What is Data in Machine Learning?

In machine learning, **data** refers to the raw information that is used to train models. It consists of **features (inputs)** and **labels (outputs, if available)** and can be structured (e.g., tables, databases) or unstructured (e.g., images, text, videos).

Types of Data in Machine Learning

1. Structured vs. Unstructured Data

Type	Description	Examples
Structured Data	Organized in a tabular format with rows and columns. Easy to analyze using databases.	Spreadsheets, SQL databases
Unstructured Data	Data without a predefined structure. Requires processing techniques like NLP and computer vision.	Images, videos, audio, text

2. Labeled vs. Unlabeled Data

Type	Description	Example
Labeled Data	Each data point has both input features and the corresponding correct output (label). Used in **supervised learning**.	A dataset of images with labels: ‘cat’ or ‘dog’.
Unlabeled Data	The dataset contains only input features without labels. Used in **unsupervised learning**.	A dataset of customer transactions without knowing their purchasing behavior.

Labeled Data Explanation

- In **supervised learning**, models learn from labeled data.
- Example: A dataset with **emails (features)** and labels like **"spam" or "not spam"**.

Unlabeled Data Explanation

- In **unsupervised learning**, models learn patterns without predefined labels.
- Example: Clustering customers based on their purchase behavior without predefined categories.

Significance of Dimensionality Reduction in Unsupervised Learning

Dimensionality reduction is a technique used to reduce the number of features in a dataset while preserving essential information. It is crucial in **unsupervised learning** for the following reasons:

1. Removes Redundant & Noisy Features

- Many real-world datasets have **high-dimensional features**, which can contain redundant or irrelevant information.
- Example: In customer segmentation, not all collected attributes may be useful.

2. Improves Model Performance

- High-dimensional data can cause the **curse of dimensionality**, where models perform poorly due to increased complexity.
- Reducing dimensions helps algorithms like **k-means clustering** work efficiently.

3. Enhances Visualization

- In **unsupervised learning**, data visualization is crucial for pattern recognition.
- **Techniques like PCA (Principal Component Analysis) and t-SNE** help project data into 2D or 3D for better insights.

4. Reduces Computation Cost

- Fewer dimensions mean **less storage and faster training**.
- Useful in big data applications.

5. Avoids Overfitting

- When data has too many features, models may learn noise instead of meaningful patterns.
- Reducing dimensions prevents overfitting and generalizes the model better.

Popular Dimensionality Reduction Techniques

1. Eigen value Decomposition
2. Principle Component Analysis
3. Linear Discriminant Analysis
4. Singular value Decomposition

1. Eigenvalue Decomposition (EVD)

Eigenvalue decomposition (EVD) is a fundamental matrix factorization technique used in linear algebra, particularly in machine learning, data science, and engineering applications. It expresses a square matrix as a product of its eigenvectors and eigenvalues.

1. Definition

For a square matrix **A** of size n×n, the eigenvalue decomposition is given by:

$$A=V\Lambda V^{-1}$$

where:

- A is an n×n square matrix.
- V is an n×n matrix whose columns are the **eigenvectors** of A.

- Λ (Lambda) is a diagonal $n \times n$ matrix containing the **eigenvalues** of A on its diagonal.
- V^{-1} is the inverse of V, assuming V is invertible.

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a dimensionality reduction technique that transforms high-dimensional data into a lower-dimensional space while preserving as much variance as possible.

---

1. Why Use PCA?

- **Reduce Dimensionality:** Helps simplify models and speed up computations.
- **Remove Redundancy:** Eliminates correlated features.
- **Improve Visualization:** Projects data into 2D or 3D for better pattern recognition.
- **Avoid Overfitting:** Reducing features prevents models from learning noise.

---

2. How PCA Works

PCA finds new **orthogonal** axes (principal components) that maximize variance.

