Classification

Spam or Ham

Key Points

- Data
- Data Split
- Data Cleaning
- Evaluation Metric
- Impact of category Encoding (ham as 1 or spam as 1)
- Handing Imbalance data
- Choosing right probability threshold
- Where model went wrong
- Demo with giving an example Text
- Comparison

Spam Examples

1: Free entry in 2 a wkly comp to win FA Cup final tkts 21st May 2005. Text FA to 87121 to receive entry question(std txt rate)T&C's apply 08452810075over18's

2: FreeMsg Hey there darling it's been 3 week's now and no word back! I'd like some fun you up for it still? Tb ok! XxX std chgs to send, å£1.50 to rcv

3: WINNER!! As a valued network customer you have been selected to receivea å£900 prize reward! To claim call 09061701461. Claim code KL341. Valid 12 hours only.

4: Had your mobile 11 months or more? U R entitled to Update to the latest colour mobiles with camera for Free! Call The Mobile Update Co FREE on 08002986030

5: SIX chances to win CASH! From 100 to 20,000 pounds txt> CSH11 and send to 87575. Cost 150p/day, 6days, 16+ TsandCs apply Reply HL 4 info

6: URGENT! You have won a 1 week FREE membership in our å£100,000 Prize Jackpot! Txt the word: CLAIM to No: 81010 T&C www.dbuk.net LCCLTD POBOX 4403LDNW1A7RW18

7: XXXMobileMovieClub: To use your credit, click the WAP link in the next txt message or click here>> http://wap. xxxmobilemovieclub.com?n=QJKGIGHJJGCBL

8: England v Macedonia - dont miss the goals/team news. Txt ur national team to 87077 eg ENGLAND to 87077 Try:WALES, SCOTLAND 4txt/Ì¼1.20 POBOXox36504W45WQ 16+

9: Thanks for your subscription to Ringtone UK your mobile will be charged å£5/month Please confirm by replying YES or NO. If you reply NO you will not be charged

10: 07732584351 - Rodger Burns - MSG = We tried to call you re your reply to our sms for a free nokia mobile + free camcorder. Please call now 08000930705 for delivery tomorrow

Ham Examples

1: Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...

2: Ok lar... Joking wif u oni...

3: U dun say so early hor... U c already then say...

4: Nah I don't think he goes to usf, he lives around here though

5: Even my brother is not like to speak with me. They treat me like aids patent.

6: As per your request 'Melle Melle (Oru Minnaminunginte Nurungu Vettam)' has been set as your callertune for all Callers. Press *9 to copy your friends Callertune

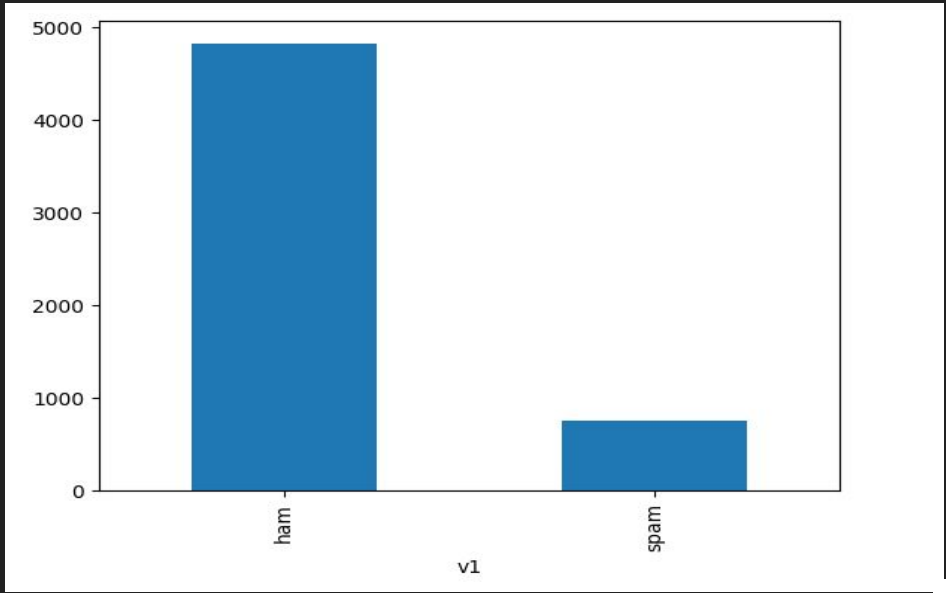
7: I'm gonna be home soon and i don't want to talk about this stuff anymore tonight, k? I've cried enough today.

