

CNN 스팸 메일 분류

21016006 김석현

data

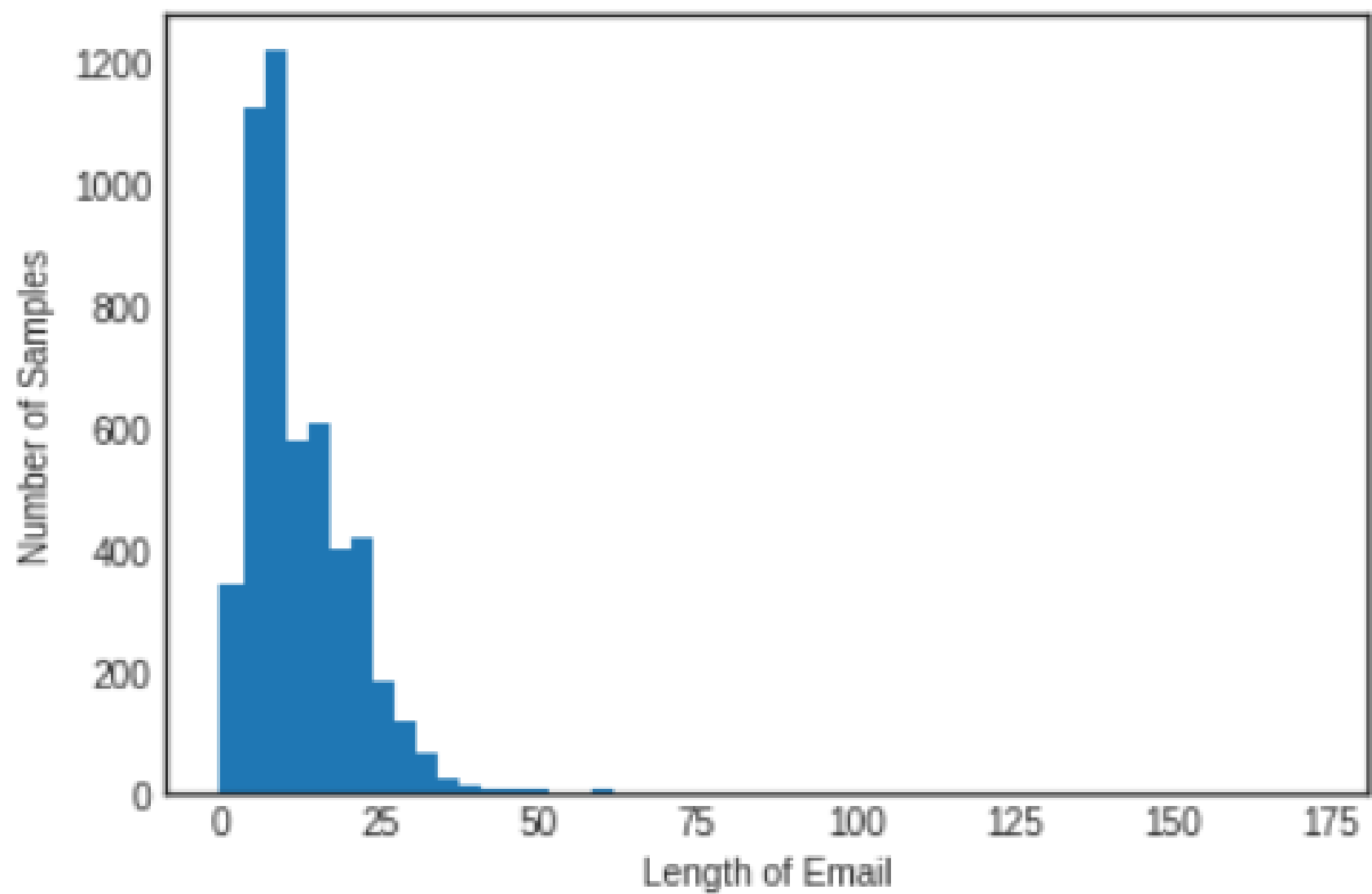
<https://raw.githubusercontent.com/mohitgupta-omg/Kaggle-SMS-Spam-Collection-Dataset/master/spam.csv>

