Special Topics in Applications (AIL861)
Artificial Intelligence for Earth Observation
Lecture 14

Instructor: Sudipan Saha

Classification: Label and Confidence

Average predicted probabilities

Label – usual arg max

Confidence – usual max

Softmax - Overconfident

"On Calibration of Modern Neural Networks" – Modern neural networks are over-confident

✔ Why modern?

✔ Over-confident – decreased entropy. Solution – Increase the entropy.

Temperature Scaling

✔ Use a single scalar parameter T>0 for all classes

✔ Divide the logit by temperature scalar before computing Softmax (**soften the Softmax**)

✔ T = 1 corresponds to the case of the actual probability

✔ T tends to infinity – all classes are assigned equal probability (maximum entropy)

Two Types of Ensemble

- Randomization

- Boosting

Randomization: Use Subset of Data

Why and why not?

Random Initialization

✔ Of neural network parameters.

✔ Ensemble as uniformly-weighted mixture model.

Dropout

✔ Different neurons dropped – different networks

✔ Each such network can be considered as a Monte Carlo sample from the space of all available models

Dropout

✔ Simply apply dropout at test time.

✔ Instead of one prediction, we get many, one for each model.

✔ Score can be averaged or distribution can be computed.