Step-by-Step Process:

1. **Standardize the Data:**
   - Ensure all features have mean = 0 and variance = 1.
2. **Compute the Covariance Matrix:**
   - Measures relationships between variables.
3. **Compute Eigenvalues & Eigenvectors:**
   - Eigenvectors define principal components.
   - Eigenvalues measure their importance.
4. **Sort & Select Top k Components:**
   - Retain only the most significant components.
5. **Transform Data:**
   - Project data onto the new feature space.

Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is a supervised dimensionality reduction technique that maximizes class separability by finding a new feature space where different classes are well-separated.

---

1. Significance of LDA

- **Class Separability:** Unlike PCA (which preserves variance), LDA maximizes class discrimination.
- **Dimensionality Reduction:** Reduces the number of features while retaining essential class information.
- **Improves Classification Accuracy:** Helps models perform better by emphasizing important features.
- **Preprocessing for Classification Algorithms:** Works well with classifiers like SVM, logistic regression, etc.

---

2. How LDA Works

LDA projects data onto a new axis that maximizes the distance between different classes while minimizing variance within each class.

Step-by-Step Process:

1. **Compute the Mean of Each Class**
    - Compute the mean vector for each class.
2. **Compute the Scatter Matrices:**
    - **Within-Class Scatter Matrix (Sw):** Measures variance within each class.
    - **Between-Class Scatter Matrix (Sb):** Measures variance between classes.
3. **Compute Eigenvalues & Eigenvectors:**
    - Solve the **generalized eigenvalue problem** for Sw−1SbSw^{-1}SbSw−1Sb.
4. **Select Top Discriminant Components:**
    - The eigenvectors corresponding to the largest eigenvalues form the new feature space.
5. **Transform Data:**
    - Project data onto the selected LDA components.

---

3. LDA vs PCA

Feature	LDA	PCA
Type	Supervised	Unsupervised
Objective	Maximizes class separability	Maximizes variance
When to Use?	When class labels are available	When class labels are not available
Applications	Classification tasks	Feature extraction, compression

Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) is a **matrix factorization** technique used in linear algebra, machine learning, and data science. It decomposes a matrix into three other matrices, revealing important properties of the original data.

---

1. Definition

For an m ×n matrix **A**, the SVD is given by:

$$A = U\Sigma V^T$$

where:

- **U (Left Singular Vectors)**: An m×m orthogonal matrix.
- **Σ(Sigma) (Singular Values)**: A diagonal m×n matrix containing singular values.
- **V^T (Right Singular Vectors)**: The transpose of an n ×n orthogonal matrix.

Key Properties:

- The singular values in Σ(Sigma) are **non-negative** and arranged in **descending order**.
- The **rank** of A is equal to the number of **nonzero singular values** in Σ.
- If A is **square and symmetric**, SVD is equivalent to **Eigenvalue Decomposition**.

---

2. Applications of SVD

🔹 **Dimensionality Reduction**: Used in **Latent Semantic Analysis (LSA)** and **Principal Component Analysis (PCA)**.
🔹 **Image Compression**: Helps store images efficiently by keeping only dominant singular values.
🔹 **Noise Reduction**: Removes less important components while keeping the main structure.
🔹 **Recommendation Systems**: Netflix, Spotify use SVD for **collaborative filtering**.

Example:- $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$ Find out Eigenvalues and Eigen vectors of this matrix

→ by theorem $A\vec{u} = \lambda \vec{u}$

Where $\vec{u}$ = eigen vector λ = Eigen value

So, $|A - \lambda I| = 0$

$$\left| \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

$$\begin{vmatrix} -\lambda & 0 \\ -2 & -(3+\lambda) \end{vmatrix} = 0$$

$(3+\lambda)\lambda + 2 = 0$

$\lambda^2 + 3\lambda + 2 = 0$

$(\lambda+2)(\lambda+1) = 0$

So either $\lambda = -2$ or $\lambda = -1$

from $A\vec{u} = \lambda\vec{u}$ we get

$$\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} \vec{u}_1 \\ \vec{u}_2 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} u_2 \\ -2u_1 - 3u_2 \end{bmatrix} = \begin{bmatrix} \lambda u_1 \\ \lambda u_2 \end{bmatrix}$$

