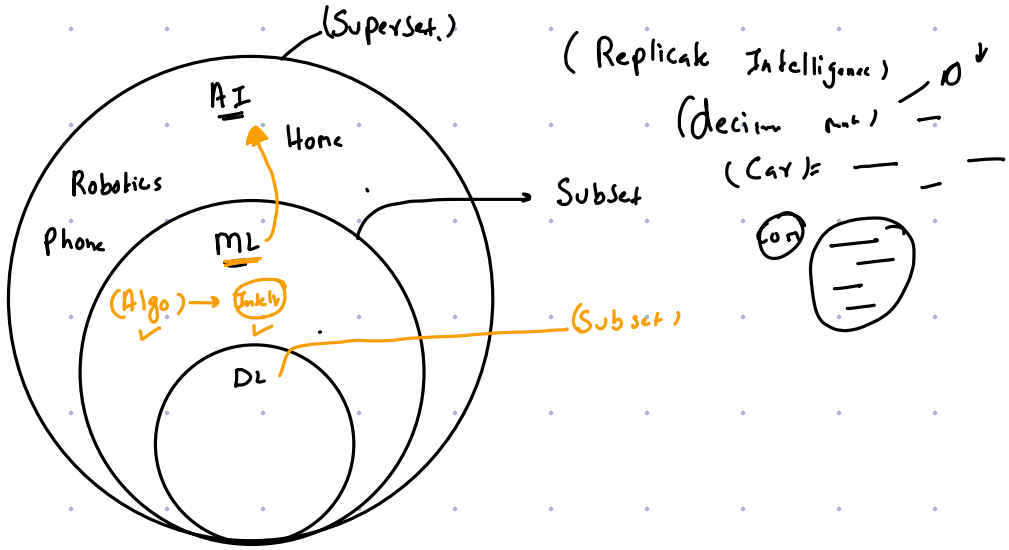
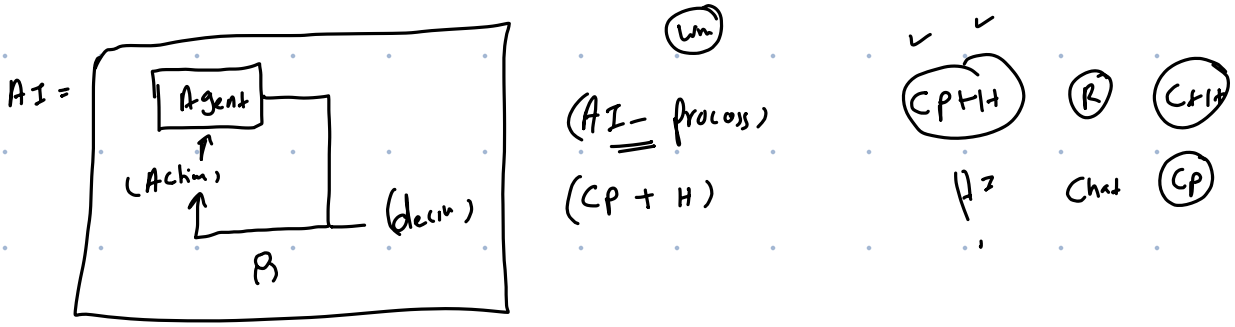
Namaste folks,

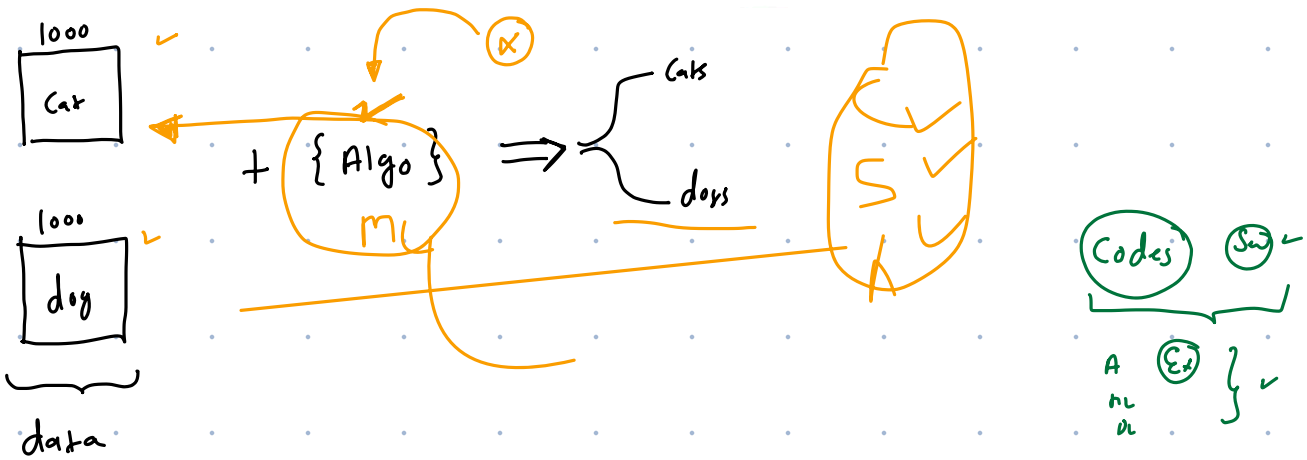
letls Start in 2-4 minutes

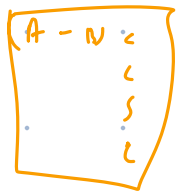


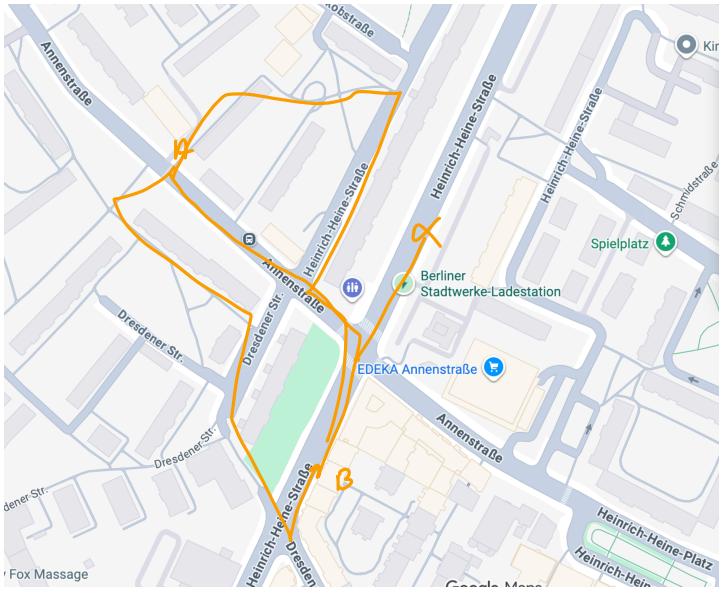


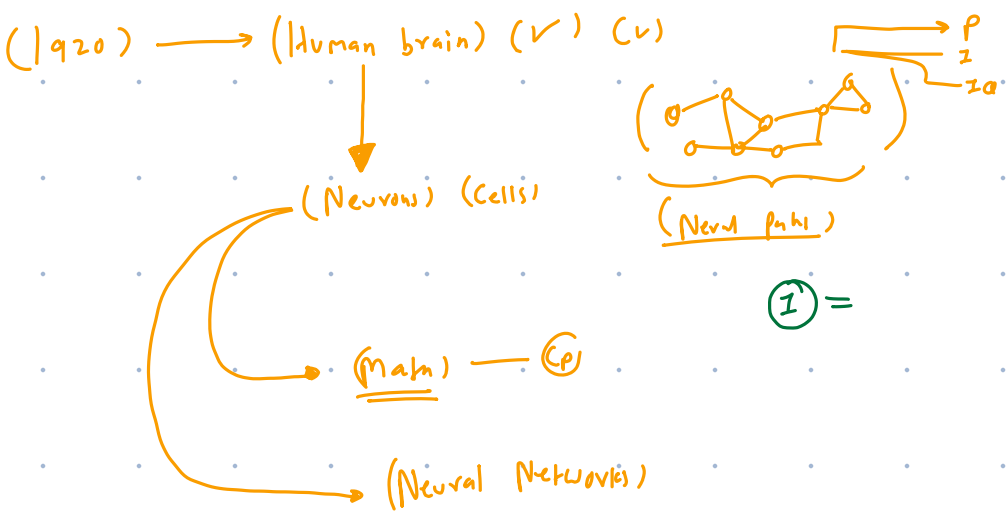
- Machine Learning is a subset of artificial intelligence which gives a machine the ability to learn without being explicitly programmed.

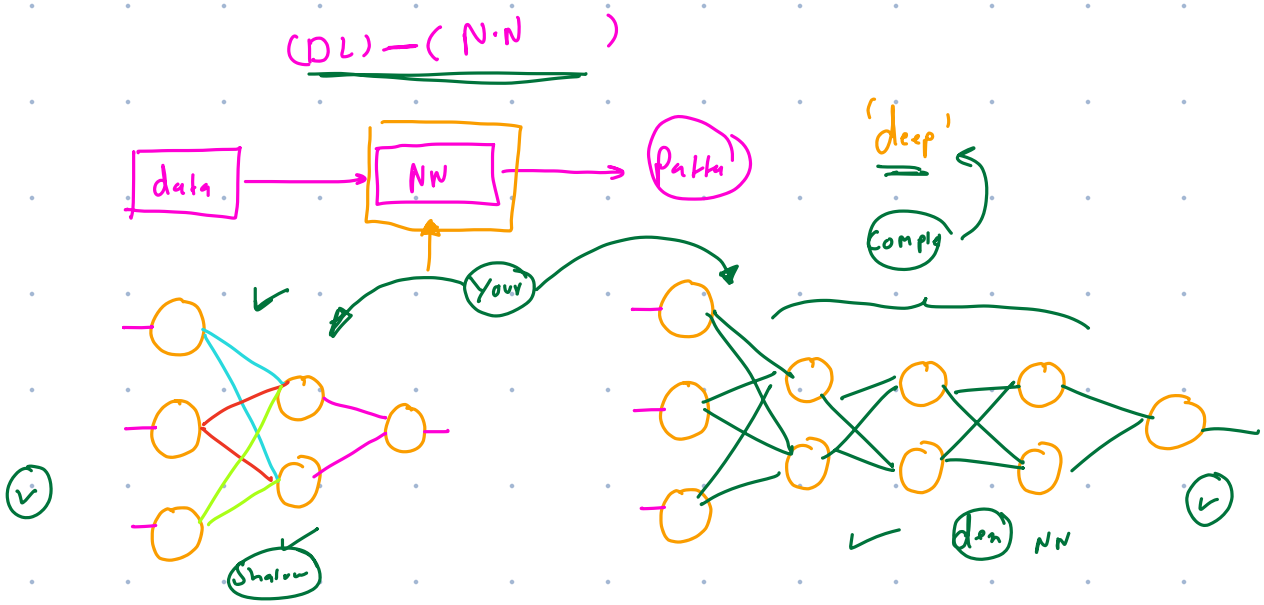
- Data is the key and the learning algorithm





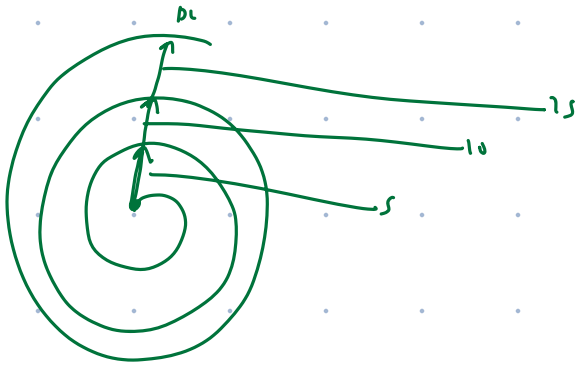

Annenstraße
Heinrich-Heine-Straße
Dresdener-Str.
Berliner
Stadtwerke-Ladestation
Spielplatz
EDEKA Annenstraße
Fox Massage
Heinrich-Heine-Platz





Algo = ((Stats)/ old) Prob — Algo

DL = NW — data



DL - (lot) { P C S }

ML - { }

(How about Now?)

(ML) → ① Supervised learning
② Unsupervised learning
③ Reinforcement

(SL): (data - label)

(Bank loan) Data

Age	Salary	Cibil	loan?
23	1.2	780	Yes
30	90	600	No
45	50	830	Yes
25	78	800	(Yes)

Yes / No (2)

Yes / No

hotel → Reviews → Pos, Neg, Neg

labels (3)

help

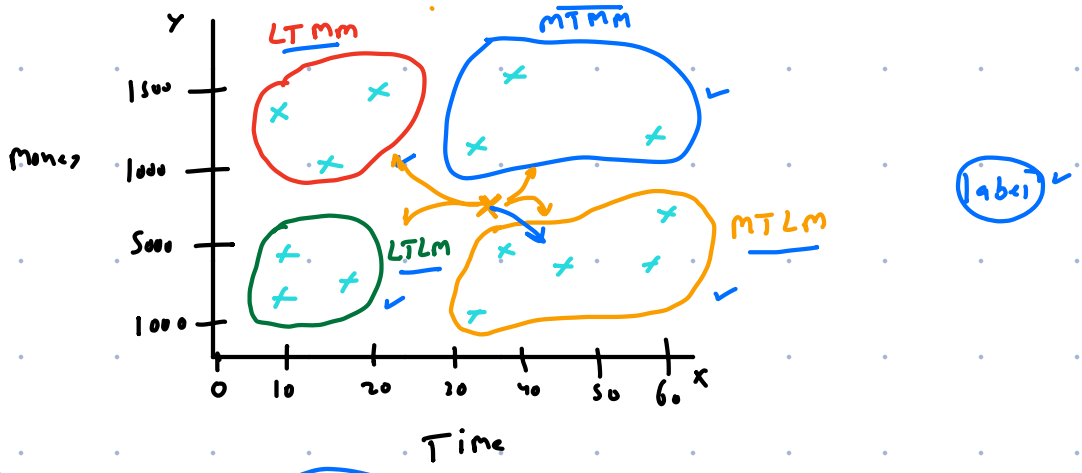
Age	Salary	Cibil
23	70	780

(data - No label) = USL

(Mall)

Time	money

Cluster



label

(USL) = Clustering / (groups)

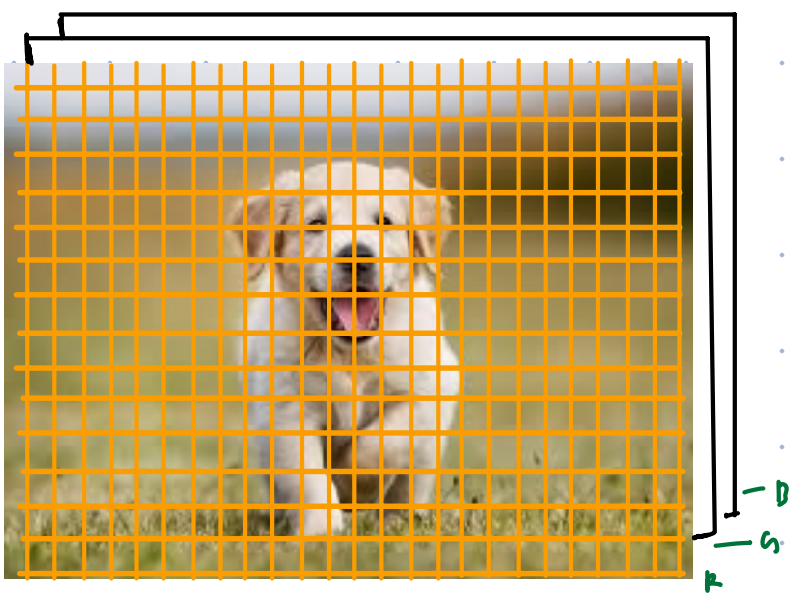
data / ij → labels

(RFE + NN) ⇒ (ANN) (one Arch)

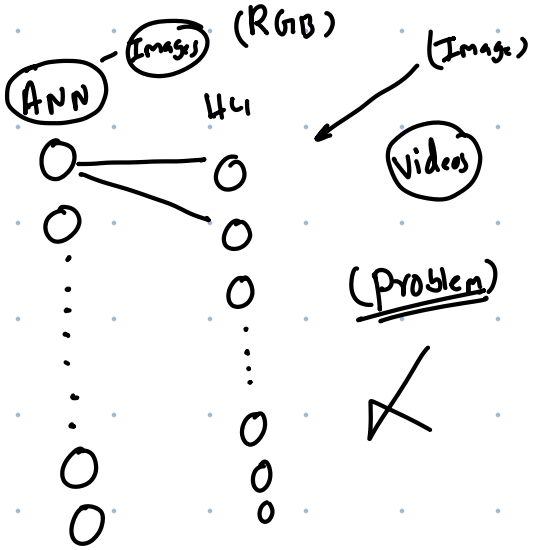
(O O O + O O + O → easy)

(Disadvantage) ⇒ (Image)

⇒ Dog/Cat



900 × 900 × 3 ⇒ 24, 30, 000/-



{24, 30, 000/-} + {800}

24, 30, 000 × 800 = (1.94×10^9)

w + b

(Image) => {(Image Kernels / Filters)} - Why?

(Color) => (Pixels) => (0 - 255)

White - 255

Black - 0

(Channels)

(8-bit) → (1 byte)

$2^8 = 256$ → (0-255)

$(2)^n = (2^1/2^2/2^3) \Rightarrow (2, 4, 8, 16, 32, 64, 128)$

(What = ?)

