

Your project title here

First Author
email@domain

Second Author
email@domain

Third Author
email@domain

Fourth Author
email@domain

1 Problem statement

Summarize the goal of your project and its motivations in this section.

Some general guidelines for the final report: MINIMUM of 4 pages in this format (not including references). Upload your codes into a public github repository. It should have a detailed README and be reasonable clean and commented. If you're submitting a Github link, make sure it goes to a *public* repository. This report should obviously be more detailed than your proposal report. It should contain multiple tables and/or plots that convey numerical information – for example, statistics about your data or annotations, accuracy or other results of running an algorithm on the data, or something else. Please use this LaTeX template to write your report. Note that some sections may not be relevant for your project; feel free to delete them and add other sections if you wish! If you're unsure whether to include something or not, please contact the instructors.

2 What you proposed vs. what you accomplished

Provide a short list of things you proposed to do (in the project proposal) and whether you actually completed these things or not. If you failed to do something, *briefly* explain why! If you made significant changes to your project since the proposal, also note that here. You can go into more detail later in the report. This section is more of a summary and should be no longer than a column, as in the example list below:

- ~~Collect and preprocess dataset~~
- ~~Build and train (specific baseline model) on collected dataset and examine its performance~~

- ~~Build and train fancy model and examine its performance~~
- *Make fancy model perform better than baseline model:* We failed to do this because we only started working on our project on May 11
- ~~Perform in-depth error analysis to figure out what kinds of examples our approach struggles with~~

3 Related work

Provide a survey of prior work that is related to your project. **Make sure to cite AT LEAST 10 papers in this section.** You don't have to describe all of them in detail, but for a couple of the most related papers, briefly describe what they did and how/why your approach differs. For a good example of a related work section, see Section 3 of <https://arxiv.org/pdf/1804.08199.pdf>, which tells a story about the history of semantic role labeling and how particular papers relate to the proposed work. Your related work section should not read like a list of paper descriptions (if it does, you will lose points) but rather form a cohesive section. To look for relevant papers, check out the top NLP conferences (e.g., ACL, EMNLP, NAACL, TACL). Make sure to properly cite them. You can cite a paper parenthetically like this (Andrew and Gao, 2007) or use the citation as a proper noun, as in “Borschinger and Johnson (2011) show that...” If you're not familiar with LaTeX, you'll have to add entries to yourbib.bib to get them to show up when you cite them. Your related work section should be no longer than 1.5 columns in this format!

4 Your dataset

The most important rule of NLP: look at your data! Provide us with examples from your

dataset, and describe your task in a coherent manner. Explain what properties of the data make your task challenging. Report the source of the dataset, its basic statistics (e.g., size, number of words/sentences/documents) and some other statistics that are specifically relevant to your task. Show a couple input / output pairs to make it clear what you're doing (but don't use up too much space in doing so!).

4.1 Data preprocessing

If you did any preprocessing, explain what you did (and why) here!

4.2 Data annotation

If your project involves annotation, you may have started a pilot annotation experiment, annotating a few dozen or few hundred examples. What major issues have come up? Do you and your project partners agree or disagree on examples? Report interannotator agreement if applicable.

5 Baselines

What are your baselines, how do they work, and what are their results? Why did you choose these baselines over other models? Additionally, explain how each one works, and list the hyperparameters you are using and how you tuned them! Describe your train/validation/test split. If you have tuned any hyperparameters on your test set, expect a major point deduction!

6 Your approach

What is your approach and how does it work? Do you expect it to fail in similar ways to your baselines? Did you manage to complete a working implementation? What libraries did you use to accomplish this? Did you rely on help from any existing implementations? If so, please link to them here. **What models did you implement yourself, and what files in your uploaded code are associated with these models?** What kind of computers are you running your experiments on? Are there any issues that you could not solve? If you used Colab, were there any Colab-specific hacks you needed to make to train your model? What results did your model achieve, and how do these results compare to your baselines? Be specific!!! Note that there could be many other important details specific to your approach that you should include here if appropriate.

7 Error analysis

What kinds of inputs do your baselines fail at? What about your approach? Are there any semantic or syntactic commonalities between these difficult examples? **We would like to see a manual error analysis (e.g., annotate 100-200 failed examples for various properties, and then discuss the results and hypothesize about why the).**

8 Contributions of group members

List what each member of the group contributed to this project here. For example:

- member 1: did data collection / processing and lots of writing
- member 2: built and trained models
- member 3: error analysis and annotations

If you would like to privately share more information about the workload division that may have caused extenuating circumstances (e.g., a member of the group was unreachable and did no work), please send a detailed note to the instructors GMail account. We will take these notes into account when assigning individual grades.

9 Conclusion

You've now gotten your hands dirty with NLP tools and techniques! What takeaways do you have about your project? What proved surprisingly difficult to accomplish? Were you surprised by your results? If you could continue working on your project in the future, what directions would you pursue?

References

- Andrew, G. and Gao, J. (2007). Scalable training of L1-regularized log-linear models. In *Proceedings of the 24th International Conference on Machine Learning*, pages 33–40.
- Borschinger, B. and Johnson, M. (2011). A particle filter algorithm for Bayesian wordsegmentation. In *Proceedings of the Australasian Language Technology Association Workshop 2011*, pages 10–18, Canberra, Australia.