Lecture 2

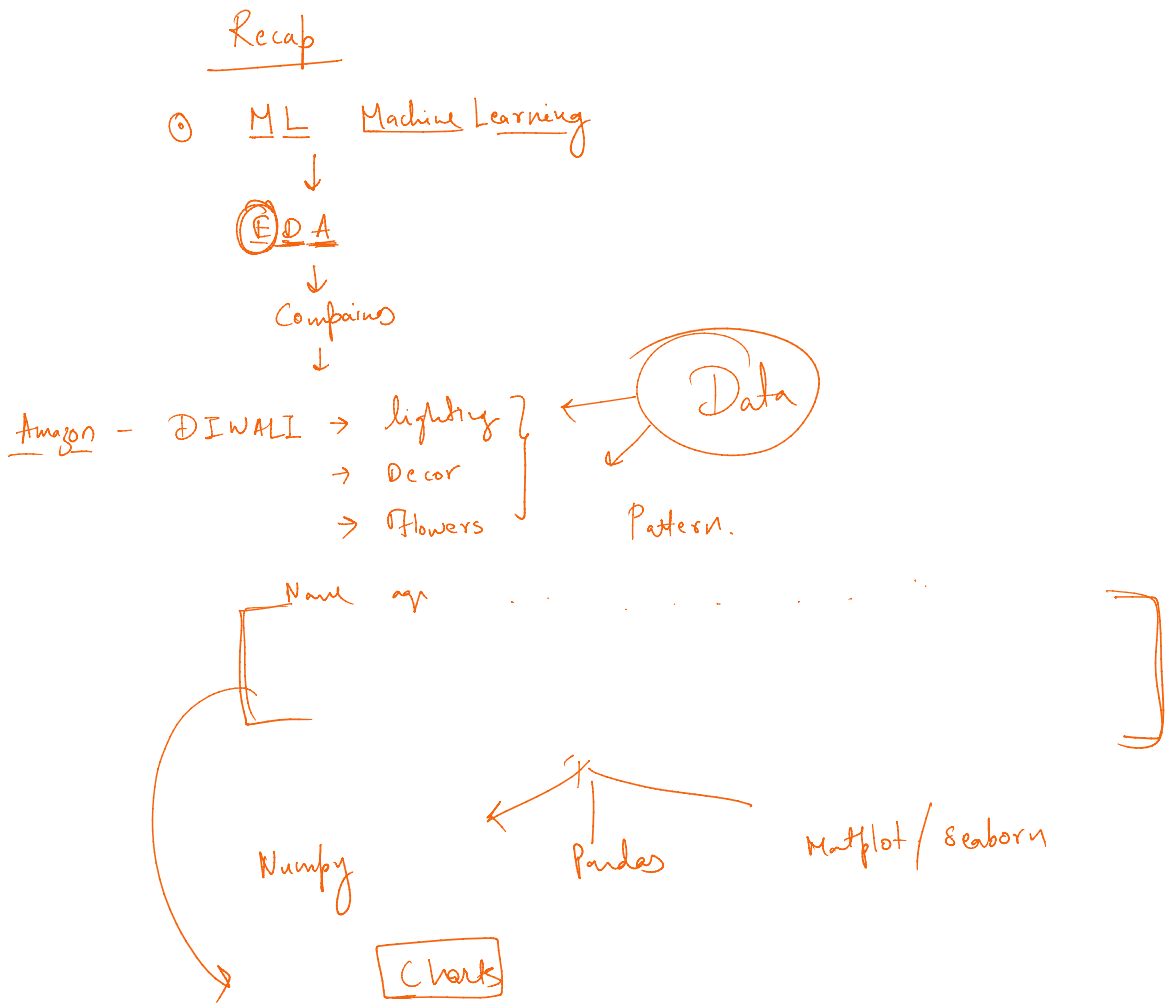
Friday, October 24, 2025 8:57 PM

Agenda

1. Np.arange
2. Fancy Indexing
3. 2d matrix
   a. Reshape
   b. Indexing and slicing
   c. Masking
4. Aggregate functions
   a. Axis
5. Np.all
6. Np.any
7. Np.where

1. Sort
2. Sort 2d array
3. Matrix multiplication
   a. Np.dot
   b. Np.matmul
   c. @
4. Vectorization
5. Broadcastiing
6. Np.tile
7. Split
   a. Hsplit
   b. Vsplit
8. Stacking

Recap



Zomato Dataset

41 Min 42 Min

$\frac{41 + 42}{2}$

cleaning data.

