Ramakrishna Mission Vivekananda Educational & Research Institute

Belur Math, Howrah, West Bengal

School of Mathematical Sciences, Department of Computer Science

End Semester Examination - **CS411: Applications of Computer Vision and Deep Learning**

M.Sc. Computer Science and Big Data Analytics Date: 18 May 2024

Time: **Part 1 (MB-207)** 11:00 AM to 01:00 PM & **Part 2 (Arya lab)** - 2:00 PM to 6:00 PM

Instructor: Jimut Bahan Pal Max marks: 100

Instructions: **Please answer all the questions**. The question set is divided into two parts, i.e., Part (i) Theory (40 marks) and Part (ii) Application (60 marks). The theory questions are divided into MCQs where only one answer is correct; **any wrong answer will fetch a -ve mark of 1**, MSQs where multiple answers might be correct, **selecting any wrong answer will fetch a -ve mark of 2**, and lastly, subjective questions, where you are supposed to answer descriptively in page; **no -ve marks for this part**. The Application question is divided into multiple interrelated parts, where you solve a regression problem using Python and Pytorch **no -ve marks for this part**. You are allowed a maximum of 2 hours for Part (i) and a maximum of 4 hours for Part (ii). Part (i) **is open notes (no internet is allowed)**, and Part (ii) **is open internet (no chatting is allowed)**. Leaving the hall before 1 hour is not allowed in any of the parts. Students can take a 1-hour break in the middle of the parts if they want to but, not more than 4 hours will be provided for the second part. If one completes Part (i), please start Part (ii), and never return to Part (i); this is only allowed by submitting the answer script for Part (i). Bio breaks are allowed. **Please note that a maximum of 6 hours will be allowed for the whole test, and copies should be submitted after that, no extra time will be given**. Please write the answer in the answer scripts as numbers only, like for example if the answer is 1(e), then just write 1(e) for MCQs. If the answer is 1(a) and 1(b) then write 1-(a),(b) for MSQs and write full answers for the subjective questions. Please be short and to the point for the Subjective questions. Talking and discussions are strictly prohibited, in case of doubt, please make suitable assumptions and proceed with the test. **You can take a photocopy of the answer script and send it as a PDF once you submit the scripts (for Part -1). You can send the jupyter notebook as a reply to the email, once you complete the Part-2. Please send it as a reply to the final instructions email that was sent to you. Don't panic... All the best!**

Part - 1 (Theory - MCQ, MSQ and Subjective) — 40 Marks

Multiple Choice Questions (MCQs) — 10 Marks

(Only open notes allowed; no internet allowed) Here only one answer is correct; **any wrong answer will fetch a -ve mark of 1**. Please answer only if you are over 90% confident that this will be the answer!

1. We have a PRIMITIVE retrieval augmented question answering system that works as follows. The questions are all MCQs with M options for each answer. There is a database of questions and corresponding answers, which has been created over some time. The QA system is trained to either retrieve an answer or to "guess" it (there is no option other than these two modes of answering). When a question is input, the QA system refers to the database of previously answered questions with probability P and retrieves the answer. When "guessing", the probability that the QA system will be successful is $1/M$. (2)

   What is the conditional probability that the answer to a question is a RETRIEVED one, given that the answer is correct? Let M be 4 and P be 0.5. Choose from the below options:

(a) 5/6

(b) 4/5

(c) 3/4

(d) 2/3

2. Choose the weight change expression for the weight w_1 as in the figure below from the options given. Assume the loss function is total sum square and both O and H are sigmoid neurons. Let the output of O and H also be denoted as O and H. The target output is T. The learning rate is η. (2)

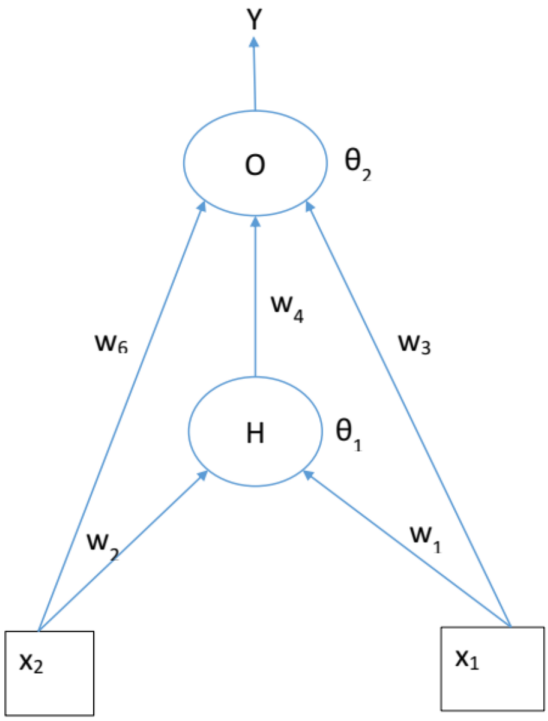


Figure of a neural network having sigmoid as activation.

