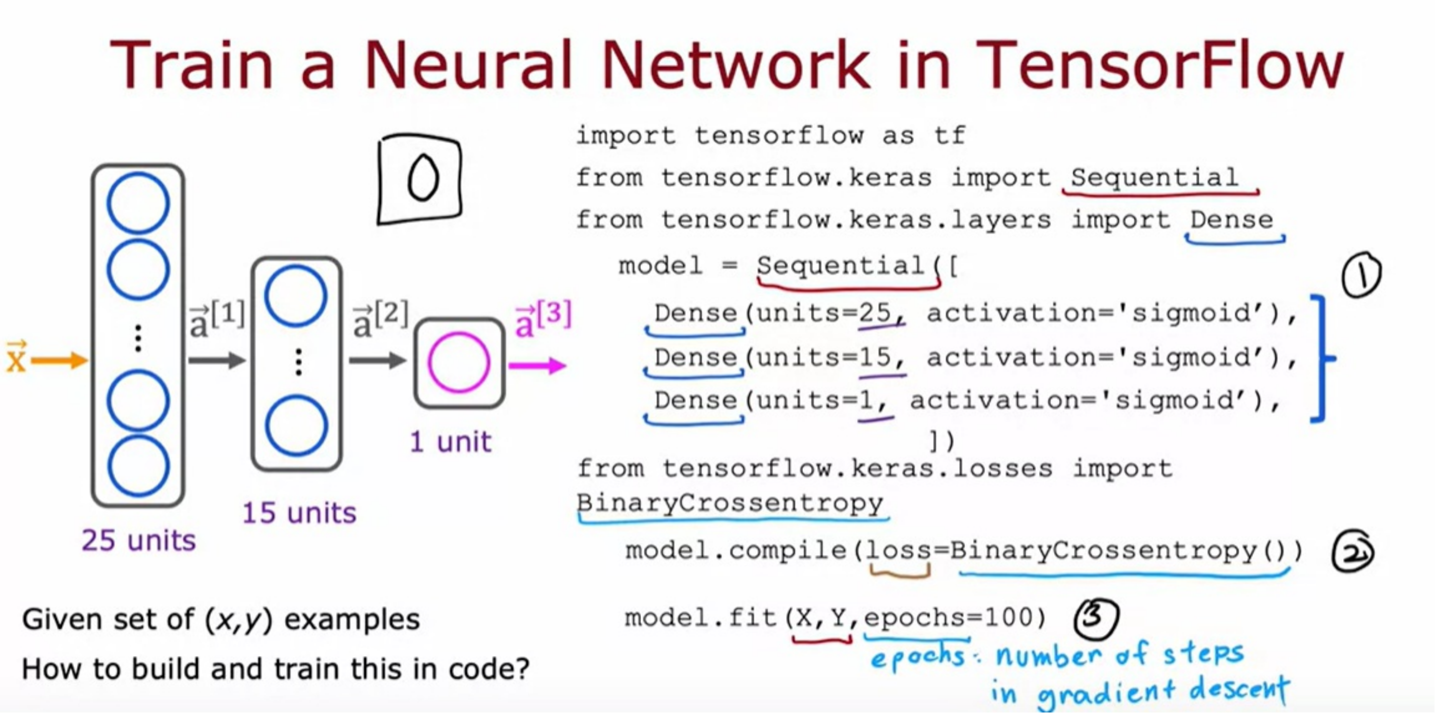
Week 2

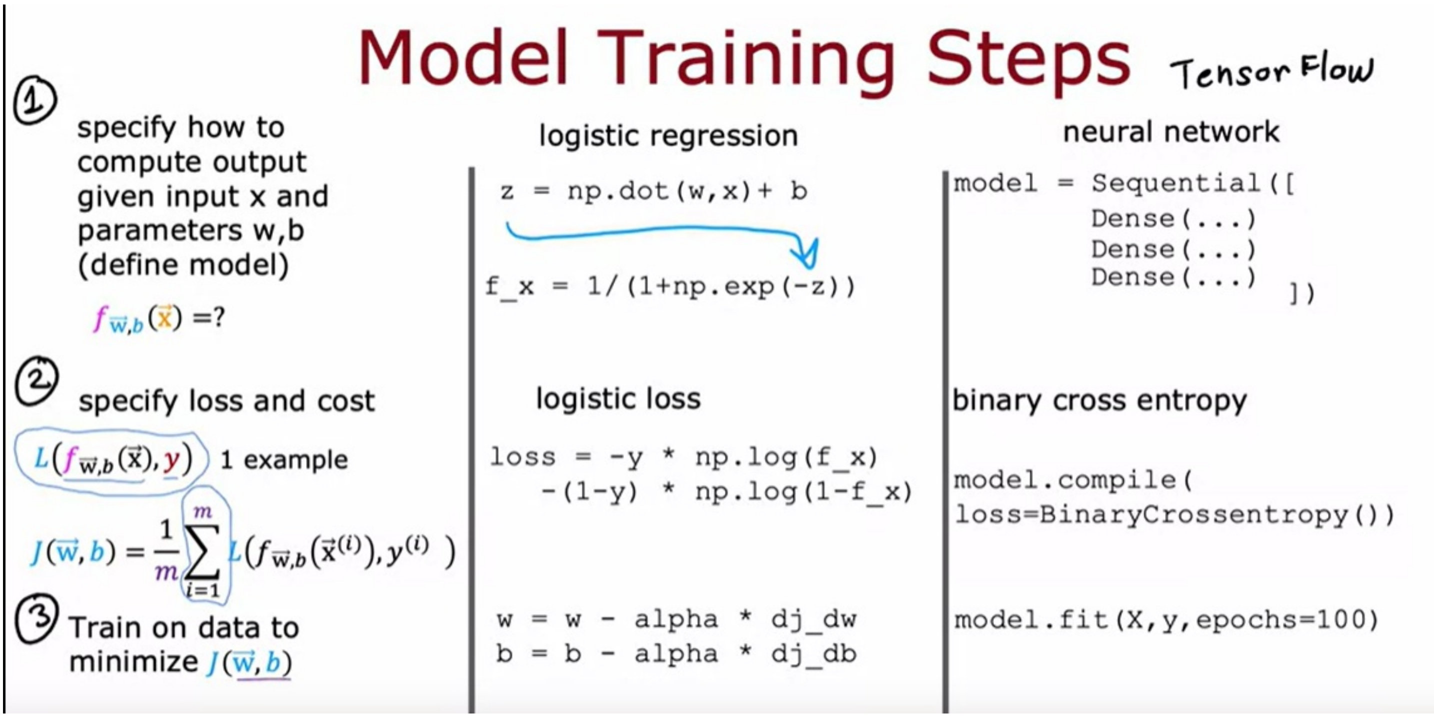
Neural Network Training

- Mode.compile → states the loss function

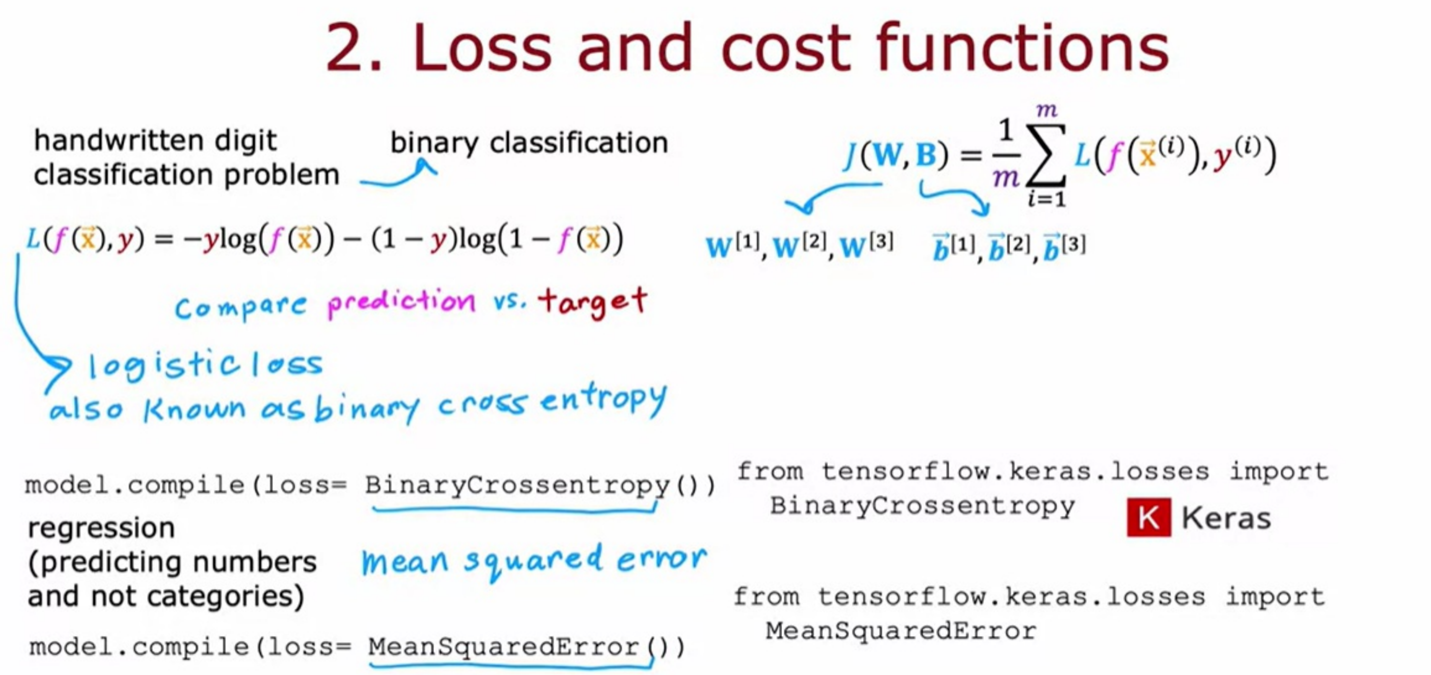


Model Training Steps