(Logic)

(Matrix (X x Y x 3)

(Edges)

(Numbers)

(Vertical Edges / Horiz Edges)

x (VEd) [m]

x HEd [m]

=> (Image)

(CNN) = (Convolution Neural Networks)

Arch = (Silver bullet) - (One Size fit all)

(Convolution operation) => (Hadamard prod) (Math)

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ 3 & 1 & 3 \end{bmatrix}$$

x → (Hadamard prod)

(Elem wise x^n)

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 1 \end{bmatrix}$$

R

C

=> dot prod

$A \odot B$ (Buns eye)

(Convolution) ① / ③

(4x6) Gry (3x5)

4x4 / 5x5 / 6x6 - (n+n) Pad

1

Image (6x6)

232	13	21	192	39	153
137	237	4	122	118	74
124	227	82	200	200	62
144	128	90	22	58	63
48	21	135	29	86	71
90	245	26	244	147	157

*

(3x3) (Kernel) ⇒

1	0	-1
1	0	-1
1	0	-1

3 × 3

$= \begin{bmatrix} 48 & 92 & 87 & 107 \end{bmatrix}$ 4 × 4

I/p (n x n) * (m x m) (3x3) = (n-m+1) x (n-m+1)

(6x6) * (3x3) = (6-3+1) x (6-3+1)

(Size) (n x m) → (900 x 200)

4 x 4

Clear

2D

10 30

(H x w) (K-h x K-w)

I/p

(n x c) * (x x z)

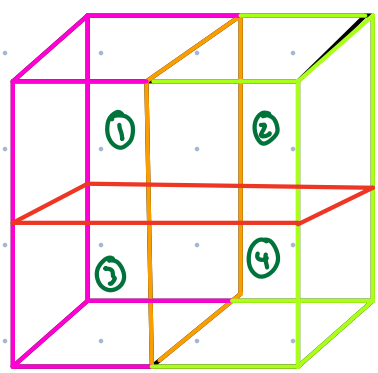
= (n-x+1) x (k-z+1)

O/p = (H - k-h + 1)

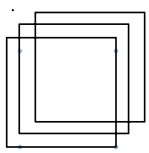
O/p = (W - k-w + 1)

1 0 -1
1 0 -1
1 0 -1

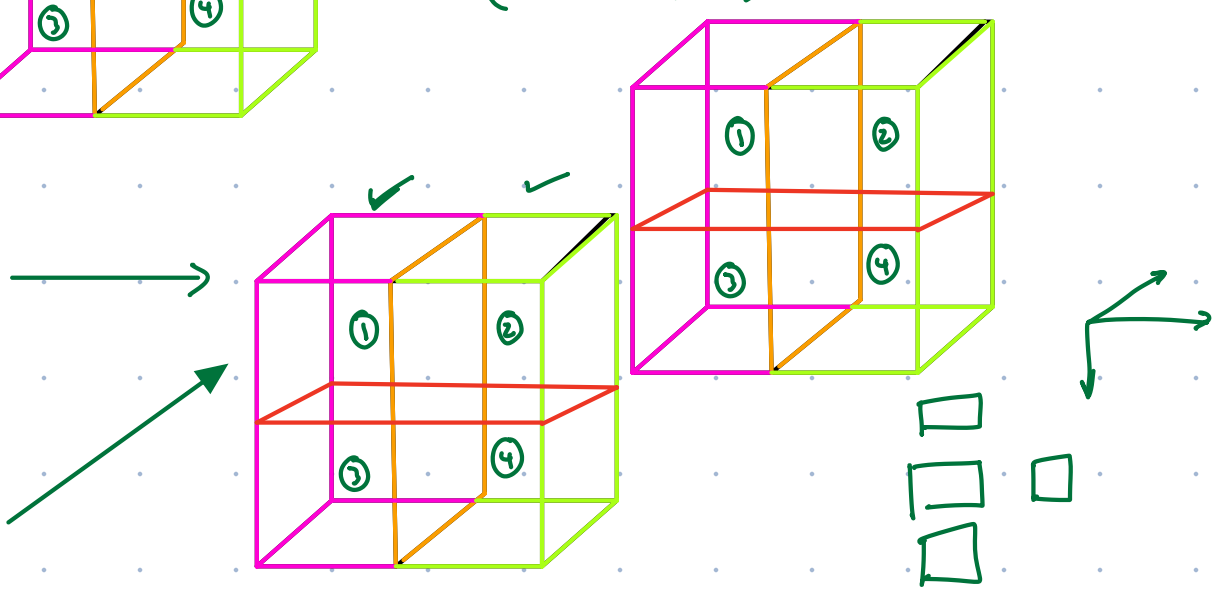
(RGB)

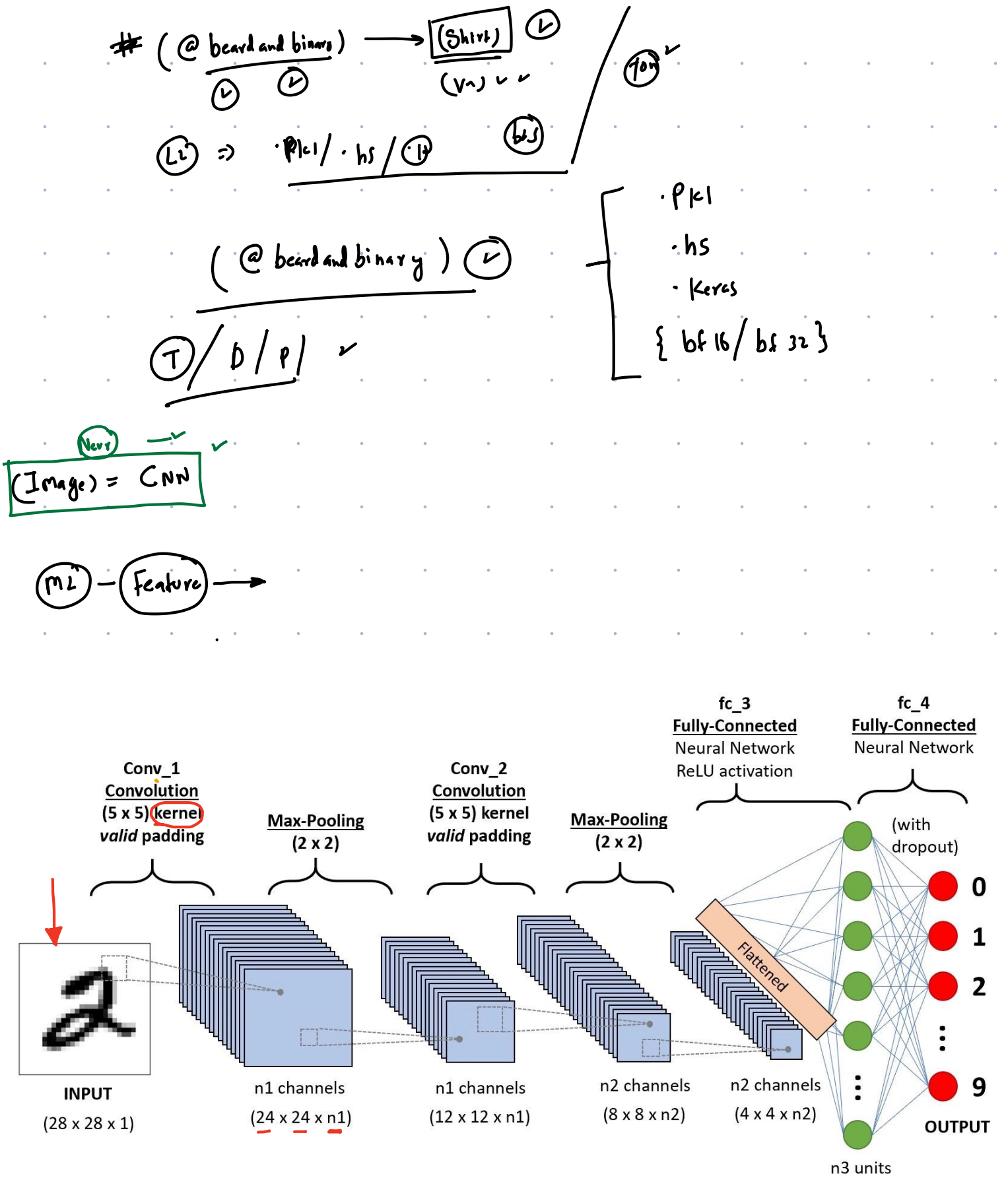


*



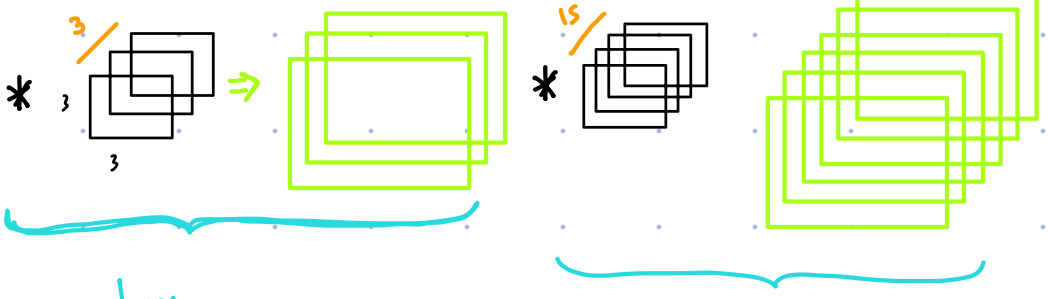
(Rubix cube)

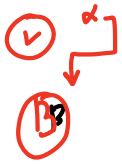



fc_3
Fully-Connected
Neural Network
ReLU activation
fc_4
Fully-Connected
Neural Network
Conv_1
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
Conv_2
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
(with
dropout)
Flattened
0
1
2
9
INPUT
(28 x 28 x 1)
n1 channels
(24 x 24 x n1)
n1 channels
(12 x 12 x n1)
n2 channels
(8 x 8 x n2)
n2 channels
(4 x 4 x n2)
n3 units
OUTPUT

26

26





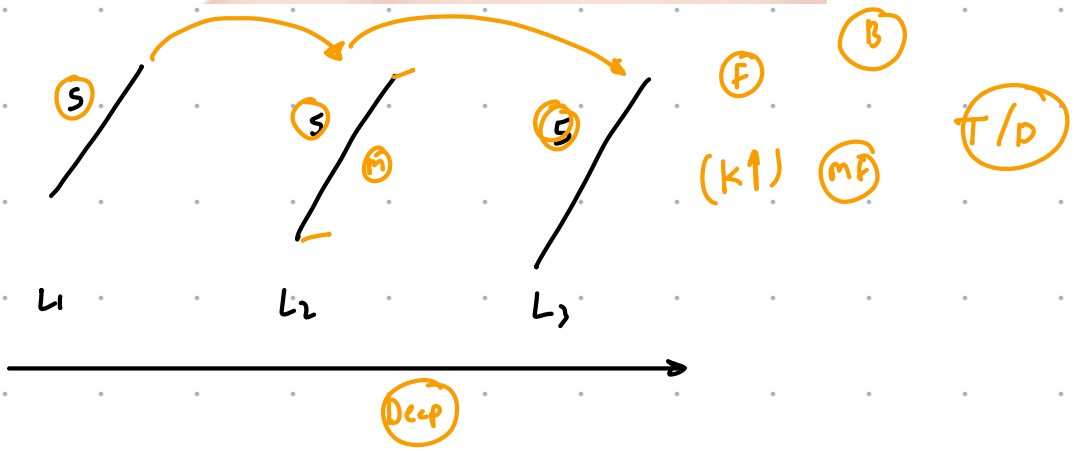
layer-1 = Basic (VE/HF/ etc)

layer-2 = Sep - V

layer-3 =

ANN - HL H_2

3) - 4)



Image

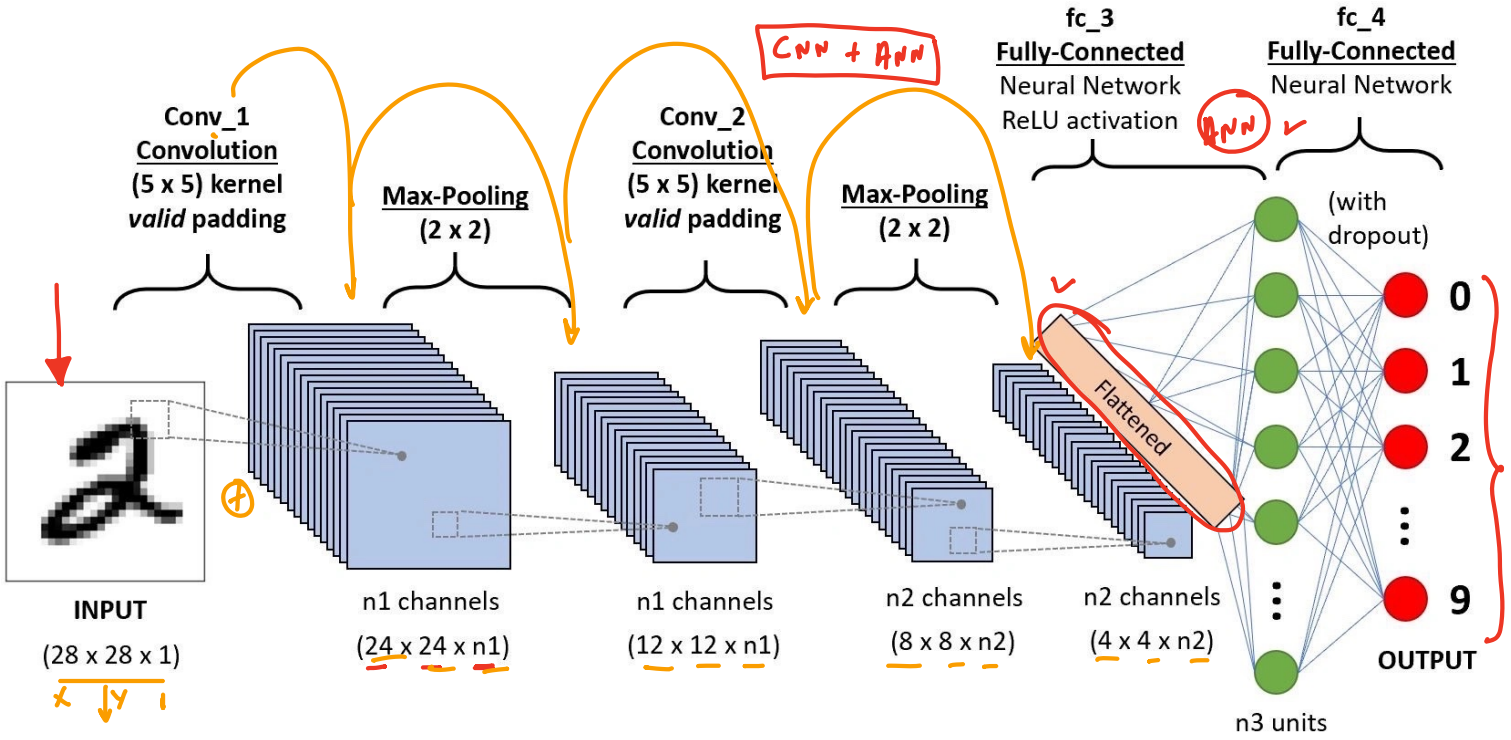
B

F

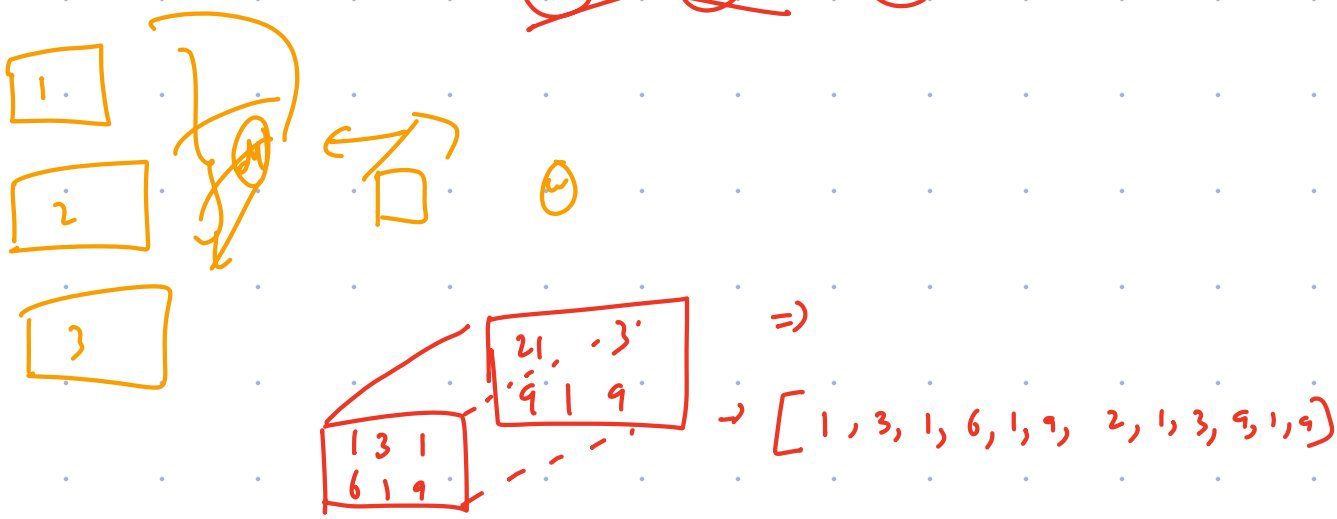
(K1)

