(** Tentative Curriculum - Can be subject to minor changes)

Introduction to Machine Learning

Total Course Hours: 10 hours
[Tuition assistance available based on eligibility] -
Number of Sessions: 10
Session Length: 55 minutes
(35 min Lecture, 15 min Activity, 5 min Take-Home Lab Assignment Discussion)
Mode: In-person OR Synchronous Online Session options available.
One session Every week starting Sept 9th (Day/Time to be determined)

Session 1: What is Machine Learning?

- **Lecture (35 min):**
  - Motivation & Purpose: Why should I learn Machine Learning?
  - Defining Machine Learning: How computers learn from data without explicit programming.
  - ML vs. AI: Understanding the relationship and differences.
  - Brief historical context of ML.
  - Types of Machine Learning: Supervised, Unsupervised, and Reinforcement Learning (conceptual overview).
- **Activity (15 min):**
  - Brainstorming: Students list 3-5 everyday applications where they suspect Machine Learning is used.
  - Group Share: Discuss examples and identify common characteristics. What would you do differently?
- **Take-Home Lab Assignment (5 min):**
  - Find an example of Machine Learning in a product or service you use daily. Briefly describe what it does and how you think it might "learn."

Session 2: Data - The Fuel for Machine Learning

- **Lecture (35 min):**
  - The critical role of “data” in Machine Learning & Predictions.
  - Real-world applications & Data sets of ML in action (e.g., recommendation systems, image recognition, spam filters).
  - Types of Data: Numerical (continuous, discrete) and Categorical (nominal, ordinal).

    - Understanding Features and Labels (Target Variables).
    - Basic concepts of data collection and preparation (e.g., why clean data is important).
    - Simple examples of messy data (missing values, outliers) and the need for cleaning.
- **Python Coding Activity (15 min):**
    - Data Identification: Given a few small, simple datasets (e.g., student grades, animal characteristics), students identify features and potential labels.
    - Discussion: How might missing data impact a model?
- **Take-Home Lab Assignment (5 min):**
    - Find a small public dataset online (e.g., from Kaggle or a government data portal). Identify at least 3 features and one potential label you could try to predict.

Session 3: Supervised Learning: Regression Models

- **Lecture (35 min):**
    - Introduction to Supervised Learning: Learning from labeled data.
    - What is Regression? Predicting a continuous numerical value.
    - Linear Regression (conceptual, no complex math): The idea of finding a "best fit line."
    - Common use cases for Regression (e.g., predicting house prices based on size, predicting temperature, stock prices).
- **Activity (15 min):**
    - Graphing Prediction: Students are given a sample data set and asked to use python ML library to implement a regression model that best predicts future scores.
    - Discussion: What makes a "good" prediction line?
- **Take-Home Lab Assignment (5 min):**
    - Collect 5 data sources for two related numerical variables (e.g., your height and your arm span, or the number of push-ups you can do and your current age). Try to predict a new value using your "line of best fit" from the activity.

Session 4: Supervised Learning: Classification Models

- **Lecture (35 min):**
  - What is Classification?
  - Binary Classification (two categories) vs. Multi-class Classification (more than two categories).
  - Applications & How Classification Models work: Spam detection (spam/not spam), image classification (cat/dog), disease diagnosis (sick/healthy).
  - The concept of decision boundaries
- **Activity (15 min):**
  - Categorization Game: Students are given a set of items with various characteristics and asked to sort them into predefined categories based on rules they infer.
  - Discussion: How did you decide which category each item belonged to?
- **Take-Home Lab Assignment (5 min):**
  - Imagine you have data on fruits (e.g., color, size, shape). Given a new fruit with specific characteristics, describe how a classification model might determine if it's an apple or an orange.
  - Python Coding Assignment

Session 5: Model Training and Evaluation

- **Lecture (35 min):**
  - The importance of splitting data: Training Set vs. Testing Set.
  - Why we need a testing set: To evaluate how well the model generalizes to new, unseen data.
  - Concept of Overfitting (model too complex, memorizes training data) and Underfitting (model too simple, doesn't capture patterns).
  - Basic evaluation metric: Accuracy (simple definition: percentage of correct predictions).
- **Activity (15 min):**
  - Data Splitting Exercise: Given a small dataset, students divide them into a training set and a testing set with different possibilities (50:50, 60:40, 70:30,etc) and observe the prediction change.
  - Discussion: Why is it bad if a model only performs well on the training data?
- **Take-Home Lab Assignment (5 min):**
  - Think about a simple prediction you make regularly (e.g., predicting if it will rain based on clouds). How would you "test" your prediction method to see if it's accurate?