$u_2 = \lambda u_1$

$-2u_1 - 3u_2 = \lambda u_2$

or

$u_2 = \lambda u_1$

$(\lambda+3)u_2 = -2u_1$

$(\lambda+3-1)u_2 = (-2-\lambda)u_1$

$$\frac{u_1}{(\lambda+2)} = \frac{u_2}{(-\lambda-2)}$$

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} (\lambda+2) \\ (-\lambda-2) \end{bmatrix}$$

for λ_1 : $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{0}{\sqrt{0^2+16}} \\ \frac{-4}{\sqrt{0^2+16}} \end{bmatrix}$

for λ_2 $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{1+9}} \\ \frac{-3}{\sqrt{1+9}} \end{bmatrix}$

so eigenvectors

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \qquad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ -3/\sqrt{10} \end{bmatrix}$$

We can also use λ_1 & λ_2 individually to find out Eigen vector like

$$\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} u_2 \\ -2u_1 - 3u_2 \end{bmatrix} = \begin{bmatrix} -2u_1 \\ -2u_2 \end{bmatrix}$$

as well as

$$\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} u_2 \\ -2u_1 - 3u_2 \end{bmatrix} = \begin{bmatrix} -u_1 \\ -u_2 \end{bmatrix}$$

Principle Component Analysis.

[N.B: you can consider $x_1 = x$ & $x_2 = y$ also]

Example

F	x_1	x_2
1	4	11
2	8	4
3	13	5
4	7	14

Features are x_1 & x_2

Step 1: Calculate mean.

$$\bar{x}_1 = \mu_1 = \frac{1}{4}(4+8+13+7) = 8$$

$$\bar{x}_2 = \mu_2 = \frac{1}{4}(11+4+5+14) = 8.5$$

Step 2: Calculate Covariance matrix

$$S = \begin{bmatrix} Cov(x_1, x_1) & Cov(x_1, x_2) \\ Cov(x_2, x_1) & Cov(x_2, x_2) \end{bmatrix}$$

Formula of cov matrix $= \frac{1}{N-1} \sum_{k=1}^{n} (x_{ak} - \mu_a) \times (x_{bk} - \mu_b)$

where $a =$ ~~no. of data on~~ the types of data.

So $Cov(x_1, x_1) = \frac{1}{N-1} \sum_{k=1}^{N} (x_{1k} - \mu_1)(x_{1k} - \mu_1)$

$$= \frac{1}{3}\left[(4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2\right]$$

$$= 14$$

$$Cov(x_1, x_2) = \frac{1}{N-1} \sum_{k=1}^{N} (x_{1k} - \mu_1)(x_{2k} - \mu_2)$$

$$= \frac{1}{3}\left[(4-8)(11-8.5) + (8-8)(4-8.5) + (13-8)(5-8.5) + (7-8)\cdot(14-8.5)\right]$$

$$= -11$$

$$Cov(x_2, x_1) = Cov(x_1, x_2)$$
$$= -11$$

$$Cov(x_2, x_2) = \frac{1}{N-1} \sum_{k=1}^{N} (x_{2k} - \mu_2)(x_{2k} - \mu_2)$$

$$= \frac{1}{3}\left[(11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2\right]$$

$$= 23$$

So $S = \begin{bmatrix} Cov(x_1, x_1) & Cov(x_1, x_2) \\ Cov(x_2, x_1) & Cov(x_2, x_2) \end{bmatrix}$

$$= \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

<u>Step 3:</u> Calculate Eigen value of Covarience matrix. determinent of $(S - \lambda I) = 0$

where λ = scalar Eigen value.

So, $\det(S - \lambda I) = 0$

$$\begin{vmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{vmatrix} = 0$$

or, $(14-\lambda)(23-\lambda) - (-11)(-11) = 0$

or, $\lambda^2 - 37\lambda + 201 = 0$.

