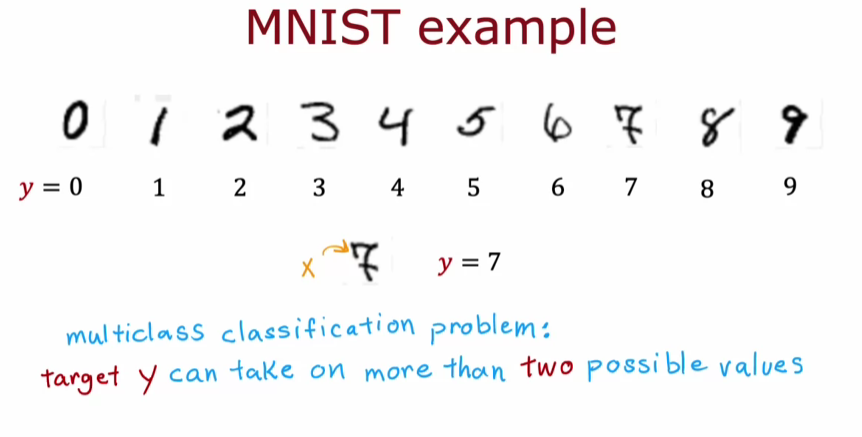
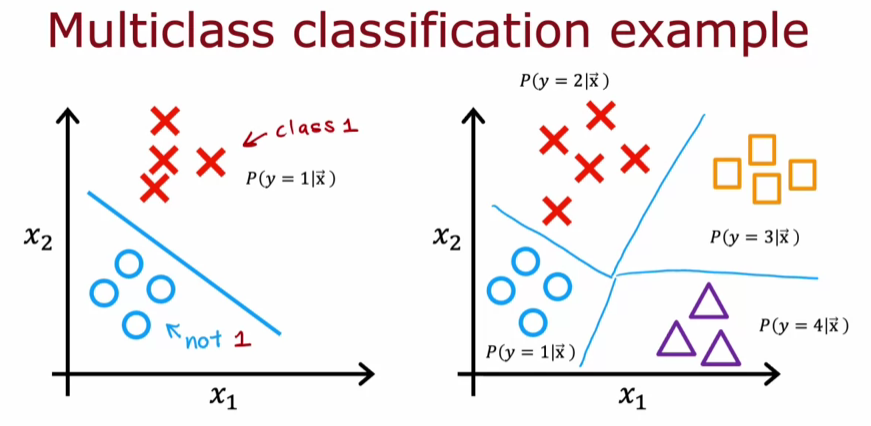
**MULTICLASS CLASSIFICATION**

**Understanding multiclass classification**

* **Multiclass classification is used when there are multiple discrete categories to classify, such as recognizing digits (0-9) or diagnosing multiple diseases.**
* **Unlike binary classification, where the output is limited to two classes (0 or 1), multiclass classification can involve several classes, such as four different types of defects in a product.**

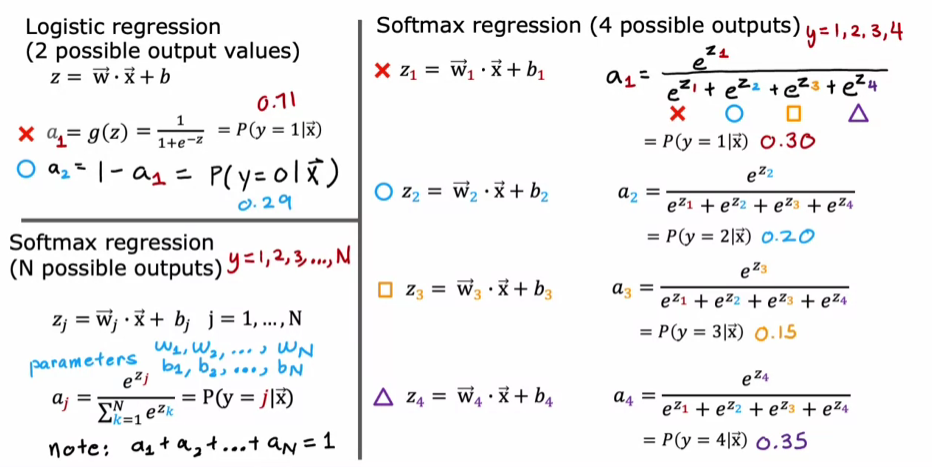
**Data representation in multiclass classification**

