
AI/ML Interview Questions (5 per section)

1. Core Concepts

Purpose: Test understanding of algorithms, models, and data handling.

- Q:** Explain the difference between supervised and unsupervised learning.
A: Supervised learning uses labeled data to train models (e.g., classification, regression), while unsupervised learning works on unlabeled data to find patterns (e.g., clustering, dimensionality reduction).
- Q:** What is overfitting and how can you prevent it?
A: Overfitting occurs when a model learns noise in the training data. Prevention methods: regularization (L1/L2), dropout, cross-validation, pruning, more data.
- Q:** What is the difference between precision, recall, and F1-score?
A: Precision = $TP / (TP + FP)$, recall = $TP / (TP + FN)$, F1-score = harmonic mean of precision & recall. F1 balances both metrics for imbalanced datasets.
- Q:** Explain gradient descent and its variants.
A: Gradient descent is an optimization algorithm that updates model weights using the loss gradient. Variants: Stochastic (SGD), Mini-batch, Adam, RMSprop for faster convergence and stability.
- Q:** What are embeddings in machine learning?
A: Embeddings are dense vector representations of categorical variables, words, or entities, capturing semantic relationships in lower-dimensional space (e.g., Word2Vec, BERT embeddings).

2. Project Experience

Purpose: Assess practical application and problem-solving.

- Q:** Describe an ML project you implemented end-to-end.
A: Candidate should explain problem, data preprocessing, model selection, training, evaluation, and deployment. Example: "Built a spam email classifier using NLP, TF-IDF, and Logistic Regression."
- Q:** How did you handle missing or noisy data in your project?
A: Methods: imputation (mean/median/mode), removal of rows/columns, interpolation, or using models robust to missing data.
- Q:** Which model evaluation metrics did you choose and why?
A: Example: For classification, accuracy, precision, recall, F1-score; for regression, RMSE, MAE, R^2 . Choice depends on problem type and dataset balance.
- Q:** How did you optimize your ML model's performance?
A: Hyperparameter tuning (GridSearch/RandomSearch), feature engineering, normalization, ensemble methods, or model selection.

5. **Q:** Describe a challenge you faced in your ML project and how you solved it.
A: Example: "Class imbalance in dataset—used SMOTE oversampling to balance classes, improving F1-score by 12%."
-

3. Teaching / Mentoring Ability

Purpose: Evaluate communication and mentoring skills.

1. **Q:** Have you explained ML concepts to peers? How?
A: Example: Conducted workshops on Python ML libraries or explained neural networks using visualizations and simple examples.
 2. **Q:** How would you teach someone the difference between regression and classification?
A: Regression predicts continuous values (e.g., house prices), classification predicts discrete classes (e.g., spam/not spam). Use examples and visual plots.
 3. **Q:** How do you simplify complex ML topics like deep learning?
A: Use analogies, step-by-step visualizations, simple datasets, and interactive coding demos.
 4. **Q:** Have you contributed to tutorials or open-source educational content?
A: Example: Created GitHub notebooks, blog posts, or YouTube tutorials on ML projects.
 5. **Q:** How do you evaluate if someone has understood your ML teaching?
A: By giving exercises, coding tasks, quizzes, or asking them to explain back concepts in their own words.
-

4. Google Technologies

Purpose: Test familiarity with Google Cloud and related ML tools.

1. **Q:** Have you used Google Colab for ML experiments?
A: Yes—used for training models with GPU acceleration and sharing notebooks collaboratively.
 2. **Q:** Explain how you would deploy an ML model on Google Cloud.
A: Use Vertex AI: train or upload a model, create endpoint, and deploy it for predictions; optionally integrate with Cloud Functions or App Engine.
 3. **Q:** Have you used TensorFlow or TensorFlow Extended (TFX)?
A: TensorFlow for model building; TFX for production pipelines including data validation, transformation, and serving.
 4. **Q:** How would you use BigQuery for ML tasks?
A: Use BigQuery ML to run SQL-based model training directly on large datasets without exporting them, e.g., linear regression, classification.
 5. **Q:** How do you leverage Google APIs for AI/ML projects?
A: Example: Cloud Vision API for image recognition, Natural Language API for sentiment analysis, Speech-to-Text API for audio processing.
-

5. Management / Leadership Skills

Purpose: Assess project planning, teamwork, and leadership.

1. **Q:** How do you plan an ML project from scratch?
A: Define problem → gather data → preprocess → select model → train/evaluate → deploy → monitor. Use milestones and version control.
2. **Q:** How do you handle deadlines for multiple ML tasks?
A: Prioritize tasks based on impact, break into sprints, track progress using tools like Jira or Trello.
3. **Q:** How do you collaborate with data engineers or backend teams?
A: Coordinate on data pipelines, APIs, deployment, and monitoring; maintain clear documentation and code standards.
4. **Q:** How do you ensure reproducibility in ML projects?
A: Use version control (Git), track dependencies, save datasets, set random seeds, containerization (Docker).
5. **Q:** How do you mentor junior team members in ML projects?
A: Provide guidance on data preprocessing, coding standards, model selection, encourage pair programming, and code reviews.