Name :- Subhankar Dutta

Sec :- J

Year :- 3rd

Stream :- CSE

C. Roll :- 68

E. Roll :- 12019009001207.

Subject :- Deep Learning

Subject code :- PEC - CS50

8 i@ . A special class of this problem it called a seq2seq for modeling problem where input as well as output are sequence

In sequence learning the two problems FCNN & CNN do not hold as the output at any time step depends on previous step i/p & len of the input is not fixed

(b) The recurrent neural network however is able to remember those character because of its internal memory it produces output copies that output & loops it back into the network.

1(i)

CNN	RNN
* CNN is more Powerfull the RNN	* RNN is contain less feature compair to CNN
* CNN are ideals for images and video Processing	* RNN are ideal for text and Spaches analysis.

(ii) LSTM in Artificial RNN arch used in the field of deep Learning. Unlike Standard feed forward neural Network LSTM has feedback connections it can not only Process single data point but also entire seuend

(iv) (i) RNN connot be stackup
(ii) Slow and Complex traning

5

(5Aii) Regularization is needed as this technique discourages learning more complex or flexible model so, as avoide the risk of overfitting

(5Aii) Dropout is a tech where randomly selected neurons are ignored during Training. They are dropped-out randomly. That means. They are contributing no activation

~~2) Model such as Learning Rate~~ shorter of smaller

2) find S = 1 ; Padding = P = 1

$$w_2 = \frac{w_1 - F + 2P}{S} + 1$$

$$\frac{32 - 5 + 2}{1} = 29 + 1 = 30$$

$$H_2 = \frac{H_1 - F + 2P}{S} + 1$$

$$= 30$$

output of Dimension $H_1 \times W_1 \times D_1$

$$30 \times 30 \times 6 =$$

Stride = 1, Padding = 1

$$W_2 = \frac{W_1 - F + 2P}{S} + 1$$

$$\frac{16 - 3 + 2}{1} + 1 = 16$$

$$H_2 = \frac{16 - 3 + 2}{1} + 1 = 16$$

$$H_2 \times W_2 \times S$$

$$= 16 \times 16 \times 8$$

6 B Input:

No. of Parameter =

Convolution

No. of Parameter in a convolution

$(C_{in} \times K \times K) + 1) \times K$

$26 \times 6 = 156$

Max Pool → No. of Parameter =

Convolutions = $((5 \times 5 \times 6) + 1) \times 16$

$= 2416$

Total $2416 + 6 + 116 + 10 = 2572$

(11 6.i) Back propagation is ~~the~~ used to train the neural network of the Chain rule method in Simple terms after each feed-forward Pass to adjust the model's Parameter based on weight and biased

(1.6ii) momentum is an extention to the gradient Descent Optimization algo. that allow the Search ~~pass~~ oscillations of noisy gradient and could ~~not~~ across flat

7 II) Need of Auto Encoder:

* Dimensional reduction
* use for detecting anomalies and for reducing data
* To deep arch to learn complex relationships.