracy: 0.8217 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.2214 - accuracy: 0.9138 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.1137 - accuracy: 0.9642 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.0530 - accuracy: 0.9890 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.0263 - accuracy: 0.9968 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.0161 - accuracy: 0.9981 - 2s/epoch - 9ms/step
Epoch 7/69
245/245 - 2s - loss: 0.0115 - accuracy: 0.9983 - 2s/epoch - 9ms/step
Epoch 8/69
245/245 - 2s - loss: 0.0097 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.0087 - accuracy: 0.9985 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0072 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0071 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0061 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0064 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0058 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0059 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0051 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0035 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0036 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0032 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0037 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0058 - accuracy: 0.9994 - 2s/epoch - 9ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0021 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0019 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0021 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0018 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0040 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0011 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0012 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 35/69
245/245 - 2s - loss: 5.9474e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 5.4591e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 37/69
245/245 - 2s - loss: 5.2392e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 38/69
245/245 - 2s - loss: 5.0584e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 4.8764e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 40/69
245/245 - 2s - loss: 4.7316e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 41/69
245/245 - 2s - loss: 4.5380e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 4.3784e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 43/69
245/245 - 2s - loss: 4.2235e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 2s - loss: 4.0768e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 3.9347e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 3.8051e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 3.6569e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 48/69
245/245 - 2s - loss: 3.5370e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 49/69
245/245 - 2s - loss: 3.4034e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 3.3023e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 3.2630e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0011 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 53/69
245/245 - 2s - loss: 0.0023 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 54/69
245/245 - 2s - loss: 7.1626e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0024 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0086 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0045 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 3.5124e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 2.9786e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 2.8510e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 62/69
245/245 - 2s - loss: 2.7596e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 2.6843e-04 - accuracy: 1.0000 - 2s/epoch - 9ms/step
Epoch 64/69
245/245 - 2s - loss: 2.6122e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 2.5655e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 2.4778e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 67/69
245/245 - 2s - loss: 2.4193e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 68/69
245/245 - 2s - loss: 2.3516e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Epoch 69/69
245/245 - 2s - loss: 2.2787e-04 - accuracy: 1.0000 - 2s/epoch - 8ms/step
Test Accuracy: 0.8427335023880005
```

[82]:
```
#21 SMOTE Bag of n grams 1000 feat [bi,tri grams]
run_models(x_train28, x_test28, y_train28, y_test28, n_words28)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 66.07
Precision : 0.6606982990152194
Recall : 0.6606982990152194
F1-score : 0.6606982990152194
ROC_AOC_Score for Naive Bayes:  0.7673964277762806

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 77.68
Precision : 0.7767830498358699
Recall : 0.7767830498358699
F1-score : 0.77678304983587
ROC_AOC_Score for Logistic Regression:  0.818328592437221

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 79.8
Precision : 0.7979707549985079
Recall : 0.7979707549985079
F1-score : 0.7979707549985079

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 81.11
Precision : 0.8111011638316921
Recall : 0.8111011638316921
F1-score : 0.8111011638316922
ROC_AOC_Score for Random Forest:  0.8733344001570209

FOR ANN:

Model: "sequential_26"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_52 (Dense)            (None, 50)                50050

