



**MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY**  
P.O. Box 972-60200 – Meru-Kenya  
Tel: +254(0) 799 529 958, +254(0) 799 529 959, + 254 (0) 712 524 293,  
Website: [info@must.ac.ke](mailto:info@must.ac.ke) Email: [info@must.ac.ke](mailto:info@must.ac.ke)

---

**University Examinations 2023/2024**

FOURTH YEAR SECOND SEMESTER EXAMINATION FOR THE DEGREE OF  
BACHELOR OF DATA SCIENCE

**CDS 3451: EMERGING TECHNOLOGIES IN DATA SCIENCE**

**DATE: APRIL 2024**

**TIME: 2 HOURS**

**INSTRUCTIONS:** Answer question *one* and any other *two* questions

---

**QUESTION ONE (30 MARKS)**

- a. Define the term virtualization and explain how it enables cloud computing (4 Marks)
- b. Explain three security challenges associated with IoT devices and sensor networks. (6 Marks)
- c. Define machine learning and distinguish between supervised, unsupervised, and reinforcement learning. (5 Marks)
- d. Discuss the importance of authentication and authorization in ensuring data security. Explain the difference between the two concepts. (6 Marks)
- e. Explain the concept of blockchain technology and four key characteristics. (6 marks)
- f. Define deep learning and explain its significance in the field of artificial intelligence. (3 Marks)



MUST is ISO 9001:2015 and



ISO/IEC 27001:2013 CERTIFIED

## QUESTION TWO (20 MARKS)

- a. Imagine you are working for a bank and tasked with developing a system to detect potential fraudulent transactions using historical transaction data.
- i. Explain how unsupervised learning can be used for anomaly detection in this scenario.
  - ii. Briefly discuss two specific anomaly detection algorithms and their suitability for this task. (4 Marks)
- b. Discuss the role of data analytics and visualization in identifying trends, patterns, and outliers in datasets. How can these insights inform business decisions? (6 Marks)
- c. Explain the concept of a firewall in network security and its role in preventing unauthorized access to data. (6 Marks)

## QUESTION THREE (20 MARKS)

- a. Explain two ethical considerations involved in developing and deploying AI models?
- i. Provide two specific examples. (6 Marks)
- b. A social media platform uses a data science model to recommend content to its users. This model utilizes user data, including demographics, interests, and past interactions. However, concerns exist regarding potential privacy breaches.
- i. Discuss two potential security threats this system could face, and explain how each threat could be mitigated through data science techniques. (6 Marks)
  - ii. Briefly mention two ethical considerations involved in such recommendation systems. (2 Marks)
- c. Explain the 3 Vs (Volume, Velocity, Variety) of sensor data and briefly discuss the challenges associated with handling them in the context of IoT. (6 Marks)



#### **QUESTION FOUR (20 MARKS)**

- a. Consider a scenario where you are designing a blockchain-based identity management system for a government agency to verify the identities of citizens in a secure and decentralized manner. Discuss how blockchain technology can be leveraged to create a tamper-proof and auditable record of citizen identities while preserving user privacy. Explain the use of cryptographic hashing and digital signatures to ensure data integrity and authentication in the blockchain network. Discuss the potential applications of blockchain-based identity management in areas such as voting systems or digital passports. (8 Marks)
- b. Explain the basic working principles of a convolutional neural network (CNN) and its key applications. (4 Marks)
- c. A public health organization wants to analyze factors contributing to the spread of a disease in a specific region. They have data on demographics, health records, and environmental factors for that region. (8 Marks)
- i. Describe the data exploration steps you would take to understand the distribution of cases and identify potential correlations with various factors.
- ii. Choose two different types of data visualization techniques you would use to present the spatial distribution of the disease and its potential association with other factors.

#### **QUESTION FIVE (20 MARKS)**

- a. Imagine you are working on a project to predict crop yield using data from various sensors in a smart farm (e.g., soil moisture, temperature, light). Describe the potential challenges associated with handling and analyzing this type of data. Briefly discuss two machine learning techniques you could use for prediction and explain their suitability for this task. (8 Marks)
- b. Explain the difference between synchronous and asynchronous communication in cloud computing. Provide examples of each. (6 Marks)
- c. Outline three key security concerns associated with sensor data and IoT devices. Briefly describe security measures employed to mitigate these concerns (6 Marks)

