Based on the **GTU Paper Format** you provided and standard Diploma Engineering trends, I have analyzed the unit-wise weightage and repetition patterns for **Unit 1: Fundamental Concepts of Data Mining**.

In the 70-mark paper, Unit 1 typically contributes around **10–12 marks**, appearing mostly in **Question 1** and occasionally in **Question 2**.

Unit 1: Predicted Question Bank (Fundamental Concepts)

1. Most Repeated / High-Probability Questions

These follow the (a) 3-mark, (b) 4-mark, and (c) 7-mark structure of your model paper.

[Short Answer Type - 3 Marks]

- **Define Data Mining.** Why is it often called "Knowledge Discovery from Data" (KDD)?
- **List any four applications** of Data Mining in the modern IT industry.
- **Explain the term 'Metadata'** in the context of a Data Warehouse.
- **What are 'Primitives'** in data mining? List any two.

[Descriptive Type - 4 Marks]

- **Differentiate between Data Mining and Data Warehousing.** (Provide at least 4 points).
- **Explain the Architecture of a typical Data Mining system** with a neat block diagram.
- **Briefly describe the major issues** faced in Data Mining (e.g., performance, diversity of data).
- **Explain the 'Classification of Data Mining Systems'** based on the types of data or knowledge to be mined.

[Long Answer / Diagram-Based - 7 Marks]

- **Describe the KDD Process in detail.** Draw a neat diagram showing all steps from Data Cleaning to Knowledge Presentation.
- **Explain the Evolution of Database Technology.** How did we move from simple file processing to modern Data Mining?
- **Discuss Data Mining Functionalities.** Explain Characterization, Discrimination, Association, and Evolution Analysis with examples.

---

2. Application & Logical Thinking Questions

These questions test your ability to apply "Theory to System" and are key for scoring 7/7 in part (c) of your paper.

1. **Scenario-based Logic:** A supermarket manager wants to know which items are frequently bought together to rearrange the store shelves. Which data mining functionality should be used? Explain the logic behind this "Market Basket Analysis."
2. **System Design:** If you are building a Data Mining system for a **Healthcare Hospital**, identify three types of data sources (e.g., SQL databases, image files) the system must integrate. Mention one "Major Issue" you might face with this sensitive data.
3. **Concept Interpretation:** "Data Mining is not just a simple transformation of technology, but an evolution." Justify this statement by explaining how it differs from traditional "Data Querying" in SQL.
4. **Problem Solving:** During the KDD process, if a dataset contains **"Noise"** (incorrect values) and **"Missing Values,"** which specific step of KDD will you focus on? Explain two methods to handle this.
5. **Comparative Reasoning:** Why is a **Data Warehouse** necessary for effective Data Mining? Could we perform data mining directly on an operational database (OLTP)? Why or why not?

Exam Strategy for Unit 1:

- **The "KDD Diagram" is non-negotiable:** It has a 90% probability of appearing as a 7-mark question. Practice drawing it in under 3 minutes.
- **Key Comparisons:** GTU examiners love "Data Mining vs. Data Warehouse" and "OLTP vs. OLAP." Memorize a table with 4–5 clear points.
- **Focus on Keywords:** When defining functionalities, use terms like *Frequent Patterns, Class Labels, and Anomalies* to get full marks.