PR INSEM PYQ

UNIT 1

1) Explain the Pattern. Pattern recognition and its application.

A **pattern** is a regularity or arrangement of elements that repeats in a predictable manner.
It can be **visual**, **numerical**, **textual**, or **behavioral**.

Examples:

- Visual: Stripes on a zebra.
- Numerical: Sequence 2, 4, 6, 8, ... (increases by 2).
- Behavioral: Customer buying habits over weekends.

In **computer science**, a pattern often refers to a *structured form of data* or *recognizable arrangement* that can be detected and analyzed.

Ans. : Pattern recognition can be defined as the categorization of input data into identifiable classes via the extraction of significant features or attributes of the data from a background of irrelevant detail.

3. Applications of Pattern Recognition

Pattern recognition is used widely in **AI, Machine Learning, and Data Analytics**.

Application Area	Example
Computer Vision	Face recognition in security systems
Speech Recognition	Voice assistants (Siri, Alexa, Google Assistant)
Medical Diagnosis	Detecting tumors in MRI scans
Biometrics	Fingerprint, iris, or retina scanning
Natural Language Processing	Spam email filtering, sentiment analysis
Robotics	Object detection for navigation
Financial Forecasting	Stock market trend prediction
Handwriting Recognition	Digitizing handwritten forms
Cybersecurity	Intrusion detection systems
Weather Prediction	Recognizing climate patterns

2) What are the different models or approaches of pattern recognition?

Statistical Approach:

- Statistical methods are mathematical formulas, models, and techniques that are used in the statistical analysis of raw research data.
- The application of statistical methods extracts information from research data and provides different ways to assess the robustness of research outputs.
- This approach is based on statistical decision theory. Pattern recognizer extracts quantitative features from the data along with the multiple samples and compares those features.

Example: Naive Bayes classifier for spam detection.

Applications: OCR (Optical Character Recognition), medical diagnosis.

Structural /Syntactical Approach:

- This approach is closer to how human perception works.
- It extracts morphological features from one data sample and checks how those are connected and related.
- The Structural Approach is a technique wherein the learner masters the pattern of sentence.
- Structures are the different arrangements of words in one accepted style or the other.
  - Types of structures: Sentence Patterns
  - Phrase Patterns
  - Formulas

Example: Using grammar rules to validate programming code syntax.

Applications: DNA sequence analysis, chemical compound structure recognition.

Neural Pattern Recognition

- In this approach, artificial neural networks are utilized.
- Compared to the ones mentioned above, it allows more flexibility in learning and is the closest to natural intelligence.
- Normally, only feed-forward networks are used for pattern recognition. Feed-forward means that there is no feedback to the input.
- Similar to the way that human beings learn from mistakes, neural networks also could learn from their mistakes by giving feedback to the input patterns.
- This kind of feedback would be used to reconstruct the input patterns and make them free from error; thus increasing the performance of the neural networks.

Example: Convolutional Neural Networks (CNNs) for face recognition.

Applications: Image classification, speech recognition.

3) Explain feature extraction?

Feature Extraction – Meaning

Feature Extraction is the process of **transforming raw data into a set of measurable characteristics** (called *features*) that can be used to identify or describe a pattern.

📌 In simple words:

> "It's like summarizing a big paragraph into keywords that tell you the most important information."

Why it's Important?

- Raw data can be **huge, noisy, and redundant**.
- Feature extraction **reduces the size of the data** while keeping the important information needed for recognition.
- Makes **pattern recognition algorithms faster and more accurate**.

Steps in Feature Extraction

1. **Data Collection** → Gather raw input (image, text, audio, etc.).
2. **Preprocessing** → Clean the data (remove noise, normalize values).
3. **Identify Features** → Select measurable attributes (edges, shapes, frequency, etc.).
4. **Feature Transformation** → Convert features into numerical form.
5. **Feature Selection** → Keep only the most relevant features.

Types of Features

- **Statistical Features** → Mean, variance, standard deviation.
- **Structural Features** → Shape, edges, curves.
- **Texture Features** → Smoothness, contrast.
- **Frequency Features** → Fourier or wavelet transforms (for signals/images).

Example

Face Recognition System:

- Raw data: Image of a face (thousands of pixels).
- Features: Distance between eyes, nose width, jawline shape, skin tone histogram.
- Output: Feature vector → Used to compare with stored faces.

