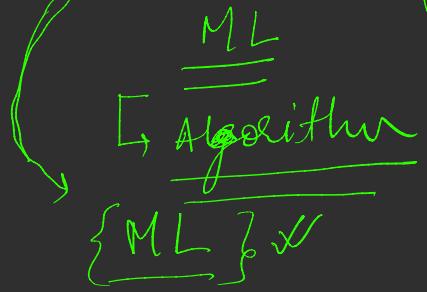
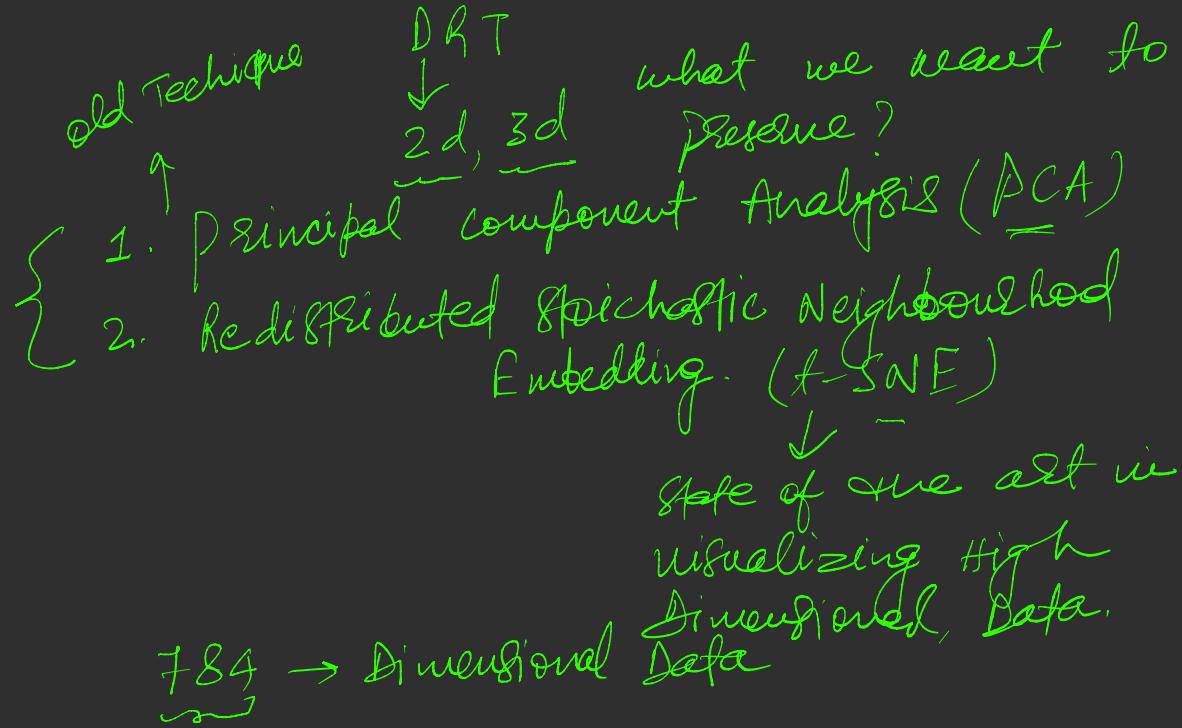


Dimensionality Reduction Techniques



what is Dimensionality Reduction?
why should we learn it?

	Name	RollNo.	code	specifi.	Netwo	Marks	In
Row1 →	Student B	✓	✓	✓	✓	✓	✓
Row2 →	Student 2						
Row3 →	:						
⋮							
RowN →	Student N						



Row vector & column vector

Iris flower dataset

for any given flower \rightarrow 4 features
 $\begin{matrix} \text{Sepal length, Sepal width, Petal length} \\ \text{& petal width} \end{matrix}$ (4 variables)

$d_i \rightarrow$	(SL)	(SW)	(PL)	PW	
-	2.3	5.6	1.2	5.1	

classify $\left\{ \begin{matrix} \text{Setosa, Versicolor,} \\ \overline{\text{Iris}} \end{matrix} \right. \left. \begin{matrix} \text{virginica} \\ \overline{\text{Iris}} \end{matrix} \right\} \rightarrow y_i$

$$\text{flowed} = \underbrace{[SL \ PL \ SW \ PW]}_{\text{real values}}$$

i -th point : $x_i \in \mathbb{R}^d \rightarrow d$ -dimensional
 \downarrow column vector

