

Machine Learning

End of Module Assessment

e-portfolio & Reflection

Murthy Kanuri

Student ID: 12696139

E-PORTFOLIO home page link

<https://m-kanuri.github.io/>

E- Portfolio Machine learning module link

<https://m-kanuri.github.io/Module3.html>

Table of Contents

1.	Introduction	3
2.	Technical Contributions (Units 6 & 11)	3
3.	Project evaluation (Units 6 & 11)	4
4.	Learning Reflections	4
5.	Experience as a Team Member	5
6.	Personal and Professional Development	5
7.	Team Learning.....	6
8.	Future Learning goals.....	6
9.	References.....	7
10.	Appendix A: e-Portfolio Links.....	8
11.	Appendix B : Development Plan	9

1. Introduction

I have learned a lot in the Machine Learning module thanks to the helpful teaching of my tutor, Liz Coulter-Smith. The lessons were engaging and motivating and gave me a solid foundation in Machine learning concepts.

We began exploring the fourth industrial revolution and its significance, followed by Exploratory data analysis (EDA), Correlation and regression techniques, linear regression with Scikit-learn, Clustering methods like K-means and hierarchical clustering, and Jaccard Coefficient, as well as Neural network architecture such as ANN (artificial neural networks) and CNN (Convolutional Neural Networks).

In addition to learning about machine learning's future, we also discovered the legal, social, ethical and professional issues faced by machine learning professionals. The course aligned with my career aspirations and provided hands-on experience, boosting my confidence in applying ML techniques to real-world scenarios.

2. Technical Contributions (Units 6 & 11)

In Unit 6, I actively contributed to the development project by implementing clustering algorithms using Python. I applied K-means clustering to segment the Airbnb dataset based on key features like price, location, and property type. I leveraged Python libraries such as pandas for data manipulation, matplotlib and seaborn for visualisation, and sci-kit for model training and evaluation. My contributions include feature selection and data preprocessing insights, improving the clustering model's accuracy.

In Unit 11, I successfully worked on Neural network models with Python-based code, using libraries such as TensorFlow and Keras to build and train deep learning models. At first, I had trouble finding the best settings for the model, which caused it to overfit. Adjusting things like the learning rate, the batch sizes, and the activation functions used in the model was also challenging, leading to results that were not as good as they could be. To address this, I experimented with different configurations, used

early stopping to prevent overfitting and used grid search for hyperparameter optimisation. This hands-on experience deepened my understanding of architecture design, hyperparameter tuning, and performance evaluation techniques such as accuracy and loss tracking.

3. Project evaluation (Units 6 & 11)

The primary difference between Unit 6 and Unit 11 is that Unit 6 is a team-based development project, whereas Unit 11 is an individual project. In Unit 6, we focused on clustering techniques for Airbnb Data analysis using Python libraries. In Unit 11, I independently worked on neural network design, focussing on architecture design and performance evaluation. Transitioning from a team project to an individual project allowed me to develop independent decision-making, problem-solving and technical skills. Also, during Unit 6, our understanding of machine learning was preliminary. In contrast, for Unit 11, there was confidence as we learnt many topics and discussed various techniques in depth.

If there are individual projects like Unit 11 in the upcoming modules, I need to allocate more time for analysis and execution.

4. Learning Reflections

I explored various ML algorithms during the machine learning module, including linear regression and convolutional neural networks (CNNs). In Unit 4, I learned how to use Scikit-learn for linear regression, focusing on creating and testing models and calculating performance metrics such as R-squared and Mean Squared Error (MSE). Unit 5 was about clustering techniques, where I learned how to group data effectively. In Unit 10, I worked with CNNs (Convolutional Neural Networks), experimenting with different settings (number of layers, filter sizes, activation functions, etc) and observing how they impacted the model's results. CNNs are more effective for image recognition tasks (Goodfellow, 2016) and clustering methods like K-means are suitable for customer segmentation.

