AC40001 - Daniel Blackley

Dissertation Plan

Aim: A comparison of Dropout and Weight Uncertainty as Bayesian Approximations in Skin Lesion classification

Full Plan: I plan to apply Dropout and Bayes by Backprop methods to identify uncertainty in Deep learning models for skin Lesion Classification; Bayesian methods allow for a framework to measure model uncertainty but usually come with a computational cost. I hope to compare these two Bayesian methods with a standard Softmax response as a baseline and identify a balance between computational complexity and test accuracy. Implementation of these Models will be achieved using the Pytorch and Pyro libraries, using Numpy and Matplotlib to record results.

It should be noted that to ensure the testing remains fair I plan to implement the Bayesian methods only in the final two layers of the network, switching between the two methods while keeping the overall architecture of the network the same. I plan on using the Adam optimiser but the loss function recommended by the Bayes by Backprop paper[5] is the ELBO loss function. I will likely therefore use the ELBO loss function for Bayes by Backprop, but use cross-entropy for the standard Softmax and Dropout Implementation.

The dataset I will use is the ISIC 2019 Dataset, This dataset is particularly suited to the task as the training data consists of 8 different categories of skin lesions, and the test data has an extra outlier class which consists of either dermoscopic images or random abstract objects, e.g. a circle. The Network must be able to express some level of uncertainty at this outlier class and choose not to guess.

The papers I will be basing my research on will be Dropout as a Bayesian Approximation[4] and Weight Uncertainty in Neural Networks [5], which both demonstrate their respective Bayesian models on classification tasks. These Papers both use the MNIST dataset, however Machine Learning can be applied to Skin Lesion Classification as discussed in the ISIC research papers [1, 2, 3]

Frameworks: Pytorch and Pyro

Dataset: ISIC 2019: <https://challenge2019.isic-archive.com/>

References/Main Papers:

[1] Tschandl P., Rosendahl C. & Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5, 180161 doi.10.1038/sdata.2018.161 (2018)

[2] Noel C. F. Codella, David Gutman, M. Emre Celebi, Brian Helba, Michael A. Marchetti, Stephen W. Dusza, Aadi Kalloo, Konstantinos Liopyris, Nabin Mishra, Harald Kittler, Allan Halpern: “Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)”, 2017; arXiv:1710.05006.

[3] Marc Combalia, Noel C. F. Codella, Veronica Rotemberg, Brian Helba, Veronica Vilaplana, Ofer Reiter, Allan C. Halpern, Susana Puig, Josep Malvehy: “BCN20000: Dermoscopic Lesions in the Wild”, 2019; arXiv:1908.02288.

[4] Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. ArXiv, abs/1506.02142.

[5] Blundell, C., Cornebise, J., Kavukcuoglu, K., & Wierstra, D. (2015). Weight Uncertainty in Neural Networks. ArXiv, abs/1505.05424.