CSE 472 (Machine Learning Sessional) - Term Project guideline

For students under the supervision of Dr. Mohammad Saifur Rahman (MDSR)

Document History

Version	Created/Updated on	Change List	Modified by
1.0	19-JAN-2022	First draft	MDSR

Project proposal (Due: 23-JAN-2023)

Prepare a single slide presentation that should cover the following aspects:

1. Problem definition
2. Dataset (including link and some basic statistics). If you can find a Bangladeshi dataset, that will be extra exciting to work on.
3. Proposed solution (architecture)
4. Performance metrics
5. Link to video presentation of your proposal - no more than 2 minutes long (strict). All members must participate in the proposal presentation. The video does not need to show the presenter(s).

You have to submit the presentation slide (PDF/power point) in Moodle.

This will bear marks.

Literature review (Due: 28-JAN-2023)

You must do a literature review as follows.

1. Is there published work on the dataset you will work on? If not, then at least there should be published work on the problem that you plan to work on.
2. Following the above criteria, find out some papers, do a basic skim through the papers to figure out
   a. What is the best result published in a reputed venue?
   b. What are some very recent works (in reputed venue), that may not necessarily have produced best results, but perhaps had offered novel architecture that allowed for an interpretive model.
3. Based on the above directions and your own analysis, choose one paper. Send that paper to me as a reference for this work.
4. Your goal should be to beat state-of-the-art methods so that we can publish this as a short research article in some reasonable conference, that has proceedings published in IEEE Xplore,

Springer, ACM ICPS or alike. Think of NSysS 2023, or alike. Being able to produce superior results is not a requirement for good grades, but thriving for it is.

This will bear marks.

Weekly updates

1. It is optional to provide weekly updates. However, regular progress check point helps in properly judging your efforts.
2. I will be available in my office room (Room 218) Tuesdays, 2:30 – 4:30 PM, for project related discussions, if you have any. Note that this is overlapping with thesis hours and your thesis supervisors' directives take precedence. You are not obligated to show up in this slot. If you have discussions, but cannot afford to attend this slot due to thesis supervisor's meetings, then talk to me to setup another time.

Weekly updates will not directly bear marks. However, marks will be given for sincere efforts and weekly updates can be an indication of sincere effort.

Project checkpoint (Due: Week# 12)

1. A check-point presentation must be made in the penultimate week of the term. This will bear marks and should include the following:
   a. Problem definition
   b. Dataset and its analysis (Statistics)
   c. Proposed solution (architecture)
   d. Loss function and its intuition.
   e. Performance report
   f. Challenges/Discussion

Project presentation (Due: Week# 13)

1. The final presentation must be made in the final week of the term. This will bear marks.
   a. Problem definition
   b. Dataset and its analysis (Statistics)
   c. Proposed solution (architecture)
   d. Loss function and its intuition.
   e. Performance report
   f. Comparison with state-of-the-art methods
   g. Challenges/Discussion

Special Notes from Supervisor

1. I love to work in bioinformatics, health informatics. If you feel excited about it, you can choose something from this field. For example –
    a. Detection of various diseases/conditions from medical images (e.g. COVID-19 detection from chest X-ray)
    b. Detection of cell/organ types from medical images (e.g. blood cell type from histology image)
    c. Protein secondary structure prediction, various attribute prediction of proteins and so on
2. Very recently, I have been working on Spatial Transcriptomics (ST) (gene expression profile at cell level granularity) which is heavily used in tumor progression studies. An interesting recent trend has been to predict ST signals from histology images. Search whether you can find some datasets around this prediction problem or any other prediction problem around spatial or single-cell transcriptomics.
3. I am also very interested to work on any deep learning project in general. So, even if you are not interested in the above, go for any interesting problem that excites you and I will be happy to help you in your pursuit of solving the problem.