

- a. It's likely because of the naive assumption. The word counts or single data that lead to incorrect classification should have been seen many more times in one type of class than the other. Hence that leads the model to misclassify. So the naive/independence assumption doesn't allow the model to capture the context/semantics of the speeches. Which as seen in this case can lead the model to make incorrect classifications even with a high confidence level.
- b. It can even leave the patient dead or be exposed to a very heavy treatment. I would argue about not using the Naive Bayes Classifier in a high stake setting. The Naive assumption it makes is a very weak/ non-realistic assumption. Although it sometimes leads to quite accurate results, as seen in this case, there could be cases where it misclassifies with high confidence.
- c. The most straightforward benefit of it is that it informs the user about the probability or confidence level of its prediction. Which provides more transparency, and I think it aligns better with our current understanding of the universe.
- d. Instead of scaling the denominator with `self.smoothing*2`. I would scale the denominator with `self.smoothing*vocab_size`. This would arbitrarily set the starting probabilities for each word to be less than .5 .