~~either~~

So $\lambda = \frac{1}{2}(37 \pm \sqrt{565})$

$\lambda_1 = 30.3849 \quad \lambda_2 = 6.6151$

Step 4: Computation of the Eigen vectors

Let's say Eigen vector $U = \begin{bmatrix} \vec{u}_1 \\ \vec{u}_2 \end{bmatrix}$ so to calculate Eigen vector

$$S\vec{u} = \lambda \vec{u}$$

$$\begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} 14u_1 - 11u_2 \\ -11u_1 + 23u_2 \end{bmatrix} = \begin{bmatrix} \lambda u_1 \\ \lambda u_2 \end{bmatrix}$$

so $14u_1 - 11u_2 = \lambda u_1$ — (1)

$-11u_1 + 23u_2 = \lambda u_2$ — (2)

~~by putting λ_1 & λ_2 equal value into eqn (1) & eqn (2) we get~~

~~$14u_1 = (14 - 30.38)$~~

$(14-\lambda)\, u_1 = 11u_2$ &

$\cancel{\lambda u_1}\ 11u_1 = (23-\lambda)\, u_2$

so, by these two eqn we can say that

$$\frac{u_1}{11} = \frac{u_2}{(14-\lambda)} \text{ — (1)} \quad \& \quad \frac{u_1}{23-\lambda} = \frac{u_2}{11} \text{ — (2)}$$

So from these 2 quation

from 1 we consider

$U_a = \begin{bmatrix} 11 \\ 14-\lambda \end{bmatrix}$ if we calculate by putting λ_1 & λ_2 we get

$U_{a_1} = \begin{bmatrix} 11 \\ -16.38 \end{bmatrix}$ & $U_{a_2} = \begin{bmatrix} 11 \\ 7.39 \end{bmatrix}$

eigen vectors for 1st scenario

$$e_{a_1} = \begin{bmatrix} \frac{11}{\sqrt{11^2+(-16.38)^2}} \\ \frac{-16.38}{\sqrt{11^2+(16.38)^2}} \end{bmatrix} = \begin{bmatrix} 0.55 \\ -0.83 \end{bmatrix}$$

$$e_{a_2} = \begin{bmatrix} \frac{11}{\sqrt{11^2+7.39^2}} \\ \frac{7.39}{\sqrt{11^2+7.39^2}} \end{bmatrix} = \begin{bmatrix} 0.83 \\ 0.55 \end{bmatrix}$$

Find out the eigen vectors for 2nd scenario using 2 equation which is

$$\frac{u_1}{23-\lambda} = \frac{u_2}{11}$$

$$U_b = \begin{bmatrix} 23-\lambda \\ 11 \end{bmatrix}$$

for λ_1, $U_{b_1} = \begin{bmatrix} 23-30.38 \\ 11 \end{bmatrix} = \begin{bmatrix} -7.38 \\ 11 \end{bmatrix}$

λ_2 $U_{b_2} = \begin{bmatrix} 16.39 \\ 11 \end{bmatrix}$

eigen vectors for 2^{nd} scenario.

$$e_{b_1} = \begin{bmatrix} \dfrac{-7.38}{\sqrt{(-7.38)^2+(11)^2}} \\ \\ \dfrac{11}{\sqrt{(-7.38)^2+11^2}} \end{bmatrix} = \begin{bmatrix} -0.55 \\ \\ 0.83 \end{bmatrix}$$

$$e_{b_2} = \begin{bmatrix} 0.83 \\ \\ 0.55 \end{bmatrix}$$

so the <u>step 5</u>: Computation of first principle component using

$$e^T \begin{bmatrix} u_{ak} - u_1 \\ u_{bk} - u_2 \end{bmatrix}$$

for first case scenario.