8: I've been searching for the right words to thank you for this breather. I promise i wont take your help for granted and will fulfil my promise. You have been wonderful and a blessing at all times.

9: I HAVE A DATE ON SUNDAY WITH WILL!!

10: Oh k...i'm watching here:)

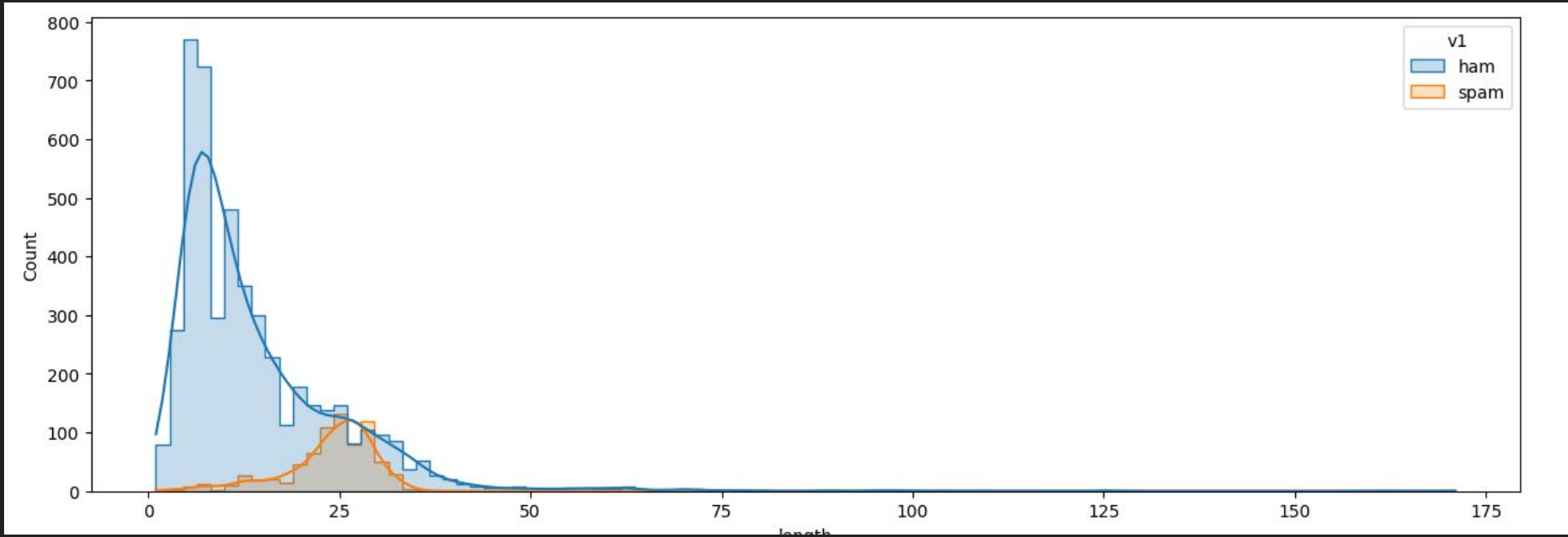
Data Distribution (87% ham data points) in total of 5572




```
train_set: ham_ratio: 0.87, spam_ratio: 0.13, Total: 4457
valid_set: ham_ratio: 0.87  spam_ratio: 0.13, Total: 558
test_set : ham_ratio: 0.87  spam_ratio: 0.13, Total: 557
```

Words in each category

Distribution of Words in each category



Evaluation Metric

I will be selecting the model which would give me the best F1-Score on test_set. Note that F1-Score in itself can vary greatly depending upon specific technique you use to calculate F1-Score. There are three variation of F1-Score

1. F1_score with macro averaging: It calculates f1-score for each class individually and then take average of all of them by giving equal weightage.