4) What is design cycle of pattern recognition?

Design Cycle of Pattern Recognition

The design cycle describes the **step-by-step process** of developing a pattern recognition system — from collecting data to recognizing and classifying patterns.

---

1. Data Collection

- **Description:** Gather raw data (images, text, signals, sensor readings, etc.) relevant to the patterns you want to recognize.
- **Example:** Taking images of handwritten digits for a digit recognition system.

---

2. Preprocessing

- **Description:** Improve the quality of raw data and remove noise.
- **Steps:** Normalization, filtering, background removal, scaling.
- **Example:** Converting images to grayscale and resizing them to a standard size.

3. Feature Extraction

- **Description:** Transform raw data into a compact and informative set of **features**.
- **Purpose:** Reduce data size and retain essential information for recognition.
- **Example:** In speech recognition, extract pitch, frequency, and energy.

---

4. Feature Selection

- **Description:** Choose the most relevant features from the extracted set to improve accuracy and efficiency.
- **Example:** Selecting only the edges and corners in object recognition, ignoring color if it's irrelevant.

---

5. Model Selection / Training

- **Description:** Choose a recognition model (statistical, syntactic, neural network, etc.) and train it using labeled data.
- **Example:** Using a Convolutional Neural Network (CNN) for image classification.

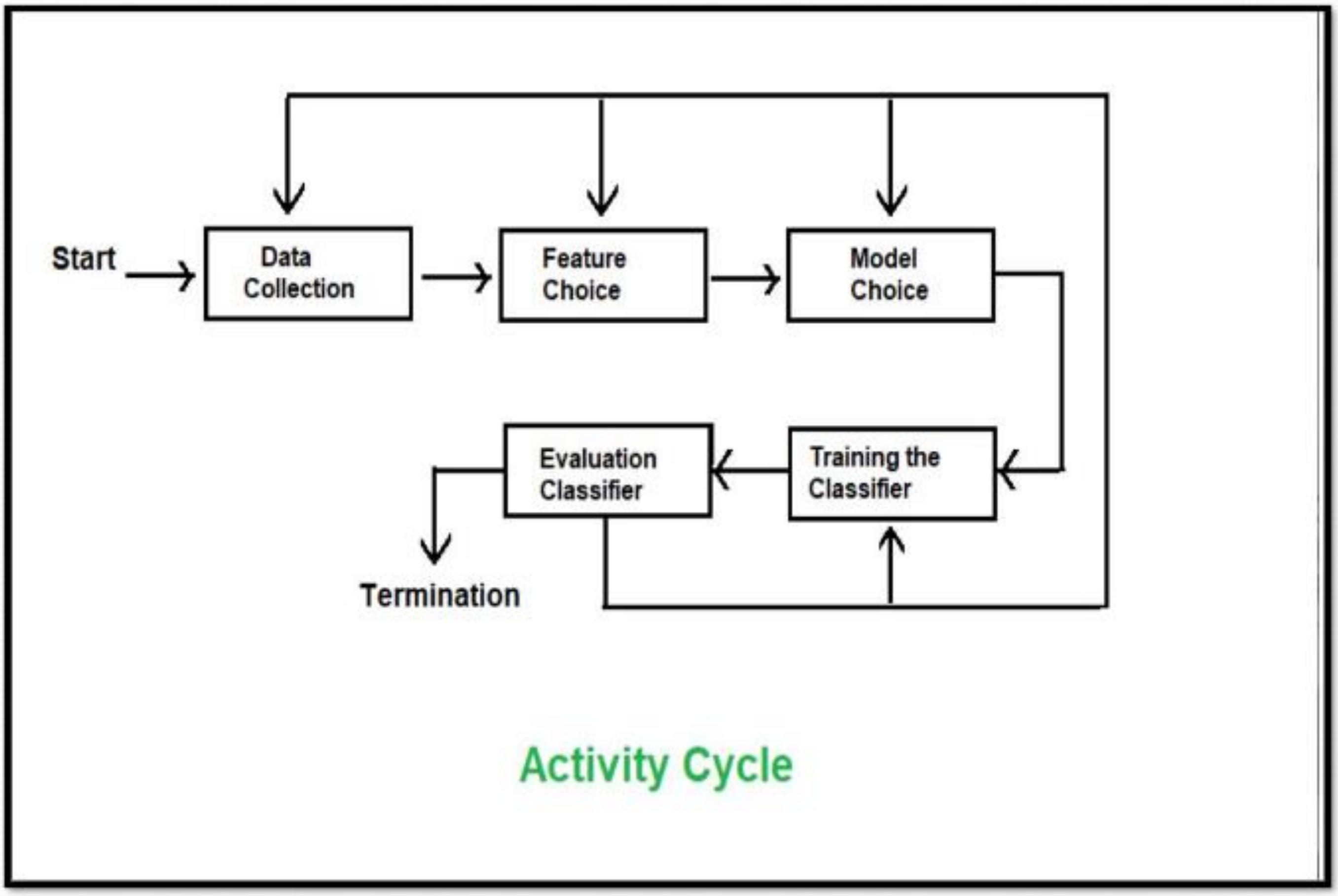
---

6. Classification / Recognition

- **Description:** Compare new input patterns with the trained model and assign them to a class.
- **Example:** Classifying a handwritten "5" as digit 5.

7. Post-processing & Evaluation

- **Description:** Improve results by refining outputs and evaluating system performance using accuracy, recall, precision, etc.
