

This document is intended to be exported as Anki flashcards.
Answers generated/derived by/from AI are coloured in blue! (Gemini/NotebookLM in use)
“Open-book” answers are marked as purple, corrections from model solutions are coloured as green.
Any arguments arising from group discussion are marked as fuchsia.

7CCSMDM1

Practice Questions

If you spot anything wrong, please let me know!

discord: @insertish

izzy · insrt.uk

Contents

1. Tutorial 1	2
---------------------	---

1. Tutorial 1

Which of the following statements about Data Mining is false?

- ✗ It can be automated.
- ✗ It can be used for making predictions.
- ✗ It extracts meaningful and useful patterns from data.
- ✓ It requires generating new data.

Which of the following statements about tabular data sets, i.e. ones represented in the form of a two-dimensional table, is false?

- ✗ Columns correspond to attributes.
- ✗ Rows correspond to instances.
- ✓ Attributes are typically assumed to be independent.
- ✗ Instances are typically assumed to be independent.

A data set from a touristic agency contains instances with information about their travelers. The ranking in number of stars (1 – 5) of the last hotel visited by a customer is an example of a:

- ✗ nominal attribute — it can be ordered
- ✓ ordinal attribute
- ✗ continuous numeric attribute — only five options
- ✗ discrete numeric attribute — values 1...5 can be represented a...b; requires we assume that the difference from 2-star hotel to 3-star hotel is the same as 4- to 5-star, etc

Which of the following data mining tasks aims to divide the instances into groups, so that instances in the same group are more similar than instances in different groups?

- ✗ regression
- ✗ classification
- ✓ clustering
- ✗ association

Which of the following are unsupervised learning tasks?

- ✓ association and clustering
- ✗ classification and regression
- ✗ association and regression
- ✗ classification and clustering

stopped at slide 21