|  |  |
| --- | --- |
| **CSCI 4366/6366 Neural Networks and Deep Learning**  **Fall 2024**  Section 80 |  |

Semester Project Template[[1]](#footnote-1)

Please make a copy of the template, review and delete the instructions and the examples in gray, and populate each of the fields in regular black font.

**Nickname:** Pick a witty name that describes that problem you’re solving.

**Project Team Members (1-3 members):**

|  |  |
| --- | --- |
| **Name** | **GWID** |
|  |  |
|  |  |
|  |  |

**Document Version History**

|  |  |
| --- | --- |
| **Date** | **Description** |
|  | Initial Draft Submitted |
|  | Instructor Feedback 1 |
|  | Edits and updates from Feedback 1 |
|  | Approval from instructor |
|  | Final Report Submitted |

|  |
| --- |
| **Project Guidelines**  The objective of the DL project is to gain hands-on experience in solving an interesting problem using Machine Learning. You may work individually or as a team (up to five).  The problem should include supervised, semi-supervised, or reinforcement learning methods.  If you are already pursuing research in DL, you may choose a particular aspect related to your research.  You may focus on a specific method or algorithm and evaluate its performance against baseline datasets (i.e., new activation function, architecture, distance measure, optimizer, etc.). Alternatively, you may apply ML methods on a specific real-world application (e.g., real-time trends analysis of streaming media).  At the end of the course, teams will present a 10-minute overview with demo to the class, and submit a final report and source code.  I am happy to address any specific questions during office hours or by appointment. |

# 1 Analysis

Add foundational background information about your project. It should be sufficiently detailed to give the reader understanding of what problem you’re trying to solve and how your approach is different from others who have tried something similar.

## 1.1 Problem Description:

Describe the context of the problem, why it’s important, and how ML might solve this problem.

**Example:**

*During Covid lockdowns and Work From Home (WFH), we observed a 30% increase in the number of chat message, and employees are increasingly struggling with keeping up with conversations across multiple chat rooms. We would like to use text summarization to summarize the dialogs to help users manage their chat conversations.*

***Problem Description****: Can abstractive text summarizers be used for chat conversation summaries?*

## 1.2 Performance Criteria:

State a few top-level project requirements your solution should fulfill (limit to no more than 5 or 6):

*Example:*

* *Summaries should contain the essential facts of the conversations.*
* *Summaries should be grammatically correct and free of offensive language.*
* *Summaries should be generated for chat session with three or more messages that occur in a window of five minutes.*
* *Summaries should be no more than 100 words.*
* *…*

## 1.3 Related Work:

Here, you should describe the domain broadly, and how others have attempted the same or similar problem. Generally, you’ll want to list papers from peer-reviewed conferences (NeurIPS, ICML, ICLR, etc.) or journals (JMLR, AI, etc.). You may use blogs and on-line sources as secondary related work. If you are trying to solve a problem for industry, you may also refer to the work other teams have done.

You’ll want to highlight and define the **Machine Learning Task**.

**Example:**

*Text summarization aims at generating accurate and concise summaries from input document(s). In contrast to extractive summarization which merely copies informative fragments from the input, abstractive summarization may generate novel words. A good abstractive summary covers principal information in the input and is linguistically fluent (Zhang, Zhao, Saleh, & Liu, 2020). …*

## 1.2 Project Objective:

The project objective is a statement of what the project is to achieve, by which success or failure of the project can be evaluated.

***Example:***

*Create a chat summarization model that can provide accurate, and syntactically correct summaries for at least 80% of chat conversations.*

# 2 Hypothesis

This section proposes a solution to achieve the project objective, which describes the method, data, and the experiment for the project.

For the method, describe the ML methods you plan on using and justify briefly why this method will achieve the objective.

For the data description, provide the source of the data, and descriptive information, such as number of examples, number and type of features, and the labels (if supervised). You may either use public data sets, or you may provide your own data. If you use your own data, please confirm in the proposal that the data contains neither confidential (e.g., trade secrets), private (e.g., medical records, SSN, etc.), or proprietary (e.g., limited access). As part of the final delivery, you should include your training and test data sets.

For the experiment, describe how you will evaluate the performance your model has achieved. You should be able to provide quantified results from your experiments.

**Example:**

***Method:*** *We will train and evaluate how well an abstract text summarization transformer model can generate chat summaries. The PEGASUS model…*

***Data:*** *We will use the SAMSum Corpus (*[*https://arxiv.org/src/1911.12237v2/anc*](https://arxiv.org/src/1911.12237v2/anc)*) with over 16k chat dialogues with manually annotated summaries. The data publicly available and free for academic use.*

***Experiment:*** *We will evaluate the summaries generated by the model by computing the ROUGE-2 score (*[*https://en.wikipedia.org/wiki/ROUGE\_(metric)*](https://en.wikipedia.org/wiki/ROUGE_(metric))*) between model’s summary and the test set summary. If we achieve an average ROUGE-2 score of 0.6 or higher, we have achieved the objective.*

*Additionally, we will perform manual evaluation of the model with no fewer than 200 chat summaries. At least 10 participants, not directly involved with the project, will be presented the chat dialog and the model’s summary, and will be asked to grade the summary on a five-point scale from* ***completely wrong, slightly wrong, irrelevant, somewhat accurate, to completely accurate.*** *We will present aggregate results as the distribution of counts and percentage on for each bin. Our goal is to achieve at least 60% in the* ***somewhat accurate*** *and* ***completely accurate*** *bins.*

# 3 Synthesis

Discuss the most important software, ML libraries, and other open-source tools that you plan on using in the project. How will you tune hyperparameters, etc.

Subdivide the project into phases and/or subcomponents and estimate the number of days of effort required to accomplish the task.

While its not required, I strongly encourage including a project work plan with tasks and expected dates to help track your project.

# 4 Validation

## 4.1 Results:

Summarize how you will compute the raw results for each experiment and describe visualizations and tables you intend to include in the final report.

You should also demonstrate learning as loss function curves (e.g., Tensorboard) that converge over many training iterations.

## 4.2 Conclusions:

**Formal Conclusions**: Formal conclusions determine whether your project objective was achieved and require statistical measures to quantify uncertainty and confidence. Propose and justify a statistical test/method (e.g., Mann-Whitney, Kolmogorov-Smirnov, etc.) to determine if the original research hypotheses passes or fails. For formal conclusions, you should provide confidence statistics (i.e., p-values) from your statistical tests.

**Informal Observations**: Include additional conclusions relevant to your work that may be based on rigorous analysis or may be insightful and speculative. There is no need to include any remarks about informal observations in the proposal, but please include them in the final report.

**Future Work**: Not added in the proposal, but in the final report you should include a discussion about what promising avenues your work may lead to, even if they are not immediately aligned with your original objectives. If you didn’t achieve the original objective, address what might be done with more time and effort to achieve the objective.

1. Template derived from Bock, P. (2001) Getting it Right: R&D Methods for Science and Engineering, Academic Press [↑](#footnote-ref-1)