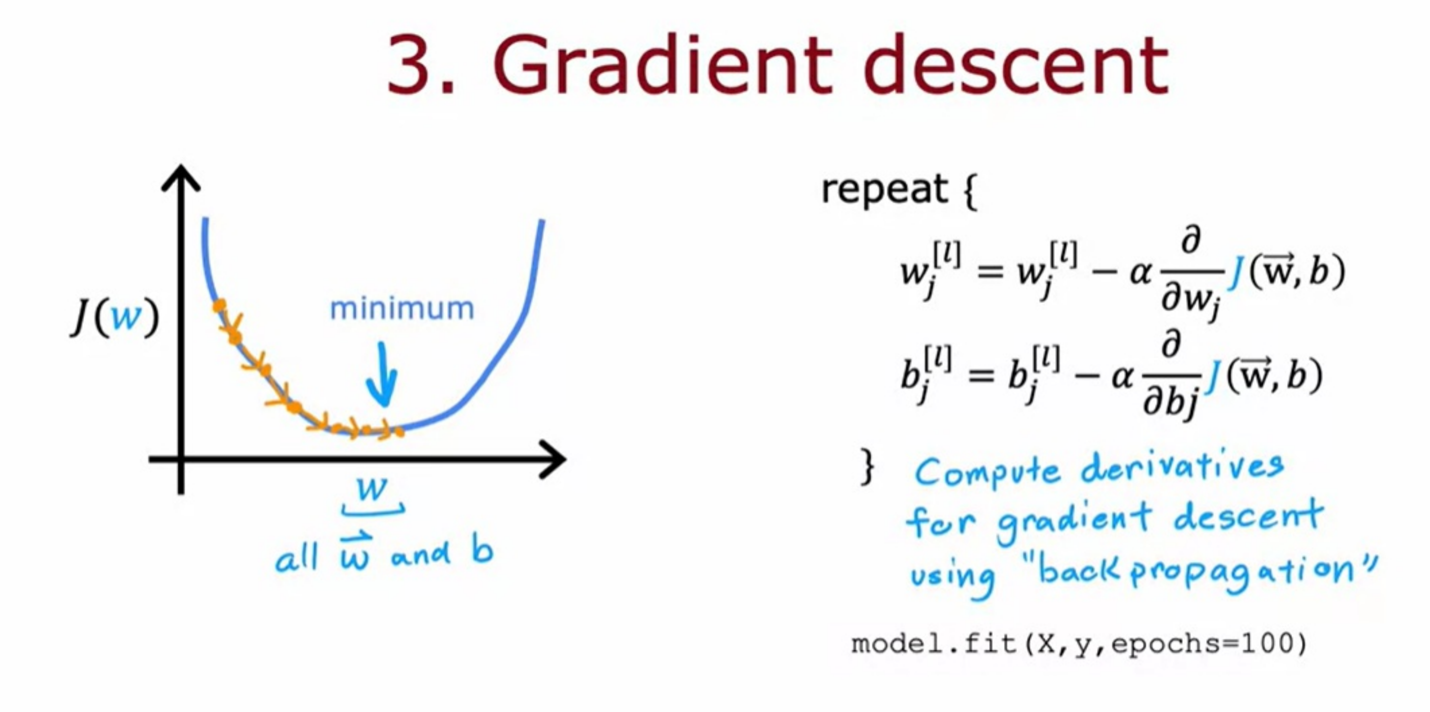
- Recap



- Step 1 → Create the model (The first slide)
- Step 2 → specify loss



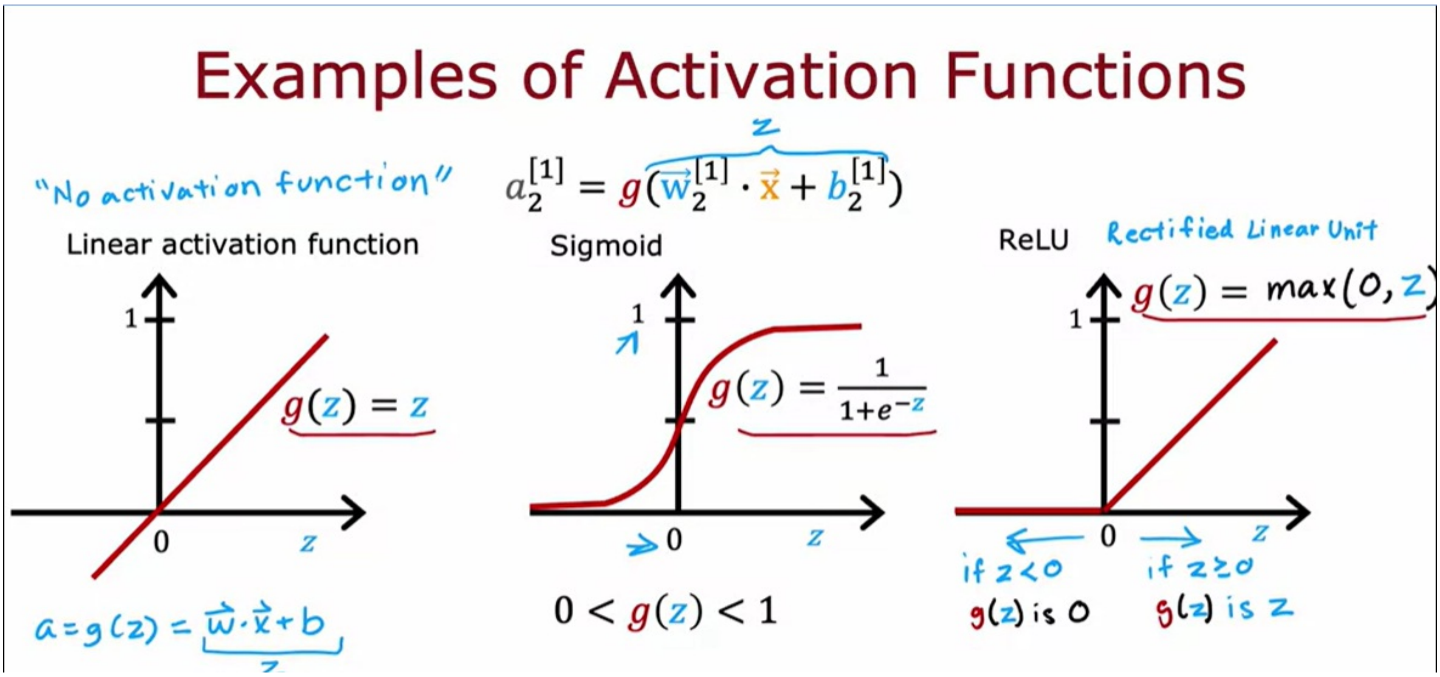
- Step 3 → Gradient descent (Minimize cost)



