MiniProject 3: Multi-label Classification of Image Data

COMP 551 (001/002), Fall 2020, McGill University

Please read this entire document before beginning the assignment.

Preamble

- **Due on December 10th at 11:59pm**. Can be submitted late until December 15th at 11:59pm with a 20% penalty.

- To be completed in groups of three, all members of a group will receive the same grade. You can work with new group members or same members that you have collaborated before with, whichever works best.

- To be submitted through MyCourses as a group. You need to register your new group on MyCourses and any group member can submit the deliverables, which are the following two files:

    1. **code.zip**: Your data processing, classification and evaluation code (.py and .ipynb files).

    2. **writeup.pdf**: Your (**max 2-page**) project write-up as a pdf (details below).

- TA: Samin Yeasar Arnob, samin.arnob@mail.mcgill.ca

- We recommend to use **Overleaf** for writing your report and **Google colab** for coding and running the experiments. The latter also gives access to the required computational resources. Both platforms enable remote collaborations.

- This project also includes a Kaggle competition. **It is very important that you follow the following rules**:

    1. **Do not** register at Kaggle unless you have finalized your group in MyCourses.

    2. Create a group at Kaggle, where the group name should be the same as the registered group in MyCourses.

    3. Double registration or submissions from multiple group members at Kaggle is not allowed.

    4. You can submit up to 5 solutions each day.

Problem definition

In this mini-project, we will develop models to classify image data. We will use a "modified" MNIST dataset: `https://drive.google.com/file/d/1LcKqf1d7bctw5lx0YZf31kCUF0zEYOsi/view?usp=sharing` where each image contains 1-5 digits. The task is to implement a model for this multi-label classification task.

If the number of digits in an image is less than 5 then the remaining labels are associated with a special class called "no-digit". As can be seen in the figure-1, no-digit is class "10". This means we have a total of 11 classes (0-10 digits and no-digit), and each image is associated with 5 such classes. The dataset contains both training and test data, where the label for the test set is not provided.

You are free to use any Python libraries you like to extract features and pre-process the data, evaluate your model, and to tune the hyper-parameters, etc. Your model should be a neural network and it should be trained by automatic differentiation.
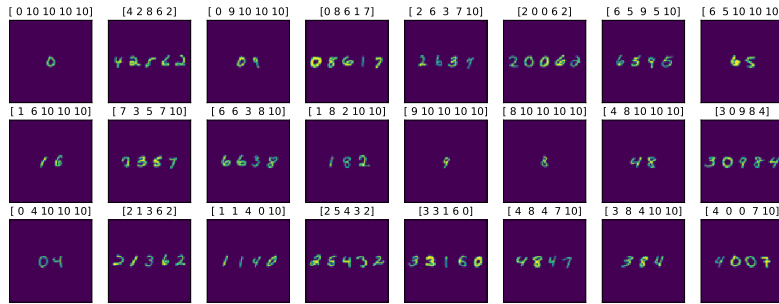
Figure 1: Sample data from modified MNIST dataset

This project can be divided into two part:

1. Implement a deep neural network model for the task and report the followings:

   - Describe the proposed model and training objective for the task in detail.

   - Report the best performing model architecture and its hyper-parameter in a table in the Appendix.

   - Justify the of choice of hyper-parameters (model architecture, optimizer and its parameters) that lead to this final model.

   - Compare and report the validation and train performance of your best performing model as a function of training epochs.

2. Participate in COMP 551 Kaggle competition:

   - Once you have formed a group in MyCourses, go to the **Multi-label Classification of Image Data** page below and join the competition using the same group name:
     `https://www.kaggle.com/t/162b731e49e44d7e914aabc04f449a8f`

   - 50% of the grade is allocated based on the performance on the kaggle competition. Once you have a trained model, you need to predict the test label and submit at the Kaggle competition.

   - The evaluation is split private and public score (based on a further split of the test data). The private will be revealed after the deadline and it will be used for grading.

   - Here's how the submission file should look like, **"sample.csv"**:
     `https://drive.google.com/file/d/1xcmkOcshHTtROxYf9ca8Q7WFJ6R6OHIa/view?usp=sharing.`
     It should have two columns ("Id","Label"), where the Label will be the sequence of predicted labels. For example, have a look at the sample table, where predicted $[0, 1, 2, 8, 10]$ has been converted as a sequence of $[012810]$.

Id	Label
0	14923
1	27238
2	65192
3	012810

Table 1: An example of submission file

Evaluation

Evaluation has two components, where the first component is similar to previous mini-projects:

- 10% completeness
- 15% correctness
- 10% writing and code quality
- 15% originality and creativity

The second component is your based on your group's performance on the private part of the test set in the Kaggle competition (50%):

- 5% for successful submission of your results in the competition
- 20% for obtaining the performance better than the baseline
- 25% based on ranking

Deliverables

In addition to participating in the competition, please submit:

1. **code.zip**: Your data processing, classification and evaluation code (as a combination of .py and .ipynb files).

2. **writeup.pdf**: Your (**max 2-page**) project write-up as a pdf (appendices may be used for additional graphs and tables if really needed). In contrast to previous mini-projects there is no need to analyze the dataset unless you find it important to discuss to justify your proposed method.