LIGN 167
Deep Learning for Natural Language Understanding

Fall 2023

Instructor: Prof. Leon Bergen (lbergen@ucsd.edu)

Teaching Assistant: Nico Tedeschi (njtedeschi@ucsd.edu)

Teaching Assistant: Xingyuan Zhao (xiz103@ucsd.edu)

Instructional Assistant: Xiaoyue Wang (xiw027@ucsd.edu)

Instructional Assistant: Gurman Dhaliwal (gdhaliwa@ucsd.edu)

Overview

An introduction to neural network methods for analyzing linguistic data. Basic neural network architectures and optimization through backpropagation and stochastic gradient descent. Attention, transformers, masked/autoregressive language modeling, and encoder-decoder architectures. Large language models and their future.

Prerequisites

Multivariable calculus (MATH 10C or MATH 20C or MATH 31BH or equivalent). Some background with programming is strongly recommended, as well as some linear algebra.

GPT-4

You are required to purchase a subscription to ChatGPT Plus ($20 per month, for two months). We will be using GPT-4 (only accessible through ChatGPT Plus) extensively in class.

Readings

Optional textbooks:

- Goldberg, Y. (2017). Neural Network Methods for Natural Language Processing. Morgan & Claypool.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning. MIT Press.

- Jurafsky, D., and Martin, J. (2023). Speech and Language Processing.

These are the standard textbooks in the field, and have another important advantage: they are available for **free**.

The books can be accessed at the following links:

- Goldberg (on-campus/VPN access only): `https://link.springer.com/book/10.1007/978-3-031-02165-7`
- Goodfellow et al.: `http://www.deeplearningbook.org`
- Jurafsky and Martin: `https://web.stanford.edu/~jurafsky/slp3/`

Grading

Problem sets

There will be 5 problem sets distributed on Piazza throughout the course. These will be coding assignments in Python/PyTorch, and will be worth 70% of your grade.

Collaboration policy

You may collaborate with up to three other students on each problem set. You should submit **as a group** on Gradescope (one submission per group). You should indicate the contributions of each group member.

Regrading policy

If you would like to request a regrade, you must do so within 3 days of the return of the problem set. Regrades should be requested using Gradescope.

Lateness policy

Late assignments will receive 20% off per day. If you have a valid reason for late submission (e.g. medical), then you must submit documentation (e.g. a doctor's note) to me.

Final project

There will also be a final project, worth 30% of your grade. Projects will involve a final presentation and write-up, and can be completed in groups of up to 4 students.

Your group should submit a project proposal, consisting of a 1-page writeup of what you plan to do, and how you plan to do it.

Integrity

We adhere to the UCSD Policy on Integrity of Scholarship, so please consult it: https://academicintegrity.ucsd.edu/process/policy.html

Note that per UCSD policy, we are obligated to report instances of suspected academic dishonesty to the Academic Integrity Office.

Any students who require OSD accommodations should meet with me during the first week of class to discuss arrangements.

Provisional Schedule

- September 28: Introduction
- October 3: Supervised learning. Linear regression. Pset 1 released.
- October 5: Linear regression continued.
  Reading: Goldberg, Sections 2.1-2.6
- October 10: Logistic Regression.
  Reading: Goldberg, Sections 2.1-2.6
- October 12: Optimization through gradient descent. Pset 1 due.
  Reading: Goldberg, Sections 2.7-2.8
- October 17: Gradient descent continued. Multilayer perceptrons.
  Reading: Goldberg, Sections 3.1-3.4
- October 19: Multilayer perceptrons continued. Pset 2 released.
  Reading: Goldberg, Sections 4.1-4.5
- October 24: Multilayer perceptrons continued.
  Reading: Goldberg, Intro to Section 5, Sections 5.1.1-5.1.2
- October 26: Backpropagation. Pset 2 due.
  PyTorch 60 minute blitz: `https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html#`
- October 31: Backpropagation continued. Pset 3 released.
  Project proposal due
- November 2: Word2Vec. Reading: Goldberg, Sections 8.1-8.3
- November 7: Autoregressive models
  Reading: Jurafsky and Martin, Sections 9.1-9.2
- November 9: Attention. Pset 3 due.
  Reading: Jurafsky and Martin, Section 10.1

- November 14: Introduction to transformers. Pset 4 released.
  Reading: Jurafsky and Martin, Section 10.2
- November 16: Transformers continued.
- November 21: Autoregressive language modeling. Pset 4 due.
  Reading: Jurafsky and Martin, Section 11.1, 11.2
- November 23: Thanksgiving.
- November 28: Masked language modeling. Pset 5 released
- November 30: Encoder-decoder architectures.
- December 5: The future of language modeling. Pset 5 due
- December 7: Presentations