Activation Functions

Examples of Activation Functions

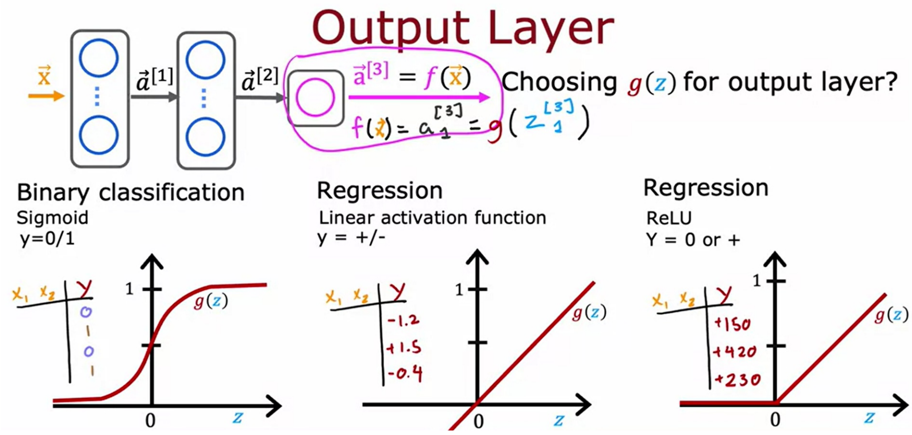
- Linear Activation function
- Sigmoid function
- Rectified Linear Unit (ReLU)
- Later: SoftMax



How to Choose Activation Function?

