**UNIVERSITY CENTRE SOMERSET**

**Computing and Digital Technologies**  
  
Assignment Coversheet and Grading Criteria  
2022 / 2023

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Qualification** | | | **Module Code and Title** | |
| BSc (Hons) Computing and Digital Technologies  FdSc Computing and Digital Technologies | | | SCDT64 Machine Learning and Artificial Intelligence | |
| **Student Name and Number** | | | **Module Tutor** | |
|  | | | Simon West | |
| **Date Issued** | | **Submission Date** | | **Return Date** |
| 10/03/2023 12:00 | | 31/03/2022 15:00 | | 01/07/2023 12:00 |
|  | | |  | |
| **Assignment Number** | 1 of 2. This assignment is worth 50% of the overall module. | | | |
| **Assignment Title** | Coursework One – Practical and Presentation | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Module Learning Outcomes**  *To achieve the outcomes the evidence must show that the learner is able to:* | |  | **Task no.** |
| A1 | To understand and critique the concepts of artificial intelligence, machine learning and decision making for data science, and compare and test a range of techniques. |  | 1 |
| C2 | Implement appropriate machine learning tools and techniques in the classification of data items. |  | 1 |
| D1 | Apply critical thinking skills and an analytical and scientific approach to problem solving. |  | 1 |

|  |  |
| --- | --- |
| **Word Count of Submission** | 2000 (+/-10%) |
| **Student Declaration** | |
| Through submitting this assignment through Turnitin you agree that the work was prepared entirely by yourself in accordance with Open University’s Prevention of Academic Dishonesty Code of Practice. | |

|  |
| --- |
| **Assignment Feedback** |
| All feedback for this assignment will be provided through Turnitin in accordance with the grading criteria below on the given return date. |



|  |  |
| --- | --- |
| **Assignment Task(s)** | |
| **Task no.** | **Task details** |
| 1 | You have been provided with a wide selection of datasets that can be used to demonstrate the tools and techniques used in machine learning classification. These datasets can be used for:   * Binary Classification * Linear / Logistic Regression * Clustering * Multiclass Classification * Natural Language Processing   For this assignment you are required to:   * Select and compare a range of tools and techniques in machine learning classification e.g. a critical comparison of Python/Scikit-Learn/Pickle/Flask against ML.NET/ASP.NET Core/Azure.   500 words here, stick it in jupyter notebook   * Undertake each phase of the CRISP-DM model in the classification of data items for a problem of your choosing. You must use one of the given datasets although you can select your own tools and techniques. You must implement the following CRISP-DM phases:  * + Business Understanding (Come up with a problem yourself)   + Data Understanding   + Data Preparation   + Modelling   + Evaluation   + Deployment   You decide how many words goes here (No recommended amount)   * Evaluate the suitability of the tools and techniques used to solve your chosen machine learning classification problem. * Evaluate the suitability of using the CRISP-DM standard process for machine learning projects and where suitable, make suggestions for improvement.   You should produce a written report and software artefact e.g. a Jupyter Notebook and Flask application; ML.NET and C# application etc. You will also be required to give a technical demonstration and code walk-through of both your report and software artefact in place of a traditional presentation.  When submitting it all files including PDF must be submitted.  Save/Export it as a PDF file when done.  Don’t be descriptive, evaluation is important  References goes in notebook  **It’s a presentation** |
|  |  |

|  |
| --- |
| **Sources of Information** |
| Machine Learning for Beginners, Microsoft (2021): <https://github.com/microsoft/ML-For-Beginners>  Practical Machine Learning with Python A Problem-Solver’s Guide to Building Real-World Intelligent Systems (2018) Sarkar, Bali and Sharma |

|  |
| --- |
| **Submission Requirements** |
| 1. Check the grading criteria below to ensure your assignment document meets the demands of the above task(s). 2. If the assignment contains any practical work, place any relevant additional files (i.e. software) into a folder and zip-up the entire folder into a single zip file. 3. Name the assignment document and any relevant ZIP file using the following format:  SCDTXX \_CWX\_*StudentNumber\_FirstName\_LastName*.docx/zip  (replace the *placeholders* with your student number, first and last name respectively) 4. Go to the Turnitin and use the upload facility to submit your assignment and any required ZIP file to the relevant module. There is no need to submit this assignment brief. |

