Regular Project

Abstract

We will follow the proposal's grading method, and the grade will be based on the original quality of the proposal and potential improvements.

- Project description
- Importance of project
- Argued about novelty of their project
- Discussed potential findings

Introduction

We will follow the proposal's grading method, and the grade will be based on the original quality of the proposal and potential improvements.

- Started with general trend in NLP/high-level related works
- Scope down narrowly to what will be proposed/work on
- Point major limitations of existing works
- State clearly their plan

Related Works

We will follow the proposal's grading method, and the grade will be based on the original quality of the proposal and potential improvements.

- Identify key papers relevant to work
- Talk about them in detailed and crisp manner
- Exhaustive (no missing paper)
- Organized based on aspects of the problem (see guidelines)
- Avoid false claims about cited papers
- Objective look at other papers (do not make them look bad)

Modeling

- Describe the models you implemented

- State a hypothesis and use the models to prove it
- Contains two subsections: Baselines, proposed model
- Baseline(s): Each baseline follow the following criteria
    - Describe the baseline
    - Explain how it works
    - Provide mathematical details (when relevant)
    - Use figures when appropriate
    - Mentions if it was reimplemented or used existing code
    - Provide a fair comparison by using the same resources
- Proposed model
    - Describes the proposed model
    - Use figures when appropriate
    - Explain the improvement over baseline(s)

Dataset and evaluation

- Contains two subsections: Data and Evaluation Metric

For data:

- Cite and give name of the datasets used
- Explain why the chosen datasets are appropriate
- If you have new datasets since proposal, update and mention them explicitly
- Describe dataset statistics:
    - Train/validation/test split
    - Average length of inputs/outputs
    - Number of class labels
    - Distribution of class labels
- Meaningful analysis of dataset

For metric:

- Cite, and give full name of metric
- Explain why the chosen datasets are for the dataset and task
- Provide intuition about what the metric is computing
- If it's uncommon, define it (mathematically or verbally, whichever is more appropriate)

Experiments

Describe experimental details:

- Hyperparameters

- Number of training epochs
- Number of runs (e.g. 3+)
- Software details (e.g. Pytorch, huggingface, Scikit-learn)
- Relevant implementation details (e.g. chose a batch size due to memory limit, used half precision for faster training)
- link to code repository - add @xhluca and @ncmeade if it's private)

Results and discussions

Results:

- Describe results of baselines
- Describe results of preliminary models
- Create a table of results per the guidelines

Discussion:

- Is your model better than the baseline?
- Does it confirm your hypothesis?
- If numbers are lower, describe why it's the case

Analysis:

- Show examples of predictions
- Qualitative analysis of examples
- Discuss cases where baseline model fails but main model succeed
- Discuss cases where baseline models succeed but main model fails

Conclusion

- Discuss main takeaways
- Summarize what you did
- Summarize the findings
- Take about future works

Code

- Release usable code
- Avoid instructions that only work on specific platforms (e.g. SLURM)
- Provide self-contained instructions that are straightforward to follow
- Release code on GitHub (if it’s private, invite @xhluca or @ncmeade)

- Include instructions in Readme, or link to separate markdown file with full instructions
- Indicate what the TA should be reading

Quality of Writing

- Clear and concise writing.
- Coherent sentences that follow a natural and logical flow
- Define important technical terms clearly such that any student in the class can understand
- If you have abbreviation, define it the first time you mention it
- Avoid repeating the same statement/idea multiple time when unnecessary
- Writing grammatically correct sentences
- Avoid typo/missing words and phrases that are difficult to read
- If you have text in different language (e.g. in a figure, table), translate it
- If you have math equations, define them clearly right before or after the equation.
- Each figure/table should be referred to from the text

Formatting

- When citing, the link should be clickable and link to the bibliography
- Avoid including pre-print citations when the paper has a published version available (e.g. don't cite "ArXiv" if it's available on ACL anthology)
- Use \citet for textual citations, \citep for parenthetical citations. For example, "Devlin et al. (2018) proposed a..." vs "We used the BERT model (Devlin et al. (2018)"
- Math equations should be correctly formatted
- Figures should be clearly readable (e.g. use PDF, SVG instead of PNG when possible, otherwise use high DPI)
- Tables should not be images
- You should have clickable links to figures and tables

Contributions

- Listed the contributions of the members

Reproducibility Project

Reproducibility Summary

- Scope of reproducibility: State main claim of reproduced paper
- Scope of reproducibility: Place work in context to tell reader the objective
- Methodology: Describe what you did
- Results: Describe overall conclusions
- Results: Use specific and precise language
- Results: Used judgment to decide if results support original claim
- What was easy: Summarize which part of original paper are easy to apply
- What was difficult: Indicate which parts of original paper are difficult to re-use
- What was difficult: Indicate which parts of original paper requirement significant work/resources to verify
- Communication with original authors: Describe concisely contact with original authors

Introduction

- Introduce and motivate the problem and discuss its importance.
- At a high-level, describe the contributions of the original paper and their significance.
- In addition, at a high-level, describe the additional experiments or analysis that were performed. Briefly describe the conclusions drawn.

Scope of reproducibility

- State the main claims of the paper they are reproducing.
- State whether they are using the original codebase or not.
- State, specifically, which results from the original paper they are replicating.