 dense_53 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5836 - accuracy: 0.7003 - 3s/epoch - 11ms/step
Epoch 2/69
245/245 - 2s - loss: 0.4657 - accuracy: 0.8147 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.4027 - accuracy: 0.8344 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.3567 - accuracy: 0.8574 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.3172 - accuracy: 0.8782 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.2812 - accuracy: 0.8964 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.2488 - accuracy: 0.9161 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.2183 - accuracy: 0.9303 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.1917 - accuracy: 0.9414 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.1683 - accuracy: 0.9499 - 2s/epoch - 10ms/step
Epoch 11/69
245/245 - 2s - loss: 0.1483 - accuracy: 0.9592 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.1311 - accuracy: 0.9639 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.1178 - accuracy: 0.9693 - 2s/epoch - 8ms/step
Epoch 14/69
245/245 - 2s - loss: 0.1056 - accuracy: 0.9725 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0964 - accuracy: 0.9752 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0879 - accuracy: 0.9765 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0816 - accuracy: 0.9789 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0758 - accuracy: 0.9808 - 2s/epoch - 8ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0718 - accuracy: 0.9806 - 2s/epoch - 7ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0676 - accuracy: 0.9817 - 2s/epoch - 7ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0648 - accuracy: 0.9812 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0618 - accuracy: 0.9824 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0596 - accuracy: 0.9831 - 2s/epoch - 9ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0574 - accuracy: 0.9824 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0557 - accuracy: 0.9827 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0545 - accuracy: 0.9825 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0526 - accuracy: 0.9838 - 2s/epoch - 8ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0518 - accuracy: 0.9830 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0507 - accuracy: 0.9834 - 2s/epoch - 8ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0495 - accuracy: 0.9840 - 2s/epoch - 8ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0484 - accuracy: 0.9838 - 2s/epoch - 9ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0477 - accuracy: 0.9844 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0470 - accuracy: 0.9840 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0462 - accuracy: 0.9844 - 2s/epoch - 8ms/step
Epoch 35/69
245/245 - 2s - loss: 0.0458 - accuracy: 0.9850 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0449 - accuracy: 0.9848 - 2s/epoch - 8ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0445 - accuracy: 0.9843 - 2s/epoch - 8ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0443 - accuracy: 0.9847 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0434 - accuracy: 0.9848 - 2s/epoch - 9ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0434 - accuracy: 0.9852 - 2s/epoch - 8ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0431 - accuracy: 0.9857 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0426 - accuracy: 0.9852 - 2s/epoch - 8ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0428 - accuracy: 0.9852 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 2s - loss: 0.0423 - accuracy: 0.9852 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 0.0417 - accuracy: 0.9853 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0414 - accuracy: 0.9854 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 0.0414 - accuracy: 0.9852 - 2s/epoch - 8ms/step
Epoch 48/69
245/245 - 2s - loss: 0.0409 - accuracy: 0.9858 - 2s/epoch - 8ms/step
Epoch 49/69
245/245 - 2s - loss: 0.0411 - accuracy: 0.9858 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 0.0408 - accuracy: 0.9857 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 0.0407 - accuracy: 0.9852 - 2s/epoch - 8ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0404 - accuracy: 0.9859 - 2s/epoch - 9ms/step
Epoch 53/69
245/245 - 2s - loss: 0.0409 - accuracy: 0.9855 - 2s/epoch - 8ms/step
Epoch 54/69
245/245 - 2s - loss: 0.0405 - accuracy: 0.9857 - 2s/epoch - 8ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0402 - accuracy: 0.9857 - 2s/epoch - 8ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0397 - accuracy: 0.9858 - 2s/epoch - 9ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0403 - accuracy: 0.9859 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 0.0399 - accuracy: 0.9855 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0397 - accuracy: 0.9854 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 0.0395 - accuracy: 0.9862 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 0.0396 - accuracy: 0.9862 - 2s/epoch - 8ms/step
Epoch 62/69
245/245 - 2s - loss: 0.0393 - accuracy: 0.9862 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 0.0396 - accuracy: 0.9853 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 0.0396 - accuracy: 0.9863 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 0.0398 - accuracy: 0.9857 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 0.0395 - accuracy: 0.9861 - 2s/epoch - 8ms/step
Epoch 67/69
245/245 - 2s - loss: 0.0394 - accuracy: 0.9858 - 2s/epoch - 9ms/step
Epoch 68/69
245/245 - 2s - loss: 0.0394 - accuracy: 0.9864 - 2s/epoch - 9ms/step
Epoch 69/69
245/245 - 2s - loss: 0.0393 - accuracy: 0.9864 - 2s/epoch - 8ms/step
Test Accuracy: 0.8024470210075378
```

[83]:
```
#22 SMOTE Bag of n grams 5000 feat [bi,tri grams]
run_models(x_train29, x_test29, y_train29, y_test29, n_words29)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 63.83
Precision : 0.6383169203222918
Recall : 0.6383169203222918
F1-score : 0.6383169203222918
ROC_AOC_Score for Naive Bayes:  0.7699203690575235

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 80.27
Precision : 0.8027454491196657
Recall : 0.8027454491196657
F1-score : 0.8027454491196657
ROC_AOC_Score for Logistic Regression:  0.8464136414769332

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 83.65
Precision : 0.8364667263503431
Recall : 0.8364667263503431
F1-score : 0.8364667263503431

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 80.51
Precision : 0.8051327961802447
Recall : 0.8051327961802447
F1-score : 0.8051327961802448
ROC_AOC_Score for Random Forest:  0.8777590330915526

FOR ANN:

Model: "sequential_27"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_54 (Dense)            (None, 50)                250050