|  |  |
| --- | --- |
| **IMPORTANT INFORMATION** | |
| * Please stay within the limits of the word count stated at the top of assignment brief. Any additional content over the word count limit (plus or minus 10%) will be disregarded and not be assessed at all.  All work should be submitted online via Turnitin.Please ensure that you submit your assignment on the right submission slot for each module.It is your responsibility to check that you can access Turnitin and Teams properly. If your college student account is locked, please contact ITU on 01823 366 354 or email them at ITHelpdesk@btc.ac.uk and request to have your account unlocked, but please ensure you allow plenty of time to do this, do not leave everything until the last day of your deadline.If there are circumstances where you need to submit your assignment other than online, please discuss your needs with the module tutor and alternative arrangements could be made so that you can submit your coursework within the set deadline.Regulations allow you to submit coursework up to 6 working days late. A penalty of deducting 10% will be applied for each day an assignment is late, with a maximum penalty of deducting 60% from your final mark for the late assignment. Any assignment submitted later than 6 days with be awarded a mark of zero. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Numeric Grade | Descriptor  (Class Band) | Undergraduate Grading Criteria | | | | |
| Comparison and Evaluation of Tools and Techniques used in ML Classification | Implementation and Analysis of Business and Data Understanding | Implementation and Analysis of Data Preparation, Modelling and Evaluation | Deployment and Evaluation of the Suitability of the CRISP-DM Model | Overall presentation, Structure and Sources |
| 80-100 | Outstanding  (Upper Distinction) | A comparison and evaluation of tools and techniques for ML classification have been explored and evaluated with suitable use cases to an outstanding standard. The tools and techniques selected are to be industry standard, acceptable practices. | An outstanding implementation of the business understanding, and data understanding phases has been undertaken. Industry standard tools and methods have been employed to define and communicate business requirements for a specific ML problem. A wide variety of industry standard tools and techniques have been employed in a highly professional way to understand and explain the chosen dataset including statistics, charts, checking for missing values etc. | An outstanding implementation of the data preparation and modelling has been undertaken. Industry standard tools and methods have been employed to clean, wrangle, normalise and transform the raw data. A variety of ML models/algorithms have been trained, tested and evaluated in a highly professional manner. Each model has been tested and evaluated for performance using various metrics e.g. F1, confusion matrices etc. | An outstanding implementation and use case of ML model deployment has been demonstrated. Industry standard tools and methods have been employed in the deployment of the ML model and provided an outstanding end-user UI. An outstanding evaluation has been conducted on the suitability of the CRISP-DM model for ML, highly relevant and well-justified improvements have been suggested. | Referencing is of professional publication quality.  The structure of the report follows all UCS guidelines and is over publication quality. |
| 70-79 | Excellent  (Lower Distinction) | A comparison and evaluation of tools and techniques used for ML classification within industry, a selection of tools and techniques have been analysed with research from a variety of sources and use cases from which excellent conclusions have been drawn. | An excellent implementation of the business understanding, and data understanding phases has been undertaken. Mostly industry standard tools and methods have been employed to define and communicate business requirements for a specific ML problem. A variety of industry standard tools and techniques have been employed in a professional way to understand and explain the chosen dataset including statistics, charts, checking for missing values etc. | An excellent implementation of the data preparation and modelling has been undertaken. Mostly industry standard tools and methods have been employed to clean, wrangle, normalise and transform the raw data. A variety of ML models/algorithms have been trained, tested and evaluated in a professional manner. Each model has been tested and evaluated for performance using various metrics e.g. F1, confusion matrices etc. | An excellent implementation and use case of ML model deployment has been demonstrated. Mostly industry standard tools and methods have been employed in the deployment of the ML model and provided an excellent end-user UI. An excellent evaluation has been conducted on the suitability of the CRISP-DM model for ML, relevant and well-justified improvements have been suggested. | All citations have been incorporated properly into the text. All references listed properly in the reference list.  The structure of the report follows all UCS guidelines and is over very high quality. |