Methodology

- Describe the methods you used for reproducing the results of the original paper.
- Describe the computational resources used to reproduce the results.
- Describe if the authors' original code was used. If only parts of the authors' code was used, describe which parts.
- State the compute budget required for reproduction (e.g., how many GPU hours used).

Model descriptions

- Provide a brief description of each model and algorithm used in reproduction.
    - Relevant citations should be included.
- For each model used, state the following information:
    - How large is it (i.e., how many parameters)?
    - What data was the model trained on?
    - Was the model pre-trained?
    - Are there any major limitations?

Datasets

- For each dataset used, include the dataset statistics:
    - How many examples?
    - Average example length (e.g., number of tokens).
    - Distribution of labels (in the case of a classification dataset).
- For each dataset used, state the train/validation/test split sizes. If the dataset does not have a training split, mention this.
- Include a link or reference for each dataset used.
- At a high-level, describe how each dataset was collected (e.g., was it crowdsourced?).

Hyperparameters

- Describe the settings of the hyperparameters used (e.g., we used a learning rate of 2e-5).
- If a hyperparameter search was conducted, describe how it was done:
    - How did you search the hyperparameter space? Linear search? Random search? Grid search?
    - State how many repetitions (i.e., seeds) were used to select the best hyperparameter configuration.
    - State what evaluation metric was used to select the best hyperparameter configuration (e.g., we determined the best hyperparameter configuration by computing the validation accuracy).

Experimental setup and code

- Describe how experiments were set up
- Written clearly enough for readers to replicate
- Include description of evaluation measures
- Provide link to code (see code requirements in section “Code”)

Computational requirements

- Description of hardware used (GPU, CPU, other)

- For each model, include a measure of the average runtime
- For each experiment, include the total computational requirements
- Considered the perspective of a reader that wants to use the approach reproduced
- List what a reader would find useful

Results

- High-level overview of your results
- Discuss objectively if results support the main claims of the original paper
- Section is factual and precise
- Did not include judgment and discussion (which should be in the “Discussions” section)

Results reproducing original paper

- Had sufficient number of experiments with respect to the original paper
- For each experiment:
    - Indicate which claim in Section 2 it supports
    - Indicate if it successfully reproduced the associated experiment in the original paper
- Logically group related results into sections

Results beyond original paper

- Indicate whether the original paper failed to fully specify some experiment(s)
- Indicate if additional experimentation was necessary
- Include results of any additional experiments (only if it’s necessary for this reproduction)

Discussion

- Judge if experimental results support the claims of the paper
- Discuss strengths of your approach
- Discuss weaknesses of your approach
- Indicate if you did not have time to run all experiments and why
- Explain how additional experiments further strengthen the claims in the paper

What was easy

- Judge what was easy to reproduce
    - Author’s code is clearly written
    - Author’s code was easy to run
- Avoid giving sweeping generalizations
- Explain why something was easy (e.g. code had good API and lots of examples)

What was difficult

- List part of the study that:
    - took more time than anticipated
    - felt were difficult
- Put discussion in context (e.g., avoid saying “math was difficult to follow”, instead “math require advanced knowledge of X to follow”)

Communication with original authors

- Documented the extend of communication with the original authors
    - If they responded, summarize their responses
    - If they did not, indicate which other means you used to reach out (email, social media, GitHub issues)
- Either:
    - List specific questions that were asked, or
    - Sent full report to get their feedback

Code

- Release usable code
- Avoid instructions that only work on specific platforms (e.g. SLURM)
- Provide self-contained instructions that are straightforward to follow
- Release code on GitHub (if it’s private, invite @xhluca or @ncmeade)
- Include instructions in Readme, or link to separate markdown file with full instructions
- Indicate what the TA should be reading
- Wrote the code in Python
- Released a PyPi package with instructions (if it’s not relevant, explain why in the GitHub readme, in a section called “Notes about release”)
    - For others to run the experiments
    - To import and reuse parts of the code as a standalone library
    - Can be installed via `pip` from the official PyPi

Quality of Writing

- Clear and concise writing.
- Coherent sentences that follow a natural and logical flow
- Define important technical terms clearly such that any student in the class can understand
- If you have abbreviation, define it the first time you mention it
- Avoid repeating the same statement/idea multiple time when unnecessary

- Writing grammatically correct sentences
- Avoid typo/missing words and phrases that are difficult to read
- If you have text in different language (e.g. in a figure, table), translate it
- If you have math equations, define them clearly right before or after the equation.
- Each figure/table should be referred to from the text

Formatting

- When citing, the link should be clickable and link to the bibliography
- Avoid including pre-print citations when the paper has a published version available (e.g. don't cite "ArXiv" if it's available on ACL anthology)
- Use \citet for textual citations, \citep for parenthetical citations. For example, "Devlin et al. (2018) proposed a..." vs "We used the BERT model (Devlin et al. (2018)"
- Math equations should be correctly formatted
- Figures should be clearly readable (e.g. use PDF, SVG instead of PNG when possible, otherwise use high DPI)
- Tables should not be images
- You should have clickable links to figures and tables

Contributions

- Listed the contributions of the members