

# CSE708: Application of Data Analytics and Engineering

## Fall 2023

### Midterm Exam

Due on Tuesday 10/03/2023 at 9:59AM

1. **(50pts)** Using cosine similarity compute the similarity between these two sentences sentence1 and sentence2:

**Sentence 1:** “the cat sat on the mat.”

**Sentence 2:** “the cat slept on the bed.”

Note: Please, this is a hand computing question. Don’t use any programming language.

2. **(30 pts)** Use one of techniques we have learned in CSE708 to normalize the following data. Please indicate what technique you chose.

Age= [12, 16,10, 20, 15]

3. Given the following dataset.

Coffee	Day	Temperature	Course	Observation
No	Monday	Hot	CSE708	Sleepy
Yes	Tuesday	Hot	CSE708	Sleepy
Yes	Thursday	Hot	CSE708	Sleepy
No	Monday	Cold	CSE708	No sleepy
Yes	Friday	Cold	CSE708	No Sleepy

- a. **(10 pts)** Today is Wednesday. George forgot to drink his coffee and it is cold outside, what do you think will be the observation about him during CSE708 class? Explain your opinion in one sentence.
- b. **(10 pts)** How about Marie? It was Friday, she drank her coffee, and it was cold outside. Explain in one or two words.