ePortFolio web page address: https://nithya8483.github.io/ePortfolio/

Github source url: https://github.com/nithya8483/ePortfolio

Individual reflection

1.0 Introduction

This module, Machine Learning (ML) is one of the core modules in the course Artificial Intelligence (AI). Throughout this module, I devoted much time to learn about the basics and equip myself to independently work on practical aspects of ML. This report provides a reflective piece of my learning in this module (UoEO Short Guide to Reflective Writing).

Summary of learning outcomes:

1. Articulate the legal, social, ethical and professional issues faced by machine learning professionals.
2. Understand the applicability and challenges associated with different datasets for the use of machine learning algorithms.
3. Apply and critically appraise machine learning techniques to real-world problems, particularly where technical risk and uncertainty is involved.
4. Systematically develop and implement the skills required to be effective member of a development team in a virtual professional environment, adopting real-life perspectives on team roles and organisation.

2.0 Knowledge gained throughout this module

Throughout this module, I have acquired the following data analysis and ML models knowledge and hand-on experience on with Python libraries:

- Good understanding of the exploratory data analysis, visualisation of univariate, bivariate and multivariate relationship between variables through boxplot, scatterplot, histograms, bar charts, acquire descriptive statistical data of variables, discover correlation and regression between variables through correlation heat map, thereby discover pattern, trends, spotting missing values and testing hypothesis, dimensionality reduction, estimating Skewness and Kurtosis, categorical feature exploration, etc

- Importance of correlation and regression in ML, quantifying correlation with covariance and Pearson correlation coefficient, using linear regression analysis and logistic regression for classification problems, using Scikit-Learn to model a linear relationship, developing a multivariate linear regression model.

- Experience in importing Python libraries and coding using Jupyter notebook.

- Learned about different ML algorithms: clustering, artificial neural network (ANN) and convolutional neural network (CNN).

- Clustering: segmentation technique used to partition data using base and descriptor variables, Similarity/dissimilarity measurements using Euclidean distance, simple matching coefficient, Jaccard coefficient, cosine similarity, Pearson correlation coefficient and Bregman divergence, types of clusters and clustering methods such as k-means, hierarchial agglomerative clustering, DBSCAN, cluster evaluation methods such as cohesion and separation with Silhouette method.

- ANN: function of perceptron, binary classification with OR, AND, XOR operator, Multi-layer perceptron with hidden layers, use of functions such as activation function, weight updating, error functions, gradient decent, back

propagation for error reduction, delta rule, calculating bias and applications of ANN.

- CNN: Object detection by feature learning with CNN and classification with ANN, multi-channel filtering with weighted kernels in convolutional layer, activation rectifier function for detection of features, average or maximum pooling, flattening with multi-layer perceptron with back propagation, softmax classification.

- Model selection and evaluation: avoiding overfitting and under fitting, bias variance problem, resampling methods to estimate model performance such as random train/test split, cross-validation and bootstrap methods, k-fold and leave one out cross-validation, performance metrics sensitivity and specificity.

- Industry 4.0 revolution, descriptive, predictive, prescriptive machine learning model used in industry 4.0.

3.0 Activities carried out

In this module, I carried out these activities both independently and as a team to acquire the knowledge detailed in section 2.0 that helped me to achieve all the 4 learning outcomes mentioned in section 1.0.

Reading: I read the relevant reading material in the reading list and made notes in each unit, which provided the basis of my learning.

Lectures: I actively attended and made notes on the following lecture casts which were very valuable in understanding principal points in each unit.

Lecture casts	
Unit 1	Introduction to Machine Learning
Unit 3	Correlation and Regression
Unit 5	Clustering
Unit 7	Artificial Neural Network
Unit 8	Backpropagation of Artificial Neural Network
Unit 9	Convolutional neural network
Unit 11	Model selection and evaluation

Seminars: I watched offline Tutor-led seminars which were very useful in all aspects, especially, in the eportfolio activities and the team projects.

ePortFolio and seminar preparation activities: I independently performed the eportfolio and seminar preparation activities in each unit. These activities provided me confidence to independently carry out data preprocessing, data visualization and EDA, ML modelling with clustering and CNN in a very practical way. Screenshots of activities performed attached in Appendix section

Team activities: As a team member, I shared the work required for team projects and individually performed the work and discussed periodically with team mates

Individual contribution to team activities:

Team project 1: I contributed towards action plan for project completion outlining the tasks required to carry out. I performed data preprocessing such as encoding categorical variable with numerical values, created correlation heat map to find out the variables with high correlation and carried out the clustering model for the dataset. I also provided valuable points for the discussion and conclusion.

Team project 2: I prepared the project plan and created first draft of the presentation, detailing the tasks required for the project. I prepared text for the introduction, about CNN, contributed to data pre-processing, neural network modelling, optimising

model parameters such as kernel size, number of kernels and convolutional layers, and early stopping for better model accuracy and to discussion and conclusion.

