Task: “Build & Judge a Mini AI”

Part 1 — Chronology of AI

Write one real-world example for each stage:

1. **Machine Learning** → Netflix
2. **Deep Learning** → DALL-E
3. **Computer Vision** → Face-ID in iPhones
4. **NLP** → Chatbots
5. **LLMs** → ChatGPT

Part 2 — Deep Learning Architectures

Match the model to the use case:	Use cases:
1. RNN	a. Image recognition
2. LSTM	b. Text translation (old Google Translate)
3. CNN	c. Predicting the next word in ChatGPT
4. Transformer	d. Early speech-to-text systems

Answer: 1-d, 2-b, 3-a, 4-c

Part 3 — Frameworks

Choose one framework (PyTorch / TensorFlow / Keras).

In one sentence, explain why you would use it if you were a student making a cat-vs-dog classifier.

Answer: Keras - It is beginner friendly and offers a easy to understand approach and is well suited for student-based-approach.

Part 4 — Evaluation Metrics

Imagine you built a spam filter.

1. Precision: If it marks 10 emails as spam and 7 are truly spam → what’s Precision?

Answer: Precision would be 0.7 i.e. 70%.

Precision = True Positives / Number of Predicted Positives

Precision = 7/10 = 0.7

2. Recall: If there were 12 spam emails in total, how many did it catch? (use same example)

Answer: Recall = True Positives/ Actual Positives = 7/12 =~ 0.58 = 58%

3. **F1 Score: Use the formula and calculate (round to 2 decimals).**

Answer: F1 = 2 x (Precision*Recall) / (Precision + Recall)

F1 = 2. (0.7*0.58)/ (0.7+0.58)

F1 = 2. (0.406/1.28) =~ 0.64

4. **MSE/MAE: Predict your friend's age (actual = 15, prediction = 18). Which metric punishes the error more?**

Answer: Error = prediction - actual

error = 18-15 = 3

MSE = (18-15) *(18-15) = 9

MAE = |18-15| = 3

Therefore, MSE punishes the error more because the error is squared amplifying the mistake is larger

5. BLEU/ROUGE: AI translated "The cat sat on the mat" as "Cat is on the mat." Which metric (BLEU/ROUGE) do you think would give a high score?

Answer: ROGUE (Recall oriented Understudy for Gisting Evaluation)- it focuses on meaning/words. It is recall focused metric. It checks if the important content is captured, which here is captured unlike BLEU (Bilingual Evaluation Understudy) which measures how accurate the generated text is.

Part 5 — Responsible AI & Explainability

You built an AI that predicts loan approvals. A customer asks, "Why was my loan rejected?"

Write one simple way to explain the decision fairly (e.g., "Your income was too low compared to the loan size").

Answer: Your income was lower than the loan amount requested, therefore the loan request is rejected.