Email_Classifier

November 18, 2024

0.1 Approach/Potential Outline:

1. Finding better features based on the email text. Some example features are:
    1. Number of characters in the subject/body
    2. Number of words in the subject/body
    3. Use of punctuation (e.g., how many '!'s were there?)
    4. Number/percentage of capital letters
    5. Whether the email is a reply to an earlier email or a forwarded email
2. Finding better words to use as features. Which words are the best at distinguishing emails? This requires digging into the email text itself. Alternatively, identify misclassified emails and see which relevant words are missing in the model.
3. Reducing dimensionality and/or multicollinearity. few methods to achieve this:
    1. Implement PCA.
    2. Interpret the model coefficients. Note that a feature will be more valuable in classification if its coefficient has a larger **absolute** value. If the coefficient has a lower **absolute** value, the feature likely isn't valuable in classifying emails.
4. Better data processing. For example, many emails contain HTML as well as text. Extracting the text from the HTML to help find better words. Or can match HTML tags themselves, or even some combination of the two.
5. Model selection. Adjust the parameters of the model (e.g. the penalty type, the regularization parameter, or any arguments in `LogisticRegression`) to achieve higher accuracy. Should be using cross-validation for feature and model selection! Otherwise, it will likely overfit to the training data.
    1. Implementing L1 regularization. The documentation for `LogisticRegression` may be helpful here.
    2. imported `GridSearchCV`and use sklearn's `GridSearchCV` (documentation) class to perform cross-validation.

Note 1: use the **validation data** to evaluate the model and get a better sense of how it will perform on the test set. However, overfiting to in the validation set if it try to optimize the validation accuracy too much. Alternatively, can perform cross-validation on the entire training set.

[36]:
```
import numpy as np
import pandas as pd
import sys

import matplotlib.pyplot as plt
%matplotlib inline
```

```
import seaborn as sns
sns.set(style = "whitegrid",
        color_codes = True,
        font_scale = 1.5)

from datetime import datetime
from IPython.display import display, HTML
```

0.2 Loading and Cleaning Data

[37]:
```
import zipfile
with zipfile.ZipFile('spam_ham_data.zip') as item:
    item.extractall()
```

[38]:
```
original_training_data = pd.read_csv('train.csv')
test = pd.read_csv('test.csv')

# Convert the emails to lowercase as the first step of text processing.
original_training_data['email'] = original_training_data['email'].str.lower()
test['email'] = test['email'].str.lower()

original_training_data.head()
```

[38]:

	id	subject	email	spam
0	0	Subject: A&L Daily to be auctioned in bankrupt…	url: http://boingboing.net/#85534171\n date: n…	0
1	1	Subject: Wired: "Stronger ties between ISPs an…	url: http://scriptingnews.userland.com/backiss…	0
2	2	Subject: It's just too small …	<html>\n <head>\n </head>\n <body>\n <font siz…	1
3	3	Subject: liberal defnitions\n	depends on how much over spending vs. how much…	0
4	4	Subject: RE: [ILUG] Newbie seeks advice - Suse…	hehe sorry but if you hit caps lock twice the …	0

[39]:
```
# Fill any missing or NAN values.
print('Before imputation:')
print(original_training_data.isnull().sum())
original_training_data = original_training_data.fillna('')
print('------------')
print('After imputation:')
print(original_training_data.isnull().sum())
```

```
Before imputation:
```

```
id         0
subject    6
email      0
spam       0
dtype: int64
------------
After imputation:
id         0
subject    0
email      0
spam       0
dtype: int64
```

0.3 Training/Validation Split

[40]:
```
# This creates a 90/10 train-validation split on our labeled data.
from sklearn.model_selection import train_test_split
train, val = train_test_split(original_training_data, test_size = 0.1,␣
 ↪random_state = 42)

# We must do this in order to preserve the ordering of emails to labels for␣
 ↪words_in_texts.
train = train.reset_index(drop = True)
```

0.4 Feature Engineering

[41]:
```
from projB2_utils import words_in_texts

words_in_texts(['hello', 'bye', 'world'], pd.Series(['hello', 'hello␣
 ↪worldhello']))
```

[41]:
```
array([[1, 0, 0],
       [1, 0, 1]])
```

0.5 EDA and Basic Classification

[42]:
```
some_words = ['drug', 'bank', 'prescription', 'memo', 'private']

X_train = words_in_texts(some_words, train['email'])
Y_train = np.array(train['spam'])

X_train[:5], Y_train[:5]
```

[42]:
```
(array([[0, 0, 0, 0, 0],
        [0, 0, 0, 0, 0],
        [0, 0, 0, 0, 0],
        [0, 0, 0, 0, 0],
```

```
        [0, 0, 0, 1, 0]]),
 array([0, 0, 0, 0, 0]))
```

[43]:
```
from sklearn.linear_model import LogisticRegression

simple_model = LogisticRegression()
simple_model.fit(X_train, Y_train)

training_accuracy = simple_model.score(X_train, Y_train)
print("Training Accuracy: ", training_accuracy)
```

```
Training Accuracy:  0.7576201251164648
```

0.6 Evaluating Classifiers

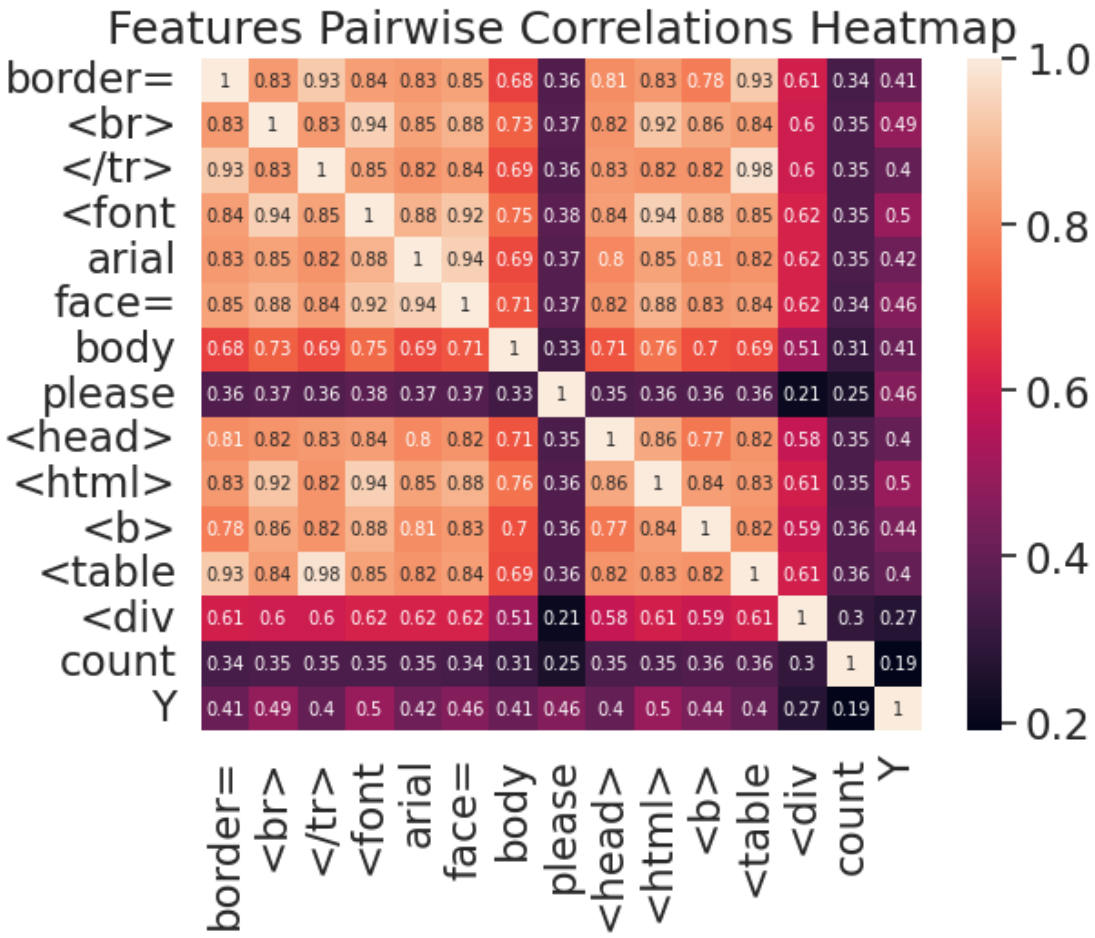
Presumably, the classifier will be used for **filtering**, or preventing messages labeled spam from reaching someone's inbox. There are two kinds of errors we can make: - **False positive (FP)**: A ham email gets flagged as spam and filtered out of the inbox. - **False negative (FN)**: A spam email gets mislabeled as ham and ends up in the inbox.

