Group: G10 - Strawberry Dumplings

Member: Tian Zebai_FW83IY, ChenSiyu_D8WCT3, WangFei_IIONHK,

Topic: Smart Ordering in a Restaurant

Evaluation Plan

Evaluation team composition

Our target user group comprises accessible smartphone users with one or more instances of utilizing a mobile application for booking and ordering food services. For the evaluation process, eight users were selected as evaluators, responsible for conducting testing tasks and sharing valuable feedback. A team member oversaw the evaluation to ensure its smooth progression and collected relevant data. To encourage active participation in this sub-evaluation, we offered each evaluator a drink token as a reward, fostering motivation and cooperation.

Content and preparation of assessments

Software version: the latest current stable version of the smart ordering system will be selected for testing.

Test equipment: the test will be conducted on Android and iOS platforms (mainly on the iOS platform as the model was created based on the iphone13), including smartphones and tablets.

Task case: mainly test search and intelligent recommendation scenarios, includes some basic usage scenarios.

Evaluation arrangements and timetable

During the schedule creation process, tests are administered throughout 3-4 weeks, with an estimated duration of 30 minutes per assessor. To ensure assessment efficiency, the start and end times of each task are recorded. Additionally, a buffer of a few minutes is allocated after each assessor's test to accommodate potential delays. The total time for each assessment session is anticipated to range between 30 and 60 minutes, depending on the complexity of the task.

Assessment sites and settings

To ensure a comprehensive evaluation process, we are offering location options that allow assessors to participate both online and offline, providing flexibility for participation from any location with an Internet connection. Accessibility is a priority, and we are taking steps to ensure that all assessors have easy access to the assessment platform. Additionally, assessor segregation measures will be implemented to prevent contact between assessors, thereby safeguarding the objectivity of the results. This approach aims to create a logical and systematic assessment procedure while prioritizing accessibility and maintaining the integrity of the evaluation process.

Overall testing process

During the evaluation process, the evaluator will perform a series of tasks, including

searching, selecting and adding items to the shopping cart to recreate real-life scenarios. Throughout the task execution, user experience and system performance will be monitored, with metrics assessed via questionnaires and in-app tracking tools.

Before the evaluation, the evaluators will understand the test case demo, and after the test, provide feedback by filling out a questionnaire. This pre- and post-test preparation ensures a thorough understanding of the assessment process. To maintain consistency and fairness, the system goes through a recovery phase after each evaluation, resetting it to its initial state for subsequent evaluators. This approach ensures a standardized starting point for each assessment.

After the evaluation phase, the collected data will be carefully processed and analyzed. This data-driven approach is the foundation for identifying areas for system improvement, allowing informed decisions to be made that improve overall performance and user satisfaction.

The following is an introduction to test cases：

Task 1：Fuzzy search with precise recommendations:

1. Evaluators imagined that they wanted to eat a "spicy dish" and then opened the app and searched for it using voice input.
2. The app returns specific dish suggestions and corresponding merchant options using ElasticSearch and ChatGPT technology.
3. The evaluator selects a dish based on the returned results, but does not place an order.
4. The relevance and accuracy of the search results are recorded and evaluated, as well as the accuracy of the voice recognition.

Task 2: Specific Needs Search and Selection:

1. The evaluator uses voice to search for a specific dish, such as "hot pot".
2. The app displays a list of nearby merchants where hotpot can be ordered.
3. The evaluator selects a merchant, views the details of the dish, and attempts to add it to the shopping cart.
4. No payment is made, and the evaluator checks that the items and prices in the shopping cart are correct before exiting.
5. Evaluate the smoothness of the whole operation process and the response time of the system.

Questionnaire

A Google questionnaire was used, here is the link to the questionnaire to click through to view it: https://forms.gle/Hziaff5nkWKinVgQA

Through this detailed evaluation programme, we were able to ensure that we were able to efficiently collect a large amount of useful user feedback and system performance data within a limited timeframe. This information will help us gain insights into the user experience when using the smart ordering system and provide empirical evidence for future product iterations.