Defining Success Metrics

Objective

Develop a comprehensive plan to define, measure, and monitor success metrics for the Mango Oasis app, ensuring it provides accurate, helpful, and unbiased responses to customers interested in mango-related products.

Selected Metrics

Metric	Description	Target Value
Faithfulness	Assess how accurately the chatbot's responses reflect factual data and customer intent	90%+
Mango Recipe Accuracy	Measure the percentage of chatbot responses that provide accurate and relevant mango recipes	95%+
Bias	Monitor for any gender, racial, political, or geographical biases in the chatbot’s responses	0%
Efficiency	Measure how quickly and resource-efficiently the chatbot responds	Avg. 2 seconds or less
Answer Relevancy	Ensure that the chatbot's responses directly address the questions posed by the customers	85%+

Measurement Plan

- Regularly review a sample of responses for factual accuracy and alignment with customer intent.
- Implement a tagging system for recipe-related queries and responses to easily track accuracy.
- Conduct quarterly bias audits where diverse sample queries are reviewed for any identifiable bias.
- Track response times using system logs, ensuring the average response time stays within the target.
- Regularly review random samples of chatbot conversations, rating each response for relevance.

Monitoring and Iteration Plan

To ensure the Mango Oasis chatbot consistently delivers accurate, relevant, and helpful responses, we'll implement a thoughtful and proactive monitoring plan:

Regular Audits: Each month or quarter, our review team will evaluate a sample of responses for faithfulness, rating them on their alignment with factual information and customer intent.

Customer Feedback: We'll continually gather customer feedback and ratings on response helpfulness to identify trends and refine performance over time.

Recipe Accuracy Reviews: Our subject matter experts, along with a user feedback mechanism, will review recipe responses weekly to ensure they're accurate and relevant. This regular review process allows us to keep the knowledge base up-to-date and precise.

Efficiency Monitoring: We'll closely monitor resource usage to maintain fast response times and set up alerts for any instances where response times exceed our targets, so we can quickly address any potential inefficiencies.

Automated Relevance Scoring: Where possible, we'll utilize automated tools, such as natural language processing, to assess response relevance. In addition, a simple thumbs up/down rating system will allow users to directly share their perception of response relevance.

This ongoing monitoring approach helps us uphold high standards in both accuracy and efficiency, creating a consistently positive experience for our users.