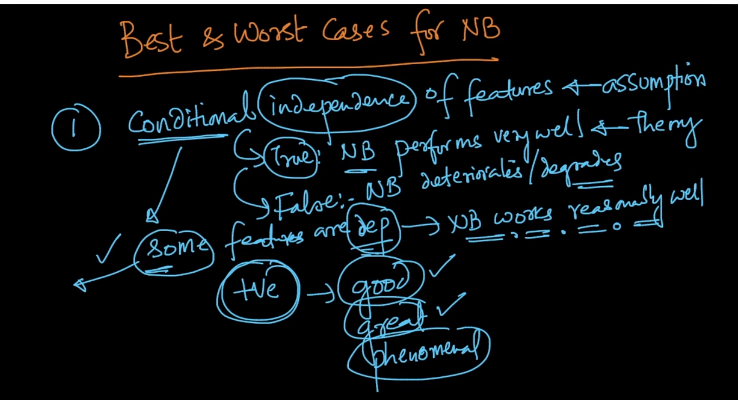
**Best and Worst cases.**

The first and most basic assumption for NB is feature independence and when this holds true NB performs very well.

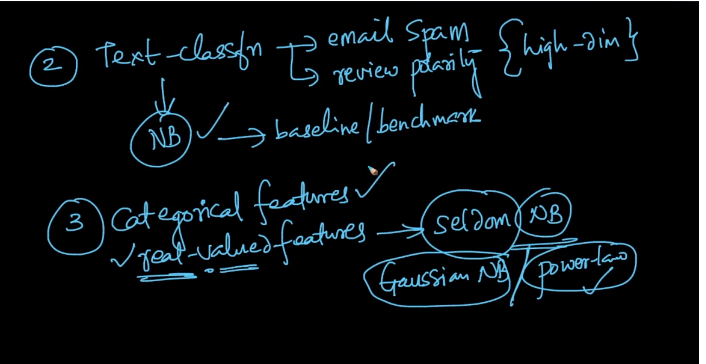
But in practical even when some feature are dependent NB performs reasonably well and it has some scientific proof behind it which is not required to be known.



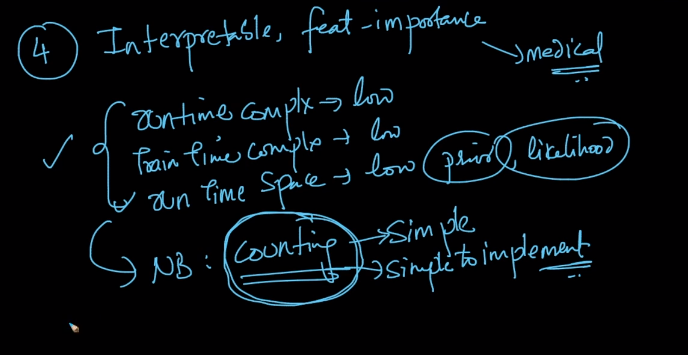
Secondly as we have seen NB performs exceptionally well in Text classifications like Email classification or Review polarity .

It has actually became benchmark for all other algorithms.

And we have also seen that NB is seldom used with numerical features,



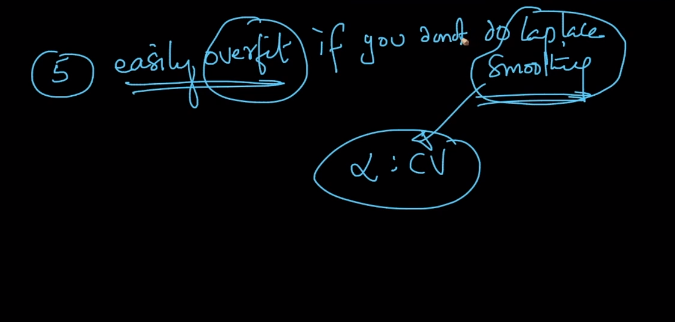
Another thing in NB is that it is super easy to understand , is very much interpretable and have very low run time and low space complexity at runtime.



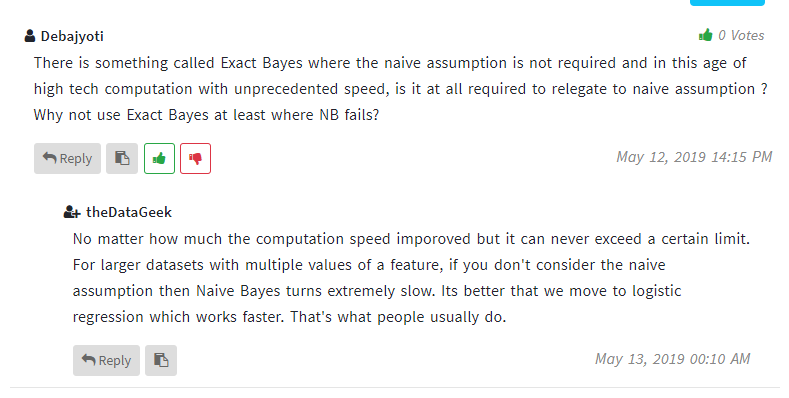
Very simple to implement also .

