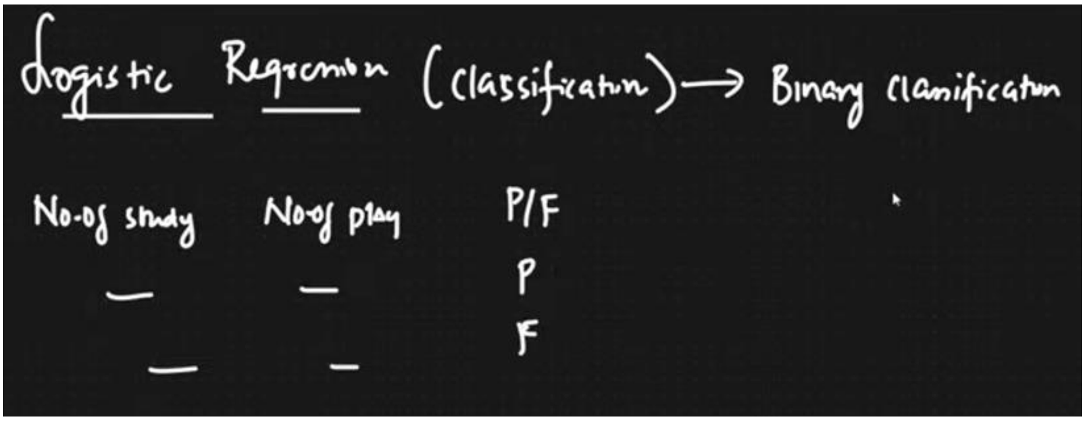
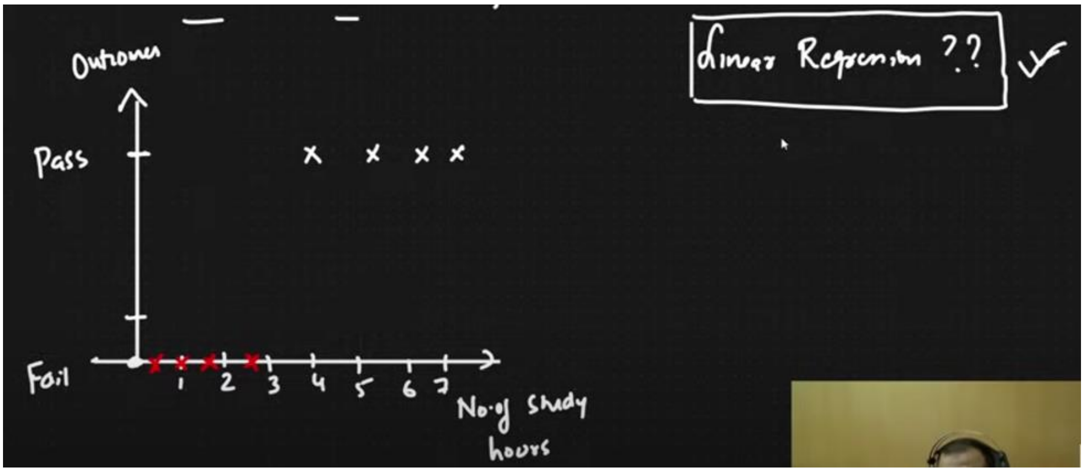
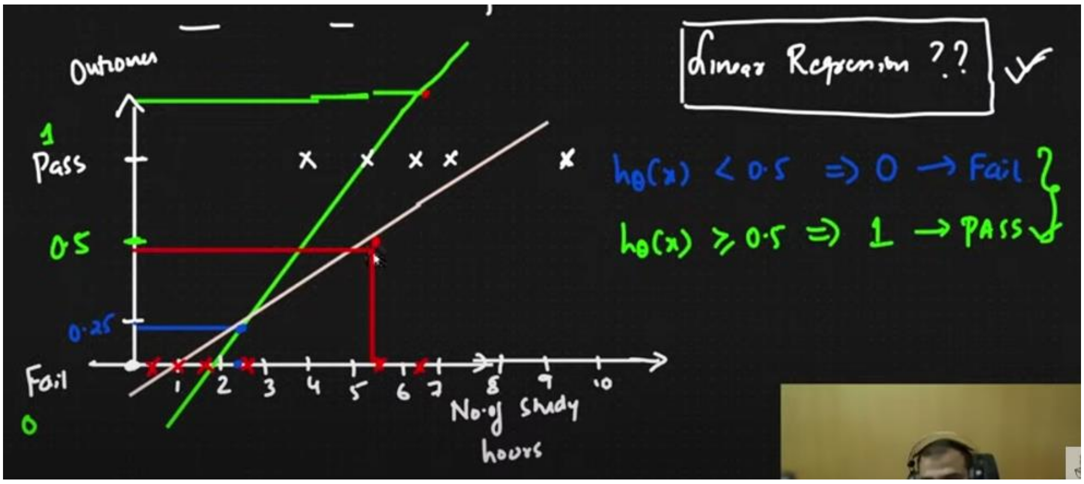