Numpy
Pandas

Tools to perform EDA.

1. Np.arange : id
2. ~~Fancy Indexing~~
3. ~~2d matrix~~
   a. Reshape
   b. Indexing and slicing
   c. Masking
4. Aggregate functions

1
2
3
4
5
6

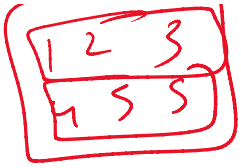
a. Reshape
b. Indexing and slicing
c. Masking
4. Aggregate functions
   a. Axis
5. Np.all
6. Np.any
7. Np.where

1. Sort
2. Sort 2d array
3. Matrix multiplication
   a. Np.dot
   b. Np.matmul
   c. @
4. Vectorization
5. Broadcastiing
6. Np.tile
7. Split
   a. Hsplit
   b. Vsplit
8. Stacking

Interview

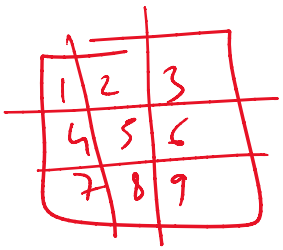
5

5
6
7
8
9



5, 2 → 2, 5

T – R – C
C – R



reshape (R, C)

10, 1
1, 10
1, 1, 10

import numpy as np

① np.arange (start, end, step)

np.arange (1, 1000) →

1
2
.
.
.
999

end

np.arange (1, 1000, 2)

1
3
5
7
9
11
13

10, 20, 30 90

np.arange (10, 100, 10)

10, 20, 30 90

(10, 91, 10)

10, 20 90

np.arange(1, 3, 0.2)

1, 1.2, 1.4 . . . 2.8

~~10 MIA~~ 10 M for 1 — 10 M.

Fancy Indexing

Zomato

array ⇒ Votes → [700, 900, 400, 200, 300]

Votes + 2

Votes ≥ 500 cond

True True false false false

≥ 4.0

Votes = [700, 600, 1200, 300, 400]

Votes [[1, -1]]

600 200

fancy indexing

votes [votes > 500] → T F T F

It will only give you True element

band

2000 700

Happy

Heartbeat

Happy → Walking More

Sad

2000 3000 1500 ↑

1500 1200 700 ↓

200 400 600

x Miss x Miss ✓

Column

2D

R, C

Votes Cost

700 "1700"

800 "1900"

730 "2000"

350 "500"

150 "1700"

700 "1200"

Rows

Shape 6, 2

↓

Reshape (2, 6)

700 1700 800 1900 730 2000

350 500 150 1700 700 1200

1700 1700 | 1800 1900 | 1930 2000

350 500 150 1700 700 1200

Initial — Starting

np.array

arr = np.arange(10, 100, 10)

10, 20, 30, 40, 50, 60, 70, 80, 90

Neural Network

arr.reshape(3, 3)

{10, 20, 30
40, 50, 60
70, 80, 90}

92 →

Exp	Sal
1	10k
2	20k
3	30k
4	40k
5	50k

→ 10k × Exp = Sal

2.3k

5k

10k

10 100

10 20 30
40 50 60
70 80 90 100

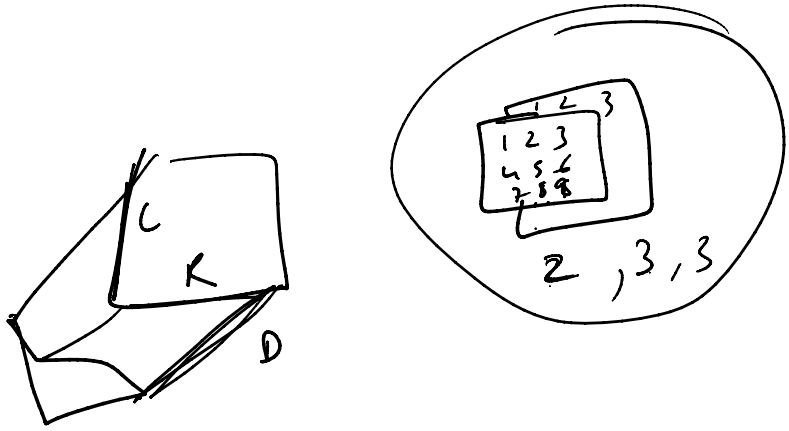
2×4

reshape 90

array (90, 15)

[ ]

astype.