To be clear, we label spam emails as 1 and ham emails as 0. These definitions depend both on the true labels and the predicted labels. False positives and false negatives may be of differing importance, leading us to consider more ways of evaluating a classifier in addition to overall accuracy:

Note that a True Positive (TP) is a spam email that is classified as spam, and a True Negative (TN) is a ham email that is classified as ham.

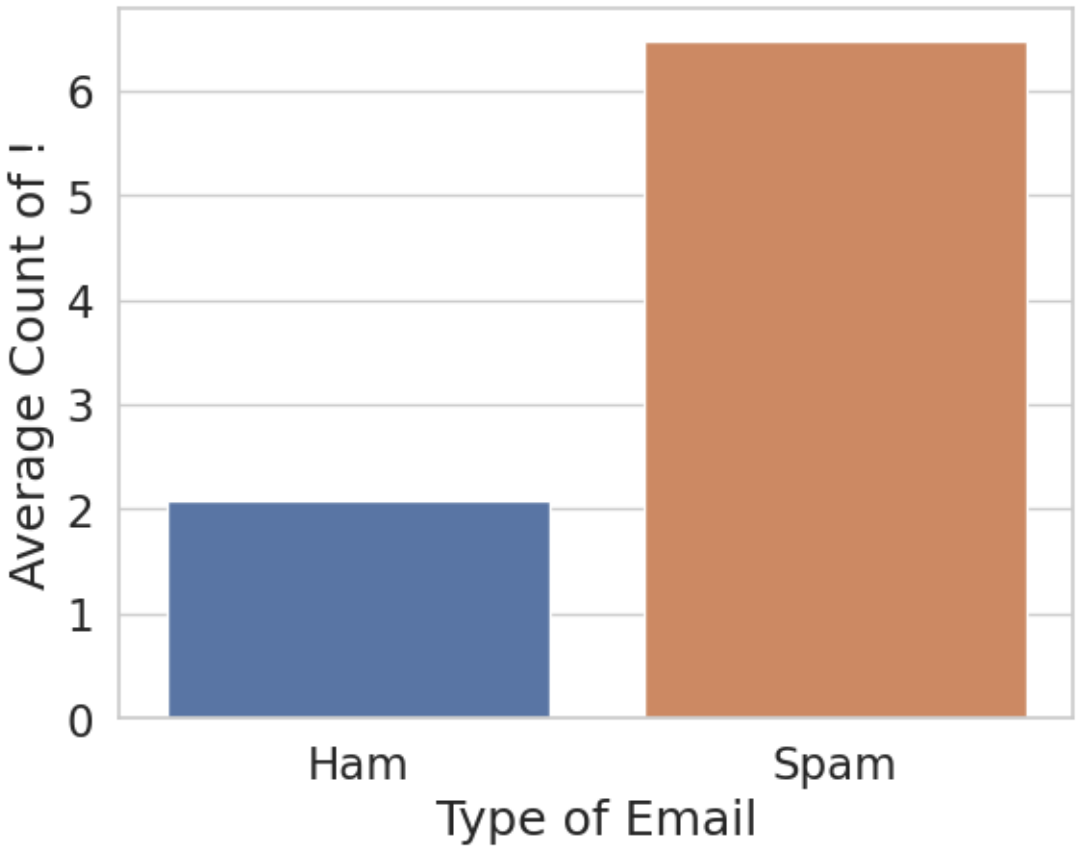
[44]:
```
#Visual of each email's count of characters, numbers, symbols, etc...
#Finding better features based on the email text. Some example features are:
#Number of characters in the subject/body
#Number of words in the subject/body
#Use of punctuation (e.g., how many '!'s were there?)
#Number/percentage of capital letters
#Whether the email is a reply to an earlier email or a forwarded email
words = ['border=', '<br>', '</tr>', '<font', 'arial',␣
 ↪'face=','body','please','<head>','<html>', '<b>', '<table', '<div']
X_train_vis = words_in_texts(words, train['email'])
X_df = pd.DataFrame(columns = words)
for i, row in enumerate(X_train_vis):
    X_df.loc[i] = row
X_df['count'] = train['email'].str.count('!')
X_df['Y'] = train['spam']
display(sns.heatmap(data = np.round(X_df.corr(), 2), annot = True, annot_kws =␣
 ↪{'size':7}).set(title='Features Pairwise Correlations Heatmap'))
```

```
[Text(0.5, 1.0, 'Features Pairwise Correlations Heatmap')]
```



[45]:
```
bar = train['email'].str.count('!').groupby(train['spam']).mean().to_frame().
↪reset_index()
bar = bar.replace({0:'Ham', 1:'Spam'})
sns.barplot(data = bar,x = 'spam', y = 'email')
plt.xlabel('Type of Email')
plt.ylabel('Average Count of !')
```

[45]:
```
Text(0, 0.5, 'Average Count of !')
```



From the heatmap plotted above we can observe that common html tags are associated not only with other html tags but also with determining our response variable 'spam'. The second barplot plots the count of exclamations in each ham and spam email texts, proving that exclamation points are on average more prevalent within spam emails than ham.

1 Building Own Model!

[46]:
```
train[train['spam'] == 1][40:60].head()
```

[46]:

	id	subject	email	spam
133	7486	Subject: [] \n	<!-- saved from url=3d(0022)http://internet.e-…	1
134	4126	Subject: ****Already own a satellite? Need a …	recieve all channels on your satellite system!…	1
138	1087	Subject: FREE Cell Phone + $50 Cash Back!\n		
144	3303	Subject: Thanksgiving Sale\n		
145	17	Subject: National Charity Suffering Since 9/11\n		

```
138  <html><head><title>free motorola cell phone wi…     1
144  \n this is a multi-part message in mime format…     1
145  <body bgcolor=#ffffff>\n <div><font face=arial…     1
```

[47]:
```
# import libraries
# You may use any of these to create features.
from sklearn.preprocessing import OneHotEncoder
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, roc_curve, confusion_matrix
from sklearn.model_selection import GridSearchCV
from sklearn.decomposition import PCA
import re
from collections import Counter
```

[48]:
```
print('Ham Replies Prop: ', train.loc[(train['subject'].str.contains('Re:')) &␣
 ↪train['spam'] == 0,:].count()[1]/train.shape[0])
print('Spam Replies Prop: ', train.loc[(train['subject'].str.contains('Re:')) &␣
 ↪train['spam'] == 1,:].count()[1]/train.shape[0])
```

```
Ham Replies Prop:  0.9912152269399708
Spam Replies Prop:  0.008784773060029283
```

[49]:
```
X_df = train.copy()
#train['email'].str.count(r'/?[2]').groupby(train['spam']).mean().to_frame()
#X_df['has reply'] = train['subject'].str.contains(r'Subject:\s(Re:)\s', regex=␣
 ↪True)
X_df['X_prop'] = (X_df['subject'].str.count(r'[A-Z]')/(X_df['subject'].str.
 ↪strip(' ').str.len()))
X_df
np.mean(X_df[X_df['spam'] ==True]['X_prop'])
```

