

Journal Submission 1

Student name: Fathima Zaineb Ismath

Student ID: H00385662

Student email address: fzi2000@hw.ac.uk

Date of submission: 07 October 2024

Title of dissertation and brief description	<p>Title: Predicting Cardiovascular Diseases using Machine Learning</p> <p>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.</p>
Communicating with supervisor	<p><i>I messaged Professor Cristina expressing interest for her supervision for dissertation, shared the research proposal.</i></p> <p><i>Supervisor agreed, said she found the topic interesting and updated the allocation in the project system. We agreed on a preferred time and location for our first meeting.</i></p> <p><i>First meeting with Professor Cristina to get started with Dissertation (In person) – 20/09/2024.</i></p> <p><i>We discussed about the D1 approaches, professor explained the details about the ethics form, project handbook & asked me to find a dataset used in previously research papers. She suggested me to start writing from the background chapter of D1 and submit a draft to her by the end of Week 4 or start of Week 5. I presented my research done during summer on the proposed research topic. Supervisor advised to include 5-7 research papers due to word limit constraints.</i></p>
References consulted	<p><i>I have researched on some papers found online through Google Scholar and Discovery.</i></p> <p>Baghdadi, N.A. et al. (2023) 'Advanced machine learning techniques for cardiovascular disease early detection and diagnosis', <i>Journal of Big Data</i>, 10(1). doi:10.1186/s40537-023-00817-1.</p> <p>Bizimana, P.C. et al. (2024) 'Automated heart disease prediction using improved explainable learning-based technique', <i>Neural Computing and Applications</i>, 36(26), pp. 16289–16318. doi:10.1007/s00521-024-09967-6.</p> <p>Careervira (2023) Machine learning for beginners: A comprehensive 2023 guide, Medium. Available at: https://medium.com/@careervira.community/machine-learning-for-beginners-a-comprehensive-2023-guide-4ec02d5caab7 (Accessed: 22 September 2024).</p> <p>García-Ordás, M.T. et al. (2023) 'Heart disease risk prediction using Deep Learning techniques with feature augmentation', <i>Multimedia Tools and Applications</i>, 82(20), pp. 31759–31773. doi:10.1007/s11042-023-14817-z.</p> <p>Holdsworth, J. and Scapicchio, & Mark (2024) What is deep learning? IBM. Available at: https://www.ibm.com/topics/deep-learning (Accessed: 22 September 2024).</p> <p>"Hospitalization for Heart Attack, Stroke, or Congestive Heart Failure among Persons with Diabetes", Special report: 2001 – 2003, New Mexico</p>

	<p>Kavitha, M. <i>et al.</i> (2021) 'Heart disease prediction using Hybrid Machine Learning Model', <i>2021 6th International Conference on Inventive Computation Technologies (ICICT)</i> [Preprint]. doi:10.1109/iciict50816.2021.9358597.</p> <p>Krishnani, D. <i>et al.</i> (2019) 'Prediction of coronary heart disease using supervised machine learning algorithms', <i>TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)</i> [Preprint]. doi:10.1109/tencon.2019.8929434.</p> <p>Latha, C.B. and Jeeva, S.C. (2019) 'Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques', <i>Informatics in Medicine Unlocked</i>, 16, p. 100203. doi:10.1016/j.imu.2019.100203.</p> <p>Mohan, S., Thirumalai, C. and Srivastava, G., 2019. Effective heart disease prediction using hybrid machine learning techniques. <i>IEEE access</i>, 7, pp.81542-81554.</p> <p>Schlesinger, D.E. and Stultz, C.M. (2020) 'Deep Learning for Cardiovascular Risk Stratification', <i>Current Treatment Options in Cardiovascular Medicine</i>, 22(8). doi:10.1007/s11936-020-00814-0.</p> <p>Shah, D., Patel, S. and Bharti, S.K. (2020) 'Heart disease prediction using Machine Learning Techniques', <i>SN Computer Science</i>, 1(6). doi:10.1007/s42979-020-00365-y.</p> <p>Sudhakar, K. and Manimekalai, D.M., 2014. Study of heart disease prediction using data mining. <i>International journal of advanced research in computer science and software engineering</i>, 4(1), pp.1157-1160.</p> <p>UCI: Heart Failure Prediction Dataset. UCI Machine Learning Repository. (Accessed: October 2024). https://archive.ics.uci.edu/dataset/45/heart+disease</p> <p><i>Shankar et al. (2020) Heart Disease Prediction Using CNN Algorithm</i></p> <p><i>Multi Disease Prediction System using Random Forest Algorithm in Healthcare System</i></p> <p><i>A Systematic Review of Artificial Intelligence Models for Time-to-Event Outcome Applied in Cardiovascular Disease Risk Prediction</i></p> <p><i>Artificial intelligence in the risk prediction models of cardiovascular disease and development of an independent validation screening tool: a systematic review</i></p> <p><i>Predicting Heart Attack through Explainable Artificial Intelligence</i></p> <p><i>Automated heart disease prediction using improved explainable learning-based technique</i></p> <p><i>Optimized Ensemble Learning Approach with Explainable AI for Improved Heart Disease Prediction</i></p> <p>https://link.springer.com/article/10.1007/s00521-024-09967-6</p> <p>https://www.iraj.in/journal/journal_file/journal_pdf/6-71-140490825388-92.pdf</p> <p>https://ieeexplore.ieee.org/abstract/document/8971374</p> <p>https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6164626</p> <p>https://link.springer.com/article/10.1007/s11042-022-14305-w</p> <p>https://link.springer.com/article/10.1007/s00500-022-07788-0?fromPaywallRec=true</p> <p>https://ieeexplore.ieee.org/document/71003400</p>
--	--

