Assignment Week 5

Due: 2/19/2024

Shimon Greengart

Softmax is a function used to help solve a multiclass classification problem in machine learning, predicting which class any given datapoint is in. It takes as inputs a logit for each class, using a different set of weights and bias for each class. The function is a ratio between each logit given and all logits. However, each logit is e raised to that power to accentuate smaller differences, making it more likely for the model to pick one class as much more likely than the others. The output is a probability for each class; since this is a multiclass problem and not a multilabel one, they all add up to 1, meaning the model can only predict each datapoint being in a single class. The largest probability determines which class the model predicts. A model using softmax is trained using categorical cross-entropy.