Ryan MacKellar

HNR 499: Final Project

31 March 2025

Senior Project Proposal: A Beginner's Guide to AI Development

Central Focus:

The goal of my senior project is to comprehensively explore the mathematics underlying artificial intelligence (AI) and machine learning (ML) algorithms, specifically neural networks and linear algebra, and to demonstrate how these mathematical principles translate directly into practical coding and real-world deployment. This project will involve building machine learning algorithms from scratch using Python and subsequently applying standard Python libraries for advanced implementations. Ultimately, my work will culminate in an informational manual designed to clearly connect mathematical theories, coding practices, and future-oriented AI strategies.

Rationale:

Artificial intelligence continues to rapidly transform various sectors, demanding skilled individuals who deeply understand both the theoretical mathematics and practical coding foundations. My project addresses this need by offering detailed insights into the process of developing AI algorithms from fundamental principles. Unlike many existing resources that either focus solely on theory or only on high-level coding, my manual uniquely bridges these domains, providing thorough mathematical explanations alongside practical coding examples and real-world applications. This knowledge is invaluable, not only for my personal interest but also to substantially enhance my employability and entrepreneurial potential within the AI technology sector.

Ryan MacKellar

HNR 499: Final Project

31 March 2025

Methodology:

My approach is structured into a few primary phases:

1. **Weeks 1-3:** Conduct a detailed literature review and summarize foundational mathematics relevant to machine learning, emphasizing neural networks and linear algebra, and explore how these translate into coding structures. Select datasets (either one large or a few varying in complexity and size) to work with later on to distinguish the project from other resources online.
2. **Weeks 4-7:** Utilize PyCharm on my MacBook to program neural network algorithms from scratch, connecting mathematical concepts explicitly with the code.
3. **Weeks 8-11:** Implement advanced ML algorithms utilizing Python libraries (such as TensorFlow, Keras, PyTorch, NumPy, and Pandas) to engage with more complex, real-world AI challenges and applications.
4. **Weeks 12-13:** Research the future trajectory of AI, including potential societal impacts and strategic skill sets required to maintain relevance in a rapidly evolving technological landscape.

Weeks 14-15: Synthesize all research, analyses, coding examples, and insights into a cohesive, educational AI manual. This manual will include annotated Python code segments, screenshots demonstrating coding processes, and explicit explanations connecting mathematical theory and coding practice to real-world contexts.

Ethical Considerations: No human or animal subjects will be involved, thus no additional university approval is necessary.

Ryan MacKellar

HNR 499: Final Project

31 March 2025

Final Deliverable:

The final project submission will be an informational manual of approximately 25 pages formatted in MLA style. It will detail mathematical foundations, coding processes, advanced applications, and forward-looking insights into artificial intelligence, supported by comprehensive visual aids such as screenshots, code annotations, diagrams, and explanatory narratives.

Timeline and Advisor Meetings:

- Weeks 1-3: Mathematical and coding foundation research.
- Weeks 4-7: Development of neural networks from scratch.
- Weeks 8-11: Algorithm implementation using Python libraries.
- Weeks 12-13: Analysis of AI's future and strategic skill enhancement.
- Weeks 14-15: Compilation, editing, and final writing of the manual.

Advisor checkpoints:

- Week 3: Initial literature and research review.
- Week 6: Neural network coding progress review.
- Week 9: Intermediate Python library coding review.
- Week 12: Review of findings related to future AI trends and outline of the final manual.
- Week 14: Format information, compile, and fine tune the final deliverable.

Distinctiveness

This project does not duplicate any existing coursework or academic obligations but significantly

contributes to my professional goals by enhancing my qualifications for careers or entrepreneurial opportunities within the AI technology industry. As mentioned, specific data sets will be selected to demonstrate how models can be tuned to work in specific situations.

Ryan MacKellar

HNR 499: Final Project

31 March 2025

"Build a Machine Learning Model." *Codecademy*,

www.codecademy.com/learn/paths/machine-learning. Accessed 19 Mar. 2025.

Codecademy's machine learning path walks users through the process of building models using Python, starting with an introduction to key concepts and the role of data in model training. The course dives into supervised learning, covering regression and classification with hands-on exercises using algorithms like linear regression and decision trees. It also introduces unsupervised learning techniques, such as clustering, to uncover patterns in unlabeled data. Throughout the program, learners work with popular libraries like scikit-learn and TensorFlow to apply what they learn in real-world-inspired projects. By the end, participants gain practical experience in developing, testing, and deploying machine learning models.

"Build Your First Machine Learning Model in Python." *YouTube*, uploaded by

freeCodeCamp.org, 2 Sep. 2023, www.youtube.com/watch?v=29ZQ3TDGgRQ.

Accessed 19 Mar. 2025.

This tutorial walks viewers through building their first machine learning model using Python's scikit-learn library. It starts by introducing key concepts like datasets, features, and labels before moving on to data loading and preprocessing to prepare it for training. The video then demonstrates how to implement a simple linear regression model, breaking down each step in a clear and accessible way. Along the way, viewers learn how to assess the model's performance using mean squared error and visualize the results with matplotlib. The tutorial also highlights best practices for model development, such as data splitting and validation. By the end, viewers will have hands-on experience constructing and evaluating a basic machine learning model.

