CS542 (Fall 2023) Initial Self-Assessment

Due August 27, 2023

Answer all the questions below to the best of your understanding. Do not use the internet to look up answers. "I don't know" is a perfectly acceptable answer. This homework is *not* graded but will help tailor the class to your level as much as possible.

Your answers must be legible. Do not put your name anywhere on this assignment. You may print and write on this PDF and scan it, or you may answer the questions in a separate text file and submit that.

1. What is the relationship between regular expressions and finite state automata?

2. What is an *n-gram*?

3. Write down Bayes' Theorem.

4. Define precision, recall, and F-score.

5. What is a “gradient” in machine learning?

6. What is hidden in a Hidden Markov Model?

7. What “type” is a *context-free grammar* in the Chomsky hierarchy?

8. What does it mean to “parse” a sentence?

9. Who said “You shall know a word by the company it keeps?”

   (a) (Bonus for bragging rights) Who is frequently (and incorrectly) credited with the above quote?

10. Define “neural network.”

11. What is a convolution in a CNN?

12. What is the vanishing gradient problem?

13. What is the curse of dimensionality?

14. What is an "embedding"?

15. What is an "encoder"? What is a "decoder"? (in the context of neural networks/Transformers)

16. What is an "autoregressive" model?

17. According to the developers of the Transformer architecture, what is all you need?

Submission Instructions

Please submit your answers (in PDF format - printed and scanned images are OK) to the drop box on Canvas.