First principle component

$$= \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} 4-8 \\ \\ 11-8.5 \end{bmatrix}$$

$$= -4.30535$$

Like this for every component will be

3.73, 5.69 & -5.12

Example 2 (PCA)

X	Y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2

$\bar{x} = \mu_1 = \frac{1}{4}(2.5 + 0.5 + 2.2 + 1.9) = 1.77$

$\bar{y} = \mu_2 = \frac{1}{4}(2.4 + 0.7 + 2.9 + 2.2)$
$= 2.05$

Covarience matrix

$$S = \begin{bmatrix} Cov(X, X) & Cov(X, Y) \\ Cov(Y, X) & Cov(Y, Y) \end{bmatrix}$$

$$Cov(X, X) = \frac{1}{N-1} \sum_{k=1}^{n} (x - \mu_1)(x - \mu_1)$$

$$= \frac{1}{3}\left[(2.5 - 1.77)^2 + (0.5 - 1.77)^2 + (2.2 - 1.77)^2 + (1.9 - 1.77)^2\right]$$

$$= \frac{1}{3}\left[(0.73)^2 + (-1.27)^2 + (0.43)^2 + (0.13)^2\right]$$

$$= 0.78$$

$$Cov(X, Y) = Cov(Y, X)$$

$$= \frac{1}{N-1} \sum_{k=1}^{n} (x - \mu_1)(y - \mu_2)$$

$$= \frac{1}{3}\left[(2.5 - 1.77)(2.4 - 2.05) + (0.5 - 1.77)(0.7 - 2.05) + (2.2 - 1.77)(2.9 - 2.05) + (1.9 - 1.77)(2.2 - 2.05)\right]$$

$$= \frac{1}{3}\left[0.25 + 1.71 + 0.36 + 0.15\right] = 0.82$$

$$Cor(Y, Y) = \frac{1}{N-1} \sum_{k=1}^{N} (y - \mu_2)(y - \mu_2)$$

$$= \frac{1}{3}\left[(2.4 - 2.05)^2 + (0.7 - 2.05)^2 + (2.9 - 2.05)^2 + (2.2 - 2.05)^2\right]$$

$$= \frac{1}{3}\left[0.12 + 1.82 + 0.72 + 0.02\right]$$

$$= 0.89$$

so, $S = \begin{bmatrix} 0.78 & 0.82 \\ 0.82 & 0.89 \end{bmatrix}$

so, by $\det|S - \lambda I| = 0$ we get

$$(0.78 - \lambda)(0.89 - \lambda) - 0.672 = 0$$

$$\lambda^2 - 1.67\lambda - 1.36 = 0$$

$$\lambda = \frac{+1.67 \pm \sqrt{8.2289}}{2}$$

$$= \cancel{\frac{1.67 \pm 1.43}{2}}$$

$$\lambda = 0.83 \pm 1.43$$

so $\lambda_1 = 2.26$, $\lambda_2 = 0.6$

so the eigen vector ~~value~~ willbe feature matrix

$$S\vec{u} = \lambda \vec{u}$$

$$\begin{bmatrix} 0.78 & 0.82 \\ 0.82 & 0.89 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$0.78 u_1 + 0.82 u_2 = \lambda u_1 \quad \& \text{ — (I)}$$

$$0.82 u_1 + 0.89 u_2 = \lambda_2 u_2 \text{ — (II)}$$

~~Q.8~~ from substracting ① & ②.

$$0.04u_1 + 0.07u_2 = \lambda u_1 - \lambda u_2$$

$$(0.04 - \lambda)u_1 = (-0.07 - \lambda)u_2$$

or, $(\lambda - 0.04)u_1 = (\lambda + 0.07)u_2$

$$\frac{u_1}{(\lambda + 0.07)} = \frac{u_2}{(\lambda - 0.04)}$$

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} (\lambda + 0.07) \\ (\lambda - 0.04) \end{bmatrix}$$ so for λ_1 $\begin{bmatrix} 2.33 \\ 2.22 \end{bmatrix}$ for λ_2 $\begin{bmatrix} 0.67 \\ 0.56 \end{bmatrix}$

eigen vectors

$$\begin{bmatrix} \frac{2.33}{\sqrt{2.33^2 + 2.22^2}} \\ \frac{2.22}{\sqrt{2.33^2 + 2.22^2}} \end{bmatrix} = \begin{bmatrix} 0.72 \\ 0.68 \end{bmatrix}$$ these are feature vectors

$$\begin{bmatrix} \frac{0.67}{\sqrt{0.67^2 + 0.56^2}} \\ \frac{0.56}{\sqrt{0.67^2 + 0.56^2}} \end{bmatrix} = \begin{bmatrix} 0.77 \\ 0.64 \end{bmatrix}$$

we can find out also the principle components also.