[49]:
```
0.16631139763467886
```

[50]:
```
# Define  processing function, processed data, and model here.
# You may find it helpful to look through the rest of the questions first!
def feature_process(df):
    words = ['border=', '<br>', '</tr>', '<font','arial',␣
 ↪'face=','body','please','<head>','<html>', '<b>',  '<table','<div',
         '<p', 'sans-serif', 'free', 'cash',
            'offer', 'business', 'verdana', '<option', 'address', 'need',␣
 ↪'money', 'credit card', '<center', 'align=', 'emailing list',
           'urgent', 'dear', 'congratulation', 'dollar', 'confidential',␣
 ↪'drug']
    X = words_in_texts(words, df['email'])
    X_reg = pd.DataFrame(X,columns = words)
    #for i, row in enumerate(X):
        #X_reg.loc[i] = row
```

```
## Add additional Features below this
    X_reg['net_sub'] =  df['subject'].str.findall(r'[!=&;$%#"]').str.len().
 ↪fillna(0)
    X_reg['net_email'] =  df['email'].str.findall(r'[!=&;$#<]').str.len().
 ↪fillna(0)
    X_reg['multi_dashes'] =  df['email'].str.findall(r'/-{3,}').str.len().
 ↪fillna(0)
    X_reg['repeating_chars'] =  df['email'].str.findall(r'/[a-z]{5,}').str.
 ↪len().fillna(0)
    X_reg['text_text'] = df['email'].str.split(' ').str.len()/df['email'].str.
 ↪len().fillna(0)
    X_reg['text_subject'] = df['subject'].str.split(' ').str.len().fillna(0)
    X_reg['Is Reply'] = df['subject'].str.contains(r'\Subject: Re:', regex=␣
 ↪True).fillna(0)
    avg = np.mean((df['subject'].str.count(r'[A-Z]')/df['subject'].str.
 ↪len())**2)
    #X_reg['Capital'] = ((df['subject'].str.count(r'[A-Z]')/df['subject'].str.
 ↪strip(' ').str.len())**2).fillna(avg)
    X_reg['shouting'] = ((df['subject'].str.count(r'[A-Z]')/df['subject'].str.
 ↪strip(' ').str.len())**2).fillna(avg)
    #X_reg['Capital'] = df['Capital'].fillna(np.mean(X_reg['Capital']))
    sub_list = ['please', 'address', 'money', 'time', 'free', '$', 'email',␣
 ↪'credit card', ' need', 'information', '#', 'asap', 'assistance',
                '"', 'urgent', 'need', 'cash', 'congrats', 'congratulation']
    X_sub = words_in_texts(sub_list, df['subject'].str.lower())
    X_sub_df = pd.DataFrame(X_sub, columns =sub_list)
    #for i, row in enumerate(X_sub):
    #    X_sub_df.loc[i] = row
    #return X_reg.merge(X_sub_df)
    return X_reg.merge(X_sub_df, left_index=True, right_index=True, how='left').
 ↪fillna(0)
#-----------------------

X_reg = feature_process(train)
simple = LogisticRegression(penalty = 'l1', solver = 'liblinear')
X_reg['Y'] = train['spam']
simple.fit(X_reg.iloc[:,:X_reg.shape[1]-1], X_reg['Y'])
```

[50]:
```
LogisticRegression(penalty='l1', solver='liblinear')
```

[62]:
```
import warnings

# Ignore all warnings
warnings.filterwarnings('ignore')

# To ignore specific warnings, e.g., ConvergenceWarning
```

```
from sklearn.exceptions import ConvergenceWarning
warnings.filterwarnings('ignore', category=ConvergenceWarning)
```

[63]:
```
parameters = {'solver':['lbfgs', 'liblinear', 'newton-cg', 'saga']}
grid = GridSearchCV(estimator=simple, param_grid=parameters)
grid_result = grid.fit(X_reg.iloc[:,:(X_reg.shape[1]-1)], X_reg['Y'])
grid_result.cv_results_
```

[63]:
```
{'mean_fit_time': array([0.002246  , 0.17111087, 0.01595397, 0.78244791]),
 'std_fit_time': array([0.000787  , 0.06369757, 0.02645998, 0.01798326]),
 'mean_score_time': array([0.        , 0.00426412, 0.        , 0.00430465]),
 'std_score_time': array([0.00000000e+00, 2.78523519e-05, 0.00000000e+00,
1.78273019e-04]),
 'param_solver': masked_array(data=['lbfgs', 'liblinear', 'newton-cg', 'saga'],
              mask=[False, False, False, False],
        fill_value='?',
             dtype=object),
 'params': [{'solver': 'lbfgs'},
  {'solver': 'liblinear'},
  {'solver': 'newton-cg'},
  {'solver': 'saga'}],
 'split0_test_score': array([       nan, 0.92348636,        nan, 0.80705256]),
 'split1_test_score': array([       nan, 0.92880905,        nan, 0.79574185]),
 'split2_test_score': array([       nan, 0.90818363,        nan, 0.80239521]),
 'split3_test_score': array([       nan, 0.92010652,        nan, 0.80159787]),
 'split4_test_score': array([       nan, 0.91011984,        nan, 0.7976032 ]),
 'mean_test_score': array([       nan, 0.91814108,        nan, 0.80087814]),
 'std_test_score': array([       nan, 0.0078706 ,        nan, 0.00395094]),
 'rank_test_score': array([3, 1, 3, 2], dtype=int32)}
```

[64]:
```
train_predictions = simple.predict(feature_process(train))

# Print  training accuracy.
training_accuracy = np.mean(train_predictions == train["spam"])
training_accuracy
```

[64]:
```
0.9241315053906562
```

1.1 Test Predictions on Test Set (Unlabeled)

[65]:
```
train.iloc[:,:3]
```

[65]:
```
        id                                             subject  \
0     7657            Subject: Patch to enable/disable log\n
1     6911       Subject: When an engineer flaps his wings\n
2     6074  Subject: Re: [Razor-users] razor plugins for m…
3     4376  Subject: NYTimes.com Article: Stop Those Press…
```

```
4     5766  Subject: What's facing FBI's new CIO? (Tech Up…
…     …                                                   …
7508  5734  Subject: [Spambayes] understanding high false …
7509  5191        Subject: Reach millions on the internet!!\n
7510  5390                          Subject: Facts about sex.\n
7511   860  Subject: Re: Zoot apt/openssh & new DVD playin…
7512  7270  Subject: Re: Internet radio - example from a c…

                                                     email
0     while i was playing with the past issues, it a…
1     url: http://diveintomark.org/archives/2002/10/…
2     no, please post a link!\n \n fox\n ----- origi…
3     this article from nytimes.com \n has been sent…
4     <html>\n <head>\n <title>tech update today</ti…
…                                                     …
7508  >>>>> "tp" == tim peters <tim.one@comcast.net>…
7509  \n dear consumers, increase your business sale…
7510  \n forwarded-by: flower\n \n did you know that…
7511  on tue, oct 08, 2002 at 04:36:13pm +0200, matt…
7512  chris haun wrote:\n > \n > we would need someo…

[7513 rows x 3 columns]
```

[66]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(train.iloc[:,:3],␣
 ↪train['spam'], test_size = .33,
                                                    random_state = 42)
test_predictions = simple.predict(feature_process(X_test))
np.mean(test_predictions == y_test)
```

[66]:
```
0.8471774193548387
```

[67]:
```
test_predictions = simple.predict(feature_process(test))
# np.mean(test_predictions == test['spam'])
```

[68]:
```
# Assuming that the predictions on the test set are stored in a 1-dimensional␣
 ↪array called
# test_predictions.
submission_df = pd.DataFrame({
    "Id": test['id'],
    "Class": test_predictions,
}, columns=['Id', 'Class'])
timestamp = datetime.now().strftime("%Y%m%d_%H%M%S")
filename = "submission_{}.csv".format(timestamp)
submission_df.to_csv(filename, index=False)

print('Created a CSV file: {}.'.format("submission_{}.csv".format(timestamp)))
```

```
display(HTML("Download test prediction <a href='" + filename + "'␣
↪download>here</a>."))
```