* **In multiclass problems, datasets can represent multiple classes visually, with different symbols (like circles, squares, and triangles) indicating different categories.**
* **The goal is to estimate the probability of each class, rather than just one, allowing for a more complex decision boundary.**

****

**SOFTMAX**

* **Softmax regression extends logistic regression by allowing the output variable (y) to take on multiple values (e.g., 1, 2, 3, or 4) instead of just binary outcomes.**
* **It computes a set of values ( ) for each class, using the formula ( ), where (  ) and (  ) are parameters for each class.**

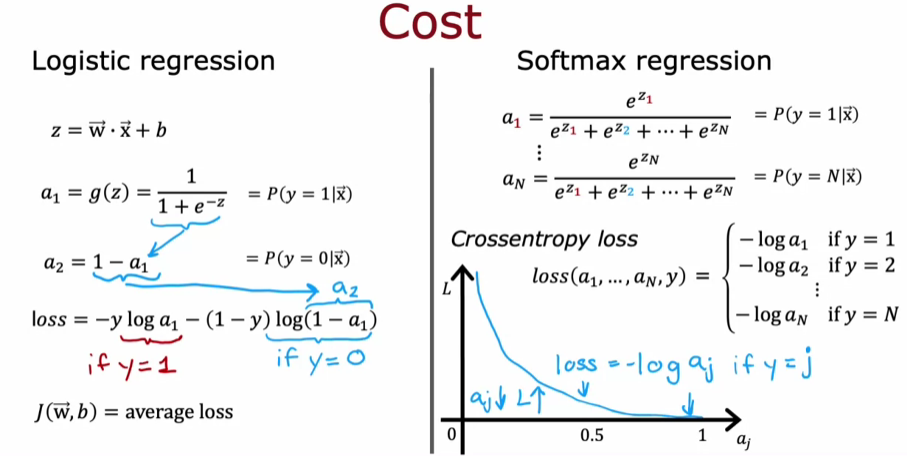
****

**Calculating Probabilities**

* **The probabilities for each class are calculated using the softmax function: (  ), ensuring that all probabilities sum to 1.**
* **This means that if you know the probabilities for classes 1, 2, and 3, you can easily find the probability for class 4 by subtracting the sum of the known probabilities from 1.**

**Cost Function for Softmax Regression**

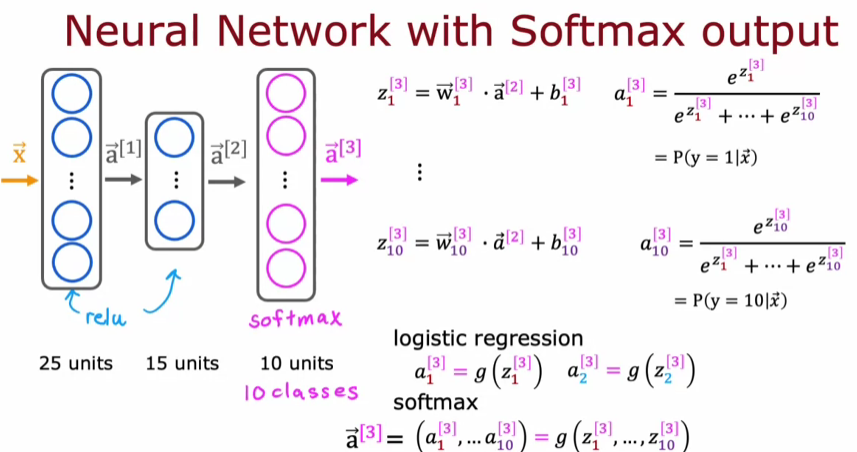
* **The loss function for softmax regression is defined as the negative log of the predicted probability for the true class label, incentivizing the model to maximize the predicted probability for the correct class.**
* **For each training example, only the loss corresponding to the actual class label is computed, which helps in optimizing the model effectively.**