Linear Discriminent Analysis

Example:- class1 = {(4,2) (2,4), (2,3)}
class 2 = {(9,10), (6,8), (9,5)}

Step: 1 Segregate classes and Find out mean of classes after implementing 2D data in 1D data.

$$\mu_1 = \frac{1}{N_1} \sum_{class \in x_1} class1 = \frac{1}{3}\left[\begin{pmatrix}4\\2\end{pmatrix} + \begin{pmatrix}2\\4\end{pmatrix} + \begin{pmatrix}2\\3\end{pmatrix}\right]$$

$$= \begin{bmatrix} 8/3 \\ 3 \end{bmatrix} = \begin{bmatrix} 2.6 \\ 3 \end{bmatrix}$$

$$\mu_2 = \frac{1}{N_2} \sum_{class \in x_2} class2 = \frac{1}{3}\left[\begin{pmatrix}9\\10\end{pmatrix} + \begin{pmatrix}6\\8\end{pmatrix} + \begin{pmatrix}9\\5\end{pmatrix}\right]$$

$$= \begin{bmatrix} 8 \\ \frac{23}{3} \end{bmatrix} = \begin{bmatrix} 8 \\ 7.6 \end{bmatrix}$$

Step2: Find out the covarmeen matrix of each classes which introduces scatter plot of each classes:

$$S_1 = \frac{1}{N-1} \sum_{x=1}^{n} (x-\mu_1)(x-\mu_1)^T$$

$$= \frac{1}{2}\left[\begin{pmatrix}4\\2\end{pmatrix} - \begin{pmatrix}2.6\\3\end{pmatrix}\right]\left[\begin{pmatrix}4\\2\end{pmatrix} - \begin{pmatrix}2.6\\3\end{pmatrix}\right]^T$$

$$+ \left[\begin{pmatrix}2\\4\end{pmatrix} - \begin{pmatrix}2.6\\3\end{pmatrix}\right]\left[\begin{pmatrix}2\\4\end{pmatrix} - \begin{pmatrix}2.6\\3\end{pmatrix}\right]^T$$

$$+ \left[\begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2.6 \\ 3 \end{pmatrix}\right]\left[\begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2.6 \\ 3 \end{pmatrix}\right]^T$$

$$= \frac{1}{2}\left[\left[\begin{pmatrix} 1.4 \\ -1 \end{pmatrix}\begin{pmatrix} 1.4 & -1 \end{pmatrix}\right] + \left[\begin{pmatrix} -0.6 \\ 1 \end{pmatrix}\begin{pmatrix} -0.6 & 1 \end{pmatrix}\right] + \left[\begin{pmatrix} -0.6 \\ 0 \end{pmatrix}\begin{pmatrix} -0.6 & 0 \end{pmatrix}\right]\right]$$

$$= \frac{1}{2}\left[\begin{bmatrix} 1.96 & -1.4 \\ -1.4 & 1 \end{bmatrix} + \begin{bmatrix} 0.36 & -0.6 \\ -0.6 & 1 \end{bmatrix} + \begin{bmatrix} 0.6 & 0 \\ 0 & 0 \end{bmatrix}\right]$$

$$= \begin{bmatrix} \frac{2.92}{2} & \frac{-2}{2} \\ \frac{-2}{2} & \frac{2}{2} \end{bmatrix} = \begin{bmatrix} 1.46 & -1 \\ 1 & 1 \end{bmatrix}$$

