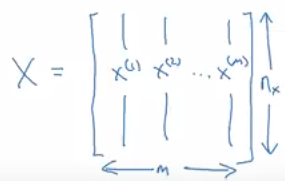
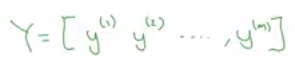
**Lec 5: Binary Classification in Deep Learning**

We are looking at logistic regression.

For a 64 x 64 image with three channels, it must be converted to a feature vector (np.flatten) of size 64 x 64 x 3.

*Notations used:*

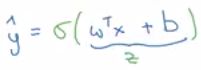
* (x, y) is the training example along with its label.
* Consider ‘m’ training examples: 
* For neural networks it is always advised to stack the ‘m’ input vectors and its labels separately:  

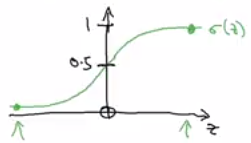
**Lec 6: Logistic Regression**

Given a training example X, we want to predict y’ (y-hat):  where x is .

Its parameters weights and bias are: 

The output cannot be a linear equation because what we want is a probability output, hence we fit a sigmoid function.



Sigmoid function:  

Note:

* If value is large then output is close to 1
* If value is small then output is close to 0

**Lec 7: Logistic Regression Cost Function**

Cost function is used to train the parameters ‘W’ and ‘b’.

The loss function determines how good the predicted output (y’) is against actual output (y). Squared error is an absolute choice but in logistic regression the gradient descent would have many minima (not convex), hence it is not used. 

The loss function used is: 



Loss function is for a single training example.

The cost function is used for the entire training set: 

