

ML CS503: Lab Assignment 4

Project Assignment

Due Date: 24th April 2022 11:59 PM | M.M.: 100

Instructions:

1. This lab assignment is for those students who are not doing the term project.
 2. This lab assignment needs to be done individually and it carries the 12% weightage.
 3. Start by going through the research paper first before you start coding.
 4. For queries you should comment on the Google classroom. Alternatively, you may email the course TAs or course instructors.
 5. All code must be written using a python notebook (preferably using google colab).
 6. Submission must be done in the Google classroom. The code as well as the accompanying observations should be made part of the colab notebook.
 7. Code Readability is very important. Modularize your code by making use of classes and functions that can be flexibly reused wherever necessary. Also use self explanatory variable names and add comments to describe your approach wherever necessary.
 8. Students are expected to follow the honor code of the class. Discussions and interactions with your classmates to take help in this assignment is not allowed.
 9. Use the following naming convention: Name_rollnumber_lab4.ipynb for submission.
-

1 Predictions using Human-AI team

This assignment is based on the following research article appeared in 34th edition of Advances in Neural Information Processing Systems (NeurIPS 2021).

Kerrigan, G., Smyth, P., Steyvers, M. (2021). Combining Human Predictions with Model Probabilities via Confusion Matrices and Calibration. Advances in Neural Information Processing Systems, 34.

This paper has been provided as part of this assignment. In this paper, the authors proposed a set of algorithms that combine the probabilistic output of a machine learning model with the human predictions on the same input. This paper shows how the accuracy of predictions are driven by the confidence of machine learning model in its output. Empirical studies of the proposed model have been performed using image classification task with CIFAR-10 and a subset of ImageNet datasets.

The authors have released code for this project here https://github.com/gavinkerrigan/conf_matrix_and_calibration

You are allowed to use the available code of this paper to complete this assignment.

2 The Tasks

For the purpose of running experiments, you may only consider CIFAR-10 dataset.

Q.1. Reproduce Results: [25 points]

Reproduce the results depicted in the paper only for CIFAR-10 dataset using the authors' code. Out of the four pre-trained CNN models considered in the paper for CIFAR-10 experiments, it will okay if you consider any two dissimilar approaches.

Are you able to get the same results as reported by the authors? Comment on the challenges faced. List all the hyper-parameters values that authors didn't mention and you need to make your own choice for the same. [Any new insights from the paper will be subjected to additional bonus marks.](#)

Q.2. Model Multiple Humans: [35 points]

Consider an approach where more than one human decision makers (say 3) are available. Each human provides one label for the given input image. The final human predicted label can then be chosen based on majority voting rule. If none of the human provides the same label, then a random label from the three labels are chosen. To understand this, consider the following example:

1. If human 1 suggest that the output is cat, human 2 suggest that the output is deer, and human 3 suggests that the output is bird, then one of the output from cat, deer and bird will be selected uniformly at random.
2. If human 1 suggests cat, human 2 suggests bird, and human 3 suggests cat, then the human label will be considered as cat.

Mention your approach for modelling more than one human on the CIFAR-10 dataset. Now, regenerate all the results generated in the previous section using the multiple human labels.

Comment on your observations. How does the error rates perform with respect to different error rates of human models? Also, mention the use-cases or application with such an approach where there are more than one human decision makers in a human-AI team.

[\[Bonus Marks:\] Provide a better mathematical model to incorporate multiple humans based on the theory proposed in the paper. \[Extra 20 points\]](#)

Q.3. Neural Network for Calibrated Probabilities: [35 points]

Now consider another case where the final "calibrated probabilities" are produced using a neural network model. Let's call this neural network model as **Team model**. Here, the output of the Team model will have same units as the output of the machine learning model however the input to the Team model will be the output of machine learning model and human predicted label (consider the base case of only one human).

Train the Team model and regenerate the results as in the base-case (Q.1). Compare your results with the base-case and list your observations. Do you find the Team model a better approach to combine human and machine model outputs?

***** All the Best*****