Project Report

New Car Buy Recommender System

Masters of Technology in Intelligent Systems

Module: Machine Reasoning

Ajay Vikram Singh (A0020986B) Rahul Jalan (A0195299H)

TABLE OF CONTENTS

EXECUTIVE SUMMARY	4
PROBLEM	5
Definition	5
Description	
SCOPING FOR THE PROJECT	
Inclusions	
Business	
Technical	
Exclusions	6
SOLUTION	7
Knowledge Discovery & Acquisition	7
Table 1: Female drivers a CF features economy = 1	
KNOWLEDGE STRUCTURING — DEPENDENCY DIAGRAM	
Knowledge Modelling & Representation	
Buying budget behaviour and calculation: Derived Rules	
Table 2: Mapping of Age to CF 70/30	
Certainty factors	
Table 3: Certainty factor for the car recommender system	
User inputs:	
Table 4: User inputs table for the car recommender system	11
Car recommendation system application outputs:	11
Defined Rules	
Example 1: user survey derived CF recommendation	
Table 5: >50 years old users car make preference	11
Example 2: user survey derived CF recommendation	
Table 6: Males prefer safety and security	
Table 7: Flowchart for Rule Definition engine to update the Certainty Factor	
CAR RECOMMENDATION SYSTEM OUTPUTS	
Table 8: Output setup for the car recommendation system (in current version)	
System Design:	
System Architecture considerations:	15
INPUTS & OUTPUT DIALOGUE SCREENS	16
USER INPUT FORM	16
RECOMMENDATION SCREEN	
PROJECT CRITICAL ANALYSIS	17
Key Survey and Expert Findings	17
FUNCTIONALITY AND SYSTEM SHORT COMINGS, PENDING ISSUES	17
System Limitations	
FUTURE PROJECT ENHANCEMENTS	19
Scope Enhancements	10
FUNCTIONAL ENHANCEMENTS	
TECHNICAL ENHANCEMENTS	
	_
REFERENCES	20
APPENDICES	
A = = A - C	2.4

INSTITUTE OF SYSTEMS SCIENCE, NATIONAL UNIVERSITY OF SINGAPORE

Survey Form:	21
Survey raw data:	
Survey cleaned data:	23
Survey Google analytics report	
APPENDIX B: CAR RECOMMENDATION SYSTEM USER GUIDE	
Appendix C: Sample Scenario	
APPENDIX D: GLOSSARY	28

Executive Summary

We, Singapore citizens and residents, are known globally for our 5 C desires- car, condo, cash, credit card, and country club. Car ownership in Singapore for Singaporean Residents is a cherished aspiration. [Ref 1] captures the sentiment in international press. Closer home, Singaporean love affair for car ownership is evident from weekend newspaper advertisements.

However, Car ownership is a costly affair for us, primarily because of the need to control traffic, ranking Singapore as one of the costliest places to own and maintain a car. [Ref 2]

Our "love" affair with car ownership is so skewed that MAS (monetary authority of Singapore) has to step in to enforce financial prudence in Singaporeans. [Ref 3]

A wide range of choices of car brands offering a bewildering range of possibilities across ride, comfort, safety, gadgetry etc are available to Singaporeans. This is thanks to our open and services driven economy.

Above all aspects, makes the buying decision process a very confusing and suboptimal affair, leaving much room for improvement.

Our project, New car buy recommendation system, is an effort to improve the car buy decisioning system in Singapore. Our recommendation system today suggests car budget and car make/type to look for, thus narrowing down the choice process quite significantly.

We used a combination of expert interviews as well as carefully designed surveys to define a set of rules that help make optimal decision on car buying. To capture the fuzziness in buying decision, we reinforced thus defined/mined rules with the concept of certainty factor.

We used the KIE for user input capture, KIE DROOLS for decision rule, and additionally Orange pack as well as Excel, Tableau and Google sheet analytics for offline data exploration and data mining.

The solution we have presented is a first attempt to bring objectivity and optimality to car the decisioning process. Hence, we prepared a roadmap to capture further objectivity and augmented intelligence in car buying process.

Our achievement is happy Singapore car buyers, where they find the best ride suited to their needs and aspirations that the money they can realistically afford!

