**CCS 4403: Machine Learning**

Contact Hours: 45 hours

Prerequisites: AMA 4107 Probability and Statistics I, CCS 4302 Principles of Artificial Intelligence.

Purpose of the Course: This course is designed to give the intuition behind modern machine learning methods as well as a bit more formal understanding of how, why, and when they work.

Expected Learning Outcomes of the Course:

Students who successfully complete this unit of study will be able to:

i. Develop Proficiency in Machine Learning Fundamentals:

By the end of the course, students should be able to demonstrate a solid understanding of the fundamental concepts of machine learning, including key terminology, types of machine learning, and the underlying mathematical principles.

ii. Apply Supervised Learning Techniques:

Students should be able to implement and deploy supervised learning algorithms, such as linear regression and logistic regression, for both regression and classification tasks. This objective includes the ability to preprocess and analyze data, train models, and assess their performance.

iii. Utilize Unsupervised Learning Approaches:

Students should gain proficiency in unsupervised learning techniques, specifically in clustering (e.g., K-means) and dimensionality reduction (e.g., PCA). They should be capable of applying these methods to explore patterns and relationships within datasets without labeled outputs.

iv. Evaluate and Interpret Model Performance:

Upon completion of the course, students should be skilled in evaluating the performance of machine learning models using appropriate metrics. They should also be able to interpret the results, identify areas for improvement, and make informed decisions about model selection.

v. Apply Machine Learning to Real-World Problems:

The course aims to equip students with practical skills by engaging them in a hands-on capstone project. By the end of the semester, students should be able to formulate a real-world problem, apply appropriate machine learning techniques, and present a solution, showcasing their ability to apply theoretical knowledge to practical scenarios.

**Course Content**

Fundamentals of machine learning, covering key mathematical concepts, programming skills using Python and relevant libraries (NumPy, pandas), and the core principles of supervised and unsupervised learning. The curriculum includes an overview of regression, classification, clustering techniques (such as K-means and hierarchical clustering), and dimensionality reduction methods (e.g., PCA, t-SNE). Emphasis is placed on model evaluation using cross-validation and performance metrics. The course also incorporates a special topics segment, allowing students to explore advanced areas like neural networks or natural language processing. The semester culminates in a hands-on capstone project, providing students with the opportunity to apply their acquired skills to a real-world problem, reinforcing theoretical concepts through practical application.

**Course Outline**

**Weeks 1-2**: Introduction to Machine Learning

- Definition and types of machine learning

- Applications of machine learning in computer science

- Overview of supervised, unsupervised, and reinforcement learning

**Weeks 3-4**: Mathematical Foundations for Machine Learning

- Linear algebra essentials

- Basics of probability and statistics for machine learning

**Weeks 5-6:** Programming Foundations

- Introduction to Python for machine learning

- Basic libraries: NumPy, pandas, matplotlib

**Weeks 7-8:** Supervised Learning

- Introduction to regression and classification

- Linear regression and logistic regression

**Weeks 9-10:** Unsupervised Learning

- Clustering algorithms: K-means, hierarchical clustering

- Dimensionality reduction: PCA, t-SNE

**Weeks 11-12:** Model Evaluation and Selection

- Cross-validation

- Performance metrics: accuracy, precision, recall, F1 score

**Weeks 13-14**: Special Topics and Project

- Choose a special topic based on student interest (e.g., neural networks, natural language processing)

- Capstone project: Apply machine learning concepts to a real-world problem

**Week 1-2 Notes**

https://gamma.app/docs/Introduction-to-Machine-Learning-l27ibp2hw1tp1rx?mode=doc

**Week 1-2 Assessment**

Section 1: Multiple Choice Questions (10 points)

1. What is the primary goal of machine learning?

a. Automating programming

b. Analyzing historical data

c. Improving machine performance

d. Enhancing human intelligence

2. Which type of machine learning involves learning from labeled data?

a. Unsupervised learning

b. Reinforcement learning

c. Semi-supervised learning

d. Supervised learning

3. What is the main challenge in machine learning associated with an algorithm learning the training data too well but failing on new data?

a. Underfitting

b. Overfitting

c. Bias

d. Variance

4. In unsupervised learning, what is the primary task?

a. Regression

b. Classification

c. Clustering

d. Dimensionality reduction

5. Which library is commonly used for data manipulation and analysis in Python?

a. TensorFlow

b. PyTorch

c. NumPy

d. Scikit-learn

Section 2: Short Answer Questions (20 points)

6. Explain the difference between supervised and unsupervised learning, providing an example for each.

7. Describe the role of training data in machine learning. Why is it essential?

8. Provide a brief overview of the machine learning workflow, including the main stages.

9. Briefly discuss one challenge associated with bias in machine learning models and propose a strategy to address it.

10. Name and briefly explain two popular machine learning algorithms and their typical applications.