R C 2, 3

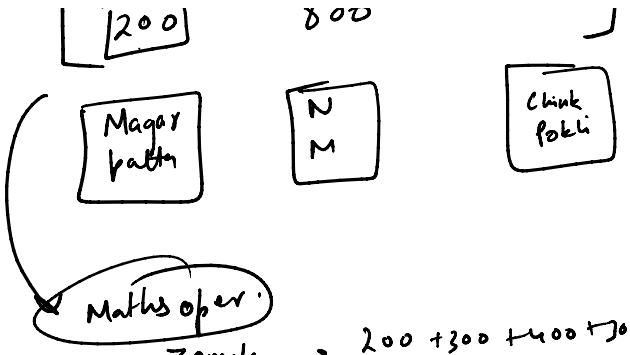
10, 20
30, 40 →
50, 60

3, 2

Before

	area1	area2	area3
a=	200	300	400
	300	700	300
	200	800	900

Votes

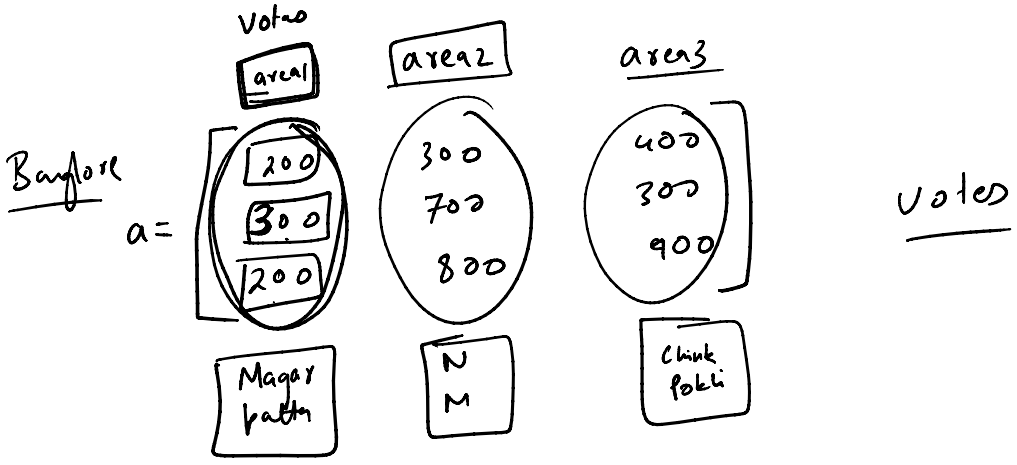


a. mean()
a. sum()

8000 + Student

a. max() — 900

a. min() — 200



area1 > area2 > area3

200 + 300 + 200 / 3

300 + 700 + 800 / 3 —

R L

Root spot — 3

3

R, C

axis — 0 — col
1 — Row

np. All → Veg nonVeg

np.all ([T, F, F, T, F]) all?

False

If all the values in array is True or Not

np a = 4.9 4.8 4.7 4.3 4.7 4.95
T T T T T T

np.all (a > 4.2) = True.