- **Example:** Applying spell correction after OCR to fix misrecognized words.



5) What is Neural Pattern Recognition Approach?

Neural Pattern Recognition Approach

Definition

The **Neural Pattern Recognition Approach** uses **Artificial Neural Networks (ANNs)** — computational models inspired by the human brain — to learn and recognize patterns from data.

> It works by training interconnected layers of artificial "neurons" to detect patterns and classify them.

How It Works

1. **Input Layer** → Receives features extracted from raw data.
2. **Hidden Layers** → Perform processing through weighted connections and activation functions to detect complex relationships.
3. **Output Layer** → Produces the final classification or recognition result.
4. **Learning Process** → Adjusts connection weights based on errors (e.g., using backpropagation).

Key Characteristics

- **Data-driven:** Learns patterns from examples, not hard-coded rules.
- **Non-linear Mapping:** Can model complex relationships between features.
- **Adaptive:** Improves performance as more data is provided.
- **Robust to Noise:** Can still work well even with imperfect data.

Example

- **Face Recognition:**
  - Input → Pixel values of an image.
  - Network learns to detect features like eyes, nose, mouth.
  - Output → Person's identity.

Applications

- **Computer Vision:** Object and face recognition.
- **Speech Recognition:** Voice assistants like Siri and Alexa.
- **Medical Imaging:** Tumor detection in X-rays or MRIs.
- **Handwriting Recognition:** Digitizing handwritten notes.
- **Fraud Detection:** Recognizing unusual patterns in transactions.

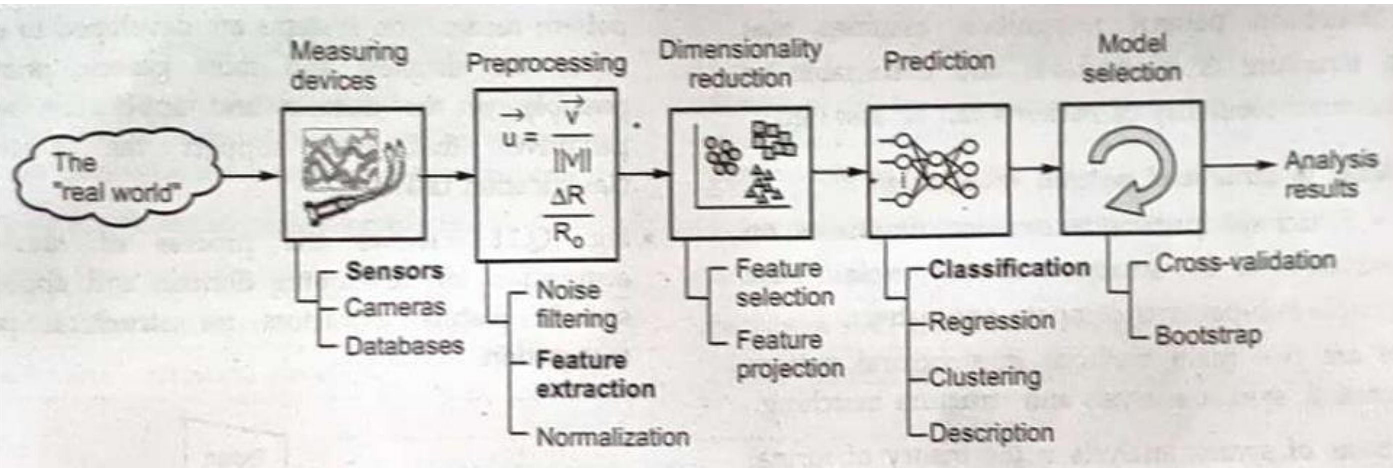
Advantages

- Can handle large and complex datasets.
- Automatically extracts important features.
- High accuracy after sufficient training.