Tutorials: I watched several YouTube tutorials on topics such as EDA, CNN, ANN and model evaluation and read about python libraries.

Discussion with others: I had verbal discussion with my team mates and work colleagues about data analysis tools and python libraries.

4.0 Learning reflection

Throughout this module the learning materials and all the activities were very engaging and provided me great insight to machine learning models. Being a Radiotherapy Physicist, I have previously learned about convolutional model based algorithm in radiotherapy treatment planning systems such as the pencil beam algorithm, collapsed cone algorithm, analytical anisotropic algorithm using backpropagation, gradient decent methods to simulate radiation dose calculation in patient CT images (Oncology Medical Systems). I found this has a close analogy with CNN in ML and this module helped me to understand the basics of these algorithms very well, thereby had a positive impact on my professional front.

In terms of coding, I have learned more new python libraries and confident to use them for any given problem.

I feel I am fully equipped to build ML models for any given data independently. Also I am confident to talk about ML algorithm to professionals in any field.

5.0 References

UoEO Short Guide to Reflective Writing. Available from: https://www.my-course.co.uk/course/view.php?id=13&section=4 [Accessed 12 December 2022]

Oncology Medical Physics, Dose calculation algorithms Available from: https://oncologymedicalphysics.com/dose-calculation-algorithms/ [Accessed 12 December 2022]

8.0 Appendix:

Feedback from peers

Feedback from Sandip Biswas, one of my team members:

Nithya has been a team player by presenting her ideas and solutions to problems. Her approach towards the Exploratory Data Analysis while resolving the Airbnb business case moved the analysis towards an informative closure. Further down, another group project about Neural Analysis using CNN and resolving the overfitting issue gave us a breather to ensure that we could conclude with a valid summary and information.

Nithya made sure during the group discussions and assignments that they were available before the next teamwork so that they could be reviewed for the next step.

Feedback from Vasilisa Lukashevich, one of my team members:

Nithya Kanakavelu offered a couple of good additional ideas for preventing overfitting the model while we were working at the CNN team project. Actually, for both team projects she showed great support for every step: from the code-compiling to the presentation itself. 10:47

Evidence of activities done:

All the evidences of individual and team activities performed are included in the ePortfolio. Please see the link to ePortfolio and Github source url in Page 1. Below are the screenshots of ePortfolio Machine Learning module page.
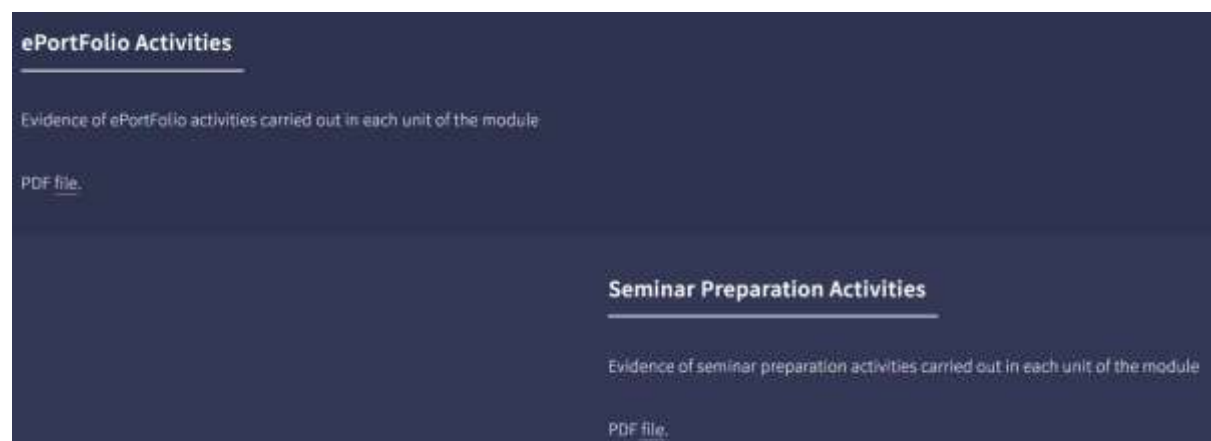
Professional skills matrix

Skills acquired	Evidence	Learning outcome covered
Time management	Team projects deadlines	1,2,3,4
Commercial awareness	Article review in ePortfolio and seminar preparation activities	1,2,3,4
Critical thinking and analysis	Team projects and ePortfolio and seminar preparation activities	1,2,3,4
Communication and Literacy skills	Team projects and discussions	1,2,3,4
IT and digital	Team projects and ePortfolio and seminar preparation activities	1,2,3,4
Numeracy	Team projects and ePortfolio and seminar preparation activities	1,2,3,4
Research	Team projects and ePortfolio and seminar preparation activities	1,2,3,4
Interpersonal	Team presentation	1,2,3,4
Problem solving	Team projects and ePortfolio and seminar preparation activities	1,2,3,4
Ethical awareness	Reading materials	1,2,3,4
Teamwork	Team projects	1,2,3,4
Critical reflection	Reflective writing	1,2,3,4