mF

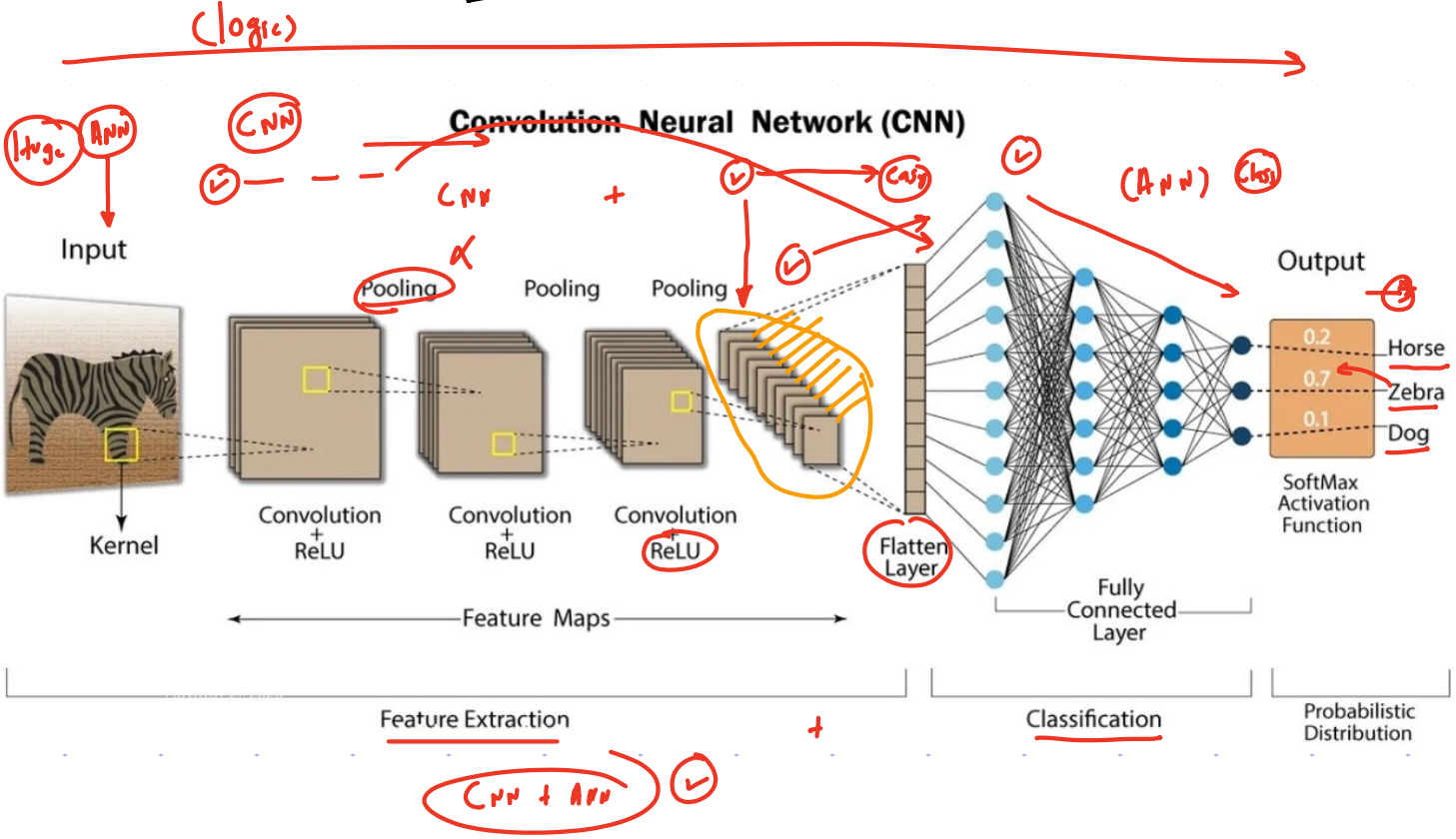
F/D

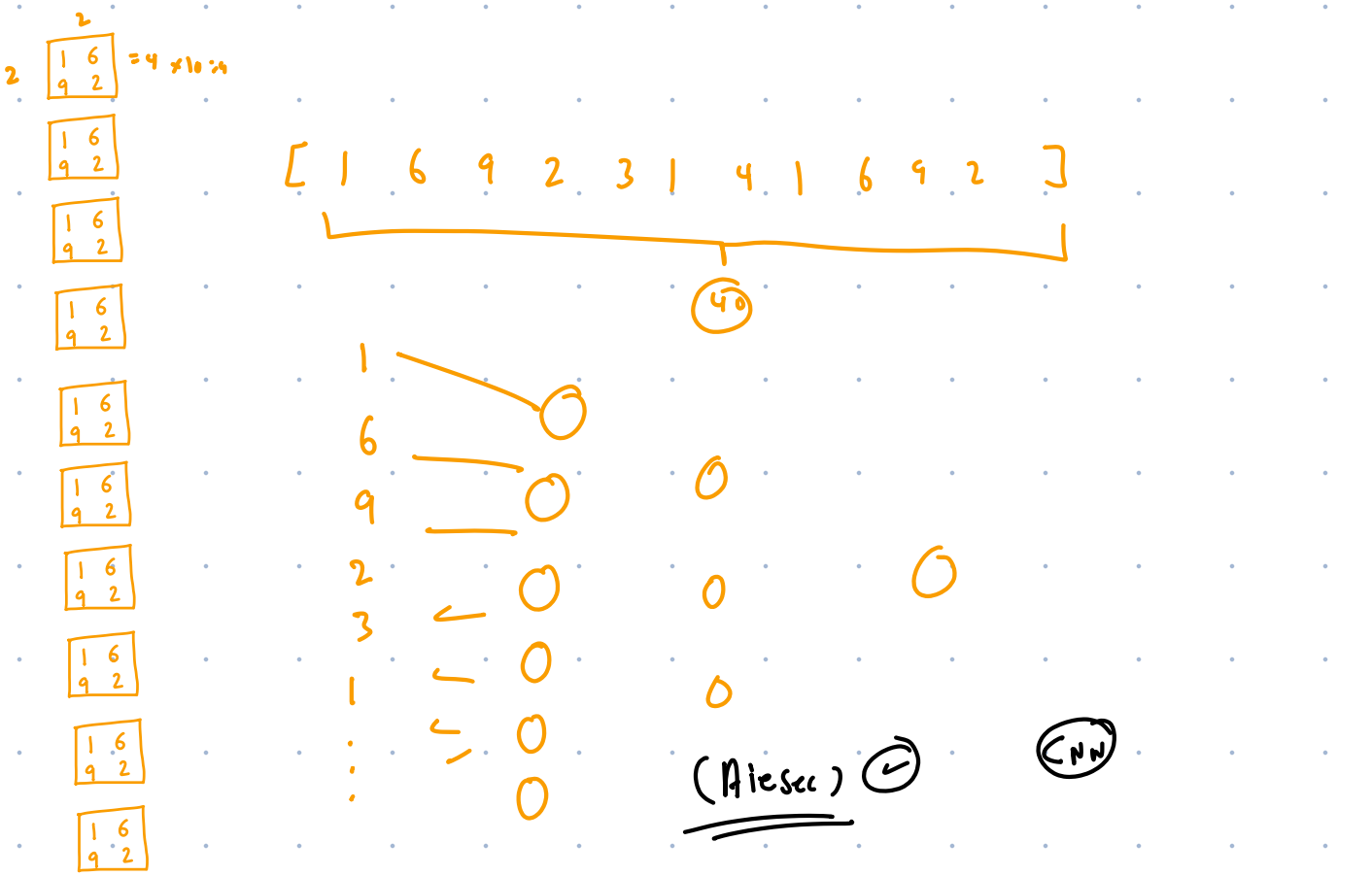

CNN + ANN
Conv_1
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
Conv_2
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
fc_3
Fully-Connected
Neural Network
ReLU activation
ANN
fc_4
Fully-Connected
Neural Network
(with
dropout)
Flattened
INPUT
(28 x 28 x 1)
n1 channels
(24 x 24 x n1)
n1 channels
(12 x 12 x n1)
n2 channels
(8 x 8 x n2)
n2 channels
(4 x 4 x n2)
n3 units
0
1
2
9
OUTPUT


CNN
Poly
Conv


1
2
3
=>

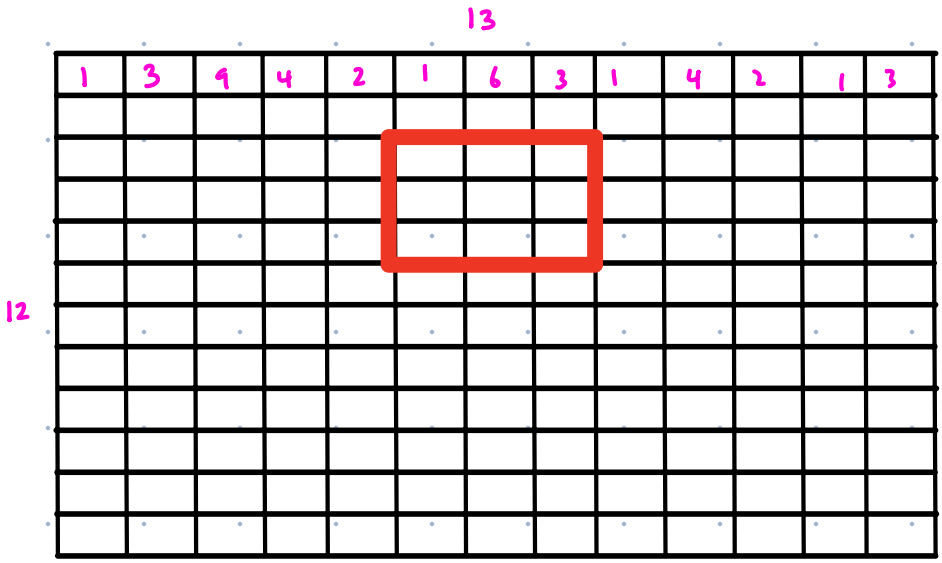
Arch :(basic)


Convolution Neural Network (CNN)
Input
Kernel
Pooling
Pooling
Pooling
Convolution + ReLU
Convolution + ReLU
Convolution + ReLU
Feature Maps
Flatten Layer
Fully Connected Layer
Output
0.2
0.7
0.1
Horse
Zebra
Dog
SoftMax Activation Function
Feature Extraction
Classification
Probabilistic Distribution


[1 6 9 2 3 1 4 1 6 9 2]

① Striding ✓

CNN = Stride = 2 — Jump



Conv = St = 1

(Reading a book) — ✓ ✓ ✓

↳ E/T/ → (go trough quickly)

Detail - und

↳ lost

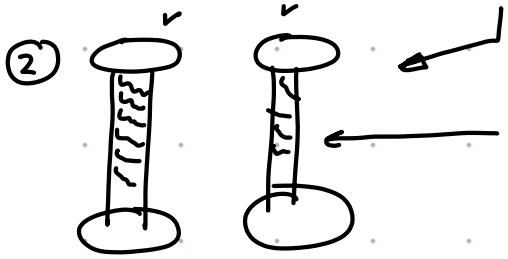
∝ — k

Why

Read Every word	go trough fast	Speed read.
Time	less	
understand	less precise	
S = 1	S = 2	S = 3

① (Dim Red) → (x-x') → (y-y') ↓ (less Comp cost)

②




gettyimages
Credit: SimonSkafar
2157121002

(Padding) ⇒ What - ?

Why - ?

(IP/VP/NLP/AP)

Rep

163	9	104	46	177	28
123	252	52	241	216	88
218	122	214	150	111	40
31	24	208	75	0	253
148	30	84	81	156	110
202	31	254	44	227	110

① 163 - 1

252 - 1, 1

214 - 1, 1, 1

more

P - S

Padding

C → P

0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	50	100	150	100	50	0	0
0	0	100	200	250	200	100	0	0
0	0	150	250	255	250	150	0	0
0	0	100	200	250	200	100	0	0
0	0	50	100	150	100	50	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

S=1 9

S=2 ~~50~~ 4

S=2

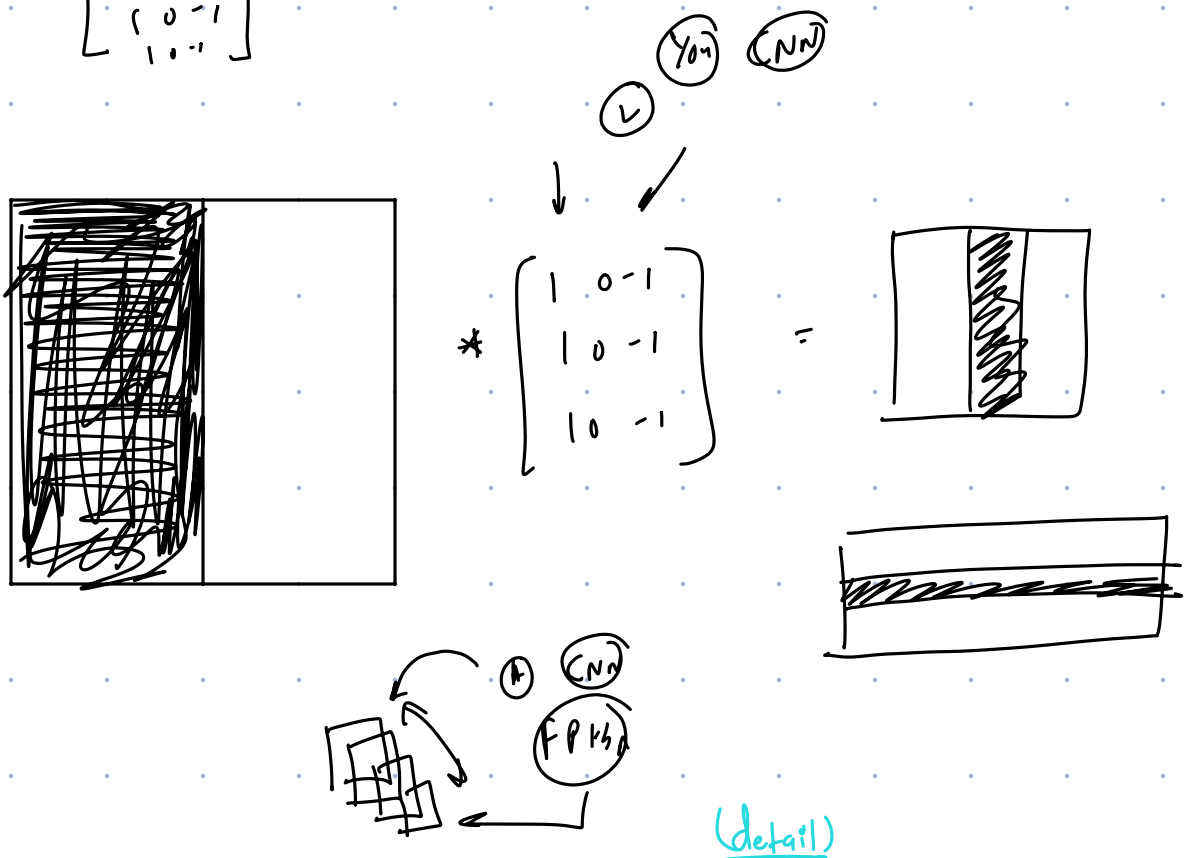
CNN = (Linkedin)

(Theory)

(CNN)

(Predefined) = (vc)

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} = (S)$$



(detail)

{CNN + Striding + Padding} ⟶ L

(Pooling) ⟹ (Min / Avg / Max) (0-255)

(Why?)

Max

(more imp)

0-255

CNN

(Window Size)

(image)

3	9	6	4	1	16
6	8	14	18	19	20
35	255	0	16	9	8
155	124	79	82	87	19
21	175	14	35	91	71
26	133	103	87	42	81

2

2

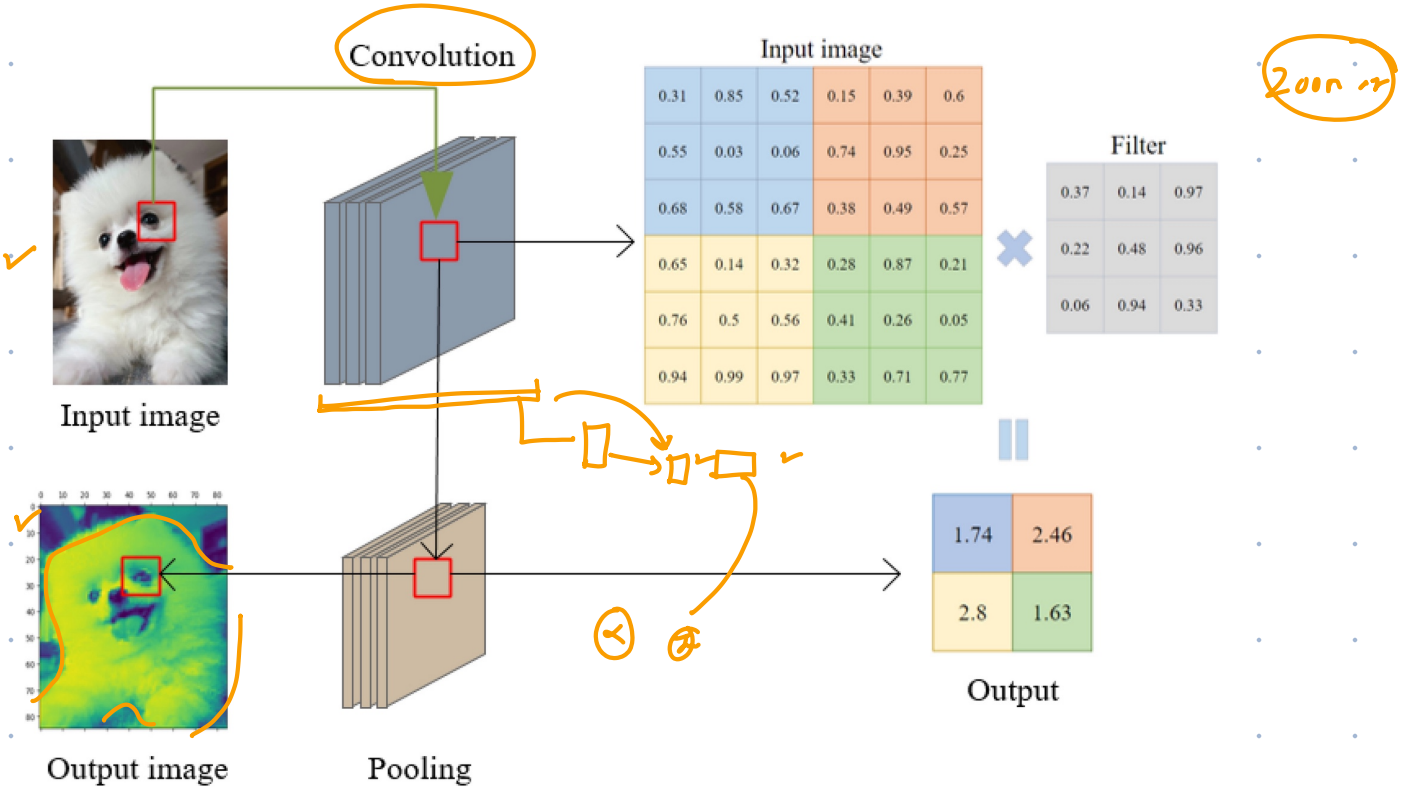
9	18	20
255	79	87
175	124	91

min

(min / max)

21

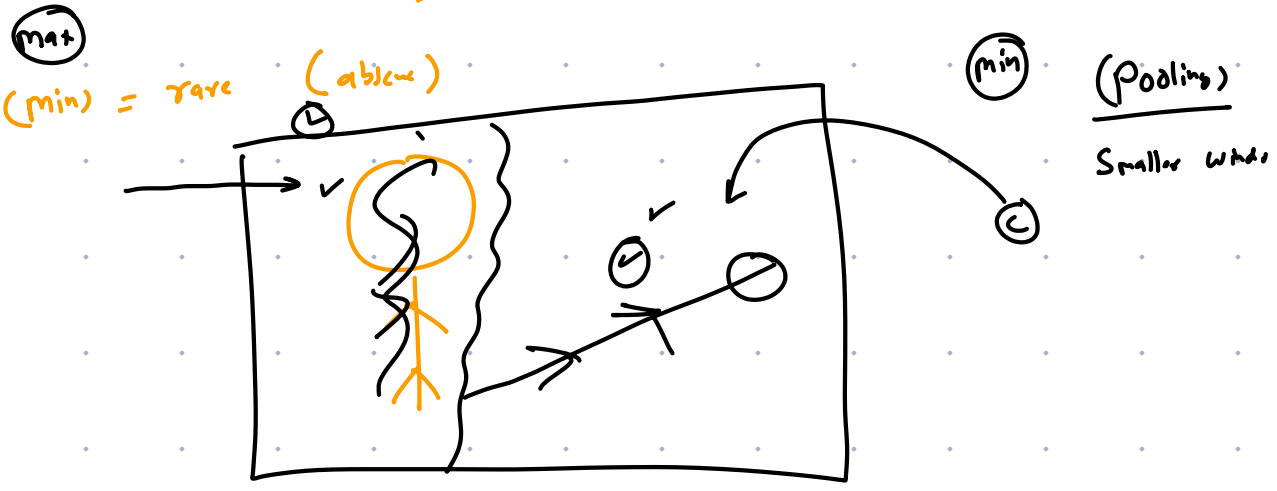
Pooling : " down sampling operation that reduce spatial dimension image (H - W) such that import features in the image are retained, "



(max) - (features) (from) = (HD) 0-255 - 21 - (175)

(downsamp)

(Irr) α (θ)



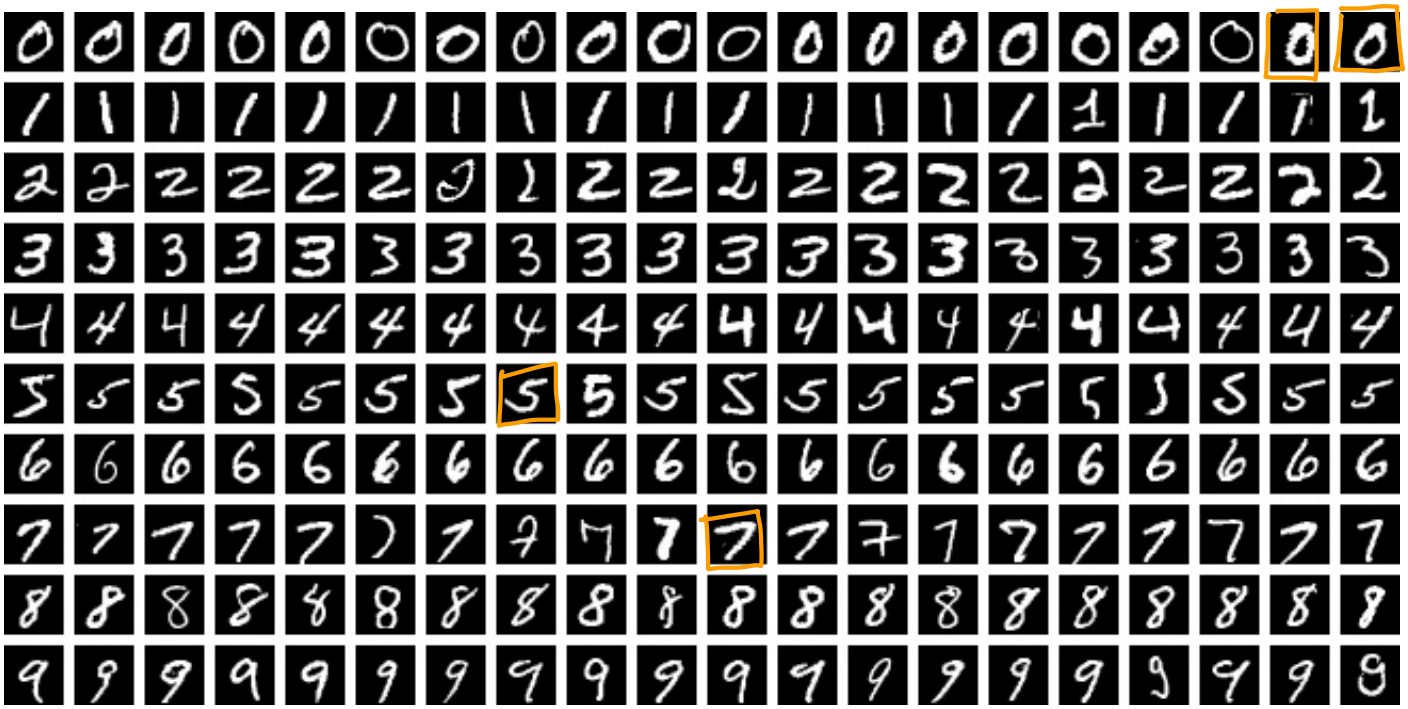
(Good phones) - HD

CNN

2001

(MNIST)

