TORRENS UNIVERSITY AUSTRALIA

ASSESSMENT 2 BRIEF

Subject Code and Title	ISY503 Intelligent Systems
Assessment	Programming Task
Individual/Group	Individual
Length	Code + Manual (500 words +/- 10%)
Learning Outcomes	The Subject Learning Outcomes demonstrated by successful completion of the task below include: a) Determine suitable approaches towards the construction of AI systems. c) Apply knowledge based or learning based methods to solve problems in complex environments that attempt to simulate human thought and decision making processes, allowing modern society to make further advancements. e) Apply the foundational principles of AI learnt throughout the course and apply it to the different areas of Natural Language Processing, Speech Recognition, Computer Vision and Machine Learning.
Submission	Due by 11:55pm AEST Sunday end of Module 8 (Week 8).
Weighting	35%
Total Marks	100 marks

Task Summary

This individual assessment provides you an opportunity to explore the impact of applying various Machine Learning techniques on a dataset in a sandbox environment. You will complete the Programming Exercise from Google that will introduce you to modelling in the Machine Learning world. Note that this exercise is limited to exploring the application of Linear Regression in great detail, however, the feature engineering, transformations and hyperparameter tuning involved in applying different implementations of the regression algorithm are investigated. There is an emphasis on understanding the impacts of various feature transformations as well. Although a simple data set has been provided in this task, there will be opportunity to apply normalisation techniques.

You should follow the task instructions set out in the Google lab to setup and run the various libraries and environments as well as loading the dataset. The instructions will take you through various tasks including identifying different applicable ML models, appropriate hyperparameter and feature transformation exploration. While writing your own models, think outside the box and see if a custom ML model can be made. As there is no "one right answer" to this task, the assessment is seeking to help you explore the impacts of various possible options to further your own understanding. The task instructions and rubric outline in detail what each grade assigned to students will demonstrate.

The assessment also requires you to write a manual of approximately 500 words, explaining the models and ML techniques utilised, what impact they had on the data exploration and visualisation task and provide an evaluation of their efficiency. Once again, this is an exploration task and your analysis and conclusions of the effectiveness of various models you've investigated will be the subject of the marking criteria.

Context

This assessment presents an opportunity to study Linear Regression as the ML technique to study in more detail. The various models available to solve a regression problem form a significant part of a data scientist's toolset as many use cases in academia and industry will need knowledge on how to process datasets where a continuous range is used to predict outcomes. This task will help you explore different models available to solve such problems, utilise the wide variety of properties including hyperparameters and techniques like feature engineering to help you understand the impact of these changes. You will also be presented with an opportunity to visualise and analyse the impact of your changes.

Task Instructions

You need a Google account to do this assessment. You can create a free Google account here: https://myaccount.google.com/.

Once created, you need to navigate to the Google created lab: "Intro to Modelling" here: https://colab.research.google.com/github/google/eng-edu/blob/master/ml/fe/exercises/intro_to_modeling.ipynb?utm_source=ss-data-prep&utm_campaign=colab-external&utm_medium=referral&utm_content=intro_to_modeling

In addition to following the instructions outlined in the lab, you must:

- Implement a possible solution to each of the tasks outlined in the lab
- Add appropriate comments to your code created, following machine learning best practices for clean coding: https://towardsdatascience.com/clean-machine-learning-code-bd32bd0e9212
- Identify various different models that would be appropriate to use as alternatives for the tasks presented by the lab by varying hyperparameters and features. There is also an opportunity for you to create your own custom model by using different regressor functions

within TensorFlow. For more details, see: https://www.tensorflow.org/tutorials/customization/basics

- Familiarise yourself with the assessment's rubric to understand how the various assignment grades are assigned.

- Produce a manual of 500 words in length outlining:

    o The answers to the questions posed in each of the tasks within the lab.
    o The choice of models you made during your assessment including the various hyperparameters you chose and feature engineering performed for the appropriate task.
    o An analysis of the various models created and an evaluation of their efficiency.

Referencing

It is essential that you use appropriate APA style for citing and referencing research. Please see more information on referencing here http://library.laureate.net.au/research_skills/referencing

Submission Instructions

Submit your code as a Jupiter notebook (.ipnyb) and manual (as a .docx) via the **Assessment 2** link in the main navigation menu in ISY503: Intelligent Systems. The Learning Facilitator will provide feedback via the Grade Centre in the LMS portal. Feedback can be viewed in My Grades.

