Laptop Recommendation System

Group members:

CHEN JIAJUN
LI YUAN
WANG DING
ZUO ZHEN

2020.11.1

CONTENTS

1 EXECUTIVE SUMMARY

The computer has been the 21st century and work and learning inseparable, closely related tools. The requirement of replacing the computer is long-term, because of the iteration of the hardware, the iteration of computer products, and the complexity of software, the requirement of computer configuration is also improving. We chose an application scenario to think about, that is, freshmen buy computers. But can students choose their laptops correctly? As we know, computer configuration is mainly CPU, graphics card, hard disk, and then some features such as screen, keyboard, interface, design, and so on. But computers at the same price often have different priorities, and how to make decisions that are more in line with their needs after budgeting is a difficult problem for those who have little knowledge of computers. There is also a problem, even people familiar with computers, it is difficult to identify their needs and make smarter decisions.

We want to solve this problem, so we plan to build a computer-based recommendation system. We crawled the computer notebook data from the http://www.zol.com.cn, cleaned and deleted the data, and SQLitE3 established the computer information database. We used vue.js + vuetify.js UI for the front end design.

We use vue.js as a page building framework, and adopt a vuetify.js UI framework, save time for page structure design and UI design, fast implementation of a simple multi-page APP. Including language switching, questionnaire filling, output, and other functions... (currently used locally in vue.js app), parse rule files using a self-developed rule engine (python). txt) Build a rule base from user preference settings to laptop configuration. Backstage uses Flask framework to build interfaces, Using python post method to resolve user preference data sent back by the front end, And derived preference information by the rule engine, Get the laptop recommended configuration list. After matching through a background SQL - based matching system, Returns the details of several computers with the highest matching score to the front end in JSON format.

Our team works very happily on this project and hopes that the project can be used and feedback by users. This project can become more international, if we can find a more standardized notebook computer shopping site, enrich the database, Maybe we can identify the differences in more goods and recommend them to users more accurately.

2 PROBLEM DESCRIPTION

I believe most people don't have a good experience in buying a computer. The purchase behavior may be due to brand loyalty that causes consumers to pay attention to their new products, or because a feature in the advertisement moves the consumer, or because the consumer sees a particular feature of the product, such as a keyboard, screen, etc. But in this way, consumer demand is not satisfied.

Even in the current market, some computer sales, evaluation company consultants, and so on will provide advice to consumers or share commentary videos, but often biased, misleading, and lack enough two-way communication. Or some sales and consultants cannot really understand consumer needs, so to some extent, consumers are difficult to make unrepentant shopping behavior. On the other hand, even if consumers understand their needs, they need to search for and compare configuration information for different notebooks.

2.1 PROJECT OBJECTIVE

We hope to build an expert system to eliminate the bias of salespeople and evaluation agencies as much as possible, but to absorb their more relevant professional knowledge so that users with vague needs can obtain recommended products suitable for themselves. Solve this aspect of computer knowledge lack of users to buy and purchase problems. We hope that through the questionnaire design, the complex, computer configuration issues into a more life-oriented dialogue, user-oriented issues. For example, we transform the related problems of CPU, graphics card through computer use, whether we often come out to transform the problem of endurance and configure the high-performance problem through the pursuit of game experience. We will guide customers to tap their own needs, verify their own needs, and even prioritize their needs (because the budget is fixed). When many needs of user conflict, the system can recommend the products closest to their needs under the guidance of expert advice. The richer the data about notebook computers in the system database, the better.

Considering the iteration of the computer, our database only contains notebook computer products for nearly two years. Including a variety of brands, in the implementation of the recommended function is sufficient. Commercially, we hope that the system can be deployed on the web page, users can use it directly through the access link, and finally through the output link of the recommended product to understand and buy the product carefully. In this way, the system can also obtain some profits through distribution channels. And after completing the questionnaire and obtaining the portrait, we can accurately put the corresponding computer advertisement in the output interface.

3 KNOWLEDGE MODELING

Knowledge modeling can be decomposed into three main stages
(1) Knowledge identification
(2) Knowledge specification
(3) Knowledge Acquisition

- Schreiber, et al.,2001

Various activities are carried out during each of these stages and the crux of model construction lies in.

3.1 KNOWLEDGE IDENTIFICATION

Data acquisition is very important, we hope to get standardized data so that we can directly obtain structural specification data, a better expression of knowledge, so we climbed from the shopping website computer commodity-related data, At the same time from the evaluation agency website to get computer recommendation related recommendation knowledge.

Source of information	**Insights from information source**	**Knowledge acquisition technique**
E-shopping website (JD, ZOL) It provides general information on computer	Brand, price, disk, GPU, CPU, Color, weight, Characteristics of available feature, thickness, battery life	Web scraping to obtain current computer information.
Rating agent of computer (Dong Li/Zhongzhen/Da li/Lao Ba/Renrating agent)	Spec required by GAME PC/STUDY PC/WORKING PC Brand level rating comprehensively Spec contrast in CPU, GPU, Disk, Screen, Price	By go through the video and paper published by the agent

3.2 KNOWLEDGE SPECIFICATION

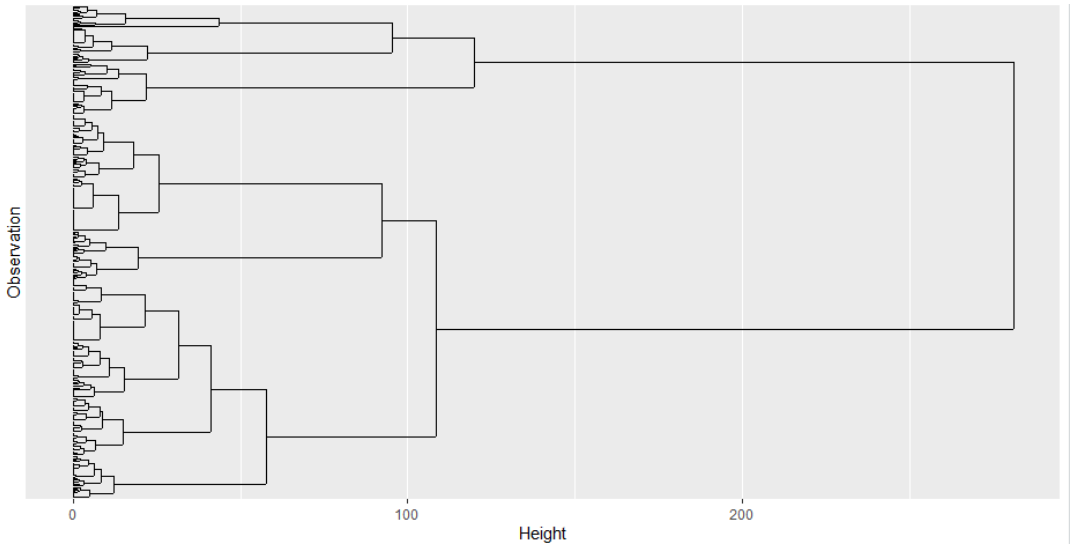
Data cleaning from structured data, data mining (clustering) rules, and weight calculations from expert recommendations (weights)

We need to use knowledge to build databases, bring insights out from the data, design rule files, formalize problem-solving domain knowledge. in this section We use the existing data to build the background database of the system. To avoid the bias of the

evaluation organization, we use the method of data mining to automatically find the relationship between the computer configuration.

