a. How would you approach writing test cases to ensure lead scoring engine is tested for all possible scenarios of the workflow with all possible types of data collected about the prospect?

1. First need to understand the workflow and data points :-

- **Workflow Stages:**

- New Lead Generation
- Scheduling Test Drive
- Dealer Visit
- Test Drive,Loan Application,Pricing,Negotiation,Booking
- Follow-up
- Final Booking and Delivery

- **Data Points :**

- Basic Information (Name,Contact,Location)
- Requirements (Car Model,Budget,Features)
- Purchase Timeline
- Interaction History(Test Drive,Loan details,Price Negotiations)

1. Test Case Categories:-

- **Functional Test Cases:**Ensure that the lead scoring engine process data at each stage and assign a lead score to prospect.

- **Boundary Test Cases**:To test the maximum,minimum and edge cases for inputs(eg:high or low budgets,long or short timelines)

- **Integration Test Cases**:To verify that the lead scoring engine properly integrates with sales app

- **Performance Test Cases**:Ensuring the lead scoring engine efficiently process the data under different loads.(eg:large number of leads being processed parallely)

- **Data Test Cases**:To check whether the lead scoring engine uses complete and accurate data to calculate lead score

- **Scenario Based Test Cases**:To create test cases for different customer journey(eg: A customer may book after one visit vs the customer who books after multiple follow-ups)

- **Negative Test Cases**:Including scenarios where data is incorrect or incomplete to check how the system handles that values.

b. What kind of information from the model will be required to build test cases?

- Every data points that the model uses to score leads.

- Formula used to calculate the lead score.

- To know how the model reacts to different combinations of input data.

- Edge cases where model may behave unpredictably(while handling missing data).

- To guage performance need to know how quickly the model will process data and generate a lead score under different load situations.

- To know how score was derived ,need interpertablitiy mechanisms of the model,which would be helpful in testing and validation.

- Algorithm that is used by model when there is missing data.

- With the above information from the model,can ensure the test cases will be covering all scenarios.

c. What kind of historical data will be required to build test cases?

- Historical records of data that have gone through the entire sales funnel, from initial contact to final purchase or dropout.

- Data on the accuracy of predictions made by models, compared with actual outcomes.

- Instances where the lead score was high or low and the outcome did not align with the score.

- Incomplete and inconsistent data that need to be tested for robustness.

- With this above information from the model, can create test cases covering all input scenarios under different conditions.

d.What kind of information from the engineering team will be useful?

- Architecture of how lead scoring engine integrates with other systems.

- API's used by the lead scoring engine which request and response formats.

- Performance metrics such as repsonse time and data processing duration.

- Load testing results that has been already performed.

e. What kind of tests will be required to deploy the system into production?

- **Unit Testing** : To verify that components of the system work properly

- **Integration Testing**:To ensure that different components in the system work together in smooth manner.

- **System Testing**:Testing the whole system to check whether it meets the functional requirements.

- **Performance Testin**g:To guage system's performance under different load conditions.

- **Security Testing**:To indentify the vulnerabilities in the system.

- **Regression Testing** : To ensure that the new updates do not break the present funnctionalities.

- **Usablility Testing** : To assess the system user's interface and experience

f. How would you measure effectiveness of the model?

- Collect feedback from the sales team and try to understand how useful the model's predicitions are important in their day-to-day operations.

- From which we can get to know whether the leads prioritized by the model are likely to convert or not.