Disadvantages

- Requires a lot of training data.
- Computationally expensive.
- Often works as a "black box" — hard to interpret how it makes decisions.

6) Explain in details of Pattern Recognition System & what need of pattern recognition system



Refer Q. 4

3. Need for Pattern Recognition Systems

Pattern recognition systems are essential because:

1. **Automation**
   - Reduces human effort in tasks like sorting, detecting, or classifying data.
   - Example: Automated postal address reading.
2. **Speed and Efficiency**
   - Processes large volumes of data much faster than humans.
   - Example: Real-time face detection in surveillance cameras.
3. **Accuracy**
   - Advanced algorithms can often detect patterns more reliably than humans, especially in noisy or complex data.
4. **Handling Complex Data**
   - Can deal with high-dimensional, large-scale, and complex datasets that are difficult to analyze manually.
5. **Cost-effectiveness**
   - Once developed, systems reduce labor costs in repetitive recognition tasks.
6. **Consistency**
   - Unlike humans, they do not suffer from fatigue or bias; the output is consistent.
7. **Enabling Intelligent Applications**
   - Pattern recognition is the foundation for AI applications like self-driving cars, medical imaging systems, and fraud detection.

7) Describe Pattern and feature with suitable example.

1. Pattern – Definition

A **pattern** is an arrangement or structure of data that follows certain rules or regularities, which can be observed and recognized.

In pattern recognition, a pattern represents an **object, event, or signal** to be classified.

📌 **Example:**

- Handwritten digit "5" (as an image) is a pattern in a digit recognition system.
- Voice sample of a person in speech recognition.

2. Feature – Definition

A **feature** is a measurable property or characteristic of a pattern that helps in distinguishing it from other patterns.

Features are used as inputs for pattern classification.

📌 **Example:**

- In handwritten digit recognition: number of curves, pixel intensity, height-to-width ratio.
- In face recognition: distance between eyes, nose width, skin tone.

Features may be represented as continuous, discrete, or discrete binary variables. A feature is a function of one or more measurements, computed so that it quantifies some significant characteristics of the object.

Example: consider our face then eyes, ears, nose, etc are features of the face.
A set of features that are taken together, forms the **features vector**.

Example: In the above example of a face, if all the features (eyes, ears, nose, etc) are taken together then the sequence is a feature vector([eyes, ears, nose]).

8) Explain the feature selection process in detail

1. Definition

Feature Selection is the process of selecting the most **relevant and important features** from the extracted set, in order to improve recognition accuracy and reduce computational cost.

📌 In short:

> From a big list of possible features, we keep only the ones that matter most for classification.

---

2. Need for Feature Selection

- Reduces **dimensionality** of data (less storage & processing).
- Removes **irrelevant or redundant** features.
- Improves **speed** of the recognition system.
- Often increases **accuracy** by removing noise features.

✅ Steps in Feature Selection Process:
Feature Generation:

All potential features are extracted from data (e.g., 100 features from an image).

Feature Ranking/Scoring:

Each feature is ranked based on criteria like:

Information gain

Mutual information

Correlation with output labels

Subset Selection:

Select the top-ranked features or use algorithms (e.g., Recursive Feature Elimination) to select optimal subset.

Model Evaluation:

Evaluate the model performance using selected features via metrics like accuracy or F1-score.

Finalization:

Best feature subset is finalized for the pattern recognition system.

9) Compare Classification and Recognition in Pattern Recognition?

1. Classification – Definition

Classification is the process of **assigning an input pattern to one of several predefined classes** based on its features.

📌 Example: Classifying an email as *spam* or *not spam*.

---

2. Recognition – Definition

Recognition is the **overall process of identifying a pattern**, which may involve **classification, description, or matching**.