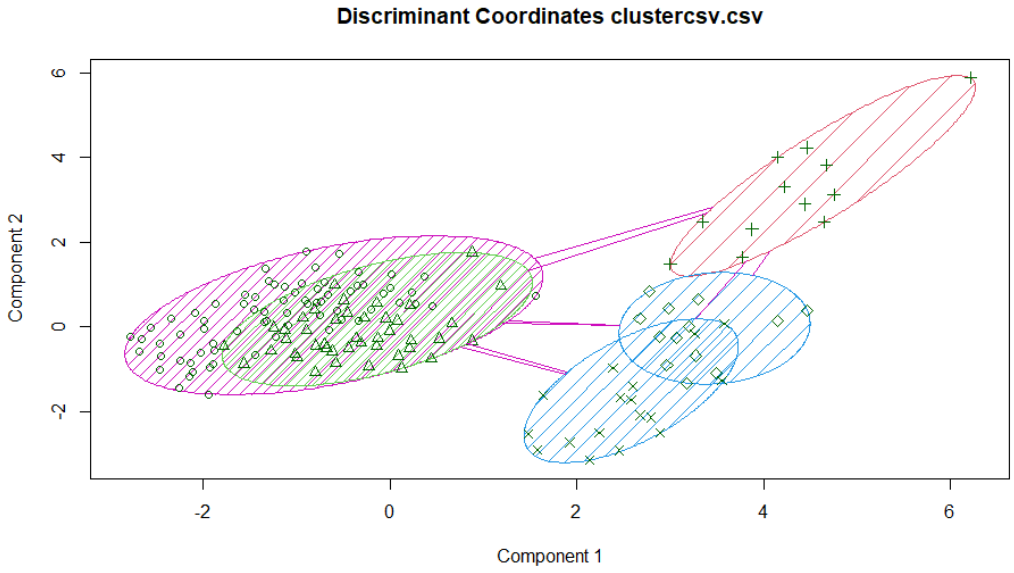
After getting the web crawler data, we found that the data provided by the website is more standardized and structured, but some of the data items have no data, such as price, memory, graphics card, screen resolution. So I collate or delete incomplete data, leaving 500 computer information (enough).

After determining the computer data, we begin to analyze the computer data. After the standard normalized data, we use the hierarchical cluster method to observe.



Pic.1 - hierarchical cluster

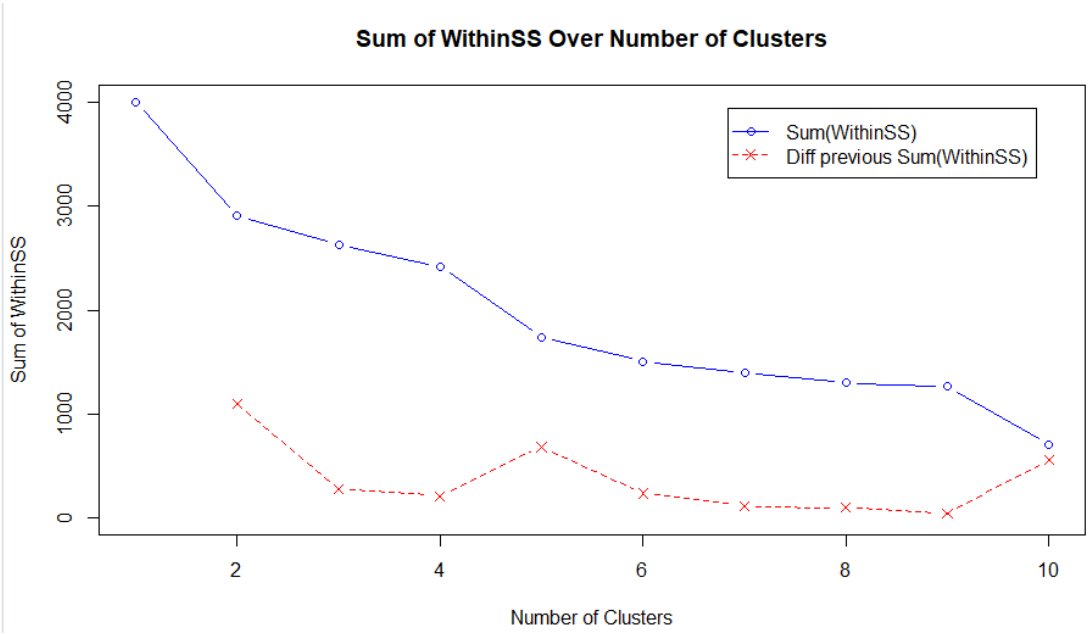
Then we use the k-means clustering algorithm and perform different verification to find the appropriate number of clusters.



Pic.2 - K-means cluster

According to the suggestion, we divide the cluster into a game version, audio version, light version, business version, learning version. By standardizing the data, we adjust

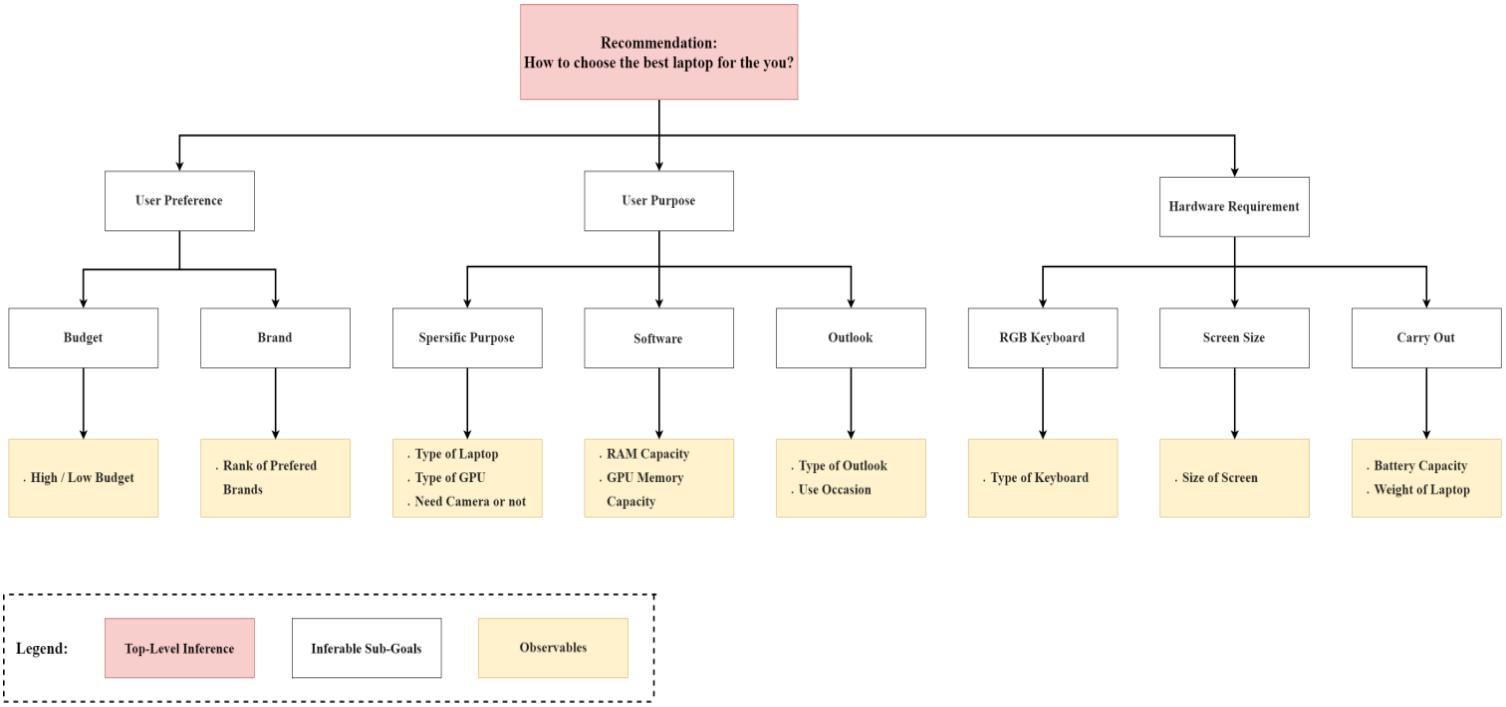
the weights of some feature properly. To make these four clustering more obvious, we also verify that the five classifications have relatively good performance.