 dense_55 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5638 - accuracy: 0.7177 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.3460 - accuracy: 0.8817 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.2069 - accuracy: 0.9320 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.1219 - accuracy: 0.9661 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.0760 - accuracy: 0.9817 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.0504 - accuracy: 0.9902 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.0370 - accuracy: 0.9921 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.0293 - accuracy: 0.9948 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.0239 - accuracy: 0.9953 - 2s/epoch - 9ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0206 - accuracy: 0.9965 - 2s/epoch - 9ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0190 - accuracy: 0.9969 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0180 - accuracy: 0.9968 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0159 - accuracy: 0.9974 - 2s/epoch - 8ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0152 - accuracy: 0.9976 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0155 - accuracy: 0.9973 - 2s/epoch - 9ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0143 - accuracy: 0.9976 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0139 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0136 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0133 - accuracy: 0.9976 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0129 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0132 - accuracy: 0.9974 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0127 - accuracy: 0.9974 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0122 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0126 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0117 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0115 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0117 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0119 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0113 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0113 - accuracy: 0.9980 - 2s/epoch - 9ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0111 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0110 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0112 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0110 - accuracy: 0.9980 - 2s/epoch - 9ms/step
Epoch 35/69
245/245 - 2s - loss: 0.0110 - accuracy: 0.9981 - 2s/epoch - 9ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0108 - accuracy: 0.9978 - 2s/epoch - 10ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0109 - accuracy: 0.9980 - 2s/epoch - 9ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0108 - accuracy: 0.9980 - 2s/epoch - 9ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0108 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0108 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0107 - accuracy: 0.9980 - 2s/epoch - 9ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0135 - accuracy: 0.9977 - 2s/epoch - 9ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0108 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 44/69
245/245 - 2s - loss: 0.0109 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 3s - loss: 0.0109 - accuracy: 0.9980 - 3s/epoch - 12ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0107 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 0.0105 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 48/69
245/245 - 2s - loss: 0.0106 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 49/69
245/245 - 2s - loss: 0.0106 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 0.0106 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 53/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9981 - 2s/epoch - 10ms/step
Epoch 54/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0104 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0107 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 0.0136 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 0.0110 - accuracy: 0.9977 - 2s/epoch - 9ms/step
Epoch 62/69
245/245 - 2s - loss: 0.0111 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 0.0108 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 0.0118 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 0.0107 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 66/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9978 - 2s/epoch - 8ms/step
Epoch 67/69
245/245 - 2s - loss: 0.0104 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Epoch 68/69
245/245 - 2s - loss: 0.0104 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 69/69
245/245 - 2s - loss: 0.0105 - accuracy: 0.9981 - 2s/epoch - 8ms/step
Test Accuracy: 0.8191584348678589
```

[84]:
```
#23 SMOTE TF-IDF 1000 feat
run_models(x_train30, x_test30, y_train30, y_test30, n_words30)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 70.37
Precision : 0.7036705461056401
Recall : 0.7036705461056401
F1-score : 0.7036705461056401
ROC_AOC_Score for Naive Bayes:  0.7677552906642918

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 74.81
Precision : 0.7481348851089227
Recall : 0.7481348851089227
F1-score : 0.7481348851089226
ROC_AOC_Score for Logistic Regression:  0.8224086437619376

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 89.35
Precision : 0.8934646374216652
Recall : 0.8934646374216652
F1-score : 0.8934646374216652

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 89.29
Precision : 0.8928678006565205
Recall : 0.8928678006565205
F1-score : 0.8928678006565205
ROC_AOC_Score for Random Forest:  0.9479630503052532

FOR ANN:

Model: "sequential_28"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_56 (Dense)            (None, 50)                50050

 dense_57 (Dense)            (None, 1)                 51

=================================================================
Total params: 50,101
```