Problem

Definition

Which Car should I buy? Car Features, Japanese/European, fuel efficiency, second car, family needs are just some of the fuzzy questions that a car buyer goes through? What is the best decision considering various individual, family, societal needs, financial constraints Vs car ownership costs.

Above is the predicament household question in Singaporean car buyer minds. Unfortunately, no clear and robust decisioning system at hand.

Hence, we formulated the problem statement, thus:

A prospective car buyer weighs in various questions like above, and makes a subjective decision which we observe is not the most optimal.

Description

To further describe the problem involves an appreciation of following aspects for the problem space:

- 1. The buying process is subjective and hence fuzzy. The concept of Certainty Factor has to be introduced.
- 2. There is sufficient synthesized human knowledge on car buying that must be captured as part of decisioning system
- 3. Buying process has "crowd wisdom" and nuggets of knowledge have to be unearthed. Given that the buying process entails whole population of Singapore, hence surveys are clearly must and needed.

Scoping for the project

Inclusions

Business

- 1. Only Singapore market car buying behaviour and decisioning constraints are captured. Application formulation does not apply to other countries.
- 2. Singapore New car buying process is captured
- 3. Today's financial regulations on car loan
- 4. Expert interview and survey data mining

Technical

- 1. Application relies on making strong inferences from the survey and single expert interview (this is clearly an inferential limitation)
- 2. Application is offline today, based on KIE rules engine
- 3. Offline inclusion of survey data mining results

Exclusions

- 1. Not recommending individual model make (E.g. Toyota Camry 2.5 Hybrid Vs Mercedes Benz C-Class 180 Avantgarde). Models pricing and Model inventory/life cycle are volatile.
- 2. resale Car buying process
- 3. web interactive and public domain
- 4. Differentiation between car owners and prospective car buyers not made

Please review the roadmap recommendations to address some of the exclusions listed above.

Solution

The solution involves a three-step knowledge definition and synthesis process:

- 1. Acquisition & Discovery
- 2. Structuring
- 3. Modelling & Representation

Knowledge Discovery & Acquisition

Knowledge Acquisition constituted following 3 step process:

- a) Step 1: Self-study and self-articulation: There are two areas of knowledge acquisition here:
 - a. Individual and family/Friends buying experience and post facto (post buying challenges). This formed the dimensions of decision making process, the dimensions nature whether definite or fuzzy contributors to buying logic. This step also involved reviewing government rules and regulations regarding loan eligibility [Ref 3] and minimum cash down for car buying.
 - b. Leading car portals: [Ref 5]. We leveraged SGCarmart portal and paid special attention to their car comparison and car catalogue pages. These give curated car buyer knowledge acquisition process
- b) Step 2: Expert feedback: Equipped with our understanding and appreciation of car buying process, its dimensions and challenges, we sought feedback of a Subject Matter Expert. [Ref 4]. We conducted an in-depth interview conducted with Jay, who carries a sound decade plus continued experience in Singapore car buying market as a senior new car sales manager. Synthesizing the interview outcomes, we:
 - a. Confirmed evident rules in step 1
 - b. un-earthed un-conventional buying rules, amounting to 100% certainty factor. We also identified "grey" buying areas in our interview. Besides, we got a good understanding on finances and finance behaviour related to car buying process. E.g. women buyers predominantly look for fuel-efficient cars.

Row Labels	Sum of Fuel	Count of Gender	
Female	15	16	94%
Male	30	42	
(blank)			
Grand Total	45	58	

Table 1: Female drivers a CF_features_economy = 1

(see table 3 for the whole list of CFs)

c) Step 3: End-user Survey: Having understood key matters related to car buying process, and also knowing grey areas in our thus acquired knowledge above, we went to survey creation. Sample Survey questionnaire is in [Ref 6]

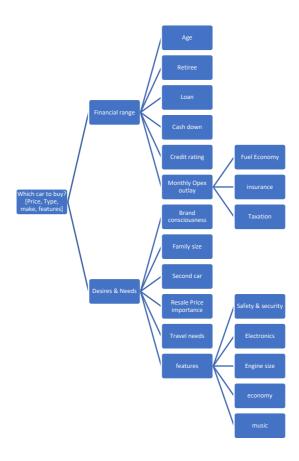
In Summary, our project's knowledge acquisition tree looks like below:



(Figure 1: Knowledge Acquisition for car buying recommendation system)

Knowledge Structuring – Dependency diagram

Synthesized Knowledge from Knowledge discovery and acquisition process is represented by dependency diagram (figure 2) below.



