ML PREVIOUS PAPER

UNIT 1

1A) HOW DOES THE CENTRAL NERVOUS SYSTEM HELP IN CREATING INTELLIGENT MACHINES? (6M)

The central nervous system (CNS) provides a biological model for creating intelligent machines. By studying the CNS, researchers gain insights into how neurons, synapses, and neural networks process information, learn from experiences, and adapt over time.

Key aspects include:

- **Neural Networks:** Inspired by the brain's structure, artificial neural networks (ANNs) are designed to simulate the way the brain processes information, enabling machines to recognize patterns and make decisions.
- **Learning and Adaptation:** The CNS learns from experiences through synaptic plasticity. Similarly, machine learning algorithms adjust weights within artificial networks based on training data to improve performance.
- **Parallel Processing:** The brain processes multiple pieces of information simultaneously. ANNs leverage parallel processing capabilities to handle large datasets and complex computations efficiently.
- **Cognitive Functions:** Understanding functions like memory, perception, and problem-solving helps in developing AI systems capable of similar tasks.

1B) WHAT IS MACHINE LEARNING? EXPLAIN ANY THREE BUSINESS APPLICATIONS OF MACHINE LEARNING. (8M)

Machine learning (ML) is a subset of artificial intelligence (AI) that involves the development of algorithms that allow computers to learn from and make predictions or decisions based on data. It enables systems to improve their performance over time without being explicitly programmed.

Three business applications of machine learning:

- **Predictive Maintenance:** In manufacturing, ML algorithms analyze data from machinery to predict when maintenance is needed, reducing downtime and saving costs.
- **Customer Segmentation:** Retail and marketing sectors use ML to segment customers based on their behavior and preferences, allowing for targeted marketing and personalized recommendations.
- **Fraud Detection:** Financial institutions employ ML to detect fraudulent transactions by identifying unusual patterns and anomalies in transaction data, enhancing security and reducing financial losses.

2A) WHAT IS A WELL-POSED LEARNING PROBLEM? SPECIFY TASK, PERFORMANCE MEASURE, AND TRAINING EXPERIENCE FOR THE FOLLOWING LEARNING PROBLEMS. (7M)

A well-posed learning problem is one where the objective is clearly defined, and the performance can be measured. It includes:

Task (T): What the system needs to learn.

Performance measure (P): How the system's performance is evaluated.

Training experience (E): The data and experiences the system learns from.

(i) A checkers learning problem

- **Task (T):** Play checkers.
- **Performance measure (P):** Win/loss ratio or score.
- **Training experience (E):** Historical games of checkers, including moves and outcomes.

(ii) A handwritten recognition learning problem

- **Task (T):** Recognize handwritten characters.
- **Performance measure (P):** Accuracy in recognizing characters.
- **Training experience (E**): A dataset of handwritten characters with corresponding labels.

