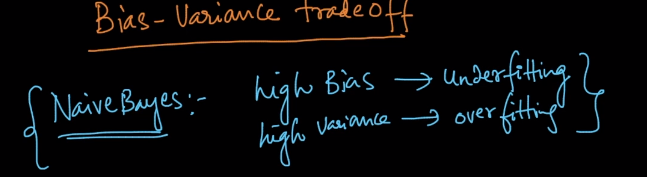
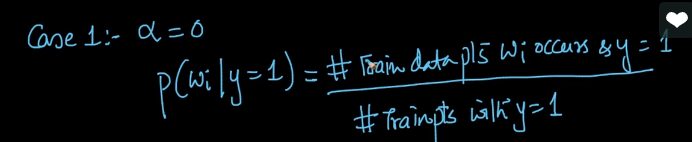
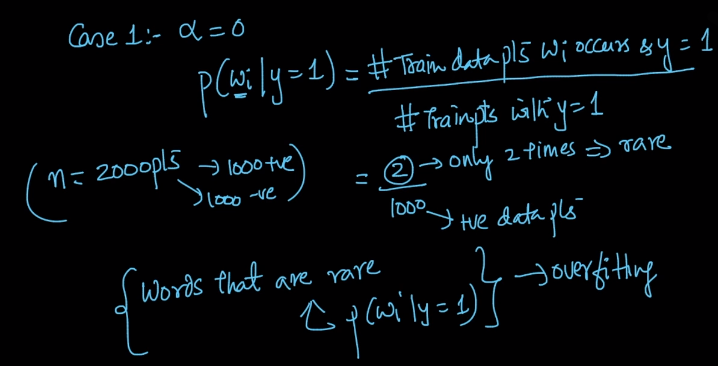
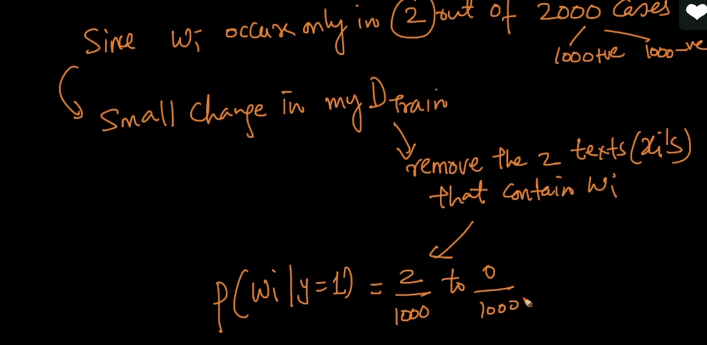
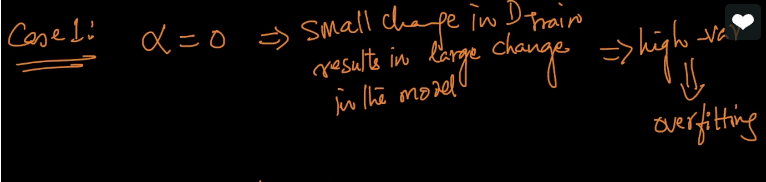
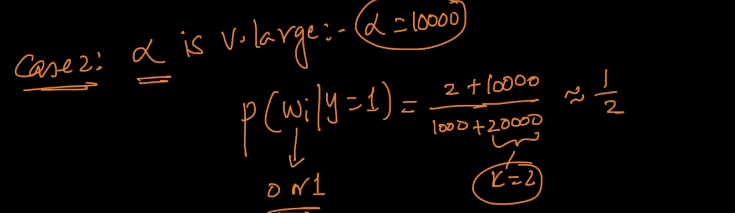
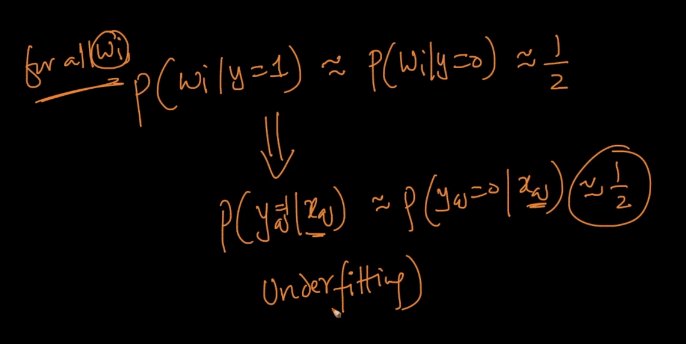
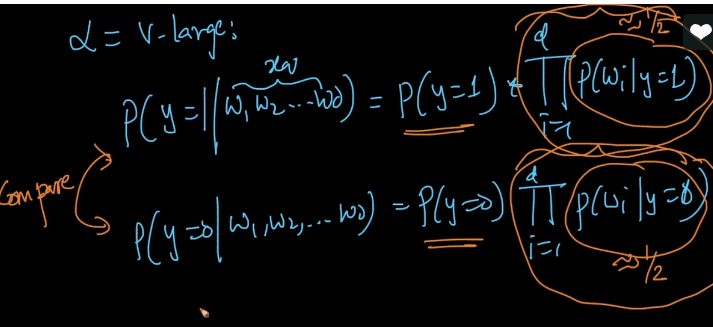
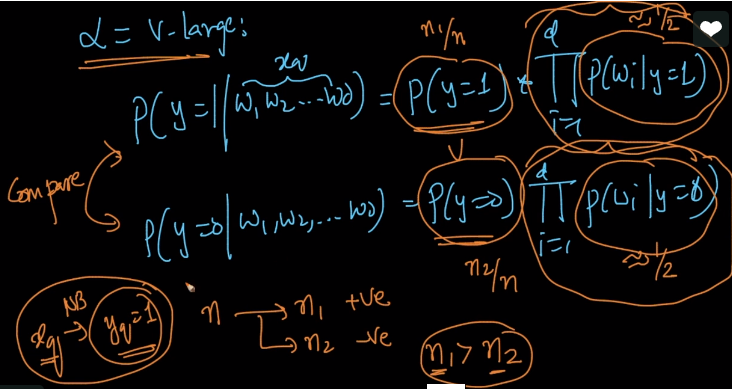
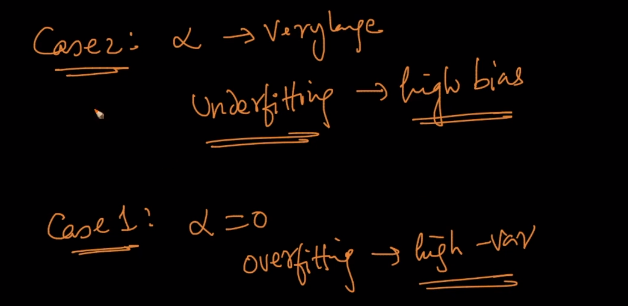
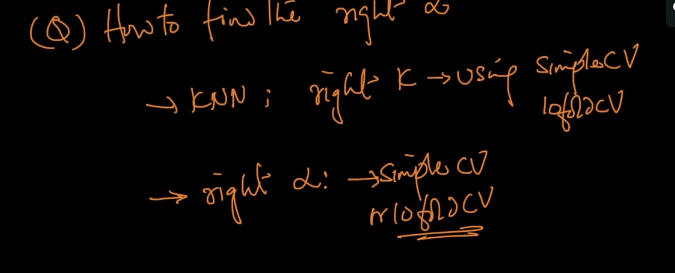
**Bias and Variance tradeoff**🡪high bias = under fitting  
🡪high variance = over fitting  
  
