**Polynomial Kernel**

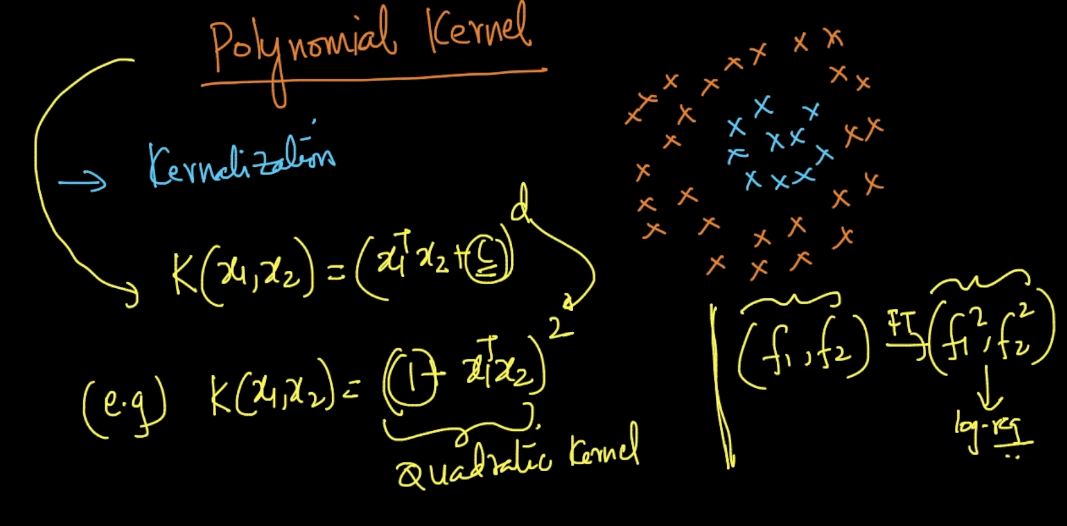
As shown in below dataset we can linearly separate it by applying feature transformation in logistic regression.

In general polynomial kernel can be represented as

K(x1, x2) = (x1Tx2 + c)d

Here c and d is constant.

For eg. To linearly separate dataset we use quadratic kernel(because d = 2)



In below image we first put value of X1 and X2 in quadratic kernel

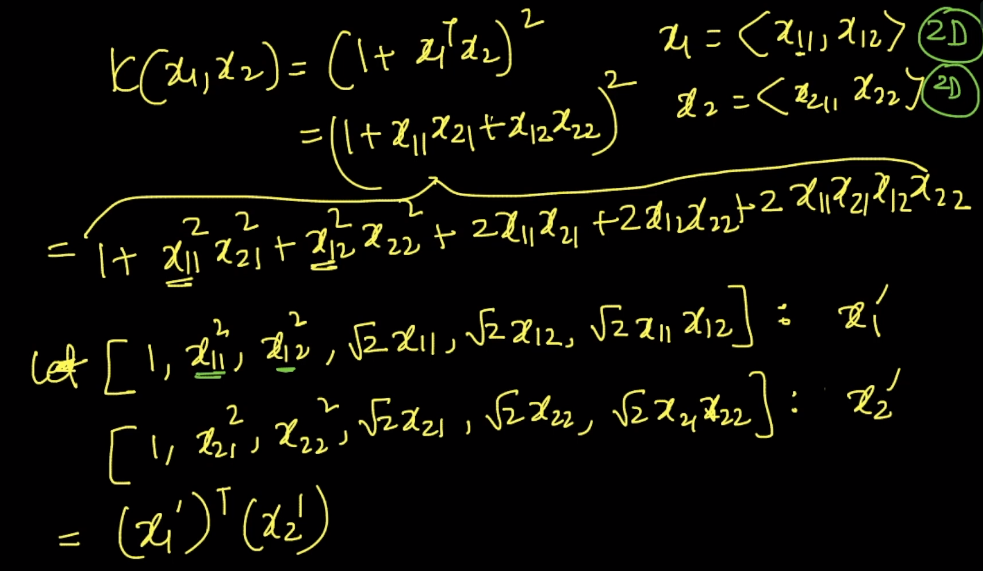
Then we apply (a+b+c)2­and we get value of K(X1,X2)

Now after the value we get linear seperates our data.