False

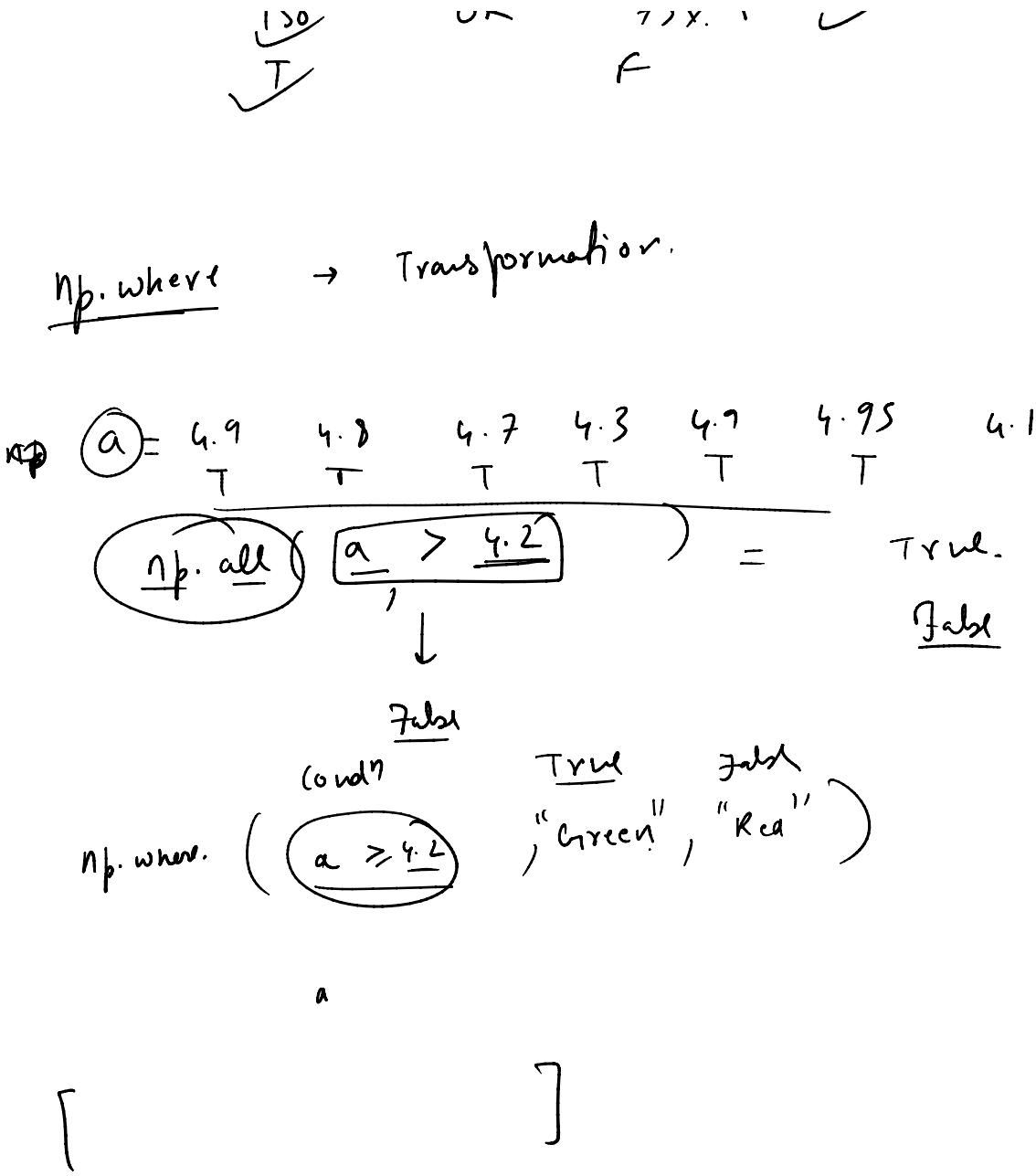
python and Notes

np.any (T F T T F F F)

If any one value is T

True.

ISO OK 75x.
T F



Sorting

np.sort (array)

0	1	2	3	4	5	6	7	8	9
array([775,	787,	918,	88,	166,	286,	2556,	324,	504,	402])
[3,	4,	5,	7,	9,	8,	0,	1,	2,	6]

np.sort (~~[illegible]~~,)

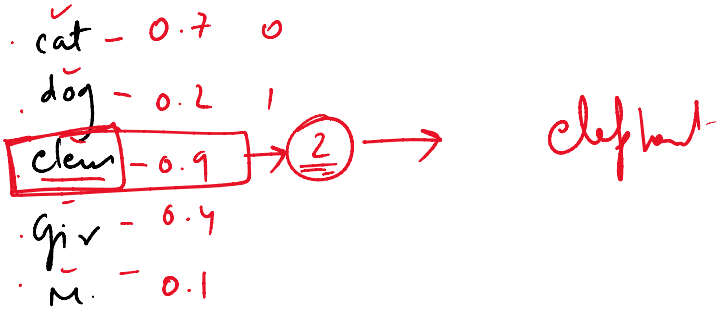
↓

88, 166, 286

np.argsort

argument

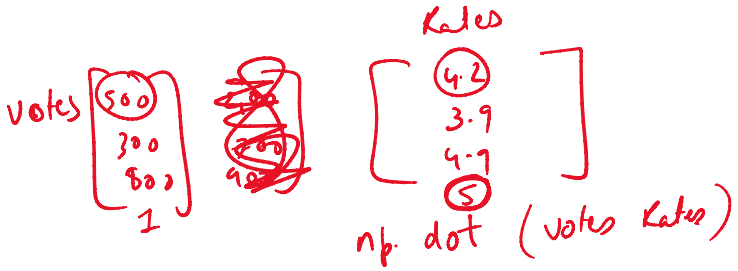
index



np.matmul(a,b)

np.dot (a,b)

a@b



Vectorization

Broadcasting

hsplit

vsplit

conc

— Next

conc

Pandas
Adwan

PSP } Extra Credit
{ fitbit analysis }

Doubt.

9545126535