# CS 4410: Mid Exam FA2024

***Exam Time: 60 Minutes*** **Total Marks: 30**

**Must write your section A/B, Troy ID, and Name on every sheet.**

---

Q1 [10 points] Briefly define and differentiate between the following (if possible with an example):

A) Bias and Variance B) Underfitting and Overfitting C) Validation Set and Test Set D) Regression and Classification

Q2 [5 points] Provide at least three different measures of classification. Differentiate with examples where a particular measure would make more sense than the others.

Q3 [5 points] Past data tells us that 10% of patients entering the clinic have liver disease. Five percent of the clinic's patients are alcoholics. Among those patients diagnosed with liver disease, 7% are alcoholics. The probability that a patient is alcoholic, given that they have liver disease, is 7%. What is the chance of having liver disease if the patient is an alcoholic?

Q4 [5 points] Compute the Maximum Likelihood Estimate (MLE) of the mean ($\mu$) and variance ($\sigma^2$), assuming $X$ (the data points) is Gaussian Distributed. Show your steps.

Q5 [5 points] Observe the following transaction data. Is there any frequent itemset of size 3 with a minimum support of 15%? What rule(s), if any, can you derive with a minimum confidence of 60%? Explain your answer.

| **Transaction** | **Items** |
|---|---|
| 1 | milk, bananas, chocolate |
| 2 | milk, chocolate |
| 3 | milk, bananas |
| 4 | chocolate |
| 5 | chocolate |
| 6 | milk, chocolate |
| 7 | milk, bananas, chocolate |