1. Which one is better – random weight assignment or assigning same weights to the units in the hidden layer?
2. Generally, how do we choose optimum hidden layers to get efficient accuracies.
3. When to use dropout and which layers to use dropouts.
4. Sometimes deep learning algorithm are less efficient than regular machine learning problems why is that?
5. How is the concept of VC dimensionality is used in regular problems?
6. Many algorithms like LMS how they are exactly used in working of current NN.
7. How do we know the backpropagation algorithm is not stuck local minimum?
8. **Suppose that you have to minimize the cost function by changing the parameters. How do we do it?**
9. **Would the network that uses a dimensionality reduction technique always give same output as network with hidden layer?**
10. **Which gradient technique is more advantageous when the data is too big to handle in RAM simultaneously?**
11. **“People who bought this, also bought….” recommendations on Amazon, what algorithm is used here?**
12. **Why is Manhattan distance not used in kNN machine learning algorithm to calculate the distance between nearest neighbours?**
13. Is the given deep learning model prove to over fitting? If so, what can you do about this?
14. What approach will you follow to suggest followers on Twitter?
15. How will you sample a stream of data to match the distribution with real data?
16. Google and Facebook are **translating text** into hundreds of languages at a time. This is being done through some deep learning models being applied to NLP tasks. How exactly this is different from Pucho?
17. **When you have an appropriate type of deep learning to solve the problem.**
18. **deep learning models require clear and informative data (and mostly big data) to train. What to do when we have not much of data?**
19. **Will I be using GPUs, CPUs or both? Am I going to rely on a single-system GPU or a distributed system?**
20. What’s my data pipeline? How do I plan to extract, transform and load the data (ETL)? Is it in an Oracle DB? Is it on a Hadoop cluster? Is it local or in the cloud?
21. **What kind of non-linearity, loss function and weight initialization will I use?**
22. Why are CNNs used primarily in imaging and not so much other tasks?
23. Is it OK to connect from a Layer 4 output back to a Layer 2 input?
24. the approach in SSD and YOLO for object detection. How these approaches differ from Faster-RCNN. When will you use one over the other?
25. Why do RNNs have a tendency to suffer from exploding/vanishing gradient?
26. Does using full batch means that the convergence is always better given unlimited power?
27. How to find the best hyper parameters?
28. Can they abstract the forward, backward, update operations as matrix operations, to leverage BLAS and GPU?
29. What is the problem with sigmoid during backpropagation?
30. Given a black box machine learning algorithm that you can’t modify, how could you improve its error?
31. How RPNs are trained for prediction of region proposals in Faster R-CNN?
32. What are the best practices for data cleaning?
33. How would you build a system that auto corrects text that has been generated by a speech recognition system?
34. What is latent semantic indexing and where can it be applied?
35. What are stop words? Describe an application in which stop words should be removed.
36. What is a regular grammar? Does this differ in power to a regular expression and if so, in what way?
37. What are the difficulties in building and using an annotated corpus of text such as the Brown Corpus and what can be done to mitigate them?
38. How would you build a POS tagger from scratch given a corpus of annotated sentences? How would you deal with unknown words?
39. **In Latent Dirichlet Allocation model for text classification purposes, what does alpha and beta hyperparameter represent-**
40. **Word2Vec model is a machine learning model used to create vector notations of text objects. Does Word2vec contains multiple deep neural networks?**
41. **What the deep learning models used in pucho?**
42. **Does pucho give flexibility to employees to try out something new?**
43. **In pucho does R&D based on deep learning take place ?**
44. **How much of understanding of every NN model used is stressed?**
45. **During the course of internship are will we get to work with real time model that pucho works?**
46. **Pucho specialises in NLP, Is it trying to explore other domains of AI?**