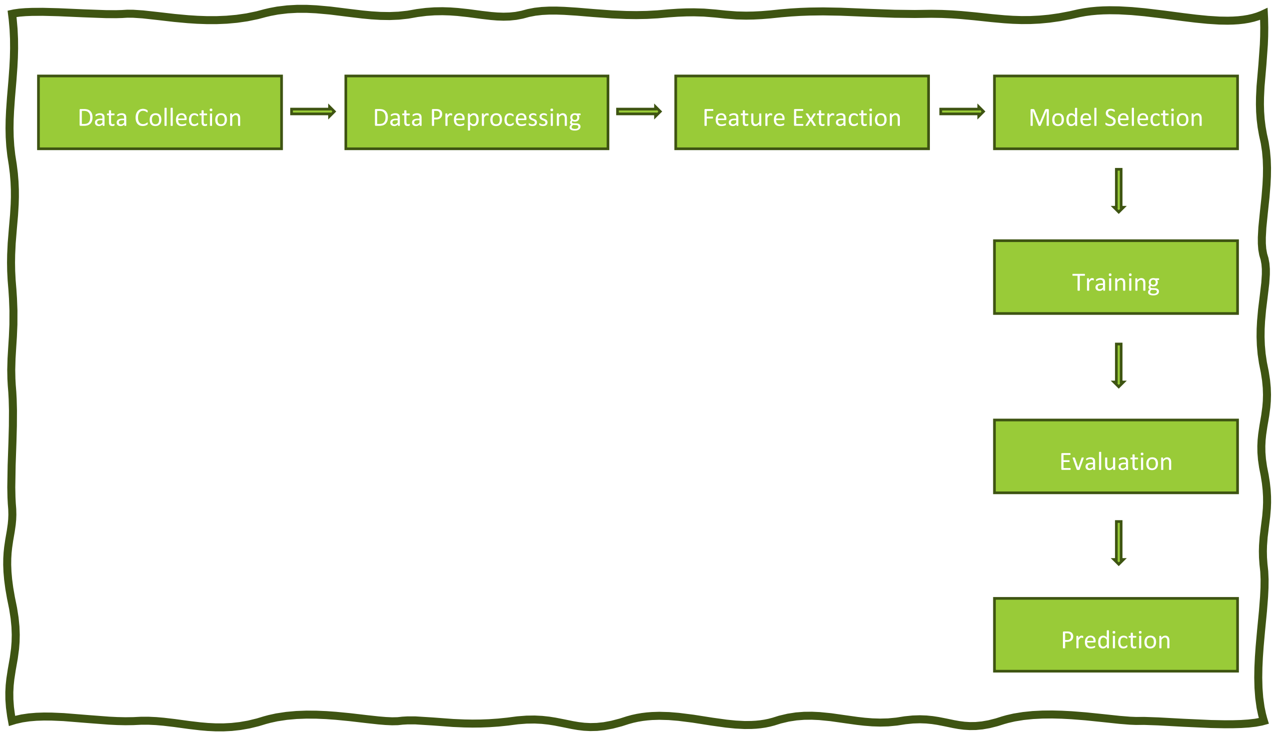
(iii) A robot driving learning problem

- **Task (T):** Drive a robot autonomously.
- **Performance measure (P):** Safety and efficiency of driving (e.g., collision rate, time to destination).
- **Training experience (E):** Recorded driving experiences, including sensor data and driving actions.

2B) WITH A NEAT DIAGRAM EXPLAIN THE BLOCK DIAGRAM OF MACHINE LEARNING. (7M)

A typical block diagram of a machine learning system includes the following components:

- **Data Collection:** Gathering relevant data for training.
- **Data Preprocessing:** Cleaning and preparing data for analysis.
- **Feature Extraction:** Identifying important features from the data.
- **Model Selection:** Choosing an appropriate machine learning algorithm.
- **Training:** Feeding the preprocessed data into the model and adjusting parameters.
- **Evaluation:** Measuring the model's performance using metrics.
- **Prediction:** Using the trained model to make predictions on new data.

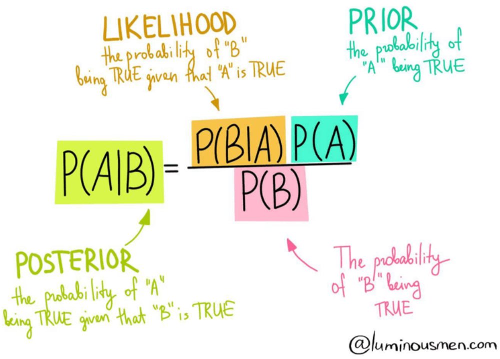


UNIT 2

3A) HOW BAYES' THEOREM SUPPORTS THE CLASSIFICATION TASK? EXPLAIN WITH AN EXAMPLE. (7M)

Bayes' theorem helps in updating the probability estimate for a hypothesis as additional evidence is acquired. It's foundational for the Naive Bayes classifier, which applies Bayes' theorem with strong independence assumptions.

Bayes' Theorem Formula:



YouTube Video Link: Bayes theorem and Example

3B) EXPLAIN THE CONCEPT OF LINEAR REGRESSION WITH LEAST SQUARE ERROR CRITERION. (7M)

Linear Regression: Linear regression is a data analysis technique that predicts the value of unknown data by using another related and known data value. It mathematically models the unknown or dependent variable and the known or independent variable as a linear equation.