2. F1_score with weighted averaging: It calculates f1-score for each class individually and then take the weighted average of each class's f1-score

3. F1_Score: It calculates the f1 score for the positive class(the class which you have encoded as 1). Now this particular technique very much depends on which class you have encoded as 1. For example: In this task if we encode ham as 1 then we will get a higher f1_score

Impact of choosing wrong evaluation metric

Total Data Points = 100

Ham = 90

Spam = 10

Your Model = Model which predict every data point as HAM. Now-

- Scores for Ham
  - Precision = 90 / 100
  - Recall = 90 / 90
- Scores for Spam
  - Precision = 0 / 100
  - Recall = 0 / 10

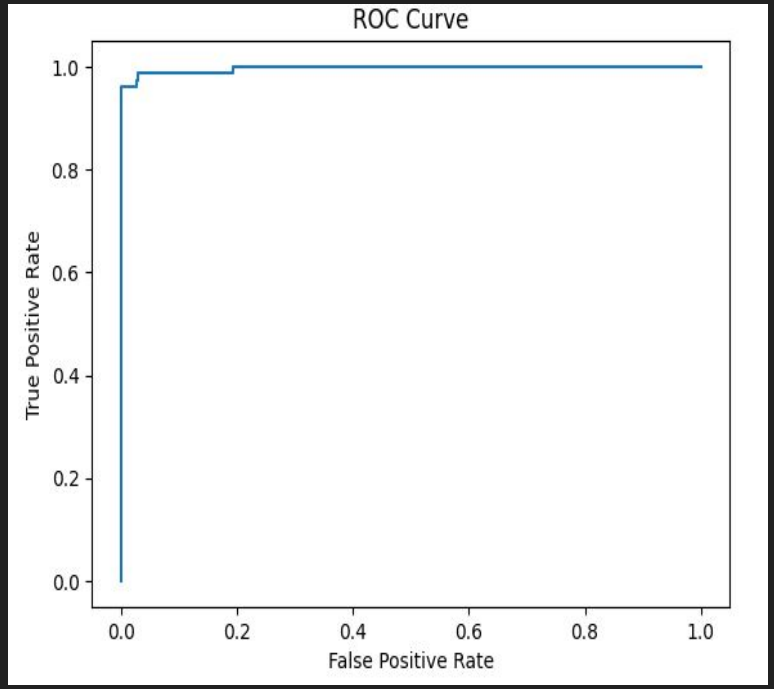
BERT performance on validation set

```
y_pred_ = list(map(lambda x: 1 if x>0.5 else 0, logits))
score = {
'f1':[f1_score(val_y_t, y_pred_, average='macro')],
'precision':[precision_score(val_y_t, y_pred_, average='macro')],
'recall':[recall_score(val_y_t, y_pred_, average='macro')],
'accuracy':[accuracy_score(val_y_t, y_pred_)]
}

val_res = pd.DataFrame(score)
val_res.index = ['BERT']
val_res
```