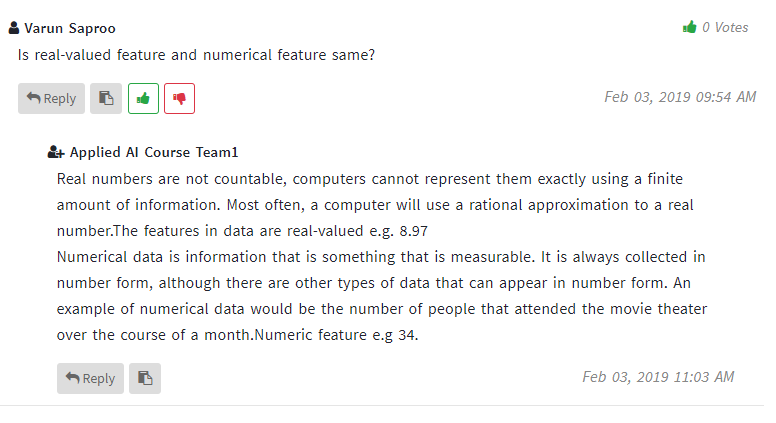
There is just one catch to NB, that it gets overfit very easily if we don’t apply Laplace Smoothing.

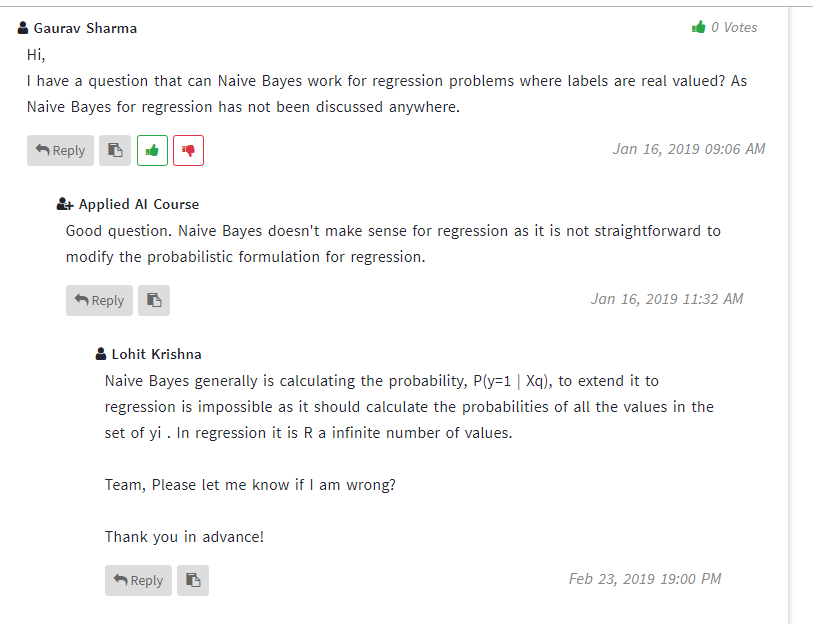
So always apply Laplace Smoothing while using NB>

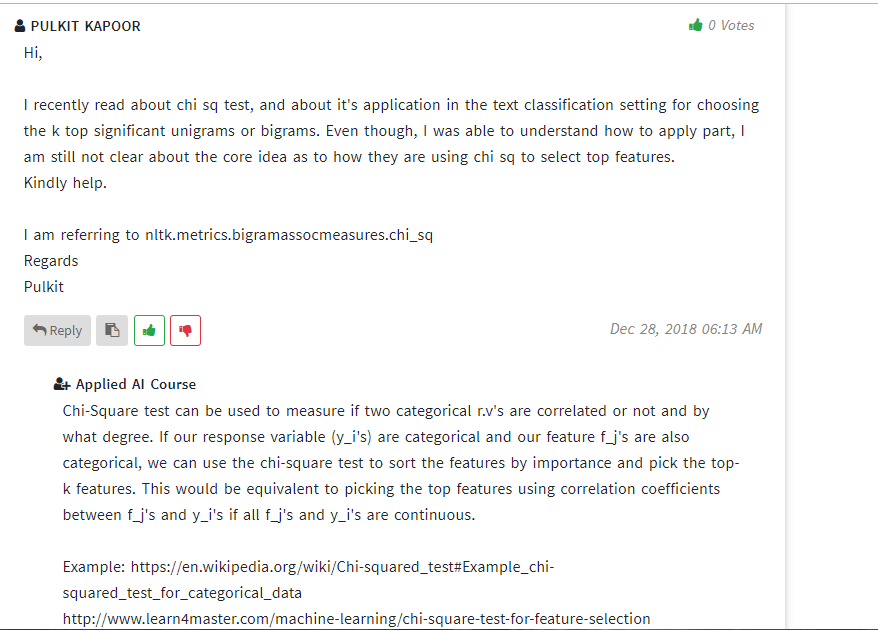


Comments:









Example: https://en.wikipedia.org/wiki/Chi-squared\_test#Example\_chi-squared\_test\_for\_categorical\_data  
http://www.learn4master.com/machine-learning/chi-square-test-for-feature-selection  
https://stats.stackexchange.com/questions/24179/how-exactly-does-chi-square-feature-selection-work