**4. How do you evaluate the performance of a machine learning model?**

**Please explain some commonly used performance metrics.**

To assess a machine learning model's performance, follow these steps:

1. **Prepare data:** Divide dataset into three parts: training data to build the model, validation data to fine-tune settings, and test data to measure final performance.
2. **Choose appropriate metrics:** Select evaluation measures based on your problem. Common options include error rates for predictions (like Mean Squared Error) and classification metrics (like accuracy, precision, and recall).
3. **Test the model:** Apply the chosen metrics to your test data to determine how well the model performs on unseen data.
4. **Compare models:** If you have multiple models, use statistical methods to decide which one is best.

**Key points:**

* Avoid overfitting or underfitting your model.
* Balance between bias (error due to assumptions) and variance (error due to model sensitivity).
* Use domain knowledge to interpret results and identify potential issues.

**And:**

* **Confusion matrix** provides a detailed breakdown of correct and incorrect predictions.
* **ROC curve** visualizes the trade-off between true positive rate and false positive rate.
* **AUC** is the area under the ROC curve, representing overall model performance.
* **mAP** considers multiple thresholds for object detection evaluation.
* **Task-specific metrics** vary depending on the specific language model task (e.g., summarization, translation).

**Commonly used performance metrics**

**(This table is converted to image format to fit the page limits)**

A close-up of a document

Description automatically generated

A screenshot of a computer

Description automatically generated