```
Trainable params: 50,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5813 - accuracy: 0.6868 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.4886 - accuracy: 0.7615 - 2s/epoch - 7ms/step
Epoch 3/69
245/245 - 2s - loss: 0.4494 - accuracy: 0.7883 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.4232 - accuracy: 0.8025 - 2s/epoch - 7ms/step
Epoch 5/69
245/245 - 2s - loss: 0.4023 - accuracy: 0.8186 - 2s/epoch - 7ms/step
Epoch 6/69
245/245 - 2s - loss: 0.3836 - accuracy: 0.8239 - 2s/epoch - 7ms/step
Epoch 7/69
245/245 - 2s - loss: 0.3658 - accuracy: 0.8374 - 2s/epoch - 7ms/step
Epoch 8/69
245/245 - 2s - loss: 0.3473 - accuracy: 0.8493 - 2s/epoch - 7ms/step
Epoch 9/69
245/245 - 2s - loss: 0.3282 - accuracy: 0.8634 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.3089 - accuracy: 0.8742 - 2s/epoch - 8ms/step
Epoch 11/69
245/245 - 2s - loss: 0.2887 - accuracy: 0.8871 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.2678 - accuracy: 0.9011 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.2464 - accuracy: 0.9169 - 2s/epoch - 8ms/step
Epoch 14/69
245/245 - 2s - loss: 0.2241 - accuracy: 0.9277 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.2021 - accuracy: 0.9386 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.1812 - accuracy: 0.9511 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.1620 - accuracy: 0.9582 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.1420 - accuracy: 0.9682 - 2s/epoch - 7ms/step
Epoch 19/69
245/245 - 2s - loss: 0.1244 - accuracy: 0.9760 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.1083 - accuracy: 0.9825 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0936 - accuracy: 0.9862 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0814 - accuracy: 0.9902 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0697 - accuracy: 0.9932 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0599 - accuracy: 0.9944 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0512 - accuracy: 0.9964 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0443 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0380 - accuracy: 0.9977 - 2s/epoch - 8ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0328 - accuracy: 0.9980 - 2s/epoch - 8ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0282 - accuracy: 0.9983 - 2s/epoch - 8ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0242 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0213 - accuracy: 0.9986 - 2s/epoch - 7ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0188 - accuracy: 0.9983 - 2s/epoch - 7ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0165 - accuracy: 0.9986 - 2s/epoch - 7ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0147 - accuracy: 0.9983 - 2s/epoch - 7ms/step
Epoch 35/69
245/245 - 2s - loss: 0.0125 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0113 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0103 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0091 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0085 - accuracy: 0.9988 - 2s/epoch - 7ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0078 - accuracy: 0.9986 - 2s/epoch - 7ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0066 - accuracy: 0.9990 - 2s/epoch - 7ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0062 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0059 - accuracy: 0.9987 - 2s/epoch - 7ms/step
Epoch 44/69
245/245 - 2s - loss: 0.0053 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 0.0044 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0037 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 0.0036 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 48/69
245/245 - 2s - loss: 0.0031 - accuracy: 0.9995 - 2s/epoch - 7ms/step
Epoch 49/69
245/245 - 2s - loss: 0.0039 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 52/69
245/245 - 2s - loss: 0.0029 - accuracy: 0.9994 - 2s/epoch - 9ms/step
Epoch 53/69
245/245 - 2s - loss: 0.0029 - accuracy: 0.9994 - 2s/epoch - 9ms/step
Epoch 54/69
245/245 - 2s - loss: 0.0022 - accuracy: 0.9994 - 2s/epoch - 9ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9997 - 2s/epoch - 10ms/step
Epoch 56/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 58/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0018 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 61/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 62/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 63/69
245/245 - 2s - loss: 0.0018 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 64/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 0.0023 - accuracy: 0.9992 - 2s/epoch - 7ms/step
Epoch 66/69
245/245 - 2s - loss: 0.0013 - accuracy: 0.9995 - 2s/epoch - 7ms/step
Epoch 67/69
245/245 - 2s - loss: 5.8824e-04 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Epoch 68/69
245/245 - 2s - loss: 4.8561e-04 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Epoch 69/69
245/245 - 2s - loss: 4.6013e-04 - accuracy: 0.9999 - 2s/epoch - 7ms/step
Test Accuracy: 0.8675022125244141
```

[85]:
```
#24 SMOTE TF-IDF 5000 feat
run_models(x_train31, x_test31, y_train31, y_test31, n_words31)
```

```
FOR NAIVE BAYES:

Test Accuracy Score of Basic Naive Bayes Model: 72.37
Precision : 0.7236645777379886
Recall : 0.7236645777379886
F1-score : 0.7236645777379885
ROC_AOC_Score for Naive Bayes:  0.8041893946495741

FOR LOGISTIC REGRESSION:

Test Accuracy Score of Basic Logistic Regression Model: 78.66
Precision : 0.786630856460758
Recall : 0.786630856460758
F1-score : 0.786630856460758
ROC_AOC_Score for Logistic Regression:  0.8712993180140383

FOR LINEAR SVC:

Test Accuracy Score of Basic Linear SVC Model: 89.67
Precision : 0.8967472396299612
Recall : 0.8967472396299612
F1-score : 0.8967472396299612

FOR RANDOM FOREST:

Test Accuracy Score of Basic Random Forest Model: 88.87
Precision : 0.8886899433005073
Recall : 0.8886899433005073
F1-score : 0.8886899433005073
ROC_AOC_Score for Random Forest:  0.9517036466328408

FOR ANN:

Model: "sequential_29"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_58 (Dense)            (None, 50)                250050

 dense_59 (Dense)            (None, 1)                 51

=================================================================
Total params: 250,101
```