	f1	precision	recall	accuracy
BERT	0.98442	0.990058	0.978965	0.992832

Selecting right probability threshold

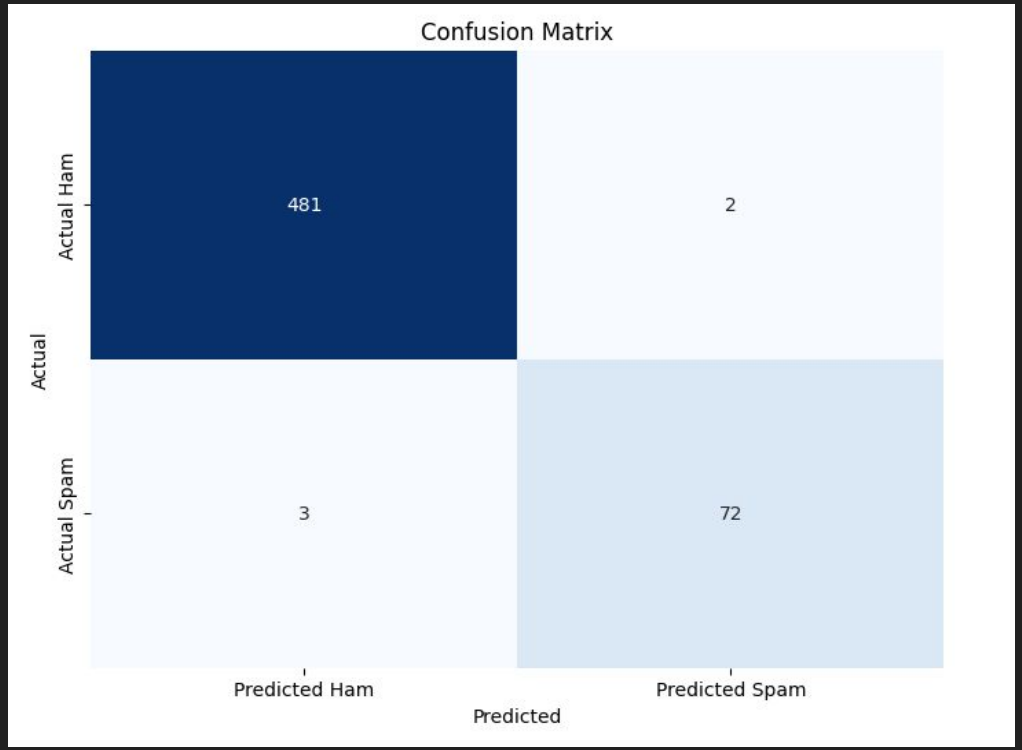



```
y_pred_ = list(map(lambda x: 1 if x>0.25 else 0, logits))
score = {
'f1':[f1_score(val_y_t, y_pred_, average='macro')],
'precision':[precision_score(val_y_t, y_pred_, average='macro')],
'recall':[recall_score(val_y_t, y_pred_, average='macro')],
'accuracy':[accuracy_score(val_y_t, y_pred_)]
}

val_res = pd.DataFrame(score)
val_res.index = ['BERT']
val_res
```

	f1	precision	recall	accuracy
BERT	0.980636	0.983387	0.97793	0.991039

Where model went wrong (5 incorrect predictions out of 558)



Text that are predicted as Ham, but they were actually Spam. ¶

+ Code + Markdown

```
]:
val_df = val_data['v2'].reset_index(drop=True)
val_y = val_y_t.reset_index(drop=True)
val_pred = pd.Series(y_pred_)

temp = val_df[(val_pred==0) & (val_y==1)]
for index in temp.index:
    #print(f"Original Text : {u_train[index]}")
    print(f"Text : {temp[index]}")
    print("="*150)
```

```
Text : Bought one ringtone and now getting texts costing 3 pound offering more tones etc
======================================================================================================================================================
Text : RCT' THNQ Adrian for U text. Rgds Vatian
======================================================================================================================================================
Text : Do you realize that in about 40 years, we'll have thousands of old ladies running around with tattoos?
======================================================================================================================================================
```

Text that are predicted as Spam, but they were actually Ham

+ Code + Markdown

```
val_df = val_data['v2'].reset_index(drop=True)
val_y = val_y_t.reset_index(drop=True)
val_pred = pd.Series(y_pred_)

temp = val_df[(val_pred==1) & (val_y==0)]
for index in temp.index:
    #print(f"Original Text : {u_train[index]}")
    print(f"Text : {temp[index]}")
    print("="*150)
```

```
Text : Somebody set up a website where you can play hold em using eve online spacebucks
======================================================================================================================================================
Text : HCL chennai requires FRESHERS for voice process.Excellent english needed.Salary upto  &lt;#&gt; .Call Ms.Suman  &lt;#&gt;  for Telephonic interview -via Indyarocks.com
======================================================================================================================================================
```

Performance on test set

	f1	precision	recall	accuracy
BERT	0.980408	0.983202	0.977659	0.991023

```
print(classification_report(y_test_, y_pred_))
```

```
              precision    recall  f1-score   support

           0       0.99      1.00      0.99       483
           1       0.97      0.96      0.97        74

    accuracy                           0.99       557
   macro avg       0.98      0.98      0.98       557
weighted avg       0.99      0.99      0.99       557
```

Thank you