In Naïve Bayes, alpha value in Laplace smoothing determines over fit or underfit. Let’ see this below  
**Case#1: when alpha = 0(No Laplace smoothing applied )**  
If we consider the likelihood   


Let’s say a word occur only 2 times, which is very rare. Even for words that are rare, we are giving probability which mean over fitting meaning **high variance**.   
  
High variance meaning, even if there is small change in training data, result in model changes dramatically.  
Let’s see an example: a word occurs only 2 times out of 2000 cases(+1000,-1000) , and I do small change in training data(remove 2 words which are rare),then there is real change in model output. Probability will completely become 0 and model changes completely   
  
From this we can conclude small change in Dtrain results in large change in model which mean high variance, which is over fitting . find below pic

  
  
**Case#2: when alpha is very large:**Here, we consider alpha as a large value, let’s say 1000   
  
then probability will be half for y1 and y0 points meaning model cannot say to which class label it belongs to, this is called underfitting.  
  
  
See highlighted part in below pic, which is likelihood becomes half when class label is 1 or 0, Meaning likelihood is equal for both the class label.  
  
Since the likelihood is same for both the classes,   
now It drills down to “class prior” part P(y=1) = n1/n and p(y=2) = n2/n. now if we calculate the probability, then which has a greater number of points, will win. If n1 has more, n1 will win and vice versa   
   
  
Now immdeiate question is how to find the right alpha?  
In KNN: we found K using simple cross validation or ten cross validation  
In NB: we need to follow same cross validation or ten cross validation  
  
  
Alpha and k are called as hyperparameters  
