

University of Pittsburgh
CS 2731
Introduction to Natural Language Processing
Midterm

Spring 2023

This exam has 4 questions for a total of 100 points.

Name: _____

Student ID: _____

Instructions:

- You will have 90 minutes to answer the questions.
- The exam is closed-book, closed-slides, and closed internet.
- Write clearly and legibly. Please show your steps, and make sure you are not writing just a number. Please write brief 2-5 sentence paragraphs to open-ended questions.
- You do not need your laptop or calculator for the exam.
- Please silence your phones and put them away.
- Good luck!

No	Topic	Max Score	Your Score
1	Machine Learning Evaluation Basics: Precision, Recall & F-Score	25	
2	AI Ethics: Bias	25	
3	Language Modeling	25	
4	Discourse	25	
Total Score			

[PAGE INTENTIONALLY LEFT BLANK]

Mert Inan merterm.github.io

1) [25 Points] Machine Learning Evaluation Basics: Precision, Recall & F-Score

In this task, politician names are automatically extracted from a collection of documents and classified with a label (left, right, or center) or are unclassified. A correct instance is one in which the system output and answer key agree on a label. If no label is assigned in the answer key or the system output, it will not be listed below or considered in the score.

Example	System Output	Answer Key
Barack Obama	center	left
Bill Clinton	center	left
Bugs Bunny	left	
George W. Bush	right	right
Hillary Clinton	center	left
Howard Stern		center
Jason Brown	center	
John Major	right	center
Jonathan Swift	left	
Karl Marx	left	
Mitt Romney	right	center
Noam Chomsky	left	left
Pippi Longstocking	center	
Ralph Nader	left	left
Richard Cheney	right	right
Rush Limbaugh		right
Sarah Palin	right	right
Shaquille O'Neal	center	

- a) Calculate precision, recall, and f-score (f-measure) for the given system output and answer key.

2) [25 points] AI Ethics: Bias

- a) Describe the bias network effect or bias laundering. (10 pts)

- b) Name the stages that bias might be introduced to a system. (10 pts)

- c) Give an example of bias laundering (5 pts)

Mert Inan merterm.github.io

3) [25 points] Language Modeling

- a) What is a language model? (5 pts)

- b) What are language models used for? (5 pts)

- c) Name three applications. (15 pt)

Mert Inan merterm.github.io

ent discourse frameworks. (10 pts)

- ent discourse frameworks. (10 pts)