INTERNAL ASSESSMENT ANSWER BOOKLET

UG / PG

ಗ್ಲೋಬಲ್ ಅಕ್ಯಾಡೆಮಿ ಆಫ್ ಟೆಕ್ನಾಲಜಿ, ಬೆಂಗಳೂರು

Global Academy of Technology, Bengaluru

NAAC A

TEST – I ○	TEST – II ●	TEST - III ○	Script No. 310224

Date	21/02/24	Semester	5th	USN	1GA21A1006

Subject Name	Deep learning principles & practices	Subject Code	21AML53

I here with abide by the rules and regulations of the Institute.

AISHWARYA·H·M

Aishwarya·H·M

Name and Signature of the Candidate

I have verified the data fillied by the candidate

LP

Room Superintendent's Signature

# of Graph Sheets / Drawing Sheets Attached

Internal Assessment Marks

	a	b	c	Total
1	10			10
2				
3	5	5		10
4				
5				

	a	b	c	Total
6	5	5		10
7	10			10
8				
9				
10				

Total Marks obtained	Out of 40 Marks	40

Very good

LP 21/2/24

Name and Signature of Faculty with Date

AISHWARYA·H·M

Aishwarya·H·M 23/2/24

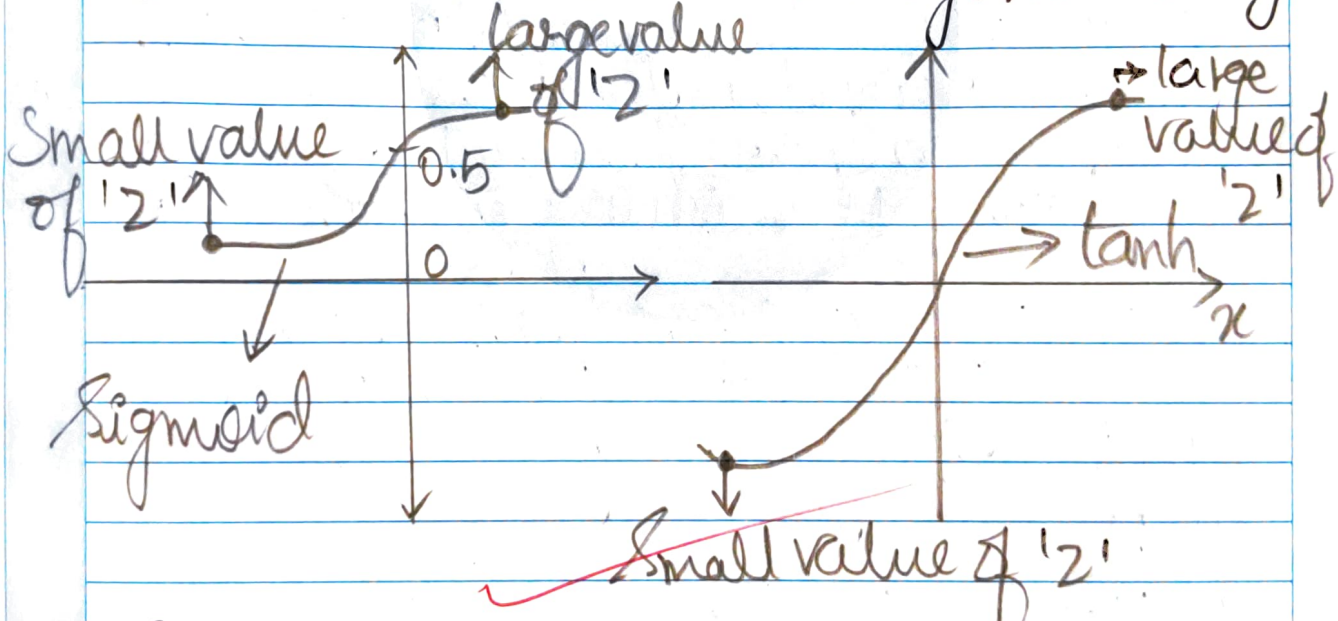
Name and Signature of the Student with Date

Q. Nos.

1) The weight initializations can be made through by two techniques using random values:

a] Normal Distribution
b] Uniform Distribution

⇒ The weights when large values are initialized, the z value increases or it is large.
⇒ The learning rate also takes huge/more amount of time.
⇒ We can observe these changes or in the sigmoid/tanh graphs, where the slope of the graph (gradient descent decreases and and learning rate is high



⇒ The above diagram depicts the sigmoid/tanh graphs.
⇒ As in the same case if the value of Weights initialized are small, in this case also we get 'z' value as small

and learning rate also takes large amount of time.