(Figure 2: Dependency diagram: how to buy a new car in Singapore)

Knowledge Modelling & Representation

Buying budget behaviour and calculation: Derived Rules

Car budget recommendation is based on knowledge rules extracted from 3 step processes of Knowledge discovery and acquisition. Here following facts are important for defining the recommendation on car buying budget:

- Singapore regulation stipulates a maximum 70% loan. Min 30% must be cash down
- Financial lending companies and car companies ecosystem brings more incentives if a minimum 30% percentage of car value is taken as loan
- Younger people generally prefer maximum loan
- Older people generally prefer no to minimum loan

Accordingly, we defined, confidence factor (CF) to model end buyer fuzziness in financial clarity Vs the observed behaviour in field and government rules. The CF defined are:

- 1. CF to take 70% loan = CF 70 loan
- 2. CF to take 30% loan = CF_30_loan

Based on user inputs, CFs are updated based on a combination of survey and expert rules.

For the three identified age brackets, the mappings are as follows:

	, 11 6	,	
Scenario	Customer age group	CF_70_loan	CF_30_loan
	mapping		
Scenario 1	Age > 40 & Age < 50	0	0
Scenario 2	Age > 50	-1	8.0
Scenario 3	Age < 40	1	-0.5

Table 2: Mapping of Age to CF 70/30

Associated car budget logic for each of the scenarios:

Scenario 1:

```
IF

(CF_70_loan == 0 && CF_30_loan == 0)

THEN

Price_Max = Cash_Max/(1- loan_max)

Price_Min = Cash_Min/(1-loan_min)
```

Where Max/Min represents user inputs/gut feel on max min loan and cash down

Scenario 2:

Certainty factors

As discussed in earlier sections, the buying process is inherently fuzzy and is inclusive of multiple factors. Hence, we have defined the following variables to capture the fuzziness.

Certainty Factor (CF)	Explanation (leaning towards)	min	max	initialized to
CF_30_loan	taking 30% loan	-1	1	0
CF_70_loan	taking 70% loan	-1	1	0
CF_Make_EUR	a European make car - e.g. Mercedes	-1	1	0
CF_Make_JP	a Japanese make car - e.g. toyota	-1	1	0
CF_Make_MY	a malayisa make car	-1	1	0
CF_Make_KOR	a korean make car	-1	1	0
CF_Make_American	an American make car	-1	1	0
CF_Type_Hatch	a hatchbatch e.g. Honda Jazz	-1	1	0
CF_Type_SUV	an SUV e.g. Subaru forester	-1	1	0
CF_Type_MPV	an MPV like Honda Stream	-1	1	0
CF_Type_Sedan	a Sedan e.g. Toyota camry	-1	1	0
CF_Type_Stationwagon	a stationwagon e.g. Volvo V60	-1	1	0
CF_features_SS	safety and security leading car	-1	1	0
CF_features_economy	higher fuel economy car	-1	1	0
CF_features_engine	higher engine performance i.e. high Torque,	-1	1	0
CF_features_gadgets	electronic gadgets like vision, car parking aid	-1	1	0
CF_features_music	higher quality music instruments	-1	1	0

Table 3: Certainty factor for the car recommender system

User inputs:

Car recommendation system end user inputs are captured as follows:

Input	Description	Data type	Initiatized to	Values
Age	Age of buyer	numeric	18	18-80
Gender	Gender of buyer	Binary	0	M = 1, F= 0
Num_People_family	Number of family members in the hous	Numeric	0	
is_second_car	is the car being bought second car	Binary	0	0= 1st car, 1= 2nd car
Cash_min	Min cash user thinks has for car buy	Numeric	0	0 to 500000
Cash_max	Max cash user think has for car buy	Numeric	0	0 to 500000
Loan_min	Min loan user think shall need for car be	percent	0%	0%-70%
Loan_max	Max loan user think shall need for car b	percent	70%	0%-70%
Weekly_Usage	Weekly car usage	Binary	0	0 = low usage, 1 = high usage
Brand_Consciousness	Is brand consciousness a factor in decisi	Binary	0	1=brand conscious, 0 = not
monthly_car_expenses	Monthly expense include fuel, maintena	Binary	0	1= high, 0 = low

Table 4: User inputs table for the car recommender system

Car recommendation system application outputs:

Car recommendation system outputs are captured by following table:

Defined Rules

Based on Expert interview and Survey results (Certainty Factor), below diagram's capture the evolution of the CFs based on inputs (table 4) given by an end user of the car recommendation system application.

