Journal Submission 2

Student name: Fathima Zaineb Ismath

Student ID: H00385662

Student email address: fzi2000@hw.ac.uk

Date of submission: 24 October 2024

Title of dissertation and brief description	Title: Predicting Cardiovascular Diseases using Machine Learning Description: This project aims to advance the prediction of heart disease through the application of machine learning algorithms, with an emphasis on integrating explainable artificial intelligence (XAI) techniques and risk stratification models. By incorporating XAI, the project seeks to enhance the interpretability of predictive models, enabling both healthcare professionals and non-specialists to better understand the underlying factors contributing to the predictions and support in accurate clinical decision-making and treatment. Description (Updated) This project aims to advance the prediction of heart disease through the application of machine learning algorithms, deep learning techniques with an emphasis on integrating explainable artificial intelligence (XAI) techniques and risk stratification models. By incorporating XAI, the project seeks to enhance the interpretability of predictive models, enabling both healthcare professionals and non-specialists to better understand the underlying factors contributing to the predictions and support in accurate clinical decision-making and treatment.
Communicating with supervisor	*I messaged Professor Cristina about the Ethics form submission and Supervisor agreed for an online meeting where she guided me with the submissions.* *After I submitted the Ethics form, the supervisor approved.* *Expressed interest in exploring both machine learning techniques and deep learning. Supervisor encouraged me to do it provided it doesn't complicate the dissertation, owing to other courses and integration of explainable AI.* *I informed the supervisor that I'll be submitting a draft of some chapters towards the end of October.*
References consulted	*New references consulted between Week 4 to Week 7.* • Tsao, C.W. *et al.* (2023) 'Heart Disease and Stroke Statistics—2023 Update: A Report From the American Heart Association,' *Circulation*, 147(8). https://doi.org/10.1161/cir.0000000000001123. (Tsao *et al.*, 2023) • *Heart Disease Facts* (2024). https://www.cdc.gov/heart-disease/data-research/facts-stats/index.html. • World Health Organization: WHO (2019) *Cardiovascular diseases*. https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1 (Accessed: 11 October, 2024) • Ouyang, S. (2022) 'Research of Heart Disease Prediction Based on Machine Learning,' *2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)* [Preprint]. https://doi.org/10.1109/aemcse55572.2022.00071. • Lee, YG., Oh, JY., Kim, D. et al. (2023) *SHAP Value-Based Feature Importance Analysis for Short-Term Load Forecasting*. J. Electr. Eng. Technol., 18, 579–588. https://doi.org/10.1007/s42835-022-01161-9. • Yassine, I., Mousa, D. & Zayed, N., 2014. *Automatic Cardiac MRI Localization Method*. Proceedings of the 7th Cairo International Biomedical Engineering

	Conference, CIBEC 2014. doi:10.1109/CIBEC.2014.7020942.
Tools explored/used	• *Github* • -Jupyter Notebook for practicing data cleaning and feature selection. • Overleaf latex editor • Anaconda python environment. • Libraries such as numpy, scikit-learn, pandas and other data mining libraries were explored. • Experimented by creating a basic model that can denote 0 and 1 for predicting heart disease. • Used TensorFlow library, Gradio, Shapley • *Experimented with sklearn.metrics and evaluation strategies such as accuracy, precision etc* • *Logistic Regression and random forests.*
Other work carried out	• *Submitted the Ethics form for approval* • *Combined datasets from UCI and integrated into a single excel file* • *Expressed interest in exploring deep learning models.* • *Completed writing project journal 2.* • *Explored various python libraries available online and referred to similar implementation on GitHub.* • *Refined Chapters 1 and 2.* • *Drafted Chapter 3,4 and 5.* • *Completed risk analysis and the Gantt Chart for project management which are included in the Appendix.* • *Formatted the report in latex editor*

Plan the next 2 to 3 weeks	• *Complete chapter 3,4 and 5 of D1 and submit the draft to the supervisor for reviewing by the end of week 8.*
	• *Complete the draft of the Requirements Analysis and Methodology chapter.*
	  - *Refine Functional and Non-Functional Requirements*
	  - *Connect them with the objectives using a traceability matrix.*
	  - *Formulate research questions for the project.*
	• *Modify and refine the draft using feedback provided by the supervisor.*
	• *Familiarize with deep learning and machine learning techniques.*
	• *Perform data preprocessing and understand the finalized dataset.*
	• *Finally, draft other sections of D1.*
	• *Draft the PLES section.*
	• *Add references.*
	• *Feedback Iterations: After getting feedback from the supervisor for the draft, add her insights and refine the report accordingly.*

Overall Reflection	*The deliverable 1 is progressing well. I have completed chapters 1 and 2 and refined them. Chapters 3,4 and 5 are drafted and will be refined for clarity. I converted the report from word format to latex for better presentation. The appendix section has been drafted. It has been a little challenging with all other courses having submissions within the same week, but with better time management plan, the D1 will successfully be completed.*
Additional section	*You may want to add a section relevant to your specific project.*