Tools explored/used	<ul style="list-style-type: none"> • <i>Github</i> • <i>-Jupyter Notebook for practicing data cleaning and feature selection.</i> • <i>Overleaf latex editor</i> • <i>Anaconda python environment.</i> • <i>Libraries such as numpy, scikit-learn, pandas and other data mining libraries were explored.</i> • <i>Experimented by creating a basic model that can denote 0 and 1 for predicting heart disease.</i> • <i>Used TensorFlow library, Gradio, Shapley</i> • <i>Experimented with sklearn.metrics and evaluation strategies such as accuracy, precision etc</i> • <i>Logistic Regression and random forests</i>
Other work carried out	<ul style="list-style-type: none"> • <i>Structure of Chapter 1 and 2 is planned out.</i> • <i>Analysis of datasets from Kaggle and other resources.</i> • <i>Completed writing project journal 1.</i> • <i>Made an excel spreadsheet on relevant prior works on the topic.</i> • <i>Explored various python libraries available online and referred to similar implementation on GitHub.</i> • <i>Narrowed down the literature review to 7 research papers.</i> • <i>Wrote a rough draft for Chapter 2 Background and cited various sources using Harvard citation style.</i>
Plan the next 2 to 3 weeks	<ul style="list-style-type: none"> • <i>Finalize on the datasets, methodologies, research questions and objectives of the project.</i> • <i>Analyse existing literature and choose 5-7 research papers for critical analysis. Once the papers are chosen, complete writing the critical analysis and relevant work sections in Chapter 2.</i> • <i>Research on machine learning algorithms used in prior work and identify the gaps in existing literature.</i> • <i>Inform the supervisor about the dataset, methodologies and objectives of the project.</i> • <i>Complete chapter 2 of D1 and submit the draft to the supervisor for reviewing by the end of week 4 or week 5.</i> • <i>Modify and refine the draft using feedback provided by the supervisor.</i> • <i>Familiarize with deep learning and machine learning techniques.</i> • <i>Perform data preprocessing and understand the finalized dataset. Familiarize with the attributes and get an overview of the data.</i> • <i>Submit the ethics approval form in the project system.</i> • <i>Finally, draft other sections of D1.</i>
Overall Reflection	<p><i>The dissertation work is progressing well, with significant progress made in the initial stages of the project. Communication with my supervisor has been very productive and has helped me clarify the direction of the research. The project title, objectives, and a basic research proposal have been decided, and the supervisor provided valuable feedback during our first meeting, particularly in relation to structuring the dissertation and the ethical considerations involved.</i></p> <p><i>I have explored various tools, datasets, and machine learning libraries, successfully building a preliminary model for heart disease prediction. The exploration of Python libraries such as NumPy, scikit-learn, and TensorFlow has helped me gain hands-on experience in data preprocessing and model development. However, selecting the most appropriate dataset for the final project was a challenge due to the diversity of available datasets. After careful consideration to ensure the chosen dataset aligns</i></p>

	<p><i>with the project's goals, I have narrowed it down to 3 datasets.</i></p> <p><i>The literature review is coming together, with 12 research papers identified, and I have begun drafting Chapter 2 on the background, narrowing down the review to a smaller set of key papers to meet the word limit constraint. Time management has been a challenge, particularly balancing background research with hands-on technical work and other courses. However, breaking down tasks into smaller milestones, such as completing a rough draft of Chapter 2 by the end of Week 4, has helped in staying on track.</i></p> <p><i>Looking ahead, my primary focus will be on finalizing the dataset, methodologies, and research questions, as well as critically analyzing 5-7 research papers for the literature review. Ensuring that my supervisor is kept updated on the project's progress and incorporating feedback promptly will also be key to refining the work. The project will be fruitful through continued research and structured planning.</i></p>
Additional section	<p><i>You may want to add a section relevant to your specific project.</i></p>