

Question and Answers:

1. What made you decide on a SMS spam classification for this project?
 - a. Well, email classification has a bit of a head start and the techniques are a bit more refined. I figured spam text messages were not quite as sophisticated and we had very clearly defined buckets we could sort messages in – spam or ham. It was less subjective than some categories we might define.
2. What made you decide on the two models you implemented?
 - a. Naïve Bayes is the classic classification algorithm – logistic regression as well. Both were fairly easy to implement, ran quickly, and gave us surprisingly good results given the above!
3. Why the word clouds?
 - a. They were in vogue for a few years and I thought they were a great look at the different word choices between ham and spam messages. Obviously, if we were trying to sell you something, we would choose different language. But I thought this hammered home that difference and so I felt including it was important.
4. What was the F-Score, and why was it used as a metric?
 - a. It is a numerical score for the harmonic mean of the precision and recall. This numerical score makes it easy to quantify how well a model performs. It can be hard to translate “well, this one got this many right and this one got this many right” and compare the two especially if they were inaccurate in different ways. Essentially, it tries to capture the measure of how accurate and precise the model is.
5. What was the most difficult challenge with the project?
 - a. Actually, removing the stopwords from the list! It took a very long time for my poor computer to run it – so much that I borrowed a timing technique I used in a previous simulation. In that project, I simulated time series data in real time – the reading from 20 seconds after start in our historical data was transmitted exactly 20 seconds after start. I used this to estimate how much longer the stopwords process would run, including elapsed time, estimated time remaining, and progress so far. I am hoping it was somewhat more accurate than Windows 95!
6. Are there other uses for this classification that you identify?
 - a. Sure! I talked a lot about customer service interactions. Many people I know will do anything they can to avoid talking on the phone with a person. Shooting off a text message to ask what is up with my Amazon order? I could get behind that. However, any time you have an open channel of communication like that you will want to have a system that can identify legitimate contacts.
7. Did the models you create work solely for spam or can they be adapted?
 - a. Absolutely! Our inputs were text strings with a classification of either ham or spam, but to the model it was assigning either a 1 or a 0 as the category. I could feed it grandma’s pie recipes mixed with shopping lists and it would do a pretty good job of sussing out which rows lead to secret family recipes and which resulted from a trip to Publix.
8. Any plans to extend the project further?
 - a. I alluded to it a bit in the end, but I think it would be a nice project to dabble in blending models with. We already have two models, that have very good results. We could see if we could push that accuracy or F-score even higher by blending the two models!

9. What other projects might you use the models for?
 - a. I really want to look into predictive text generation as well. I would love to generate paragraphs of text at different heat levels, combine it with actual paragraphs from the source text and see how well the model picks out the generated text compared to the source text. I feel that often the level is set sort of by convention and “yeah, that looks about right” but I would love to be able to graph the F-Score versus that to see if we had a nice elbow or point at which the model performs the best.
10. How do you think the model can be refined in the future?
 - a. Our test data set was a little on the smaller side and was not exclusively US based text messages. I feel like once we had this one in the wild, we could store the resulting text and classifications in order to add to our dataset. Especially if we had a moderating team, a quick yes or no flag could quickly add data to our collection.