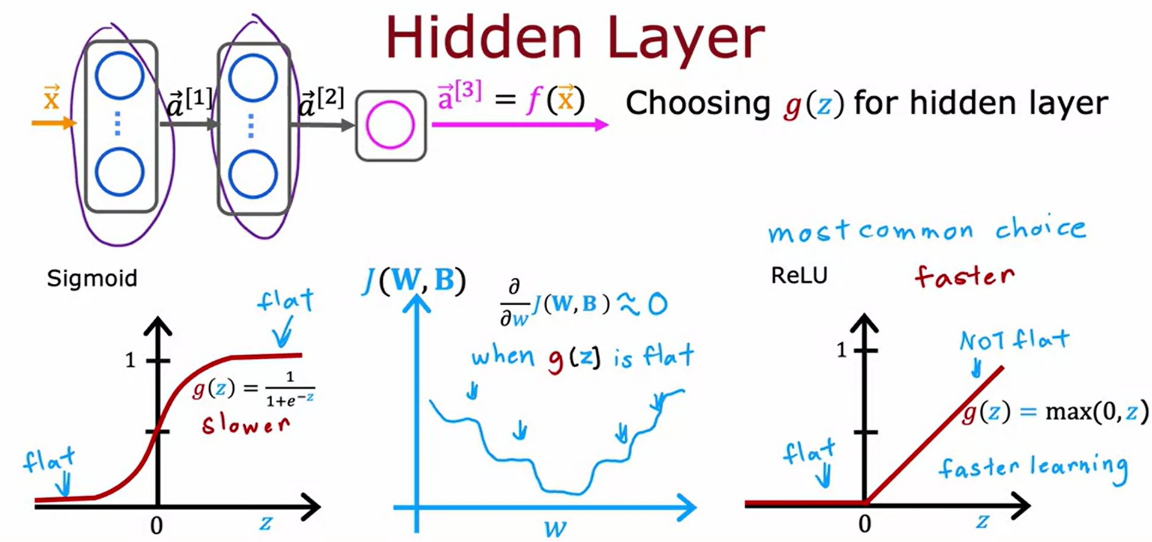
Output layer

- It depends on the type of output you want



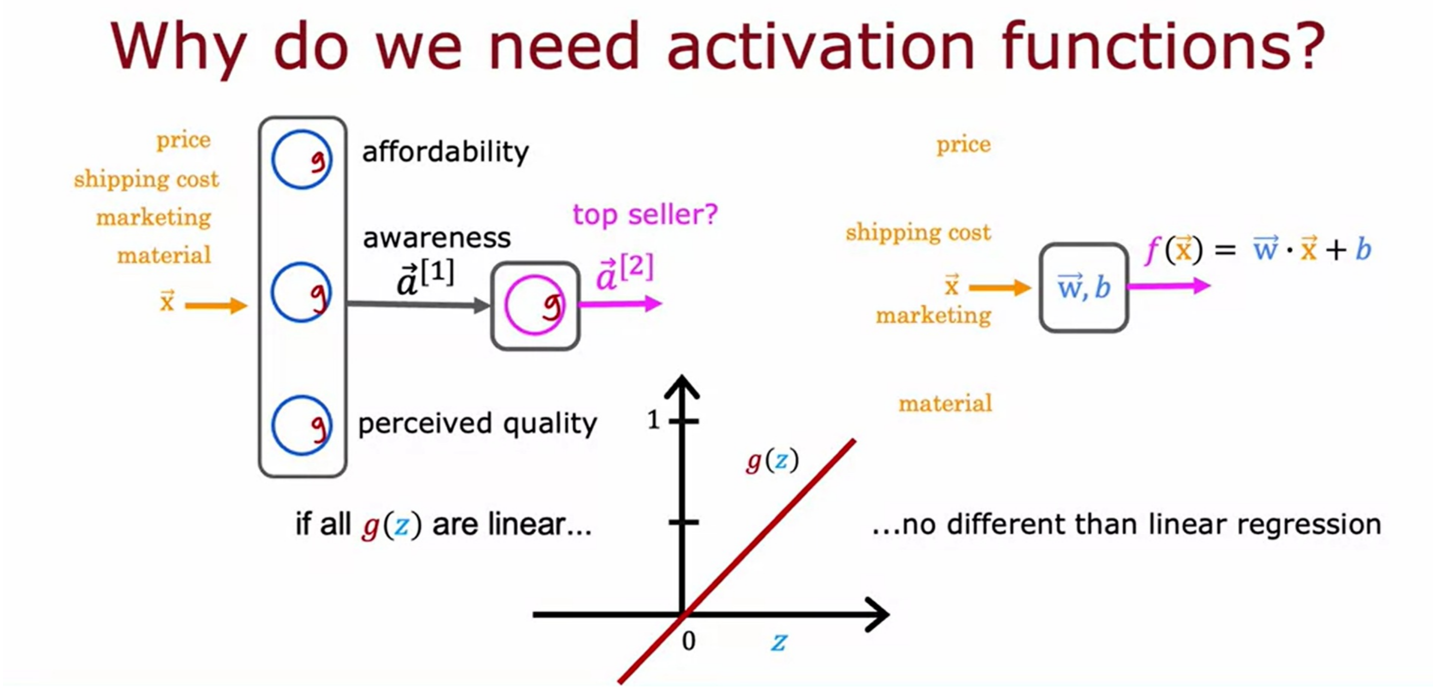
Hidden Layer

- When the output is flat, it will be slower to reach the minimum point



Why do we need Activation Functions?

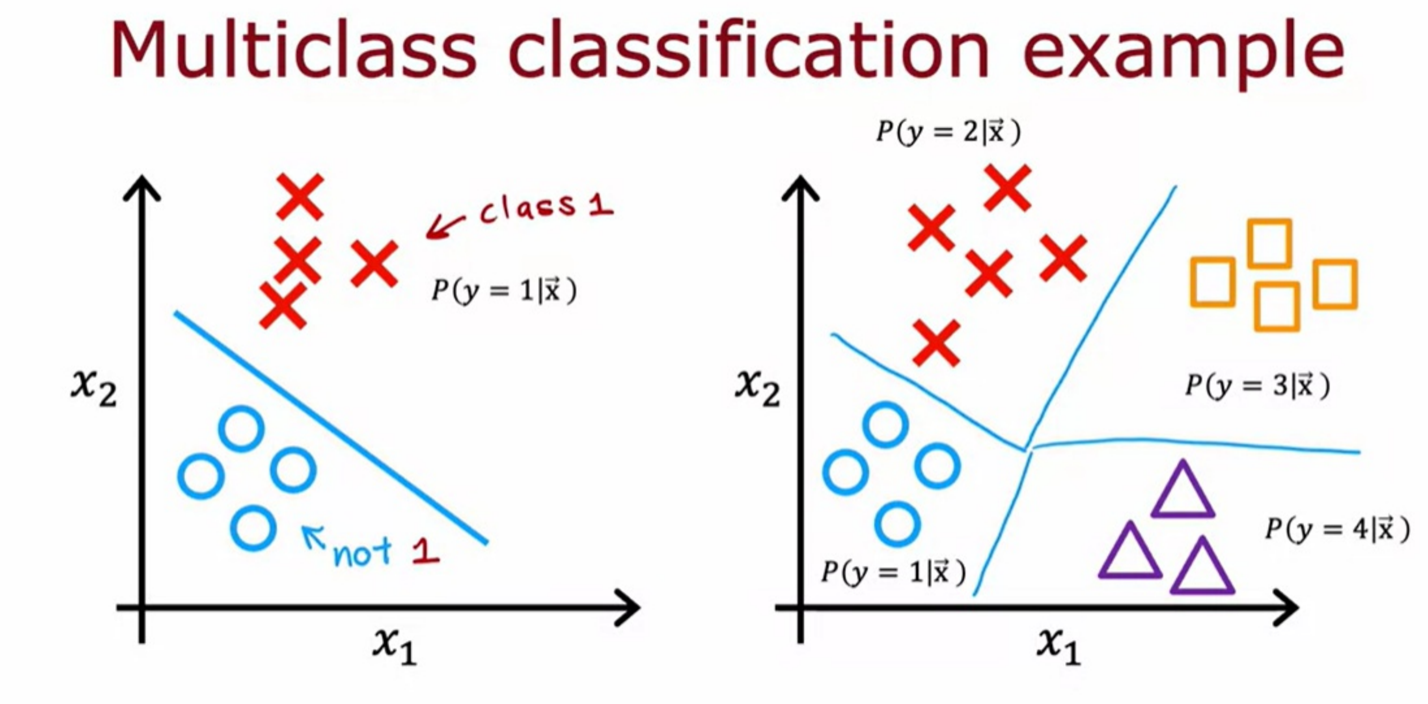
- Why do they not work otherwise?
- In this case, having multiple layers will not make calculations any different