```
Created a CSV file: submission_20241118_000602.csv.
```

```
<IPython.core.display.HTML object>
```

---

2 Analysis/ Key Takeaways

1. How did better features worked for the model?
2. What worked and didnt?
3. Surprising findings?

1.) My model progressively reduced its error and increased its accuracy as I implemented more general html tags that I found are common in programming along with perfroming EDA such as counting the unique values of each word/tag split by spaces and sentence ending characters (.!,:, etc..). From this EDA and filtering for spam email I was able to gain a good understadning of what sort of words/tags spam emails tend to contain outside of commonly used words that overlap both types of emails such as 'the', 'of'. The heat map derived above also provided a good sense to how these different words would perform in our model as we were able to consoliate pairwise correlations which inicated that the html tags correlated highly with not only each other but with the our response variable too ('spam'). Taking the suggestions offered above I also performed regex parsing to count characters such as !,#,-,&,= as I found these to common amongst spam emails, especially groups of them which signaled my capture groups being multiples of these aggregated counts. Additionally, I applied our words_in_texts function to the subject column and merged that to act as another set of features (although it did not contribute much). Furthermore, I added a column that distinguished reply emails emails versus regular emails. Lastly, the behavior of different characters was key to allude spam emails, thus I computed the ratio of capital letters versus total characters and found that spam emails contained higher ratios than ham emails (i.e spam emails and subjects tended to be more in a "shouting" tone). All in all, the key features that mitigated overlap and helped with interpretibility of the model was the incorportation of excessive symbols and html tags in the model as it was highly correlated with accurately distinguishing.

2.) Commonly used overlapping words tended to disprupt the accuracy, thus calling for adjustments and removals of my words list by steering away from commonly used vocabulary. Adding the < against the html tags helped ensure we were indeed targeting the html commands prevalent amongst spam emails. I was initially for looping an enumerated matrix into a data frame which took a long execution time, thus I adjusted and was able to input the matrix directly into a df as entries with the corresponding columns without perfromance errors. The regression would fail when Nan values were encountered during the feature process, thus ensuring the Na values were imputed was a major step.

3.) I was surprised at the excessive use of symbols and html commands used in spam emails. Additionally I realized the majority of replied emails were practically regular ham emails and were almost never replies. Upon computing the feature of the ratio of upper case letters versus all letters in the subject columns, I was intrigued to find that raising the value by 2 improved the accuracy of the model. On a final note of this feature, the feature contribution was significantly more notable than if I had not computed the ratio and simply processed the sole count.

3 Prediction Performance Evaluation with ROC Curve

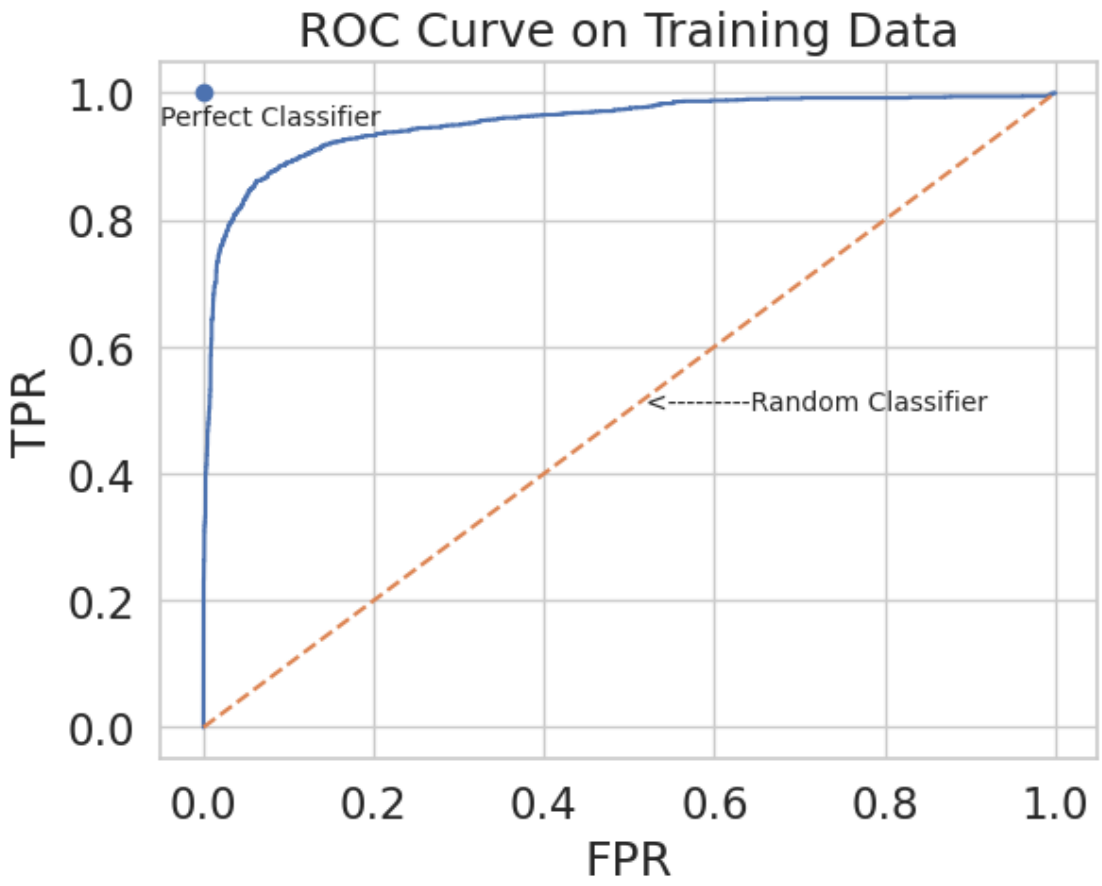
In most cases, we won't be able to get 0 false positives and 0 false negatives, so we have to compromise. For example, in the case of cancer screenings, false negatives are comparatively worse than false positives — a false negative means that a patient might not discover that they have cancer until it's too late. In contrast, a patient can receive another screening for a false positive.

Recall that logistic regression calculates the probability that an example belongs to a particular class. To classify an example, we say that an email is spam if our classifier gives it ≥ 0.5 probability of being spam. However, **we can adjust that cutoff threshold**. We can say that an email is spam only if our classifier gives it ≥ 0.7 probability of being spam, for example.

[69]:
```
train1 = feature_process(train)
```

[70]:
```
fpr, tpr, t = roc_curve(train['spam'], simple.predict_proba(train1)[:,1])
plt.plot(fpr, tpr)
plt.plot(fpr, fpr, label = 'Random Classifier', linestyle = 'dashed')
plt.xlabel("FPR")
plt.ylabel("TPR")
plt.title('ROC Curve on Training Data')
plt.text(0.52, 0.5, '<---------Random Classifier', size = 10)
plt.scatter(0,1)
plt.text(-.05, .95, 'Perfect Classifier', size = 10)

plt.show();
```



3.0.1 Assessing Example

[71]:
```
# Assessing Example
print("spam: " + str(train.loc[1092]["spam"]))
print("\nemail:\n" + train.loc[1092]["email"])
```

```
spam: 0

email:
this is a multi part message in mime format.

 --_nextpart_1_bvfoditvghtocxfdvjnkcuwblfv
 content-type: text/plain; charset="us-ascii"
 content-transfer-encoding: 7bit

 … with our telecoms partner bumblebee !