Pic.3 - Sum of WithinSS

3.3 KNOWLEDGE ACQUISITION

Knowledge Acquisition - Dependency Diagrams How to Make Rules



Pic.4 - Dependency Diagram

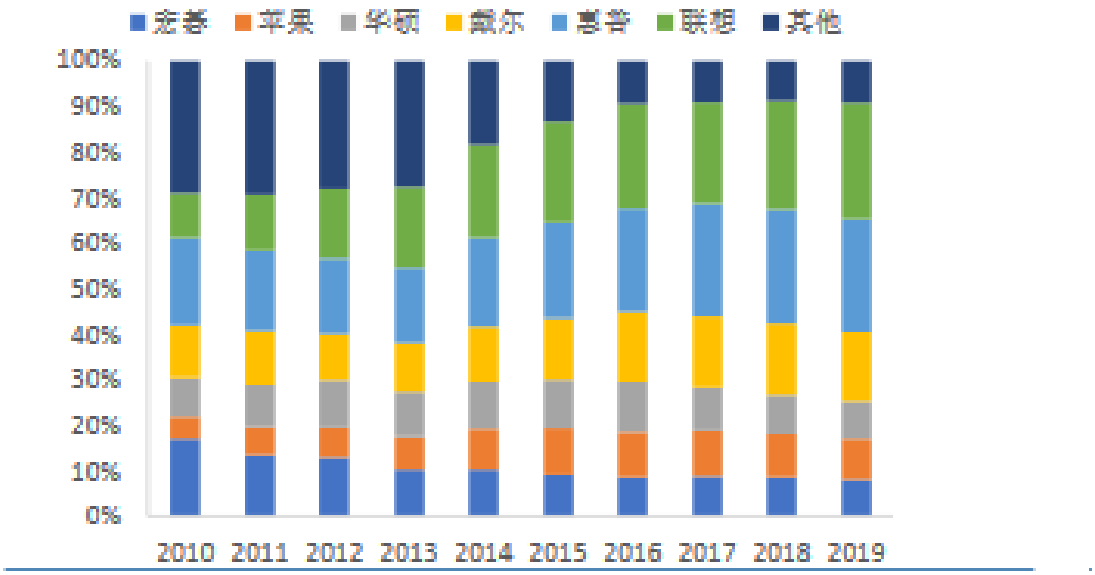
After watching the videos of several evaluation organizations, we conducted a group discussion and got the dependency diagram of systematic reasoning according to the

expert suggestion and market phenomenon.

We divide the computer into four categories: game, office, audio and video, and learning because each category has a different focus on hardware configuration requirements. So we make it easier for the system to recommend computer types that match its category. Budget is often the first choice made by a consumer and plays a vital role in the whole shopping, so we set the price to a relatively high recommended weight.

As we know, consumers have different needs, some of them are more personalized, and experts suggest that they play a small role in personalized needs, so we reach a consensus with users by designing questionnaires. Whether to set some personalized needs to hard constraints, such as brand, screen size, weight, price, battery life.

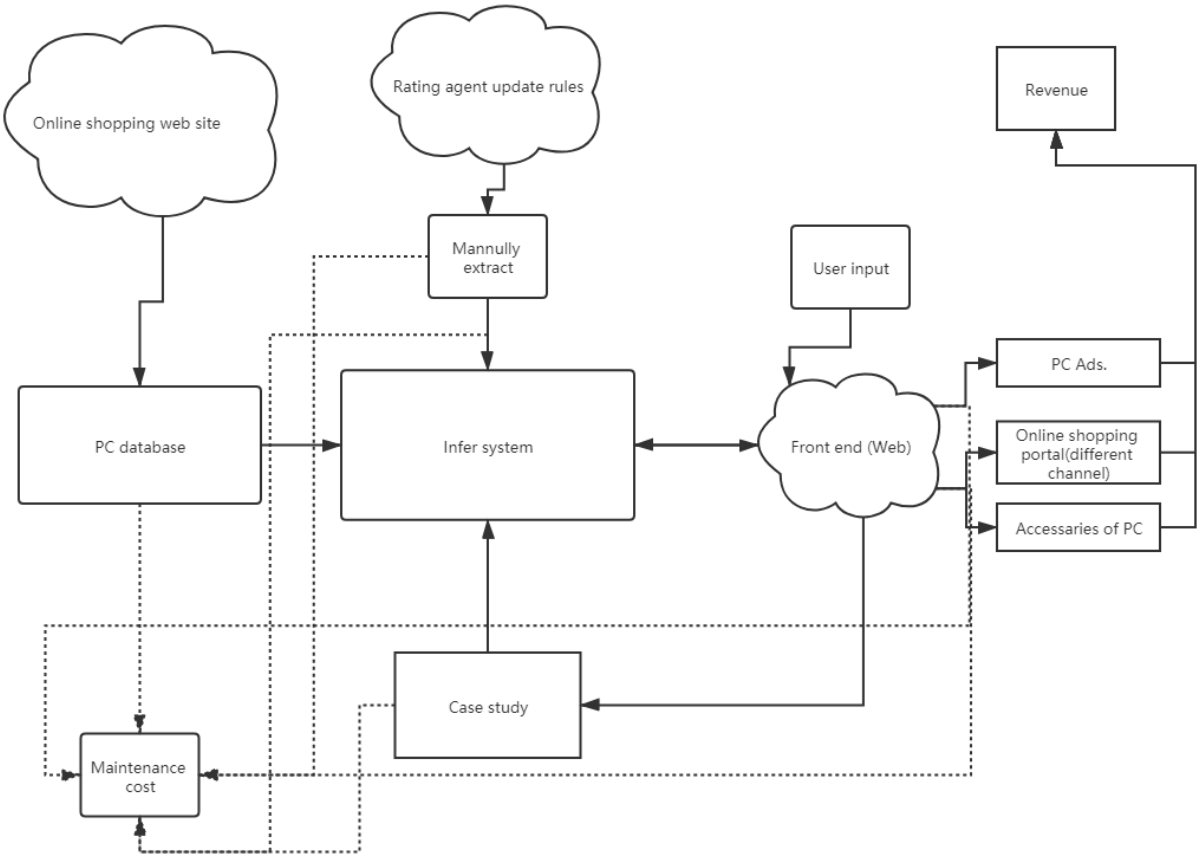
In the analysis of the characteristics of the brand, we have fully considered the Chinese computer market. From this table, we can see that the market is still circulating several leading brands, but other computers still account for a lot of shares. There are a variety of niche brands. Here we take into account a 2020 evaluation report on the notebook brand effect, the first-line brands often have the design ability of notebook mold and have their own factory, the after-sales network spread throughout the country; The second line is more to label the products of other manufacturers to sell; third-line brands often provide higher configuration at lower prices. So according to the evaluation report, we recommend different brands to different degrees.



Pic.5 - Market share by brand

The design of the rule engine is attached for the reference of interested people.

