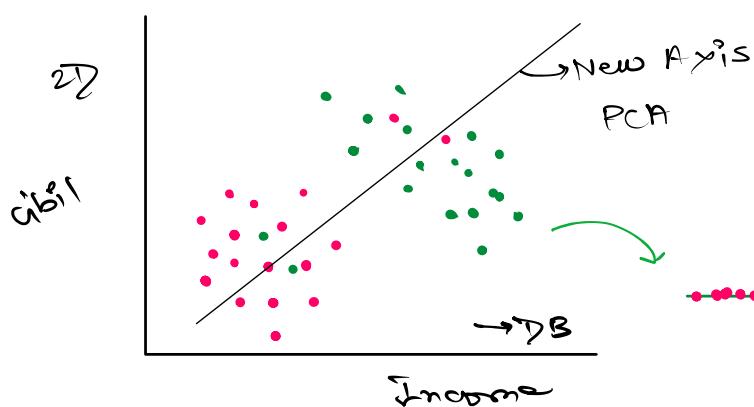


Also used for Dimensionality Reduction like PCA-

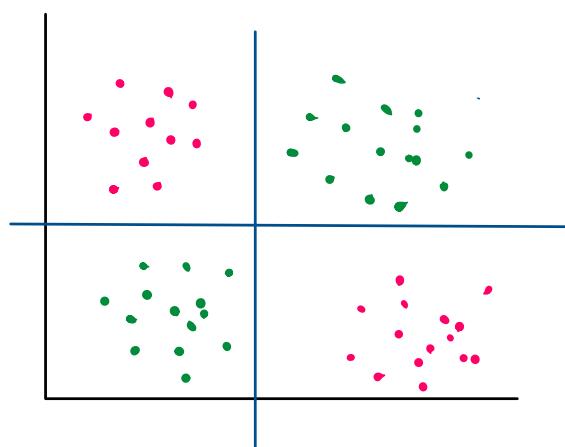
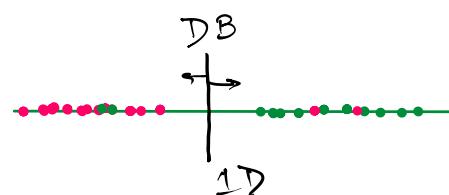
only used for Classification Data.

Income	Cibil score	Defaulted or Not?
—	—	Def
—	—	Def
—	—	ND
—	—	ND
—	—	ND

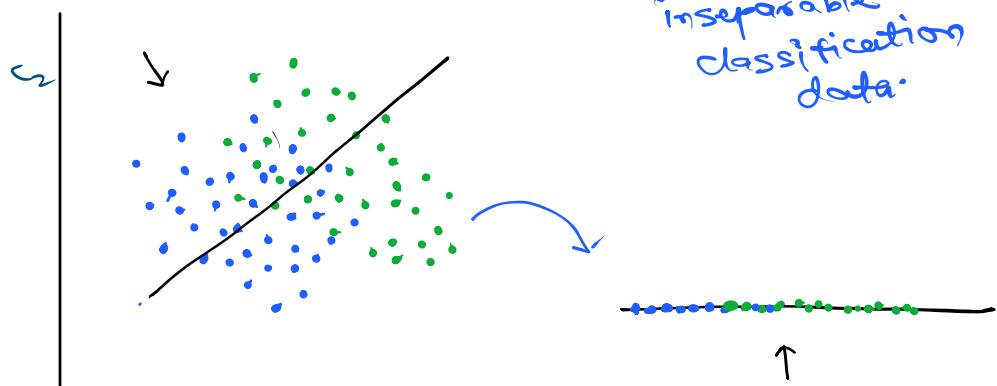


DB \rightarrow 1D Alg^o

Axis \rightarrow PCA or LDA



inseparable
classification
data.



$$\frac{C_1}{C_2}$$

To solve the problem of inseparable classification data, we use LDA.

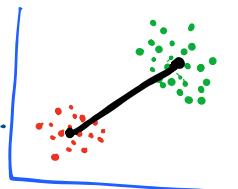
LDA will project the data onto a new axis to make the data easily separable. And while doing so it will automatically reduce the dimension of the data.

The new axis that LDA will create should follow these 2 criterias (both considered simultaneously) →

- ① Maximize the distance b/w the means of two classes / categories.
- ② Minimize the variation (spread) within each class.

μ_R = Mean of Red points No

S_R^2 = variation of Red Datapoints.



μ_G = Mean of yes green points.

