

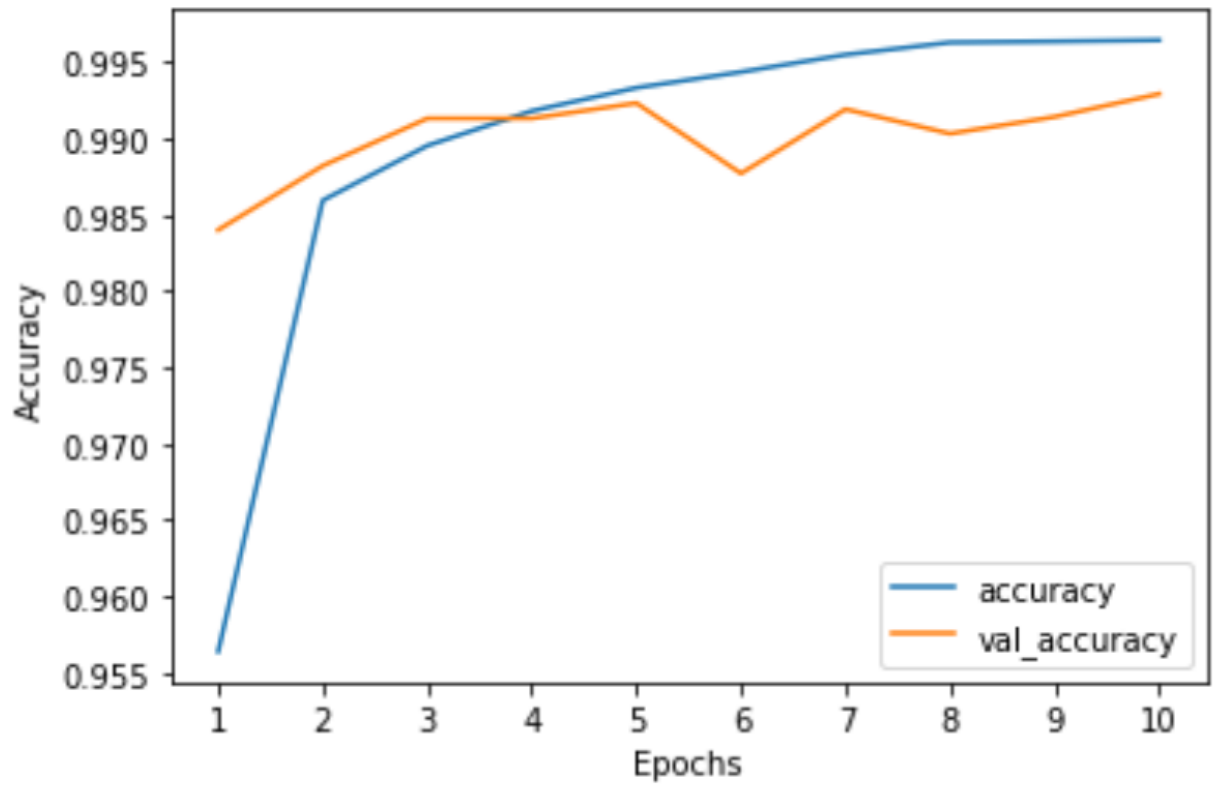
Error Decision: Categorical Cross Entropy

Definition:

Cross entropy is a loss function that calculates a score based on the difference between the actual and predicted probability distributions for each class in a given problem. The score is minimized and an ideal cross entropy value is 0, meaning the expected and actual probability distributions are identical.

Reasoning:

We chose to use cross entropy as our loss function because it yielded the best accuracy results in our limited testing runs. Cross entropy is an industry standard model for multi-class classification problems, and we did not feel a need to stray from it given the small number of classes in the mnist dataset. This did require us to use one hot encoding for our labels, but doing so did not meaningfully impact runtime and we did get 99% accuracy measurements as opposed to sparse cross entropy, which was only 98%. We thought that this tradeoff was worth the increased accuracy, so we stuck with standard cross entropy.



As seen in the above graph, our model achieved accuracy results well within the acceptable range, with an accuracy above 99% by the 10th epoch.