

1. Clarifying Requirements

A short conversation with the interviewers, asking about some general questions.

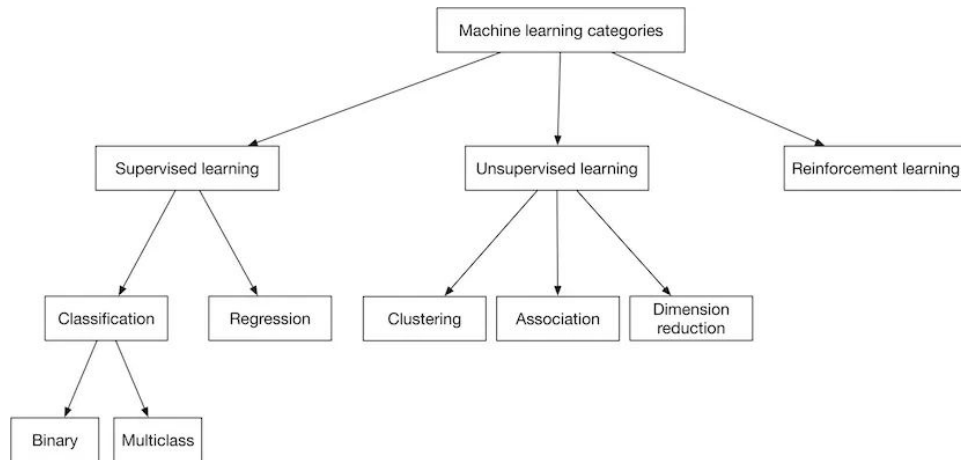
Topics could include but not limit to:

- business objective
- Features
- Data
- Constraints
- Scale
- Performance

After the conversation, it is good to summarize the problem in several sentences.

2. Frame the problem as ML

- Define ML objectives (be concrete, such as maximize ad click rate, increase user engagement)
 - translate business objective to ml objective
- Specify system's input and output (simple diagram)
- choose right ML category



3. Data Prepare

- Manifest raw data (table form)
- Feature engineering
 - Specify important features, why they are important?
 - Preprocess features into ML readable data
 - Text: BOW, Word2Vec, BERT
 - Image/Video: CNN
 - Category: One-Hot
 - IDs: embedding layer

4. Model Development

- Propose at least two options
 - Use simple diagram to show how each works
 - Discuss the pros/cons of each option

5. Model Training

- Discuss how to build the dataset from scratch
 - how to label data (for ad click, positive samples could be ads clicked by users)
 - be careful about imbalance dataset
- Choose loss function
- Discuss if fine-tuning is possible

6. Evaluation

- Offline evaluation
 - Precision
 - Recall
 - mAP
 - ROC
 - Cross-Entropy
- Online evaluation
 - Click-through rate (CTR)
 - Dwell time (time spent on posts/ads)
 - User satisfaction rate found in a user survey
 - Conversion rate
 - Reaction rate
 - Revenue lift

7. Deployment

- Concrete diagram showing the overall system pipeline
 - Data generation pipeline
 - Prediction pipeline