⇒ As we propagate back from output layers, the weights are multiplied across each neuron in the hiddenlayers.

⇒ Let us Assume 15 Hidden layers, the weight matrix is multiplied across all layers. And we get very large number in the earlier layers. This creates a exploding gradient descent.

⇒ If the small values of 'z' are fixed the slope of the graph decreases progressively, and weight matrix is updated slowly. This leads to the vanishing gradient descent.

⇒ We came across these challenges and got two methods of initializing namely:

1) He
2) Xavier Glorot.

⇒ He: Initialization methods takes random initialization methods using normal distribution.

Q. Nos.

⇒ We have the formula: $\sqrt{\frac{2}{Size^{[l-1]}}}$

⇒ To the layers of Network, He can be written as:

$W^{[l]}$ = random_matrix * $\sqrt{\frac{2}{Size^{[l-1]}}}$

2 ⇒ Xavier initialization takes the weights for each neurons in the layers same as he, but slightly different as:

Formula: $\sqrt{\frac{1}{Size^{[l-1]}}}$, it is bit change in initialization as: $\sqrt{\frac{2}{Size^{[l-1]} + Size^{[l]}}}$

∴ The Xavier initialization for layers of Network can be written as:

10 $W^{[l]}$ = random_matrix * $\sqrt{\frac{2}{Size^{[l-1]} + Size^{[l]}}}$

⇒ The potential challenges were overcomed due to some of the techniques.

⇒ The initialization of cost function also played crucial role in weight initialization where if cost value of z is 0, the cost function would not

Q. Nos.

converge.

⇒ If z was very small and the cost function converged in beginning ~~and diverged later~~.

⇒ If z was very large the cost function would converge and diverge.

⇒ If z was probable estimated the learning rate was at pace and cost function converges.

⇒ Also due to multiple neurons redundancy values, we cannot initialize weight to '0', but bias can be '0'.

⇒ So the neurons could be symmetric.

Q. Nos.

3] a] Traditional ANN to deal with image Dataset.

⇒ The very popular MNIST dataset was used in Deep neural Network.
⇒ The traditional ANN could classify the MNIST dataset pretty good.
⇒ The MNIST dataset is very specific where images are uniformly spaced and aligned to center.
⇒ In real world we do not get image dataset as these specific.
⇒ The traditional ANN could not highlight the most trainable features
⇒ The ANN took very long time and memory space to calculate these features
⇒ It did not consciess the spatial features.
⇒ ANN could not deeply identify the image location features dimensions.
⇒ CNN architecture has the core components to address the spatial features.
⇒ The above challenges where faced with also respect to
* Overfitting
* Spatial features
* Dimensionality reductions.

Q. Nos.

⇒ The Building Blocks of CNN are:

1. Kernel / filter
2. Convolution layer
3. Pooling layer.

1⇒ Kernel: It is a matrix of special type applied to image to extract features and underlying patterns of objects in images.

⇒ Kernel/filter can be of many varities so we can identify different types of patterns.

Eg: Sobel filter which identifies vertical and Horizontal edges in images.

* Outline filter which identifies or outlines the prominent features.

2. Convolution layer:

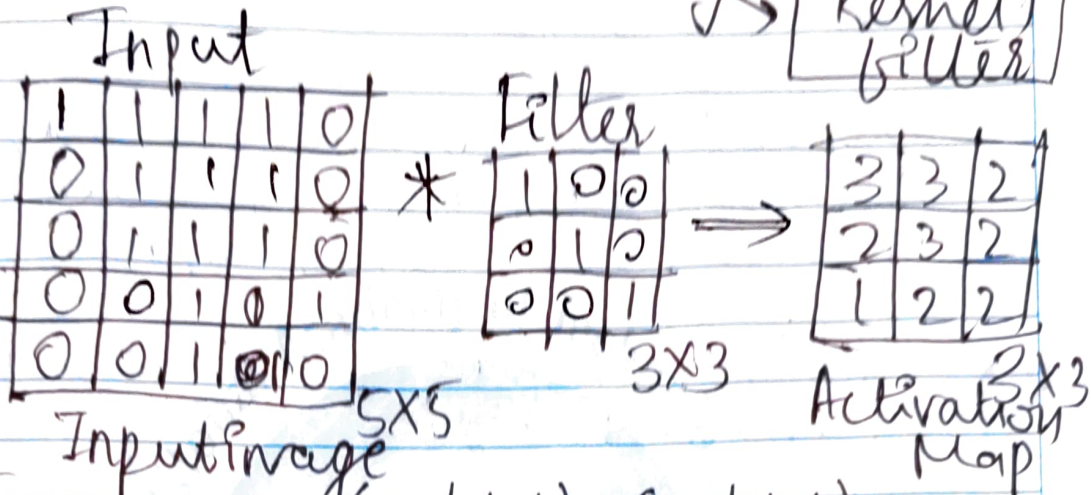
⇒ The convolution layer is identified as the important layer to classify the different challenges.

