ROHINI
COLLEGE OF ENGINEERING & TECHNOLOGY
Approved by AICTE and Affiliated to Anna University (An ISO Certified Institution) | Accredited with A+ Grade by NAAC
Recognized under Section 2(f) of University Grants Commission, UGC ACT 1956
(AUTONOMOUS)

STATISTICAL PATTERN RECOGNITION

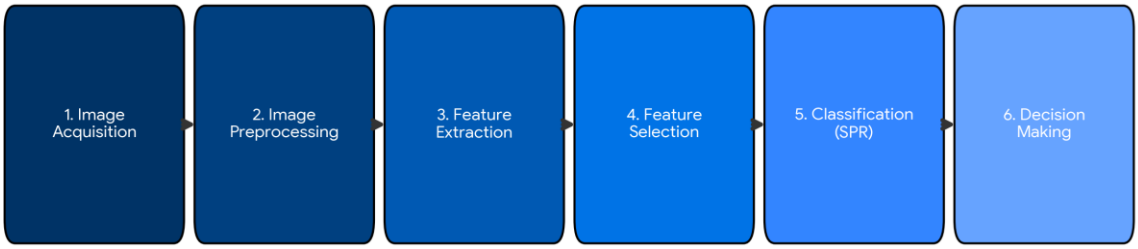
1. Introduction to Statistical Pattern Recognition (SPR)

Statistical Pattern Recognition (SPR) is a core technique in medical image analysis used to classify tissue types, detect anatomical structures, and diagnose pathologies automatically using mathematical models.

- **Pattern:** Quantitative descriptors derived from medical images (e.g., shape, intensity, texture).
- **Feature Vector:** A numerical vector representing the key visual features of an image region.
- **Supervised Classification:** Learning decision rules using labeled medical training data.

2. Medical Image Processing & Analysis Pipeline

The typical pipeline for processing and classifying medical images using SPR contains 6 key stages:



Stage 1: Image Acquisition

Capture raw images via X-Ray, CT, MRI, Ultrasound, or PET scans.

Stage 2: Preprocessing

Noise filtering (Gaussian, Median), contrast enhancement (CLAHE), and spatial normalization.

Stage 3: Feature Extraction

Extracting shape, texture (GLCM), intensity, and edge features from regions of interest (ROI).

Stage 4: Feature Selection

Reducing dimensionality via PCA or LDA to retain only discriminative features.

Stage 5: Classification

Applying statistical decision classifiers (Bayesian, KNN, LDA, SVM).

Stage 6: Decision Making

Final medical evaluation: Normal, Benign, Malignant, or Disease Severity Grade.

3. Statistical Classifiers in Medical Imaging

3.1 Bayesian Classifier

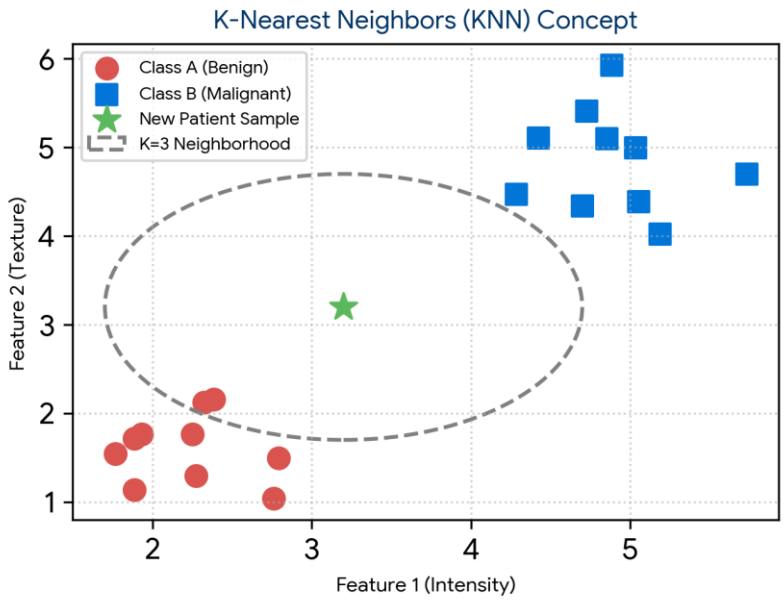
Calculates posterior probability using Bayes' Theorem:

> **Posterior Probability Formula:**
>
> **P(C | X) = [P(X | C) * P(C)] / P(X)**
>
> **Posterior = (Prior * Likelihood) / Evidence**