📌 Example: Recognizing a handwritten digit as "5" (involves identifying and confirming it matches known patterns).

Aspect	Classification	Recognition
Definition	Assigns input to a specific class	Identifies and interprets the input pattern
Scope	Narrow, focused only on class assignment	Broad, may include classification and other analysis
Process	Works on extracted features to decide class	May include preprocessing, feature extraction, classification
Example	Classifying an image as "cat"	Recognizing that the image is a *domestic short-haired cat* in a garden
Role	Step in recognition process	Complete identification process

10) Comparison: Syntactic, Statistical, and Neural Pattern Recognition Approaches

Aspect	Syntactic Pattern Recognition (SyntPR)	Statistical Pattern Recognition (StatPR)	Neural Pattern Recognition (NeurPR)
Basic Idea	Recognizes patterns using grammar-like rules and structural relationships.	Recognizes patterns by measuring statistical properties (probabilities, distributions).	Recognizes patterns by learning from data using Artificial Neural Networks.
Representation	Symbols, strings, and production rules.	Feature vectors and probability models.	Layers of neurons with weights and activation functions.
Best For	Patterns with a well-defined structure (syntax).	Problems where patterns can be described with numerical features and statistics.	Complex, high-dimensional, and unstructured patterns.
Example	Chemical structure analysis, language syntax checking.	Handwritten digit recognition (using probability of features), medical diagnosis.	Face recognition, voice recognition, object detection.
Advantages	Interpretable, good for structured data.	Works well with limited data, mathematically sound.	Handles complex, noisy data, high accuracy with enough data.
Limitations	Fails with noisy/unstructured data.	Assumes data follows certain statistical models.	Requires large datasets, high computation.

UNIT 2

11) Explain the statistical pattern recognition.

Statistical Pattern Recognition

Definition

Statistical Pattern Recognition is a method of identifying and classifying patterns based on **statistical properties** of their features.
It uses probability theory to decide which class a given pattern most likely belongs to.

How It Works

1. **Feature Extraction** – Measure important characteristics of the pattern (e.g., length, width, color intensity).
2. **Statistical Modeling** – Represent these features using probability distributions (often Gaussian / Normal distribution).
3. **Classification** – Assign the pattern to the class with the highest probability (using decision rules like Bayes' theorem).

Key Concepts

- **Feature Vector**: A list of measured attributes for a pattern.
- **Probability Density Function (PDF)**: Describes how likely certain feature values are.
- **Decision Boundary**: A line or surface that separates different classes in feature space based on probability values.

Example

Suppose we are classifying fruits:

- Features: weight, color, diameter.
- If an unknown fruit has a high probability of matching the "apple" distribution, it is classified as an apple.

Advantages

- Based on solid mathematical theory.
- Works well when data distribution is known or can be estimated.
- Can handle noise in measurements.

Limitations

- Requires a large dataset to estimate accurate probability models.
- Assumes features follow certain distributions, which may not always be true.

12) What are the different types of pattern classification?

Difference Between Types of Pattern Classification

Aspect	Supervised Classification	Unsupervised Classification	Semi-supervised Classification	Statistical Classification
Training Data	Uses **labeled data** (each example has a known class).	Uses **unlabeled data** (no class labels).	Uses **few labeled** + **many unlabeled** examples.	Uses labeled data with **statistical models** of feature distributions.
Goal	Learn a mapping from features to known classes.	Group similar patterns into clusters without knowing classes.	Improve classification using limited labeled data.	Assign a class based on highest probability.
Learning Method	Algorithms like Decision Trees, SVM, Neural Networks.	Clustering algorithms like K-means, Hierarchical clustering.	Combination of supervised & unsupervised techniques.	Bayesian classifiers, Gaussian models, discriminant analysis.
Example	Spam vs Non-spam email detection.	Grouping news articles by topic without labels.	Classifying medical images with only few labeled examples.	Handwriting recognition using Gaussian probability models.
Advantages	High accuracy with enough labeled data.	Useful when labels are unavailable.	Improves performance when labeling is costly.	Based on solid mathematical theory; interpretable.
Limitations	Needs large labeled datasets.	May group incorrectly without prior info.	Still needs some labeled data.	Assumes features follow a known distribution.