Least Squares Error Criterion: The least squares criterion is a formula used to measure the accuracy of a straight line in depicting the data that was used to generate it. That is, the formula determines the line of best fit.

YouTube Video Link: Linear Regression with Least Square Error Criterion

4A) CONSIDER THE DATASET D, CALCULATE PROBABILITY FOR UNSEEN PATTERN (X: {M, 1.85}) USING NAÏVE BAYES CLASSIFIER. (7M)

Naïve Bayes Classifier: Naive Bayes algorithm (NB) is Bayesian graphical model that has nodes corresponding to each of the columns or features. It is called **naive** because, it ignores prior distribution of parameters and assume independence of all features and all rows.

4A-> SOLUTION

YouTube Video Link: Naïve Bayes Classifier

4B) IN A PARTICULAR PAIN CLINIC, 10% OF PATIENTS ARE PRESCRIBED NARCOTIC PAINKILLERS. OVERALL, 5% OF THE CLINIC'S PATIENTS ARE ADDICTED TO NARCOTICS. OUT OF ALL THE PEOPLE PRESCRIBED PAIN PILLS, 8% ARE ADDICTS. IF A PATIENT IS AN ADDICT, WHAT IS THE PROBABILITY THAT THEY WILL BE PRESCRIBED PAIN PILLS? (USE BAYES' THEOREM) (7M)

4B-> SOLUTION

UNIT 3

5) CONSIDER THE FOLLOWING POSITIVELY AND NEGATIVELY LABELED DATA POINTS.

POSITIVELY LABELED DATA POINTS: **(3, 1), (3, -1), (6, 1), (6, -1)**

NEGATIVELY LABELED DATA POINTS: **(1, 0), (0, 1), (0, -1), (-1, 0)**

1)DOES THESE POINTS LINEARLY SEPARABLE?

2)DETERMINE THE HYPER PLANE THAT CLASSIFIES THESE POINTS.

- **Are these points linearly separable?**
  Yes, these points are linearly separable. A linear classifier can separate the positive and negative points.

- **Determine the hyperplane that classifies these points.**
  A hyperplane in two dimensions is a line described by $ax + by + c = 0$
  We can use the points to determine a suitable hyperplane.

  Let's use the points (3,1) and (1,0) for the positive and negative classes, respectively. One possible separating hyperplane is: $3x + y - 10 = 0$

6A) SUPPOSE THAT WE HAVE TRAINED A CLASSIFIER TO CLASSIFY MEDICAL DATA TUPLES, WHERE THE CLASS LABEL ATTRIBUTE IS CANCER AND THE POSSIBLE CLASS VALUES ARE TRUE AND FALSE. THE CONFUSION MATRIX OF THE MODEL IS AS GIVEN BELOW. COMPUTE THE FOLLOWING: (8M)

	Predicted +ve	Predicted -ve
Actual +ve	4	7
Actual -ve	4	185

1. SUCCESS RATE

2. MISS CLASSIFICATION RATE

3. SENSITIVITY

4. SPECIFICITY

6A-> SOLUTION

6B) IN SUPPORT VECTOR MACHINES FIXING THE MAXIMUM MARGIN SOMETIMES LEAD TO MISCLASSIFICATION OF SOME OF THE SAMPLE. HOW TO OVERCOME THIS PROBLEM?

In support vector machines (SVMs), the goal is to find a decision boundary that maximizes the margin, which is the distance between the boundary and the nearest data points (support vectors). However, in some cases, this maximum margin boundary might not perfectly separate all the training data points correctly. Here’s how we can handle this issue:

1. **Soft Margin SVM:**
   - Instead of insisting on a perfect separation of classes, we allow some samples to be misclassified. This approach is called a soft margin SVM.

- The idea is to find a balance where the margin is still maximized but some data points are allowed to be on the wrong side of the decision boundary.
- This is particularly useful when the data is not perfectly separable or when there is noise in the data.