4 SOLUTION OUTLINE



Pic.6 - Commercial design

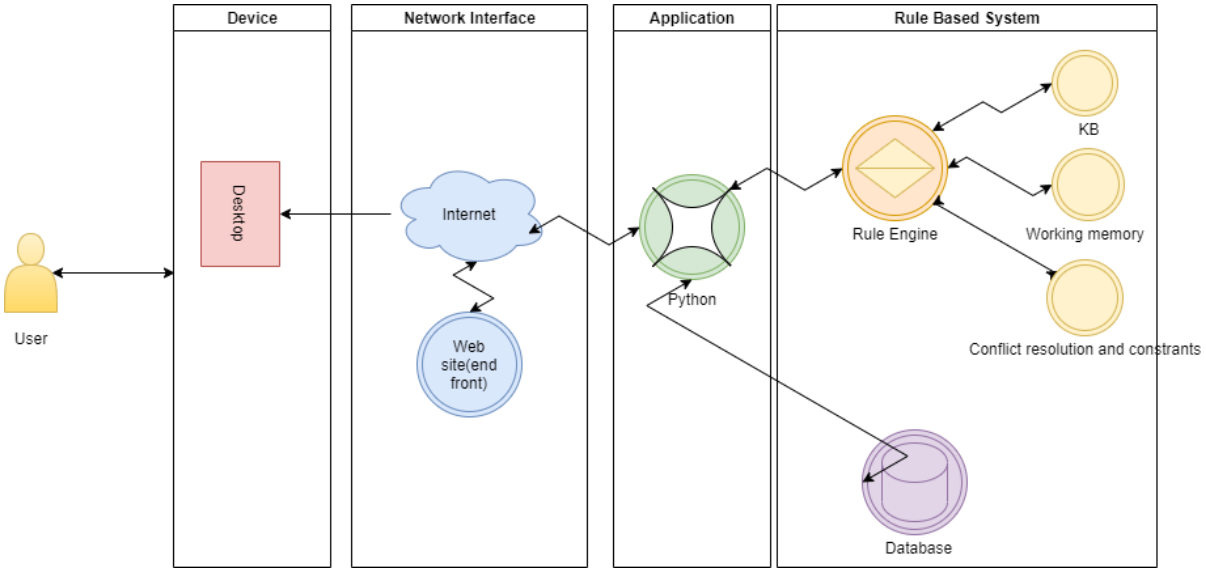
From the block diagram, we can see that there are two external inputs inside the system, one is to evaluate the website knowledge update input to the rule engine, the other is the website crawler crawling data to the knowledge base. The input of these two systems needs to be executed regularly. It belongs to the main cost of the product. User input through the front-end web page to collect user needs in the form of questionnaires, through the reasoning engine to get the results, to the online shopping portal and accessories of PC links, and in the front-end will also display computer advertising, These three will become the main way to revenue the system's profitability: computer advertising, computer peripheral hardware, distribution channels of the E-shopping website. Worth mentioning, the system is expected to record the results of reasoning and learning improvements.

The system is based on the problem-solving strategy of an expert system. Considering that the computer update speed is too fast, the database data is a new computer product in the past two years, which needs to be updated manually. Experts suggest that based on the video and articles of the evaluation organization, we design the rule engine, and design the questionnaire structure corresponding to the rule engine, which also meets the individual needs of users in the universal reasoning mechanism, enables the structuring of a rule-base which specifies the knowledge and reasoning

requirements of selecting a PC unit.

In order to avoid the influence of mechanism bias, we use the clustering data mining method for computer classification.

4.1 SYSTEM ARCHITECTURE



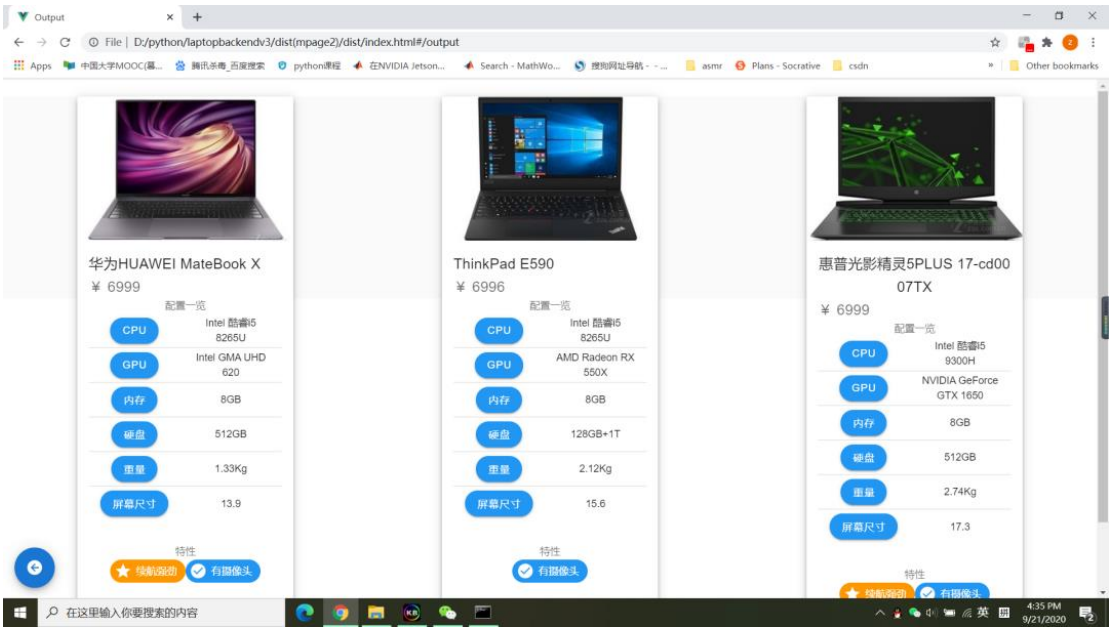
Pic.7 - System architecture

Web (Front end):
We used vue.js + vuetify.js UI for the front-end design.

Vue.js was used as the framework for page building and take vuetify.js UI framework, which saved the time of page structure design and UI design and quickly realized a simple multi-page APP, including language switching, questionnaire filling, output other functions and so on. (Apply in local within vue.js app so far)

Backstage (Backend):
The crawler technology and sqlite3 are used to establish the database of computer information. Use the self-developed rules engine (python) parsing rules file (.txt to build a rule library from user preferences to the computer.

Flask framework is used to build the background interface, python is used to parse the user preference data sent back by the front end using the post method, and the preference information is deduced by the rule engine to get the recommended configuration list of the computer. After matching through the background SQL-based matching system, the detailed information of several computers with the highest matching score is returned to the front end in JSON format.



Pic.8 - Output and interpreter

4.2 PROJECT SCOPE

The scope of the project is divided into four parts:
(1) Need to crawl data and standardize its structure
(2) Need to design expert advice into rule engine
(3) Design Front
(4) Market research for business analysis

4.3 SYSTEM'S FEATURES

A similar computer recommendation system has not been found on the market. At present, some of the more similar artificial customer services in the market will introduce the computer products of the shop through brief dialogue with the users, or some evaluation agencies will directly provide suggestions to introduce the computer products. But obviously, from the functional point of view, these two ways, are more subjective, not objective. But from the consideration of product design, brand, marketing, and other factors, subjective views can also capture the hearts of consumers, evaluation agencies can also make their own community shape.