Layman's Explanation

- **Supervised** → Like learning to identify fruits after being shown pictures with names.
- **Unsupervised** → Like sorting fruits into groups based on looks, without knowing their names.
- **Semi-supervised** → You know a few fruits' names, guess the rest by similarity.
- **Statistical** → You guess the fruit based on the probability of its size, color, and shape matching known fruits.

13) What is an unsupervised learning approach in pattern recognition?

Unsupervised learning

- Unsupervised learning analyzes datasets that don't have any labels or any secondary information (such as metadata) that could be used as labels.
- Instead, an unsupervised machine learning system looks across the breadth of unprocessed data for recurrent patterns and attempts to perform classification and labeling tasks without human supervision.

Advantages of unsupervised learning

- Data discoverability
- Improved anomaly detection
- Automated preprocessing

Clustering Types of Unsupervised Learning Algorithms

- **Exclusive (partitioning)**:In this clustering method, Data are grouped in such a way that one data can belong to one cluster only. Ex: K-means
- **Agglomerative:** In this clustering technique, every data is a cluster. The iterative unions between the two nearest clusters reduce the number of clusters. Ex.Hierarchical clustering
- **Overlapping**: fuzzy sets is used to cluster data. Each point may belong to two or more clusters with separate degrees of membership.Ex: Fuzzy C-Means
- **Probabilistic**:This technique uses probability distribution to create the clusters
- Example: “man’s shoe.”,“women’s shoe.”,“women’s glove.”,“man’s glove.”
- can be clustered into two categories “shoe” and “glove” or “man” and “women.”

📌 Quick Analogy:

- **Exclusive** → One student is in only one classroom.
- **Agglomerative** → Start with each student alone, keep grouping them into bigger classes.
- **Overlapping** → A student can be in both the music club and sports club.
- **Probabilistic** → A student has a 60% chance of being in science club and 40% in art club — we choose the higher one.

14) What is the Gaussian distribution classification?

Gaussian distribution classification is a statistical method that classifies patterns based on the assumption that their features follow a Gaussian (Normal) distribution — the classic *bell curve* shape.

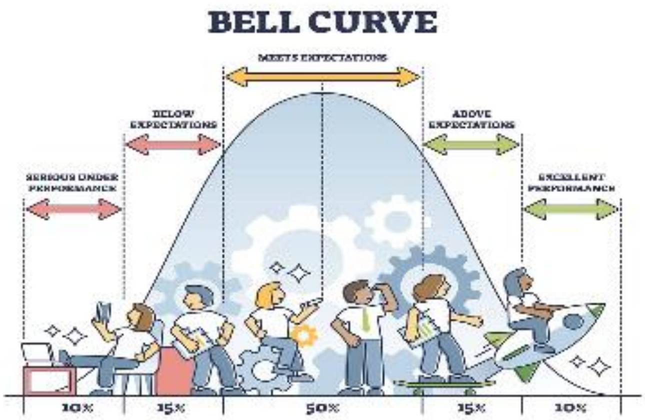
How It Works

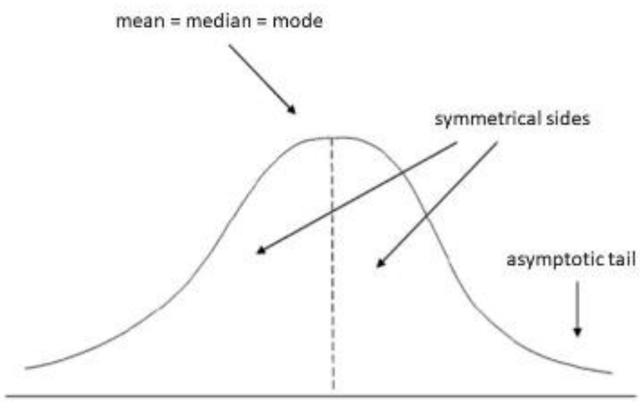
1. **Assume Data Distribution** – For each class, the features are assumed to be normally distributed.
2. **Estimate Parameters** – Find the **mean** (μ) and **standard deviation** (σ) for each class from training data.
3. **Probability Calculation** – Use the Gaussian probability density function (PDF) to find how likely a new data point belongs to each class:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

4. **Classification Decision** – Assign the data point to the class with the **highest probability**.

- Definition of Gaussian:-**being or having the shape of a normal curve or a normal distribution.**





What Is a Normal Distribution?

- Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean.

- In graphical form, the normal distribution appears as a "bell curve".

- In a normal distribution the mean is zero and the standard deviation is 1. It has zero skew and a kurtosis of 3.

- Normal distributions are symmetrical, but not all symmetrical distributions are normal.

- Many naturally-occurring phenomena tend to approximate the normal distribution.

- In finance, most pricing distributions are not, however, perfectly normal.

Properties of the Normal Distribution

- its mean (average), median (midpoint), and mode (most frequent observation) are all equal to one another.
- Moreover, these values all represent the peak, or highest point, of the distribution. The distribution then falls symmetrically around the mean, the width of which is defined by the standard deviation.
- *"All normal distributions can be described by just two parameters: the mean and the standard deviation."*

15) What are discriminant functions and its uses?

Definition

A **discriminant function** is a mathematical function used in pattern recognition to **separate different classes** in a feature space.
For a given input pattern x, each class i has its own discriminant function $g_i(x)$.
The input is assigned to the class with the **highest value** of $g_i(x)$.

Purpose

- To decide **which class a pattern belongs to.**
- To define **decision boundaries** between classes.

How It Works

1. **Input**: Feature vector of the pattern.
2. **Function Evaluation**: Calculate $g_i(x)$ for all classes.
3. **Decision Rule**: Assign to the class with the largest $g_i(x)$.

Common Types

- **Linear Discriminant Function (LDF)** – Decision boundary is a straight line/plane.
- **Quadratic Discriminant Function (QDF)** – Decision boundary is curved.

Uses

1. **Pattern Classification** – Assigning new samples to known categories.
2. **Face Recognition** – Deciding which person's profile matches best.
3. **Medical Diagnosis** – Classifying patients into disease categories.
4. **Document Classification** – Sorting emails into spam/non-spam.

16) What are risk and errors in measurement of classification performance?

Risk and Errors in Classification Performance

1. Errors in Classification

In classification, **error** means assigning a pattern to the wrong class.

Types of Errors:

1. **Type I Error (False Positive)** – Predicting a pattern belongs to a class when it actually does not.
   - Example: Classifying a healthy email as spam.
2. **Type II Error (False Negative)** – Predicting a pattern does not belong to a class when it actually does.
   - Example: Missing a spam email and putting it in the inbox.

Error Rate Formula:

$$\text{Error Rate} = \frac{\text{Number of Wrong Predictions}}{\text{Total Predictions}}$$

2. Risk in Classification

- **Definition:** Risk is the **expected loss** when making classification decisions, considering both the probability of errors and the cost of each type of error.
- **Formula:**

$$R = \sum_i \sum_j \lambda_{ij} \cdot P(w_i|x)$$

Where:

- λ_{ij} = Loss for classifying an object from class i into class j
- $P(w_i|x)$ = Probability that the object belongs to class i given input x

---

3. Measurement Metrics

- **Accuracy** – Proportion of correctly classified patterns.
- **Precision** – Correct positive predictions out of all positive predictions.
- **Recall (Sensitivity)** – Correct positive predictions out of all actual positives.
- **F1-Score** – Harmonic mean of precision and recall.
- **Confusion Matrix** – Table showing correct and incorrect predictions for each class.

✅ **In short:**

- Error is about how often we're wrong.
- Risk is about how bad it is to be wrong.

17) Describe the different approaches to developing Statistical Pattern Recognition (StatPR) Classifiers.

Different Approaches to Developing Statistical Pattern Recognition Classifiers

Statistical Pattern Recognition (StatPR) classifiers are developed based on probability models of pattern features.
The main approaches are:

---

1. Parametric Approach

- **Idea:**
  Assumes that the probability distribution of features is known (or can be approximated) with a fixed number of parameters.
- **Steps:**
  1. Choose a probability model (e.g., Gaussian distribution).
  2. Estimate parameters like **mean (μ)** and **covariance (Σ)** from training data.
  3. Apply decision rules (e.g., Bayes' theorem) for classification.
- **Example:** Gaussian classifier with estimated μ and Σ.

2. Non-Parametric Approach

- **Idea:**
  Makes no assumption about the underlying probability distribution.
- **Steps:**
  1. Use training data directly to estimate probabilities.
  2. Classify by measuring similarity/distance to stored samples.
- **Example:**
  - **k-Nearest Neighbour (k-NN)** – Classifies based on majority class among *k* nearest samples.
  - **Parzen Windows** – Estimates probability density without assuming a model.