5. Experience as a Team Member

In the development team project ("Airbnb business analysis using a data science approach"), I played a key role by bringing the team together, setting up the project board in Trello, scheduling and chairing regular meetings, circulating minutes, and following up with the team members to track progress. My technical contributions include using Python libraries and performing data preparations and visualisation activities. The project helped bring the best out of each team member and helped me learn from others. Despite challenges such as coordination across different time zones and reaching an agreement on business analysis questions, the visualisations overcame them by setting up WhatsApp groups for communication and having evidence-based discussions to reach an agreement. The project also helped me contribute to and improve my presentation and reporting skill set, which are important for my professional growth.

6. Personal and Professional Development

The Machine learning model has contributed to my career goals by providing me with practical skills and theoretical knowledge essential for a career in Artificial intelligence. My hands-on experience with Python, Scikit-learn, etc., has built my confidence in applying machine learning techniques to real-world problems. Knowing ethical and legal considerations has also helped me approach AI projects responsibly and professionally. Throughout the module, I developed skills in identifying biases, data preprocessing, model optimisation, and programming, which are also crucial for my professional development. The data preparation techniques, including decision-making, that I learned could help in the retail and healthcare projects that I work on.

This module has changed my approach to professional networking and how I can better engage in AI forums and discussions to stay updated on industry trends.

7. Team Learning

In this module, we had an excellent opportunity to collaborate with fellow students from different backgrounds, which enhanced my teamwork, programming and communication skills. We had two collaborative discussions in this module: The Fourth Industrial Revolution and Legal and Ethical Views on ANN Applications. My post on the 4th industrial revolution was on f was the 2017 ransomware attack (known as wannacry) on the National Health Service (NHS).

These discussions allowed each student to share their understanding, and the peer evaluation responses further improved our understanding of the topics. Additionally, we contributed and shared information on a few topics through the Module Wiki. These exercises reinforced my understanding and provided insights I would not have gained through self-study.

8. Future Learning goals

I plan to continuously learn by staying connected with advancements in Machine Learning, such as deep learning and reinforcement learning. I am more interested in specialising in reinforcement learning to explore its applications in automation and robotics. AI optimisation in the professional environment. I have already started discussing AI optimisation with colleagues, and I plan to conduct a workshop to raise awareness of ethical AI concerns by the end of February 2025. Within six months, I plan to pursue AI ethics certifications (e.g., Coursera, edx) to deepen my understanding of responsible AI practices in future projects.

My detailed personal / professional development plan can be found in the appendix below.

9. References

- Géron, A., 2019. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. 2nd ed. Sebastopol: O'Reilly Media.
- Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep Learning. Cambridge: MIT Press.
- Chollet, F., 2018. Deep Learning with Python. Shelter Island: Manning Publications.
- National Audit Office (2018) 'Investigation: WannaCry cyber-attack and the NHS', National Audit Office Report, 27 March. Available at: <https://www.nao.org.uk> [Accessed 15th January 2024]
- Schwab, K. (2016) The Fourth Industrial Revolution: what it means and how to respond World Economic Forum. <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/> [Accessed 15th January 2024]
- Scikit-learn developers, 2024. *Scikit-learn: Machine Learning in Python*. [online] Available at: <https://scikit-learn.org> [Accessed 15 January 2025].
- European Commission, 2021. *Ethics Guidelines for Trustworthy AI*. [online] Available at: <https://digital-strategy.ec.europa.eu> [Accessed 15 January 2025].
- The Royal Society, 2019. *AI Ethics and Governance*. [online] Available at: <https://royalsociety.org> [Accessed 15 January 2025].
- Airbnb Engineering, 2021. *Using Data Science to Understand Airbnb's Business*. [online] Available at: <https://medium.com/airbnb-engineering> [Accessed 15 January 2025].
- Kaggle (2021) *New York City Airbnb Open Data*, Kaggle. Available from: <https://www.kaggle.com/code/whyalwaysme/ab-nyc> 2019 [Accessed 15 January 2025].
- Mukhiya, S.K. & Ahmed, U. (2020) Hands-On Exploratory Data Analysis with Python. Packt Publishing.

