

Assessment Rubric COMP3425 Data Mining

This rubric will be used to mark your assignment. You are advised to use it to supplement your understanding of what is expected for the assignment and to direct your effort towards the most rewarding parts of the work. Your assignment will be marked out of 100, and marks will be scaled back to contribute to the defined weighting for assessment of the course.

Review Criteria	Max Mark	Exemplary	Excellent	Good	Acceptable	Unsatisfactory
1. Platform & 2. Data	10	<p>9-10</p> <p>1. Platform description complete (memory, CPU, operating system, software).</p> <p>2a Demonstrates understanding of the purposes and process sufficient to frame report.</p> <p>2b All correlations for mentioned variables clearly explained in terms of the data semantics, in the correct directions and for correct or plausible domain reasons.</p>		<p>7-8</p> <p>1. Platform description complete (memory, CPU, operating system, software).</p> <p>2a Clear description of the the data domajn.</p> <p>2b Partially clear and correct explanation in terms of data semantics</p>	<p>5-6</p> <p>1. Platform description incomplete (memory, CPU, operating system, software).</p> <p>2a Attempt but unclear</p> <p>2b Partial description of variables or unclear</p> <p>2b Partial explanation in data context</p>	<p>0-4</p> <p>1. Platform description incomplete.</p> <p>2a Incomplete or faulty</p> <p>2b Description unrelated to correlation of variables.</p> <p>2b Explanation unrelated to data source</p>

Criteria	Max	Exemplary	Excellent	Good	Acceptable	Unsatisfactory
3. Association mining 7	10	9-10 a. Answers demonstrate deep understanding of association mining, by the careful selection of interesting and differentiated rules and clear rationale for interestingness. b. Comment shows original and insightful analysis of association mining on the problem.		7-8 a Support and confidence clear a 3 rules given a objective interestingness is given for all 3 a subjective interestingness attempted b Comment makes sense.	5-6 a Support or confidence not clear a < 3 rules given a objective interestingness is incomplete a subjective interestingness is incomplete b Comment cursory.	0-4 Required information not provided and/or incorrect or misleading, demonstrating lack of engagement with the problem
4. Simple classification 7	10	9-10 Explanation of <i>Opinionated</i> demonstrates understanding of problem. Deep understanding of the 4 models demonstrated thorough analysis of performance on the task.		7-8 a correctly explains why definition of <i>Opinionated</i> makes it seem easy b 4 confusion matrixes given b confusion matrixes explained in terms of the data and the method and the model learnt. c evidence of understanding what the models are doing c reasoning for comparative performance demonstrating understanding of the methods behind them	5-6 a partially explains why definition of <i>Opinionated</i> makes it seem easy b 4 confusion matrixes given b confusion matrixes explained at face value only c partial understanding of learnt models c comparative performance only cursorily presented c reason for comparative performance is shallow	0-4 a inadequate explanation b confusion matrix missing or misunderstood. c interpretation of confusion matrix missing or faulty c little understanding of what the models are doing c missing or unexplained comparative analysis

Criteria	Max	Exemplary	Excellent	Good	Acceptable	Unsatisfactory
5. Prediction 13	20	<p>17-20</p> <p>Approach to problem demonstrates serious effort to produce good results and a deep understanding of the relative benefits of the 2 methods in the context of the problem domain.</p> <p>Results are interpreted in the context of the problem domain.</p>	<p>14-16</p> <p>a justification for choice shows understanding of the comparative benefits of each and extensive experiments.</p> <p>b parameter variations shows a combination of experimentation and understanding of the parameters</p> <p>c several subjective and objective evaluation measures used as appropriate to method including synthesised evaluation</p> <p>c justification for stopping demonstrates awareness of appropriateness of best result and scope of potential for further improvement</p>	<p>12-13</p> <p>a justification for choice shows understanding of the comparative benefits of each and experiments with performance.</p> <p>b parameter variations shows a combination of experimentation and understanding of the parameters.</p> <p>c multiple subjective and objective evaluation measures used as appropriate to method</p> <p>c justification for stopping demonstrates awareness of appropriateness of best result</p>	<p>10-11</p> <p>a justification for choice shows some understanding of the comparative benefits of each or experiments with performance.</p> <p>b parameter variation demonstrates some experimentation</p> <p>c cursory evaluation given</p> <p>c justification for stopping perfunctory</p>	<p>0-9</p> <p>a weak justification for choice</p> <p>b variation insufficient</p> <p>c evaluation fails to demonstrate effort or understanding of evaluation</p> <p>c justification for stopping effectively absent</p>