1) Unlike, linear regression, here our dependent variable (target variable) is classified in no. of categories! In the above example, it's binary: Pass & Fail!





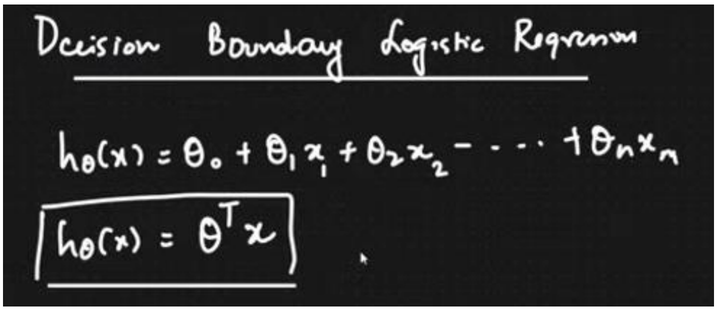
Over here, we can observe that, linear regression can't be worked on the logistic regression because:

When there's a new output and we draw a new line, the intersection meets at a particular pts which gives us the output FALSE but in reality, it's TRUE!

Also, it can be greater than 1 and lesser than 0 (-ve), which cause us the problems, should be 0 to 1.

Hence, linear regression won't work! Then? Will squash (straight line) called as **Sigmoid function**!

2) Decision boundary in the case of logistic regression:

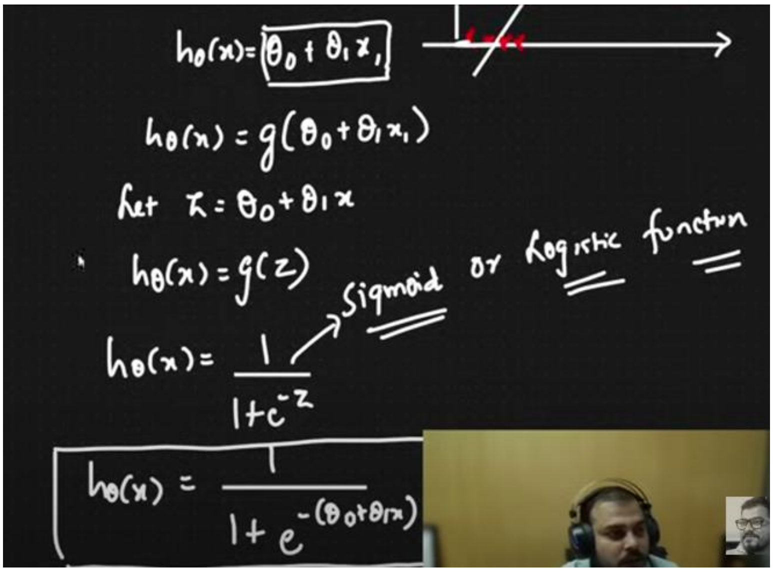


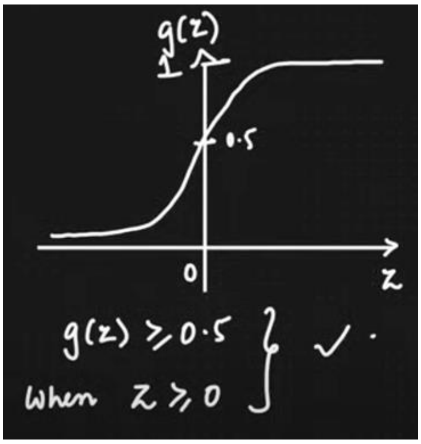
1) create best fit line,

2) squash (straight line on both the side 0 and 1)

