

Self-Assessment Details

The following checklists provide an overview of my self-assessment for this unit.

	Pass (P)	Credit (C)	Distinction (D)	High Distinction (HD)
Self-Assessment (please tick)				✓

Self-assessment Statement

	Included (please tick)
Learning summary report	✓
Weekly tasks signed off	✓

Pass Checklist

	Included (please tick)
Assignment (bird species classification) signed off	✓

Credit Checklist, in addition to Pass Checklist

	Included (please tick)
Project (face recognition) submitted	✓

Minimum Distinction / High Distinction Checklist, in addition to Credit Checklist

Declaration

I declare that this portfolio is my individual work. I have not copied from any other student's work or from any other source except where due acknowledgment is made explicitly in the text, nor has any part of this submission been written for me by another person.

Signature: Tran Duc Anh Dang

Introduction

This report summarises what I learnt in COS30082 Applied Machine Learning. It includes a self-assessment against the criteria described in the unit outline, a justification of the pieces included, details of the coverage of the unit's intended learning outcomes, and a reflection on my learning.

Overview of Pieces Included

This section outlines the pieces that I have included in my portfolio.

1. **lab01:**
 - Predicting Engagement
 - Predicting Insurance Premiums
 - Lab01.ipynb
 - This tutorial covers data visualization, variable encoding, decision trees, linear regression, polynomial regression, random forests, and SVM.
2. **lab02:**
 - RegressionModel.ipynb
 - HousePricePrediction.ipynb
 - This tutorial covers linear regression analysis, types of linear regressions, optimization techniques in regression analysis, and regularization.
3. **lab03:**
 - Gender Prediction
 - Titanic Survival Prediction
 - This tutorial covers logistic regression analysis, cost functions, gradient descent types, and decision boundaries for binary classification.
4. **lab04:**
 - Basic SVM
 - SVM with Kernels
 - This tutorial covers support vector machines (SVM), artificial neural networks (ANN), their unique features, optimization algorithms, and use cases for choosing between neural networks and SVMs.
5. **lab05:**
 - cnn.ipynb
 - overfit_and_underfit.ipynb
 - This tutorial covers the theoretical and mathematical concepts of convolutional neural networks (CNN), types of CNNs, practical design processes for deep learning, and strategies for debugging training issues.
6. **lab06:**
 - transfer_learning.ipynb
 - This tutorial covers the concept of transfer learning, different types of transfer learning methods, their functions, the training process for various techniques, and the applicability of transfer learning.
7. **lab07:**
 - ObjectDetection.ipynb
 - This tutorial covers the concept of object detection, different types of object detection algorithms, their functions, mathematical optimization algorithms, and implementation of object detection techniques using TensorFlow API, Ultralytics.
8. **lab08:**
 - facematching.ipynb
 - This tutorial covers the face recognition problem, the difference between face verification and face identification, the automatic face recognition process, the algorithms used, and implementing face recognition using deep learning models.
9. **lab09:**
 - GAN.ipynb
 - This tutorial covers the difference between discriminative and generative models, the mathematical optimization algorithm of GANs, different variants of GANs, and their applications.

10. lab10:

- Airline Tweets Sentiment Analysis.ipynb
- FinanceSentiment.ipynb
- This tutorial covers the fundamentals of Natural Language Processing (NLP), including text classification, machine translation, speech recognition, chatbots, information extraction, text summarization, challenges in NLP, history of NLP, and key concepts like tokenization, stemming, lemmatization, part of speech tagging, named entity recognition, syntax trees, semantic analysis, word embeddings, and N-grams.

11. lab11:

- 3_prompt_engineering.ipynb
- This tutorial covers the fundamentals of large language models (LLMs), including their architecture, pre-training and fine-tuning processes, different categories (autoregressive, autoencoding, and combined models), key terminologies, transfer learning, popular models like BERT, GPT, and T5, and their applications in various NLP tasks.

12. Assignment (for students who are aiming for a credit grade or higher):

- This assignment involves using transfer learning with ResNet-50, finetuning, dropout, selective layer training, and data augmentation techniques. The goal is to classify bird species, employing techniques to handle overfitting and improve model performance through various regularization methods, evaluation metrics, and optimization strategies.

13. HD Project

- This project showcases an advanced face recognition attendance system with integrated anti-spoofing measures, leveraging state of the art machine learning techniques and innovative methodologies to achieve high accuracy, security, and user friendly interaction.

Coverage of the Intended Learning Outcomes (ILO)

This section outlines how the pieces I have included above demonstrate the depth of my understanding in relation to each of the unit's intended learning outcomes.

ILO 1. Explain machine learning life cycle

I have demonstrated through comprehensive tutorials covering model selection, training, validation, and deployment processes across various labs.

ILO 2. Use appropriate data engineering techniques for data preparation

I have illustrated these by data preprocessing steps in tutorials, including variable encoding, data augmentation, and transformation techniques.

ILO 3. Analyse and apply advanced machine learning algorithms to solve real-world problems

I have showcased how I analysed and apply advanced ML algorithms in lab projects and assignments, utilizing algorithms such as SVM, CNN, GAN, and transfer learning for tasks like prediction, classification, and recognition.

ILO 4. Evaluate, deploy and optimise machine learning project outcomes to domain specific users

The evidence that I have is in the BirdClassifier assignment and HD Project, where model performance is optimized and tailored to specific applications with detailed evaluation metrics.

ILO 5. Interpret and effectively communicate machine learning project outcomes to domain specific users

I achieved through detailed reports and visualizations in the provided tutorials and assignments, effectively communicating results and insights.

Reflection

The most important things I learnt:

I gained a deep understanding of ML algorithms, data preprocessing techniques, and model evaluation methods. Additionally, I learned to implement complex models such as CNNs, GANs, and LLMs, and the practical aspects of deploying these models.

The things that helped me most were:

The hands on lab exercises and assignments were invaluable in reinforcing theoretical knowledge through practical application. Access to detailed tutorials and real world datasets also played a crucial role in my learning.

I found the following topics particularly challenging:

While I did not find the lab tasks particularly challenging due to my extensive experience, optimizing for triplet loss training in the HD project was tricky and required careful tuning to achieve the desired results.

I found the following topics particularly interesting:

I found the topics of transfer learning and large language models (LLMs) especially fascinating. The ability to leverage pretrained models and adapt them to specific tasks opened up new possibilities in machine learning applications.

I feel I learnt these topics, concepts, and/or tools really well:

I feel confident in my understanding and application of convolutional neural networks (CNNs) and transfer learning techniques. The lab exercises on these topics provided ample practice and clarity.

I still need to work on the following areas:

As I've been working on my side project and I found out that there are a few things that I need to further improve my skills in finetuning large language models (LLMs) and optimizing their performance for specific tasks and also further improving on training RL Agent for Quantitative Trading. This area remains complex and requires more practice and deeper understanding.

This unit will help me in the future:

The knowledge and skills acquired in this unit will be invaluable in my future studies and career, particularly in the fields of AI and machine learning. Understanding advanced algorithms and practical deployment will enhance my capability to tackle real world problems.

If I did this unit again I would do the following things differently:

I would allocate more time to revising challenging concepts like GANs and LLMs or RL Agent. Additionally, I would engage more in discussions and seek feedback from peers and instructors to gain different perspectives and insights.

Other...:

The unit has significantly enhanced my problem-solving skills and my ability to approach machine learning projects systematically. It has also increased my confidence in using advanced tools and techniques in AI.