Kumar, Ravish. "Detailed Maths Topics and Their Direct Use in Machine Learning." *Medium*, 9

Aug. 2022,

medium.com/enjoy-algorithm/detailed-maths-topics-in-machine-learning-ca55cd537709.

Accessed 19 Mar. 2025.

Ryan MacKellar

HNR 499: Final Project

31 March 2025

This article explores the key mathematical principles that form the backbone of machine learning. It highlights the role of linear algebra, explaining how vectors and matrices are essential for structuring data and performing computations within algorithms. Probability concepts, including probability distributions, are covered to illustrate their importance in handling uncertainty and making predictions. The discussion also extends to statistics, emphasizing its role in analyzing and interpreting data for training and validating models. By linking these mathematical foundations directly to real-world machine learning applications, the article helps readers grasp the theoretical concepts that shape algorithm development.

"Machine Learning with Python Tutorial." *GeeksforGeeks*, 12 Oct. 2023,

www.geeksforgeeks.org/machine-learning-with-python/. Accessed 19 Mar. 2025.

This tutorial introduces the fundamentals of implementing machine learning models in Python. It starts by covering the essential prerequisites, including Python programming skills and key mathematical concepts like linear algebra and statistics. Readers are guided through the seven core steps of starting an ML project: collecting and preparing data, selecting and training a model, evaluating its performance, fine-tuning parameters, and making predictions. Each phase is explained in detail to help beginners navigate the process effectively. The tutorial also emphasizes the role of widely used Python libraries, such as NumPy, Pandas, scikit-learn, TensorFlow, and Keras, in streamlining machine learning tasks. Additionally, it answers common questions and offers tips on staying updated with industry advancements. By the end, readers will have a solid foundation for building their own machine learning projects.

Saxena, Sharoon. "Mathematics Behind Machine Learning." *Analytics Vidhya*, 10 Oct. 2019,

www.analyticsvidhya.com/blog/2019/10/mathematics-behind-machine-learning/.

Accessed 19 Mar. 2025.

This article delves into the fundamental mathematical principles that drive machine learning algorithms. It explains how linear algebra plays a key role in data representation and transformation, with vectors and matrices enabling complex computations within models. Calculus is explored in the context of optimization, particularly its use in training neural networks through gradient descent. The discussion also covers probability theory, emphasizing its importance in handling uncertainty and making predictions. Additionally,

the article highlights the role of statistics in analyzing data and conducting hypothesis testing to assess model performance. By breaking down these core mathematical concepts, the piece helps readers gain a clearer understanding of how machine learning algorithms are built and refined.

"Mathematics for Machine Learning." *GeeksforGeeks*, 5 Jan. 2024,

www.geeksforgeeks.org/machine-learning-mathematics/. Accessed 19 Mar. 2025.

This article explores the essential mathematical foundations of machine learning, highlighting key concepts necessary for understanding and building algorithms. It explains how linear algebra—specifically vectors, matrices, and matrix operations—plays a crucial role in data representation and transformations. The discussion also covers dimensionality reduction techniques like Principal Component Analysis (PCA) and Singular Value Decomposition (SVD), which leverage linear algebra to simplify complex datasets. Additionally, the article examines the role of probability and statistics in modeling data distributions and improving predictive accuracy. By breaking down these fundamental topics, the piece emphasizes how mathematics enhances machine learning models and optimizes their performance.

"Python Machine Learning." *W3Schools*,

www.w3schools.com/python/python_ml_getting_started.asp. Accessed 19 Mar. 2025.

W3Schools provides an easy-to-follow introduction to machine learning using Python. The tutorial explains machine learning as a way for computers to learn from data and statistics, positioning it as a step toward artificial intelligence. It highlights Python's advantages, such as its simplicity and the wide range of available libraries like scikit-learn. Readers are guided through key steps, from preparing data to training simple models, with practical examples that demonstrate real-world applications. The tutorial also emphasizes the role of data visualization in interpreting and presenting results. By working through this guide, beginners can gain hands-on experience with basic machine learning models, laying the groundwork for more advanced studies in the field.

"Scikit-learn." *Wikipedia: The Free Encyclopedia*, Wikimedia Foundation, 18 Feb. 2025,

en.wikipedia.org/wiki/Scikit-learn. Accessed 19 Mar. 2025.

Ryan MacKellar

HNR 499: Final Project

31 March 2025

This Wikipedia article provides an overview of scikit-learn, a free and open-source machine learning library for Python. It outlines the library's key features, including a variety of classification, regression, and clustering algorithms such as support-vector machines, random forests, gradient boosting, k-means, and DBSCAN. The article also highlights scikit-learn's compatibility with Python's scientific computing stack, noting its integration with libraries like NumPy and SciPy. A section on its history traces the library's origins as a Google Summer of Code project and its growth into a widely used tool in both academia and industry. Additionally, the implementation details explain that while scikit-learn is primarily written in Python, its core algorithms leverage Cython for optimized performance.