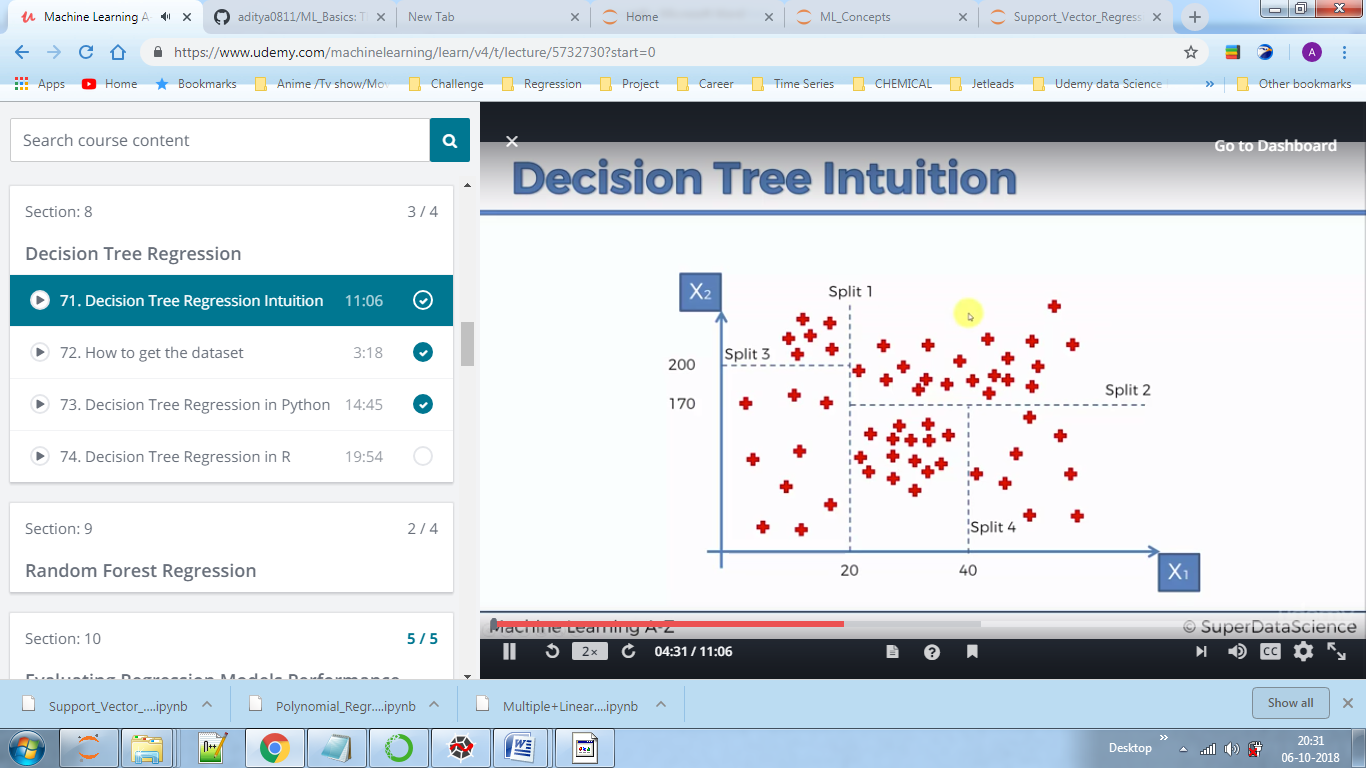


We predict y



The algorithm splits data i.e groups data such that we have more information about that group (information entropy some complex maths concept)

Now suppose we have new point (x1=30 , x2=50 which falls between split 1,2,4) to predict after splitting data somehow by an algorithm.

Now we predict average y value of all the points in that region and assign it to the new point.

Here remember the new point can only have one of the 4 values.