Final Project Description

1 Project Topics

In the CS7150, you will learn about majority of deep learning applications. Part of the learning will be online, during in-class lectures and when completing assignments, but you will really experience hands-on work in your final project. We would like you to choose wisely a project that fits your interests. One that would be both motivating and technically challenging.

- **Application project:** This is by far the most common: Pick an application that interests you, and explore how best to apply learning algorithms to solve it.

- **Algorithmic project:** Pick a problem or family of problems, and develop a new learning algorithm, or a novel variant of an existing algorithm, to solve it.

- **Theoretical project:** Prove some interesting/non-trivial properties of a new or an existing learning algorithm. (This is often quite difficult, and so very few, if any, projects will be purely theoretical.) Some projects will also combine elements of applications and algorithms.

Many fantastic class projects come from students picking either an application area that they're interested in, or picking some subfield of machine learning that they want to explore more. So, pick something that you can get excited and passionate about! Be brave rather than timid, and do feel free to propose ambitious things that you're excited about. (Just be sure to ask us for help if you're uncertain how to best get started.) Alternatively, if you're already working on a research or industry project that deep learning might apply to, then you may already have a great project idea.

2 Hint

A very good CS7150 project will be a publishable or nearly-publishable piece of work. Each year, some number of students continue working on their projects after completing CS7150, submitting their work to a conferences or journals. Thus, for inspiration, you might also look at some recent deep learning research papers. You may also want to look at class projects from previous years online (other machine learning/deep learning classes) is a good way to get ideas. Finally, here is a list of topics can be used:

1. Healthcare: Predicting Disease from Medical Images

2. Finance: Stock Price Prediction

3. Natural Language Processing (NLP): Text Summarization

4. Computer Vision: Object Detection in Drone Footage

5. Autonomous Vehicles: Lane Detection and Tracking

6. Autonomous Vehicles: Lane Detection and Tracking

7. E-commerce: Product Recommendation System

8. Cybersecurity: Anomaly Detection in Network Traffic

9. Social Media: Sentiment Analysis for Brand Monitoring

10. Environmental Science: Weather Prediction

Once you have identified a topic of interest, it can be useful to look up existing research on relevant topics by searching related keywords on an academic search engine such as: http://scholar.google.com. Another important aspect of designing your project is to identify one or several datasets suitable for your topic of interest. If that data needs considerable pre-processing to suit your task, or that you intend to collect the needed data yourself, keep in mind that this is only one part of the expected project work, but can often take considerable time. We still expect a solid methodology and discussion of results, so pace your project accordingly.

3   Project Deliverables

This section contains the detailed instructions for the different parts of your project.

Individual: The project is done individually;

Submission: We will be using Canvas for submission of all four parts of the final project. We'll announce when submissions are open for each part. You should submit on Canvas.

Evaluation: Please refer to syllabus.

The technical quality of the work: Does the technical material make sense? Are the things tried reasonable? Are the proposed algorithms or applications clever and interesting? Do the authors convey novel insight about the problem and/or algorithms?

Significance: Did the authors choose an interesting or a "real" problem to work on, or only a small "toy" problem? Is this work likely to be useful and/or have impact?

The novelty of the work: Is this project applying a common technique to a well-studied problem, or is the problem or method relatively unexplored?

In order to highlight these components, it is important you present a solid discussion regarding the learnings from the development of your method, and summarizing how your work compares to existing approaches.

4 Proposals

Deadline: Refer to Syllabus

In the project proposal, you will pick a project idea to work on early and receive feedback from the instructor. If your proposed project will be done jointly with a different class' project, you should obtain approval from the other instructor and approval from us. Please come to the instructor office hours to discuss if you would like to do a joint project. You should submit your proposals on Canvas.

In the proposal, below your project title, include the project category. The category can be one of:

1. Computer Vision

2. Natural Language Processing

3. Generative Modeling

4. Speech Recognition

5. Reinforcement Learning

6. Healthcare

7. Others (Please Specify!)

Your project proposal should include the following information:

What is the problem that you will be investigating? Why is it interesting? What are the challenges of this project? What dataset are you using? How do you plan to collect it? What method or algorithm are you proposing? If there are existing implementations, will you use them and how? How do you plan to improve or modify such implementations? What reading will you examine to provide context and background? If relevant, what papers do you refer to? How will you evaluate your results? Qualitatively, what kind of results do you expect (e.g. plots or figures)? Quantitatively, what kind of analysis will you use to evaluate and/or compare your results (e.g. what performance metrics or statistical tests)?

5 Milestone

The milestone will help you make sure you're on track, and should describe what you've accomplished so far, and very briefly say what else you plan to do. You should write it as if it's an "early draft" of what will turn into your final project. You can write it as if you're writing the first few pages of your final project report, so that you can re-use most of the milestone text in your final report. Please write the milestone (and final report) keeping in mind that the intended audience is knowing deep learning. Thus, for example,

you should not spend two pages explaining what logistic regression is. Your milestone should include the full name and email, the full title of your project.

Note: We will expect your final writeup to be on the same topic as your milestone. In order to help you the most, we expect you to submit your running code. Your code should contain a baseline model for your application. Along with your baseline model, you are welcome to submit additional parts of your code such as data pre-processing, data augmentation, accuracy matric(s), and/or other models you have tried.

Please clean your code before submitting, comment on it, and cite any resources you used. Please do not submit your dataset. However, you may include a few samples of your data in the report if you wish.

6   Format of Milestone

our milestone should be at most 3 pages, excluding references. Similar to to the proposal, it should include: Title, Author(s) Introduction: this section introduces your project, why it's important or interesting. Make sure to submit your code (or Github URL). Do not submit your dataset. It is okay to include a few samples though.

Details on the dataset Approach: Describe the current steps you have done. If you are implementing an algorithm, you should have started implementation and ideally have some early stage results. Describe precisely the remaining work you expect to complete. We ideally would like to see a model description and a training strategy (loss function for instance).

7   Final Project

The final report should contain a comprehensive account of your project. We expect the report to be thorough, yet concise. Broadly, the final project should be in the format of IEEE conferences. Details will be shared later through Canvas. In general, we will be looking for the following:

A Any hyperparameter and architecture choices that were explored Presentation of results Analysis of results Any insights and discussions relevant to the project References

- Good motivation for the project and an explanation of the problem statement

- Description of the data

- Any hyperparameter and architecture choices that were explored

- Presentation of results

- Analysis of the results

- Any insights and discussions relevant to the project

- References