

## Hand phone Recommendation Assistant

### Business problem:

There's a high saturation of phones from different brands with many features and specification. As such it can be hard for customer to find and research the phone they want. Hence, the need for a hand phone recommendation assistant that helps customer narrow down the phone they want and available in the Singtel shop.

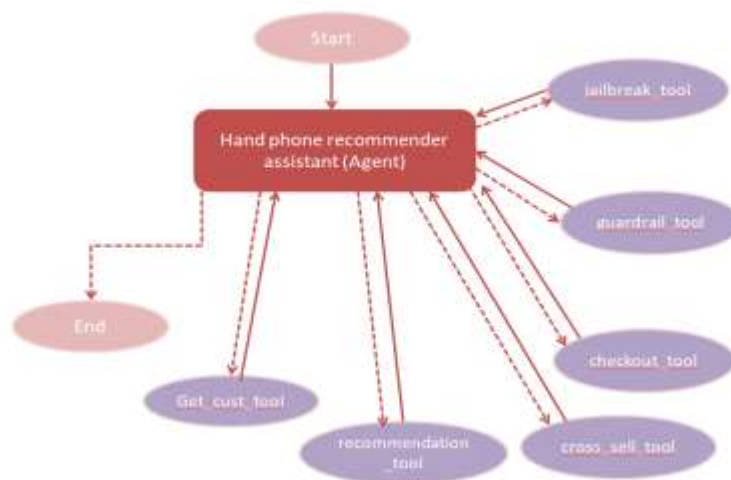
The hand phone recommendation assistant currently takes in Customer ID to extract the phone model customer is using to recommend a higher model than what customer possess. Customer can still tell the assistant their preferences in order for assistant to retrieve phone models that are of customer's interest. Finally if customer tells assistant that they interested in the phone model, assistant will direct customer to the URL of the phone model on the Singtel shop.

This will help customer save time in doing their own search and increase chance of conversion.

### Agent Model:

The hand phone recommendation assistant is built using Langgraph prebuilt agent, powered by Gemini 2.0 Flash (Model is free up to 200 requests per day and can take 1,000,0000 tokens per minute). The agent is equipped with custom tools:

- [get\\_cust\\_tool](#): retrieves customer's information given the customer ID from customer database
- [recommendation\\_tool](#): retrieves the phones in handset database specified by customer, defaults to recommending the latest phone based on launch date in the database.
- [cross\\_sell\\_tool](#): Cross sell XPower B10H power bank to customer
- [checkout\\_tool](#): directs customer to the URL of the phone model on the Singtel shop
- [guardrail\\_tool](#): stops customer from asking things not related to hand phones
- [jailbreak\\_tool](#): stops the agent if jailbreaking detected



Future tools for the agent:

- **customer\_authentication\_tool**: Currently there is a prompt in the agent that tells it how to handle asking for customer ID, this should be converted to a tool with better authentication method
- **upsell\_tool**: Persuades customer with reasoning, to take up for example, higher phone storage (“why not get the 1TB instead of 512GB that is also available on Iphone 16? You will be able to store more memories of your friends and loved ones” ) when they have decided on a phone they want
- **discount\_tool**: Provide X% of discount to customer based on their customer tier that can be stored in customer database
- **receipt\_tool**: If customer uses the assistant at a kiosk/tablet in physical shop, print out the phone information so that customer can pass it to staff to deliver phone to them and make payment

The current tools can also be refined further with more sophisticated algorithms. For example, the cross sell tool can retrieve the result of market basket or recommender system instead on the accessories that can be cross sold to customers instead.

## Data:

The hand phone recommendation assistant uses customer’s data and handset data. For this demo, mocked data were generated:

### Prompt to generate customer data

create a table of a customer database with 10 customers. the columns should have customer\_id from 1 to 10, customer\_names at random, customer\_mobile\_plan with values A, B and C randomly assigned, customer\_phone\_brand with values apple, samsung and oppo randomly assigned and customer\_phone\_sku randomly assigned and needs to tally with the customer\_phone\_brand .

Here's a randomly generated customer database table with 10 customers. The `customer_phone_sku` values are consistent with the `customer_phone_brand`:

customer_id	customer_name	customer_mobile_plan	customer_phone_brand	customer_phone_sku
1	Alice Tan	A	Apple	iPhone 15 Pro
2	Benjamin Lee	C	Samsung	Galaxy S24 Ultra
3	Chloe Ng	B	Oppo	Oppo Find X6 Pro
4	Daniel Goh	A	Apple	iPhone 14
5	Eva Lim	B	Samsung	Galaxy Z Flip 5
6	Faris Ahmad	A	Oppo	Oppo Reno11 F
7	Grace Yeo	C	Apple	iPhone SE (2022)
8	Haziq Hassan	C	Samsung	Galaxy A54
9	Iris Tan	B	Oppo	Oppo A98
10	Joshua Chua	A	Samsung	Galaxy S23



## Prompt to handphone data

The screenshot shows the Gemini AI interface. On the left, a text box contains a prompt: "help me generate a table of the specifications of apple, samsung oppo handsets, the columns should have the phone brand, phone model, phone colour, phone screen size, phone chip, phone operating system, phone storage size, phone weight, phone price in SGD. Each row should be for a single SKU for each brand. give me all available SKUs in the market." Below the prompt, a button labeled "Ask Gemini" is visible. On the right, the AI's response is displayed, titled "Comprehensive Specifications of Apple, Samsung, and OPPO Handsets in the Singapore...". The response includes a table with columns: Phone Brand, Phone Model, Phone Colour, Phone Screen Size, Phone Chip, Phone Operating System, Phone Storage Size, Phone Weight, Phone Price in SGD, and Launch Date. The table lists three Apple iPhone 16 Pro models: Desert Titanium, Natural Titanium, and White Titanium, all with 6.3-inch Super Retina XDR OLED screens, A18 Pro chips, 128GB storage, and a price of \$1,599.

