A picture containing icon

Description automatically generated

**CCT College Dublin Continuous Assessment**

|  |  |  |  |
| --- | --- | --- | --- |
| **Programme Title:** | *BSc(Hons)CompIT Year 4* | | |
| **Cohort:** | *Sept 2020* | | |
| **Module Title(s)**: | *Machine Learning for AI* | | |
| **Assignment Type:** | *Individual* | **Weighting(s)**: | *50%* |
| **Assignment Title:** | *ML\_CA2* | | |
| **Lecturer(s)**: | *David McQuaid* | | |
| **Issue Date:** | *13th May 2024* | | |
| **Submission Deadline Date:** | *31st May 2024* | | |
| **Late Submission Penalty:** | Late submissions will be accepted up to **5** calendar days after the deadline. All late submissions are subject to a penalty of **10%** of the mark awarded.  Submissions received more than 5 calendar days after the deadline above **will not** be accepted and a mark of 0% will be awarded. | | |
| **Method of Submission:** | **Moodle** | | |
| **Instructions for Submission:** | *Assessment must be submitted before 11.55pm 31st May 2024 as a Jupyter Notebook file*  *You must include your dataset used for semantic analysis* | | |
| **Feedback Method:** | **Results released on Moodle** | | |
| **Feedback Date:** | *After the Exam Boards June 2024* | | |

**Learning Outcomes:**

Please note this is not the assessment task. The task to be completed is detailed on the next page.

This CA will assess student attainment of the following minimum intended learning outcomes:

MLO 2 - Understand how to use text analytics for AI with the inclusion of labelled and

unlabelled data.

(Linked to PLO 3 (Stage 4 SLO 3))

MLO 4 - Understand Neural Networks and how this relates to the type of machine learning directed at a proposed A.I. implementation for a chosen domain.

(Linked to PLO 3 (Stage 4 SLO 3))

Attainment of the learning outcomes is the minimum requirement to achieve a Pass mark (40%). Higher marks are awarded where there is evidence of achievement beyond this, in accordance with QQI *Assessment and Standards, Revised 2013*, and summarised in the following table:

|  |  |  |
| --- | --- | --- |
| **Percentage Range** | **CCT Performance Description** | **QQI Description of Attainment** |
| **Level 6, 7 & 8 awards** |
| 90% + | Exceptional | Achievement includes that required for a Pass and in **most** respects is significantly and consistently beyond this |
| 80 – 89% | Outstanding |
| 70 – 79% | Excellent |
| 60 – 69% | Very Good | Achievement includes that required for a Pass and in **many** respects is significantly beyond this |
| 50 – 59% | Good | Achievement includes that required for a Pass and in **some** respects is significantly beyond this |
| 40 – 49% | Acceptable | Attains all the minimum intended programme learning outcomes |
| 35 – 39% | Fail | Nearly (but not quite) attains the relevant minimum intended learning outcomes |
| 0 – 34% | Fail | Does not attain some or all of the minimum intended learning outcomes |

Please review the CCT Grade Descriptor available on the module Moodle page for a detailed description of the standard of work required for each grade band.

The grading system in CCT is the QQI percentage grading system and is in common use in higher education institutions in Ireland. The pass mark and thresholds for different grade bands may be different from what you have experience of in the higher education system in other countries. CCT grades must be considered in the context of the grading system in Irish higher education and not assumed to represent the same standard the percentage grade reflects when awarded in an international context.

**Assessment Task**

Students are advised to review and adhere to the submission requirements documented after the assessment task.

**Scenario:**

This CA consists of 2 parts. Neural Networks and Semantic analysis.

**Neural Networks**

You are required to take data stored file “BankRecords.csv”, available on Moodle and process the data into a DataFrame.

You are then required to train a Neural Network to predict the “Income(Thousands’)” of the customers, including tuning the network to achieve the best results.

You must also compare your neural network to a standard ML regressor of your choosing and discuss your findings in the context of the problem at hand.

**Semantic Analysis**

You are required to source text data from any social media platform on any topic that you choose and perform semantic analysis on the text. This analysis should provide a visualization of the overall sentiment of your text data , showing the positive, neutral, and negative sentiment expressed. You will require at least 1000 text observations.