First, a few examples of user survey derived Certainty factors:

Example 1: user survey derived CF recommendation

Sum of Fuel	Column Labels								
Row Labels	American		European	Japanese	Korean	Others	(blank)	Grand Total	
25-40			8	7	4	5		24	
40-50		1	2	9	3	2		17	
50 and above			1	3				4	75%
(blank)									
Grand Total		1	11	19	7	7		45	

Table 5: >50 years old users car make preference

Example 2: user survey derived CF recommendation

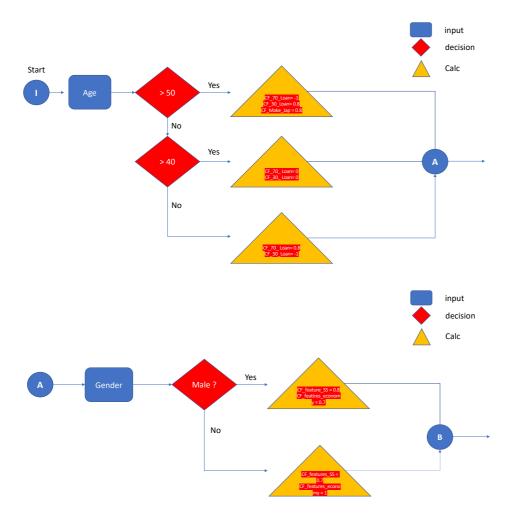
Older people prefer Japanese, resulting in derived rule CF_Make_Jap = 0.8 (see detailed flowchart in tab

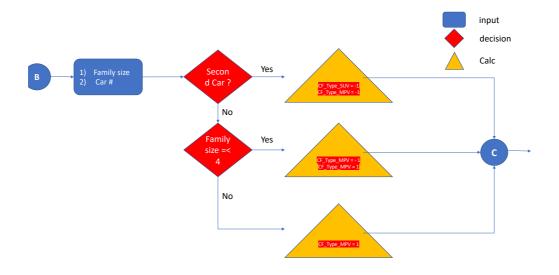
Row Labels	Sum of SS	Count of Gender	
Female	11	16	
Male	36	42	86%
(blank)			
Grand Total	47	58	

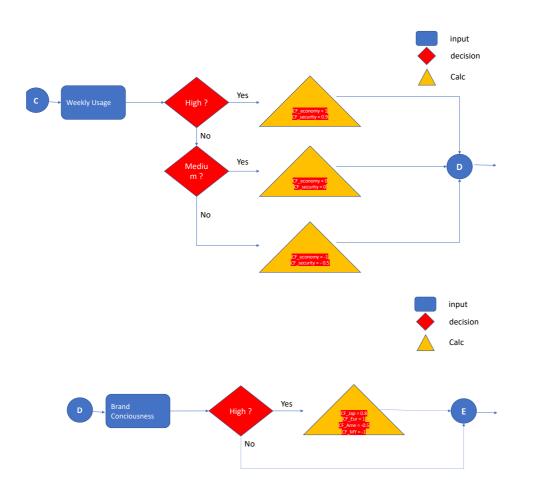
Table 6: Males prefer safety and security.

Refer table 1 for female preference

Finally, here is the flowchart for user input based CF influence.







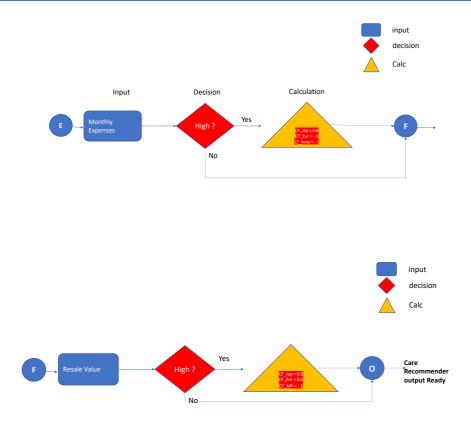


Table 7: Flowchart for Rule Definition engine to update the Certainty Factor

Finalized CFs will then generate the recommendation to the car recommendation system application end user.

