

# MACHINE LEARNING – UNIT II

## NOTES

### RGPV Exam-Oriented Notes

 Easy + Detailed + One-Day Before Exam Preparation

---

## UNIT II SYLLABUS

### Topics Covered

1. Learning a Class from Examples
  2. Classification
  3. Linear Classification
  4. Non-Linear Classification
  5. Multi-Class Classification
  6. Multi-Label Classification
  7. Decision Trees (ID3)
  8. CART
  9. Linear Regression
  10. Multiple Linear Regression
  11. Logistic Regression
  12. Neural Networks
  13. Perceptron
  14. Multilayer Perceptron
  15. Support Vector Machine (SVM)
  16. Kernel Functions
  17. K-Nearest Neighbors (KNN)
-

# 1. LEARNING A CLASS FROM EXAMPLES

## Definition

Learning a class from examples means training a machine using examples so it can classify new data correctly.

---

## Easy Explanation

The machine studies already given examples and learns patterns.

Example:

Input → Student Study Hours

Output → Pass / Fail

Machine learns:

More study hours → Higher chance of passing

---

## Steps

Training Data



Learning Algorithm



Model Creation



Prediction

---

## **Advantages**

- ✓ Learns automatically
  - ✓ Improves prediction accuracy
- 

## **Applications**

- Email spam filtering
  - Face recognition
  - Disease prediction
- 

## **Conclusion**

Learning from examples helps machines classify future data intelligently.

---

## **2. CLASSIFICATION**

### **Definition**

Classification is a supervised learning technique used to predict categorical outputs.

---

## Example

Email → Spam / Not Spam

---

## Types of Classification

Type	Meaning
Binary Classification	Two classes
Multi-Class Classification	More than two classes
Multi-Label Classification	Multiple labels together

---

## Applications

- Fraud detection
  - Medical diagnosis
  - Sentiment analysis
- 

## Conclusion

Classification predicts categories based on training data.

---

## 3. LINEAR CLASSIFICATION

### Definition

Linear Classification separates data using a straight line.

---

## Diagram

Class A ● ● ●  
----- Line  
Class B ▲ ▲ ▲

---

## Example

Pass or Fail prediction using marks.

---

## Advantages

- Simple
  - Fast training
- 

## Disadvantages

- Cannot handle complex data
- 

## Conclusion

Linear classification works well when data can be separated using a straight line.

---

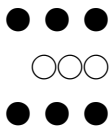
# 4. NON-LINEAR CLASSIFICATION

## Definition

Non-linear classification separates data using curves or complex boundaries.

---

## Diagram



Straight line cannot separate data.

---

## Example

Face recognition systems.

---

## Advantages

Handles complex patterns

---

## Disadvantages

More computation required

---

## **Conclusion**

Non-linear classification is used for complex real-world data.

---

## **5. MULTI-CLASS CLASSIFICATION**

### **Definition**

Classification involving more than two classes.

---

### **Example**

Animal → Cat / Dog / Horse

---

### **Applications**

- Image recognition
  - Language detection
- 

### **Conclusion**

Multi-class classification predicts one class among many possible classes.

---

## **6. MULTI-LABEL CLASSIFICATION**

## **Definition**

An object can belong to multiple classes simultaneously.

---

## **Example**

Movie categories:

Action + Comedy + Thriller

---

## **Applications**

- Image tagging
  - Music recommendation
- 

## **Conclusion**

Multi-label classification assigns multiple labels to a single instance.

---

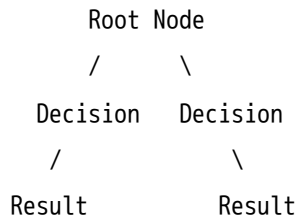
# **7. DECISION TREE**

## **Definition**

Decision Tree is a tree-structured supervised learning algorithm used for classification and prediction.

---

## Structure



## ID3 Algorithm

ID3 builds tree using:

Information Gain

Feature with highest information gain becomes root node.

---

## Advantages

- ✓ Easy to understand
  - ✓ Fast prediction
  - ✓ Handles categorical data
- 

## Disadvantages

- ✗ Overfitting problem
-

## Applications

- Medical diagnosis
  - Customer prediction
- 

## Conclusion

Decision Tree is simple and widely used for classification problems.

---

## 8. CART (Classification and Regression Tree)

### Definition

CART is a decision tree algorithm used for both classification and regression.

---

### Features

- ✓ Uses binary splits
  - ✓ Handles numerical data
  - ✓ Produces regression trees
- 

### Difference Between ID3 and CART

ID3	CART
Uses Information Gain	Uses Gini Index

ID3	CART
Only Classification	Classification + Regression
Multi-way split	Binary split

---

## Conclusion

CART is a powerful tree-based algorithm for classification and regression tasks.

---

## 9. LINEAR REGRESSION

### Definition

Linear Regression predicts continuous output using a straight-line relationship.

---

### Equation

$$y = mx + c$$

mmm

1.0

bbb

5.0

---

### Example

Predicting house price based on area.

---

## **Advantages**

- ✓ Simple
  - ✓ Easy interpretation
- 

## **Disadvantages**

- ✗ Works only for linear data
- 

## **Applications**

- Price prediction
  - Sales forecasting
- 

## **Conclusion**

Linear Regression predicts continuous values using linear relationships.

---

# **10. MULTIPLE LINEAR REGRESSION**

## **Definition**

Uses multiple input variables for prediction.

---

## Equation

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 = a + b_1x_1 + b_2x_2 + b_3x_3$$

---

## Example

House price prediction using:

- Area
  - Rooms
  - Location
- 

## Conclusion

Multiple Linear Regression improves prediction accuracy using multiple variables.