$$S_2 = \frac{1}{N-1}\left[\left[\begin{pmatrix} 9 \\ 10 \end{pmatrix} - \begin{pmatrix} 8 \\ 7.6 \end{pmatrix}\right]\left[\begin{pmatrix} 9 \\ 10 \end{pmatrix} - \begin{pmatrix} 8 \\ 7.6 \end{pmatrix}\right]^T + \left[\begin{pmatrix} 6 \\ 8 \end{pmatrix} - \begin{pmatrix} 8 \\ 7.6 \end{pmatrix}\right]\left[\begin{pmatrix} 6 \\ 8 \end{pmatrix} - \begin{pmatrix} 8 \\ 7.6 \end{pmatrix}\right]^T + \left[\begin{pmatrix} 9 \\ 5 \end{pmatrix} - \begin{pmatrix} 8 \\ 7.6 \end{pmatrix}\right]\left[\begin{pmatrix} 9 \\ 5 \end{pmatrix} - \begin{pmatrix} 8 \\ 7.6 \end{pmatrix}\right]^T\right]$$

$$S_2 = \frac{1}{2}\left[\left[\begin{pmatrix} 1 \\ 2.4 \end{pmatrix}\begin{pmatrix} 1 & 2.4 \end{pmatrix}\right] + \left[\begin{pmatrix} -2 \\ 0.4 \end{pmatrix}\begin{pmatrix} -2 & 0.4 \end{pmatrix}\right] + \left[\begin{pmatrix} 1 \\ -2.6 \end{pmatrix}\begin{pmatrix} 1 & -2.6 \end{pmatrix}\right]\right]$$

$$= \frac{1}{2}\left[\begin{bmatrix} 1 & 2.4 \\ 2.4 & 5.76 \end{bmatrix} + \begin{bmatrix} 4 & -0.8 \\ -0.8 & 0.16 \end{bmatrix} + \begin{bmatrix} 1 & -2.6 \\ -2.6 & 6.76 \end{bmatrix}\right]$$

$$= \frac{1}{2}\begin{bmatrix} 6 & -1 \\ -1 & 12.68 \end{bmatrix} = \begin{bmatrix} 3 & -0.5 \\ -0.5 & 6.34 \end{bmatrix}$$

Step 3: Find out the total scatter value of data $S = S_1 + S_2$ or the within class scatter matrix

$$S_w = \begin{bmatrix} 1.46 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 3 & -0.5 \\ -0.5 & 6.34 \end{bmatrix}$$

$$= \begin{bmatrix} 3.46 & -1.5 \\ -1.5 & 7.34 \end{bmatrix}$$

Step 4: Find out b/w scatter value of data or middle class scatter matrix

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$$

$$= \left[\begin{pmatrix} 2.6 \\ 3 \end{pmatrix} - \begin{pmatrix} 8 \\ 7.6 \end{pmatrix}\right]\left[\begin{pmatrix} 2.6 \\ 3 \end{pmatrix} - \begin{pmatrix} 8 \\ 7.6 \end{pmatrix}\right]^T$$

$$= \begin{bmatrix} -5.4 \\ -4.6 \end{bmatrix}\begin{bmatrix} -5.4 & -4.6 \end{bmatrix} = \begin{bmatrix} 29.16 & 24.84 \\ 24.84 & 21.16 \end{bmatrix}$$

Step 5 Find out the $S_w^{-1} S_B \vec{u} = \lambda \vec{u}$

so $S_w^{-1} S_B = \begin{bmatrix} 3.46 & -1.5 \\ -1.5 & 7.34 \end{bmatrix}^{-1} \begin{bmatrix} 29.16 & 24.84 \\ 24.84 & 21.16 \end{bmatrix}$

Find this ~~in~~

then

Step 6: Calculate the eigen value & ~~mat~~ eigen vector of $S_w^{-1} S_B$ using

$|S_w^{-1} S_B - \lambda I| = 0$ & $S_w^{-1} S_B \vec{u} = \lambda \vec{u}$ like PCA.

~~Step 7~~:- or Find out this eigen vector using

$\omega^* = S_w^{-1} (\mu_1 - \mu_2)$ this eq^n.

Step 7: After finding the ~~imp~~ the eigen vectors Mutiply each ~~data~~ ~~with~~ data of each class es with the eigen vectors you will find out the LDA value of this Data.