Car recommendation system Outputs

Ouput	Output bracket	Possible values	Comments
Price	Price Min	Minimum car pric	car price bracket
	Price Max	Maximum car price	car price bracket
Car country make	Country of car make	Japenese	In order of priority
		European	in order of priority
		Korean	in order of priority
		American	in order of priority
		Malaysian	in order of priority
Car Type	Type of car	Sedan	In order of priority
		SUV	In order of priority
		MPV	In order of priority
		Hatchback	In order of priority
		Stationwagon	In order of priority
Car Features	Recommended car features	Safety and securit	yes no
		Fuel Economy	yes no
		Engine	Yes no

Table 8: Output setup for the car recommendation system (in current version)

System Design:

In today's MVP stage, the system logic is defined in KIE. This includes definition of the

- User input form (as per the user input table defined in Table 4)
- Task definitions and Data Objects
- Process flow including input corrections/data correction
- And finally guided decision tables to provide the end recommendations (as per table 8 for car recommendation system output)

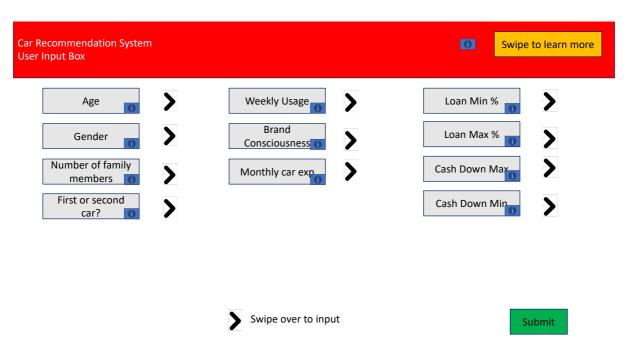
System Architecture considerations:

There are no system architecture considerations as we are in MVP stage. For the application to be launched to end users, following developments have to be defined:

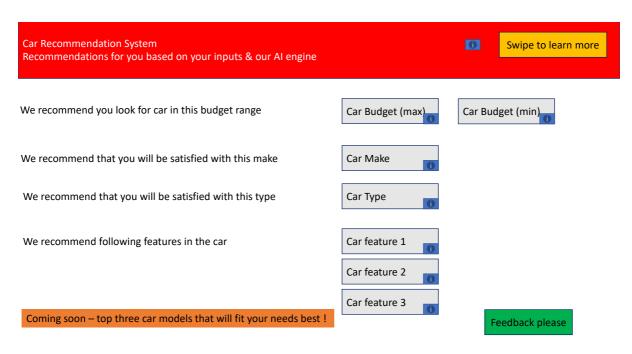
- Mobile and web interface for user interactivity (Input and Output)
- Web server hosting (for the MVP functionality and logic)
- Security provisions for unauthorized access, DDoS etc attacks

Inputs & Output dialogue screens

User Input Form



Recommendation Screen



Project Critical Analysis

Key Survey and Expert Findings

- Housing type does not matter in car selection as was based on expert feedback
- Gadgets and Music does not matter much to the survey population. We believe that
 it's due to a reasonably OK music system in most of the cars in market, and is not a
 buying factor anymore.
- Women overwhelmingly look for 'fuel economy' of car
- Men overwhelmingly look for 'safety and security' in the car.
- Retirees overwhelmingly look for Japanese car
- The application does not consider the working status because it was explicitly suggested by expert that it does not influence car buying behaviour. However, detailed data based studies can be further taken to explore alternate hypothesis.
- Most of the buyer base is 40-50 years bracket
- Retirees mostly want to buy car with full cash down payment
- Young overwhelmingly take maximum loans

Functionality and system short comings, pending issues

- Distinction of European cars wide range of quality difference between European car brands
- Distinction between sedan categories ideally luxury and mass market sedans are different markets
- specific tastes like Racing, Vintage cars
- Resale value of the car was not captured in the survey. While heuristics is available,
 i.e. Toyota and Mercedes fair well on resales value in Singapore market, since there
 was no supporting data from survey we have excluded it from user inputs. However,
 based on user inputs, we factor based on calibration.
- Next best alternative (NBA), top car make, type suggestions not implemented. E.g.
 Toyota Camry 2.5 Hybrid is the top recommendation based on inputs and knowledge
 decisioning system (i.e. Car recommendation system)
- Weekly/Daily fluctuations due to COE, flash sales
- OPC (off-peak car) or full COE based recommendations
- Ability to give meaningful recommendation with users providing only partial information
- System handles only male and female preferences (Transgender and other social /family structure preferences not curated)
- For retirees, we have not considered loan guarantors for loan. This is based on expert advise as well that most (~100%+) do cash down buying.

