Name:
NetID:

INFO 4940: Worksheet 00

1. Describe, in your own words, what the Softmax function does.

The Softmax function rescales inputs such that all elements lie in the range [0, 1] and sum to 1.0.

2. log (xy) =

log(xy) = log(x) + log(y)

3. Say you have a language model with a 10,000 token vocabulary. If you pass this model 5 input sequences that are all 10 tokens long, what will be the shape of the model's output?

The output shape will be (Batch, Length, Logits) -> (5, 10, 10000)