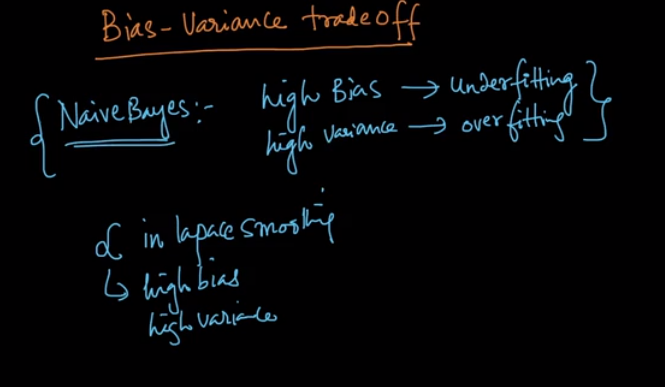
**BIAS AND VARIANCE TRADOFF:**

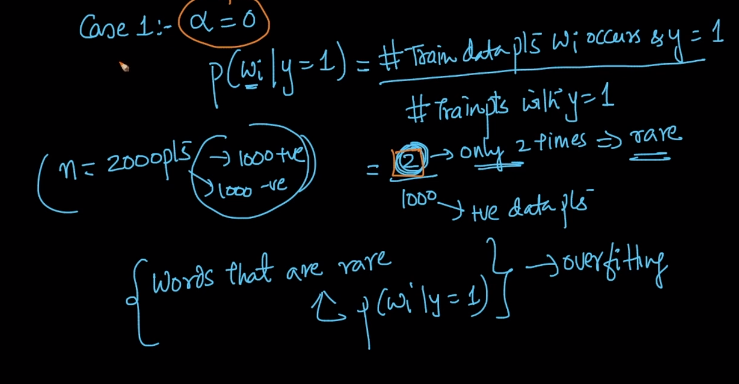
Since we already know that High bias means underfitting and High Variance means Overfitting.

In Naïve Bayes there is only one parameter that is (ALPHA) in Laplace Smoothing which results in overfitting or underfitting.



Lets see how change in value affects variance and bias

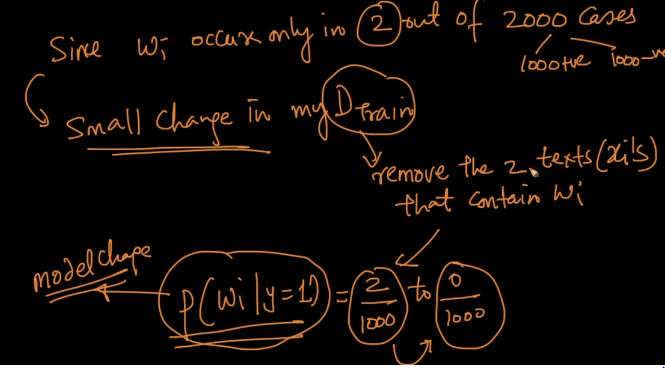
Case 1: Lets take Alpha = 0



As we can see that we are even giving the probabilities of words which are occurring very rarely means it is an overfitting problem and as we know overfitting problem have high variance

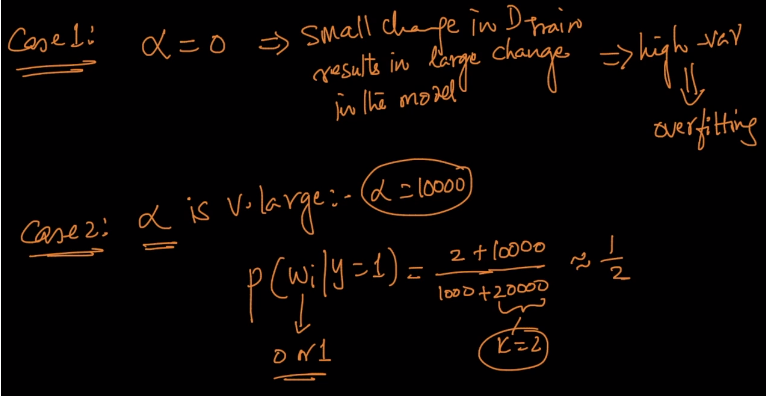
That means small change in my training data will affect a lot in my model

And so lets suppose if we remove those two sentence which contains the word the probability will fall to 0 which is really drastically change the model as it nothing but product of all the probabilities.