```
Trainable params: 250,101
Non-trainable params: 0
_________________________________________________________________
Epoch 1/69
245/245 - 3s - loss: 0.5504 - accuracy: 0.7229 - 3s/epoch - 12ms/step
Epoch 2/69
245/245 - 2s - loss: 0.3895 - accuracy: 0.8302 - 2s/epoch - 8ms/step
Epoch 3/69
245/245 - 2s - loss: 0.2908 - accuracy: 0.8922 - 2s/epoch - 8ms/step
Epoch 4/69
245/245 - 2s - loss: 0.2240 - accuracy: 0.9233 - 2s/epoch - 8ms/step
Epoch 5/69
245/245 - 2s - loss: 0.1750 - accuracy: 0.9445 - 2s/epoch - 8ms/step
Epoch 6/69
245/245 - 2s - loss: 0.1353 - accuracy: 0.9646 - 2s/epoch - 8ms/step
Epoch 7/69
245/245 - 2s - loss: 0.1053 - accuracy: 0.9760 - 2s/epoch - 8ms/step
Epoch 8/69
245/245 - 2s - loss: 0.0815 - accuracy: 0.9844 - 2s/epoch - 8ms/step
Epoch 9/69
245/245 - 2s - loss: 0.0620 - accuracy: 0.9903 - 2s/epoch - 8ms/step
Epoch 10/69
245/245 - 2s - loss: 0.0473 - accuracy: 0.9954 - 2s/epoch - 8ms/step
Epoch 11/69
245/245 - 2s - loss: 0.0362 - accuracy: 0.9971 - 2s/epoch - 8ms/step
Epoch 12/69
245/245 - 2s - loss: 0.0280 - accuracy: 0.9987 - 2s/epoch - 8ms/step
Epoch 13/69
245/245 - 2s - loss: 0.0223 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 14/69
245/245 - 2s - loss: 0.0173 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 15/69
245/245 - 2s - loss: 0.0147 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 16/69
245/245 - 2s - loss: 0.0119 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 17/69
245/245 - 2s - loss: 0.0110 - accuracy: 0.9988 - 2s/epoch - 8ms/step
Epoch 18/69
245/245 - 2s - loss: 0.0093 - accuracy: 0.9990 - 2s/epoch - 8ms/step
Epoch 19/69
245/245 - 2s - loss: 0.0073 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 20/69
245/245 - 2s - loss: 0.0061 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 21/69
245/245 - 2s - loss: 0.0053 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 22/69
245/245 - 2s - loss: 0.0050 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 23/69
```

```
245/245 - 2s - loss: 0.0050 - accuracy: 0.9991 - 2s/epoch - 8ms/step
Epoch 24/69
245/245 - 2s - loss: 0.0049 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 25/69
245/245 - 2s - loss: 0.0037 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 26/69
245/245 - 2s - loss: 0.0028 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 27/69
245/245 - 2s - loss: 0.0032 - accuracy: 0.9994 - 2s/epoch - 9ms/step
Epoch 28/69
245/245 - 2s - loss: 0.0036 - accuracy: 0.9994 - 2s/epoch - 10ms/step
Epoch 29/69
245/245 - 2s - loss: 0.0030 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 30/69
245/245 - 2s - loss: 0.0026 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 31/69
245/245 - 2s - loss: 0.0036 - accuracy: 0.9994 - 2s/epoch - 10ms/step
Epoch 32/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9994 - 2s/epoch - 10ms/step
Epoch 33/69
245/245 - 2s - loss: 0.0024 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 34/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9994 - 2s/epoch - 9ms/step
Epoch 35/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9994 - 2s/epoch - 8ms/step
Epoch 36/69
245/245 - 2s - loss: 0.0022 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 37/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 38/69
245/245 - 2s - loss: 0.0017 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 39/69
245/245 - 2s - loss: 0.0018 - accuracy: 0.9996 - 2s/epoch - 10ms/step
Epoch 40/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 41/69
245/245 - 2s - loss: 0.0015 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 42/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 43/69
245/245 - 2s - loss: 0.0023 - accuracy: 0.9995 - 2s/epoch - 9ms/step
Epoch 44/69
245/245 - 2s - loss: 0.0018 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 45/69
245/245 - 2s - loss: 0.0016 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 46/69
245/245 - 2s - loss: 0.0013 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 47/69
```