System Limitations

• Lacks Visual appeal due to using KIE forms.

- User interactivity limited to capability afforded by KIE forms.
- System can't go live online (public internet domain) due to KIE capabilities and Public domain/public server hosting and budgets
- Inputs that shall have a finite range to represent, hence choices are made binary (e.g. security preference that can range from low, medium, high or on a scale of 1-10)

Future Project Enhancements

Scope Enhancements

- Singapore Used car market inclusion
- Expanding recommendations from car make, type, features, to top matching 3 cars (NBA) available in the Singapore market at that given point in time
- Expanding to nearby countries, e.g. South East Asian markets

Functional Enhancements

- Representative surveys for Singapore car buying target population
- Machine learning module augmentation to identify
- Additional experts feedback
- Hypothesis testing scenarios

Technical Enhancements

- Launching the mobile interface and web server (AWS) with a scale up on demand backend i.e. AWS laaS/PaaS.
- Machine Learning module for offline integration based on model re-runs (surveys and user feedbacks) with AWS sagemaker.
- Integration with car dealers (e.g. C&C, Borneo motors etc) and car information aggregation websites like sgcarmart.com

References

[Ref 1]: Singaporean love for cars

http://money.com/money/4716340/worlds-most-expensive-city-cost-of-living-carownership/

[Ref 2]: Singapore is consistently the costliest place to buy car

https://www.startrescue.co.uk/news/motoring-news-2/10-most-expensive-countries-to-buy-and-run-a-car

https://www.sgcarmart.com/news/article.php?AID=10371

[Ref 3]:

Singapore government regulations on car loans

http://www.mas.gov.sg/News-and-Publications/Media-Releases/2016/MAS-Eases-Rules-on-Motor-Vehicle-Financing.aspx

http://www.ifaq.gov.sg/MAS/TOPICS/BANKING/Vehicle Loans/7419#FAQ 176191

[Ref 4]:

Expert interview: Jay, Singaporean Car Sales manager with more than a decade of successful car selling experience in Singapore. Audio transcripts submitted.

[Ref 5]:

SGCarMart.com: is the leading Singapore car information portal.

[Ref 6]:

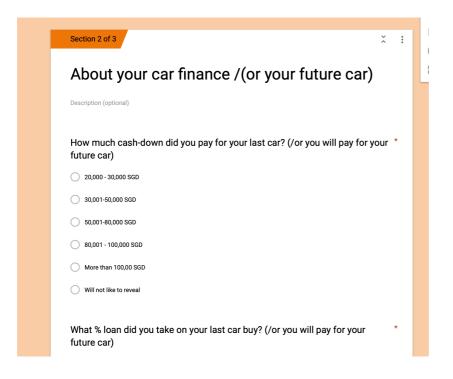
Survey to collect end buyer behaviour https://goo.gl/forms/Frt1Rh3q9ImSXKkY2

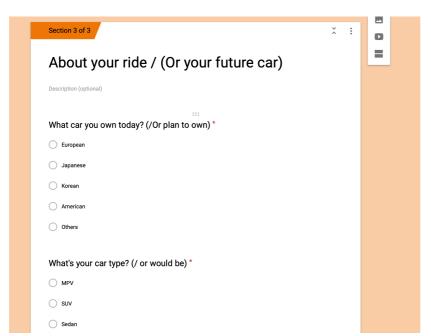
Appendices

Appendix A: Survey

Survey Form:

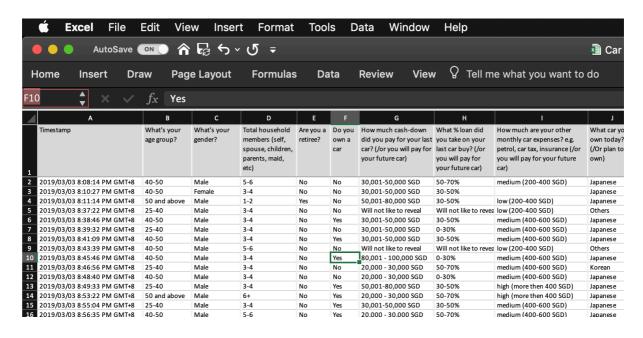




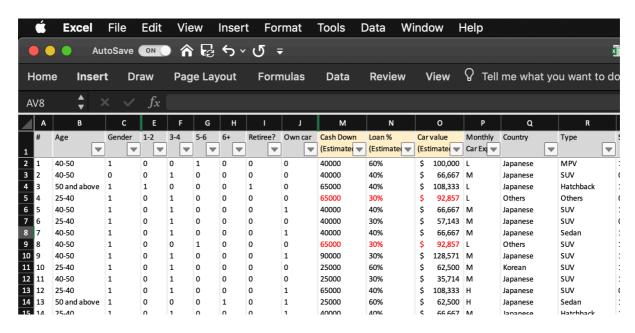


Survey raw data:

Details attached in project zip folder



Survey cleaned data:



Details attached in project zip folder

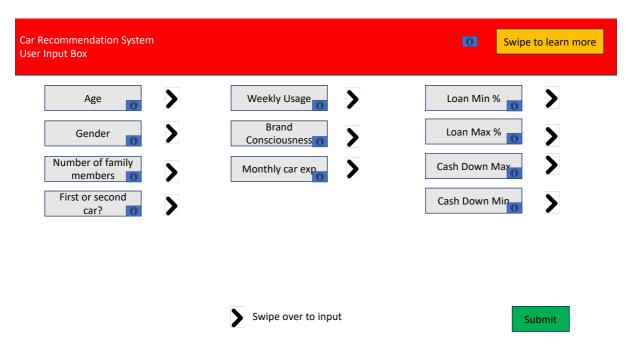
Survey Google analytics report

Details attached in project zip folder

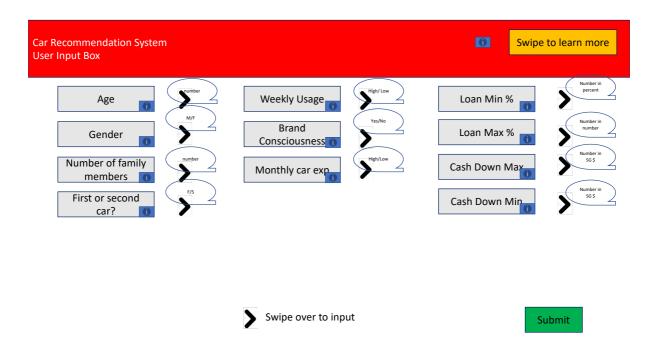
Appendix B: Car Recommendation System User Guide

User is guide to a dialogue box upon open the application on mobile or through a public URL As discussed previously, the MVP is planned for KIE only. Functionality has to be exposed through a web API.

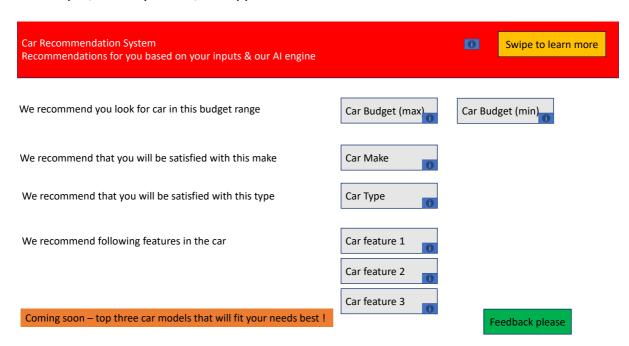
The input screen will appear like below:



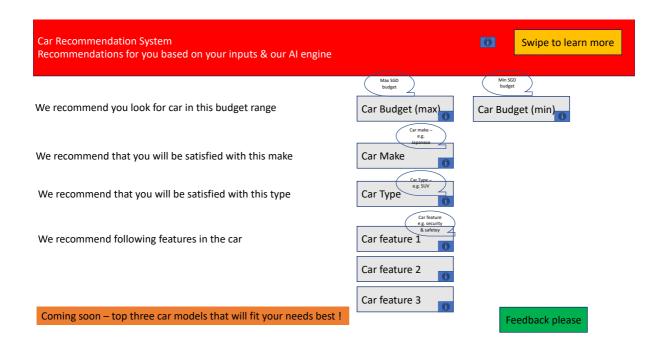
At the input screen, please input all the input fields by swiping on the forward arrow mark. The inputs are either true/false or based on some numeric range, like highlighted below!



