Reflective Journal

Name: Ethan Phan

Course: ITAI 2376

Date: 02.25.2025

Lab Module: AWS Machine Learning University - Module 2

Learning Insights

Throughout the AWS Machine Learning University Module 2 labs, I explored fundamental machine learning concepts, including supervised and unsupervised learning, data preprocessing, feature engineering, and model evaluation. One key takeaway was the significance of data quality and preprocessing, which directly impacts model performance.

These labs connected to broader machine learning principles by emphasizing iterative model improvement, bias-variance tradeoff, and the importance of choosing the right algorithm for different tasks. The labs provided practical exposure to common workflows in ML, reinforcing theoretical concepts learned in class.

The most impactful learning moment was implementing feature selection and observing how small adjustments in input features dramatically affected model accuracy. Additionally, understanding hyperparameter tuning and its effect on model generalization was a crucial insight.

Challenges and Struggles

One of the most significant technical challenges I encountered was handling missing data and outliers. Initially, I struggled with choosing the right imputation strategies without introducing

bias. I overcame this by experimenting with different techniques, such as mean imputation and KNN imputation, and evaluating their effects on model performance.

Conceptually, understanding the underlying mechanics of gradient descent was challenging, especially when interpreting the impact of learning rates on convergence. To tackle this, I referenced external resources and visualizations, which helped clarify the concept.

For problem-solving, I developed a strategy of systematic debugging—carefully analyzing error messages, using print statements, and leveraging documentation. I also adopted a more structured approach to experimenting with hyperparameter tuning, which made the process more efficient.

Personal Growth

My understanding of machine learning has evolved from seeing it as an abstract concept to appreciating its structured, methodical nature. Initially, I assumed ML was largely about selecting an algorithm, but I now understand that data preparation and feature engineering are just as critical.

One surprising discovery was how minor adjustments in feature selection and data preprocessing can significantly impact results. This reinforced the idea that a model is only as good as the data it learns from.

These skills will be valuable for both academic and professional goals. Understanding the machine learning pipeline prepares me for more complex projects in AI applications and data science. Professionally, this knowledge strengthens my ability to analyze and improve predictive models, an essential skill in many industries.

Critical Reflection

If I were to repeat these labs, I would dedicate more time to experimenting with different algorithms and feature engineering techniques rather than strictly following lab instructions. This would provide deeper insights into model behavior and optimization strategies.

These labs sparked additional questions, particularly around model interpretability and ethical AI. Understanding how to make models explainable and mitigate bias remains a crucial area of interest.

Overall, these labs provided a strong foundation in practical machine learning. They fit into the broader landscape by demonstrating how real-world ML applications require continuous refinement, validation, and understanding of both data and models. Moving forward, I plan to apply these concepts to personal projects, reinforcing and expanding my knowledge in the field.