(HQ) → Simple (24 x 24)



(Pros)

ds

Clear

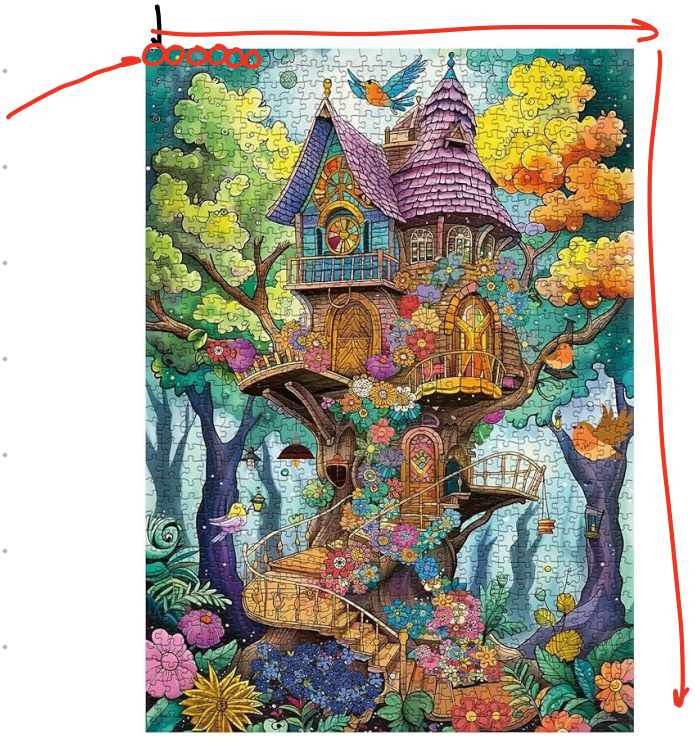
MM →

lw

lw

8 → 96%

Irl

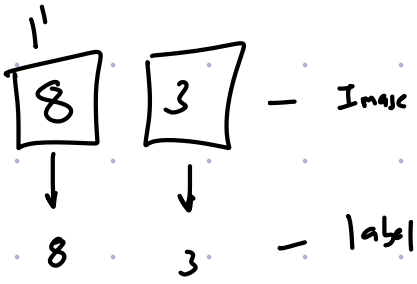


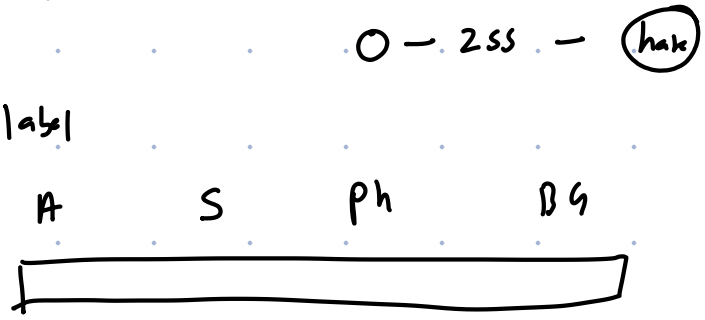
→ (BB) = (Roi)

Regr or

CP

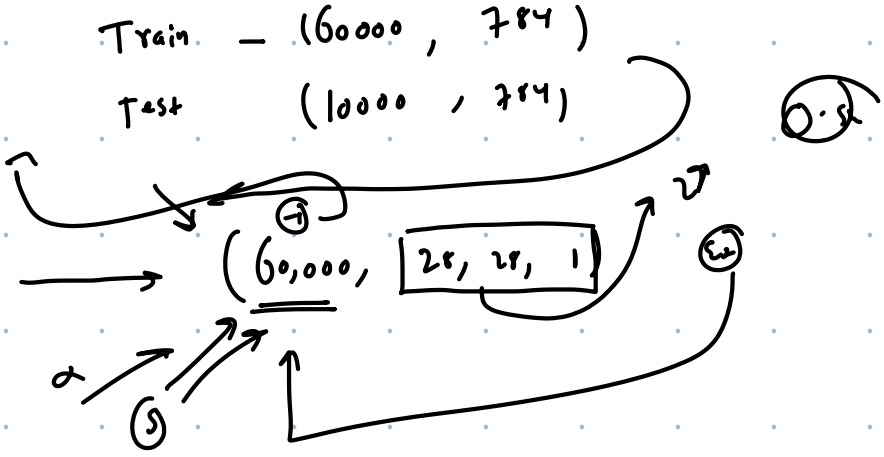
CL





0 - 255 - hair

0 - 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 (labels)



(Non-Tech) = Calm!!

= SQL = (1000%.) 1/2 (great)

(CTE) → (optimize)

Data Warehouse

(SQL)

(Snowflake | Big query)

Basic

Python

I/O

Algo + Tech

Projects

(Stats) → (1-2)

(depth) end

IS/ds

(Stats + ML)

Not

Python

loops

India → F

(1-3)

Movie

Xing

Chanel

(LinkedIn) — (rec)

0 — 9 — 10

(One-hot encoding)

0 - [0 0 0 0 0 0 0 0 0 1]
1 - [0 0 0 0 0 0 0 0 1 0]
2 - [0 0 0 0 0 0 0 1 0 0]
3
4
.....
9

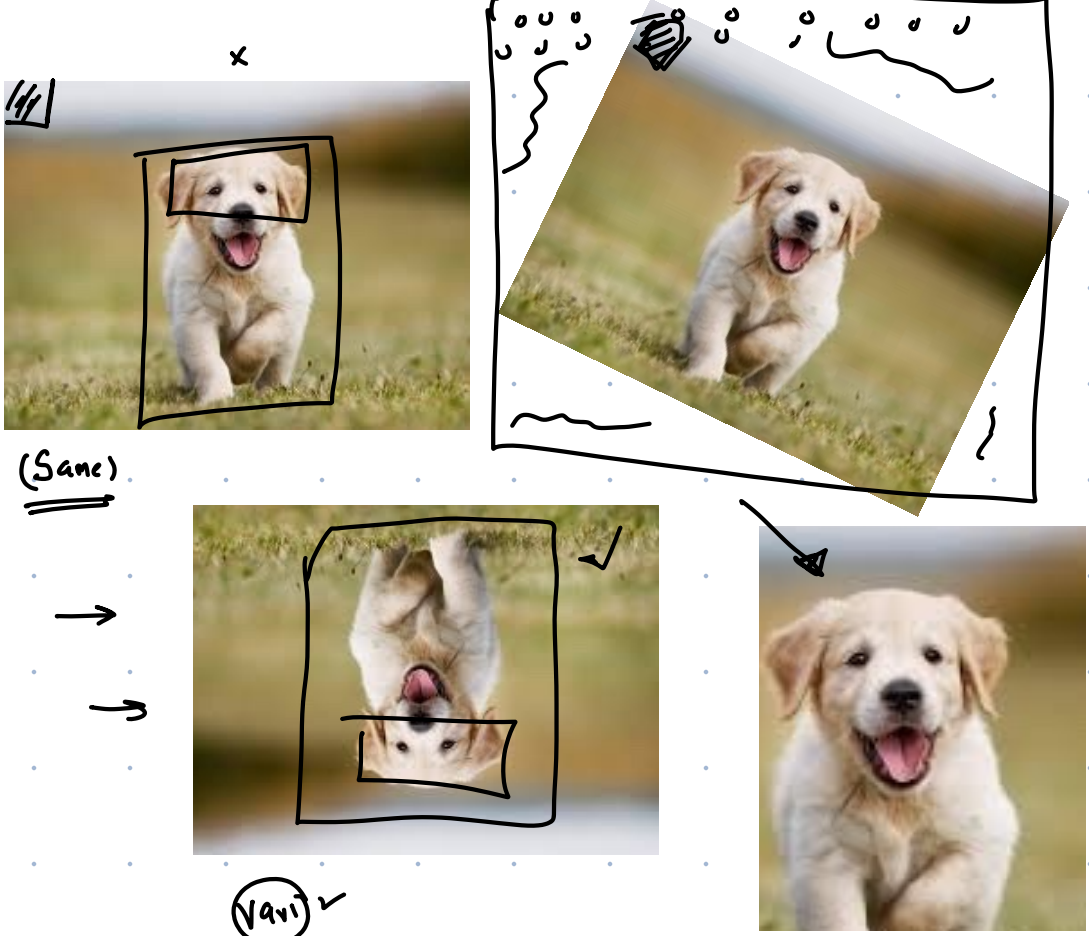
Namaste folks,

lets start in 2-4 minutes

(30-day - Challenge) (Σd) = (Post)

Source:

Backlog



(Data aug)

1000 5000

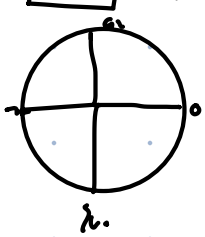
Vari

Zoom

2π radians / degree

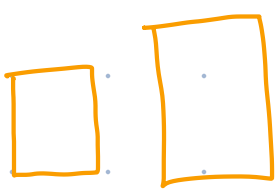
```
#data augumentation
data_augumentation = tf.keras.Sequential([
    layers.RandomRotation(0.1),
    layers.RandomZoom(0.1)
])
```

any

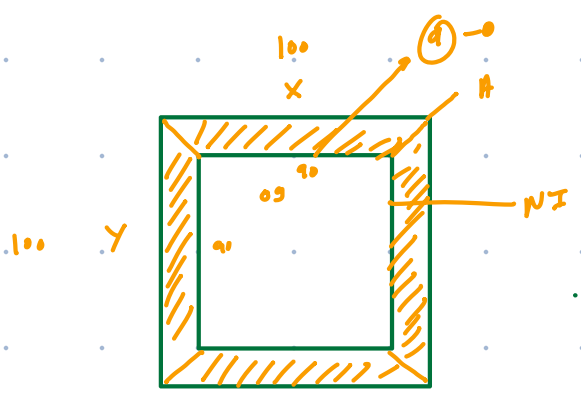


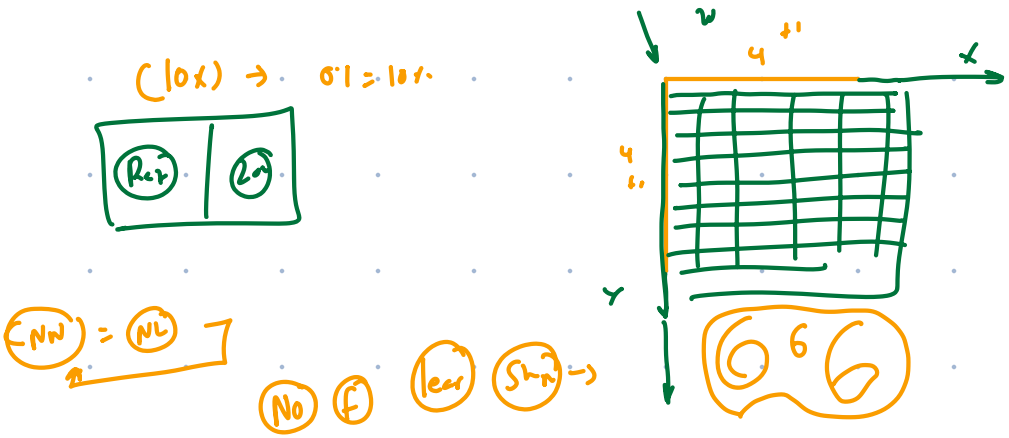
0.1 x 360 = 36° – (Random) −36° – +36

0.1 = 2P = 10%

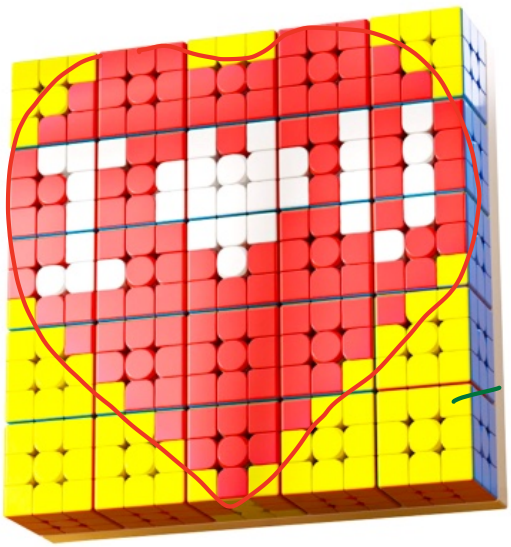


Zoom- Yes





```
#define Cnn
def create_model():
  model = models.Sequential([
      data_augumentation,
      layers.Conv2D(32, (3,3), activation='relu',
                    input_shape=(28,28,1)),
      |
  ])
```



(activation) = (non-linearity data)

Why ?

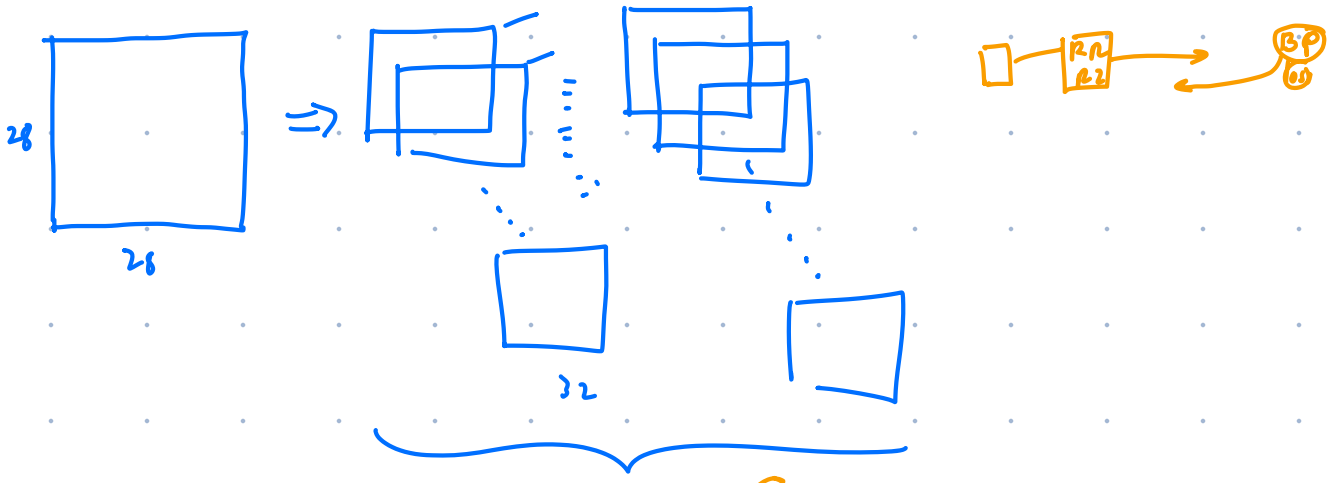
-ve → +ve

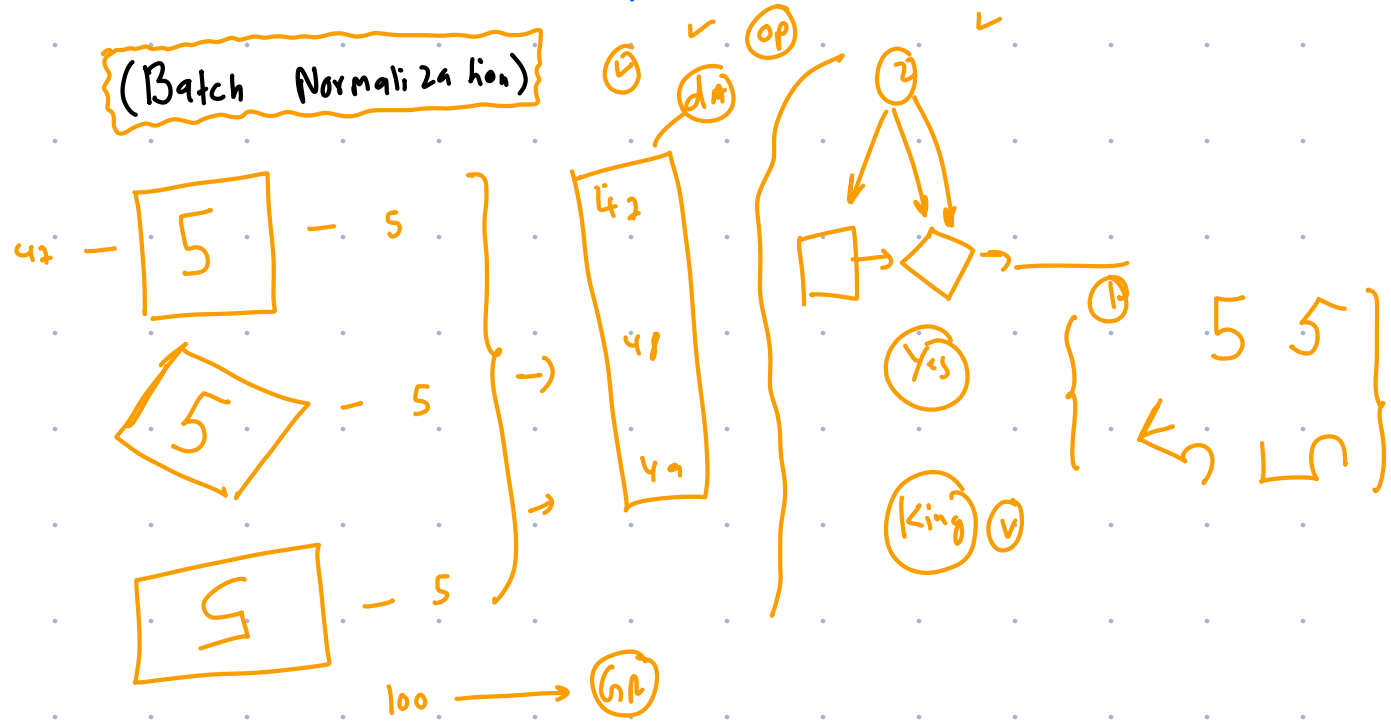
(0-255)