---

3. Semi-Parametric Approach

- **Idea:**
  Combines both parametric and non-parametric methods.
- **Example:**
  - Gaussian Mixture Models (GMMs) where each component is parametric but overall distribution is flexible.

Layman's Terms

Think of it like guessing **a person's favorite food**:

1. **Parametric Approach** –
   You already know they like food in a *specific category* (say, Italian), you just need to figure out the exact dish by asking a few questions.
   → Fixed shape, just fill in the details.
2. **Non-Parametric Approach** –
   You have **no idea** what category they like, so you keep showing them different foods and see which they choose the most.
   → No fixed shape, learn entirely from examples.
3. **Semi-Parametric Approach** –
   You know they like *spicy* food, but not which cuisine — so you try dishes from various cuisines but only spicy ones.
   → Partly fixed, partly flexible.

18) Discuss how to generate a Feature Vector with Density Function

A **feature vector** is a numerical representation of a pattern, consisting of its measurable characteristics (features).
A **density function** describes the probability distribution of these features.

Steps to generate a feature vector using a density function:

1. **Feature Measurement**
   - Collect measurable attributes from the pattern (e.g., height, width, color intensity).
2. **Probability Density Estimation**
   - Estimate the **probability density function** (PDF) for each feature using techniques like:
     - **Parametric**: Gaussian (Normal) distribution, Exponential distribution.
     - **Non-parametric**: Histogram method, Parzen windows, k-NN density estimation.
3. **Feature Vector Formation**
   - Convert each feature's density value into a vector form:

   $$\mathbf{f} = [p_1(x_1), p_2(x_2), \ldots, p_n(x_n)]$$

   where $p_i(x_i)$ is the density value for feature i.
4. **Normalization**
   - Normalize the vector so all features contribute fairly (avoid domination by large values).
5. **Use in Classifier**
   - Feed the feature vector into a classifier for pattern recognition.

Example:
In handwriting recognition:

- Features: stroke length, curve angle, pixel density.
- Each feature's probability density is computed (e.g., using Gaussian PDF).
- Combine into a feature vector for classification.

Layman's Terms

Imagine you're describing a fruit to someone who has never seen it:

1. You measure things like **weight**, color, and **sweetness level**.
2. For each measurement, you figure out **how common that value is** in known fruits (this is the density function).
3. You write these values down in a row — that's your **feature vector**.
4. Now, anyone can use this row of numbers to guess what fruit it is.

19) Describe any simple example of Statistical Model

A **statistical model** is a mathematical representation of data using probability distributions and parameters. It defines how data is generated and how it can be analyzed.

Example — Gaussian (Normal) Distribution Model:

- **Model Definition:**

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where:

- μ = mean (average)
- σ = standard deviation (spread of data)
- **Steps:**
  1. Collect training data.
  2. Calculate mean μ and standard deviation σ.
  3. Use the model to calculate the probability of new data points.
  4. Classify the point into the class with the highest probability.

Application Example:
Classifying exam scores into Pass/Fail categories based on their probability under a Gaussian model.

Layman's Terms

Think of a statistical model like **a recipe for guessing outcomes**:

- If you know most apples weigh around 150g (mean) and usually vary by ±20g (standard deviation), you can predict whether a new fruit is an apple by checking if its weight fits that pattern.
  It's basically a **math-based rulebook** for how things usually happen.

20) Describe Classifier Performance and Measurement

Definition:

Classifier performance refers to how well a classification model predicts the correct class labels for given inputs.
It is measured using various metrics that evaluate accuracy, error rate, and reliability.

Key Performance Metrics

Metric	Definition	Formula
Accuracy	Proportion of correctly classified samples.	$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$
Precision	Out of all predicted positives, how many are correct.	$\text{Precision} = \frac{TP}{TP+FP}$
Recall (Sensitivity)	Out of all actual positives, how many are correctly identified.	$\text{Recall} = \frac{TP}{TP+FN}$
F1-Score	Harmonic mean of precision and recall.	$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
Specificity	Out of all actual negatives, how many are correctly identified.	$\text{Specificity} = \frac{TN}{TN+FP}$

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives (Type I Error)
- FN = False Negatives (Type II Error)