(a) $\eta(T-O)O(1-O)w_3x_1$

(b) $\eta(T-O)O(1-O)w_6H(1-H)x_1$

(c) $\eta(T-O)O(1-O)H(1-H)w_4x_1$

(d) $\eta(T-O)O(1-O)H(1-H)(w_6+w_4+w_3)x_1$

3. Choose the ONE most specific entity from the list below to which all of the following apply as challenges. *Data Residency, Data Privacy, Complexities of real time information, Complexities of novel data streams, Hallucination, Lack of emotional intelligence, and Lack of empathy*. (2)

(a) LLMs

(b) GenAI

(c) Conversational AI

(d) Feedforward Network

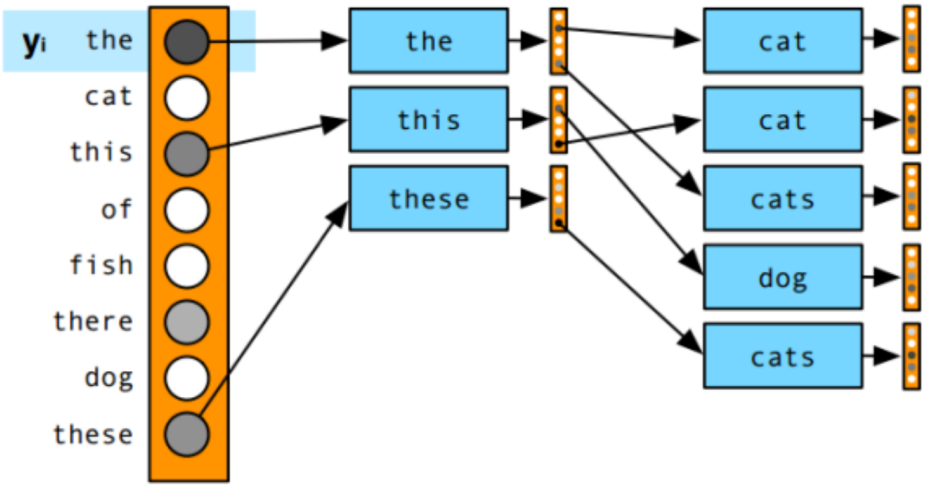
4. Which of the following best expresses the learning rule for bias terms in a feedforward neural network? η is the learning rate and δ_j is the delta value at neuron j in the network. (2)

(a) $\eta\delta_j^2$

(b) $-\eta\delta_j$

(c) $-\eta\sqrt{\delta_j}$

(d) None of the given options is correct



NMT Decoding figure.

5. The following diagram is a snapshot of what kind of search in NMT decoding? (2)
   (a) A* search
   (b) Beam Search
   (c) Metropolis search
   (d) Gibbs sampling search

Multiple Select Questions (MSQs) — 10 Marks

(Only open notes allowed; no internet allowed) MSQs where multiple answers might be correct, **selecting any wrong answer will fetch a -ve mark of 2**. Please answer only if you are over 90% confident that these will be the answer!

6. Consider a layer of neurons in a neural network. There are 5 neurons in the layer N_1, N_2, N_3, N_4 *and* N_5. The activations of the neurons for 5 data points D_1, D_2, D_3, D_4, *and* D_5 are given by the matrix below (call this *data-by-neuron matrix* as shown in figure below). (3)

	N1	N2	N3	N4	N5
D1	1	2	3	4	5
D2	6	7	8	9	10
D3	11	12	13	14	15
D4	16	17	18	19	20
D5	21	22	23	24	25

Data by neuron matrix.

That is the row D_i, gives the activations of the 5 neurons for the data instance D_i, after layer normalization, which of the following is/are true about the *data-by-neuron matrix* as shown in figure below?