| 60-69 | Very Good (Commendation) | A very good critical comparison and evaluation of tools and techniques utilised for ML classification within industry, a selection of tools and techniques have been analysed with research from a variety of sources and use cases to inform very good conclusions. | A very good implementation of the business understanding, and data understanding phases has been undertaken. Some industry standard tools and methods have been employed to communicate business requirements for a relevant ML problem. Some industry standard tools and techniques have been employed to understand and explain the chosen dataset including statistics, charts, checking for missing values etc. | A very good implementation of the data preparation and modelling has been undertaken. Some industry standard tools and methods have been employed to clean, wrangle and transform the raw data. ML models/algorithms have been trained, tested and evaluated in a professional manner. Each model has been tested and evaluated for performance using various metrics e.g. F1, confusion matrices etc. | A very good implementation and use case of ML model deployment has been demonstrated. Some industry standard tools and methods have been employed in the deployment of the ML model and provided an effective end-user UI. A very good evaluation has been conducted on the suitability of the CRISP-DM model for ML, with some relevant and justified improvements have been suggested. | Only minor errors in incorporating citations into the text. Only minor errors in incorporating references into the reference list.  The structure of the report follows all UCS guidelines and is over high quality. |
| 50-59 | Good/Satisfactory  (Upper Pass) | Good comparison and evaluation which is critical in places, of a suitable selection of tools and techniques that are often used within industry. Evidence and use cases have been used throughout to enhance conclusions of how these tools may be used in ML classification. | A good implementation of the business understanding, and data understanding phases has been undertaken. Standard tools and methods have been employed to communicate business requirements for a relevant ML problem. Standard tools and techniques have been employed to understand and explain the chosen dataset including statistics, checking for missing values etc. | A good implementation of the data preparation and modelling has been undertaken. Standard tools and methods have been employed to clean, wrangle and transform the raw data. Some suitable ML models/algorithms have been trained, tested and evaluated in a suitable manner. Each model has been tested and evaluated for performance using various metrics. | A good implementation and use case of ML model deployment has been demonstrated. Standard tools and methods have been employed in the deployment of the ML model and provided a suitable end-user UI. A good evaluation has been conducted on the suitability of the CRISP-DM model for ML, with relevant improvements being suggested. | A few errors in incorporating citations into the text. A few errors in incorporating references into the reference list. The structure of the report follows all UCS guidelines. |
| 40-49 | Marginal Pass / Satisfactory (Lower Pass) | Satisfactory comparison and evaluation of a reasonable set of tools and techniques that may be sometimes used within industry. Some evidence and use cases have been used to generate justifications of how these tools may be used. | A satisfactory implementation of the business understanding, and data understanding phases has been attempted. Tools and methods have been employed / attempted in which to communicate business requirements for a ML problem. Tools and techniques have been employed to attempt to try and understand and explain the chosen dataset including statistics, checking for missing values etc. | A satisfactory implementation of the data preparation and modelling has been attempted. Tools and methods have been employed to clean and transform the raw data. A suitable ML model/algorithm has been trained, tested and evaluated in a suitable manner. Each model has been tested and an attempt has been made to evaluate its performance. | A satisfactory implementation and use case of ML model deployment has been attempted. Tools and methods have been employed in the deployment of the ML model and provided an end-user UI. A suitable evaluation has been conducted on the CRISP-DM model for ML, with some improvements being suggested. | Some errors in incorporating citations into the text. Some errors in incorporating references into the reference list. The structure of the report follows all UCS guidelines. |
| 20-39 | Clear Fail  (Fail) | Inadequate or largely incorrect tools and techniques selected for use, with limited examples of where these may be seen the wider industry or in any relevant use case. | Inadequate or largely incorrect tools and techniques selected for use, with limited application for the business and data understanding phases. | Inadequate or largely incorrect tools and techniques selected for use, with limited application for the data preparation, modelling and evaluation phases. | Inadequate or largely incorrect tools and techniques selected for use, with limited application for the deployment phases. Little or no evidence in the evaluation of the CRISP-DM model. | Many errors in incorporating citations in text.  Many errors in incorporating and reference list. Poor structure throughout that does not report follow all UCS guidelines. |
|  | Nothing of Merit  (Fail) | No relevant tools and techniques selected for use, with no examples of where these may be seen within the wider industry/use case. | No relevant tools and techniques selected for use, with no attempt to employ them for a chosen ML problem. | No relevant tools and techniques selected for use, with no attempt to employ them for a chosen ML problem. | No relevant tools and techniques selected for use, with no attempt to employ them for a chosen ML problem. | References have not been incorporated in accordance with UCS guidelines. Very poor structure throughout. |