**Acceptable and Unacceptable Use of AI**

|  |  |
| --- | --- |
| **Acceptable and Unacceptable Use of AI** | * The use of generative AI tools (e.g. ChatGPT, Dall-e, etc.) is permitted in this assignment for the following activities:   + Brainstorming and refining your ideas;   + Fine tuning your research questions;   + Finding information on your topic;   + Drafting an outline to organise your thoughts; and   + Checking grammar and style. * The use of generative AI tools is not permitted in this course for the following activities:   + Impersonating you in classroom context   + Completing group work that your group has assigned to you   + Writing a draft of a writing assignment   + Writing entire sentences, paragraphs, papers, code fragments, functions, scripts to complete class assignments. * You are responsible for the information you submit based on an AI query. Your use of AI tools must be properly documented and cited. * Any assignment that is found to have used generative AI tools in an unauthorised way will be subject to college disciplinary procedures as outlined in the QA Manual. * When in doubt about permitted usage, please ask for clarification. |

**Requirements**

* Perform any preparation of the data (for both the Neural Network and Semantic Analysis data) , that you feel is necessary, using python in your Jupyter notebook. Explain your rationale behind your data preparation and how it will assist you.**[0-30]**
* Create and implement at least 1 Regression algorithm and 1 Neural Network that will output a Prediction of the “Income(Thousands’)” of the customers. Test these models and try to improve it as you see fit discussing your rationale. Discuss your findings and final rational for choosing a particular Regression algorithm.**[0-40]**
* Make a Prediction of a new customers income, who is not available in the original dataset **[0-5]**
* Perform Sentiment analysis on your text data and discuss your findings in detail. **[0-20]**
* Visualize your sentiment analysis showing the positive, neutral, and negative sentiment expressed**[0-5]**

**Note**

* All written work MUST be completed in Jupyter Notebook Markdown (please review “Jupyter Notebook Tutorial” Notes in Moodle if you are unsure of this).
* All data wrangling, analysis, and visualizations must be generated using python.
* All Code must be included in code blocks (As normal). No other upload will be accepted.
* All written work MUST be detailed in your Jupyter Markdown (NOT in code comments).

**Submission Requirements**

All assessment submissions must meet the minimum requirements listed below. Failure to do so may have implications for the mark awarded.

All assessment submissions must:

* Be submitted before **11.55pm 31st May 2024** as a Jupyter Notebook file.
* The Jupyter Notebook File Must be saved as “YourName\_ML\_CA2.ipynb”, and the dataset you have used.
* You must include your dataset used for semantic analysis.
* Be submitted by the deadline date specified or be subject to late submission penalties
* Be submitted via Moodle upload
* Use [Harvard Referencing](http://40.115.124.2/sp/subjects/guide.php?subject=harvardref) when citing third party material
* Be the student’s own work.
* Include the CCT assessment cover page.

**Additional Information**

* Lecturers are not required to review draft assessment submissions. This may be offered at the lecturer’s discretion.
* In accordance with CCT policy, feedback to learners may be provided in written, audio or video format and can be provided as individual learner feedback, small group feedback or whole class feedback.
* Results and feedback will only be issued when assessments have been marked and moderated / reviewed by a second examiner.
* Additional feedback may be requested by contacting Dept of Computing Manager,Additional feedback may be provided as individual, small group or whole class feedback. Lecturers are not obliged to respond to email requests for additional feedback where this is not the specified process or to respond to further requests for feedback following the additional feedback.
* Following receipt of feedback, where a student believes there has been an error in the marks or feedback received, they should avail of the recheck and review process and should not attempt to get a revised mark / feedback by directly approaching the lecturer. Lecturers are not authorised to amend published marks outside of the recheck and review process or the Board of Examiners process.
* Students are advised that disagreement with an academic judgement is not grounds for review.
* For additional support with academic writing and referencing students are advised to contact the CCT Library Service or access the [CCT Learning Space](http://learningspace.cct.ie/subjects/index.php).
* For additional support with subject matter content students are advised to contact the [CCT Student Mentoring Academy](https://moodle.cct.ie/mod/forum/view.php?id=55148)
* For additional support with IT subject content, students are advised to access the [CCT Support Hub](https://moodle.cct.ie/course/view.php?id=1861).