(a) All rows are identical

(b) The central column (N3 col) has all 0s

(c) The row sum is equal to 0 for every row

(d) The column sum is equal to 0 for every column

7. Consider the *data-by-neuron matrix* again as shown in the previous question. Now apply batch normalization for this situation. Again identify the true statement(s) from below: (3)

   (a) All rows are identical

   (b) The central column (N3 col) has all 0's

   (c) The row sum is equal to 0 for every row

   (d) The column sum is equal to 0 for every column

8. The phrase "flow of gradients" in pure feedforward neural networks involve the following (choose the correct option(s)): (2)

   (a) The context in training of the neural net parameters

   (b) Derivatives of activation functions of neurons are involved

   (c) The term "flow" particularly means the computation of derivatives at a given layer, using the derivatives found in the immediately next "upper" layer (i.e., the layer closer to the outermost layer)

   (d) The chain rule of derivatives is invoked.

9. With respect this network, the partial derivative of O_o with respect to W_{11} and that of O_1 with respect to W_{11} are (2)

   (a) different in magnitude

   (b) same in magnitude

   (c) same in sign

   (d) different in sign

Subjective Questions — 20 Marks

(Only open notes allowed; no internet allowed) You are supposed to answer descriptively in page; **no -ve marks for this part**, write to your heart's content!

10. **The Cross Entropy Loss is never negative** - prove or disprove rigorously! (10)

11. Prove or disprove rigorously **that a single neuron with ReLU activation function cannot compute XOR**. The ReLU function has the "knee" at NET=0, and after that the slope is 1. (10)

Part - 2 (Application: Coding in Python and Pytorch using Google Colaboratory) — 60 Marks

In the following task, you need to find the percentage of COVID-19 in the Computed Tomography (CT) images. The dataset can be downloaded from the following link: `https://drive.google.com/drive/folders/1T708rSHSus-jpT1caT4AceESyJb3kj2a?usp=drive_link`. You can use gdown in the Google Colaboratory platform, please use the following code:

```
! pip install --upgrade --no-cache-dir gdown
! gdown 1QPIBqmlH7qXL14zjBAPc7AfWnRW0j2SW
! unzip -qq dataset.zip
```

Code in Pytorch framework only; you may re-use code from the assignments and codes found online. Please report everything in the unseen test set divided in the dataset folder. The dataset is divided into three parts, the train has 2138 train images, and their corresponding details are present in train.csv. The validation has 306 validation images and similarly, check val.csv for their details. The test has 609 images and the details are present in the test.csv. Please use the training dataset for training and the validation dataset for validation only. Don't combine both to decrease the Mean Squared Error (MSE) of prediction on the test set. The CSV files are structured in the following way: **image_name, covid_perc, and patient_number**. Please use the proper way to load the CSV files and image data inside the pipeline.

(a) Exploratory Data Analysis (EDA): Do a bit of data analysis about the percentage of COVID distribution present in each of the train, validation, and test sets. There will be some distribution of the COVID percentage, please plot the distribution in any one of your favorite plotting tools, like seaborn, matplotlib, etc, i.e., the x-axis of the histogram will have the percentage and the y-axis will have the count of the percentage. Bin them properly to get a good visual. Do this for the train, validation and test set. Does the data distribution seem balanced/same for all the split? **(6)**

(b) Create the DataLoader/DataGenerator using pytorch. Use only an image size of 128x128, you can choose a single channel or 3-channel image as input to the model by making proper assumptions. Use information from the CSV to integrate into the pipeline. Use proper optimizer, and use proper batch-size so that the platform doesn't crash. **(6)**

(c) Create the train, validation, and test pipeline. Use L1 loss function only. Use the proper torch.no_grad function when required, use model.eval and model.train when required. For the test pipeline, you should dump the first 10 images along with their prediction and original label of the COVID percentage in a CSV. Use proper naming in CSV by dumping name, ground truth label (i.e., percentage), and predicted percentage. **(12)**

(d) Use an accurate model structure with proper activation for the regression task. Use torch.summary to list the parameters of your model. Use visualization tool to visualize your model. **(9)**

(e) Train for at least 30 epochs, save the model, and plot the variation of L1 loss during the training and validation dataset, by dumping the history generated during training and validation in proper files and folders. Save the best model by doing proper checkpointing. Find the L1 loss in the test dataset for your best created from scratch model. **(9)**

(f) Try any pre-trained model architecture Pytorch offers (like VGG pretrained on IMAGE-NET dataset), add proper layers and activation functions, and train using this model, save the history, checkpoint in a different name. Compare which of the models performs better in the test dataset and report the final L1 loss in both the models. **(12)**

(g) Use the best model to report the test L1 loss for the whole dataset, by loading it on the fly. You can save the model in google drive. You should also be able to load any particular image from the test dataset and report the original ground truth percentage and the predicted percentage for the model. **(6)**