 don't get ripped off by expensive hotel, payphone and mobile charges.
 save, save, save on international calls with ryanair's phone partner.
```

****
****

you'll save up to 70% on international phone calls when you use our
online phone card. you can use the card from any phone in any country
you visit and you won't have to worry about high phone charges
when you call home or the office.

buying a card couldn't be easier and it's totally secure. simply go to
http://www.bumblebeecommunications.com/lowcostcalls/
 to avail of this special offer for ryanair customers.

it's another great deal from ryanair and our online phone
partner, bumblebee communications.

==========

e-mail disclaimer

this e-mail and any files and attachments transmitted with it
are confidential and may be legally privileged. they are intended
solely for the use of the intended recipient. any views and
opinions expressed are those of the individual author/sender
and are not necessarily shared or endorsed by ryanair holdings plc
or any associated or related company. in particular e-mail
transmissions are not binding for the purposes of forming
a contract to sell airline seats, directly or via promotions,
and do not form a contractual obligation of any type.
such contracts can only be formed in writing by post or fax,
duly signed by a senior company executive, subject to approval
by the board of directors.

the content of this e-mail or any file or attachment transmitted
with it may have been changed or altered without the consent
of the author. if you are not the intended recipient of this e-mail,
you are hereby notified that any review, dissemination, disclosure,
alteration, printing, circulation or transmission of, or any
action taken or omitted in reliance on this e-mail or any file
or attachment transmitted with it is prohibited and may be unlawful.

if you have received this e-mail in error
please notify ryanair holdings plc by emailing postmaster@ryanair.ie
or contact ryanair holdings plc, dublin airport, co dublin, ireland.

--_nextpart_1_bvfoditvghtocxfdvjnkcuwblfv
content-type: application/ms-tnef
content-transfer-encoding: base64

ej8+ijuqaqaqcaaeaaaaaaabaaeaaqeqbgaiaaaa5aqaaaaaaadoaaeigacagaaaaelqts5nawny
b3nvznqgtwfpbc5ob3rladeiaq2abaacaaaaagacaaeegaeajwaaafnhdmugdxagdg8gnzalig9u
igludgvybmf0aw9uywwgy2fsbhmhacgnaqwaawaoaaaa0gciab4aeqaqadqabqbzaqeggamadgaa
anihcaaeabeakga0aauacwebcyabaceaaaaxnundqzu1m0zcnjvgotrcodjbote2nji0qjy5odi2
naagbwedkayayagaadeaaaalaaiaaqaaaamajgaaaaaaawa2aaaaaabaadkaohxzserqwgeead0a
aqaaaaeaaaaaaaaaagfhaaeaaaayaaaayz11czthpsa7cd1sewfuywlyo2w9q0hpvk1bsuwxltay
mdgzmde2ndi1mlotnty1ngaaab4acaabaaaajwaaafnhdmugdxagdg8gnzalig9uigludgvybmf0
aw9uywwgy2fsbhmhaaacaxeaaqaaabyaaaabwlbesgnvr3xmdo5jp7lazao6pgd8aaaeabomaqaa
aawaaabdb3lszswgu2vhbgaeab0oaqaaaccaaabtyxzlihvwihrvidcwjsbvbibpbnrlcm5hdglv
bmfsignhbgxziqaaagejeaeaaac7agaatwiaalseaabmwkz1tpw4mgmacgbyy3bnmti14jidq3rl
eavbaqmb908kgakka+mcagnocsbz8gv0mcahewkad/mauh8evghvb7irxq5raweqxzl3bgagwxhf
mwrgemks2xht2wjvcfc7gl8omdurwgxgzmmauasjawqznhfqc6yuic4dmcad8hroinuiysaosgwf
kw0eiaqxzhrueyfcdqbqhwbiuqngicekogqecobeaibqxcc5mgvazxewib0fexajgb6qasagynku
ia7aimaagxygcghvzr7hlb9hexajsb+wipcacckabgbiaxagcbdiifcgcx4wikrtqvzfjaa/jlge
kaogc4aosasgyxteaqigb0aluadabaqghlpsunkacatwcigsh1eky/sfdsxwkitplf8tby5/l0od
ikpzcgahkiiqihnhcsobdxae0g8gnzaujsdfkeuodwgj8cb5vwhgmfarib6tikqcigwlgo8qmtpf
cyaemcaxeshha6d3nqiecdauiansjkei8cru1wuaolmfohucmhii8cckvttcdgqahmakoztcdyf2
7xdwi4eymtvqcjohaaaiypk8qwlnhoaztsv0jeu0d/8ociohb4aekb7bn/eioq3grguuikof8hkl
ggckohm4fdnrbgqhhsbqiwbh/wcqeyeksh5gkzqymagqkjc7ivarigmiceegnwbtb3bbc1ai8gcy
qdixaajacja6ly93r1auyiagfr8hbtnwdean8yxahyevprhqdwwgc3qocy83ap8gpdiimcalcamg
iqa30qqa+zggimbjbzeikhmhahafwh8pjsvqnqaymaeaerbbo0l/rjuaccpaniafwanbj/ag/wea
kee4c0zmjlieojxrikt/h3ukab/4cfbisyxlskugqgtwx1drfvigab4anrabaaaasaaaadxemtng
n0mwntq3rdcxrjrdqtqwnzmyn0e4mjuxmzywmde5q0e5q0bdse9wtufjtdeuy2hvlmnvcnaucnlh
bmfpci5jb20+aamagbd/////cwdyeaeaaaafapmqaqaaafoaaabtageadgblacaadqbwacaadabv
acaanwawacuamga1acaabwbuacaaaqbuahqazqbyag4ayqb0agkabwbuageabaagagmayqbsagwa
cwahac4arqbnaewaaaaaaasa9haaaaaaqaahmn0zcuheumibqaaimjpeduheumibawdep59oaaad
ape/cqqaab4a+d8baaaadaaaaenvewxllcbtzwfuaaib+t8baaaaxqaaaaaaaadcp0diweiqgrs5
caarl+gcaqaaaaaaaaavtz1swufoqulsl09vpuzjulnuiefetulosvnuukfusvzfiedst1vql0no
pvjfq0lqsuvovfmvq049q09ztevtaaaaab4a+j8baaaafqaaafn5c3rlbsbbzg1pbmlzdhjhdg9y
aaaaaaib+z8baaahgaaaaaaaadcp0diweiqgrs5caarl+gcaqaaaaaaaaauaaaaawazqaaaaaad
abpaaaaaab4ameabaaaabwaaaenpwuxfuwaahgaxqaeaaaahaaaaq09ztevtaaaeadhaaqaaaaca
aabdt1lmrvmaab4aouabaaaagaaac4aaaadaalzaqaaaasawieiiayaaaaaamaaaaaaaabgaaaa
aa6faaaaaaaaawbwgqggbgaaaaaawaaaaaaaaeyaaaaauouaafmuaqaeahgbccagaaaaaadaaaaa
aaaargaaaabuhqaaaqaaaauaaaaxmc4waaaaaamauieiiayaaaaaamaaaaaaaabgaaaaagfaaaa
aaaaqac6gqggbgaaaaaawaaaaaaaaeyaaaayiuaaaaaaaaaaaaacwc9gqggbgaaaaaawaaaaaaa
aeyaaaaaa4uaaaaaaaadamebccagaaaaaadaaaaaaaaargaaaaaqhqaaaaaaaamazoeiiayaaaaa
amaaaaaaaabgaaaaabifaaaaaaaacwdlgqggbgaaaaaawaaaaaaaaeyaaaaabouaaaaaaaalaomb
ccagaaaaaadaaaaaaaaargaaaacchqaaaaaaaasakqaaaaaacwajaaaaaaadaayqhr1s2qmabxb/
agaaawaqeasaaaadabeqaqaaab4acbabaaaazqaaafdjvehpvvjuruxfq09nu1bbulrorvjcvu1c
tevcruvet05ur0vuuklquevet0zgqllfwfbftlnjvkvit1rftcxqqvlqse9orufore1pqklmruni

```
qvjhrvntqvzflfnbvkusu0eaaaaaagf/aaeaaabiaaaapeqxm0y3qza1nddenzfgnenbnda3mzi3
qtgynteznjawmtldqtldqenit1znqulmms5jag8uy29ycc5yewfuywlylmnvbt4aeoq=


--_nextpart_1_bvfoditvghtocxfdvjnkcuwblfv
content-type: text/plain; charset="us-ascii"
content-description: footer

---
you are currently subscribed to customers as: zzzz-ryanair@example.com
to unsubscribe send a blank email to leave-
customers-949326k@mail.ryanairmail.com

--_nextpart_1_bvfoditvghtocxfdvjnkcuwblfv--
```

3.0.2 Evaluating Examples emails from above

My classificaton would classify the email as spam due to multiple characteristics that align with what my model was interested in. This email from RyanAir, an airline provider seems to be a promotional email to gain customers on their different credit card branch of business. The email seems to contain multiple repeating use cases of symbols (i.e = and *) along with excessive repeating letters towards the end of the email which are all key indicators of spam emails. One may disagree with my argument when considering this passenger may be traveling abroad to international waters and may find this specific email vital to their foreign plans to secure a line for communication carrying mutual benefits from their airline provider. One may also argue that the user opted in for these intentionally and thus would like to be prompted with such promotions.

Before making concrete conclusions, it is pivotal to trace back the tabular data itself and the methods used to retrieve the emails as many of our preliminary assumptions are at risk of being violated conditional on the methods used. For example, in the case of data retrieval and processing of the email if symbols, html tags or letter cases were misinterpreted, then these errors can falsely disguise truly ham emails as spam.

When assessing the ambiguity of initial classifications of our training emails (deemed as true values of our response variable) we are relying that the "users" used their best judgement to classify the email, thus opening doors to ambiguity within the true identity of the email. This alludes to the question if those labeled as spam in our training were truthfully ham, but mislabeled by "users" in the preliminary stages as spam due to some organizing reason for example and thus given such assignment. Therefore, ambiguity and the origins of the training data are important and reveal that False Positives are worth investigating.

Examine how a particular feature influences how an email is classified.

[72]:
```
# Simple model introduced at the start of this notebook. Just pay attention to␣
↪the features.
some_words = ['drug', 'bank', 'prescription', 'memo', 'private']
```