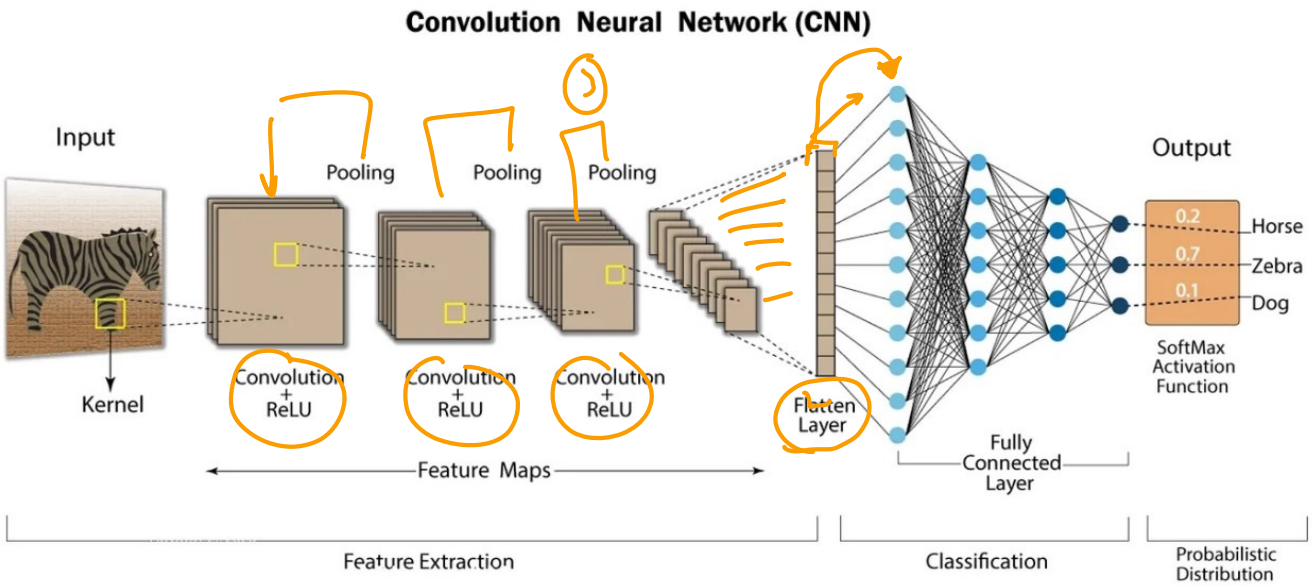
-1 0 +1

(Relu) = (Strong edges) → +ve 0 -ve

CNN + Ann

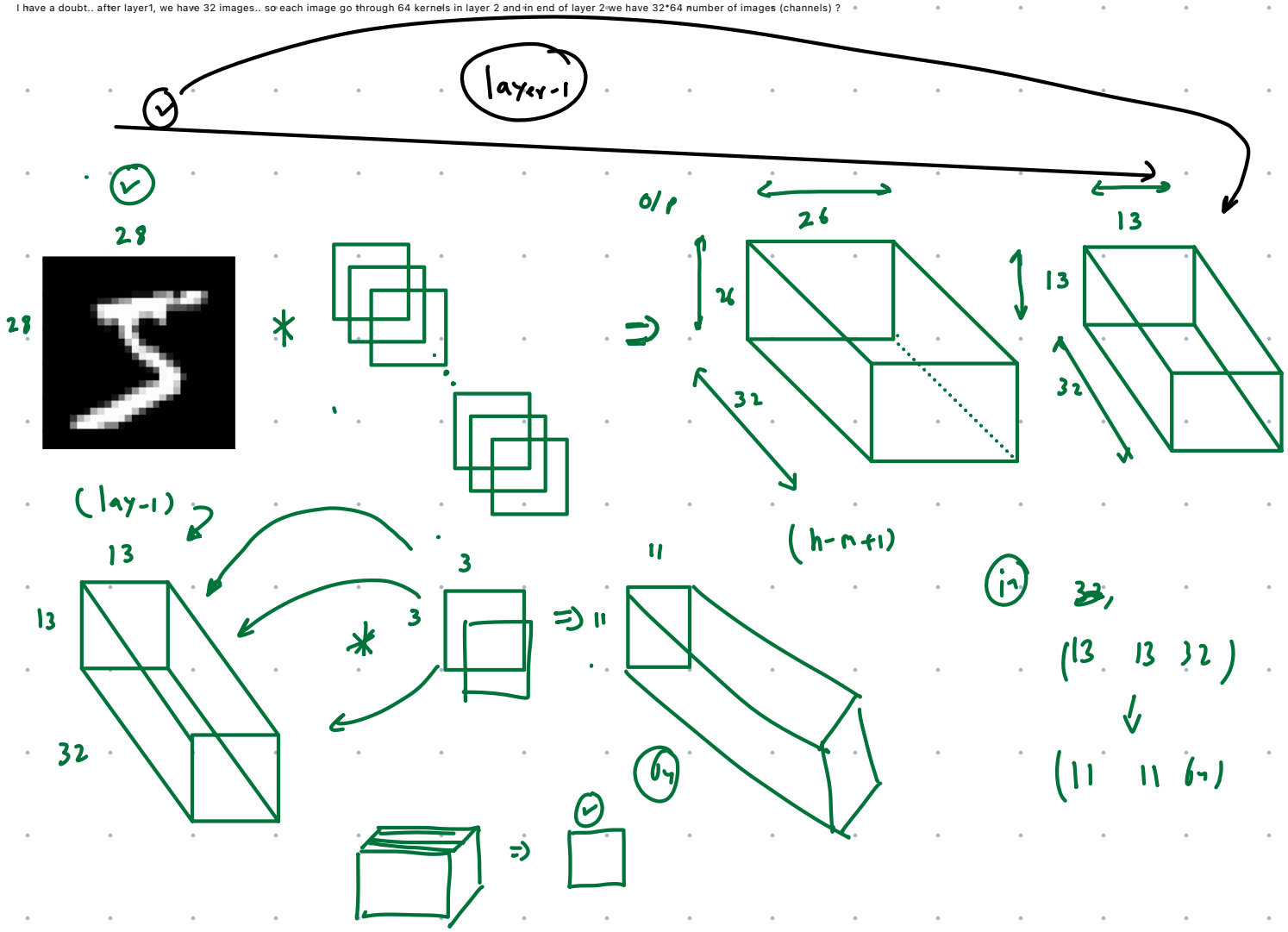

28
28
32


(Batch Normalization)
5
5
5
5
5
100
100
Problem
(Batch Normalization)
(India)
(USA)
(Germany)
(Card)
(Stable)
Relu
Batch
OG
*
OG * (K/R) → BN - O/P

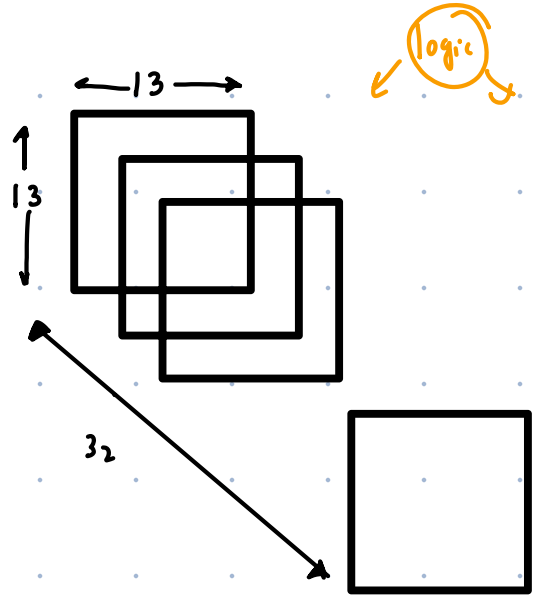


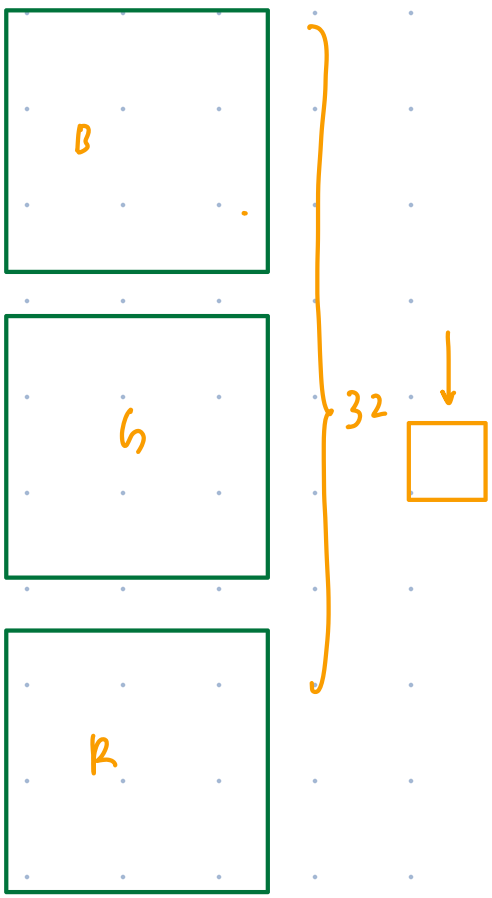
$(h-m+1) \times (h-m+1)$

I have a doubt.. after layer1, we have 32 images.. so each image go through 64 kernels in layer 2 and in end of layer 2 we have 32*64 number of images (channels) ?



(Book)

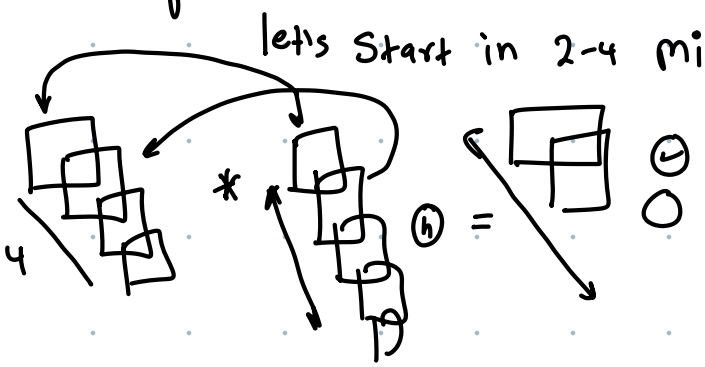




= 96 196

Namaste folks,

lets Start in 2-4 minutes



L1 — lasso ridge (Elastic-Net)

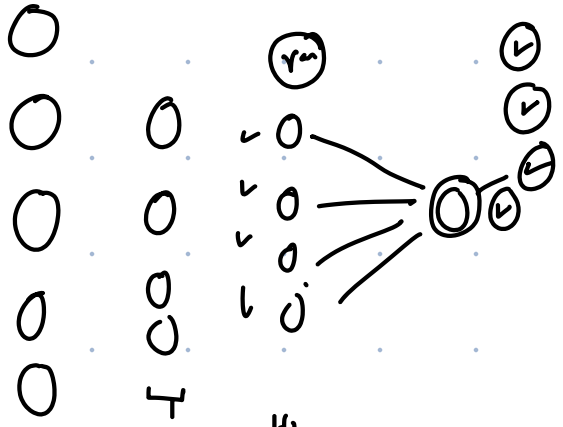
L2 —

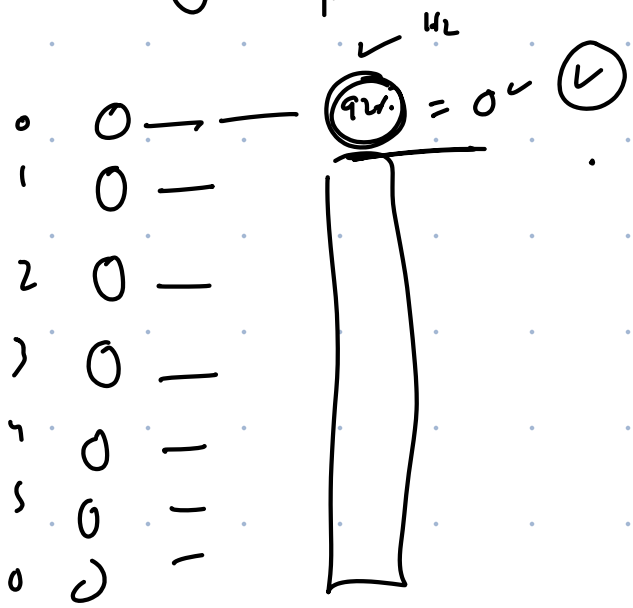
Reg → Overfitting

Observed

(20L) → (40)

(Dropout reg)

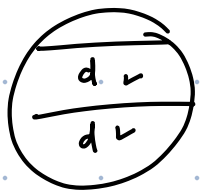




(Softmax)

(Sigmoid) → (0 -1)

(Softmax)



- Pos
- Neg
- Neut
- Tot

{ Categorical Cross entropy
Sparse Cross entropy } → (Translate)

(E - H)
(H - E)

Voice - ofc

(Voice)
→ (DSP) → (2019) =?

{ learning }

(500 +)
(Pre-req)

(gap) → (Wired)

Sale

Busin

* (Learning rate decay)

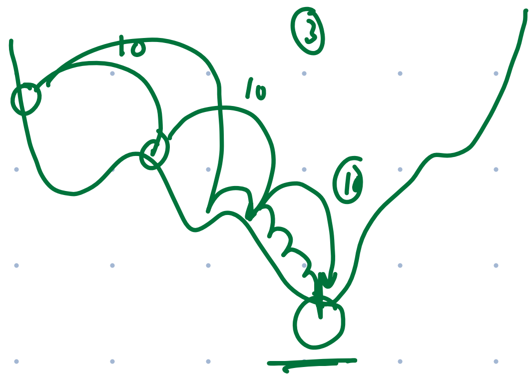
ES

optimizer = 0.001

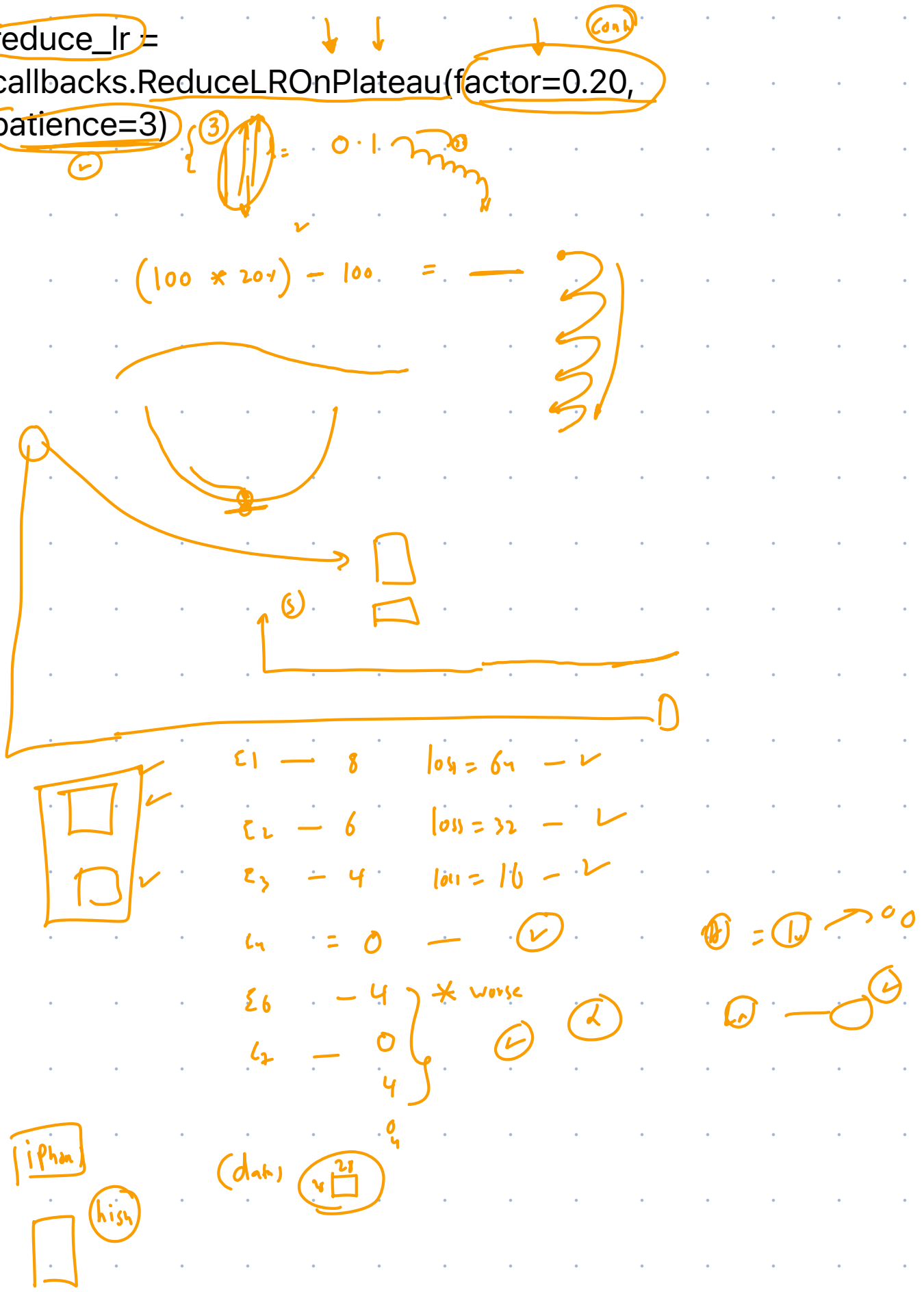
(Constant)

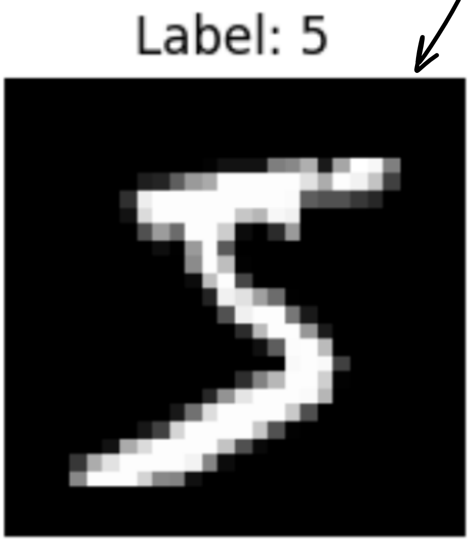
(adaptive) learning rate " reduction "

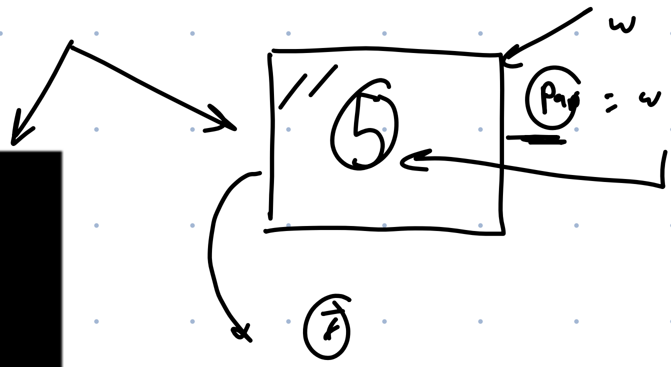
(Treasure)