****

**NEURAL NETWORK WITH SOFTMAX OUTPUT**

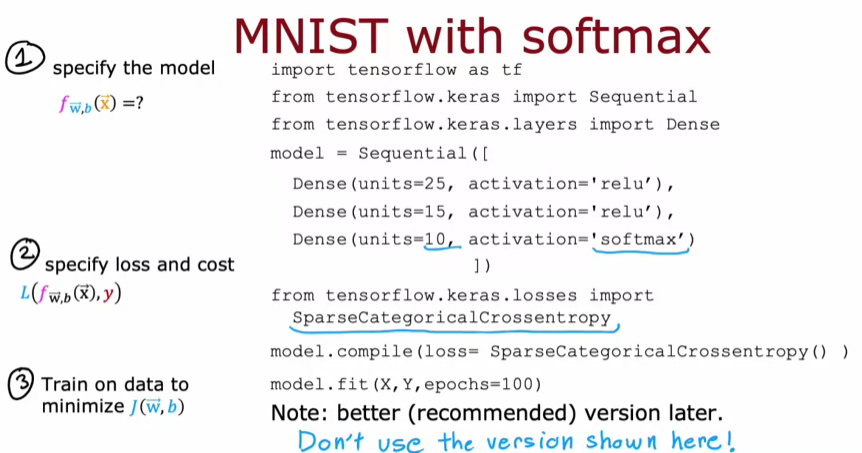
**Understanding the Softmax Output Layer**

* **To classify digits from 0 to 9, the Neural Network is modified to have 10 output units, forming a Softmax output layer.**
* **The forward propagation process remains similar, with activations computed for each layer, leading to the output layer's activations.**

****

**Implementing in TensorFlow**

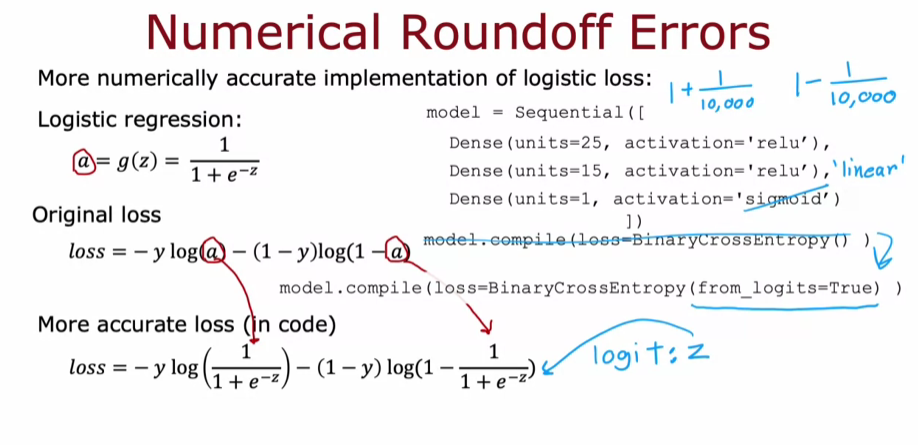
* **The implementation involves creating a model with three layers: two hidden layers with ReLU activation and a Softmax output layer for 10 classes.**
* **The cost function used is SparseCategoricalCrossentropy, which is suitable for multi-class classification where each input belongs to one category.**