3) sigmoid activation function

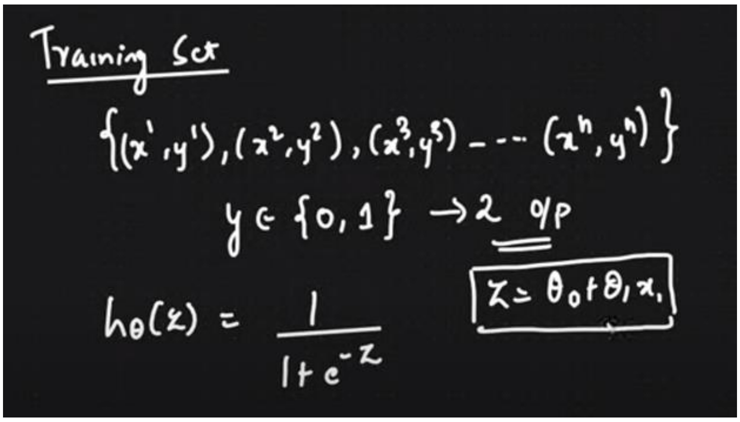
Math:





3) Why it's called logistic regression?

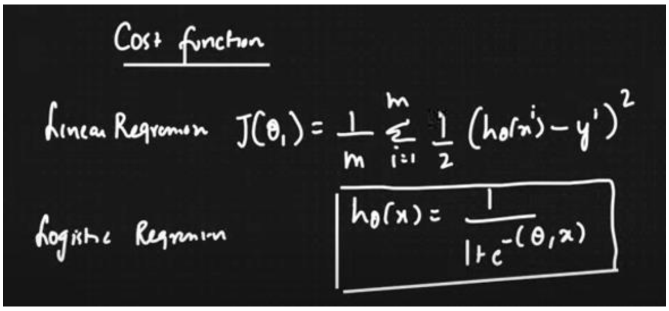
Because using regression we form a best line and due to sigmoid function, we can solve the error, making the line squash; combined it's called logistic regression!



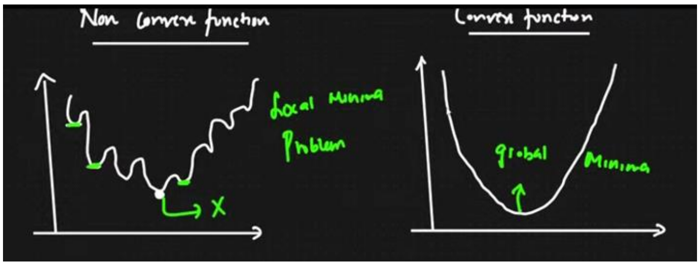
4) When, Θ0 = 0;

Change parameter Θ1 in such a way that, we get the best fit line and apply sigmoid activation function.

Will find the cost function:

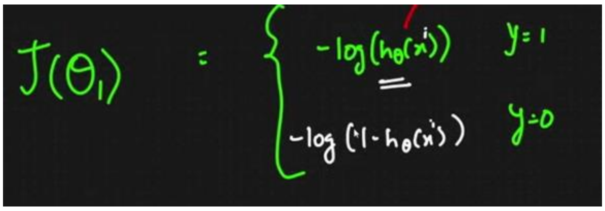


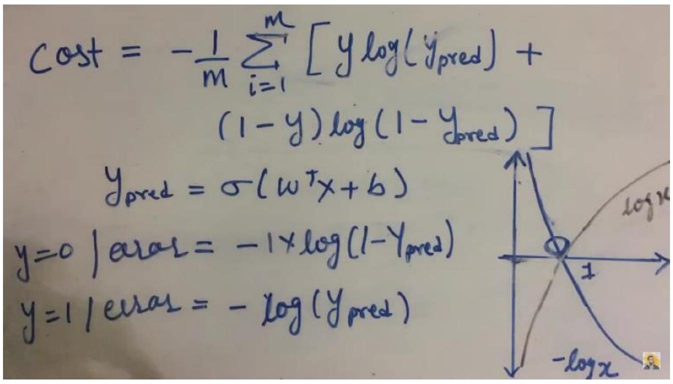
replace, hΘ(x) in the above (linear regression) equation **but wait,** we cannot use this equation as hΘ(x) is changing (can take continuous value (any number) but not binary, i.e: y = 0 / 1) and so it's non-convex function equation due to which it has lots of local minima. Reaching global minima will be difficult!



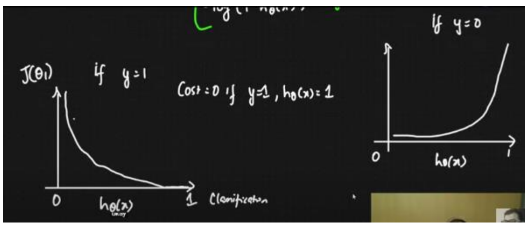
Reference!

Then? 5) Will do the following:



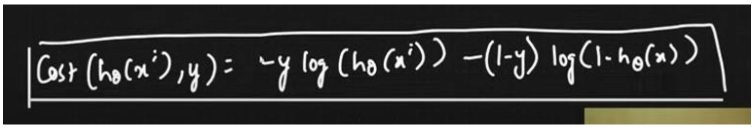


a)



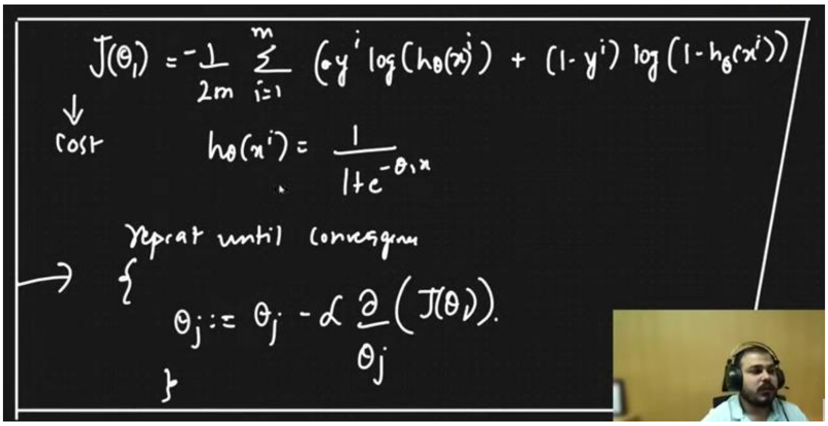
b)

c) final cost function:

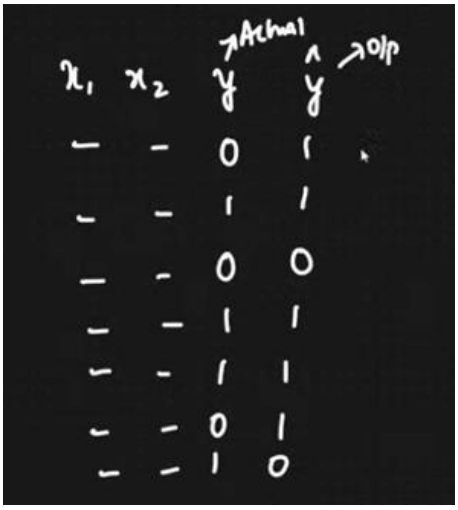


cost function & loss function with no. of parameters will always be the same!

