Project guidelines ML- 2023

Team Size: 3-4

General

1. Choosing topic: You are free to choose any topic of your choice. Can look over Kaggle, aircrwod competitions. But a suggestion would be to choose smartly so that you do not burden yourself with exploring topics that you are not familiar with. For example, do not choose a topic which requires hardcore use of CV or DL, like semantic segmentation or using GAN's etc. Even simple topics can be explored a lot using basic ML techniques. Definitely, Projects help you to explore out of your comfort zone and makes the learning process more exciting.

2. Dataset: After choosing the topic try to read a few related papers, which will help you to understand datasets used and evaluation metrics. Check whether the dataset is public or not. Many times it happens students select a topic but when they start working over it finds data is not available. Seeing the current situation creating your own data will also be challenging.

3. Analysis: This is most important. Let us understand with a very simple example, suppose for a classification problem you opted for Random Forest as a baseline and SVM performs better than RF, then you should justify why RF is not working properly or why SVM is better. Perform Grid Search to convince your evaluation committee with your results.
And for baselines, simple techniques can be considered baselines and can try to improve results over the baselines by changing your model or adding more features.
Accuracy/evaluation metrics would only be used for comparison between models and for analysis + justification. You needn't worry about optimising it, but the accuracy shouldn't be completely unreasonable also. This would be evaluated on a case-by-case basis.

4. Another question is "How at this stage we will commit individual tasks". So when you will explore a bit about your project topic you will get to know different components accordingly it can be assigned. Mention things broadly and can finalise more during mid evaluation.

5. Please note that the aim of the course project is to make you learn about machine learning and not deep learning. Hence, the use of DL methods beyond ANN is not allowed for your course projects. Please strictly adhere to this.

Format for project submission (Proposal, intermediate report/ final)

1. Use the CVPR 2019 template for all project report submissions. You may use the template provided in the link or you can refer to CVPR's website for the same.
http://cvpr2020.thecvf.com/sites/default/files/2019-09/cvpr2020AuthorKit.zip

2. You are free to change the headings according to your requirements.
3. Add diagrams/graphs/flowcharts wherever necessary.
4. Use the final version of the template for submissions.
5. Stick to the page limits. For the proposal, it's 1 page. It will be different for subsequent deadlines.

Project Proposal
One submission for each group. The submission has to be in a proper format. Please follow the template attached. Only PDF files should be submitted.
Page Limit: 1.

Since it is a machine learning course it is recommended that you should try machine learning techniques only. Project evaluation will be more on your understanding of the problem and its analysis. For example, suppose you are using Random Forest as an ML technique over a dataset and the results are not good as compared to SVM, then ablation should be WHY? Random Forest did not perform well. How it can be improved? How the performance change by changing hyperparameters. Your future reports and presentation should convince the Instructor about your understanding of different plots, figures.

What is the purpose of submitting a Proposal?

The proposal will help you to pitch your idea. Accordingly, after the evaluation groups can be notified about the shortcomings of the idea. There may be a case where "many" (more than 4-5) groups having the same project can be asked to change. Some ideas could be very simple, TA's can help you to explore some other dimensions also.
Can even look on platforms like Kaggle for project ideas.

The following components are important in the proposal. (1x5) points.

Project Title
1. Motivation: Why this project? How did you think about this?
2. Related work: Any 3 recent related works or articles. And a brief description of them.
3. Timeline: How are you going to proceed with this project? A rough timeline.
4. Individual Tasks: Tasks each member would be responsible for. Future evaluation of individual members will be based on this only.
5. Final Outcome: What are you expecting from this project? What do you want to contribute to this idea?

Interim Project PresentationTemplate
Total 10 minutes slot, which includes presentation and QnA session. So ideally the presentation should be for around 6-7 minutes. Number of slides for each section can be flexible, but all the sections are mandatory.

1. Motivation (2 slides)
2. Literature review (2 slides) Discuss at least two Research Papers, not blogs or articles.
3. Dataset description (2-3 slides)
a) Different attributes, some visualization, details regarding the dataset.
b) Details regarding any kind of preprocessing is required or not.
4. Methodology. (2-3 slides)
5. Results/Analysis/conclusion (2 slides). Based upon work done till now.
6. Timeline, are you able to follow the timeline that you proposed in the proposal and proposing future timeline. (1 slide)
7. Individual team member's contribution. (1 slide)

Final Project deliverable:
1. Final Report: Sections for the report, describing your entire work done till now.
   1. Abstract: motivation
   2. Introduction: describing problem statement.
   3. Literature Survey
   4. Dataset: Dataset details with data preprocessing techniques.
   5. Methodology, model details.
   6. Results and analysis.
   7. Conclusion
Maximum page length should be 6. Try to make your report presentable as a research paper. Format, grammatical errors, figures and tables will be thoroughly evaluated.

2. Final Presentation: Max presentation time will be 10min. your learning from the project, work left, and each member's contribution. Please edit the previously submitted progress report.

Maximum page length should be 6. Try to make your report presentable as a research paper. Format, grammatical errors, figures and tables will be thoroughly evaluated.

2. Final Presentation: Please edit the previous slide deck. Max presentation time will be 10min. Can remove the timeline section, rest all the sections should be there.

Few best projects from last year batch:
https://jainendra.in/2021/01/07/best-projects-machine-learning-course-cse343-ece343/