3.2 K-Nearest Neighbor (KNN)

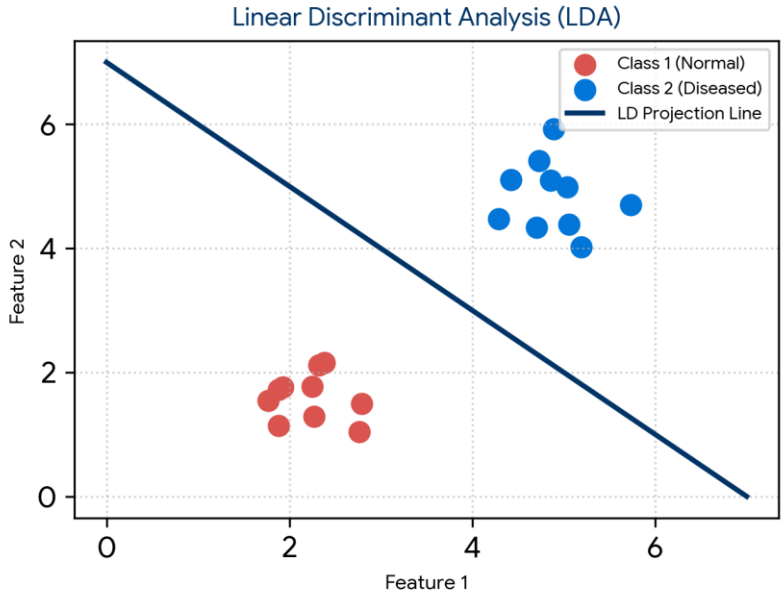
Classifies an unlabelled image sample by distance comparison with K nearest neighbors.

> **Euclidean Distance Formula:**
>
> **d(x, y) = sqrt(sum_{i=1}^{n} (x_i - y_i)^2)**



3.3 Linear Discriminant Analysis (LDA)

Projects features into lower dimensions to maximize separation between medical classes.

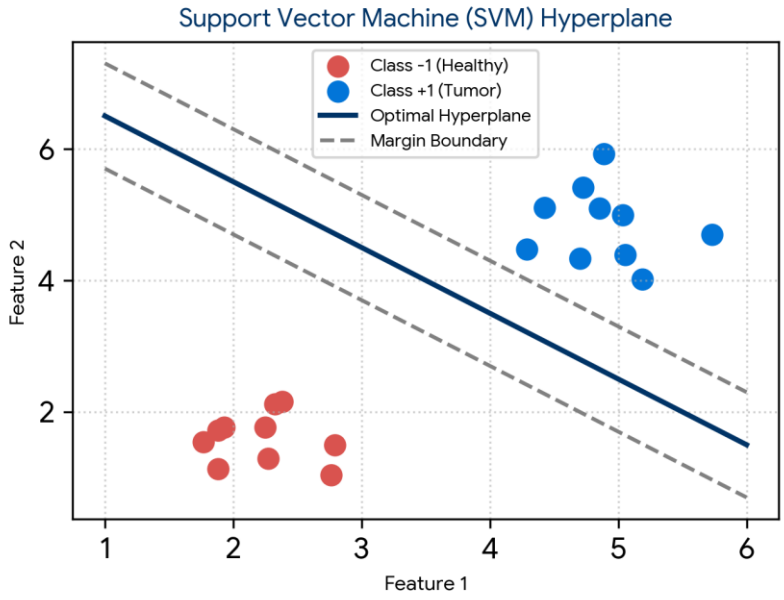


3.4 Support Vector Machine (SVM)

Constructs an optimal decision hyperplane in high dimensions with maximum margin separation.

> **Hyperplane Equation:**

> **w^T * x + b = 0**



4. Classifier Comparison Table

Classifier	Core Principle	Key Strengths	Medical Applications
Bayesian Classifier	Probabilistic Bayes' Theorem MAP rule.	Handles uncertainty well.	Mammogram microcalcification assessment.
K-Nearest Neighbor (KNN)	Distance-based voting among K neighbors.	Simple, non-parametric.	Histopathology cell pattern matching.
Linear Discriminant Analysis (LDA)	Maximizes between-class variance ratio.	Reduces dimensions.	CT scan nodule classification.
Support Vector Machine (SVM)	Optimal hyperplane with maximum margin.	Accurate in high dimensions.	Brain MRI tumor segmentation & detection.