CSC 308 Tutorial Questions 2024/2025 Session

Question 1

(a) Briefly explain what Machine Learning is 5 marks

(b) Explain the relation between hypothesis function and target function 10 marks

Question 2

(a) Briefly explain the following: 4 marks = 12 marks

i. Supervised learning ii. Unsupervised learning iii. Semi supervised learning

(b) Explain why datasets are usually split into training and test sub datasets in Machine Leaning 3 marks

Question 3

(a) Data cleaning and feature selection are part of the machine learning process. Briefly describe both and why they are necessary. 5 marks = 10 marks

(b) Explain why model should be evaluated. 4 marks

Question 4

(a) With the aid of diagram briefly explain the overall procedure of building a model 10 marks

(b) Explain what a good model is. 5 marks

Question 5

(a) Explain what a model capacity is. 6 marks

(b) Briefly explain the following with the aid of diagram

i. Underfitting ii. Good-fitting iii. Overfitting 3 marks = 9 marks

Question 6

(a) In general, the generalization error can be broken down into the following forms: 3 marks = 9 marks

$$Total\ error = Bias^2 + Variance + Unresolvable\ error$$

With the aid of diagram where applicable, briefly explain each component of the formula.

(b) Explain relationship between model complexity and error with the aid of diagram. 6 marks

Question 7

Briefly explain the following: 5 marks = 15 marks

i. Gradient Descent ii. The concept of Validation set iii. Cross Validation

Question 8

Briefly explain the following common machine learning algorithms 5 marks = 15 marks

i. Linear regression ii. Logistic regression iii. Decision Tree

Question 9

Briefly explain the following common machine learning algorithms 5 marks = 15 marks

(a) SVMs ii. KNN iii. Naïve Bayes

Question 10

(a) Briefly explain Ensemble Learning 5 marks

(b) Explain the following

i. Radom Forest ii. Gradient boosting decision tree (GBDT) 5 marks = 10 marks