```
X_train = words_in_texts(some_words, train['email'])
Y_train = np.array(train['spam'])

simple_model = LogisticRegression()
simple_model.fit(X_train, Y_train);
```

Steps: Pick an email from the training set and assign its index to `email_idx`. Then, find **one** feature used in `simple_model` such that **removing** it changes how that email is classified. Assign this feature to `feature_to_remove`.

[73]:
```
scratch = train.copy()
for word in some_words:
    scratch[word] = scratch['email'].str.count(word)
for i in some_words:
    display(scratch[scratch['spam'] == 1][~scratch['email'].str.
 ↪contains('html')].sort_values(by = i, ascending= False).head(5))
#select id: 5274
```

```
        id                                              subject  \
4125  5274  Subject: Free Excerpt;  Baby Makers, Loser Cho…
5619    36  Subject: Hey look at this, I can't believe how…
5084  4114                      Subject: Legal herb, anytime\n
5612  6785                      Subject: Legal herb, anytime\n
7211  3446                      Subject: NEW Pot Substitute!\n

                                                 email  spam  drug  bank  \
4125  \n                            foreword\n \n afte…     1     7     0
5619  \n educate yourself about everything you ever …     1     6     1
5084  ***************************\n now open seven …     1     4     0
5612  ***************************\n now open seven …     1     4     0
7211  >from the ethnobotanical herbalists who brough…     1     4     0

      prescription  memo  private
4125             0     0        0
5619             0     0        1
5084             2     0        0
5612             2     0        0
7211             2     0        0
```

```
        id                                              subject  \
5925  3798                          Subject: URGENT REPLY.\n
2347   941                       Subject: Next of kin needed?\n
2688  1222                          Subject: [ILUG] BUSINESS\n
1338   347  Subject: FW: Make Money Fast And Legal! As See…
4948  1249  Subject: BUSINESS PARTNERSHIP(URGENT/CONFIDENT…

                                                 email  spam  drug  bank  \
```

```
5925  mr.ronard tony\n wema bank plc. \n lagos/niger…     1     0     9
2347  dear sir,\n \n my name is mr. obi w, the manag…     1     0     7
2688  central bank of nigeria\n foreign remittance d…     1     0     6
1338  the ultimate way to work from home \n the best…     1     0     6
4948  mr.vincent nnaji,\n  standard trust bank ltd,\…     1     0     6

      prescription  memo  private
5925             0     0        4
2347             0     0        1
2688             0     0        2
1338             0     0        0
4948             0     0        2

        id                                              subject  \
7296  1802  Subject: The only medically proven way to lose…
4481  5036                                  Subject: Faeries\n
6132  1645                                  Subject: Faeries\n
4626  3839              Subject: Have You Never Been Mellow?\n
478   5425  Subject: Online Doctors will fill your Viagra …

                                                    email  spam  drug  bank  \
7296  below is the result of your feedback form.  it…     1     1     0
4481  uncommon exotic pleasure botanicals!\n \n feel…     1     3     0
6132  uncommon exotic pleasure botanicals!\n \n feel…     1     3     0
4626  \n greetings & blessings to you!\n \n offering…     1     0     0
478   your sex drive should never be second on the l…     1     1     0

      prescription  memo  private
7296             3     0        0
4481             3     0        0
6132             3     0        0
4626             3     0        0
478              2     0        0

        id                                              subject  \
5116  6749                           Subject: windows tips\n
927   7234                         Subject: [ILUG] ASSISTANCE\n
7371  3304  Subject: Ultimate HGH: Make you look and feel …
5686  5703      Subject: A youthful and slim summer in 2002\n
2001  2519  Subject: Hgh: safe and effective release of yo…

                                                    email  spam  drug  bank  \
5116  \n \n are you having trouble with your compute…     1     0     0
927   from: col. michael bundu. \n democratic republ…     1     0     0
7371  as seen on nbc, cbs, cnn, and even oprah! the …     1     0     0
5686  as seen on nbc, cbs, cnn, and even oprah! the …     1     0     0
2001  as seen on nbc, cbs, cnn, and even oprah! the …     1     0     0

      prescription  memo  private
```

```
5116             0     2        0
927              0     1        2
7371             0     1        0
5686             0     1        0
2001             0     1        0

        id                                                subject  \
5925  3798                              Subject: URGENT REPLY.\n
6896  2786             Subject: [ILUG-Social] urgent assistance\n
4948  1249  Subject: BUSINESS PARTNERSHIP(URGENT/CONFIDENT…
5042  3364         Subject: Assistance requested(Please Read).\n
3100  3366                       Subject: **urgent assistance**\n

                                                     email  spam  drug  bank  \
5925  mr.ronard tony\n wema bank plc. \n lagos/niger…     1     0     9
6896  attn:\n \n i am edward mulete jr. the son of m…     1     0     0
4948  mr.vincent nnaji,\n  standard trust bank ltd,\…     1     0     6
5042  dear sir/ma, \n i am hajiya maryam abacha, wif…     1     0     2
3100  5, meridian east\n leicester le3 2wz \n leices…     1     0     0

      prescription  memo  private
5925             0     0        4
6896             0     0        2
4948             0     0        2
5042             0     0        2
3100             0     0        2
```

[74]:
```
email_idx = 1338

prob_spam = simple_model.predict_proba(X_train)[:, 1]
initial_prob = prob_spam[email_idx]
initial_class = "spam" if np.round(initial_prob) else "ham"
print(f"\nPredicted probability of being spam: {np.round(initial_prob*100,␣
 ↪2)}%")
print("\nEmail:\n" + train.loc[email_idx]["email"])
```