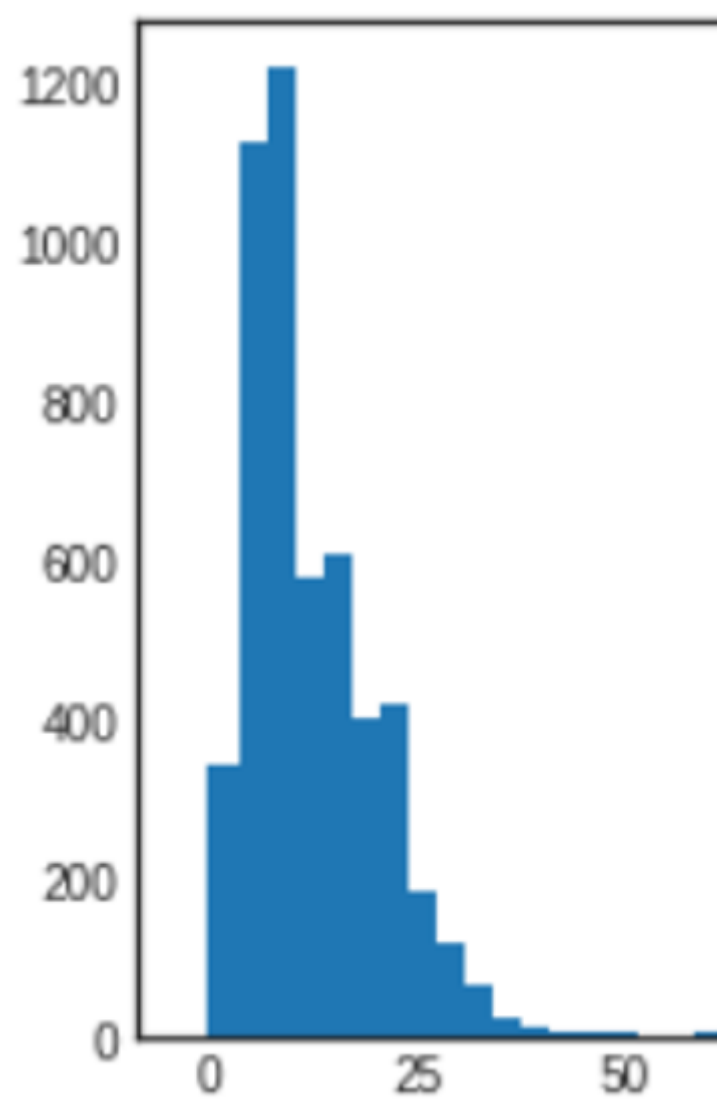
	v1	v2
0	0	Go until jurong point, crazy.. Available only ...
1	0	Ok lar... Joking wif u oni...
2	1	Free entry in 2 a wkly comp to win FA Cup fina...
3	0	U dun say so early hor... U c already then say...
4	0	Nah I don't think he goes to usf, he lives aro...
5	1	FreeMsg Hey there darling it's been 3 week's n...
6	0	Even my brother is not like to speak with me. ...
7	0	As per your request 'Melle Melle (Oru Minnamin...
8	1	WINNER!! As a valued network customer you have...
9	1	Had your mobile 11 months or more? U R entitle...

두가지 전처리

drop_duplicates

5572 ➔ 5169





모델구성


```
model=Sequential()  
model.add(Embedding(vocab_size,32))  
model.add(Dropout(0.2))  
model.add(Conv1D(32,5, strides=1, padding='valid', activation='relu'))  
model.add(GlobalMaxPooling1D())  
model.add(Dense(64, activation='relu'))  
model.add(Dropout(0.2))  
model.add(Dense(1, activation='sigmoid'))  
model.summary()  
  
model.compile(optimizer='adam',  
              loss='binary_crossentropy',  
              metrics=['acc'])
```

