ML Questions Part 2

1. Tell me something about your project you have done in past?
2. What was your Dataset size for ML Project?
3. What is type of your dataset?
4. What was frequency of your dataset? (E.g. batch, streaming etc)
5. What was source system for your dataset? (E.g. sensor, satellite Kafka, cloud, etc.)
6. What was kind of derived dataset that you have mentioned in project?
7. How you have done validation dataset?
8. Have you created any pipeline to validated this dataset or you were using any tool?
9. What do you understand by data lake?
10. What do you understand by data warehousing?
11. Can you please name some validations that you have done on top of your data?
12. How you have handled streaming dataset?
13. How many different types of environments were available in your project?
14. What was your delivery mechanism for particular project?
15. Have you used any OPS pipeline for this current project?
16. How you were doing model retraining?
17. How you have implemented model retraining in your project?
18. How frequently you have been doing model retraining and what was the strategy for model retraining?
19. What was kind of evaluation you were doing in production environment
20. What was no. of request (hits) your model was receiving on daily basis?
21. How you have implemented logging in project for any failure cases?
22. How you have integrated notification (or Alarm) system for your project?
23. How you have implemented model monitoring?
24. How you have derived final KPI (Key Performance Indicator) for your client?
25. How many dashboards were there in your project?
26. On which platform you have productionised your model?
27. What kind of API you have exposed to receive data for model?
28. What was size of your final production environment (system configuration)?
29. What an all Databases you have used in project?
30. What kind of optimization you have done in your project, till what depth & explain the example
31. Can you please talk about complete team structure and team size?
32. What was duration of your complete project?
33. What was your day to day responsibility in last 2 month?
34. What kind of change request you have been receiving after you productionised project
35. What kind of testing you have done in development, UAT, pre-pod and prod?
36. Have you used some of the predefined AI-OPS pipelines if yes explain
37. Who has implemented AI-OPS in your project?
38. What was OPS stack you have been using?
39. What do you understand by CI-CD & have you implemented those in your project. If yes, what was the tech stack you used for CI-CD pipeline?
40. What was biggest challenge you faced in project and how you have resolved it?
41. Give me one scenario where you worked as team player?

42. What was your overall learning from current project?
43. How do you keep yourself updated for new technology?
44. Have you designed an architecture for this project? If yes, define a strategy wrt to your current project.

--------- Below are some questions for people with 7+ years of Exp-----------

- What kind of discussions generally wrt client?

46. What was your contribution in team building?

- How you have defined completed tech stack for AI?
- What was kind of benefit you have given to your current company in terms of cost cutting and revenue?
- What kind of new innovation you have introduced?
- How do you push your team for research or new implementation?
- How many projects you are handling?
- How many clients you have acquired?
- If new demand comes from client, how do you evaluate that requirement?
- How do you prepare costing for project?
- What kind of skillsets you look in person, to handle delivery of upcoming project?
- On an average, how much time you took for building a new team?
- What kind of stack you involved in initial project?
- How do you decide timeline for project delivery?
- How do you keep track of project progress?
- How do you handle dependencies between the team?
- How much profit you have given to your previous organization.

Deep learning & Vision Questions

- How many images you have taken to train your DL model?
- What is size of your model that you have productionised?
- Have you tried optimizing this Vision or DL model?

65. Where you have hosted your Computer Vison model?

- What was your frame per second?
- What are the data filtration strategy you have defined for CV project in production?
- Have you used any edge device in this project, if yes, why?
- What was name of camera & camera quality?
- What was final outcome you were generating from these devices?
- Have you processed the data in local system or in cloud? Give reason.

72. How many number devices you have productionised (camera, edge devices etc.)

- Let's suppose I am trying to build solution to count no. of vehicle or to detect their no. plate or track their speed. Then what is the dependency of distance, position & angel of camera on your final model? What will happen to your model? if we change position angle.
- What was your data collection strategy in CV project, have you received data from client or you have created the data? And how you have implemented it?

- What was data labelling tool that you have used for your project?
- If I have to do OCR then which API you will use or you have used in your previous project?
- Suppose if my images data is blur, what will be your strategy to enhance the image quality?

