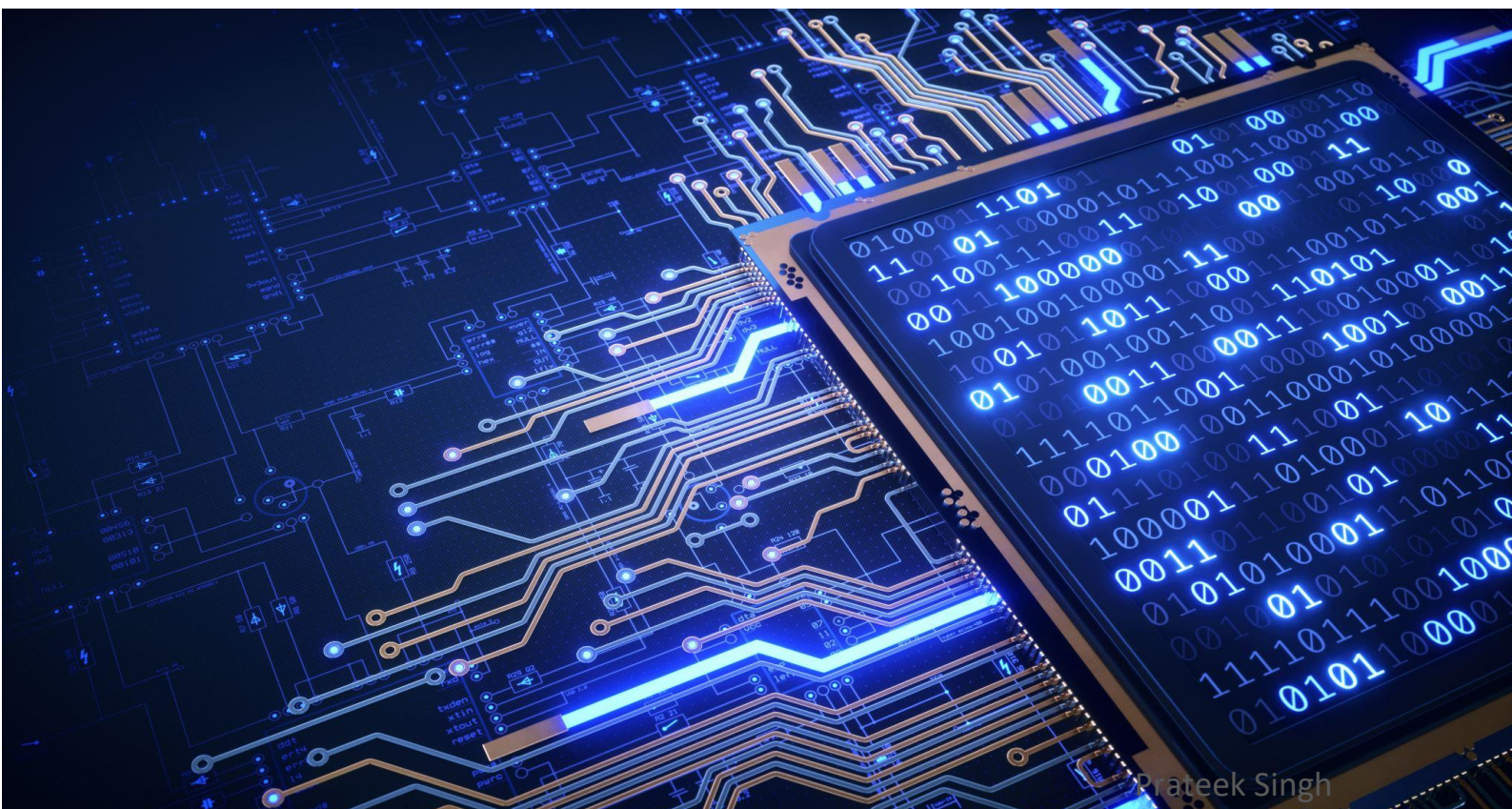


SIT799

Human Aligned Artificial Intelligence

Learning Summary Report



Prateek Singh

221218743

Self-Assessment Details

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

	Pass (D)	Credit (C)	Distinction (B)	High Distinction (A)
Self-Assessment				✓

Self-Assessment Statement

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: **Prateek Singh**

Portfolio Overview

Start with something like “This portfolio includes work that demonstrates that I have achieved all Unit Learning Outcomes for SIT799 Unit Title to a **High Distinction** level.”

I started this unit with a background of AI engineer and have learned essential concepts which are directly applicable to the solutions which I develop. I learned about developing AI solutions while considering the impact on society and ethical concerns related to it. This unit also allowed me to perform individual research while doing Distinction and High Distinction tasks, while the practical tasks allowed me to implement these concepts.

In initial tasks (Task 1.1 – Task 2.1), helped developing scripting and development skills related to AI and using relevant frameworks. These tasks allowed me to develop baseline solutions for developing basic AI tasks such as classification and allowed me to understand how these can be deployed to work on dummy datasets.