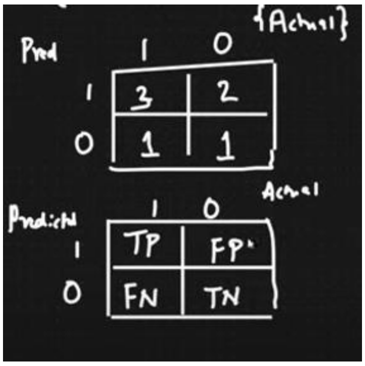
d) solves the errors w.r.t logistic regression: (log likelihood)



6) Performance Matrix: (Classification problems)

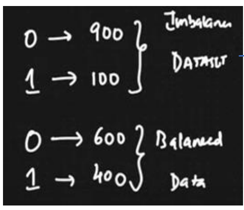


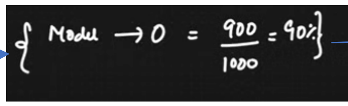
To find the accuracy and other terminology of this particular model will perform: Confusion matrix!



a) Accuracy: $\frac{TP+TN}{Total} = \frac{4}{7} = 0.57 = 57\%$

Try to reduce FP & FN!

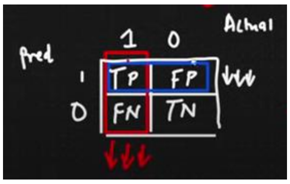




Biased Data

Need to perform other terminologies too!

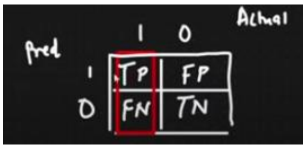
b) Precision: $\frac{TP}{TP+FP}$ (blue)



Example: Spam Classification!

GOAL: Identify which mails aren't spam & are important. Focus on FP; actual spam, predicted imp.!

c) Recall: $\frac{TP}{TP+FN}$ (red)



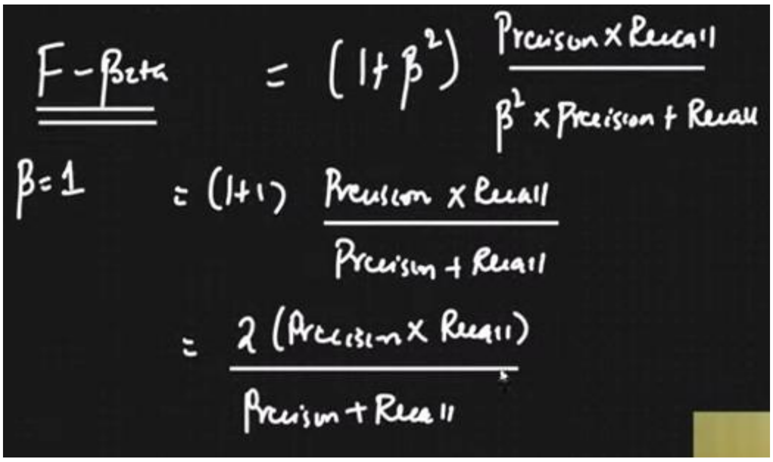
Example: Has cancer or not!

A person has detected that he/she has cancer but when we test them again, we found they don't have the cancer! Luckily, due to the repeat test, we stated they don't have cancer, just think what if on the basis of actual result, they have been started with the cancer treatment? Hell no!! They might instead get the cancer if they don't have it! So, focus more on FN, actual true, predicted no!

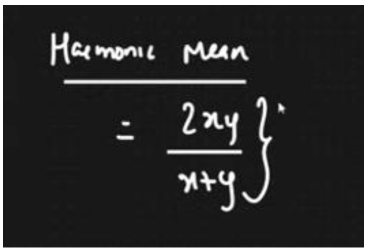
d) F1 -Score / F1 - Beta:

Example: Tomorrow stock market is going to crash

Here, both are important precision and also recall!



F1 – Score is



like:

when FP is important, will decrease β to 0.5, can say it's F0.5 score;

when FN is important, will increase β to 2, can say it's F2 score!

ON THE GIVEN PROBLEM STATEMENT, WILL GET TO KNOW WHICH TERMINOLOGY IS IMPORTANT

Reference:

1) Logistic Regression from the scratch (47:30 mins onwards)

2) Logistic Regression Cost Function