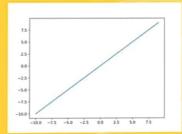
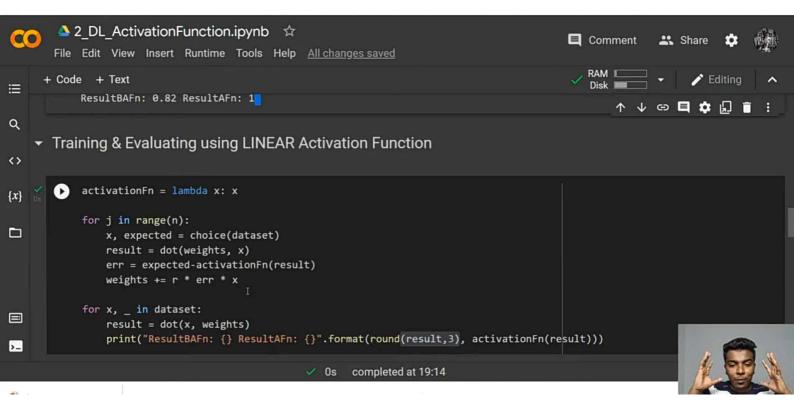




Linear functions are pretty simple. It returns what it gets as input.



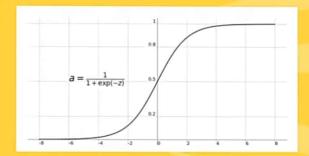




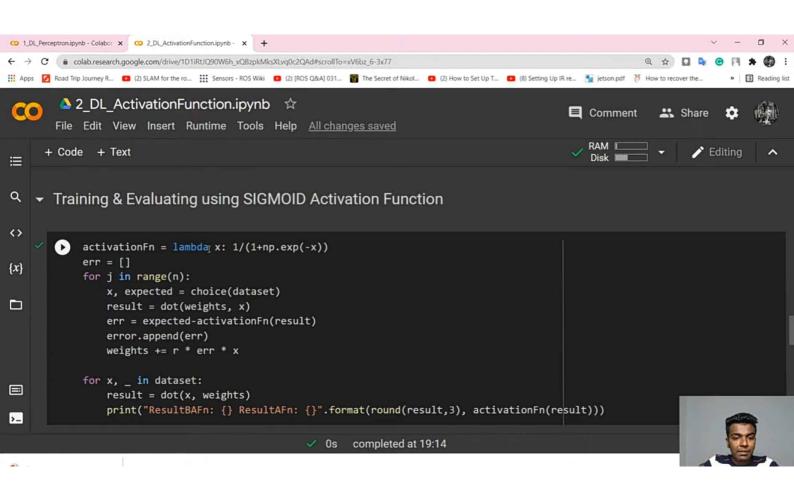
SIGMOID ACTIVATION FUNCTION

 \Diamond

- Sigmoid function returns the value between 0.0 and 1.0
- Inputs that are much larger than 1.0 are transformed to the value 1.0, similarly, values much smaller than 0.0 are snapped to 0.0
- It has a characteristic S-shaped curve



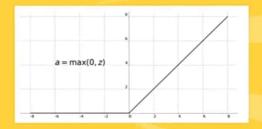




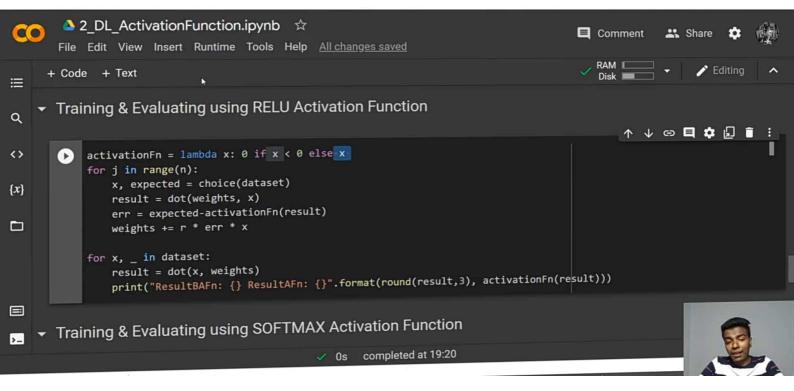
RELU ACTIVATION FUNCTION

\$

- RELU is less computational expensive than the other non linear activation functions
- RELU returns 0 if the x (input) is less than 0
 RELU returns x if the x (input) is greater than 0

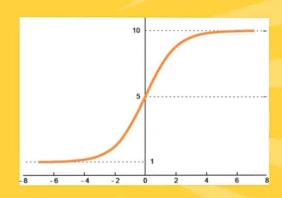




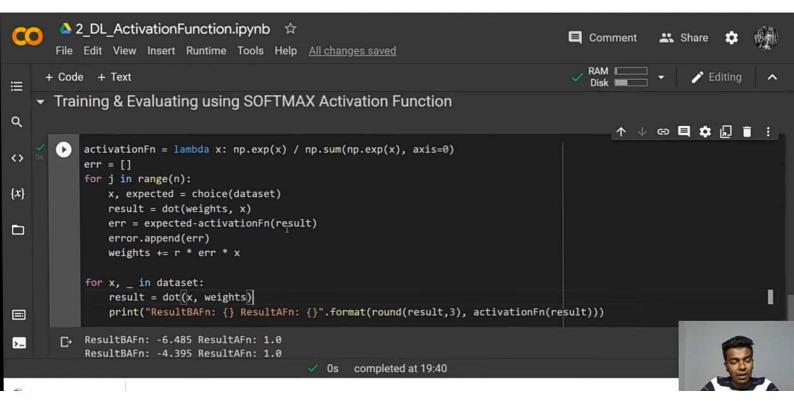


SOFTMAX ACTIVATION FUNCTION

- Softmax turns logits, the numeric output of the last linear layer of a multi-class classification neural network into probabilities
- The function is great for classification problems, especially if you're dealing with multi-class classification problems, as it will report back the "confidence score" for each class







Deep Learning Terminology - 1



Batch & Batch size

- The set of examples used in one iteration (that is, one gradient update) of model training.
- Batch Size: The number of examples in a batch



deep learning terminology - 2



Normalization

- The process of converting an actual range of values into a standard range of values, typically -1 to +1 or 0 to 1.
- For example, suppose the natural range of a certain feature is 800 to 6,000. Through subtraction and division, you can normalize those values into the range -1 to +1.



Deep Learning Terminology - 3



Batch normalization

 Normalizing the input or output of the activation functions in a hidden layer.

Benefits

- · Make neural networks more stable
- · Enable higher learning rates
- · Reduce overfitting

