

Assessment Details and Brief	
Module Code:	CI603
Module Title:	Data Mining
Author(s)/marker(s) of assessment:	Dr Gulden Uchyigit
Assessment number:	1
Assessment title:	Data Mining Coursework
Percentage contribution to module mark:	100%
Weighting of components within this assessment:	N/A
Module learning outcome(s) assessed:	<p>LO1 Demonstrate an understanding of both the theoretical and practical aspects of data mining.</p> <p>LO2 Demonstrate an understanding of the algorithmic techniques in data mining.</p> <p>LO3 Apply data mining techniques to solve real-world complex problems.</p>
The assessment is marked anonymously	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
Date of issue:	10/2/2025
Deadline for submission:	19/05/2025 (3pm)

1. A copy of your coursework submission may be made as part of the University of Brighton's and School of Architecture, Technology and Engineering procedures which aim to monitor and improve quality of teaching. You should refer to your student handbook for details.
2. All work submitted must be your own and all sources which do not fall into that category must be correctly attributed. The markers may submit the whole set of submissions to the JISC Plagiarism Detection Service.

Method of submission:	Electronic submission through MyStudies.
Date feedback will be provided:	Feedback will be provided electronically within 4 weeks of the hand-in date.

General Guidelines and requirements

Assignment: you need to (a) select a suitable dataset, (b) provide an in-depth analysis of the data, including appropriate Python implementation and (c) write a report of your findings indicating your approach, the theoretical background to any techniques you have deployed, indicating applications in a wider context and appropriate conclusions. You may choose to use a Big Data dataset but you are not required to. If you choose a Big Data dataset you will need to use Pyspark; For a *non* Big Data dataset may choose to use Python or Pyspark.

When selecting your dataset, you need to consider and discuss the nature of the data, its characteristics and properties, the quality of the data, any pre-processing tasks that you need to undertake, what you are able to achieve by analysing the data and the kinds of techniques that you are able to apply to the data.

In analysing your data, you need to select suitable techniques from those covered in the module (and beyond) and apply them to obtain appropriate outcomes. Your data set should be sufficiently large and/or complex that you will need to implement your analysis techniques, for example using Python. It may also be that you will need to apply dimensionality reduction on your data set.

Your report should include a discussion of the data itself, including any required pre-processing tasks, a clear description of the techniques that you have applied and their theoretical background and detailed outcomes of your analysis. You should research and describe the range of applications of your techniques. You should also include an executive summary in the style of a professional report suitable for a company CEO.

Total length of report: 2500 words.

Assessment criteria and marking rubric

Grade	General criteria	Evidence of understanding 25 marks	Implementation 25 marks	Research 30 marks	Report 20 marks
A+ 80-100% High Distinction	An outstanding response to the task. The work demonstrates most or all of the following characteristics beyond that expected for work at the given level of study within the discipline	Outstanding clarity of the underlying theoretical underpinnings, including research based on appropriate sources.	Outstanding implementation of the algorithms including extended features. Exceptionally well written and documented code.	Outstanding research demonstrating significant depth and breadth of knowledge, insight and identifies appropriate means of addressing issues	A very clearly and exceptionally well written and structured report covering all aspects of the assignment.
A 70-79% Distinction	An excellent response to the task. The work demonstrates most or all of the following characteristics in relation to those expected at the given level of study within the discipline	Excellent clarity of the underlying theoretical underpinnings, including some research based on appropriate sources.	Excellent implementation of the algorithms possibly including some extended features. Excellent written and documented code.	Excellent research demonstrating depth and breadth of knowledge, insight and identifies appropriate means of addressing issues	A very clearly written and well-structured report covering all aspects of the assignment.
B 60-69% Merit	A good to very good response to the task. The work demonstrates most or all of the following characteristics in relation to those expected at the given level of study within the discipline	Good clarity of the underlying theoretical underpinnings, including appropriate use of sources.	Good implementation of the algorithms. Well written and documented code.	Good research demonstrating some depth and breadth of knowledge, insight and identifies appropriate means of addressing issues	A clearly structured and generally well written report covering all aspects of the assignment

C 50-59% Pass	A sound, competent response to the task. The work demonstrates most or all of the following characteristics in relation to those expected at the given level of study within the discipline	Sound description of the underlying techniques including appropriate use of sources.	Sound implementation of the algorithms. Adequately written and documented code.	Sound research demonstrating minimal depth and breadth of knowledge and some insights.	A well structured report providing a sound explanation of most aspects of the assignment.
D 40 -49 Pass	An adequate, but weak response to the task. The work demonstrates most or all of the following characteristics in relation to those expected at the given level of study within the discipline.	An adequate but weak description of the underlying techniques, demonstrating weakness of understanding of some aspects.	Weak or partial implementation of the algorithms. Mostly adequately written and documented code but demonstrating some weaknesses	Weak or partial research demonstrating weak depth and breadth of knowledge and limited insights.	A weak and poorly structured report providing an limited explanation of some aspects of the assignment.
E <40 Fail	An unsatisfactory response to the task. The work may display some strengths, but these are outweighed by several weak features in relation to those expected at the given level of study within the discipline.	An unsatisfactory description of the underlying techniques, demonstrating significant weakness of understanding of some aspects.	An unsatisfactory implementation of the algorithms.	An unsatisfactory research demonstrating lack of depth and breadth of knowledge and insights.	An unsatisfactory report providing an inadequate explanation of the assignment.