10. Appendix A: e-Portfolio Links

e-portfolio main page	https://m-kanuri.github.io/
Machine Learning Module	https://m-kanuri.github.io/Module3.html
Unit 1	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2024/10/27/Col-dis-one.html • Initial Post: https://m-kanuri.github.io/artefacts/ML-Initial_Post.pdf • Peer Response: https://m-kanuri.github.io/artefacts/ML-Peer_Response.pdf • Summary Post: https://m-kanuri.github.io/artefacts/ML-Summary_Post.pdf
Unit 2	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2024/11/05/ML-Unit2.html • https://m-kanuri.github.io/MachineLearning/Unit02/
Unit 3	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2024/11/12/ML-Unit3.html • https://m-kanuri.github.io/MachineLearning/Unit03/
Unit 4	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2024/11/19/ML-Unit4.html • https://m-kanuri.github.io/MachineLearning/Unit04/
Unit 5	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2024/11/30/ML-Unit5.html • Clustering: https://m-kanuri.github.io/artefacts/ML-Unit5-Wiki%20Activity-%20Clustering.pdf • Jaccard Coefficient Calculations: https://m-kanuri.github.io/artefacts/ML-Unit5-Jaccard%20Coefficient%20Calculations.pdf
Unit 6	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2024/12/02/ML-Unit6.html • Final Report: Click Here • GitHub Code: https://github.com/mariaingold/AirbnbNYC • Team Contract: https://m-kanuri.github.io/artefacts/Team%20Contract%20Template.pdf • Meeting Notes: https://m-kanuri.github.io/artefacts/Team%20Meeting%20Notes.pdf
Unit 7	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2024/12/08/ML-Unit7.html • GitHub: https://m-kanuri.github.io/MachineLearning/
Unit 8	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2024/12/15/ML-Unit8.html • Gradient Cost Function: https://m-kanuri.github.io/MachineLearning/Unit08/

	<ul style="list-style-type: none"> • Initial Post: https://m-kanuri.github.io/artefacts/ML-Unit08-Initial Post.pdf • Peer Response: https://m-kanuri.github.io/artefacts/ML-Unit08-Peer Response.pdf • Summary Post: https://m-kanuri.github.io/artefacts/ML-Unit08-Summary Post.pdf
Unit 9	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2025/01/13/ML-Unit9.html • GitHub: https://m-kanuri.github.io/MachineLearning/Unit09/
Unit 10	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2025/01/16/ML-Unit10.html • GitHub: https://m-kanuri.github.io/MachineLearning/Unit10/
Unit 11	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2025/01/20/ML-Unit11.html • Model Performance Measurement: https://m-kanuri.github.io/MachineLearning/Unit11/ • New Network Presentation : Click Here • Neural Network GitHub: https://github.com/m-kanuri/m-kanuri.github.io/blob/main/NeualNetworkDesign.ipynb • Neural Network Transcript : Click Here
Unit 12	<ul style="list-style-type: none"> • Main Page: https://m-kanuri.github.io/module3/2025/01/23/ML-Unit12.html

11. Appendix B: Development Plan

Skill Area	Resources & Tools	Comments	Due Date
Machine Learning	Introduction to ML	Lecture cast (Unit1)	October 2024
	Exploratory Data Analysis	Practical experience through sample Jupyter Notebook file (Unit 2)	November 2024
	Correlation and Regression	e-portfolio activity (Unit 3)	November 2024
	Linear regression with Scikit - Learn	Practical experience sample Jupyter Notebook file (Unit 4)	November 2024
	Clustering, Jaccard Coefficient Calculations	e-portfolio activity (Unit 5)	December 2024
	K-Means Clustering Algorithm, Python	Development team project (Unit 6)	December 2024
	ANN Architecture, Perceptron, Training an ANN, Gradient Cost Function	e-portfolio activity (Unit 7 -8)	December 2024 / January 2025
	CNN Architecture	e-portfolio activity (Unit 9 & 10)	January 2025
	Model Selection and Evaluation	e-portfolio activity (Unit 11)	January 2025