```
245/245 - 2s - loss: 0.0016 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 48/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 49/69
245/245 - 2s - loss: 0.0034 - accuracy: 0.9992 - 2s/epoch - 8ms/step
Epoch 50/69
245/245 - 2s - loss: 0.0020 - accuracy: 0.9995 - 2s/epoch - 8ms/step
Epoch 51/69
245/245 - 2s - loss: 0.0013 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 52/69
245/245 - 2s - loss: 8.2434e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 53/69
245/245 - 2s - loss: 7.5093e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 54/69
245/245 - 2s - loss: 6.3677e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 55/69
245/245 - 2s - loss: 0.0013 - accuracy: 0.9997 - 2s/epoch - 9ms/step
Epoch 56/69
245/245 - 2s - loss: 6.4259e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 57/69
245/245 - 2s - loss: 3.4372e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 58/69
245/245 - 2s - loss: 0.0013 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 59/69
245/245 - 2s - loss: 0.0014 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 60/69
245/245 - 2s - loss: 8.9948e-04 - accuracy: 0.9996 - 2s/epoch - 8ms/step
Epoch 61/69
245/245 - 2s - loss: 9.5490e-04 - accuracy: 0.9997 - 2s/epoch - 9ms/step
Epoch 62/69
245/245 - 2s - loss: 0.0027 - accuracy: 0.9996 - 2s/epoch - 9ms/step
Epoch 63/69
245/245 - 2s - loss: 0.0022 - accuracy: 0.9996 - 2s/epoch - 10ms/step
Epoch 64/69
245/245 - 2s - loss: 9.3546e-04 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 65/69
245/245 - 2s - loss: 5.3727e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 66/69
245/245 - 2s - loss: 4.3442e-04 - accuracy: 0.9999 - 2s/epoch - 9ms/step
Epoch 67/69
245/245 - 2s - loss: 5.6441e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Epoch 68/69
245/245 - 2s - loss: 0.0010 - accuracy: 0.9997 - 2s/epoch - 8ms/step
Epoch 69/69
245/245 - 2s - loss: 4.6897e-04 - accuracy: 0.9999 - 2s/epoch - 8ms/step
Test Accuracy: 0.8669053912162781
```

[86]:
```
perform_list_II = perform_list1 + perform_list2 + perform_list3 + perform_list4␣
↪+ perform_list5
model_performance_II = pd.DataFrame(data=perform_list_II)
model_performance_II = model_performance_II[['Model', 'Test Accuracy',␣
↪'ROC-AUC', 'Loss', 'Precision', 'Recall', 'F1']]
model_performance_II
```

[86]:

	Model	Test Accuracy	ROC-AUC	Loss	Precision	Recall	F1
0	Naive Bayes	67.76	0.72	NaN	0.68	0.68	0.68
1	Naive Bayes	67.26	0.73	NaN	0.67	0.67	0.67
2	Naive Bayes	76.65	0.68	NaN	0.77	0.77	0.77
3	Naive Bayes	75.65	0.69	NaN	0.76	0.76	0.76
4	Naive Bayes	82.19	0.73	NaN	0.82	0.82	0.82
..	…	…	…	…	…	…	…
115	ANN	0.84	NaN	1.62	NaN	NaN	NaN
116	ANN	0.80	NaN	1.62	NaN	NaN	NaN
117	ANN	0.82	NaN	2.24	NaN	NaN	NaN
118	ANN	0.87	NaN	0.61	NaN	NaN	NaN
119	ANN	0.87	NaN	1.39	NaN	NaN	NaN

[120 rows x 7 columns]

[87]:
```
pd.set_option('display.max_rows', None)
model_performance_II
```

[87]:

	Model	Test Accuracy	ROC-AUC	Loss	Precision	Recall	\
0	Naive Bayes	67.76	0.72	NaN	0.68	0.68	
1	Naive Bayes	67.26	0.73	NaN	0.67	0.67	
2	Naive Bayes	76.65	0.68	NaN	0.77	0.77	
3	Naive Bayes	75.65	0.69	NaN	0.76	0.76	
4	Naive Bayes	82.19	0.73	NaN	0.82	0.82	
5	Naive Bayes	81.78	0.72	NaN	0.82	0.82	
6	Naive Bayes	66.96	0.73	NaN	0.67	0.67	
7	Naive Bayes	68.50	0.74	NaN	0.69	0.69	
8	Naive Bayes	66.86	0.68	NaN	0.67	0.67	
9	Naive Bayes	66.57	0.70	NaN	0.67	0.67	
10	Naive Bayes	65.99	0.73	NaN	0.66	0.66	
11	Naive Bayes	68.21	0.74	NaN	0.68	0.68	
12	Naive Bayes	67.92	0.74	NaN	0.68	0.68	
13	Naive Bayes	70.75	0.79	NaN	0.71	0.71	
14	Naive Bayes	70.13	0.73	NaN	0.70	0.70	
15	Naive Bayes	73.71	0.79	NaN	0.74	0.74	
16	Naive Bayes	68.64	0.74	NaN	0.69	0.69	
17	Naive Bayes	70.93	0.79	NaN	0.71	0.71	
18	Naive Bayes	71.44	0.77	NaN	0.71	0.71	
19	Naive Bayes	75.26	0.78	NaN	0.75	0.75	
20	Naive Bayes	66.07	0.77	NaN	0.66	0.66	

21	Naive Bayes	63.83	0.77	NaN	0.64	0.64
22	Naive Bayes	70.37	0.77	NaN	0.70	0.70
23	Naive Bayes	72.37	0.80	NaN	0.72	0.72
24	Logistic Regression	77.36	0.67	NaN	0.77	0.77
25	Logistic Regression	76.36	0.66	NaN	0.76	0.76
26	Logistic Regression	79.36	0.65	NaN	0.79	0.79
27	Logistic Regression	78.15	0.64	NaN	0.78	0.78
28	Logistic Regression	81.65	0.73	NaN	0.82	0.82
29	Logistic Regression	81.69	0.74	NaN	0.82	0.82
