

Assignment: Question and Answer

Radhika Neupane

May 14, 2024

1. ML Applications:

Question: As a ML engineer, working at a fin-tech company(e.g: khalti or esewa) you are planning to implement machine learning solutions,list out some of the challenges you might face and how would you address these challenges.

Solution: As a ML engineer, if I'm working at a fin-tech company (e.g: khalti or esewa) then the challenges which I might face would be specially related to the data quality and all. There are many fin-tech companies which deal with financial data's sensitivity and all. There are rules like GDPR and PCI DSS which make sure that the data are collected, stored and used carefully. Machine Learning itself is vulnerable when it comes to negative attacks which manipulate their output, likewise these kinds of financial data are also the main target for cyberattacks. For the solution we need to focus on choosing or building the model which can exactly explain the reason behind any predictions. Along with this data quality is very important so we must focus on data cleaning techniques and pre-processing to ensure data quality. Data plays a very important role so we can even collaborate with other related companies to enrich our collected data. Lastly we should identify and mitigate the biases from the training data before feeding it to the model.

2. Training of Machine Learning:

Question: Explain the concept and use of Cost/Loss/Error function in training of machine learning. Briefly explain the error function used to train Linear Regression.

Solution: The cost function is known as the loss function or even the error function. It is one of the concepts which finds out how well a model performed by measuring the predicted outputs and actual targets. For the linear regression, which is a method in order to predict the numeric value. As I said these functions are considered as synonyms and they can be used in changeable ways. The loss function is related to the individual data for one sample in the training set whereas the cost function is the average error of n-samples in the data for the whole training data. Going deeper loss function is used to calculate for every instance in the training set and cost function is used to calculate for the entire training set. We use an error function called Mean Squared Error (MSE) which is actually used to measure the squared difference between the predicted and the actual target values within a dataset. The formula to calculate MSE is :

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

MSE = Mean Square error

n = Number of Data points

Y_i = Observed Values

\hat{Y} = Predicted Values

3. Overfitting:

Question: Explain at least two methods used to prevent overfitting in machine learning with examples.

Solution: Overfitting basically refers to the process where the ML model learns to capture the noise and memorize the training data rather than understanding it. The two ways to prevent the problem of overfitting in machine learning is data augmentation and drop out.

Data Augmentation: It is all about making more data from the one we have. In image recognition we often slightly rotate the image, flip the images and so on which falls under data augmentation. This process is all about creating the variation in the original images so that it makes our model even better. In text classification if we are training a model to identify any kinds of spam emails then using data augmentation we can rephrase the words in our spam examples and help the model to learn the patterns of the spam emails rather than memorizing it

Dropout: It deals with the process where a neuron learns to recognize certain patterns in the data we have provided. There are many cases where neurons become too dependent on others and actually lead to overfitting so this dropout helps to prevent this by randomly ignoring during training steps. For example it's like a teacher randomly picking up a few students each time rather than the whole class.

4. Neural Network Architecture:

Question: Explain the difference between a convolutional neural network (CNN) and a recurrent neural network (RNN). When would you use one over the other? Explain in brief, the challenges faced while training the deep learning models and how can they be addressed?

Solution: Convolutional Neural Network(CNN) and Recurrent Neural Network (RNN) both are the categories of machine learning models. CNN deals with the tasks like image classification, object detection and image segmentation. They consist of the convolutional layers which learn the patterns along with the features of the imputed data using the filters across the input images. Coming towards RNN it is used for tasks like modeling languages, translation and speech recognition. In RNN they capture the memory of the input and process the sequences of data. Overfitting, Data Quality and Quantity, model complexity, hyperparameter are few challenges which we might face while training the deep learning models. There are many cases where the models may memorize the training data instead of understanding it and insufficient or low-quality training data. Batch size and learning rate is also very important so choosing these right hyperparameters can also be quite challenging. In order to solve these kinds of challenges we have various techniques like dropout and early stopping which can help us from overfitting. We can also utilize the GPUs or TPUs in training to speed up the process. Data augmentation is another best solution through which we can address the challenges.

References

- Challenges and Solutions in Training Deep Learning Models.* (n.d.). Challenges and Solutions in Training Deep Learning Models. <https://www.datasciencesociety.net/challenges-and-solutions-in-training-deep-learning-models/>
- A Complete Guide to Data Augmentation.* (2022, 11 4). A Complete Guide to Data Augmentation. <https://www.datacamp.com/tutorial/complete-guide-data-augmentation>
- convolutional neural network (CNN).* (n.d.). convolutional neural network (CNN). <https://techtarget.com/searchenterpriseai/definition/convolutional-neural-network>
- Dropout in Neural Networks.* (2022, 7 5). Dropout in Neural Networks. <https://towardsdatascience.com/dropout-in-neural-networks-47a162d621d9>
- Encord. (2023, 2 3). *Mean Square Error (MSE).* Mean Square Error (MSE). <https://encord.com/glossary/mean-square-error-mse/>
- Loss and Cost Function in Machine Learning.* (2023, 4 2). Loss and Cost Function in Machine Learning. <https://www.enjoyalgorithms.com/blog/loss-and-cost-functions-in-machine-learning>