Phone Brand	Phone Model	Phone Colour	Phone Screen Size	Phone Chip	Phone Operating System	Phone Storage Size	Phone Weight	Phone Price in SGD	Launch Date
Apple	iPhone 16 Pro	Desert Titanium	6.3-inch Super Retina XDR OLED	A18 Pro chip with 6-core GPU	iOS 18	128GB	199 grams	From \$1,599	September 20, 2024
Apple	iPhone 16 Pro	Natural Titanium	6.3-inch Super Retina XDR OLED	A18 Pro chip with 6-core GPU	iOS 18	128GB	199 grams	From \$1,599	September 20, 2024
Apple	iPhone 16 Pro	White Titanium	6.3-inch Super Retina XDR OLED	A18 Pro chip with 6-core GPU	iOS 18	128GB	199 grams	From \$1,599	September 20, 2024

Only handphone data is stored in a vector store, customer data is stored as structured table.

The RAG pipeline now only uses a simple similarity search based on cosine similarity to extract the phones that fit customers' preference. It could be further fine-tuned by tweaking the similarity score threshold or trying out other embedding models.

The handphone data itself could be further preprocessed to make the probability of retrieving the correct phones higher. For example, recoding the colour == Desert Titanium of Iphone 16 Pro to beige instead would increase the similarity score for query of "I want a beige phone" and retrieving it to recommend to customer.

The agent logs all responses by both human and assistant (AI) and also the tools that were used, see example:

```
2025-07-07 23:18:26,355 - INFO - ai Hi! I am your friendly handset recommender assistant! Before we start, can you please provide the numeric portion of your customer id?

2025-07-07 23:18:32,607 - INFO - human None : 9999
2025-07-07 23:18:34,669 - INFO - ai None
2025-07-07 23:18:35,513 - INFO - tool GetCustomerTool ('messages': [AIMessage(content='I do not have information for customer ID 9999.\nWhat phone are you interested in today?', ad
2025-07-07 23:18:36,326 - INFO - ai None I do not have information for customer ID 9999.
What phone are you interested in today?
2025-07-07 23:18:45,066 - INFO - human None oops sorry, it's 10
2025-07-07 23:18:45,802 - INFO - ai None
2025-07-07 23:18:46,929 - INFO - tool GetCustomerTool ('messages': [AIMessage(content='Hi Joshua, I see you're currently using a Samsung Galaxy S23. Are you interested in upgrading
2025-07-07 23:18:48,003 - INFO - ai None Hi Joshua, I see you're currently using a Samsung Galaxy S23. Are you interested in upgrading to the latest Samsung model?
2025-07-07 23:18:52,245 - INFO - human None yeah ok
2025-07-07 23:18:53,002 - INFO - ai None
2025-07-07 23:18:56,420 - INFO - tool RecommendationTool ('messages': [AIMessage(content='Here are the latest three Samsung phone models:\n\n1. **Samsung Galaxy S25 (Icyblue):**\n
2025-07-07 23:18:58,274 - INFO - ai None Here are the latest three Samsung phone models:

1. **Samsung Galaxy S25 (Icyblue):**
  **Launch Date:** February 7, 2025
  **Screen Size:** 6.3-inch Dynamic LTPO AMOLED 2X
  **Chip:** Qualcomm Snapdragon 8 Elite
  **Operating System:** Android 15, One UI 7
  **Storage:** 512GB (12GB RAM)
  **Weight:** 162 grams
  **Price:** $1,449
```

## **Integration Strategy and Approach:**

### **Possible deployment methods**

The hand phone recommendation assistant could be deployed on <https://shop.singtel.com>, on the tablets in the physical store for guided discovery, or on Sales Rep service tablet to better help customer choose the phone.

### **Anticipated practical integration challenges**

**Authentication:** There is an authentication whereby customer needs to input their customer ID into the chat box in order to retrieve their information for recommendation. The Jailbreak tool of the agent prevents customer from attempting to pull out customer database. Can also integrate the assistant with Singtel Account so that checkout and payment done would be more secure.

**Latency:** The recommendations now take about 5 seconds to load on the App on a small scale hand phone dataset. There would be a need to improve the RAG by tweaking similarity score threshold, reducing the number of hand phones to be retrieved or using faster embeddings. This would ensure the assistant's reply would not slow down when retrieving from actual hand phone database.

**UI design:** The interface is a simple chatbot layout for user to input their queries and the assistant to reply. The interface is designed to adjust its size to fit onto a larger laptop screen or a small mobile phone. However whenever the assistant recommends a few hand phones with their specifications, user would need to scroll the text as the chatbot screen would not be able to fit all information.

**Scalability:** There could be surge in usage of assistant during sales event or new phone launches which could slow down the assistant app. The assistant if deployed on cloud computing platform like AWS or Azure, there would be functionality to spin up more instances during peak periods and distribute traffic across compute instances.

### **Performance Metrics:**

To evaluate the effectiveness of the hand phone recommendation assistant, we can look at

- % of customers who bought phone after using the assistant (number of customers bought phone with assistant/total customers who used assistant)
- Score rating of the assistant at the end of the session (will need to implement a rating system that appears after end of interaction)
- Number of interactions with assistant (i.e. how long customer has to chat with assistant in order to get the phone that meets their requirements)
- Cost breakdown of the phones, accessories sold by the assistant