In theoretical tasks (2.2D, 3.1D, 8.1C), I performed literature survey to comprehensively sum up my understanding over presented use cases. This allowed me to pick up necessary research skills which will help take initiatives to overcome roadblocks in real world scenarios. The later practical tasks (3.2 HD, 5.2 D, 6.2HD), allowed me to develop hands-on skills to develop theoretical concepts using proof of concept examples.

While implementing these tasks I have also learned about the threats that can damage state-of-the-art systems (Task 4.1 and 4.2, adverse use of AI). AI explainability taught me the dangers of black box models and the importance of conveying underlying principles behind AI predictions. This not only makes the solution more transparent but also allows for better trained model to handle fail cases. Task 10.1 introduced me to everchanging job landscape due to the impact of AI and has made me future ready and is fundamental in picking up new employment skills.

I have shown hands-on as well as strong research capacity to look for methods beyond the coursework by conducting an independent literature survey. I believe this makes me a suitable candidate to achieve High Distinction.

Reflections

The most important things I learnt:

The most important thing which this unit has taught me is assessing the human angle involved while developing real world AI solutions. For example, in Task 6.2, AI explainability highlights why it is important for AI models to convey reasoning behind the predictions. This is a teaching which would enable me to be a better AI professional while interacting with individuals with various backgrounds.

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

This unit has made me confident about the ability to quickly prototype theoretical concepts by developing a minimal working example, this will help me convey my ideas with examples in a business setting.

I found the following topics particularly challenging:

The most challenging topic for me was task 5.2, where I had to implement a Deep Fake Comments Generator, I tried completing the task with the best of my ability but could have done better if I had a bit more time.

I found the following topics particularly interesting:

Adversarial Attacks and Explainability were the favorite topics for me, coming from a background of computer vision engineer, adversarial attacks have shown me how even state-of-the-art systems are prone to such attacks and the defensive measures required.

I still need to work on the following areas:

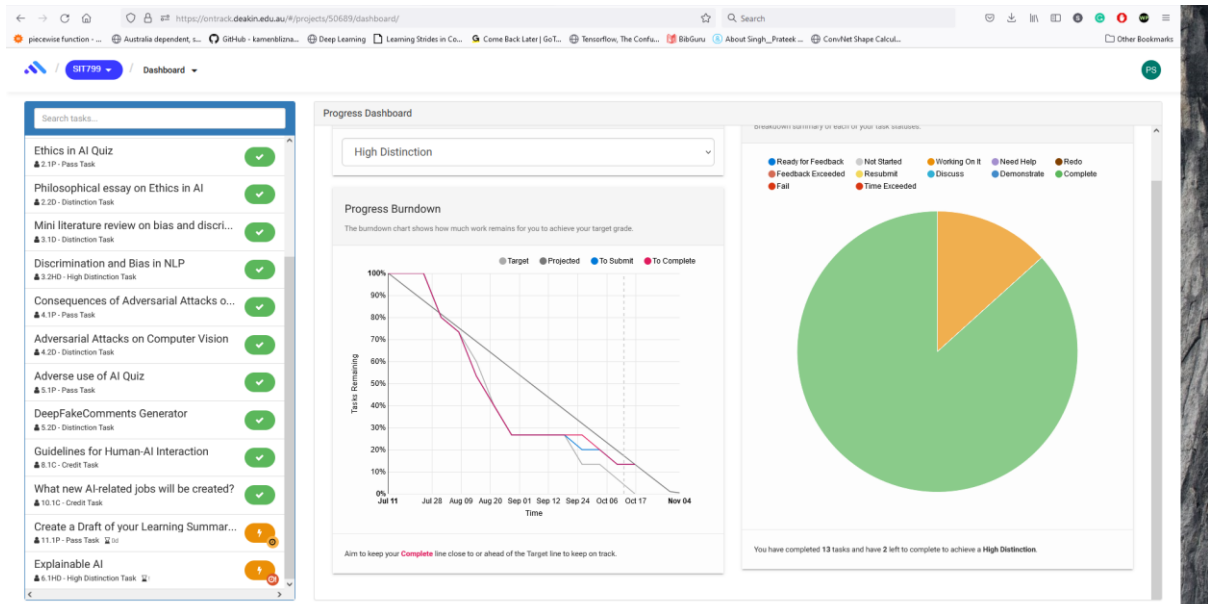
This unit has taught me concepts which are essential for me to take the next step towards being an AI leader in the industry, I would keep working on developing solutions which have a net positive impact on the society. This unit has also taught me the importance of effective time management.

The things that helped me most were:

The OnTrack tasks along with the workshops helped me the most during these 10 weeks, the tasks allowed me to go beyond the course to perform exploratory research.

My progress in this unit was ...:

The OnTrack graph shows that I have taken up tasks till the High Distinction Level and tried completing them to the best of my ability. This graph also shows that I could have done some tasks sooner.



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

Overall, there are very few things that I would do differently if I did this unit again. I would however want to manage my time more effectively so that I could do justice to all the assignments of this unit. Nevertheless, I am very satisfied by the content of this unit and real-world applications.