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

Introduction to Machine Learning

What is Machine Learning?

- 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).

Lab Assignment:

- 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."

Data - The Fuel for Machine Learning

- 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:**
  - 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?

Supervised Learning: Regression Models

- 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:**
  - 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?

Supervised Learning: Classification Models

- 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

Lab Assignment :

- 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

Model Training and Evaluation

- 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:**
  - 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?

Unsupervised Learning: Clustering

- 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).

Lab Assignment:

- 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

- 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).

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.

Introduction to Reinforcement Learning

- 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:**
  - 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?

Real-World ML Applications and Limitations

- 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:

- Pros and Cons Debate: Students discuss the advantages and disadvantages of using ML in a specific domain (e.g., using ML for diagnosing diseases).