

# CS224u experiment protocol

**Lara Thompson**

Principle Data Scientist @ Salesforce  
lara.thompson@salesforce.com

March 7, 2023

## Notes

This is a short, structured report designed to help you establish your core experimental framework. The required sections are as follows.

You are free to include additional sections as well as appendices. However, keep in mind that experiment protocols are not intended to be long documents. The goal is to help you establish for yourself what your full project will look like and expose any blockers that might prevent your project from succeeding.

The protocol document is also very useful as a basis for discussion with your project mentor concerning scope, missing pieces, and areas of concern relating to time, resources, concepts, and anything else that might come up.

## 1 Hypotheses

A statement of the project's core hypothesis or hypotheses.

These hypotheses vary widely and can be oriented towards issues not just in NLP but also in machine learning, social sciences, digital humanities, medicine, and any other field that can benefit from linguistic analysis. One might identify a particular component of the new model and hypothesize that it is crucial for success.

## 2 Data

A description of the dataset(s) that the project will use for evaluation.

Your mentor will be looking closely at this section, seeking to resolve a few crucial questions:

1. Is it clear which datasets will be used? If you just say general things like "sentiment datasets", you're likely to get some points off along with a request for the name of a specific dataset that can support your goals.

2. Is the dataset available? If access to the dataset is highly restricted, or the dataset does not

yet exist, then we will likely express concern. The course is short and the final deadline is near, so the quality of your work will suffer if you spend a lot of time waiting for data. We are likely to push you towards public datasets so that you can get started right away.

3. Is the dataset aligned with the hypotheses? If this isn't immediately evident to your mentor based on what you say, you're likely to get some points off along with a request to meet. It's therefore worth your time to make this connection in the protocol document.

## 3 Metrics

A description of the metrics that will form the basis for evaluation. We require at least one of these to be quantitative metrics, but we are very open-minded about which ones you choose. In requiring this, we are not saying that all work in NLU needs to be evaluated quantitatively, but rather just that we think it is a healthy requirement for our course.

For the quantitative evaluations, your mentor will primarily be focused on determining whether the metrics are appropriate given the data and hypotheses. If you chose something standard – e.g., F1 for a classification problem – then you probably don't need to say much. If you depart from what's standard, or you want to propose your own metric, then you'll need to justify these decisions.

## 4 Models

A description of the models that you'll be using as baselines, and a preliminary description of the model or models that will be the focus of your investigation.

At this early stage, some aspects of these models might not yet be worked out, so preliminary descriptions are fine. Your focus should be on making it clear how the models interact with your dataset and metrics to provide a clear test of your hypothesis. If your mentor can't put the pieces together in

this way, you'll get some points off and a request for a more precise description.

## **5 General Reasoning**

An explanation of how the data and models come together to inform your core hypothesis or hypotheses.

You might feel that you've already given this explanation in the above sections. That's good! It suggests that the ideas are coming together. Please restate the general reasoning in this separate section.

## **6 Summary of Progress**

What you have been done, what you still need to do, and any obstacles or concerns that might prevent your project from coming to fruition.

The more precise you are here, the more your mentor can help you. One of the most productive things they can do at this stage is help you define the scope of your project and then set priorities on that basis. For example, they might identify a core set of experiments that need to be run and suggest that you set the others aside until you have drafted your entire paper.

It's also very helpful to identify potential obstacles, so that your mentor can help you overcome them or find a strategy for avoiding them entirely.