2. **Introduction of Slack Variables:**
   - To implement a soft margin, we introduce slack variables (ξi) for each data point.
   - Slack variables allow data points to be on the wrong side of the margin or even on the wrong side of the decision boundary, but penalize such occurrences to keep them minimal.
   - The objective then becomes to minimize both the margin size and the sum of the slack variables, effectively controlling the trade-off between margin size and misclassification.

3. **Tuning the Regularization Parameter (C):**
   - In SVMs, the parameter C controls the penalty for misclassified data points and the width of the margin.
   - A larger C penalizes misclassifications more heavily, which can result in a smaller margin but fewer misclassifications.
   - A smaller C allows more misclassifications but may lead to a larger margin.

4. **Handling Non-linearly Separable Data:**
   - For data that cannot be separated by a linear boundary, SVMs can use kernel functions to map the data into a higher-dimensional space where separation is possible.
   - This allows SVMs to handle more complex decision boundaries that are not linear.

In summary, the problem of misclassification when fixing the maximum margin in SVMs is addressed by allowing a soft margin approach. This means accepting some misclassifications to achieve a wider and more realistic margin, which typically leads to better generalization on unseen data.

UNIT 4

7A) GIVEN THE TWO OBJECTS REPRESENTED BY THE TUPLES (22, 1, 42, 10) AND (20, 0, 36, 8); COMPUTE

1) EUCLIDEAN DISTANCE BETWEEN THE TWO OBJECTS.

2) MANHATTAN DISTANCE BETWEEN THE TWO OBJECTS.

3) MURKOWSKI DISTANCE BETWEEN THE TWO OBJECTS, USING=3.

7A-> SOLUTION

YouTube Video Link: K-MEANS ALGORITHM

7B) STATE THE STRENGTHS AND WEAKNESSES OF K-MEANS CLUSTERING ALGORITHM. (7M)

K-MEANS CLUSTERING ALGORITHM: K-Means Clustering is an Unsupervised Machine Learning algorithm, which groups the unlabeled dataset into different clusters.

Strengths:

1. **Simplicity**: Easy to implement and understand.
2. **Efficiency**: Computationally efficient with a time complexity of O(knT) where k is the number of clusters, n is the number of data points, and T is the number of iterations.
3. **Scalability**: Works well with large datasets.

Weaknesses:

1. **Initial Centroids**: The final result is sensitive to the initial choice of centroids.
2. **Fixed Number of Clusters**: Requires the number of clusters k to be specified in advance.
3. **Shape Assumption**: Assumes spherical cluster shapes, which may not work well for irregular shapes.
4. **Outlier Sensitivity**: Sensitive to outliers and noise.

8) SUPPOSE THAT THE DATA MINING TASK IS TO CLUSTER THE FOLLOWING EIGHT POINTS (WITH (X, Y) REPRESENTING LOCATION) INTO THREE CLUSTERS:

A1(2, 10), A2(2, 5), A3(8, 4), B1(5, 8), B2(7, 5), B3(6, 4), C1(1, 2), C2(4, 9).

THE DISTANCE FUNCTION IS MANHATTAN DISTANCE. SUPPOSE INITIALLY WE ASSIGN A1, B1, AND C1 AS THE CENTER OF EACH CLUSTER, RESPECTIVELY.

USE THE K-MEANS ALGORITHM TO SHOW ONLY

(I) THE THREE CLUSTER CENTERS AFTER THE FIRST ROUND EXECUTION.

(II) THE FINAL THREE CLUSTERS.

8-> SOLUTION

UNIT 5

9A) DESIGN A NEURAL NETWORK THAT IMPLEMENTS BOOLEAN AND GATE. (8M)

9A-> ANSWER

9B) WHAT IS ANN? BRIEFLY EXPLAIN THE PROBLEM CHARACTERISTICS FOR WHICH ANN IS MOST COMMONLY USED. (6M)

An Artificial Neural Network (ANN) is a computational model inspired by the way biological neural networks in the brain process information. It consists of interconnected groups of artificial neurons (nodes) which work together to solve specific problems.

