

Faculty of Management Studies

End Semester Examination May 2025

MS5EB05 Introduction to ML & Blockchain Technology

Programme	:	MBA	Branch/Specialisation	:	-
Duration	:	3 hours	Maximum Marks	:	60

Note: All questions are compulsory. Internal choices, if any, are indicated. Assume suitable data if necessary.
 Notations and symbols have their usual meaning.

Section 1 (Answer all question(s))

Q1.	What is the primary objective of linear regression?	Marks CO BL		
		1	1	1
Q2.	What does the 'Best Fit Line' in linear regression aim to do?			
	<input type="radio"/> To cluster similar data points <input checked="" type="radio"/> To model the relationship between independent and dependent variables <input type="radio"/> To classify categorical variables <input type="radio"/> To reduce the dimensionality of data			
Q3.	Which of the following best defines a "Problem Statement" in time series forecasting?			
	<input type="radio"/> Describing the tools used in forecasting <input type="radio"/> Clearly stating what needs to be forecasted and why <input checked="" type="radio"/> Outlining the challenges in visualizing time series data <input type="radio"/> Maximizing the number of features <input type="radio"/> Listing the components of the time series	1	2	2
Q4.	The ARIMA model consists of which three components?			
	<input type="radio"/> Auto generative, regressive, impactful, mathematical analysis <input type="radio"/> Additive, multiplicative, exponential <input checked="" type="radio"/> Autoregressive, integration, moving average <input type="radio"/> Actual, residual, integrated	1	2	2
Q5.	Which of the following is not a type of clustering?			
	<input type="radio"/> Hierarchical clustering <input type="radio"/> Partitioning clustering <input checked="" type="radio"/> Density-based clustering <input type="radio"/> Regression-based clustering	1	3	3
Q6.	The demonstration of K-Means typically involves which two steps repeatedly?			
	<input type="radio"/> Feature extraction and dimensionality reduction <input type="radio"/> Standardization and PCA <input checked="" type="radio"/> Assigning clusters and updating centroids <input type="radio"/> Prediction and labeling	1	3	3
Q7.	What is the purpose of a sigmoid curve in logistic regression?			
	<input type="radio"/> To convert categorical variables to numbers <input type="radio"/> To normalize features <input checked="" type="radio"/> To map predicted values to a probability between 0 and 1 <input type="radio"/> None of these	1	4	3
Q8.	Specificity in classification is defined as:			
	<input type="radio"/> True positives / (True Positives + False Negatives) <input checked="" type="radio"/> True negatives / (True Negatives + False Positives) <input type="radio"/> False positives / (True Negatives + False Positives) <input type="radio"/> None of the above	1	4	2

Q9. Which is a major drawback of a centralized system?

1 5 1

- Single point of failure
- Too much decentralization
- Requires consensus mechanisms
- Smart contracts cannot run

Q10. Which blockchain platform is associated with permissioned networks and enterprise use cases?

1 5 1

- Ethereum
- Bitcoin
- Ripple
- Hyperledger fabric

Section 2 (Answer all question(s))

Q11. What is multicollinearity in multiple regression?

Marks CO BL

2 1 2

Rubric	Marks
Multicollinearity	1
Multiple regression	1

Q12. (a) Describe the 8-step process for building a simple linear regression model.

8 1 3

Rubric	Marks
Each Step 1 Mark	8

(OR)

(b) What are the assumptions made in SLM? How do they affect model performance?

Marks CO BL

2 2 2

Rubric	Marks
assumptions made in SLM	4
how do they affect model performance	4

Section 3 (Answer all question(s))

Q13. Explain the key characteristics of quadratic trend in time series forecasting.

Marks CO BL

2 2 2

Rubric	Marks
characteristics of quadratic trend - 2 Marks	2

Q14. (a) Explain in detail the structure and working of an ARIMA model. Provide a use case or example.

8 2 3

Rubric	Marks
Explain in detail the structure and working of an ARIMA model.	5
Provide a use case or example.	3

(OR)

(b) Elaborate on how overfitting affects time series models and discuss techniques to detect and avoid it.

8 2 3

Rubric	Marks
Elaborate on how overfitting affects time series models	4
discuss techniques to detect and avoid it.	4

Section 4 (Answer all question(s))

Marks CO BL

Q15. Define clustering. What is the main goal of clustering in data analysis?

3 3 3

Rubric	Marks
Define clustering	1
What is the main goal of clustering in data analysis?	2

Q16. (a) Explain the K-Means clustering algorithm with an example.

7 3 2

Rubric	Marks
Explain the K-Means clustering algorithm	5
example	2

(OR)

(b) Mention the purpose of scaling and standardization in clustering.

Rubric	Marks
purpose of scaling	3.5
purpose of standardization	3.5

Section 5 (Answer all question(s))

Marks CO BL

Q17. Differentiate between classification and regression.

4 4 3

Rubric	Marks
1 MARKS FOR EACH STEPS	4

Q18. (a) Compare different types of classification techniques with suitable examples.

6 4 3

Rubric	Marks
Compare different types of classification techniques	4
examples	2

(OR)

(b) How is "purity" measured in decision trees? Describe how Gini Index or Entropy is used in the process.

Rubric	Marks
How is "purity" measured in decision trees?	3
Describe how Gini Index or Entropy is used in the process.	3

Section 6 (Answer all question(s))

Marks CO BL

Q19. Define blockchain technology. What are its core features?

4 5 2

Rubric	Marks
Define blockchain technology	2
What are its core features?	2

Q20. (a) Compare the advantages and disadvantages of a centralized trusted party vs a decentralized ledger system.

6 5 3

Rubric	Marks
advantages	3
disadvantages	3

(OR)

- (b)** What is the Hyperledger Fabric platform? How does it differ from public blockchains like Ethereum?

Rubric	Marks
What is the Hyperledger Fabric platform	3
How does it differ from public blockchains like Ethereum?	3