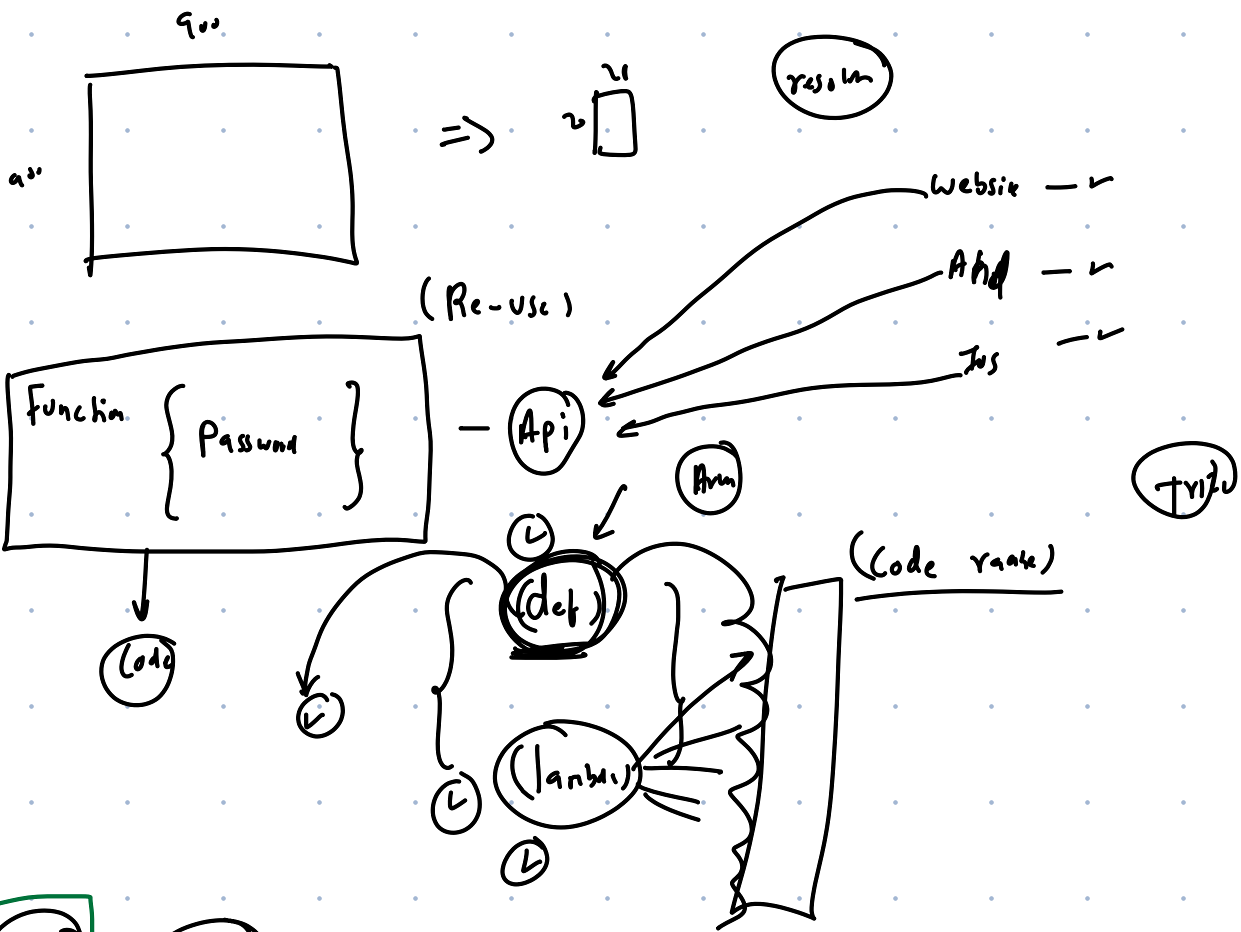


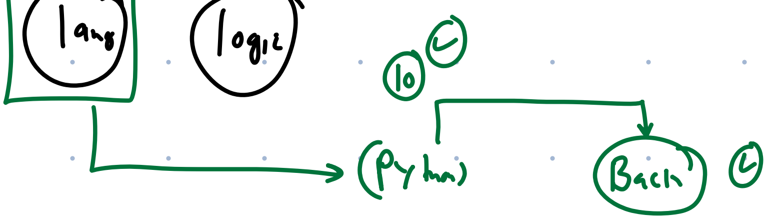
```
reduce_lr =
callbacks.ReduceLROnPlateau(factor=0.20,
patience=3)
```

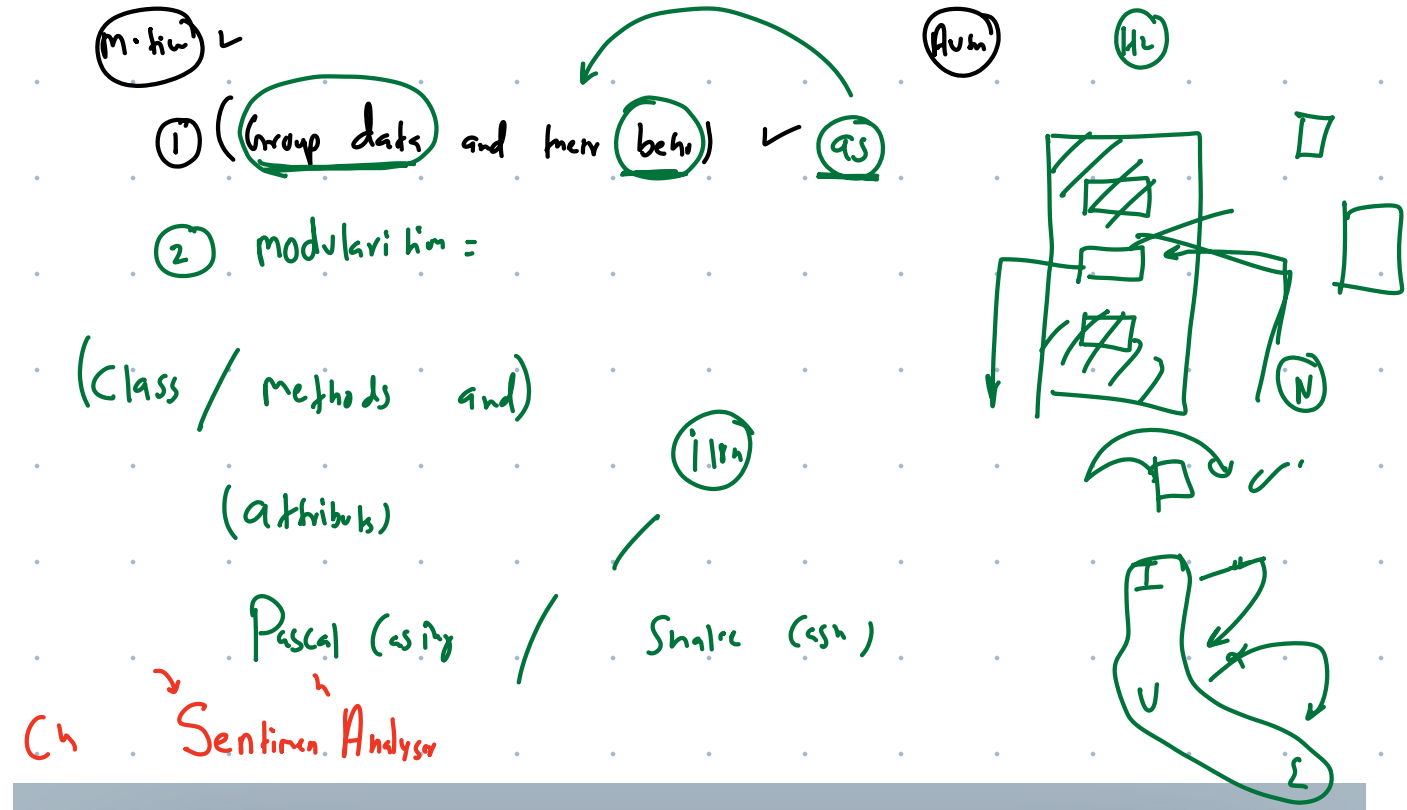



Label: 5


w
5


Function
Password
(Re-use)
Website
Api
(def)
(Code reuse)


lang
logic
(Python)
Back



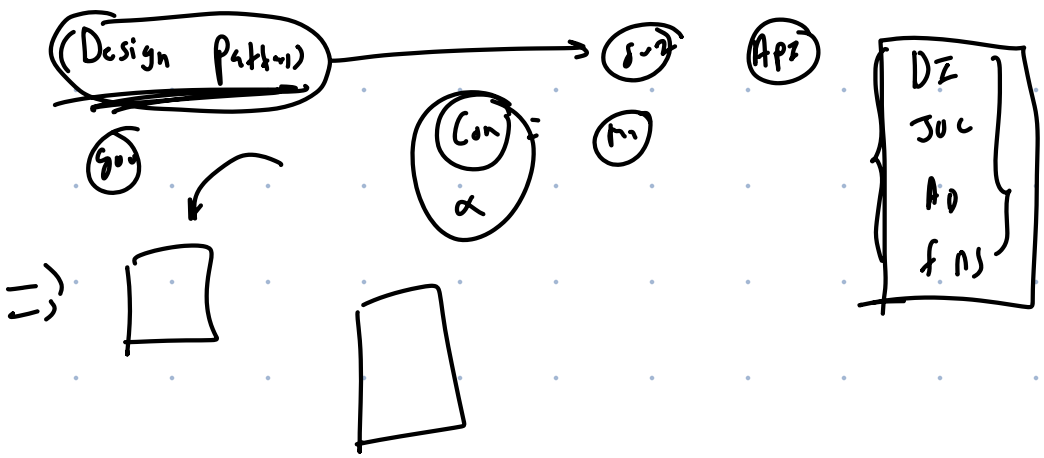

```
{
    "camelCase": "camelCase",
    "PascalCase": "PascalCase",
    "kebab-case": "kebab-case",
    "snake_case": "snake_case",
    "SCREAMING_SNAKE_CASE": "SCREAMING_SNAKE_CASE",
    "Train-Case-Name": "Train-Case-Name",
    "UPPERCASE": "UPPERCASE"
}
```

(logic)

(dunder)

```python
class Dog:
  species = "hund"

  def __init__(self, name, age):
    self.name = name
    self.age = age

  def bark(self):
    return f"{self.name} says Woof Woof!"

  def get_info(self):
    return f"{self.name} is {self.age} years old!"

dog1 = Dog("capser",3)

dog1.bark()
```

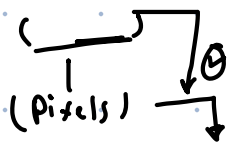
```
'capser says Woof Woof!'
```



Arch = (Why?)

Image = CNN / Generic = ANN

(NLP) = (RNN / LSTMs / GRU) = "text"

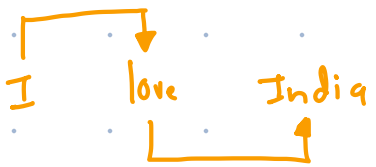


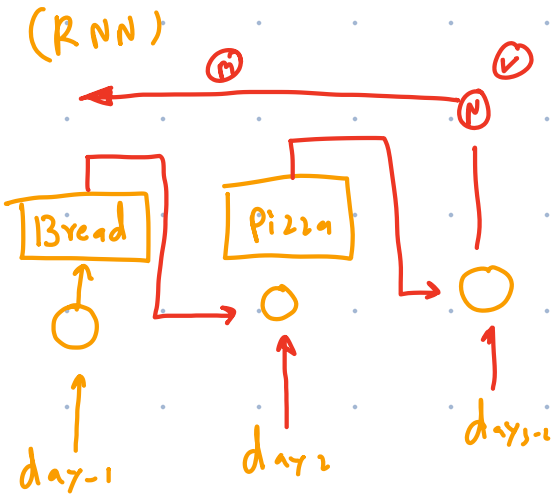
(RNN) => "Text" or "Time Series"

(Why ?) (CNN / ANN) (Man eats hot-dog / hot-dog eats man / word order is Imp

(1) (Size) = fixed or Variable => (I love movies) (I really enjoyed the movie)

(2) Order





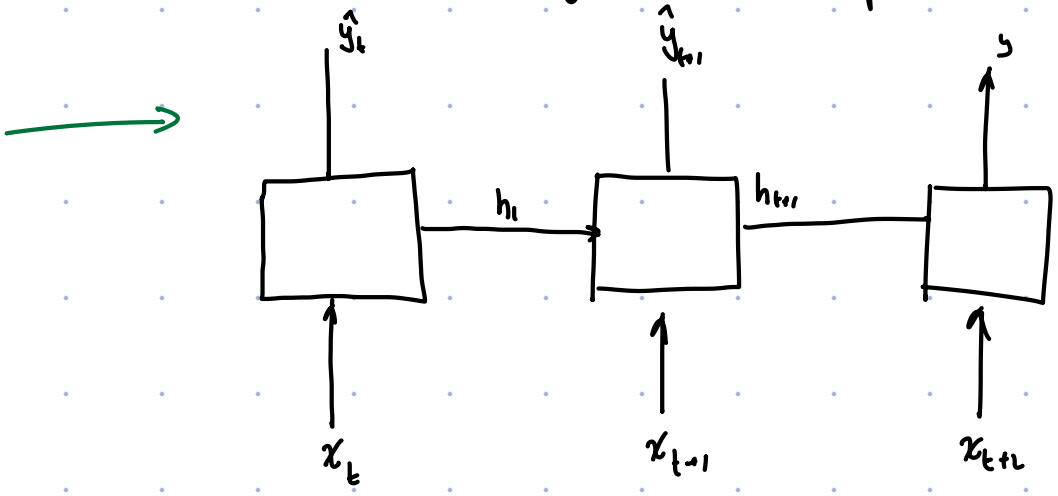
(1) O/p are dep on each other

(2) Sequential

(3) Memorize prev I/p

"Recurrent"

" Perform the same task for every element of a sequence, where in the o/p being dep on the prev comp"



(I love India)

✓ One-2-One => Image → 0/1

Abusive → "Abuse"

" " → "Abuse"

✓ One-2-Many => Image → Nature

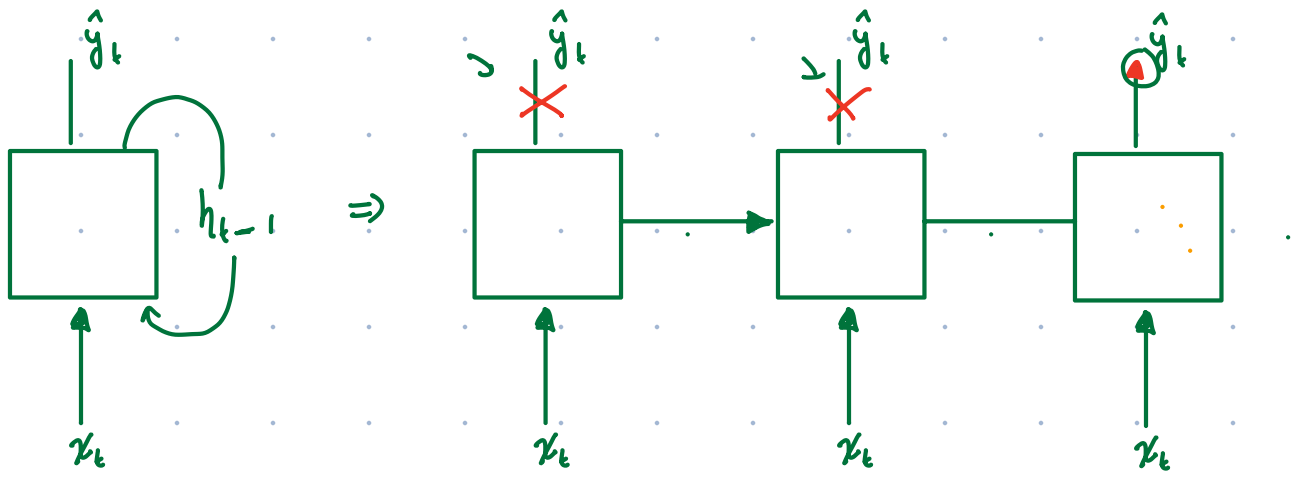
→ Forest

" " → Pos

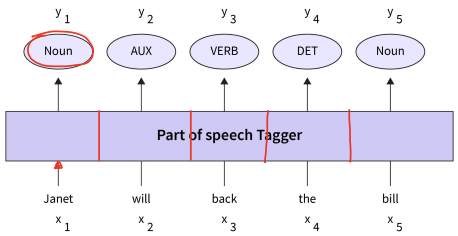
→ Food Review

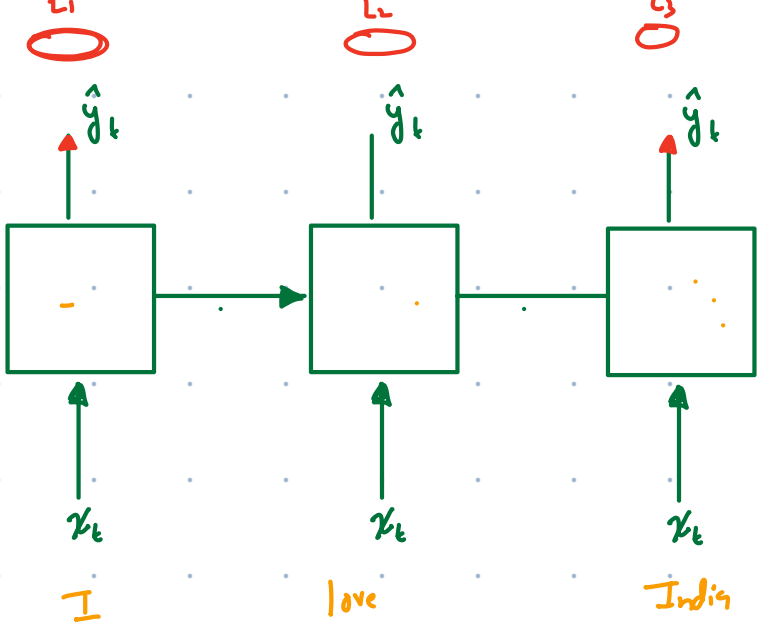
✓ Many-I-One => (Imag) → Nature

✓ Many-2-Many => (Imag) → ( )



(Parts of Speech)



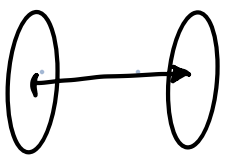


I/p O/p

One - 2 - One Mapping (lemma)

Running → run

Cats → Cat



I/p O/p

(One - 2 - many)

Bank → Finance, river edge, tilt, plane/ship

(I/p)

I/p O/p

Many - 2 - One

"Trump lost this New dreams" → (geo politics)

"LLm can be dangerous when used by kids" → Tech

(Topic modeling)