- Assignment Work Sheet to be completed

Session 6: Unsupervised Learning: Clustering

- **Lecture (35 min):**
  - Introduction to Unsupervised Learning: Finding patterns in unlabeled data.
  - What is Clustering? Grouping similar data points together.
  - Conceptual understanding of K-Means Clustering (finding natural groupings).
  - Common use cases for Clustering (e.g., customer segmentation, document organization, anomaly detection).
- **Activity (15 min):**
  - Object Grouping: Students are given a collection of diverse objects (or images of objects) and asked to group them into categories based on their own criteria.
  - Discussion: What criteria did you use to form your groups? Were there different ways to group them?
- **Take-Home Lab Assignment (5 min):**
  - Imagine you have a list of movies. How could you use clustering to group them into genres without knowing the genres beforehand? List at least three features you would use.
  - Coding: Use your ideas above and build a Movie clustering Model

Session 7: Introduction to Neural Networks

- **Lecture (35 min):**
  - Inspired by the Brain: The concept of an artificial neuron.
  - Building Blocks: Input Layer, Hidden Layers, Output Layer.
  - How Neural Networks "learn" (very simplified idea of adjusting connections/weights based on errors).
  - Brief mention of "Deep Learning" as neural networks with many hidden layers.
  - Simple applications (e.g., basic image recognition, pattern detection).
- **Activity (15 min):**
  - Draw a Simple Neural Network: Students draw a diagram of a very basic neural network with an input layer, one hidden layer, and an output layer, labeling the components.
  - Discussion: How does adding more "neurons" or "layers" change what the

network can do?

- **Take-Home Lab Assignment (5 min):**
  - Research one real-world application where neural networks are commonly used (e.g., self-driving cars, medical image analysis). Briefly describe how they are applied.

Session 8: Introduction to Reinforcement Learning

- **Lecture (35 min):**
  - What is Reinforcement Learning (RL)? Learning through trial and error, by interacting with an environment.
  - Key components: Agent, Environment, State, Action, Reward.
  - The concept of "rewards" and "penalties" guiding learning.
  - Simple examples: Training a robot to walk, playing games (e.g., AlphaGo).
  - Difference from Supervised and Unsupervised Learning.
- **Activity (15 min):**
  - Simple Game Design: Students brainstorm a very simple game (e.g., navigating a maze) and define the agent, environment, possible actions, and rewards/penalties for that game.
  - Discussion: How would an RL agent "learn" to play this game effectively?
- **Take-Home Lab Assignment (5 min):**
  - Research one real-world application of Reinforcement Learning (e.g., self-driving cars, robotics, optimizing energy consumption). Briefly describe how RL is used in that application.

Session 9: Real-World ML Applications and Limitations

- **Lecture (35 min):**
  - Diving deeper into diverse ML applications: Medical imaging analysis, fraud detection, personalized medicine, content recommendation.
  - Limitations of Machine Learning:
    - Data Dependency: ML models are only as good as their data.
    - Interpretability: Difficulty in understanding complex model decisions.
    - Robustness: Sensitivity to small changes in input data.
    - The need for human oversight.
- **Activity (15 min):**

    - Pros and Cons Debate: Students discuss the advantages and disadvantages of using ML in a specific domain (e.g., using ML for diagnosing diseases).
- **Take-Home Lab Assignment (5 min):**
    - Identify a real-world ML application you find interesting. Describe one potential limitation or challenge that application might face.

Session 10: The Future of ML and Next Steps

- **Lecture (35 min):**
    - Recap of core Machine Learning concepts: What it is, how it works, its potential, and its challenges.
    - Emerging trends in ML (e.g., Explainable AI, Responsible AI development).
    - Career paths in Machine Learning and related fields.
    - Resources for continued learning (online courses, books, communities).
    - The role of individuals in shaping the ethical development and use of ML.
- **Activity (15 min):**
    - "My ML Interest": Students share one aspect of Machine Learning that they found most interesting or one question they still have.
    - Group Discussion: What excites you most about the future of Machine Learning?
- **Take-Home Lab Assignment (5 min):**
    - Identify one specific area within Machine Learning (e.g., computer vision, recommendation systems, ethical AI) that you would like to explore further. Find one online resource (website, course, book) that could help you learn more about it.