	Future of ML	Lecture cast (Unit 12)	January 2025
Numerical Analysis	R and R Studio activities	Data Activity 1.2 & 1.3	August 2024
	R Commands	Data Activity (Unit 2)	August 2024
	R Data Management	Data Activity (Unit 3)	August 2024
	Probabilities	Probability Activities (Unit 4)	August 2024
	Producing Plots using R	Data activity (Unit 5)	September 2024
	Statistical Inference, Confidence Intervals and Hypothesis Testing	Statistics Quiz (Unit 6)	September 2024
	Parametric Tests	Data activity & Mathematics Test (Unit 7)	September 2024
	Non-Parametric Tests	Data Activity (Unit 8)	September 2024
	Chi-Square Analysis	Data Activity (Unit 9)	October 2024
	Correlation	Data Activity (Unit 10)	October 2024
	Regression Analysis	Data activity (Unit 11)	October 2024
	Bayesian Data Analysis	Activities (Unit 12)	October 2024
Understanding Artificial Intelligence	Introduction of AI, Benefits and Threats to Business, Intelligent Agents, Supervised Learning	Unit 1- 4 Watching videos	July 2024
	Learning System, Machine Learning and Deep Learning, WEKA	Unit 5-6 Watching tutorial, activities and videos	August 2024
	CRISP DM Methodology, WEKA for supervised & unsupervised learning, Data Mining, Future of AI	Unit 7-12 Watching tutorial, activities and videos	September 2024
Sharing / Improving Knowledge	Machine Learning - Collaborative Discussion	4 th Industrial revolution	November 2024
	Machine Learning - Collaborative Discussion	Legal Ethical views on ANN Applications	January 2025
	Machine Learning -WIKI activity	Clustering	November 2024
	Numerical Analysis – Collaborative Discussion	Application of Descriptive Statistics and Visualisation	September 2024
	Understanding Artificial Intelligence – Collaborative Discussion	Supervised & Unsupervised algorithms	June 2024

	Understanding Artificial Intelligence – Collaborative Discussion	AI in Financial Services	May 2024
Analytical thinking, Presentations Skills	Understanding Artificial Intelligence	Solution Implementation using WEKA software tool	July 2024
	Numerical Analysis - Statistical Analysis Presentation	Descriptive Statistics, Inferential Statistics, R Commands	October 2024
	Machine Learning – Neural Network Models for Object Recognition	Neural Network Models, activation function, Loss function.	January 2025
R Studio	Importing the JIRA data in CSV or Excel format in R for generating graphs and charts at workplace		Ongoing since August 2024
GitHub	Understanding artificial Intelligence Numerical Analysis Machine Learning	Used in all the modules for doing all the data activities	Ongoing
Presentation on AI Ethics	Presentation/ workshop on AI Ethics at workplace		February 2025
Data Preparation Techniques	Practice the techniques in the at workplace		Ongoing
Building Professional network	Meetups, Conferences, LinkedIn		Ongoing
Reinforcement Learning	Q-learning, Deep Q Networks, OpenAI Gym	Hands-on projects and simulations	April 2025
Transfer Learning	Pretrained models (e.g., ResNet, BERT)	Implement transfer learning on real-world datasets	June 2025
Bias and Fairness in AI	AI Fairness 360, Ethical AI frameworks	Techniques to identify and mitigate biases in AI models	July 2025
Certifications	AI Certification courses on AI Ethics	Coursera / Edx	August 2025