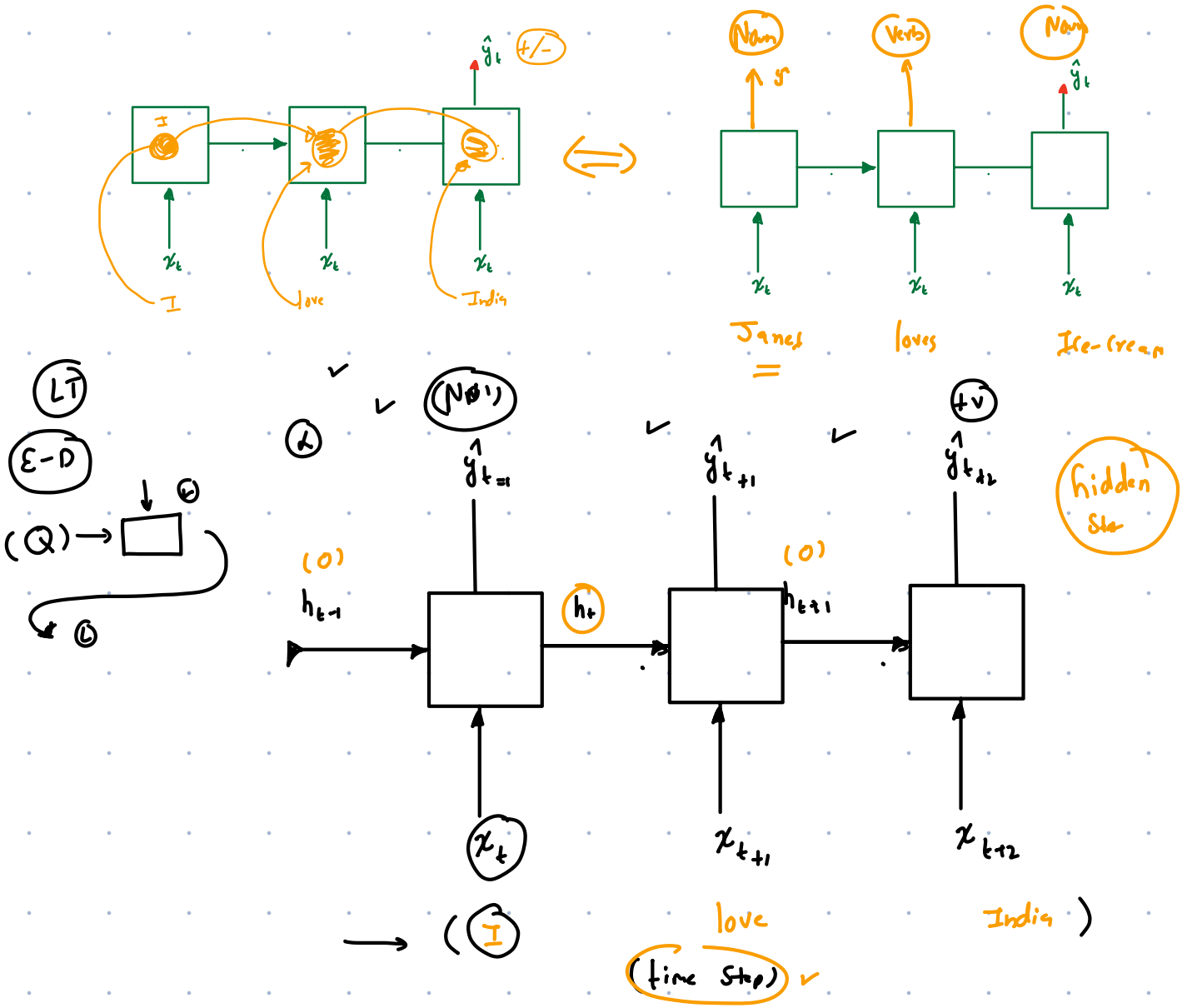
I/p I/p

Many - 2 - many

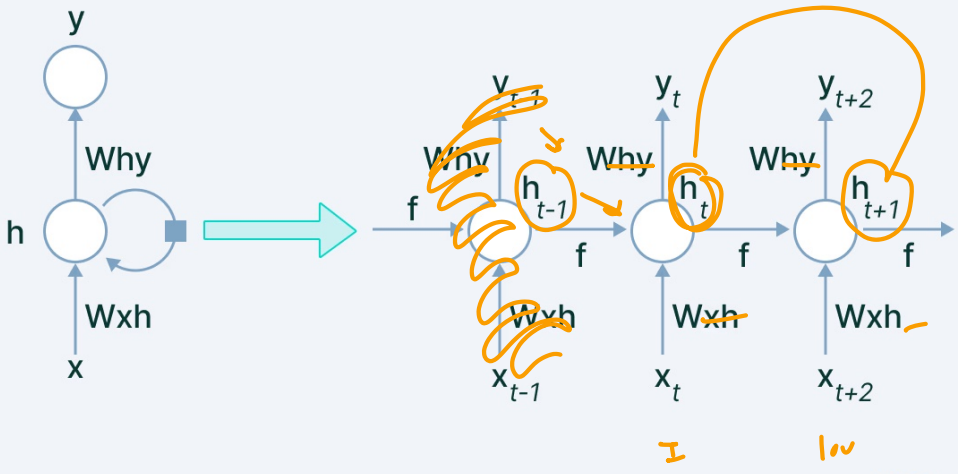
"pm modi has widrawn from Quad and decided not to attend the summit dm" ⇒ [pm modi, Quad]

(many)

(Sentiment Analysis)



The Recurrent Neural Networks (RNN)



(hidden State) (detail)

(book) = (P1 / P2 / P3) => (detailed)

(ebook) (detail)

= Hidden State = (Your mental summary book) (Short)

(Analy)

"The detective wearn a black & blue suit, heavy hair → (D − E n) = a

how h e r . n"

(Exach) = RNN = (hidden State)

(NLP) "I love India"

['I', 'love', 'India'] = (tokenisation) (Computer)

(word En)

[x x x x] [x x x x] [x x x x]

(Word Vec) = (number)

3d (-∞, ∞) 0-10

Arif = Profe	-	9	9	2
Subh Comm	-	7	6	1
Rahul Perso	-	5	5.5	3

- Arif = [9, 7, 5]
- Sub = [9, 6, 5.5]
- Rahul : [2, 1, 3]

(Word Embedding) = (Word Vec) Im

5d Not much

Arif = [9, 7, 5, 8, 9]

- Profe - 9
- Comm = 7
- Perso = 5
- Edu = 8
- Empt = 9

↑d ↑ Perso

(8B > 100B)

(Billions)

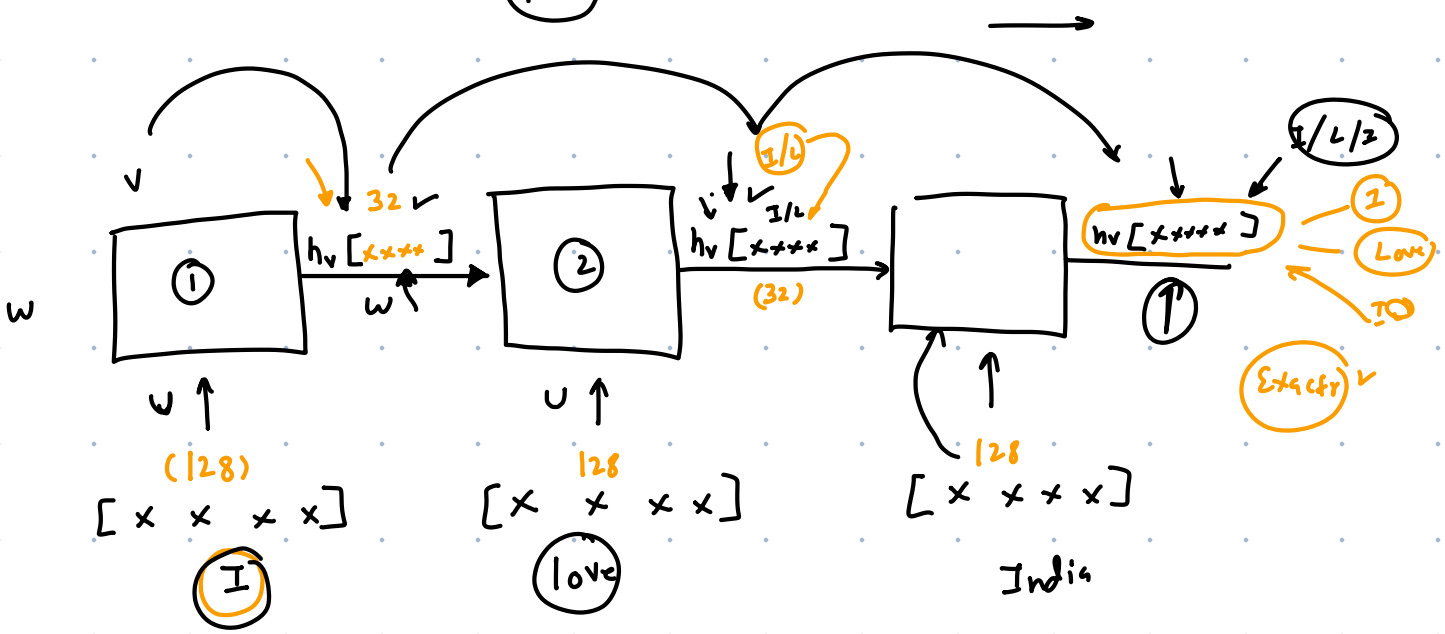
Own = N

(10 / 200)

RNN

(word vec) → (dimen)

idea

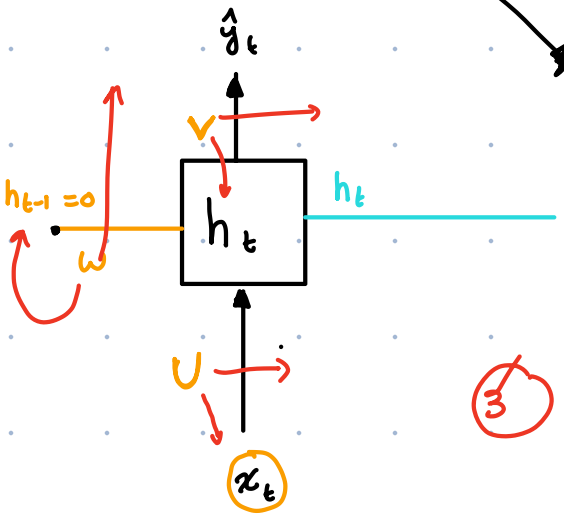


32 32 32

[ ] [ ] [ ]

(hidden layer / hidden state) NO

(back) = BP U = weights



(math) (f = Activ)

$$h_t = f(U \cdot x_t + W \cdot h_{t-1})$$

At

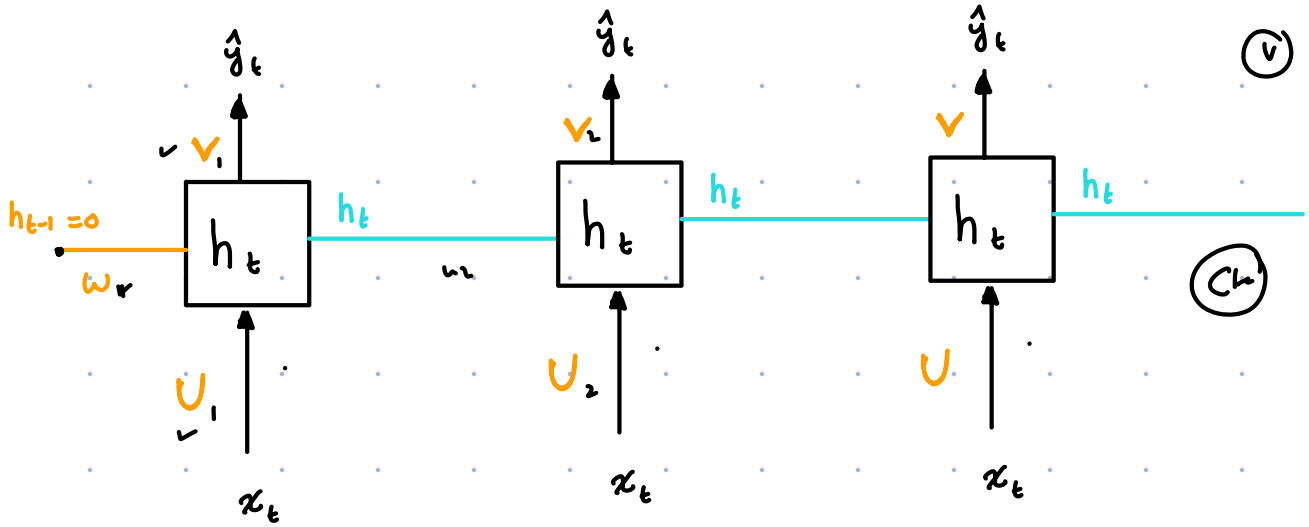
$$\hat{y}_t = g(h_t \cdot V)$$

At

logic

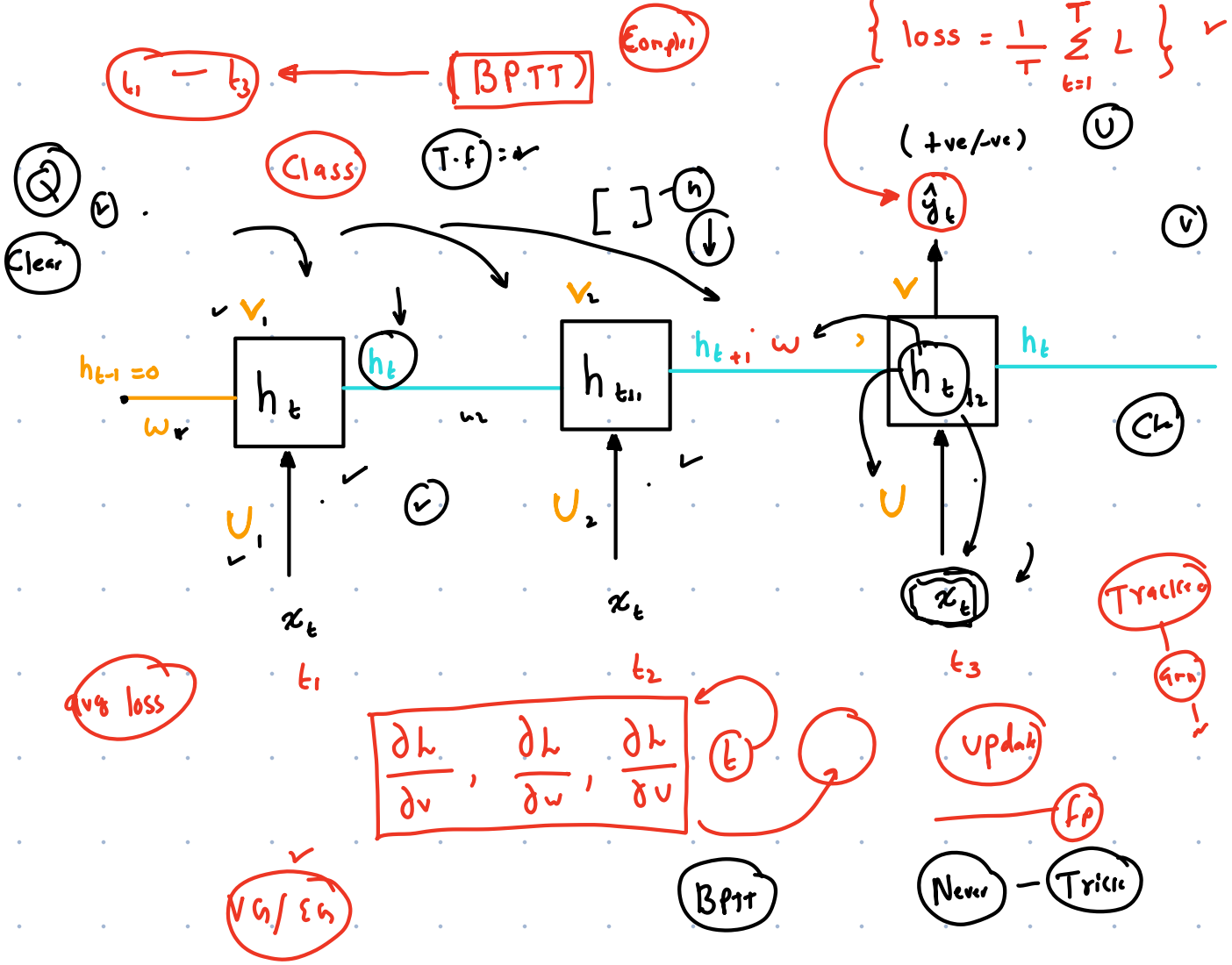
3 (U - V - W)

BPTT



No (U V W)

Same J S C Q → RNN



(NLP) = (RNN) = Q = Explain

3 time steps ⇐ I love india

I love india so much

I = (len) = (time step)

I love ice cream

It is really cold in india

I feel better now

RNN = 1

Cell - 1

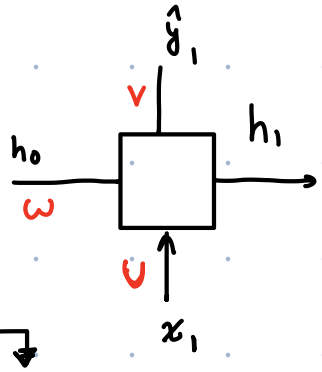
RNN = (h) = (h)

(RNN) [1, 2] → 3

DPTT

Time Step $t=1$

- i/p $= x_1$
- $h_0 = 0$



(S)

(t1)

$h_1 = \tanh(U \cdot x_1 + W \cdot h_0)$

$= \tanh(0.5 \cdot 1 + 0.4 \cdot 0)$

$= \tanh(0.5) = 0.46$

$\hat{y}_1 = V \cdot h_1$

$= 0.3\,(0.46)$

$= 0.138$

actual_value $= y_1 = 2$

$L_1 = \frac{1}{2}(\hat{y}_1 - y_1)^2$ (1 - 2)

$= \frac{1}{2}(0.138 - 2)^2$

$= \frac{1}{2}(-1.862)^2$

$L_1 = 1.73$

1 → 2

2 → 3

t_2

$x_2 = 2$

$h_1 = 0.46$

$h_2 = \tanh(U \cdot x_2 + W \cdot h_1)$

$= \tanh(0.5 \cdot 2 + 0.4 \cdot 0.46)$

$= \tanh(1.184)$

$= 0.83$

$\hat{y}_2 = V \cdot h_2$

$= 0.3 \cdot 0.83$

$= 0.249$

$L_2 = \frac{1}{2}(y_2 - y_2)^2$ $\hat{y}_2 =$ (2 - 3)

$= \frac{1}{2}(0.249 - 3)^2$

$= \frac{1}{2}(2.751)^2$

$= 3.78$

Total_loss $= (L_1 + L_2)/2 = (1.73 + 3.78)/2 = 2.76$

loss

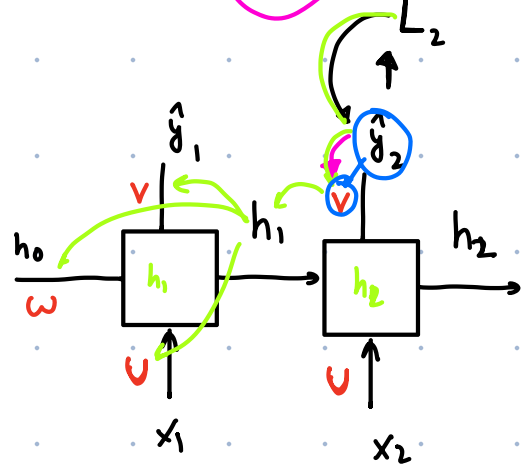
(BPTT)

$t=2$

$(\tanh)' = 1 - h_{cell}(x)^2$

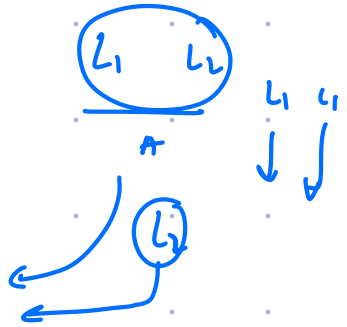
Step2: Gradient wrt Prediction

$\frac{\partial L_2}{\partial \hat{y}_2} = \hat{y}_2 - y_2 = 0.249 - 3 = -2.751$



Step2: Gradient wrt V

$\frac{\partial L_2}{\partial V}$ | $\frac{\partial L_2}{\partial h_2}$

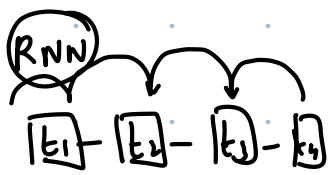


$\frac{\partial L_1}{\partial \hat{y}_1} =$ Best Small = Code Yousuf Unbeatable

(RNN) = (time bomb)

(Explode) (Vanish)

RNN = Long Sequence Vanish Gradient



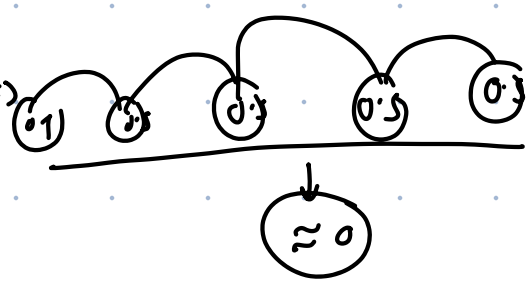
≈ 0.00000089

$= 0$

(RNN) = (long Sequence) = (gradient)