```
Predicted probability of being spam: 55.57%

Email:
the ultimate way to work from home
 the best money making system of all!!
 as seen on national tv
 as seen on 20/20 and many other credible references. this is not a scam.
 i hope this is ok that i send you this. if you aren't interested, just
 simply delete it.
 read this message if you are like me and want more than your lousy
 weekly paycheck. make more in a few months than last year at work.
```

believe it, work it.
this really works, don't make the same mistake i made. i deleted this
4-5 times before finally giving it a try. within 2 weeks the orders
(money)
started coming in just like the plan below said it would. give it a
try!!
you will be glad you did.
thanks and good luck! you won't need luck, just keep reading. don't
delete this!!!
---
first read about how a 15 year old made $71,000. see below…
as seen on national tv: this is the media report.
parents of 15 - year old - find $71,000 cash hidden
in his closet!
does this headline look familiar? of course it does.
you most likely have just seen this story recently
featured on a major nightly news program (usa).
and reported elsewhere in the world (including my neck
of the woods - new zealand). his mother was cleaning
and putting laundry away when she came across a large
brown paper bag that was suspiciously buried beneath
some clothes and a skateboard in the back of her
15-year-old sons closet.
nothing could have prepared her for the shock she got
when she opened the bag and found it was full of cash.
five-dollar bills, twenties, fifties and hundreds -
all neatly rubber-banded in labeled piles.
"my first thought was that he had robbed a bank",
says the 41-year-old woman, "there was over $71,000
dollars in that bag -- that's more than my husband
earns in a year".
the woman immediately called her husband at the
car-dealership where he worked to tell him what she
had discovered. he came home right away and they
drove together to the boys school and picked him up.
little did they suspect that where the money came from
was more shocking than actually finding it in
the closet.
as it turns out, the boy had been sending out, via
e-mail, a type of "report" to e-mail addresses that he
obtained off the internet. everyday after school for
the past 2 months, he had been doing this right on his
computer in his bedroom.
"i just got the e-mail one day and i figured what the
heck, i put my name on it like the instructions said
and i started sending it out", says the clever
15-year-old.
the e-mail letter listed 5 addresses and contained

instructions to send one $5 dollar bill to each person on the list, then delete the address at the top and move the others addresses down, and finally to add your name to the top of the list.
the letter goes on to state that you would receive several thousand dollars in five-dollar bills within 2 weeks if you sent out the letter with your name at the top of the 5-address list. "i get junk e-mail all the time, and really did not think it was going to work", the boy continues.
within the first few days of sending out the e-mail, the post office box that his parents had gotten him for his video-game magazine subscriptions began to fill up with not magazines, but envelopes containing $5 bills.
"about a week later i rode [my bike] down to the post office and my box had 1 magazine and about 300 envelops stuffed in it. there was also a yellow slip that said i had to go up to the [post office] counter. i thought i was in trouble or something (laughs)". he goes on, "i went up to the counter and they had a whole box of more mail for me. i had to ride back home and empty out my backpack because i could not carry it all".
over the next few weeks, the boy continued sending out the e-mail. "the money just kept coming in and i just kept sorting it and stashing it in the closet, barely had time for my homework". he had also been riding his bike to several of the banks in his area and exchanging the $5 bills for twenties, fifties and hundreds.
"i didn't want the banks to get suspicious so i kept riding to different banks with like five thousand at a time in my backpack. i would usually tell the lady at the bank counter that my dad had sent me in to exchange the money and he was outside waiting for me. one time the lady gave me a really strange look and told me that she would not be able to do it for me and my dad would have to come in and do it, but i just rode to the next bank down the street (laughs)."
surprisingly, the boy did not have any reason to be afraid. the reporting news team examined and investigated the so-called "chain-letter" the boy was sending out and found that it was not a chain-letter at all. in fact, it was completely legal according to us postal and lottery laws, title 18, section 1302 and 1341, or title 18, section 3005 in the us code, also in the code of federal regulations, volume 16,

sections 255 and 436, which state a product or service
must be exchanged for money received.
every five-dollar bill that he received contained a
little note that read, "please send me report number
xyx".this simple note made the letter legal because he
was exchanging a service (a report on how-to) for a
five-dollar fee.
[this is the end of the media release. if you would
like to understand how the system works and get your
$71,000 - please continue reading. what appears below
is what the 15 year old was sending out on the net -
you can use it too - just follow the simple
instructions].
thanks to the computer age and the internet !
==========
you will make over half million dollars every 4 to 5 months from your
home!!
before you say ''bull'', please read the following. this is the
letter you have been hearing about on the news lately. due to the
popularity of this letter on the internet, a national weekly news
program
recently devoted an entire show to the investigation of this program
described below, to see if it really can make people money. the show
also
investigated whether or not the program was legal.
their findings proved once and for all that there are ''absolutely no
laws prohibiting the participation in the program and if people can -follow
the simple instructions, they are bound to make some mega bucks with only
$25 out of pocket cost''.
due to the recent increase of popularity & respect this
program has
attained, it is currently working better than ever.
this is what one had to say: '' thanks to this profitable opportunity.
i was approached many times before but each time i passed on it. i am so
glad i finally joined just to see what one could expect in return for
the minimal effort and money required. to my astonishment, i received
total
$610,470.00 in 21 weeks, with money still coming in''.
pam hedland, fort lee, new jersey.
==========
=======here is
another testimonial: ''' this program has been around for a
long time but i never believed in it. but one day when i received this
again in the mail i decided to gamble my $25 on it. i followed the
simple instructions and wallaa … 3 weeks later the money started to
come
in. first month i only made $240.00 but the next 2 months after that
made a total of $290,000.00. so far, in the past 8 months by re-entering

the
program, i have made over $710,000.00 and i am playing it again. the
key to success in this program is to follow the simple steps and not change
anything.'' more testimonials later but first:
===== print this now for your future reference ========
$$$
$$$$$$$$
if you would like to make at least $500,000 every 4 to 5 months easily
and comfortably, please read the following…then read it again and
again!!!
$$$
$$$$$$$$
follow the simple instruction below and your financial
dreams will come
true, guaranteed!
instructions:
================order all 5 reports shown on the list below
================
for each report, send $5 cash, the name & number of the report
you are ordering and your e-mail address to the person whose
name appears on that list next to the report. make sure your return
address is on your envelope top left corner in case of any mail
problems.
=== when you place your order, make sure you order each of the 5
reports.
you will need all 5 reports so that you can save them on your computer
and resell them. your total cost $5 x 5=$25.00.
within a few days you will receive, via e-mail, each of the 5 reports
from these 5 different individuals. save them on your computer so they will
be accessible for you to send to the 1,000's of people who will order them
from you. also make a floppy of these reports and keep it on your desk
in case something happens to your computer.
important - do not alter the names of the people who are listed
next to each report, or their sequence on the list, in any way other
than what is instructed below in step '' 1 through 6 '' or you will
loose out on majority of your profits. once you understand the way this
works, you will also see how it does not work if you change it. remember,
this
method has been tested, and if you alter, it will not work !!! people
have tried to put their friends/relatives names on all five thinking they
could get all the money. but it does not work this way. believe us, we all
have tried to be greedy and then nothing happened. so do not try to change
anything other than what is instructed. because if you do, it will not
work for you.
remember, honesty reaps the reward!!!
1… after you have ordered all 5 reports, take this advertisement
and remove the name & address of the person in report #5.
this person has made it through the cycle and is no doubt counting

their fortune.
2… move the name & address in report # 4 down to report # 5.
3… move the name & address in report # 3 down to report # 4.
4… move the name & address in report # 2 down to report # 3.
5… move the name & address in report # 1 down to report # 2
6… insert your name & address in the report # 1 position.
please make sure you copy every name & address accurately!
==========
=======
**** take this entire letter, with the modified list of names, and save it on your computer. do not make any other changes.
save this on a disk as well just in case you loose any data. to assist you with marketing your business on the internet, the 5 reports you purchase will provide you with invaluable marketing information which includes how to send bulk e-mails legally, where to find thousands of free
classified ads and much more. there are 2 primary methods to get this venture going:
method # 1: by sending bulk e-mail legally
==========
=======
let's say that you decide to start small, just to see how it goes, and we will assume you and those involved send out only 5,000 e-mails each. let's also assume that the mailing receive only a 0.2% response (the response
could be much better but lets just say it is only 0.2%. also many people will send out hundreds of thousands e-mails instead of only 5,000 each).continuing with this example, you send out only 5,000 e-mails. with a 0.2% response, that is only 10 orders for report # 1
those 10 people responded by sending out 5,000 e-mail each for a total of 50,000.
out of those 50,000 e-mails only 0.2% responded with orders. that's=100 people responded and ordered report # 2.
those 100 people mail out 5,000 e-mails each for a total of 500,000 e-mails. the 0.2% response to that is 1000 orders for report # 3.
those 1000 people send out 5,000 e-mails each for a total of 5 million e-mails sent out. the 0.2% response to that is 10,000 orders for report! # 4. whose 10,000 people send out 5,000 e-mails each for a total of
50,000,000 (50 million) e-mails. the 0.2% response to that is 100,000 orders for
report # 5 that's 100,000 orders times $5 each=$500,000.00 (half million).
your total income in this example is: 1… $50 + 2… $500
+ 3…$5,000 + 4… $50,000 + 5… $500,000 … grand
total=$555,550.00
numbers do not lie. get a pencil & paper and figure out the worst possibleresponses and no matter how you

calculate it, you will still
make a lot of money !
==========
=======
remember friend, this is assuming only 10 people ordering
out
of 5,000 you mailed to. dare to think for a moment what would happen if
everyone or half or even one 4th of those people mailed 100,000e-mails
each or more? there are over 150 million people on the internet worldwide
and counting. believe me, many people will do just that, and more!
method # 2 : by placing free ads on the internet
==========
=======
advertising on the net is very very inexpensive and there are
hundreds of free places to advertise. placing a lot of free ads on the
internet will easily get a larger response. we strongly suggest you
start with method # 1and add method # 2 as you go along. for every $5 you
receive, all you must do is e-mail them the report they ordered. that's
it. always provide
same day service on all orders.
this will guarantee that the e-mail they send out, with your
name and address on it, will be prompt because they can not advertise
until they receive the report.
==========available reports
==========
order each report by its number & name only. notes: always
send
$5 cash (u.s. currency) for each report. checks not accepted. make sure
the cash is concealed by wrapping it in at least 2 sheets of paper. on
one of those sheets of paper, write the number & the name of the report you
are ordering,
your e-mail address and your name and postal address.
place your order for these reports now :
==========
=======
report #1 "the insider's guide to advertising for free on the net"
order report #1 from:
k.palludan
9550 summersweet ct.
las vegas, nevada 89123
usa

---
report #2 "'the insider's guide to sending bulk e-mail on the net'
order report #2 from:
kris estes
3055 casey drive #203
las vegas, nevada 89120