Characteristics of problems suited for ANN:

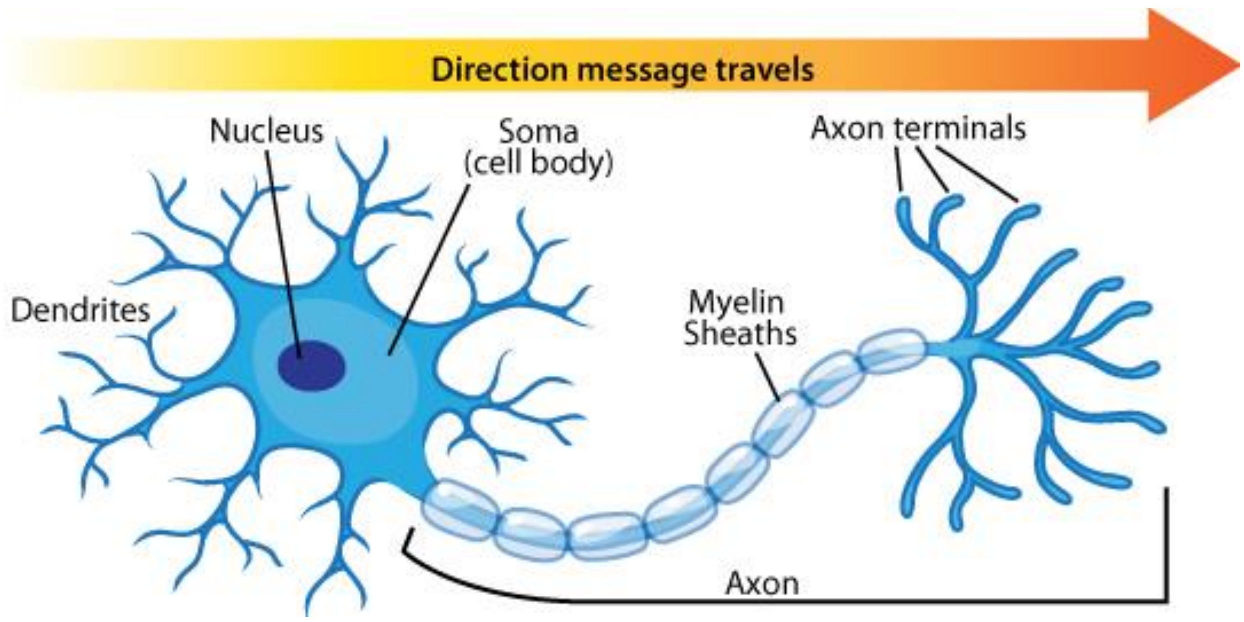
1. **Non-linear relationships**: ANNs can model complex, non-linear relationships between inputs and outputs.
2. **Pattern recognition**: Excellent for tasks like image and speech recognition where patterns are key.
3. **Data with noise**: Robust against noisy and incomplete data, making them suitable for real-world applications.
4. **Adaptation**: Capable of learning and adapting to new data through training.
5. **Generalization**: Good at generalizing from training data to unseen data.

YouTube Video Link: ANN

10A) WRITE THE BACK-PROPAGATION ALGORITHM. HOW THE ERROR IS MINIMIZED IN THE BACK-PROPAGATION ALGORITHM. (7M)

10A-> SOLUTION (PART1, PART2, PART3)

10B) DRAW THE STRUCTURE OF A BIOLOGICAL NEURON AND EXPLAIN THE FUNCTIONALITY OF AXONS AND DENDRITES OF THE NEURON. (7M)



A biological neuron consists of the following parts:

1. **Cell Body (Soma)**: Contains the nucleus and processes incoming signals.
2. **Dendrites**: Branch-like structures that receive signals from other neurons.
3. **Axon**: A long, thin structure that transmits signals away from the cell body to other neurons.
4. **Axon Terminals**: The endpoints of the axon where signals are transmitted to other neurons.
5. **Synapse**: The junction between neurons where signal transmission occurs.

Functionality:

- **Dendrites**: Receive electrical signals from other neurons and convey them towards the cell body.
- **Axon**: Conducts electrical impulses (action potentials) away from the cell body towards other neurons or effectors. Axons can be covered with a myelin sheath which speeds up signal transmission.
- **Axon Terminals**: Release neurotransmitters to transmit the signal to the next neuron across the synapse.