≈ 0 => Update = No Change

Sm =>



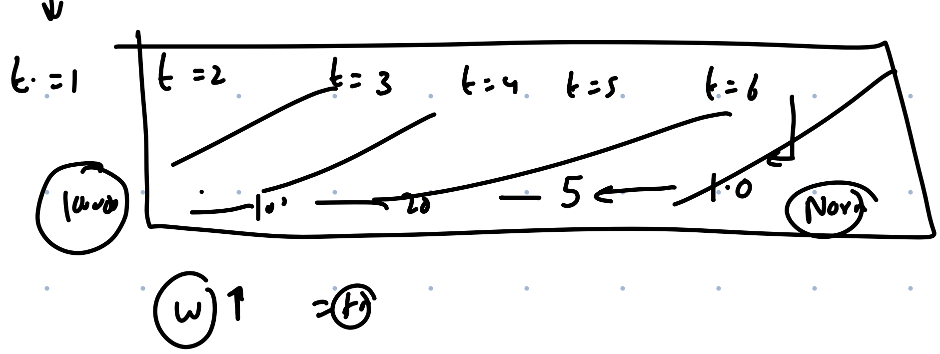
$t=1$ $t=2$ $t=3$ $t=4$ $t=5$ $t=6$

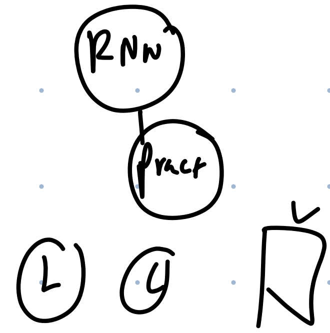
Small

$0.0001 \leftarrow 0.01 \leftarrow 0.1 \leftarrow 0.5 \leftarrow 0.8 \leftarrow 1.0$ Normal

RNN Learn

Term = L





(w) ↑ = (t)

$$\frac{\partial Loss}{\partial w_1} = \frac{\partial Loss}{h_3} \times \frac{\partial h_3}{\partial h_2} \times \frac{\partial h_2}{\partial h_1} \times \frac{\partial h_1}{\partial w_1}$$

v $w > 1$ → (2800)

(You Contro) $\frac{P}{CB}$ = (data) / P / (c)

(c) – (NN) = n (NN) (Problem)

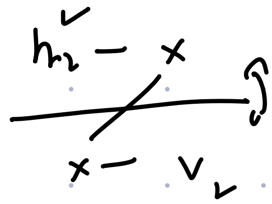
$\hat{y}_2 = ((V)(h_2))$ ↓ h_2 ↓

(0.0001 × 100) => ↑

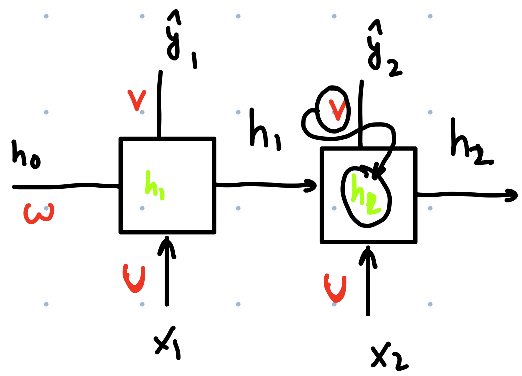
(3 × 100) = 300

$(h_2 - v)$ (direct connect)

(d) α

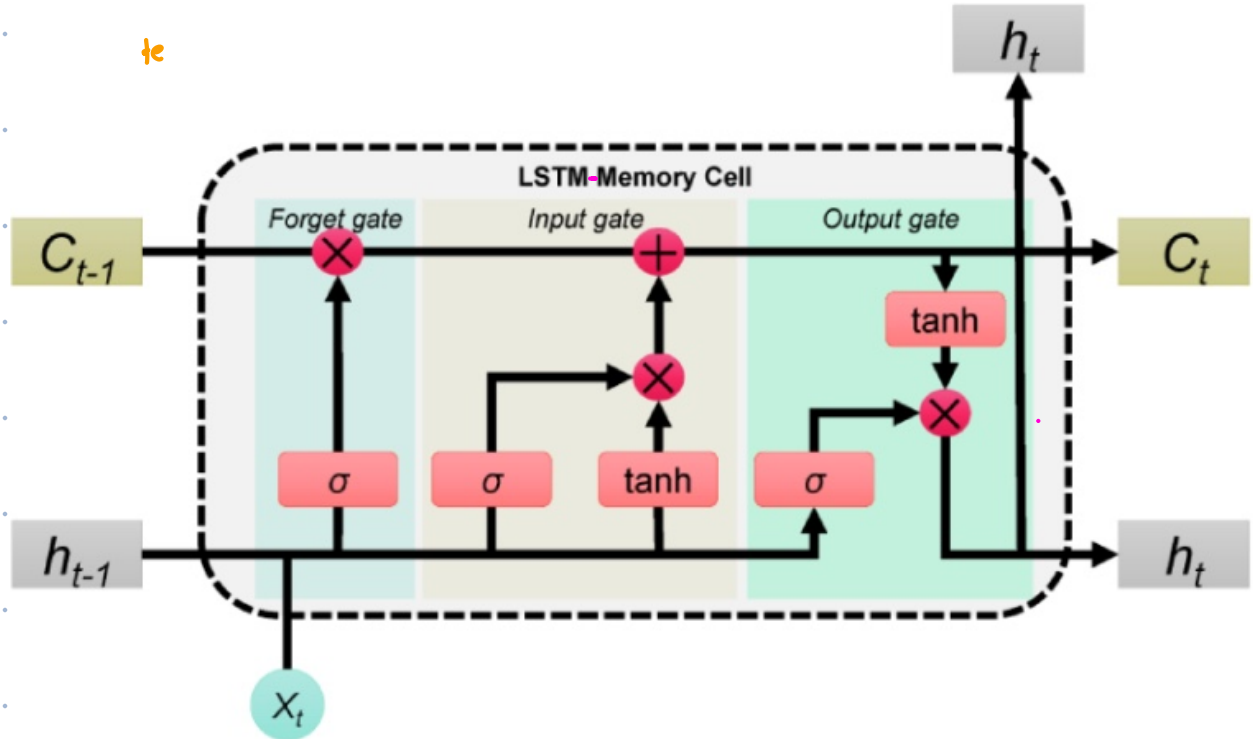


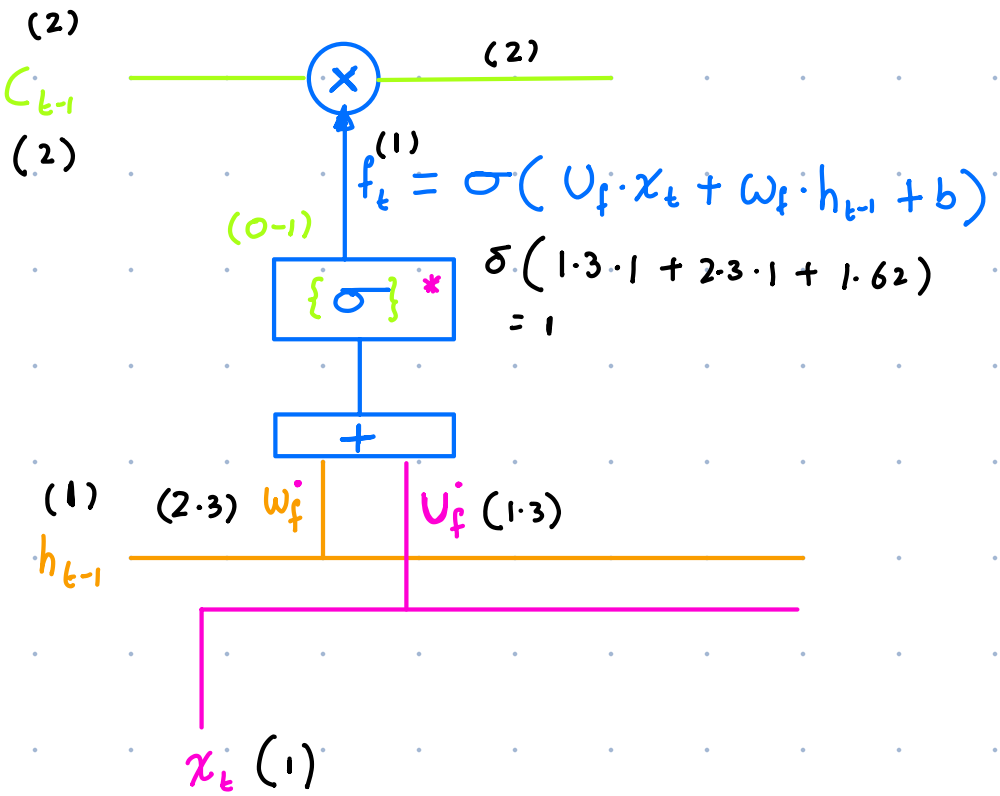
$(V \cdot h_2) =$



(100%)

(Long-Short Term memory)





(Forget gate) = Filter

•) What information from the past to KEEP

•) What information from the past to THROW

((forget) - irrelevant info)

S. A:

[i bought this phine last month.][the camera quality was amazing.]
the[battery life was good.][However, it stopped working yesterday]

Memory: [b, P, L, n] =f → (keep)

m = [b, P, ~~L~~, ~~n~~, A] = ⊕ FC

= [b, P, A] = ⊕

(b, P, A, C, L)

[deb, phone, red, pair, rain]

(0.2) (0.9) (0.7)

(Text) (Signal) (Sigmoid)

(Signal)

(State) = (New)

New 80% → (Old)

St = low. →

:)

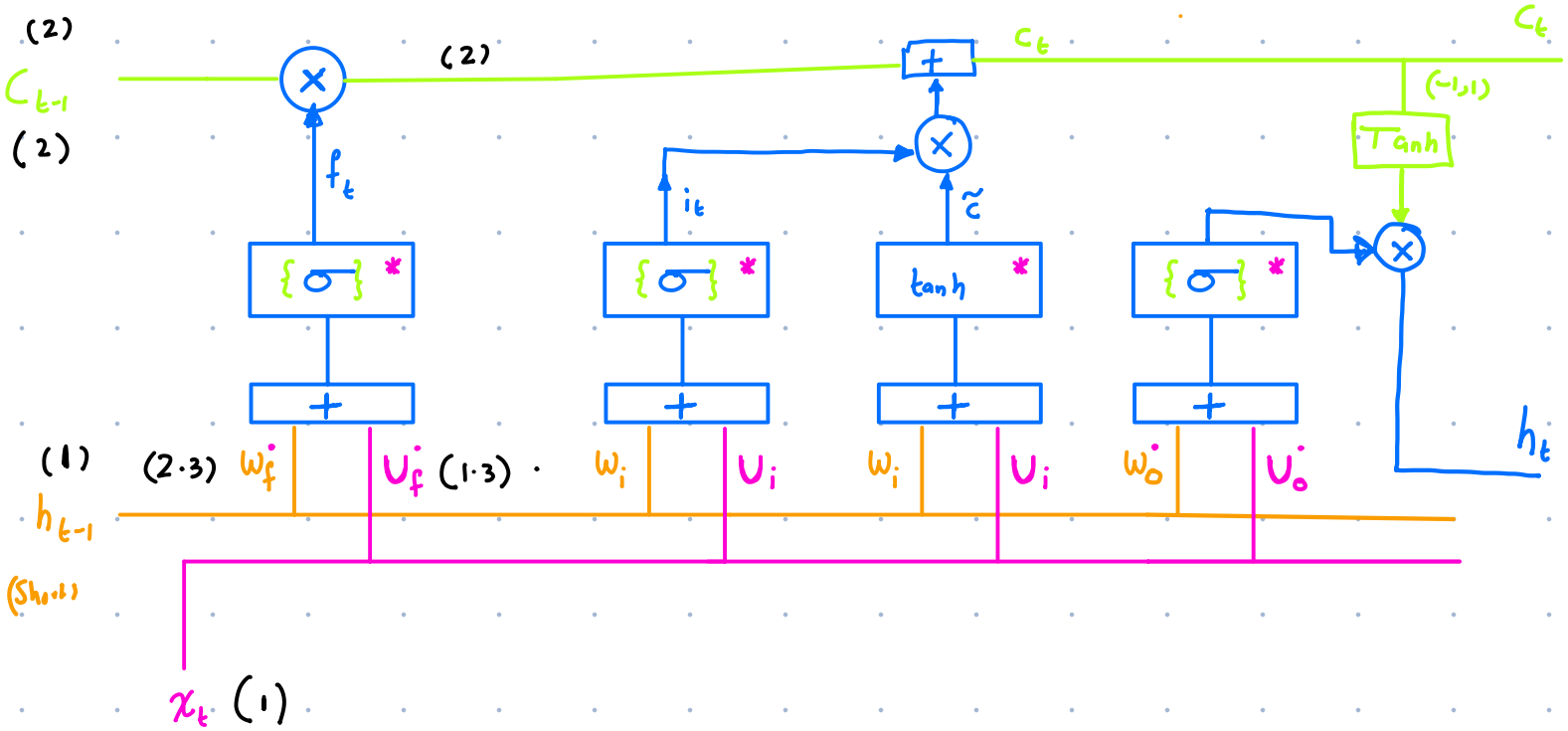
$i_t = \sigma(U_i \cdot x_t + W_i h_{t-1})$

(RNN | Trans)

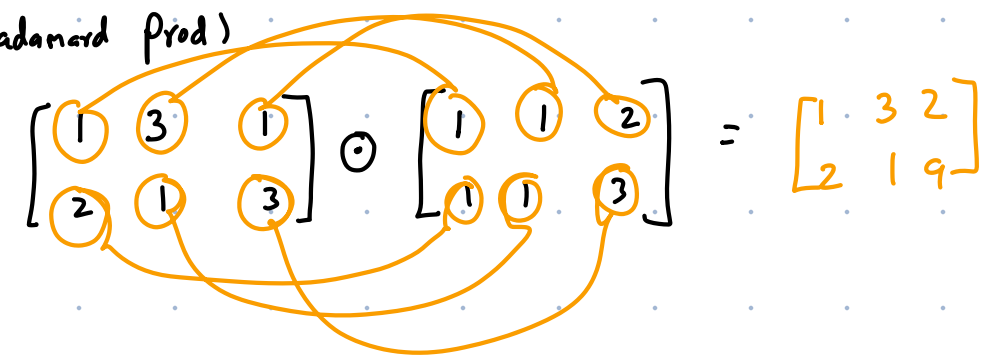
(basis)

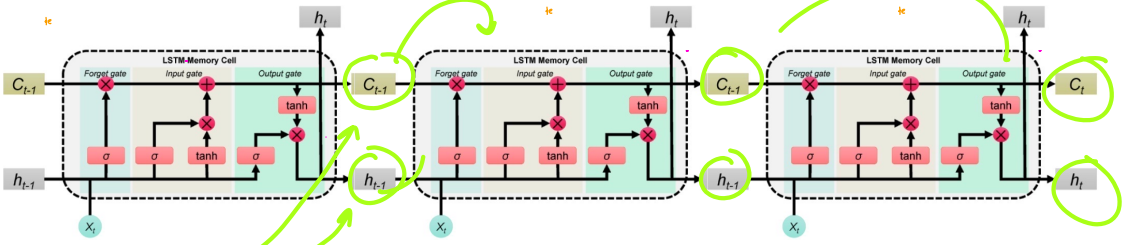
(long term) (Cell)

$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}$

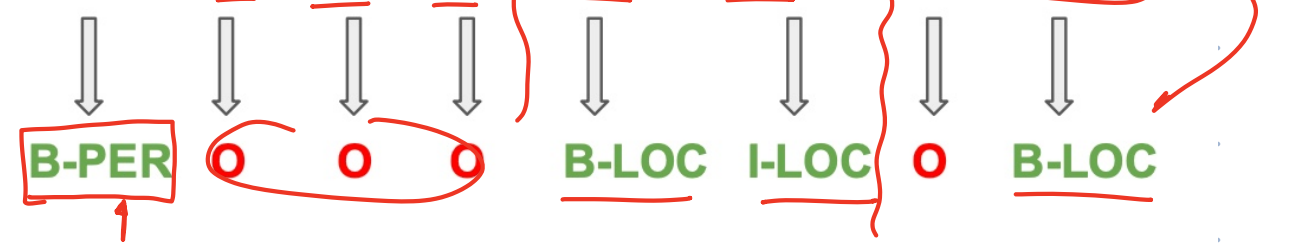


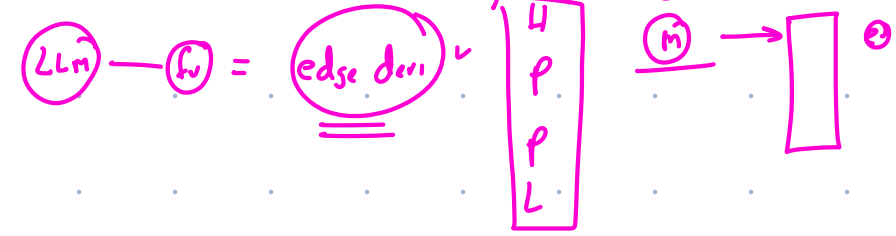
(Hadamard Prod)





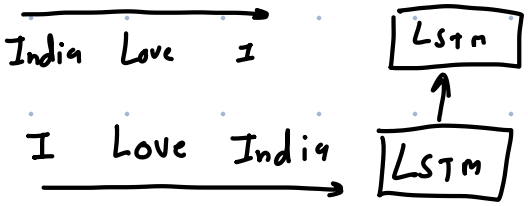
contentSkip to site indexPoliticsSubscribeLog InSubscribeLog InToday's PaperAdvertisementSupported ORG byF.B.I. Agent Peter Strzok PERSON , Who Criticized Trump PERSON in Texts, Is FiredImagePeter Strzok, a top F.B.I. GPE counterintelligence agent who was taken off the special counsel investigation after his disparaging texts about President Trump PERSON were uncovered, was fired. CreditT.J. Kirkpatrick PERSON for The New York TimesBy Adam Goldman ORG and Michael S. SchmidtAug PERSON . 13 CARDINAL , 2018WASHINGTON CARDINAL — Peter Strzok PERSON , the F.B.I. GPE senior counterintelligence agent who disparaged President Trump PERSON in inflammatory text messages and helped oversee the Hillary Clinton PERSON email and Russia GPE investigations, has been fired for violating bureau policies, Mr. Strzok PERSON 's lawyer said Monday DATE .Mr. Trump and his allies seized on the texts — exchanged during the 2016 DATE campaign with a former F.B.I. GPE lawyer, Lisa Page — in PERSON assailing the Russia GPE investigation as an illegitimate "witch hunt." Mr. Strzok PERSON , who rose over 20 years DATE at the F.B.I. GPE to become one of its most experienced counterintelligence agents, was a key figure in the early months DATE of the inquiry.Along with writing the texts, Mr. Strzok PERSON was accused of sending a highly sensitive search warrant to his personal email account.The F.B.I. GPE had been under immense political pressure by Mr. Trump PERSON to dismiss Mr. Strzok PERSON , who was removed last summer DATE from the staff of the special counsel, Robert S. Mueller III PERSON . The president has repeatedly denounced Mr. Strzok PERSON in posts on





(RNN - LSTM)

(Bi-LSTM)



Why