30	Logistic Regression	61.93	0.64	NaN	0.62	0.62
31	Logistic Regression	63.96	0.67	NaN	0.64	0.64
32	Logistic Regression	65.31	0.67	NaN	0.65	0.65
33	Logistic Regression	64.54	0.66	NaN	0.65	0.65
34	Logistic Regression	68.12	0.73	NaN	0.68	0.68
35	Logistic Regression	68.79	0.75	NaN	0.69	0.69
36	Logistic Regression	72.25	0.78	NaN	0.72	0.72
37	Logistic Regression	85.91	0.89	NaN	0.86	0.86
38	Logistic Regression	73.92	0.80	NaN	0.74	0.74
39	Logistic Regression	84.99	0.90	NaN	0.85	0.85
40	Logistic Regression	72.55	0.80	NaN	0.73	0.73
41	Logistic Regression	76.78	0.86	NaN	0.77	0.77
42	Logistic Regression	81.17	0.88	NaN	0.81	0.81
43	Logistic Regression	81.77	0.88	NaN	0.82	0.82
44	Logistic Regression	77.68	0.82	NaN	0.78	0.78
45	Logistic Regression	80.27	0.85	NaN	0.80	0.80
46	Logistic Regression	74.81	0.82	NaN	0.75	0.75
47	Logistic Regression	78.66	0.87	NaN	0.79	0.79
48	Linear SVC	81.61	NaN	NaN	0.82	0.82
49	Linear SVC	81.65	NaN	NaN	0.82	0.82
50	Linear SVC	81.90	NaN	NaN	0.82	0.82
51	Linear SVC	81.65	NaN	NaN	0.82	0.82
52	Linear SVC	81.65	NaN	NaN	0.82	0.82
53	Linear SVC	81.69	NaN	NaN	0.82	0.82
54	Linear SVC	68.02	NaN	NaN	0.68	0.68
55	Linear SVC	69.47	NaN	NaN	0.69	0.69
56	Linear SVC	68.50	NaN	NaN	0.69	0.69
57	Linear SVC	67.54	NaN	NaN	0.68	0.68
58	Linear SVC	68.50	NaN	NaN	0.69	0.69
59	Linear SVC	69.28	NaN	NaN	0.69	0.69
60	Linear SVC	84.33	NaN	NaN	0.84	0.84
61	Linear SVC	87.35	NaN	NaN	0.87	0.87
62	Linear SVC	84.00	NaN	NaN	0.84	0.84
63	Linear SVC	87.71	NaN	NaN	0.88	0.88
64	Linear SVC	88.21	NaN	NaN	0.88	0.88
65	Linear SVC	90.33	NaN	NaN	0.90	0.90
66	Linear SVC	86.66	NaN	NaN	0.87	0.87
67	Linear SVC	86.93	NaN	NaN	0.87	0.87

68	Linear SVC	79.80	NaN	NaN	0.80	0.80
69	Linear SVC	83.65	NaN	NaN	0.84	0.84
70	Linear SVC	89.35	NaN	NaN	0.89	0.89
71	Linear SVC	89.67	NaN	NaN	0.90	0.90
72	Random Forest	81.53	0.71	NaN	0.82	0.82
73	Random Forest	81.61	0.71	NaN	0.82	0.82
74	Random Forest	81.69	0.71	NaN	0.82	0.82
75	Random Forest	81.44	0.69	NaN	0.81	0.81
76	Random Forest	81.78	0.72	NaN	0.82	0.82
77	Random Forest	81.48	0.73	NaN	0.81	0.81
78	Random Forest	67.54	0.73	NaN	0.68	0.68
79	Random Forest	68.89	0.74	NaN	0.69	0.69
80	Random Forest	66.96	0.71	NaN	0.67	0.67
81	Random Forest	67.25	0.71	NaN	0.67	0.67
82	Random Forest	67.05	0.72	NaN	0.67	0.67
83	Random Forest	70.53	0.76	NaN	0.71	0.71
84	Random Forest	93.02	0.97	NaN	0.93	0.93
85	Random Forest	95.11	0.98	NaN	0.95	0.95
86	Random Forest	90.81	0.97	NaN	0.91	0.91
87	Random Forest	91.73	0.97	NaN	0.92	0.92
88	Random Forest	93.82	0.97	NaN	0.94	0.94
89	Random Forest	93.26	0.97	NaN	0.93	0.93
90	Random Forest	86.21	0.91	NaN	0.86	0.86
91	Random Forest	86.63	0.91	NaN	0.87	0.87
92	Random Forest	81.11	0.87	NaN	0.81	0.81
93	Random Forest	80.51	0.88	NaN	0.81	0.81
94	Random Forest	89.29	0.95	NaN	0.89	0.89
95	Random Forest	88.87	0.95	NaN	0.89	0.89
96	ANN	0.79	NaN	1.71	NaN	NaN
97	ANN	0.79	NaN	1.96	NaN	NaN
98	ANN	0.77	NaN	1.83	NaN	NaN
99	ANN	0.77	NaN	2.64	NaN	NaN
100	ANN	0.78	NaN	1.23	NaN	NaN
101	ANN	0.77	NaN	2.02	NaN	NaN
102	ANN	0.64	NaN	2.08	NaN	NaN
103	ANN	0.66	NaN	2.18	NaN	NaN
104	ANN	0.63	NaN	2.23	NaN	NaN
105	ANN	0.64	NaN	2.45	NaN	NaN
106	ANN	0.62	NaN	1.91	NaN	NaN
107	ANN	0.64	NaN	2.03	NaN	NaN
108	ANN	0.88	NaN	0.73	NaN	NaN
109	ANN	0.89	NaN	0.79	NaN	NaN
110	ANN	0.89	NaN	0.75	NaN	NaN
111	ANN	0.88	NaN	1.12	NaN	NaN
112	ANN	0.88	NaN	0.62	NaN	NaN
113	ANN	0.87	NaN	1.38	NaN	NaN
114	ANN	0.84	NaN	1.21	NaN	NaN

115	ANN	0.84	NaN	1.62	NaN	NaN
116	ANN	0.80	NaN	1.62	NaN	NaN
117	ANN	0.82	NaN	2.24	NaN	NaN
118	ANN	0.87	NaN	0.61	NaN	NaN
119	ANN	0.87	NaN	1.39	NaN	NaN

	F1
0	0.68
1	0.67
2	0.77
3	0.76
4	0.82
5	0.82
6	0.67
7	0.69
8	0.67
9	0.67
10	0.66
11	0.68
12	0.68
13	0.71
14	0.70
15	0.74
16	0.69
17	0.71
18	0.71
19	0.75
20	0.66
21	0.64
22	0.70
23	0.72
24	0.77
25	0.76
26	0.79
27	0.78
28	0.82
29	0.82
30	0.62
31	0.64
32	0.65
33	0.65
34	0.68
35	0.69
36	0.72
37	0.86
38	0.74
39	0.85

40	0.73
41	0.77
42	0.81
43	0.82
44	0.78
45	0.80
46	0.75
47	0.79
48	0.82
49	0.82
50	0.82
51	0.82
52	0.82
53	0.82
54	0.68
55	0.69
56	0.69
57	0.68
58	0.69
59	0.69
60	0.84
61	0.87
62	0.84
63	0.88
64	0.88
65	0.90
66	0.87
67	0.87
68	0.80
69	0.84
70	0.89
71	0.90
72	0.82
73	0.82
74	0.82
75	0.81
76	0.82
77	0.81
78	0.68
79	0.69
80	0.67
81	0.67
82	0.67
83	0.71
84	0.93
85	0.95
86	0.91

