

Interview Questions

Company Name: Blazeclan (Date : 17 Jan 2025)

Questions:

1. What is the role of randomness in the Random Forest algorithm?
2. What are the key assumptions of linear regression?
3. How does a CI/CD pipeline facilitate machine learning model deployment?
4. Why are vectors widely used in machine learning and data representation?
5. What hyperparameters have you optimized in your projects?
6. What are the key hyperparameters in Gradient Boosting, and how do they affect performance?
7. What frameworks have you used in your projects, and why?
8. How frequently and effectively do you use Git for version control?
9. What models have you implemented in your projects, and what were their purposes?
10. What is an open-source vector database, and how is it used?

Program:

Input:

```
input = 'aaabbbbcc'
```

Output:

```
output = '3a3b2c'
```

Company Name: ACL Tech (Date : 23 Jan 2025)

Questions:

1. What is polymorphism in object-oriented programming?
2. How can we group words from a list into separate groups?

Program: Group Similar Words

Input:

```
input = ["eat", "tea", "tan", "ate", "nat", "bat"]
```

Output:

```
list1 = ["eat", "tea", "ate"]
```

```
list2 = ["tan", "nat"]
```

```
list3 = ["bat"]
```

3. What is the working mechanism of the Random Forest algorithm?
4. What are the main differences between bagging and boosting?
5. What measures can ensure data security in an AI project?
6. What challenges are involved in deploying machine learning models to production?
7. Why might fine-tuning not be preferred in some machine learning scenarios?
8. How does the TF-IDF algorithm work for text representation?
9. What is an LSTM, and how does it handle sequential data?
10. How does the Transformer architecture work?

Company Name: Codincity

Date: 24 Jan 2025

Conducted by: CEO and Senior Architect (Experience: 35+ Years)

Questions:

1. Why do we use LLM?
 - What advantages do Large Language Models (LLMs) offer in real-world applications?
 - How do LLMs contribute to understanding and generating natural language?
2. Explain your Project.
 - What is the focus of your current project?
 - How do you handle the extraction of data from invoices (e.g., total amount, dates, GSTIN)?
 - What technologies and methodologies are employed in your project?
3. As this is an invoice extractor project, how do you measure the accuracy?
 - What metrics do you use to assess the model's accuracy?
 - How do you handle situations when the model provides incorrect results or predictions?
4. If accuracy is 90%, which error prevention method would you use: L1 or L2?
 - Can you explain the difference between L1 and L2 regularization?
 - When would you choose one over the other based on your project's requirements?
5. If the model is overfitted, what will you do?
 - What methods would you employ to prevent or resolve overfitting in a machine learning model?
 - Can you describe an example where you handled an overfitting problem in the past?
6. How do you maintain client data security in the project, especially when using an open-source model?
 - How do you ensure data privacy and security when using open-source models?
 - What measures are in place to safeguard sensitive client data during processing?
7. Rate yourself in Python and SQL.
 - How would you rate your proficiency in Python and SQL on a scale of 1 to 10?

- Can you provide examples where you've demonstrated advanced skills in Python and SQL?
8. Are you able to manage the team?
- How do you handle team dynamics and task delegation in a technical project?
 - What is your approach to ensuring productivity and collaboration within your team?
9. Are you able to provide training and guide the Junior Data Scientist?
- How would you approach mentoring junior data scientists in your team?
 - Can you give an example of a situation where you provided guidance or training to junior team members?
10. Are you able to directly interact with clients and give proper solutions without making fake promises?
- How do you manage client expectations and deliver realistic solutions?
 - What strategies do you use to communicate complex technical details to clients clearly and without overpromising?
-