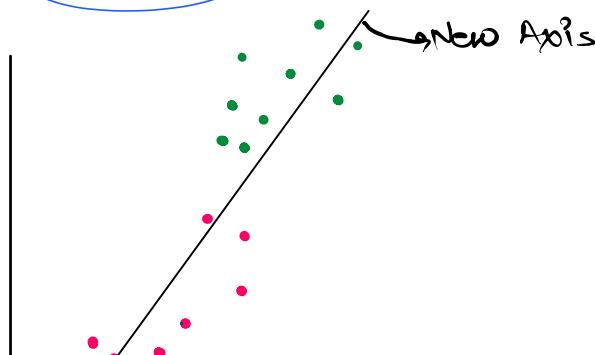
S_G^2 = Variation of green Datapoints.

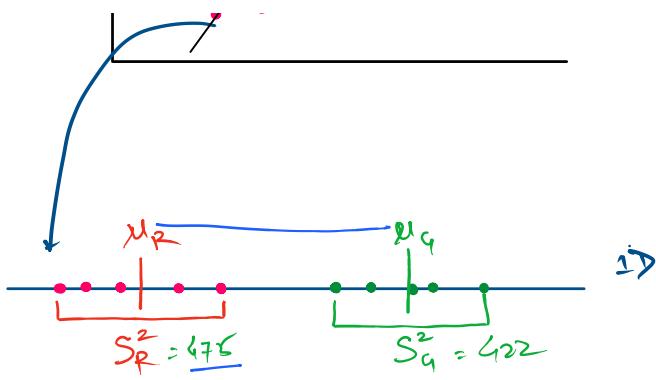
$$\frac{d_1^2 + d_2^2 + \dots}{n}$$

LDA

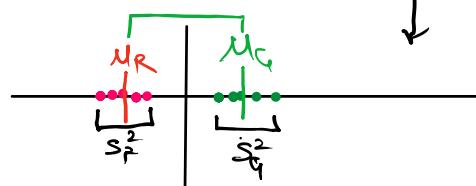
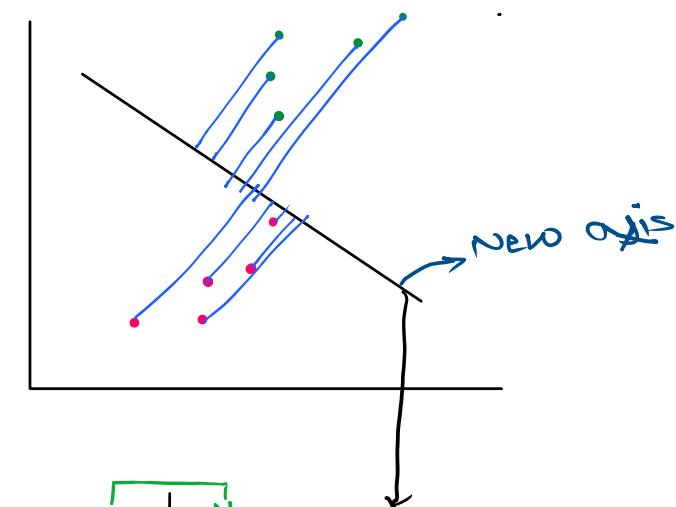
$$\frac{(\mu_R - \mu_G)^2}{S_R^2 + S_G^2}$$

Maximize
Minimize → Maximize

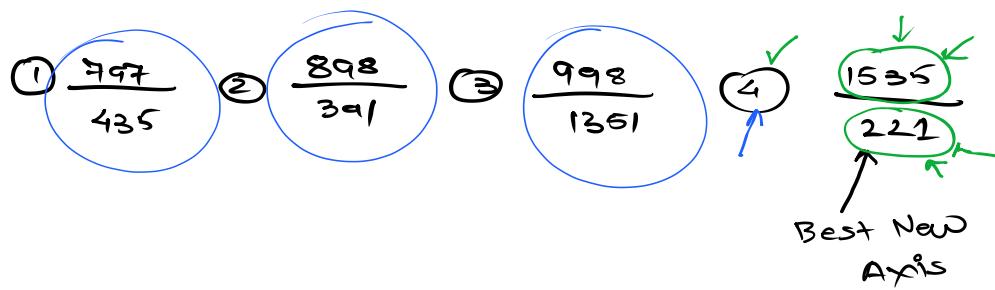




$$\frac{(\mu_R - \mu_Q)^2}{S_R^2 + S_Q^2} = \frac{980}{475 + 422}$$



$$\frac{(\mu_R - \mu_Q)^2}{S_R^2 + S_Q^2} = \frac{270}{98 + 122} \quad \text{vs} \quad \frac{980}{475 + 422}$$



Muse