At present, there are many evaluation websites, but because of its characteristics, there are certain social problems. (Due to the acceptance of promotion fees, and then make biased recommendations and evaluation, bias is too heavy)

Advantages:
(1) Using data mining in the database, the same level of the notebook is excavated directly from the hardware and hardware supporting data, which is more objective and

more suitable for the market.
(2) The expert system is applied, and the relevant information is obtained by a popular questionnaire, and reasonable recommendation is given.
(3) Background combined with the front-end questionnaire, a series of soft and hard constraints are designed to be recommended according to user personalization. Different from the recommendation of the evaluation organization, the system base on functional expert advice and computer market data mining
(4) It can be quickly recommended by importing pages containing computer configuration information.
(5) Low technology cost, fast access (by renting public domain names)

Disadvantages:
(1) Need manual regular update database, data processing.
(2) The expert system mainly starts from the function, the humanization needs to be perfected
(3) The expert system does not support automatic learning to update the rule base for the time being.

Opportunities:
According to statistics, in the first half of 2020, in order to meet the needs of home office learning, the per capita expenditure on household desktop computers and notebook computers increased by 19.4 percent and 79.6 percent, respectively.

- National Statistical Office

(1) Note computer rating agency encounter a depressed time and sharp environment. People are gradually wearing colorful glasses to watch rating agency.
(2) E-commerce channels become more mature, easy access

Improvement:
(1) In the future, we can build an expert database of automatic learning update and update it in real-time
(2) By analyzing the website of the school's professional courses, recommend computers directly

Threats:
(1) The expert system is too functional and may shrink in the marketing age.
(2) Less interaction with users, unable to accumulate users
(3) When computer configuration, we need to consider not only hardware configuration combination but user portrait, preference, product as a whole. At present, the expert system is deficient in this respect.

4.4 FUTURE IMPROVEMENTS

On the rules engine, it can have long lists to show more options, databases are richer (not only Chinese regional brands), experts recommend update the system itself.

Appendix 1: Project Proposal

GRADUATE CERTIFICATE: Intelligent Reasoning Systems (IRS)
PRACTICE MODULE: Project Proposal

Date of proposal: 28 Aug. 2020
Project Title: ISS Project – Laptop recommendation system, for intelligent course scheduling system use case
Sponsor/Client: *(Name, Address, Telephone No. and Contact Name)* Institute of Systems Science (ISS) at 25 Heng Mui Keng Terrace, Singapore NATIONAL UNIVERSITY OF SINGAPORE (NUS) Contact: Mr. GU ZHAN / Lecturer & Consultant Telephone No.: 65-6516 8021 Email: zhan.gu@nus.edu.sg
Background/Aims/Objectives: The proposed Computer recommendation system will make use of various advanced machine reasoning techniques and components to help user to make good decision on purchase PC.
Requirements Overview: • Research ability • Programming ability • System integration ability
Resource Requirements (please list Hardware, Software and any other resources) Hardware proposed for consideration: • GPU, RaspberryPi, AlphaBot, NVidia Jetson Box, etc. Software proposed for consideration: • Reasoning systems, e.g. KIE jBPM, Drools, AppFormer, OptaPlanner, Fuzzy logic, Optimization, etc • Pertained machine learning models, e.g. Vision, Speech, NLP • Machine learning use cases, e.g. Orange3, R

- Deep learning tools, e.g. Neural Network Console Sony, Python Keras
- Chat-bots, e.g. ChatterBot, DBpedia Chat-bot
- Cognitive systems, e.g. MyCroft
- Robotic Process Automation, .e.g TagUI
- Cloud computing/server, e.g. Amazon, Google, IBM, Azure, etc.
- Application container, e.g. Docker

Number of Learner Interns required: (Please specify their tasks if possible)

a team of four project members

Methods and Standards:

Procedures	Objective	Key Activities
Requirement Gathering and Analysis	The team should meet with ISS to scope the details of project and ensure the achievement of business objectives.	1. Gather & Analyze Requirements 2. Define internal and External Design 3. Prioritize & Consolidate Requirements 4. Establish Functional Baseline
Technical Construction	· To develop the source code in accordance to the design. · To perform unit testing to ensure the quality before the components are integrated as a whole project	1. Setup Development Environment 2. Understand the System Context, Design 3. Perform Coding 4. Conduct Unit Testing
Integration Testing and acceptance testing	To ensure interface compatibility and confirm that the integrated system hardware and system software meets requirements and is ready for acceptance testing.	1. Prepare System Test Specifications 2. Prepare for Test Execution 3. Conduct System Integration Testing 4. Evaluate Testing 5. Establish Product Baseline
Acceptance Testing	To obtain ISS user acceptance that the system meets the requirements.	1. Plan for Acceptance Testing 2. Conduct Training for Acceptance Testing 3. Prepare for Acceptance Test Execution 4. ISS Evaluate Testing 5. Obtain Customer Acceptance Sign-off
Delivery	To deploy the system into production (ISS	1. Software must be packed by following

	standalone server) environment.	ISS's standard 2. Deployment guideline must be provided in ISS production (ISS standalone server) format 3. Production (ISS standalone server) support and troubleshooting process must be defined.

Team Formation & Registration

Team Name: Lucky 7
Project Title (repeated): Computer recommendation system
System Name (if decided):
Team Member 1 Name: CHEN JIAJUN
Team Member 1 Matriculation Number: A0215519E
Team Member 1 Contact (Mobile/Email): chenjiajun@u.nus.edu
Team Member 2 Name: LI YUAN
Team Member 2 Matriculation Number: A0215499R

Team Member 2 Contact (Mobile/Email): 88499080/e0535589@u.nus.edu
Team Member 3 Name: ZUOZHEN
Team Member 3 Matriculation Number: A0215464H
Team Member 3 Contact (Mobile/Email): 88511594/e0535554@u.nus.edu
Team Member 4 Name: WANG DING
Team Member 3 Matriculation Number: A0216421U
Team Member 3 Contact (Mobile/Email): Wang_ding20@u.nus.edu

For ISS Use Only		
Programme Name:	**Project No:**	**Learner Batch:**
Accepted/Rejected/KIV:		
Learners Assigned:		
Advisor Assigned: Contact: Mr. GU ZHAN / Lecturer & Consultant Telephone No.: 65-6516 8021 Email: zhan.gu@nus.edu.sg		

Appendix 2: Commercial value

Computer is a must for freshmen every year, so by analyzing the number and consumption of freshmen, we can estimate the consumption scale of freshmen every year, and then estimate the business scale of purchasing computers (about 10 million notebooks, roughly estimated according to the average price of 6,000 notebooks on the market about 60 billion new notebook business, accounting for 1/6 of college students' annual consumption).

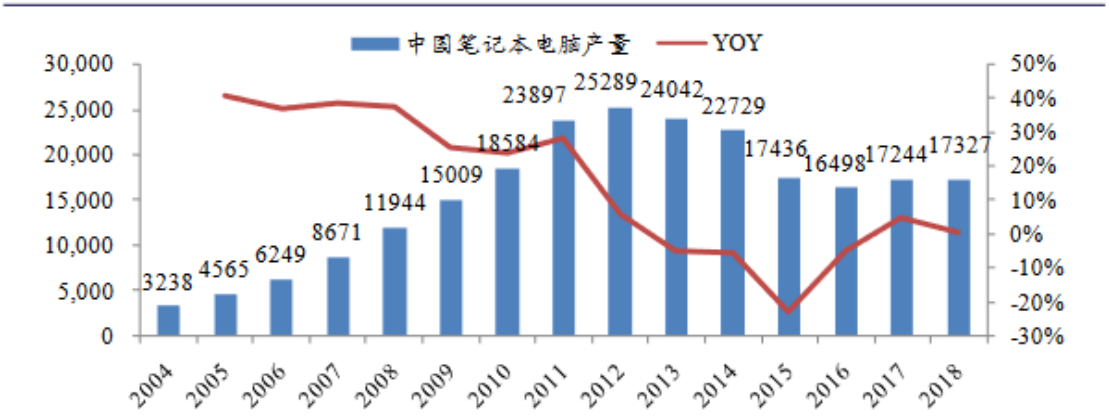
According to the Ministry of Education in China, the number of college entrance examination applicants reached 10.31 million in 2019-Ministry of Education

