Dimension Reduction

* most of the times we cant visualize high-dim data. Visualizing any kind of data that has a large number of dimensions requires feature extraction. There are methods to represent high-dim data by assuming linear relationship, and others which which assumes more complex relations.

Feature Extraction

Linear

- PCA
- LDA

Non-linear

- t-SNE

*PCA:

- We use eign vectors / values to find the features with most variance.
- Using these values we obtain transformation matrix, W, $x' = x \cdot W$.

- The main reason behind this trans. matrix is to reduce the dimensions, d, of the original matrix, X, to k which is the number we chose.

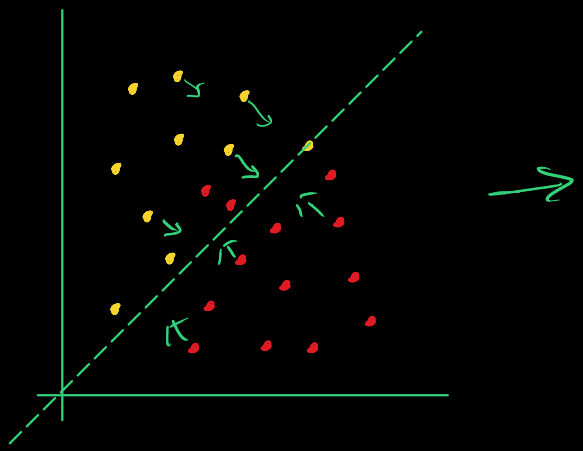
- It is Unsupervised.

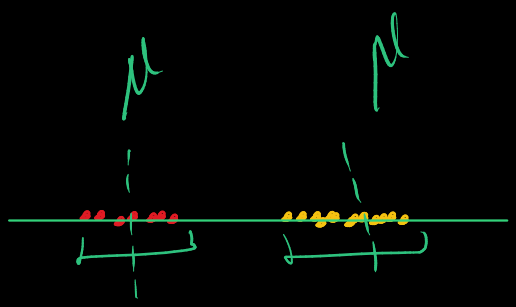
* LDA

- It's main idea is to find an axis that minimizes the variance in each dataset and maximizes the mean difference between the two.

- This axis will be used to project the data on to.

- It is Supervised.





$$\frac{(\mu - \mu)^2}{s^2 + s^2}$$

* T-SNE

- One of the most useful manifold techniques to visualize more complex related data.

- It computes the relation between data points on the original dimension. Then, tries to mimic it at lower dimensions.