Multiclass Classification

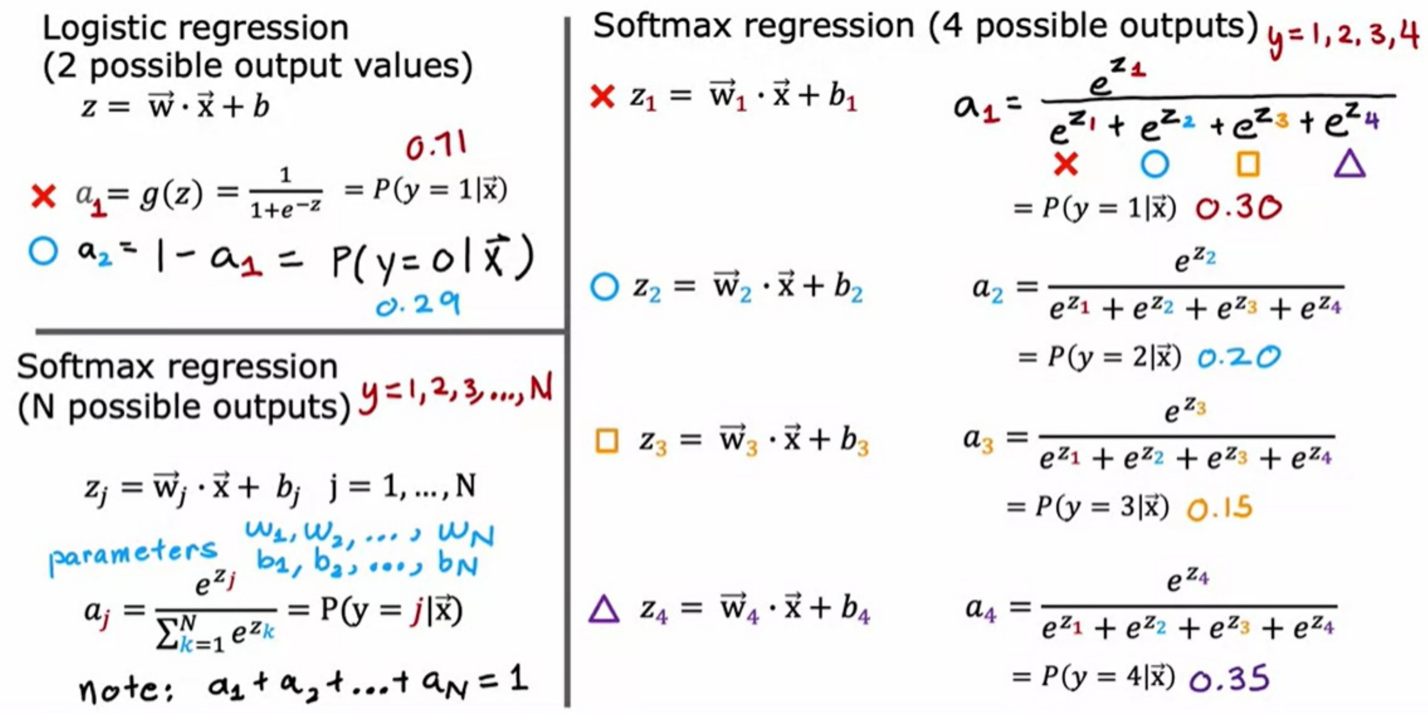
Examples

- What's the chance that Y is = 1 OR y =2 OR y =3 OR y = 4?



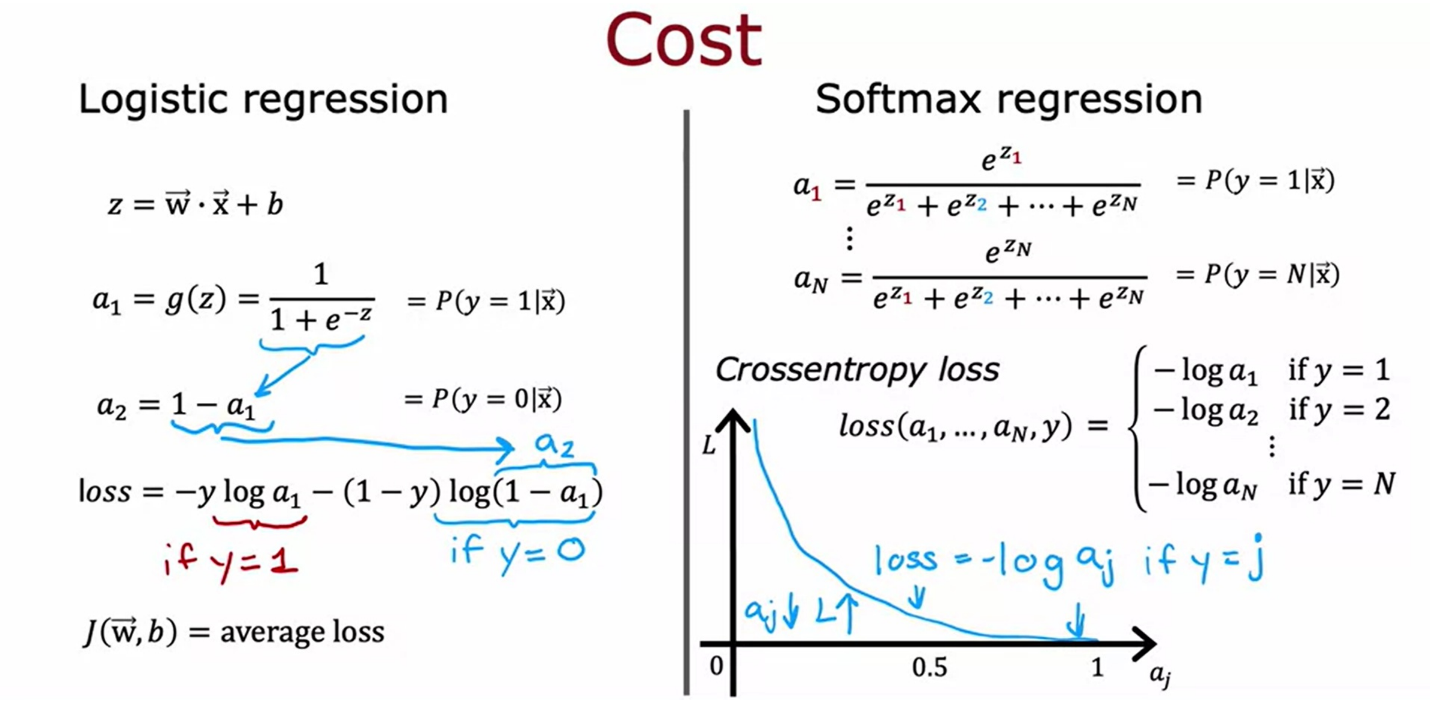
Logistic vs SoftMax regression

- SoftMax model is a generalization of the logistic regression model!