```
------------------------------------------------------------------
------
report #3 "'secret to multilevel marketing on the net"
order report #3 from :
p. clement
601 st-malo est
ile bizard, quebec h9c2p2
canada
------------------------------------------------------------------
------
report #4 "'how to become a millionaire utilizing mlm and the net"
order report #4 from:
david carpenter
13 dutch lane
hazlet, new jersey 07732
usa
------------------------------------------------------------------
-----
report #5 "'how to send out one million e-mails for free"
order report #5 from:
s. lasley
4230 edgewood circle
idaho falls, idaho 83406
usa
------------------------------------------------------------------
-----
$$$$$$$$$$$$$$$$$$$$ your success guidelines
$$$$$$$$$$$$$$$$$$$$
follow these guidelines to guarantee your success:
=== if you do not receive at least 10 orders for report #1 within 2
weeks, continue sending e-mails until you do.
=== after you have received 10 orders, 2 to 3 weeks after that you
should receive 100 orders or more for report # 2. if you did not, continue
advertising or sending e-mails until you do.
=== once you have received 100 or more orders for report # 2, you can
relax, because the system is already working for you, and the cash will
continue to roll in ! this is important to remember: every time your
name
is moved down on the list, you are placed in front of a different report.
you can keep track of your progress by watching which report
people are ordering from you. if you want to generate more income
send
another batch of e-mails and start the whole process
again. there is no
limit to the income you can generate from this business !!!
============================================================
=======
following is a note from the originator of this program:
```

you have just received information that can give you financial freedom
for the rest of your life, with no risk and just a little bit of effort.
you can make more money in the next few weeks and months than you have
ever imagined. follow the program exactly as instructed. do not change
it
in
any way. it works exceedingly well as it is now.
remember to e-mail a copy of this exciting report after you have put
your name and address in report #1 and moved others to #2 …# 5
as
instructed above. one of the people you send this to may send out
100,000 or more e-mails and your name will be on every one of them.
remember
though, the more you send out the more potential customers! you will
reach. so
my friend, i have given you the ideas, information, materials and
opportunity to become financially independent. it is up to you now !
========== more
testimonials==========
'' my name is mitchell. my wife, jody and i live in chicago. i am an
accountant with a major u.s. corporation and i make pretty good
money. when i received this program i grumbled to jody about receiving
''junk mail''. i made fun of the whole thing, spouting my knowledge of
the population and percentages involved. i ''knew'' it wouldn't work. jody
totally ignored my supposed intelligence and few days later she jumped in
with
both feet. i made merciless fun of her, and was ready to lay the old
''i told you so'' on her when the thing didn't work. well, the laugh was
on me!
within 3 weeks she had received 50 responses. within the next 45 days she
had
received total $ 147,200.00 … all cash! i was shocked. ! i have
joined
jody in her ''hobby''.
mitchell wolf m.d., chicago, illinois
==========
======= '' not
being the gambling type, it took me several weeks to make up my
mind to participate in this plan. but conservative that i am, i decided
that the initial investment was so little that there was just no way that
i
wouldn't get enough orders to at least get my money back''. '' i was
surprised
when i found my medium size post office box crammed with orders. i made
$319,210.00 in the first 12 weeks. the nice thing about this deal is
that it does not matter where people live. there simply isn't a better
investment
with a faster return and so big''.

```
dan sondstrom, alberta, canada
============================================================
======= '' i had
received this program before. i deleted it, but later i
wondered if i should have given it a try. of course, i had no idea who
to contact ! to get another copy, so i had to wait until i was e-mailed
again by someone else…11 months passed then it luckily came
again…i did not delete this one! i made more than $490,000 on my
first try and all the money came within 22 weeks''.
susan de suza, new york, n.y.
============================================================
======
'' it really is a great opportunity to make relatively easy money with
little cost to you. i followed the simple instructions carefully and
within 10 days the money started to come in. my first month i made
$20,560.00 and by the end of third month my total cash count was $
362,840.00. life is beautiful, thanx to internet''.
fred dellaca, westport, new zealand
============================================================
=======order
your reports today and get started on your road to
financial freedom !
============================================================
=======
if you have any questions of the legality of this program, contact the
office of associate director for marketing practices, federal trade
commission, bureau of consumer protection, washington, d.c.
============================================================
=======
there is no need to respond to this e-mail if you do not
wish to receive
further correspondence. this is a onetime e-mail.
good luck!
```

[75]:
```
feature_to_remove = 'bank'

changed_words = some_words.copy()
changed_words.remove(feature_to_remove)

changed_model = LogisticRegression()
X_changed = words_in_texts(changed_words, train['email'])
y = train['spam']
changed_model.fit(X_changed, y)
```

```
changed_prob = changed_model.predict_proba(X_changed[[email_idx]])[:,1][0]
changed_class = "spam" if np.round(changed_prob) else "ham"

print(f"Initially classified as {initial_class} (Probability: {np.
  ↪round(initial_prob*100, 2)}%)")
print(f"Now classified as {changed_class} (Probability: {np.
  ↪round(changed_prob*100, 2)}%)")
```

```
Initially classified as spam (Probability: 55.57%)
Now classified as ham (Probability: 24.33%)
```

How feature changed how the email that was classified. After iterating through each word and counting its occurence in each email text, I drilled down on those emails that were dependent on a specific feature word that classified it as spam and removed that word, in this case being 'bank' for the 1338th indexed email. Prior to the removal of the feature, I found that this email was at a borderline probability to being classified as ham at 55%. Considering that 'bank' was the only word that aligned the text to the words list in occurrence counts, the removal of 'bank' would shift the classification below 50% (default cutoff in the logisitc regression) and shift from spam to ham.

Interpretability Assuming the large number of word features contained in the new model prove to be more prevalent in one classification and minimze overlap, then I would expect this new model to more interpetable than simple_model as the model will now have higher confidence and probability in classifying the emails as it will be able to leverage the potential combinations and reduce dependencies on single words in the simple_model. For example, there was a high likelihood that overlap and high number of False Positives would appear in simple_model in the contexts of dealing with ham emails circulating in the finance or health industry (hence the word selection being drug, bank, prescription, etc…).

When dealing with a classification model working with hate speech, the inevitable obstacles such as diverse idioms, differing languages/phrases, misinterpration when truly raising awareness and circulation of images/memes all pose challenges in classifying hate speech. Content that would fall under such category at an amatuer level would be widely used deragatory terms tied with sensitive global conflicts, race/ethnicity, gender, etc. Tokenizing such combinations would pinpoint essential hate speeches that would be essential to remove and classify as hate speech. It is also key to contian versatility to generalize to all languages involved to accurately apply a classification tailored for a language.

3.0.3 Stakes of Misclassification

Misclassifying posts on social media can have implications on the company's image and stance towards the respective type of post removed. Further misinfomration on the platform and its stakeholders can be stunted. A false positive in this context would be a truly respective/guideline abiding post classified as hate speech, while a false negative could be a truly hate speech containing post that bypassed the classificationt test and is classified as an email that meets platform guidelines. Such issues are common amongst social media platforms and are ongoing tasks for the major players as it leads users to attack the integrity of the platform and question what it allows or restricts (i.e X contantly changing its guidelines after change in leadership).

3.0.4 Relevance of having Interpretable Model?

Having an interpretable model is pivotal to meeting social images and allowing a fair/equitable platform to all types of users that invites opinions. Thus having a model that is widely resilient against varying circumstances is important. Although the usefulness and perception of the interpretibility of the model can be skewed short term due to sudden rise in global tensions/ or rise in some form of category above that the model hasn't been trained on would signal to the developers to adjust their models to offer consieration in its classification for the repsective domain. Failing to meet such conflicts can threaten the integrity of the models and the platform itself.