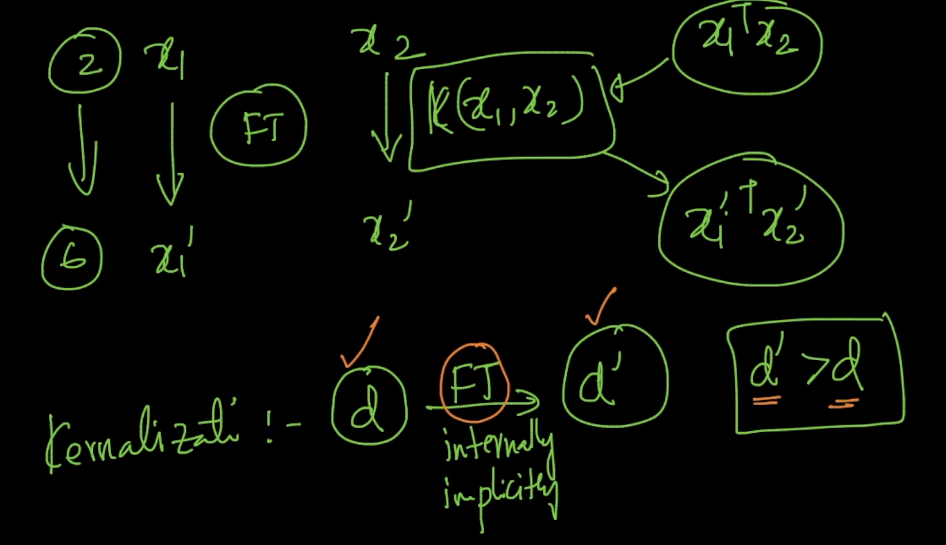
And the value we get by above of K(X1,X2) is also get by applying dot product of X1’ and X2’

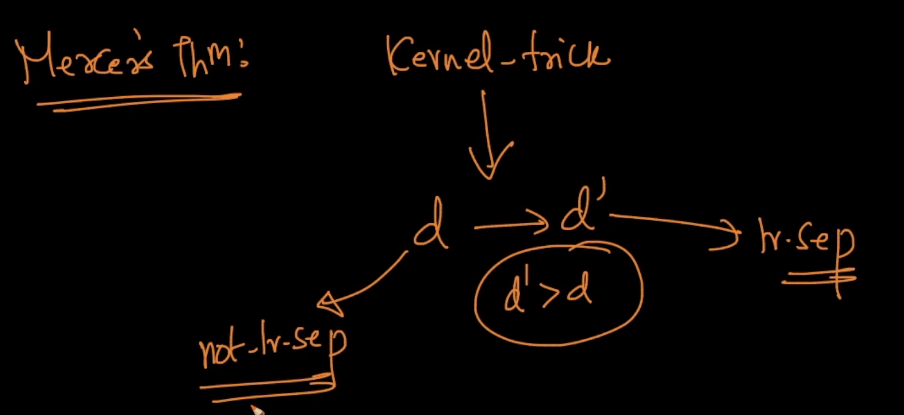
And we let the value of X1’ and X2’ such that it dot product give value of K(X1,X2).

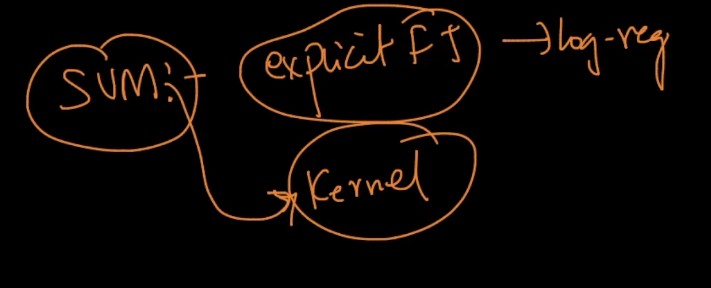
And see we get value of X1’ and X2’ is just a feature transformation of X1 and X2 in higher dimension here X1’ and X2’ is in 6-dimension.



So what kernelization do is it takes linearly inseparable data from lower dimension and then implicitly transform that data in higher dimension to linearly separate it.







One question is why we use quadratic kernel of 6-d to linearly separate below data as we linearly separate it using f12 and f22 because quadratic kernel not only linearly separates this circle data but it also linearly seperates ellipsesin short all conic-sections.

