The Naive Bayes Algorithm: Takeaways

Concepts

- When a new message "w_1, w_2, ..., w_n" comes in, the Naive Bayes algorithm classifies it as spam or non-spam based on the results of these two equations:

- To calculate $P(w_i|Spam)$ and $P(w_i|Spam^C)$, we need to use the additive smoothing technique:

- Below, we see what some of the terms in equations above mean:

Resources

- A technical intro to a few version of the Naive Bayes algorithm

- An intro to conditional independence