Criteria	Max	Exemplary	Excellent	Good	Acceptable	Unsatisfactory
6. Complex Classification 20	30	26-30 Exemplary use of classification methods with comprehensive and fit-for-purpose performance analysis on the problem that includes meaningful reflection over the three methods.	22-25 a explanation sound b,c,d parameter variation clear and extensive demonstrating understanding of effect in all 3 methods b,c,d error matrix and ROC correctly interpreted in all 3 methods b,c,d extensive use of specific evaluation methods and significance clearly explained in all 3 methods	18-21 a explanation sound b parameter variation clear and sufficient for good results b error matrix correctly interpreted b ROC correctly interpreted b some specific evaluation methods used c parameter variation clear and sufficient for good results c error matrix correctly interpreted c ROC correctly interpreted c some specific evaluation methods used d parameter variation clear and sufficient for good results d error matrix correctly interpreted d ROC correctly interpreted d some specific evaluation methods used	15-17 a satisfactory approach to dataset partitioning b parameter variation perfunctory b error matrix given b ROC given b few specific evaluation methods used well c parameter variation perfunctory c error matrix given c ROC given c few specific evaluation methods used well d parameter variation perfunctory d error matrix given d ROC given d few specific evaluation methods used well	0-14 a explanation incorrect or unsound use of training/testing/validation data b no parameter variation b no error matrix b no or faulty ROC b specific evaluation methods missing c no parameter variation c no error matrix c no or faulty ROC c specific evaluation methods missing d no parameter variation d no error matrix d no or faulty ROC d specific evaluation methods missing

Criteria	Max	Exemplary	Excellent	Good	Acceptable	Unsatisfactory
7. Clustering 8	10	<p>9-10 The application of k-means algorithm to the dataset and its evaluation demonstrates exemplary understanding of the algorithm, its evaluation, and its limitations.</p> <p>Suitable evaluation methods or clustering experiments in addition to those required here may be used.</p>		<p>7-8</p> <p>a Convincing justification for k</p> <p>b Measure calculated correctly. Discussion recognises value and limitations</p> <p>c Discussion on centres reflects numeric results and emphasises the interesting parts that relate to the significance in domain terms</p> <p>d Correct scatterplot included and description shows understanding linked to data domain</p>	<p>5-6</p> <p>a Justification offered but not clear or unconvincing</p> <p>b Measure calculated correctly</p> <p>c Discussion on centres reflects numeric results</p> <p>d Correct scatterplot included. Attempt at influences.</p>	<p>0-4 Clustering experimentation and discussion inadequate</p>
8. Qualitative Summary 7	10	<p>9-10 Many aspects of evaluation are discussed and a clear conclusion is drawn, with direct reference to the purpose of the data collection.</p> <p>Proposal for further investigation demonstrates creativity and thoughtful engagement with the problem, clearly building on the work reported.</p>	8 A clear conclusion is drawn from the work reported and a defended proposal for further investigation is proposed, with clear links to both the work reported and the domain of application.	7 A rounded, balanced summary of the work is presented with a justified proposal given.	5-6 A summary of the work is presented and a proposal made.	<p>0-4 Answer does not demonstrate adequate engagement with the problem nor a qualitative understanding of the work reported.</p>

Total: 71