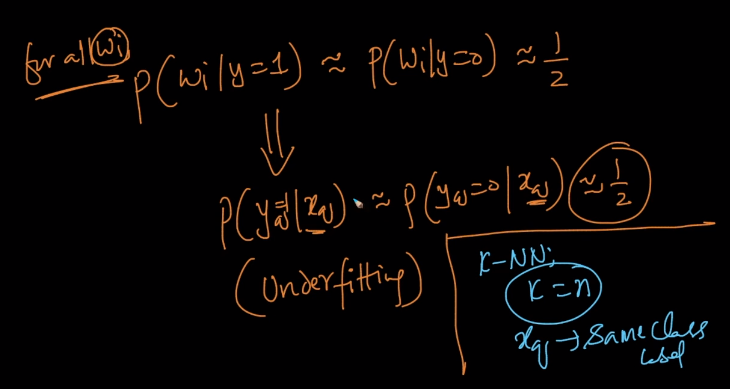


Now lets take case 2:

When value of Alpha is very large.



So when the value of Alpha becomes very large than it says the probability for every class will nearly be same and the model will not be able to tell which class our query point belongs to

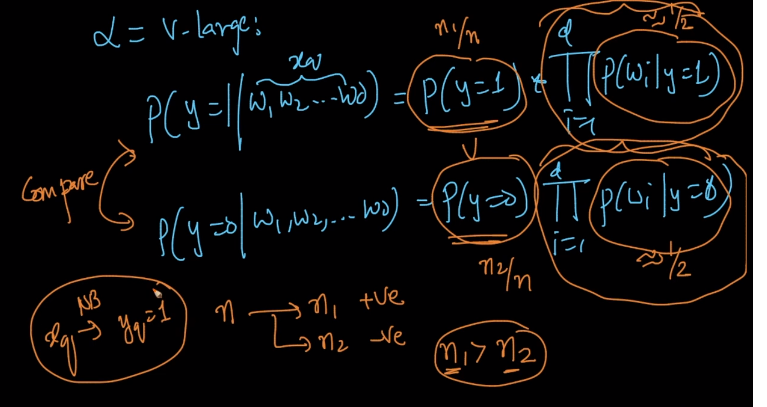


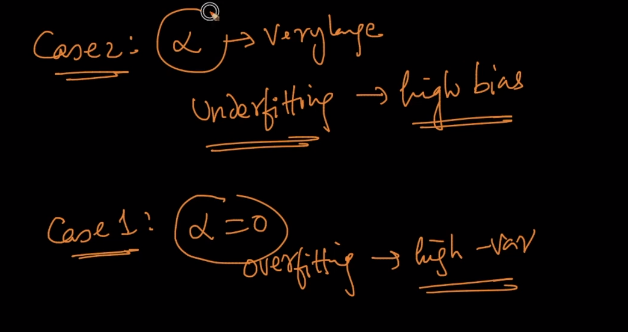
And this is the problem of underfitting which is same as KNN where K= n gives every query point a same class (majority class) and same will happen here too.

Let’s see how,

Since now all the likelihood probability have become equivalent all lies down to P(Y=0) and P(Y=1)

So now same as in KNN whichever label is majority class ,every query point will be given that label.



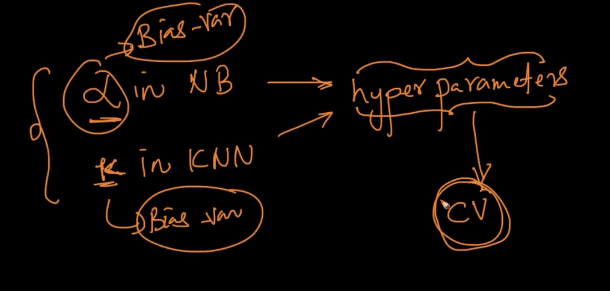


So the bias variance trade off all depends on value of ALPHA.

Now the question is how we calculate suitable Alpha ?S

It will be calculated in same manner as we calculated K in KNN .

Since both are hyperparameters which means it controls Bias Variance trade-off so the values will always be calculated by cross validation.

S