⇒ CNN is a special type of Neural Network which uses convolution instead of

Simple matrix multiplication.

⇒ Two arguments in CNN: Arg 1 → Image; Arg 2 → kernel / filter

Ex:



Formula: $(n-f+1) \times (n-f+1)$

⇒ In convolution we also have two important factors stride and padding.

⇒ Stride is the pixels moved by kernel around image subset

⇒ If it is moved by 1 pixel then strides is 1.

⇒ Padding is used to reduce the input image dimensions when it is processed to output.

⇒ Padding is nothing adding dummy pixels around the image subset of value 0.

⇒ $\left(\frac{n-f+2p+1}{s}\right) \times \left(\frac{n-f+2p+1}{s}\right)$

Q. Nos.

⇒ Padding is of two types → Valid, → Same

⇒ Valid padding is no padding is done.

⇒ Same padding is the zero padding where input image data dimension and output image data are same.

⇒ If filter f is odd, padding is symmetric.

⇒ If filter f is even padding should be asymmetric.

3) Pooling layer:

⇒ Pooling layer is mainly used to deal with trainable parameters.

⇒ This layer also downsamples the input image dimensions.

⇒ The pooling layer also has the important towards spatial feature representation.

⇒ There are three types of pooling
* Max pooling.

* Average pooling
* sum pooling.

⇒ All the above specified building blocks overall form the CNN networks.

⇒ where at last we have fully connected layer that finalizes the output to output layer.

b) Image of size 50X50
Filter size 3X3
Assumed same padding,

$$p = \frac{f-1}{2}$$

$$p = \frac{3-1}{2} = \frac{2}{2} = 1$$

∴ With padding

$$\left(\frac{n-f+2p+1}{s}\right) X \left(\frac{n-f+2p+1}{s}\right)$$

$$\Rightarrow \left(\frac{50-3+2+1}{1}\right) X \left(\frac{50-3+2+1}{1}\right)$$

$$\Rightarrow (50X50)$$

⇒ As we assumed same pad, where padding = 1 ∴ (51X51)

Q. Nos.

6) a] VGG16:

⇒ The VGG16 is the one of the most CNN model built across various types of layers.

⇒ VGG16 has 3×3 convolutional filters, where the model was used to train number of pixel images.

⇒ VGG16 has the image, with max layers and the softmax activation function.

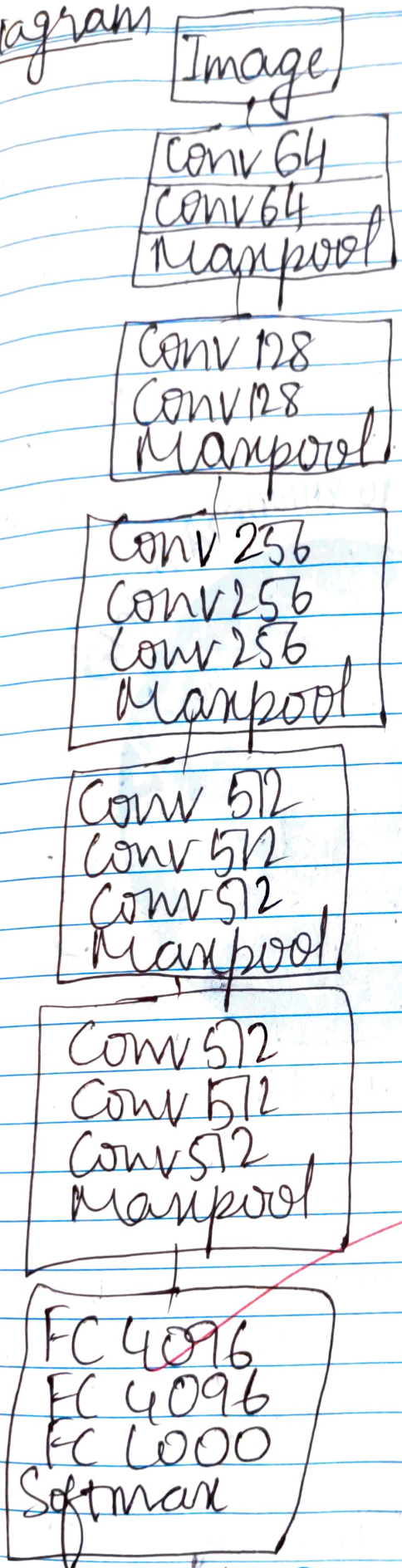
⇒ VGG16 was very accurate and convergence to the input computations, features and resolutions.

