Team name: Deep Blue Deep Learning

Team members: Haoyu Zhang, Chuanrui Liu, Dingyuan Liu, Yunzhou Liu, Yiyao Qu, Yumeng Zhang
Date:9/4/2020

(This journal has to be uploaded to your Sakai group folder by 11:59 pm every Friday **and** your team webpage till and including October 30). Change the file name to "Team_name_MM-DD-2020" where MM = month and DD= day of upload and then upload to your Sakai team folder).

Team roles for this week (write down name):

Facilitator(s): Dingyuan Liu, Yiyao Qu

Recorder(s): Yumeng Zhang, Chuanrui Liu

Deliverer(s): Yunzhou Liu

Planner(s): Haoyu Zhang

Team contact: Yiyao Qu, yiyao@live.unc.edu

See last page for description of roles. Obviously one person can take more than one role or there can be more than one person per role or make your own roles!

1. **Describe briefly what the main goal of your team is (so the peer reviewer has some context). E.g. we are working on image classification for blah de blah. Our goal is blah de blah etc. In the initial part of the semester before your proposal it is ok to put down "we are still coming up with ideas on team project".**

Our team decided to focus on building a recommendation system for a user-based online service. For example, we are looking at the Amazon movie dataset which contains users' ratings for all the movies. Our goal is to analyze the users' interactions with the interface and their feedback to build a well-defined recommendation system and improve users' overall experience at the website.

I. **What was done this week regarding the project:** If you want to include code include this in the Appendix. Describe what the group did (including contributions of individual team members) with regards to the group project this week. Give enough details so I understand what you folks have been doing over the week. Include dates of your meeting(s) and who met on these days.

We split into three groups on Monday. Yiyao Qu and Yumeng Zhang are in group one. Haoyu Zhang and Chuanrui Liu are in group two. Yuzhou Liu and Dingyuan Liu are in group three. Each team had four days to look for datasets regarding recommendation systems, and then we met on 9/4/2020 to discuss our findings about potential workable datasets.

On Tuesday, Yiyao Qu as our representative came to the Professor's office hour and asked the questions about our candidate datasets and methods for building recommendation systems in general. Afterwards, we have a better understanding of the topic we are working on.

II. **What were obstacles faced if any in working on the project?** This could be technical (like not being able to implement or understand particular techniques) or time issues (midterms for other courses etc).

Currently, we are facing the challenge of finding the proper dataset to work on, as the dataset we need for our project should contain considerable observations and substantial features to facilitate the machine learning process for this recommendation system.

For example, we have identified two similar candidate datasets on movie recommendation during our research process. The link of the first dataset is https://www.kaggle.com/roshansharma/movies-meta-data, and the link of the second dataset is https://www.kaggle.com/bandikarthik/movie-recommendation-system. According to our research on the details of these two datasets, while the two datasets contain a good amount of features to describe the movie itself, the first dataset of the movies metadata does not connect the movie information with specific users. However, the second dataset of the movie recommendation system includes the "userId" and "movieId" to relate these two features. In this way, we choose the second dataset as our final candidate datasets to compare with our further potential datasets.

III. **What is the plan for the next week including what each team member is planning to work on in the next week. Describe goals and potential timelines (" I plan to finish understanding x to see if it can be implemented for our project by Wednesday etc".)**

We are planning to finalize the dataset we are going to use and set the target for our final project. For this week, our team has already found four candidates for data sources. After comparing the advantage of each source, we suppose we will be able to finalize that in the middle of next week. And in the later half of the week, we will decide what to do with the data based on the usable information that it could provide. By next Tuesday, we will generally go through the data and revise potential problems. Also, we will finish the peer evaluation for the assigned teams. And by next Friday, we will build the setup for our project.

While in the weekly document above you will describe what your team did with regards to the team project (with proper attributions of who did what in the week) there are 4 pre-defined roles. I urge you to have different people do these jobs every week so that you gain experience in each of the jobs. There can also be more than one person per job for example 2 people recording the weekly journal.

Facilitator: Manages the group for this week including setting up times for group members to meet, making sure everyone has a say in the meetings etc.

Recorder: Person in charge of recording the meetings as well as the happenings of the week and describing what was accomplished in the meeting and writing up this report.

Deliverer: Person in charge of checking the entire journal and uploading the doc file to Sakai in the group folder as well as the representative of the group getting in touch with the instructor.

Planner: Person in charge of what will be happening next week as well as thinking about longer term goals (what more needs to be done for the project).

Team contact: yiyao@live.unc.edu.