The output, in today's MVP, will appear like this.



Point towards the information icon will provide additional details about a recommendation item.



Appendix C: Sample Scenario

Scenario: Value buyer, aged 30, male

A car buyer with following requirements captured in the user input form, ref appendix B for the user's guide, for user input dialogue box in mobile app/browser.

Input	Description	Data type	User inputs
Age	Age of buyer	numeric	30
Gender	Gender of buyer	Binary	Male
Num_People_family	Number of family members in the hous	Numeric	2
is_second_car	is the car being bought second car	Binary	No
Cash_min	Min cash user thinks has for car buy	Numeric	25000
Cash_max	Max cash user think has for car buy	Numeric	30000
Loan_min	Min loan user think shall need for car bu	percent	50
Loan_max	Max loan user think shall need for car b	percent	80
Weekly_Usage	Weekly car usage	Binary	1
Brand_Consciousness	Is brand consciousness a factor in decisi	Binary	1
monthly_car_expenses	Monthly expense include fuel, maintena	Binary	0

Results in decision process calculation, ref Table 7, flowchart for the rule definition engine to calculate CF

Input	Description	Data type	User inputs	Basis CF calculations
Age	Age of buyer	numeric	30	
Gender	Gender of buyer	Binary	Male	CF_SS == 1
Num_People_family	Number of family members in the hous	Numeric	2	CF_Type_SUV/CF_Type_MPV == -1
is_second_car	is the car being bought second car	Binary	No	not relevant for CF influence in this case
Cash_min	Min cash user thinks has for car buy	Numeric	25000	\$83,333
Cash_max	Max cash user think has for car buy	Numeric	30000	\$100,000
Loan_min	Min loan user think shall need for car bu	percent	50	CF_30_loan == -0.5
Loan_max	Max loan user think shall need for car b	percent	80	CF_70_loan == 1
Weekly_Usage	Weekly car usage	Binary	1	CF_features_economy == 1
Brand_Consciousness	Is brand consciousness a factor in decisi	Binary	1	CF_European == 1, CF_Jap == 0.8
monthly_car_expenses	Monthly expense include fuel, maintena	Binary	0	CF_European (updated) == -1, CF_Jap == 0.8

The car recommendation system, as defined in Appendix B, user guide will make following recommendations. With these recommendations, the budget is narrowed down, the car make, type and features narrowed down as well.

Ouput	Output bracket	Possible values	Recommendations
Price	Price Min	Minimum car pri	c \$83,333
	Price Max	Maximum car pr	ic \$100,000
Car country make	Country of car make	Japenese	Japanese
		European	No
		Korean	Does not matter
		American	No
		Malaysian	No
Car Type	Type of car	Sedan	Yes
		SUV	No
		MPV	No
		Hatchback	Yes
		Stationwagon	
Car Features	Recommended car features	Safety and secur	it <mark>Yes</mark>
		Fuel Economy	Yes
		Engine	Does not matter

Once we link the app to current market models and pricing, we can zero down to best 2-3 options available (future enhancement captured in Section Future Project Enhancements)

Appendix D: Glossary

MAS : Monetary Authority of Singapore

LTA: Land Transport Authority of Singapore

COE : certificate of entitlement. Granted by LTA Singapore. Each car must carry the COE.

COE is a major component of car price.

TCO: Total cost of ownership
EMI: Equal Monthly instalment

CF : Certainty factor BHP : Brake Horse Power

MVP : Minimum Viable ProductKIE : Knowledge is everythingDDoS : Distributed Denial of Service

C&C : Cycle and CarriageNBA : next best alternativesAWS : Amazon Web ServicesIaaS : Infrastructure as a Service

PaaS : Platform as a Service SUV : Sports Utility Vehicle MPV : Multi Purpose Vehicle