→ This Model was Applied to the small trainable tasks, for computations.

⇒ The Model was so accurate with accuracy and characteristical features.

⇒ This model was used to due to its simplicity and its higher memory complications.

Diagram



Applications:
→ Healthcare
→ Manufacturing Units
→ Traffic signs

VGG 16 Architectural Diagram

Q. Nos.

(b) GoogleNet:

⇒ The GoogleNet has an advanced full connected layers to deeply generalize the model of DNN.

⇒ It was introduced in 2014 and developed the Google's company to advance the enhancement of image classifications.

⇒ The images daily 3 billions are out through various medias.

⇒ As this model has Input layer many hidden layers to embark the features.

⇒ The Deep NN model is very prominent in GoogleNet.

⇒ The dimensionality Reduction and spatial features are mainly focused here.

⇒ The computations of the computer vision images are more accurated compared to earlier LeNet, Resnet.

⇒ The GoogleNet Architectural

pattern use the Softmax and Relu activation function.

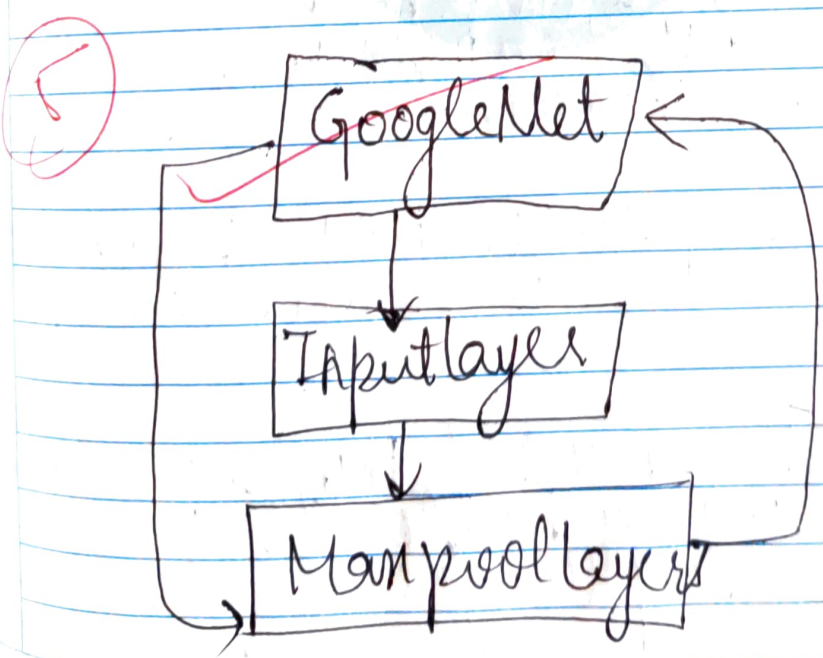
→ The GoogleNet has a fundamental deep learning approaches and a high level model to solve real-world problems.

Application:

→ Traffic Recognition

→ Sentiment Analysis

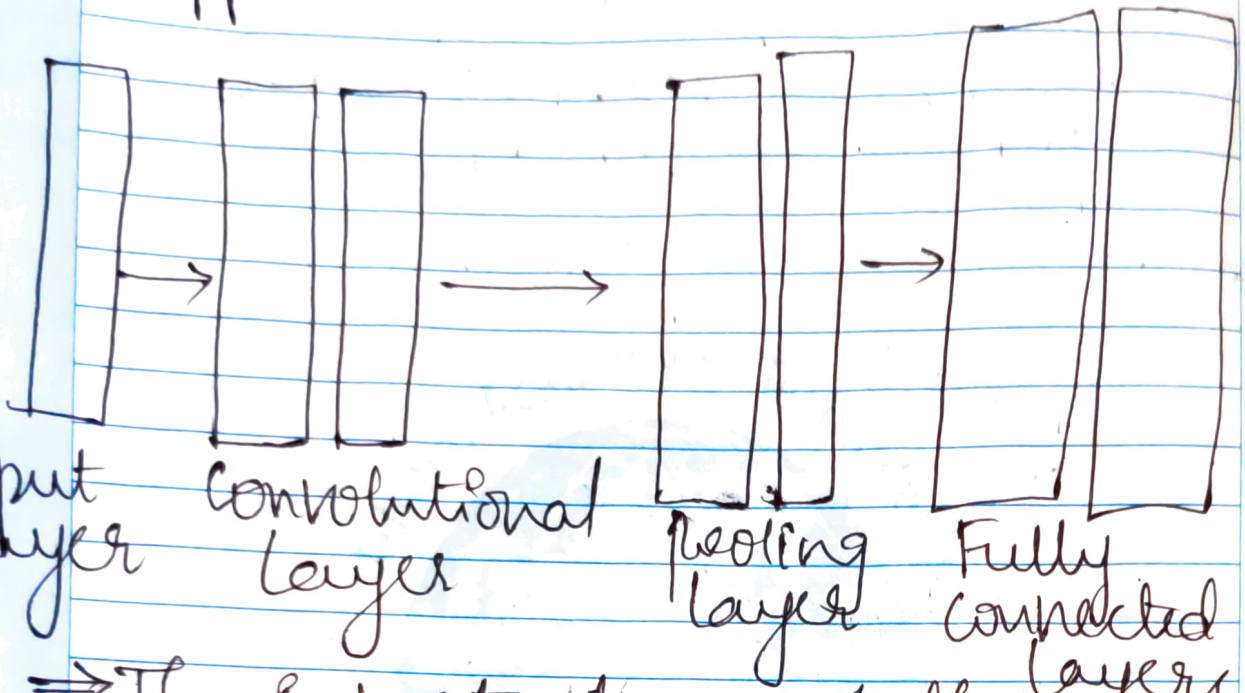
→ Language translation using deep learning techniques.



Architectural design

Q. Nos.

7) CNN architecture tailored for a specific computer vision Applications are:



⇒ The important area of the deep learning networks influence the computer vision Application.

⇒ The deep neural network has tailored the feature among the computer vision.

⇒ The computer vision is yielding better results than human vision.

⇒ The Alignment of Human phalotapy is analyzed better than Human Opthogolamagist.

→ The huge amount of data generated are analyzed by computer vision we leveraged its importance.

→ The computer vision has the major Applications:

* Healthcare:
→ ECG
→ PET scan
→ CT scan
→ MRI's

* Manufacturing:
→ Reading Barcodes
→ Manufacturing of the items in correct specific brange.
→ Placing right order of items in correct place with right quantity

* Agriculture:
→ The first Computer Vision that was adopted in Agriculture was Taarel that was Taprospera company.
→ It focused on the crops growth and Diseases.
→ Sorting of the right crops produced

Q. Nos.

→ The Design choices and Layers:

* Input layer:

→ Where the image of pixels like 5x5, 7x7 — are given to a CNN deep neural network.

→ Convolutional layer:

* This layer convolutes the input image by adding kernel or filter.

→ This layer of convolution has a pixels all around boundary in the smaller numbers.

→ So the padding is involved to as the padded pixel cannot determine the o/p image matrix.

→ The CNN convolutes the image pixels rather simply doing matrix multiplication.

→ Pooling layer:

* This layer downsamples the image dimensions also the loss of information is connected.

→ The pooling layers extracts the major spatial features of representation.

→ The pooling layer has the input given:

nh x nw x nc, with stride 's'.

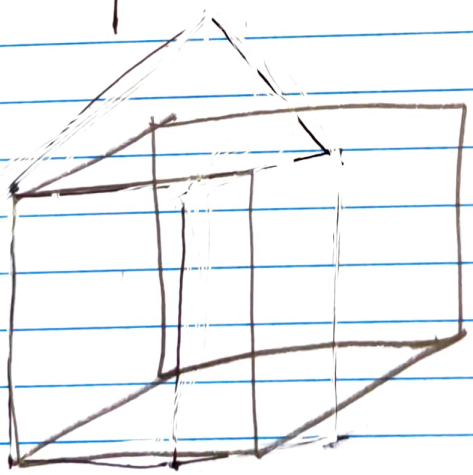
→ The fully connected layer, processing the last input kernels, gives the output matrix of image of pixels of same dimensions of input or downsampling.

→ The computer vision has the range of pixels varing 0 to 255

→ '0' is Black, 255 is white.

→ We also have RGB pixels which takes range between 0 to 255.

→ The Computer vision has:



(10)

where 'n' is width, 'y' is length '2' is channel of colors.