78. Have you implemented object tracking in any of your project, if yes, give me scenario?

- Suppose there are 2 mobile devices which are moving, so suppose if this two device position overlaps with each other then what will be your strategy to avoid any error while tracking those devices using camera?
- Have you implemented multicamera tracking, do you have Any idea about it?
- Explain me some of real-life use case of segmentation.
- What kind of AI application you will build for retail seller to increase their sell?
- Let's suppose I am trying to build AI solution to monitor a productivity of a kid, what kind of feature you would like to give in that product?

------- NLP-------

84. Have you productionised BERT Model, if yes, can you talk about hurdles that you have faced?

- How you have optimized your BERT based solution
- What kind of NLP tasks you were doing wrt BERT?
- Whether you have implemented BERT base & BERT large?
- What are the disadvantages of using BERT?
- Can please one of the lighter versions of transformer-based model?
- What was the accuracy that you were receiving wrt specific tasks?
- Let's suppose I have to build language-based model and there are online solution providers are there but they are costly, so what will be your strategy?
- Have you used hugging phase APIs?
- What is diff between BERT and GPT based models?
- There is no Decoder model in BERT then how do we get output if its just a encoder lever model?
- How masking is been implemented in BERT based model and what are its disadvantages
- How masking is been implemented in GPT based model and what are its disadvantages
- How backpropagation happens in BERT model. Explain.
- Can you explain Query, Key & Value in any Transformer based model?
- What are the main reasons behind success of transformer-based models?

100. Can we use BERT based model to generate embedding? if yes, How? NO, then Why?
101. What do you think about Open AI GPT3?
102. What is your thought about convolution auto encoder?
103. List down text summarization techniques, and which latest model you will prefer for text summarization?
104. What is the meaning of multiheaded attention
105. What do you understand by BLUE Score?
106. What is gradient clipping?
107. Can you please list down ways by which I will be able to split training across multiple GPUs?
108. Explain me difference between GRU and LSTM.
109. If I have to implement Any hour then what will be your approach? (Hint: using NLTK, Spacy or state of art model)
110. What do you understand by Uni-gram, bi-gram and tri-gram(N-gram)?
111. What do you understand by stemming and lemmatization?

112. For conversational AI solution will you use to predefined framework or will you create your own. In both cases what are adv and disadvantages?

113. Have you worked on google dialog flow, Azure-LUIS, IBM-Watson, or RASA- NLU?

114. What are the limitations of these respective platforms mentioned above?

115. If I have to build a ticket rerouting system for a banking client, how will you design this complete system?

116. Can you please how will you be able to design an app like In-Shorts? (hint: text summarization & etc)

117. If you have to build a solution that can generate a summary of the entire online class meeting then what will be your approach and what kind of hurdles you may face?

118. If I have to create a Gmail kind of text generation system then what will be your approach?

119. If I have to create a document parsing and validation system for legal, what will be your approach?

120. If you have to build a voice-based automation system, how you will design system architecture?

----- Time Series ----

121. List down, time series algorithms that you know?

122. How can we solve TS problem in deep learning?

123. Give application of TS in weather, financial, healthcare & network analysis?

124. What is diff between uptrend and downtrend in TS?

125. What do you understand by seasonality in

126. What do you understand by Cyclic pattern in your TS data?

127. How will you find Trend in TS Data?

128. Have you implemented ARCH model in TS? If yes, give scenario?

129. What is VAR (vector autoregressive) model?

130. What do you understand by univariant and multivariant TS Analysis?

131. Give example where you have created a multivariant model?

132. What do you understand by p, d, & q in ARIMA model

133. Tell me mechanism by which I can find p, d, q in ARIMA model?

134. What is SARIMA and how it's different from ARIMA?

135. What is meaning of AR, MA and I in ARIMA model?

136. Can we solve TS problems with transformers? what is your thought on that? why do you think in that way?

137. Have you ever productionised TS Based Model using LSTM? What are adv and disadvantages

138. Can we solve TS problem using Regressive algorithm, if yes, why, if no, give a reason?