

Project

Project assignment

- You will need to complete a project.
 - Groups of 2-3 are recommended
- Scope: roughly a homework assignment X size of group
 - I.e., big groups need to have a bigger scope.
- Submissions:
 - **Proposal**: by April 10
 - **Project**: by final exam date

Project assignment

- Key idea: identify a relevant **research question**, formulate and operationalize it, report results and discuss their meaning.
- The research question should be written in a way that would allow you to evaluate it (e.g., trying out different classifiers is not a research question)
- The technical plan should support this process. **Make sure you can draw a clear line between the question and your proposal (how the problem is modeled)**
- The experiments should allow you to answer the question, or at least discuss the answers. **Make sure you have relevant datasets.**
- **Evaluation** should consist of both quantitative and qualitative analysis.

Project assignment: WSD Example

- *Word sense disambiguation requires mapping words in a sentence to their sense. E.g., “**Apple** stock rose 5%” vs. “**Apple** sales rose 5%”.*
- *LSTMs can create word representation capturing the context in which words appear.*
- ***RQ:** can LSTM models capture the right word sense **without** direct supervision just based on context?*
- **Operationalize:**
 - *Identify a set of K ambiguous words.*
 - *Use a pre-trained LSTM based language model to generate a representation of these word instances in a large corpus.*
 - *Cluster the representations using K-means*
- **Evaluate**
 - *Report cluster purity.*
 - *Compare the pairwise agreement correctness to existing datasets*
 - *Qualitative: show most representative examples from each cluster (sentences closest to the centroid). Do they correspond to the same sense? Do clusters capture topic of sentences or WSD assignments?*

Project assignment: WSD Example

This example is appropriate for a group of 2 people. How should it be extended if we had 3 people?

- Essentially, **another research question!**
 - For example - given the same settings, how would LSTM compare with a different model?
 - What is captured by summing up W2Vec word vectors instead of an LSTM?
 - What is captured by using modern architectures, such as BERT, instead of an LSTM?
- The added work should be proportional to the number of people in your group.

Project proposal submission

- Short document explaining what is the research question, how it will be explored and making sure that you have the right data/tools for it.
 - **Short no more than 3 pages!**
- **General guidelines:**
 - When formulating the problem and the technical solution, be as specific and as concrete as you possibly can.
 - What are the inputs and outputs? What kind of assumptions are you making?
 - Point to specific datasets that you intend to use. Explain why they are a good match for your problem.

Project Proposal Submission

- Project Proposals should have the following sections
 - **Problem formulation and Motivation** *What is the problem you are working on? Why?*
 - **Background** (Literature review): what else has been done in this space? Look up at least 2-3 related papers and read their abstracts to get an idea.
 - **Research Question** definition. Be as formal as possible.
 - **Technical approach.** What models will you use? Explain why that's the right choice for the problem.
 - **Experimental evaluation.** Identify relevant datasets you will use and which experiments you will run (i.e., evaluation plan).
 - **Work plan** (who are the team members, how is the work distributed).
- **No more than 3 pages.**

Project ideas: Representation learning

- Take a problem that we worked on (e.g., sentiment classification) and use different techniques to improve a reasonable basic classifier that will be used as baseline (e.g., Bi-directional LSTM or BERT based model).
- Approaches can include:
 - **Multi-task learning:** train the same representation with an additional task (jointly or as pre-training). **Key question:** what are related tasks that can be trained jointly?
 - Paper reference: [A unified architecture for natural language processing: Deep neural networks with multitask learning, Collobert, Ronan and Weston, Jason](#). ICML '08
 - **Diagnose** an existing model by adding noise to the test data in a controlled way. For example, add to a positive review negative sentiment text that refers to a competing product. Check how the model deals with negations.
 - Paper reference: [Right for the Wrong Reasons: Diagnosing Syntactic Heuristics in Natural Language Inference Tom McCoy, Ellie Pavlick, Tal Linzen](#)

Project ideas: Representation learning

- Take an existing (already trained) language model and evaluate its properties. **Can be viewed as a question answering problem.**
 - “A chair has ____ legs” , candidate answers “1” , “2, “3” ,”4”
- Key question – how to come up with interesting questions?
 - Common sense (which is bigger a toaster or a house)
 - Factoid (WW2 ended in.. 1945? 1954? 1968?)
 - Evaluate long range dependencies (“the horse, that ran close to the red car, started ____”, “driving”, “galloping”)
- **Paper references:** [Do Language Models Learn Commonsense Knowledge?](https://arxiv.org/pdf/2111.00607.pdf)
<https://arxiv.org/pdf/2111.00607.pdf>
- “What you can cram into a single \$&!#* vector:” <https://aclanthology.org/P18-1198.pdf>

Project ideas: applications

- **Detecting bias** in news articles (e.g, given two reports of the same news event, predict which one has a left/right bias).
- Analyze the **language used by different social communities**. For example, what are the associations that the word “sick” has in different reddit communities.
- **Identifying emotions** in Covid related tweets and correlating them with real world events (e.g. “did the vaccine roll out reduce fear?”)
- **Predict Emoji** based on text

Papers references

Linguistic harbingers of betrayal: A case study on an online strategy game. Niculae, Vlad, Srijan Kumar, Jordan Boyd-Graber, and Cristian Danescu-Niculescu-Mizil.

More than words: Syntactic packaging and implicit sentiment. Greene, Stephan and Philip Resnik. 2009.

Argument Invention from First Principles Yonatan Bilu, Ariel Gera, Daniel Hershcovich, Benjamin Sznajder, Dan Lahav, Guy Moshkovich, Anael Malet, Assaf Gavron, Noam Slonim

Danescu-Niculescu-Mizil, Cristian, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No country for old members: User lifecycle and linguistic change in online communities.

Feuding Families and Former Friends: Unsupervised Learning for Dynamic Fictional Relationships Mohit Iyyer, Anupam Guha, Snigdha Chaturvedi, Jordan Boyd-Graber, Hal Daumé