```
87   0.92
88   0.94
89   0.93
90   0.86
91   0.87
92   0.81
93   0.81
94   0.89
95   0.89
96    NaN
97    NaN
98    NaN
99    NaN
100   NaN
101   NaN
102   NaN
103   NaN
104   NaN
105   NaN
106   NaN
107   NaN
108   NaN
109   NaN
110   NaN
111   NaN
112   NaN
113   NaN
114   NaN
115   NaN
116   NaN
117   NaN
118   NaN
119   NaN
```

#CONCLUSION

10 For the first task we can observe that, Linear SVM performs better for Bow and Bag of n grams as comapred to other models. For Tf-Idf almost all models give identical results but Random Forest or ANN can be preferred as they can be more finetuned easily and better results can be achieved from them.

For the second task, we can observe that:

- Bow (Oversampling) with Random Forest
- Bag of n gram (Oversampling) with Random Forest
- TF-IDF (Oversampling) with Random Forest

these above combinations give the best result.

When we vectorize the training data, the CountVectorizer will select the words/features/terms which occur the most frequently. It takes absolute values so if you set the 'max_features = 5000', it will select the 5000 most common words in the data. The best way to omit the word from the given corpus/data, of course, would be to use stop_words parameter, but imagine if there are plenty of such words; or words that are related to the topic but occur scarcely. In the second case, the max_features parameter will help.
We can observe from the excel table that, for our particular dataset, for the normal Bow, bag of n grams and TF-IDF cases, we get better accuracy when we used 'max_features = 1000'. This shows that 1000 is a better value for 'max_features' making feature vector for all the 3 types and hence means there are only around 1000 common and related words to our news headline. Increasing or decreasing the max_feature by 30-40% might yield even better results.

Whereas for under,over sampling and SMOTE we get better results for 'max_features = 5000' because these methods change the size of data by removing samples (undersampling) or adding samples (oversampling&SMOTE) and this results in more number of common words higher than 1000 and hence it yields better results for 'max_features = 5000'.

Overall, for classifying news headline as to be relevant or non relevant we can use Bow (Oversampling) with Random Forest model as it gives the highest accuracy of 94.57

SMOTE overall performs well and matches close to oversampling.
Undersampling, reduces the accuracy by a lot hence it is not normally preferred method of class balancing.

Undersampling involves randomly removing examples from the majority class.
Oversampling involves randomly duplicating examples from the minority class and adding them to the minority class.
SMOTE involves randomly synthesizing examples from the minority class and adding the new synthesized examples to the minority class.