Machine Learning Based Loan Approval System

QUIZ (10 MCQs)

Q1. Why did we design this project as a two-stage system?

A. To reduce training time

B. To follow academic ML standards

C. To first decide approval, then predict loan amount only if approved

D. To increase dataset size

Correct Answer: C

Explanation: In real loan workflows, amount prediction only makes sense after approval.

Q2. What does the classifier predict in this project?

A. Loan amount

B. Interest rate

C. Loan approval or rejection

D. Credit score

Correct Answer: C

Explanation: Stage 1 predicts whether the applicant is approved or rejected.

Q3. Why was the regression model trained only on approved loans?

A. To reduce computation

B. Because rejected loans do not have meaningful loan amounts

C. Because regression works only on positive labels

D. To balance the dataset

Correct Answer: B

Explanation: Loan amounts are relevant only for approved applications.

Q4. Why were preprocessing steps included inside a Pipeline?

A. To improve accuracy automatically

B. To avoid writing preprocessing code

C. To ensure the same preprocessing is applied during training and prediction

D. To reduce memory usage

Correct Answer: C

Explanation: Pipelines prevent mismatches between training and inference.

Q5. What is the role of `ColumnTransformer` in this project?

A. To train multiple models

B. To combine numeric and categorical preprocessing

C. To tune hyperparameters

D. To evaluate the model

Correct Answer: B

Explanation: It applies different transformations to numeric and categorical features.

Q6. Why did we use GridSearchCV before finalizing the model?

A. To increase dataset size

B. To automate feature selection

C. To find the best hyperparameters

D. To reduce overfitting completely

Correct Answer: C

Explanation: GridSearchCV systematically searches for optimal hyperparameters.

Q7. Why was the entire pipeline saved instead of just the model?

A. Pipelines load faster

B. To include preprocessing logic along with the model

C. Joblib cannot save models alone

D. Pipelines give better accuracy

Correct Answer: B

Explanation: Preprocessing must remain consistent during inference.

Q8. What does a high approval probability indicate?

A. Model is overfitting

B. Applicant has high income

C. Model is confident about approval

D. Regression will succeed

Correct Answer: C

Explanation: Probability reflects model confidence, not certainty.

Q9. Why did we use a YAML file in this project?

A. To store data

B. To improve model accuracy

C. To manage configuration outside the code

D. To deploy faster

Correct Answer: C

Explanation: YAML allows easy configuration changes without touching code.

Q10. Why does the system skip regression if the loan is rejected?

A. To save time

B. Regression requires approval first

C. Regression model cannot handle rejection

D. To avoid errors

Correct Answer: B

Explanation: Loan amount prediction is only meaningful for approved cases.