****

**IMPROVED IMPLEMENTATION OF SOFTMAX**

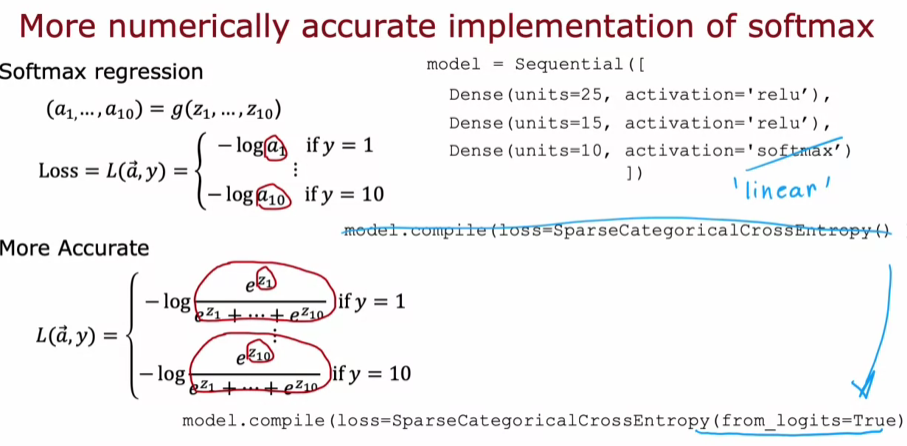
**Understanding Numerical Stability**

* **The way computations are structured can lead to numerical round-off errors, especially when dealing with very small or very large numbers in floating-point arithmetic.**
* **By rearranging terms in the computation, TensorFlow can achieve more accurate results, particularly in loss functions.**

****

**Logistic Regression Implementation**

* **In logistic regression, the output activation can be computed directly without explicitly calculating intermediate values, which helps in reducing numerical errors.**
* **The recommended approach involves using a linear activation function and specifying the loss function directly, allowing TensorFlow to optimize the computation.**

****

**Enhancing Softmax Regression**

* **Similar to logistic regression, softmax can also benefit from a more numerically stable implementation by directly specifying the loss function.**
* **The final layer of the neural network can use a linear activation function, allowing TensorFlow to compute the necessary values more accurately.**

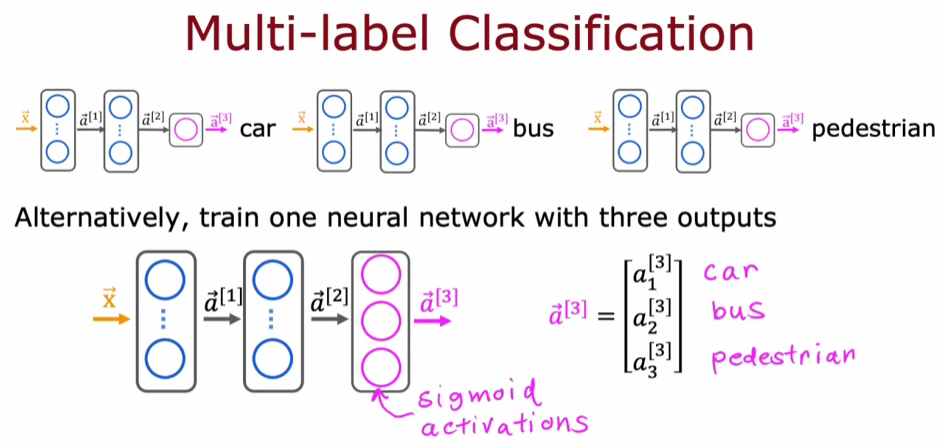
**CLASSIFICATION WITH MULTIPLE OUTPUTS**

**Understanding Multi-Class vs. Multi-Label Classification**

* **Multi-class classification involves a single output label that can belong to one of several categories, such as identifying handwritten digits (0-9).**
* **Multi-label classification allows for multiple labels to be associated with a single input, such as detecting cars, buses, and pedestrians in an image.**

**Approaches to Multi-Label Classification**

* **One approach is to create separate neural networks for each label (e.g., one for cars, one for buses, and one for pedestrians).**
* **Alternatively, a single neural network can be trained to detect all labels simultaneously, using a vector of outputs with a sigmoid activation function for each label.**

****