Course Project:
Deep Learning Research Project

CMPT 828 - 2020

Description

The course project is intended to provide an opportunity for independent study in one particular area of deep learning. It is expected that the project will be aligned with the student's research directions and hopefully will form a useful exercise or component of their thesis.

Milestone 0: Complete sign-up form with project idea – 0% (Due: Jan. 31)

Complete the google form (link on moodle) with one or more project ideas in descending order of interest. Include a relevant paper for each, if possible. Discuss these ideas with the instructor during the "scrum meeting" after a Thursday lecture before proceeding with the proposal. In particular, note which available dataset with labels will be used.

Milestone 1: Proposal – 15% (Due: Feb. 7)

A short (max 2 page) proposal setting out the specific task you will try to perform with a deep learning model. Describe the motivation for this task: why is it important to solve, in which domain will you investigate the task, e.g. images of people, and in which other domains is the task transferrable, e.g. also applicable to images of plants. Classify the task in terms we use within the course, i.e. is it a regression problem, a classification problem, or something else. What is the input to the task, what is the output? Describe in general terms the dataset(s) that are available for you to use. Describe in general terms how you will measure if the predicted outputs are good or bad.

Append a relevant research paper that you will present in for the **Research Paper Presentation**.

Milestone 2: Dataset Plan – 20% (Due: Feb. 14)

Describe in detail the dataset(s) you will use for your project, including: the type of data, the size of each sample, the label for each sample (if applicable), the number of samples in the dataset, and groupings or sub-datasets it is divided into. Perform some statistical analysis of the dataset to quantify it's distribution. For example, for a classification task plot the distribution of class examples. Note any potential problems with the dataset, e.g. inspect the data for label noise, inspect a sub-set of the data for quality and consistency. List any related public datasets that are similar to the one you will use.

Milestone 3: Evaluation Plan – 20% (Feb. 28)

Describe your plan for evaluating your task in detail. What is the loss function? What is the training, validation, test set split? Will you do cross-validation? How will you initialize? How will you

optimize hyper-parameters? Which ones? Will you evaluate different architectures? Different size architectures? What regularization scheme will you use?

Choose at least one relevant "baseline" method to compare to. Published baselines are good to include as well, but you should implement one baseline yourself. The baseline could be a simple model as well, e.g. linear regression, or a standard machine learning method with reasonable hand-tuned features. Describe the baseline(s) and the expected improvement with your deep learning approach.

Milestone 4: Final Presentation – 0% (—)

This will be omitted in lieu of the **Research Paper Presentation**.

Milestone 5: Final Report – 45% (Due: April. 7, late submissions accepted until April. 30)

A conference paper (ACM style, 8 page) formatted report that describes the work accomplished. Please use a standard organization: *Introduction*, *Methods*, *Results*, *Discussion*, and *Summary / Conclusions*.