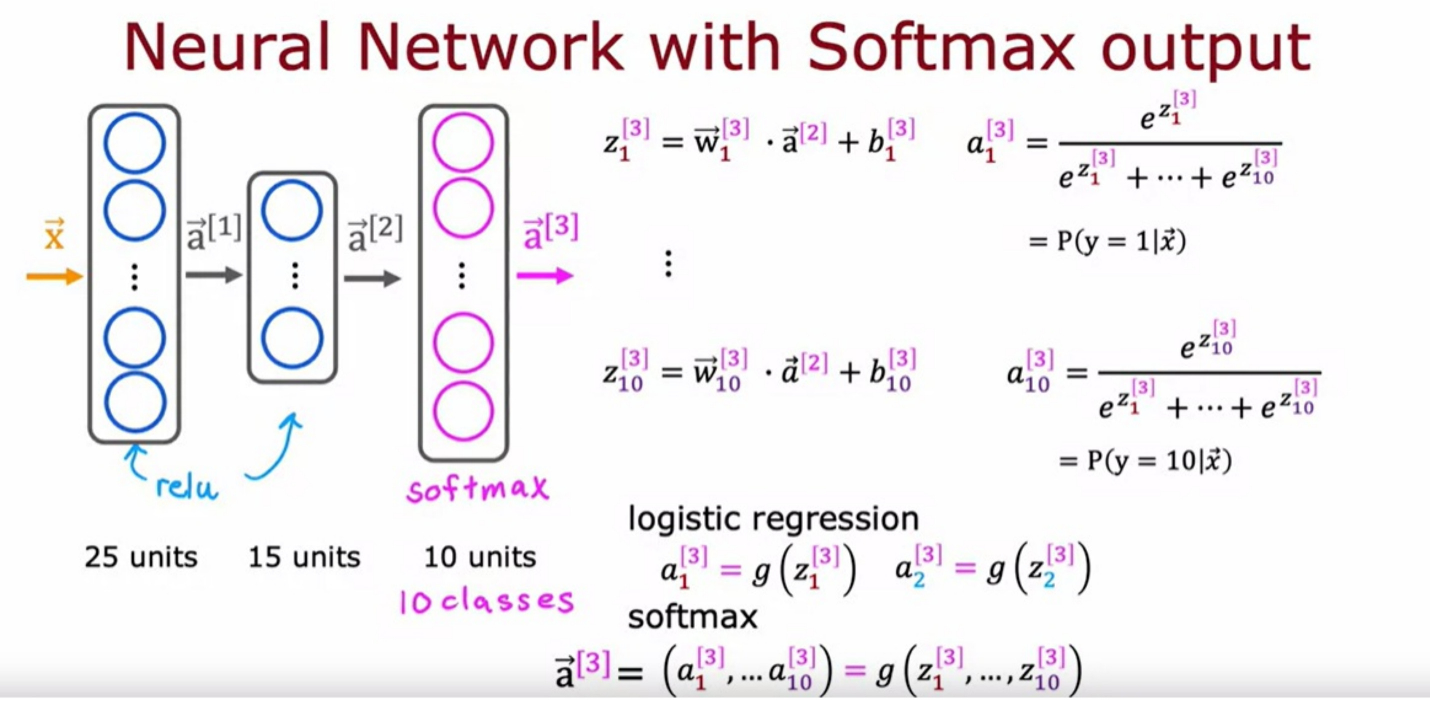
Cost function for SoftMax Regression

- The smaller the number of aja_jaj (type of class), the greater the loss impact



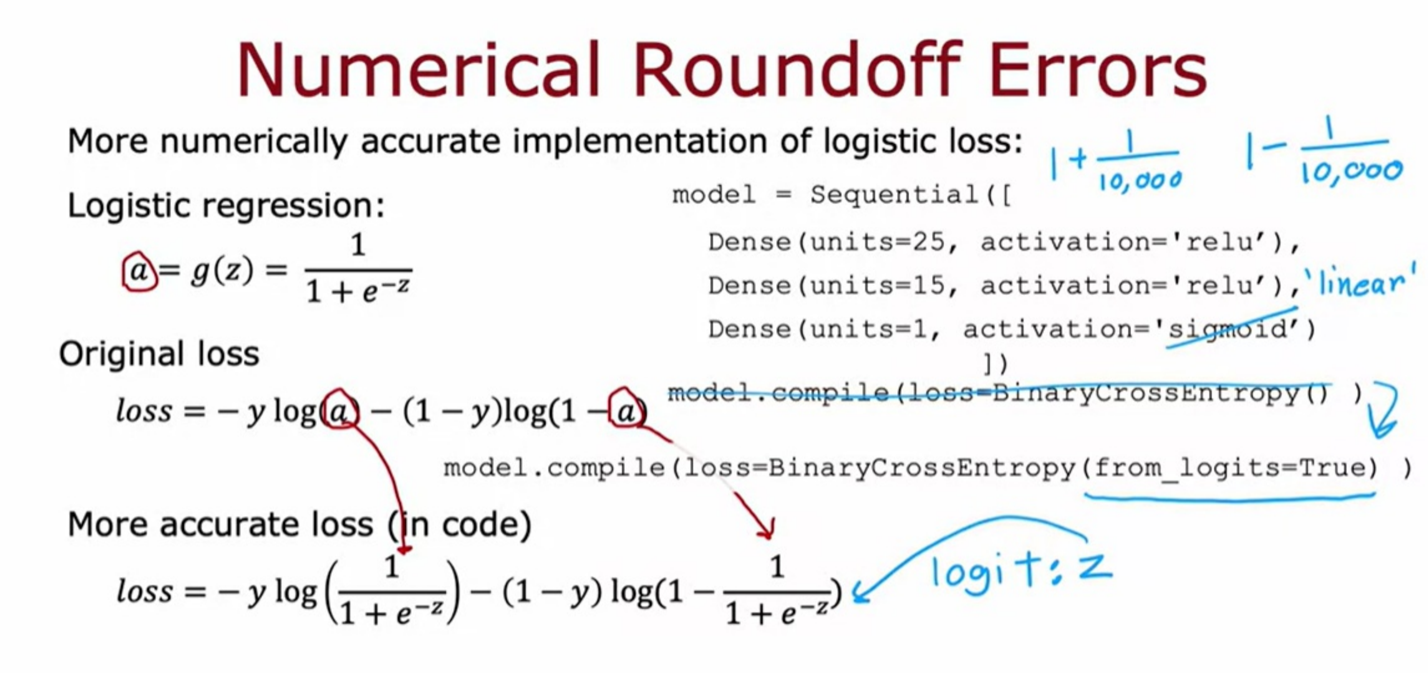
Neural Network with SoftMax Output Layer

- Output layers consists of several neurons

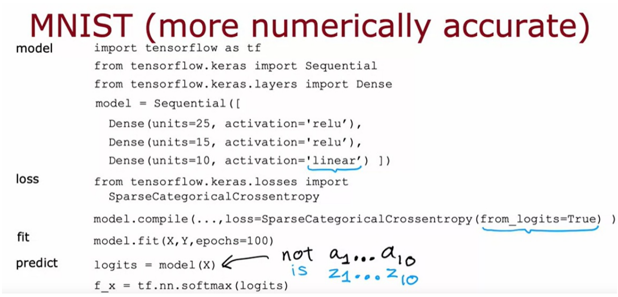


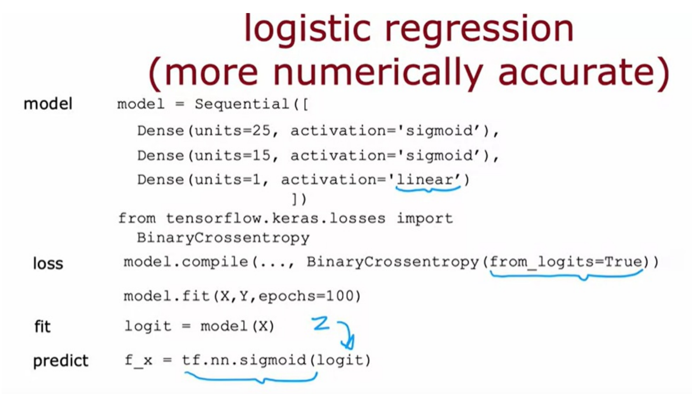
Numerical Roundoff Errors

- Tensorflow fix for logistic regression and softmax regression (though not fixing for logistic regression is generally acceptable)
- However, problem will get worse in softmax regression



Predict step changes:



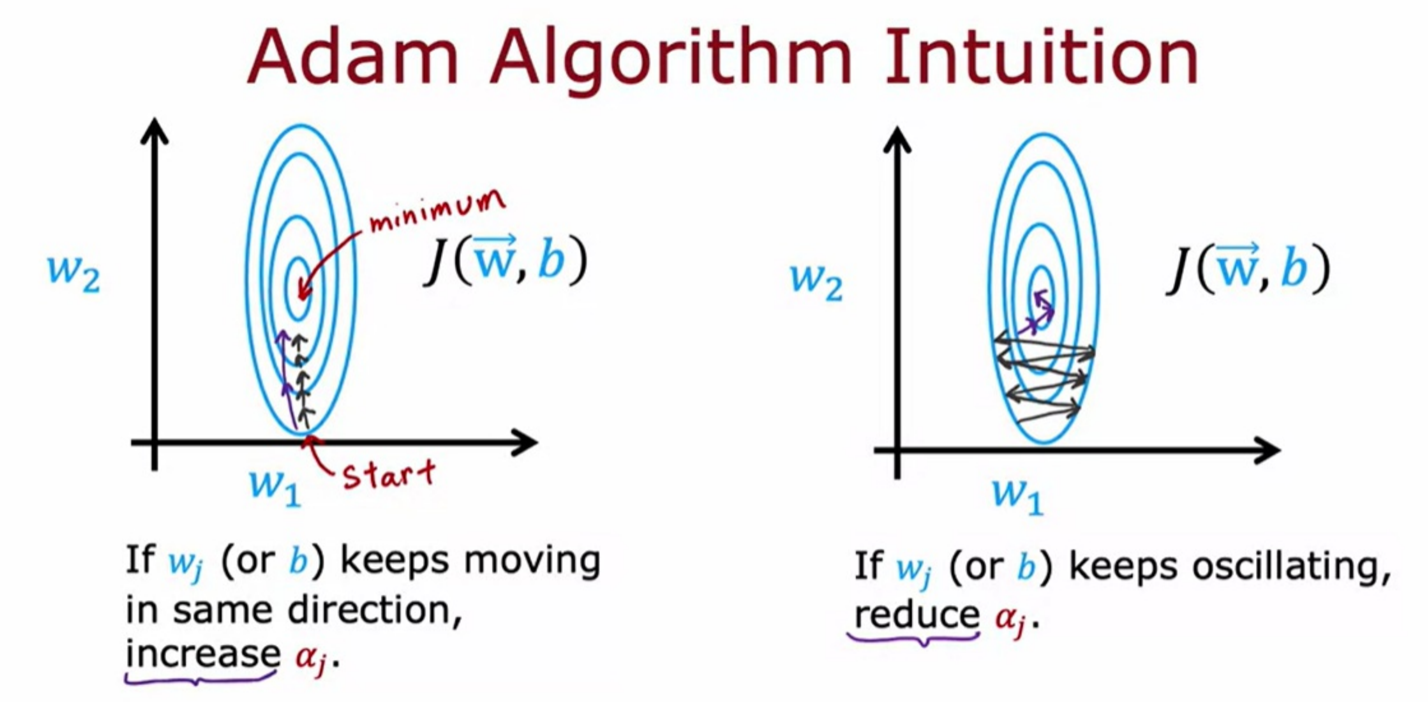


Additional Neural Network Concepts

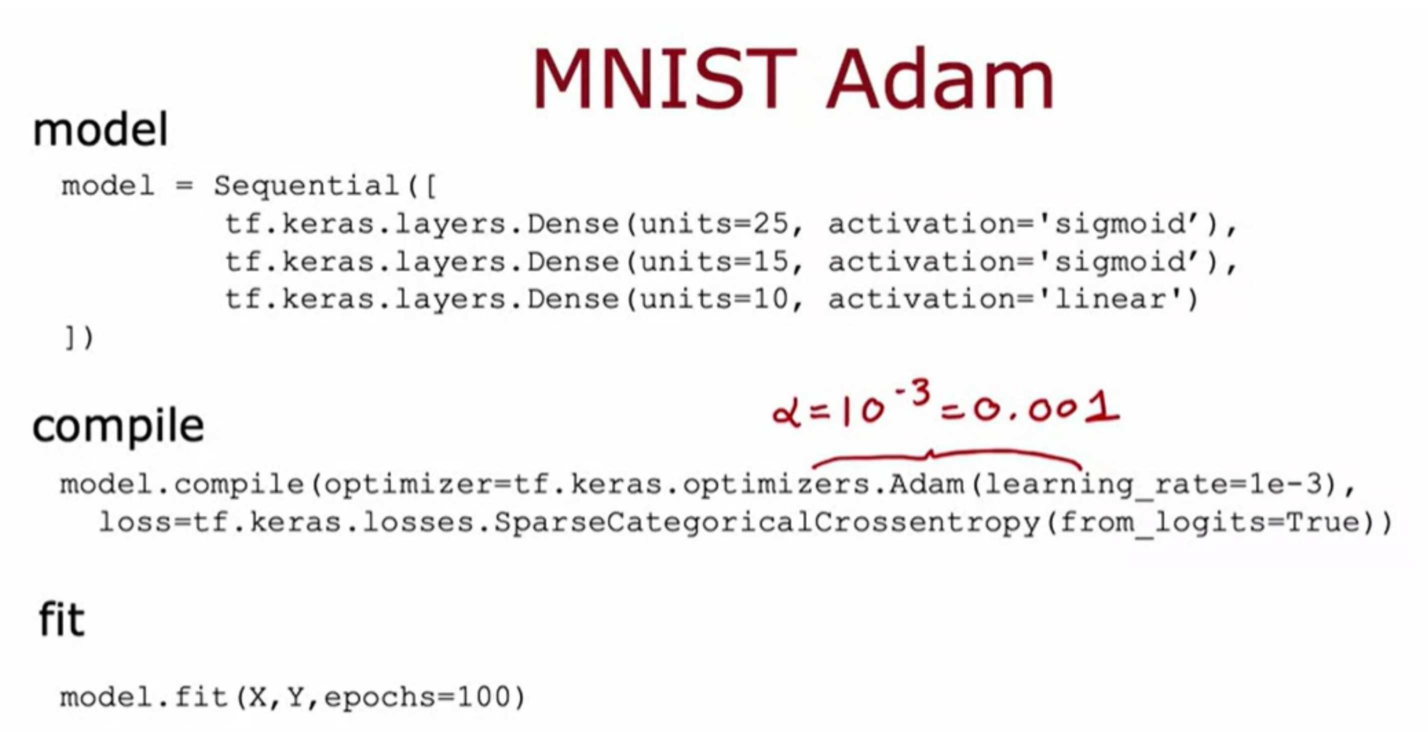
Advanced Optimization

Adaptive moment estimation (Adam) Algorithm

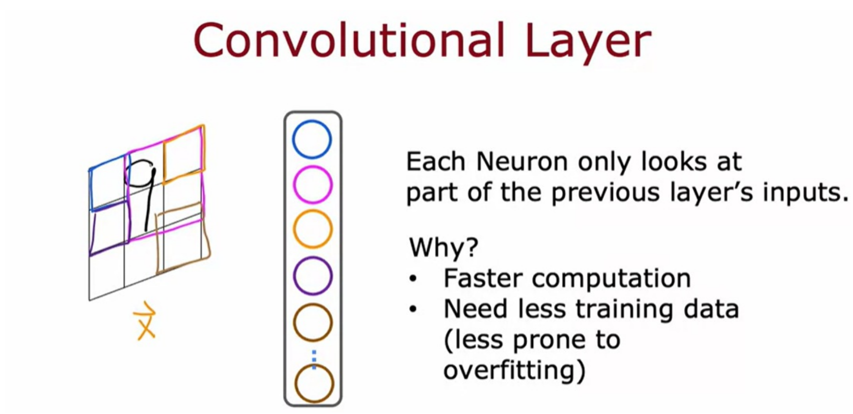
- Having different learning rates for each feature
- If weight increase s in 1 direction, increase learning rate (if going 1 direction down the hill, go faster, if zigzagging, go slower)