$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \\ \vdots \\ x_{id} \end{bmatrix} : \begin{array}{l} \text{real values} \\ \text{real space} \\ \text{column vector} \end{array}$$

Default vector notation : column vector

$$\left\{ f_1 = \begin{bmatrix} 2 & 1 & \cdot \\ 3 & 2 & \cdot \\ 1 & 9 & \cdot \\ 4 & 2 & \cdot \end{bmatrix} \right\}$$

Row vector
 $x_i = [\begin{matrix} \downarrow \\ 2 \cdot 1, \end{matrix} \quad \begin{matrix} \downarrow \\ 3 \cdot 1, \end{matrix} \quad \begin{matrix} \downarrow \\ 4 \cdot 1, \end{matrix} \quad \begin{matrix} \downarrow \\ 1 \cdot 9 \end{matrix}]_{1 \times 4}$ single row
1-row 84 column

[]_{84 columns} $x_i \in \mathbb{R}^d$

1815

Dataset → data point
 $D = \{x_i, y_i\}_{i=1}^n \Rightarrow \# \text{NO. of data points}$

da 8 fa 1 n -

$$x_i \in \mathbb{R}^{d \rightarrow} \Rightarrow x_i \in \mathbb{R}^q$$

Yi setosa, vesicole, virgin

$x_i \rightarrow$ Data point

$$\begin{bmatrix} S_L \\ S_W \\ P_L \\ P_W \end{bmatrix}$$

$y_i \rightarrow$ Class-label

{
Setosa, Versicolor, Virginica}

$$X = \begin{bmatrix} f_1 & f_2 & f_3 & \dots & f_i & \dots & f_n \\ \vdots & & & & \uparrow & & \\ f_j & & & & x_i & & \\ \vdots & & & & \downarrow & & \\ f_d & & & & & & \end{bmatrix}_{d \times n}$$

This model

$f_1 = PL$
 $f_2 = PW$
 $f_3 = SL$
 $f_4 = SW$

Column : Data - points
 Row : features / variable

~~This flower
Dataset~~

x_1

x_2

.

.

x_i

.

.

.

.

x_n

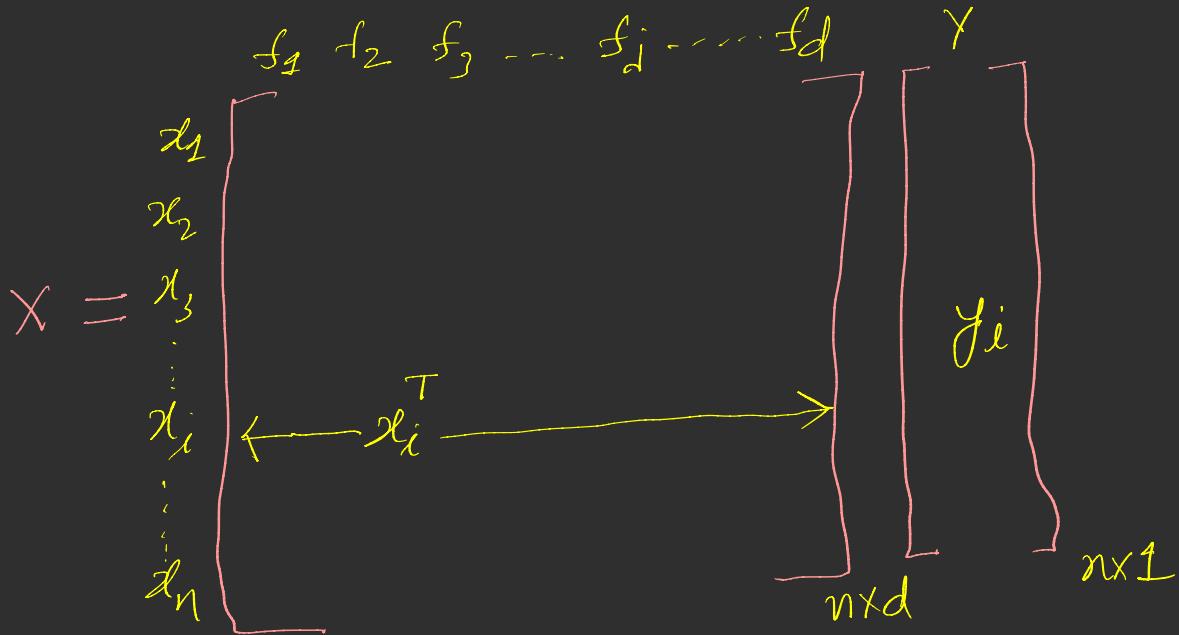
f_1
 s_L

f_2
 s_W

f_3
 p_L

f_4
 p_W

Data Pre-processing : column normalization



obtain Data \longrightarrow Pre-processing
column normalization \longrightarrow Data Modelling
(Dim- Reduction)