Academic Integrity Declaration

Please be aware of the Torrens University Australia Academic Integrity Policy and Procedure viewable online at http://www.torrens.edu.au/policies-and-forms

TORRENS UNIVERSITY AUSTRALIA

Assessment Rubric

Assessment Attributes	Fail (Yet to achieve minimum standard) 0-49%	Pass (Functional) 50-64%	Credit (Proficient) 65-74%	Distinction (Advanced) 75-84%	High Distinction (Exceptional) 85-100%
Model exploration **Percentage for this criterion = 40%**	Only one model has been trained with suggestions taken up from the lab directly.	Attempts have not been made to train the model with great variety and sample code and instructions have been blindly followed in most tasks.	An attempt has been made to create better models by using numerical and categorical models alone (including only using the same regressor function DNNRegressor outlined in the lab)	A sound attempt has been made to create multiple alternative models different to the suggestion from the lab using alternative regressors like LinearRegressor or DNNLinearCombinedRegressor or an attempt has been made to create student's own estimator.	Multiple working, comprehensive models have been attempted than the one suggestion from the lab using alternative regressors like LinearRegressor or DNNLinearCombinedRegressor and an attempt has been made to create student's own custom estimator.
Hyperparameter Tuning **Percentage for this criterion = 20%**	The project does not use hyperparameter tuning or feature transformation.	Basic attempts have been made to retune hyperparameters and feature transformations by following lab instructions.	Corresponding hyperparameters and features used in both models reflects a true effort to "train different models and visualise their impacts".	Corresponding hyperparameters and features used in all models attempted reflect a considered effort to "train different models and visualise their impacts".	Corresponding hyperparameters and features used in all models attempted reflect a sound understanding in the need to "train different models and visualise their impacts".

Assessment Attributes	Fail (Yet to achieve minimum standard) 0-49%	Pass (Functional) 50-64%	Credit (Proficient) 65-74%	Distinction (Advanced) 75-84%	High Distinction (Exceptional) 85-100%
Coding Proficiency **Percentage for this criterion = 20%**	Code has not utilised efficient practices as outlined in the Task Instructions	Code uses some Machine Learning Coding principles in the attempts made to retune models with code comments or following one of the design principles outlined within the Task Instructions.	Code uses sound Machine Learning Coding principles in the attempts to retune models with code comments and following more than one of the design principles outlined within the Task Instructions. Explanations of reasons behind choice, hyperparameters and features are explained within comments.	Code follows Machine Learning Coding principles very well in the attempts to retune models as well as new models created with code comments and following many of the design principles outlined within the Task Instructions. Explanations of reasons behind choice of regressors, hyperparameters and features are explained within comments.	Code is very efficient using the best practices including DRY principle, exception handling, no vague variable or method name, well-commented and well-organised (data preparation, model training and testing and applying, model analysis and accuracy displaying) Code follows Machine Learning Coding principles very well in the attempts to retune models as well as new models created with code comments and following many of the design principles outlined within the Task Instructions. Code comments in new models implemented use new wording completely different to the initial lab tasks i.e. no wording from the initial sample model is reused. Explanations of reasons behind choice of regressors,

Assessment Attributes	Fail (Yet to achieve minimum standard) 0-49%	Pass (Functional) 50-64%	Credit (Proficient) 65-74%	Distinction (Advanced) 75-84%	High Distinction (Exceptional) 85-100%
					hyperparameters and features are comprehensively explained within comments.
Manual Details **Percentage for this criterion = 20%**	No manual is provided or the manual is a direct copy of the code lab provided.	The manual contains a reflection of the hyperparameters tuned. A basic attempt at an analysis of the model and its efficiency is mentioned.	The manual outlines the reasons behind choosing different hyperparameters and features within the model as well as a sound analysis of the various models and their efficiency.	The manual outlines the reasons behind choosing different regressors, hyperparameters and features within the model including an opinion of which model and their efficiency is better. The wording reflects that the student understands the significance of the changes made through their analysis of the visualisation plots.	Manual is comprehensive consisting instructions on how to run the program, the expected outputs, and has explained the model and technique in details. The manual comprehensively outlines the reasons behind choosing different regressors and features within the model including an opinion of which model and their efficiency is better. The wording comprehensively reflects that the student understands the significance of the changes made through their analysis of the visualisation plots.

The following Subject Learning Outcomes are addressed in this assessment	
SLO a)	Determine suitable approaches towards the construction of AI systems.
SLO c)	Apply knowledge based or learning based methods to solve problems in complex environments that attempt to simulate human thought and decision making processes, allowing modern society to make further advancements.
SLO e)	Apply the foundational principles of AI learnt throughout the course and apply it to the different areas of Natural Language Processing, Speech Recognition, Computer Vision and Machine Learning.