---

# 11. LOGISTIC REGRESSION

## Definition

Logistic Regression is used for classification problems.

---

## Output

0 → No

1 → Yes

---

## **Example**

Disease prediction:

Disease Present / Not Present

---

## **Advantages**

Works well for binary classification

---

## **Disadvantages**

Cannot handle highly complex data

---

## **Conclusion**

Logistic Regression predicts categorical outputs.

---

# **12. NEURAL NETWORKS**

## **Definition**

Neural Networks are computing systems inspired by the human brain.

---

## Structure

Input Layer  
↓  
Hidden Layer  
↓  
Output Layer

---

## Applications

- Image recognition
  - Speech recognition
  - AI systems
- 

## Advantages

- ✓ Learns complex patterns
  - ✓ High accuracy
- 

## Disadvantages

- ✗ Requires large data
  - ✗ Training time high
- 

## Conclusion

Neural Networks are powerful models for complex machine learning problems.

---

## 13. PERCEPTRON

### Definition

Perceptron is the simplest neural network model.

---

### Working

Inputs → Weights → Summation → Output

---

### Advantages

Simple implementation

---

### Disadvantages

Cannot solve non-linear problems

---

### Conclusion

Perceptron performs binary classification using weighted inputs.

---

# 14. MULTILAYER PERCEPTRON

## (MLP)

### Definition

MLP is a neural network containing multiple hidden layers.

---

### Diagram

Input → Hidden → Hidden → Output

---

### Advantages

- ✓ Solves non-linear problems
  - ✓ Better accuracy
- 

### Disadvantages

- ✗ Computationally expensive
- 

### Conclusion

MLP is widely used for deep learning applications.

---

# 15. SUPPORT VECTOR MACHINE

## (SVM)

### Definition

SVM is a supervised learning algorithm that finds the best separating boundary between classes.

---

### Diagram

Class A ● ● ●  
----- Hyperplane  
Class B ▲ ▲ ▲

---

### Types

Type	Meaning
Linear SVM	Straight boundary
Non-linear SVM	Curved boundary

---

### Advantages

- ✓ High accuracy
  - ✓ Effective in high-dimensional data
-

## Disadvantages

 Slow for large datasets

---

## Applications

- Face recognition
  - Text classification
- 

## Conclusion

SVM is a powerful classification algorithm with strong accuracy.

---

## 16. KERNEL FUNCTIONS

### Definition

Kernel Functions convert non-linear data into higher dimensions for separation.

---

### Types

Kernel	Use
Linear	Linear data
Polynomial	Polynomial boundary
RBF	Complex data

---

## Conclusion

Kernel functions help SVM solve non-linear classification problems.

---

## 17. K-NEAREST NEIGHBORS (KNN)

### Definition

KNN classifies data based on nearest neighboring points.

---

### Working

Find nearest K neighbors

↓

Majority Voting

↓

Classification

---

### Example

If most nearby neighbors are “Red”, new point becomes “Red”.

---

### Advantages

- ✓ Simple
- ✓ No training phase

---

## **Disadvantages**

- ✗ Slow for large data
- 

## **Applications**

- Recommendation systems
  - Pattern recognition
- 

## **Conclusion**

KNN predicts class using nearest neighboring data points.

---

## **MOST IMPORTANT 7-MARK** **QUESTIONS**

1. Explain Decision Tree with ID3 Algorithm.
  2. Explain Linear Regression.
  3. Explain Logistic Regression.
  4. Explain Neural Networks.
  5. Explain Perceptron model.
  6. Explain Multilayer Perceptron.
  7. Explain SVM with example.
  8. Explain Kernel Functions.
  9. Explain KNN Algorithm.
  10. Differentiate Linear and Logistic Regression.
-



# MOST IMPORTANT 14-MARK QUESTIONS

## ★ Very Important

1. Explain Decision Tree and ID3 Algorithm in detail.
  2. Explain SVM and Kernel Functions.
  3. Explain Neural Networks and Multilayer Perceptron.
  4. Explain Linear and Multiple Linear Regression.
  5. Explain Logistic Regression with example.
  6. Explain KNN Algorithm with applications.
  7. Explain classification and its types.
- 



## PYQ-BASED EXPECTED QUESTIONS

### ★ High Probability

- Explain Decision Tree using ID3.
  - Explain SVM.
  - Explain Neural Networks.
  - Explain KNN.
  - Explain Linear Regression.
- 

### ★ Medium Probability

- Explain Logistic Regression.
  - Explain CART.
  - Explain Perceptron Learning Algorithm.
-

# ONE-NIGHT REVISION

- ✓ Decision Tree uses tree structure
  - ✓ ID3 uses Information Gain
  - ✓ Regression predicts continuous values
  - ✓ Logistic Regression predicts categories
  - ✓ Neural Network inspired by brain
  - ✓ Perceptron = single layer network
  - ✓ MLP = multiple hidden layers
  - ✓ SVM finds optimal hyperplane
  - ✓ Kernel Functions solve non-linear problems
  - ✓ KNN uses nearest neighbors
- 

## SMART STUDY PLAN

### FIRST STUDY

1. Decision Tree
  2. SVM
  3. Neural Networks
  4. Linear Regression
- 

### SECOND PRIORITY

5. Logistic Regression
6. KNN

7. Perceptron

---

## **LAST REVISION**

8. CART

9. Kernel Functions

10. Classification Types

---

## **FINAL EXAM TIP**

For 14-mark answers always write:

1. Definition
2. Easy Explanation
3. Diagram
4. Example
5. Advantages
6. Applications
7. Conclusion

This presentation style helps score higher marks in RGPV exams.