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} \quad \begin{bmatrix} f_1 & f_2 & f_3 & \cdots & f_i & \cdots & f_d \end{bmatrix}^T \quad \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{bmatrix} \quad \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix} \quad Y \quad j_i \quad n \times d \quad n \times 1$$

TSI8 flow
SL, SW, PL
PW
150 flow

$a_1, a_2, a_3, \dots, a_i, \dots, a_n \rightarrow n\text{-values of } f_j$

defn
Max $a_i = a_{\max} \geq a_i \quad (i: 1-n)$

Min $a_i = a_{\min} \leq a_i \quad (i: 1-n)$

$a_3 = 8000$
 $\rightarrow a'_1, a'_2, a'_3, \dots, a'_i, \dots, a'_n$ [New Data vector]

$$a'_i = \frac{a_i - a_{\min}}{a_{\max} - a_{\min}} \quad a'_i \in [0, 1]$$

$$a'_{\min} = \frac{a_{\min} - a_{\min}}{a_{\max} - a_{\min}} = 0 ; \quad a'_{\max} = \frac{a_{\max} - a_{\min}}{a_{\max} - a_{\min}}$$

$$\underline{a'_{\min} = 0}$$

$$\underline{a'_{\max} = 1}$$

$a_1, a_2, a_3 \dots a_i \dots a_n ; a_i \in \mathbb{R}$

↓
column - Normalization

$a'_1, a'_2, a'_3 \dots a'_i \dots a'_n$ s.t $a'_i \in [0,1]$

why?

h = height
 w = weight
student

$$f_1 = h \quad f_2 = w$$

1	162	56
2	172	72
3	182	84
4	155	95
5	160	76
.	.	.
.	.	.
.	.	.
.	.	.
n	.	.

Any real
number values

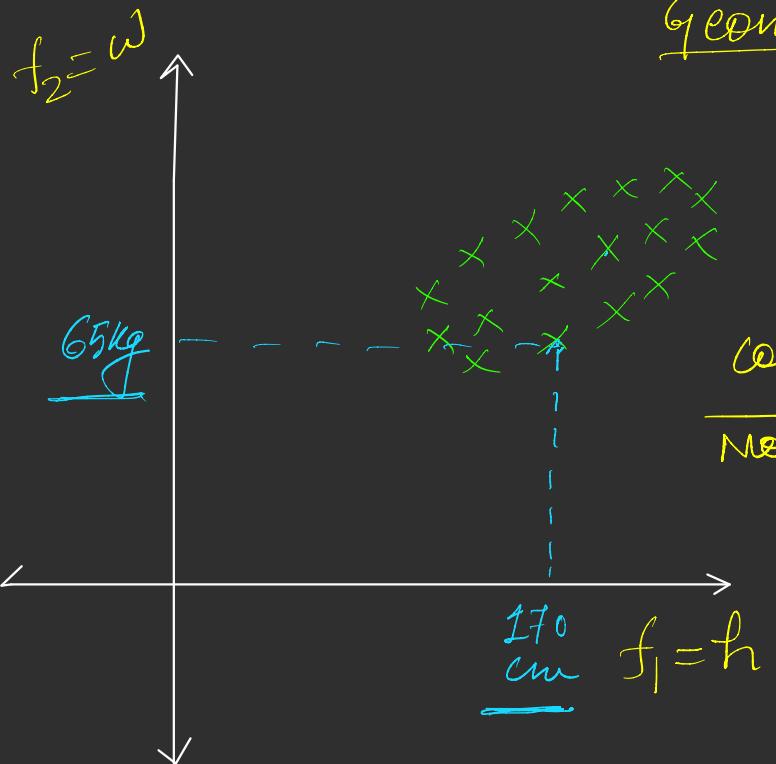
height = cm
weight = kilograms

1	$f'_1 = h'$	$f'_2 = w'$
2	.	.
3	.	.
.	.	.
.	.	.
.	.	.
n	.	.

column
Normalization

{ getting rid of
scale
Putting everything
in one scale

$[0,1]$	$[0,1]$
---------	---------



Geometric Intuition

column
Normalization