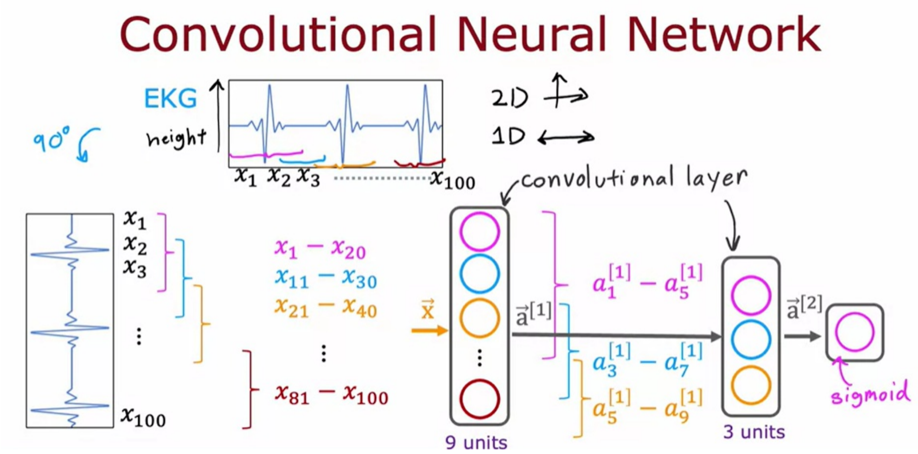


TF Implementation



Additional Layer Type

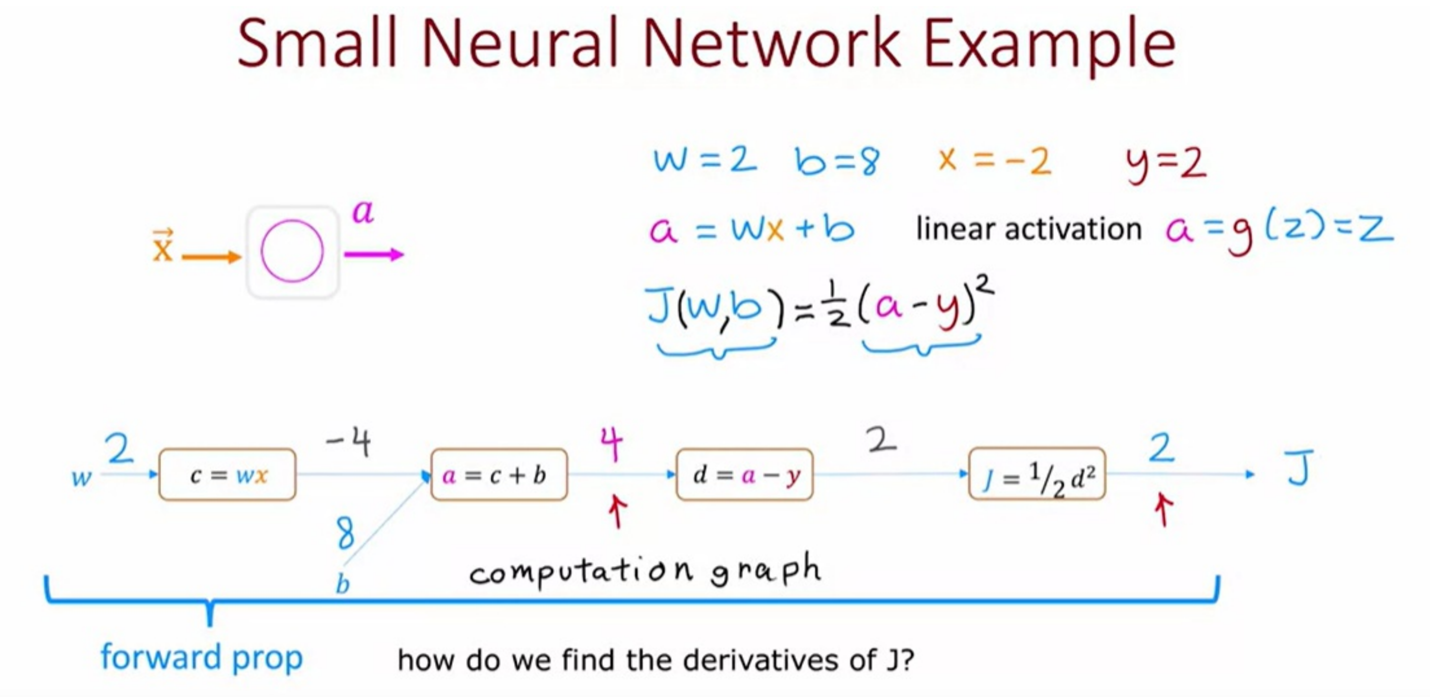




Backpropagation

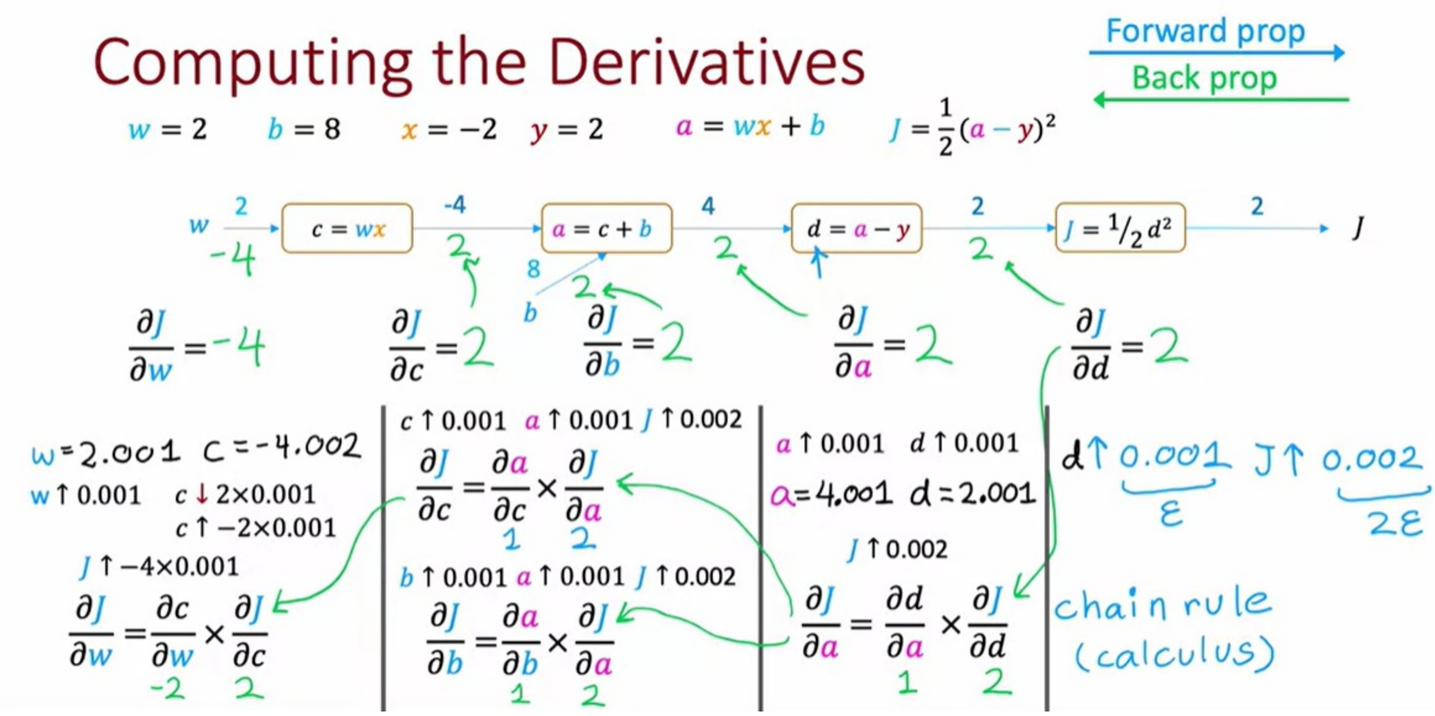
Computation Graph

Computation graph for Forward Propagation



Computation graph for Backward Propagation

- Computing the derivatives at each step is part of backward propagation
- Using the constant derived fom



Why use Backward Prop?