省/市	2019年	2018年	2017年	2016年	2015年
全国	1031万+↑	975万	940万	940万	942万

Pic.9 - Students number of graduated

According to statistics, by 2018, the daily disposable income of college students is 1405 yuan/month, and the annual consumption scale of college students is 381.6 billion yuan.
- a securities firm

According to the open market research report, the output of laptops in 2018 was 170 million, and it is forecast to reach 190 million by 2022.
- a securities firm



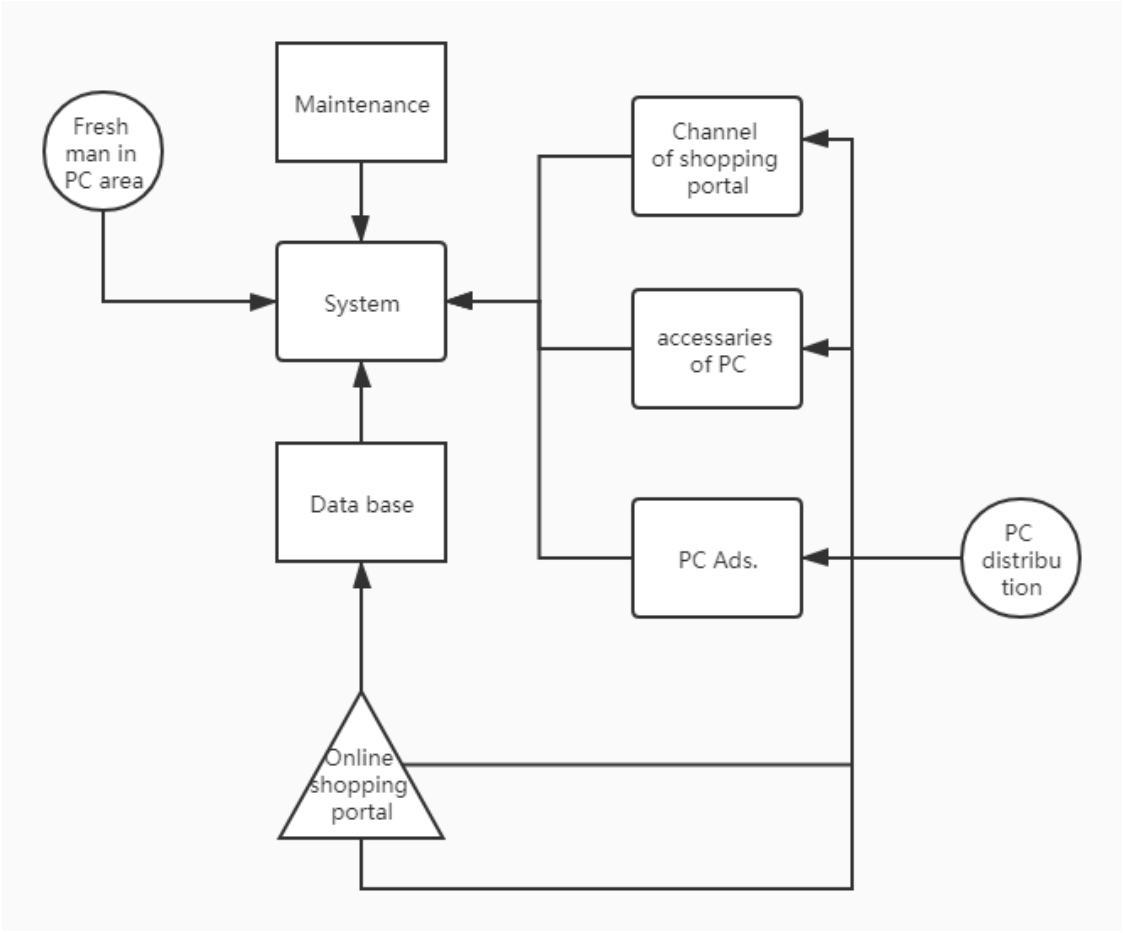
Pic.10 - PC sales volume

The number of posts that have been found on sites such as Zhihu's website asking for advice on how to buy a computer. As shown below, the problem has been viewed as high as 7.27 million, so this pain point is really there, and the user group is mostly computer white.

Pic.11 – Evidence

As shown above, we can see that even those who IT-related work still have computer white.

As a result, the user group-computer white includes the computer industry is not familiar with rookie, but also the computer configuration is not familiar with rookie, non-senior computer enthusiasts, among which the group is mainly freshmen.



Appendix 3: Installation and User Guide

Installation:

Environment Requirement (pip command):
> pip install pandas
> pip install flask_cor
> pip install flask
> pip install numpy
> pip install pysqlite3

Congratulation! The environment configuration done!

Run the systems on local machine
> python laptopbackendv4.py

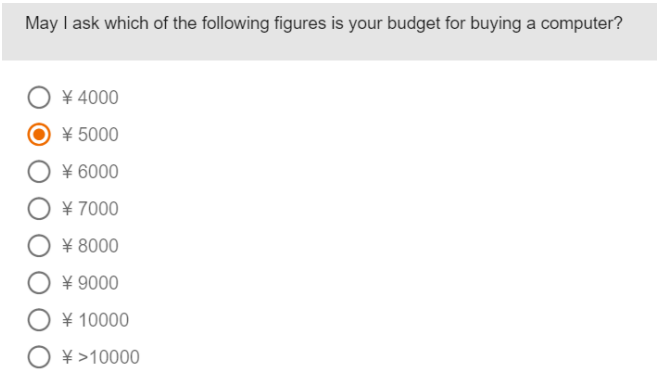
Then click into the 'pages' folder, view 'index.html' in your browser. And begin to choose your laptop!

User Guide:

The use of the system is quite simple, just complete the Questionnaire according to truth. and you can get a best recommendation for you

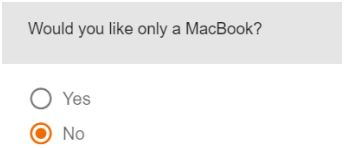
Notice:
Please complete the questions according to your situation. Because the question will change according to your choice.

About question:
Q1. About price

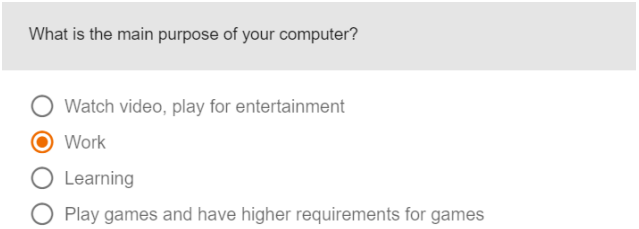


Choose one price which is most close to your budget

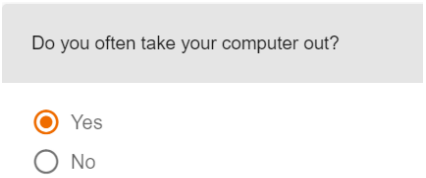
Q2. About macbook



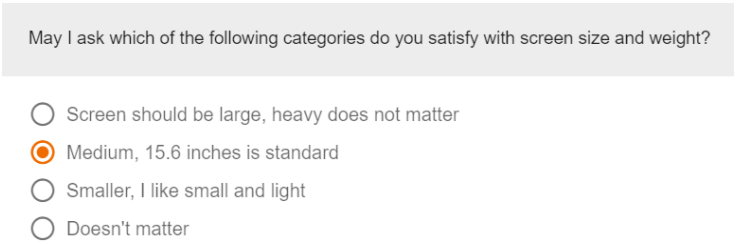
Because macbook is so different to other laptop ,so we set a individual question for it.
If you want a macbook , choose yes
Q3-4:



Choose one according the truth
Q5:

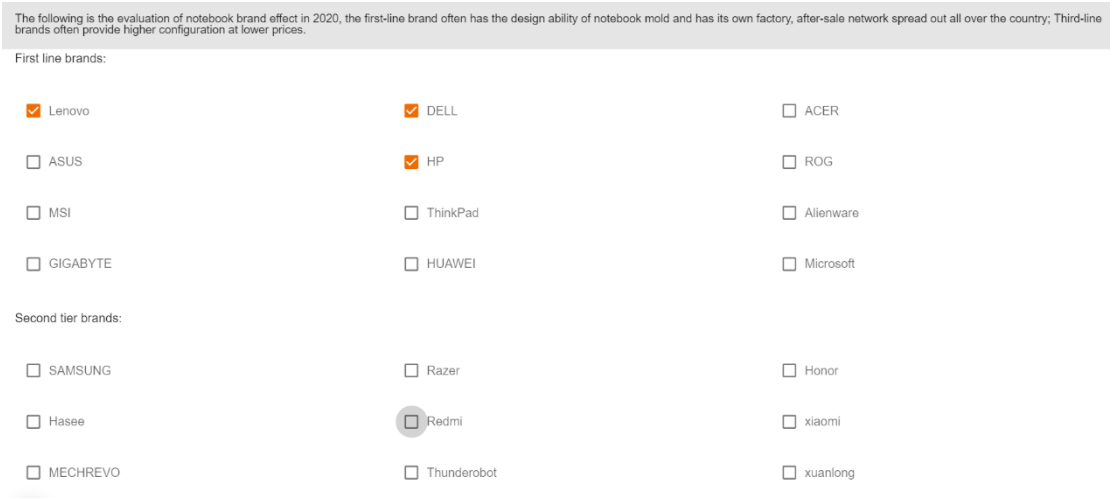


If you choose yes ,the system will consider more about the battery life.
Q6:



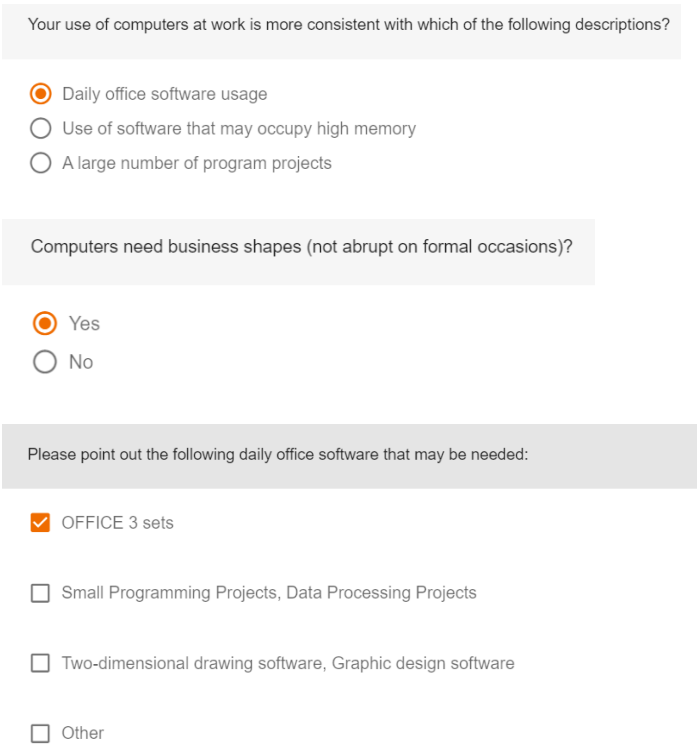
Choose one according the truth

Q7: brand:

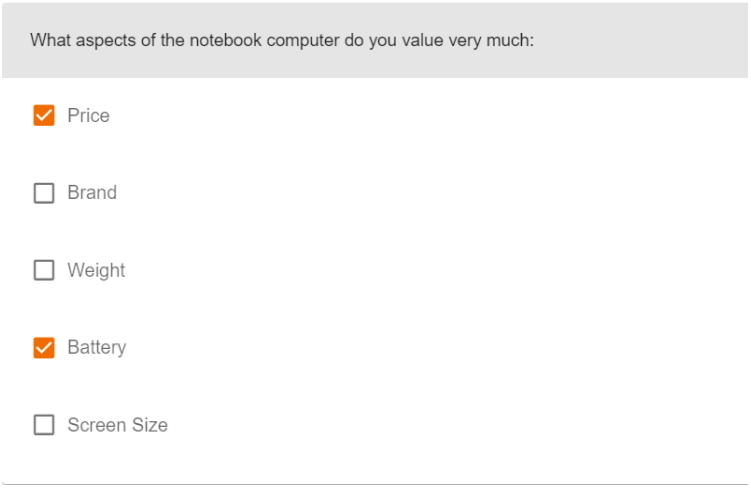


You can choose multiple among all these brands.

After Q7 there will be some question about your purpose, just choose the most similar one.

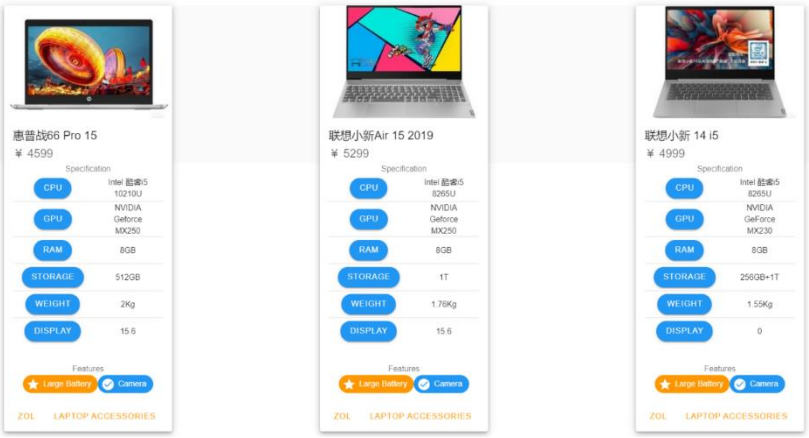


The last question is important to get the result:

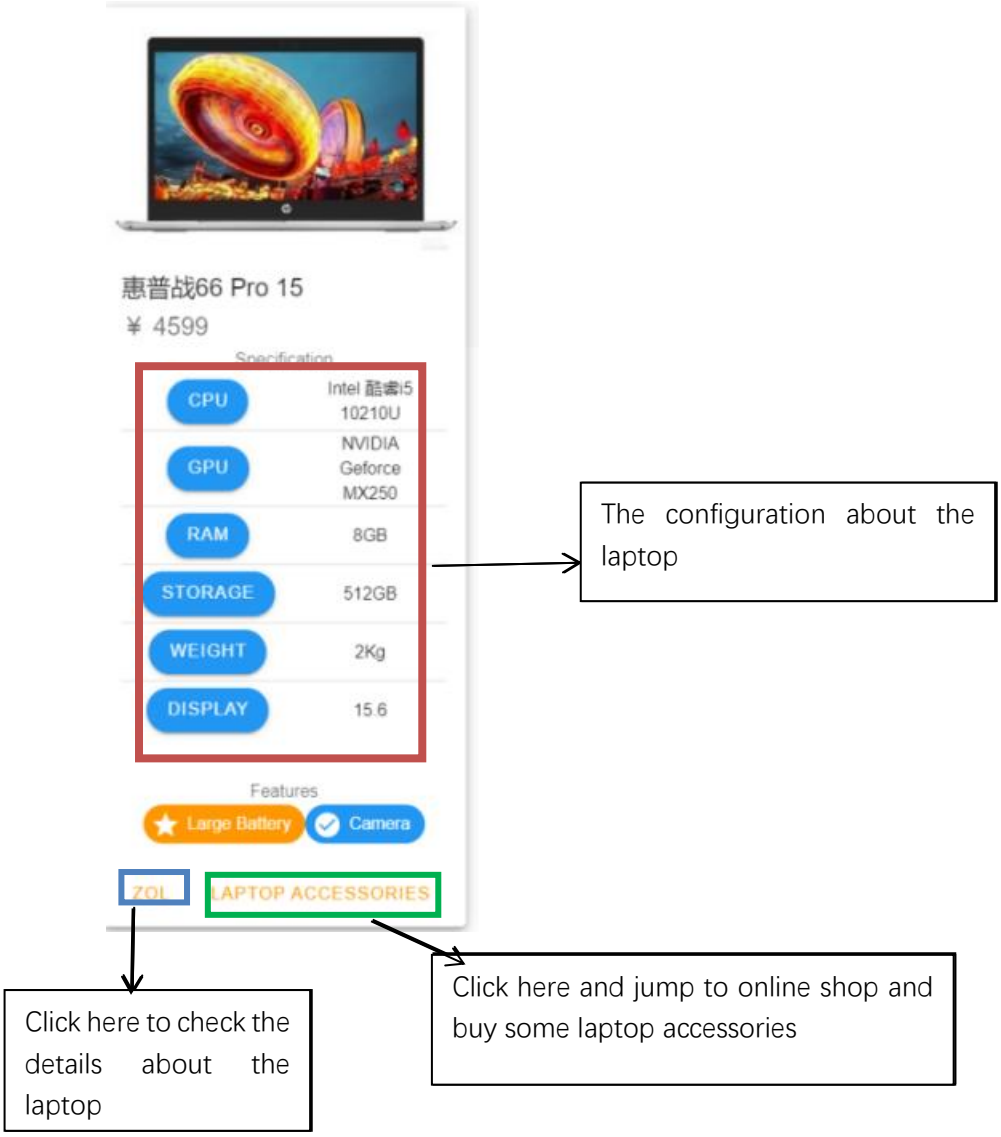


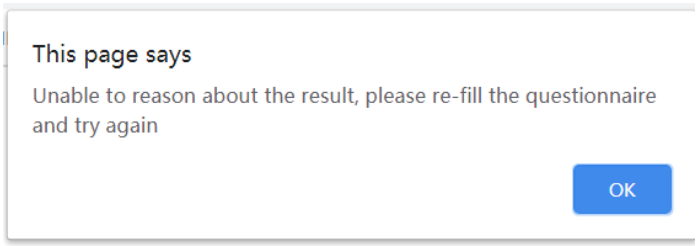
If you choose price, the result products will be definitely in the budget range.
If you choose brand , the result will be the brand you chose before(Q7)
If you choose weight ,the result will be the thin type
If you choose battery ,the result will have powerful battery
If you choose screen size ,the result will have big screen.
Notice: please avoid choosing more than three options in this question, or you will get nothing in result probably.

Result page:



Check the result page and see whether you like these laptops?





If you get this message, you should add you budget or modify the selection at the last question!

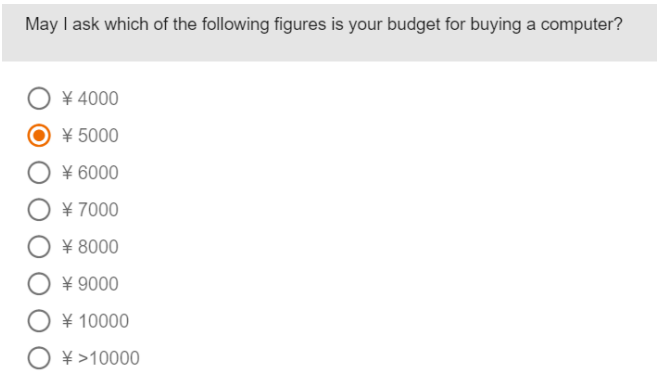
Wish you a happy use!

Appendix 4: Use Case

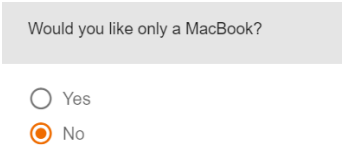
CASE 1:

Lily is a Chinese teacher in a primary school. Her old laptop was a little bit out of date. Sometimes her laptop shut down in class, which is really troublesome and low efficient. So currently, Lily wants to choose a new laptop, but she doesn't really know the computer a lot. And her budget is limited. So she tries our system.

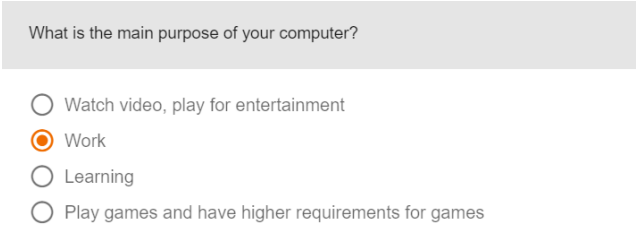
Q1: Her budget is about ¥ 5000



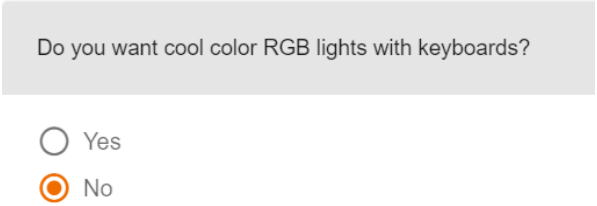
Q2: not only Macbook



Q3:



Q4:



Q5: usually take it out to classroom

Do you often take your computer out?

- (●) Yes
- ( ) No

Q6:

May I ask which of the following categories do you satisfy with screen size and weight?

- ( ) Screen should be large, heavy does not matter
- (●) Medium, 15.6 inches is standard
- ( ) Smaller, I like small and light
- ( ) Doesn't matter

Q7: She only heard about Lenovo DELL HP before so she choose these three.

The following is the evaluation of notebook brand effect in 2020, the first-line brand often has the design ability of notebook mold and has its own factory, after-sale network spread out all over the country; Third-line brands often provide higher configuration at lower prices.

First line brands:

☑ Lenovo	☑ DELL	☐ ACER
☐ ASUS	☑ HP	☐ ROG
☐ MSI	☐ ThinkPad	☐ Alienware
☐ GIGABYTE	☐ HUAWEI	☐ Microsoft

Second tier brands:

☐ SAMSUNG	☐ Razer	☐ Honor
☐ Hasee	☐ Redmi	☐ xiaomi
☐ MECHREVO	☐ Thunderobot	☐ xuanlong

Q8:

Your use of computers at work is more consistent with which of the following descriptions?

- (●) Daily office software usage
- ( ) Use of software that may occupy high memory
- ( ) A large number of program projects

Q9:

Computers need business shapes (not abrupt on formal occasions)?

- (●) Yes
- ( ) No

Q10:

Please point out the following daily office software that may be needed:

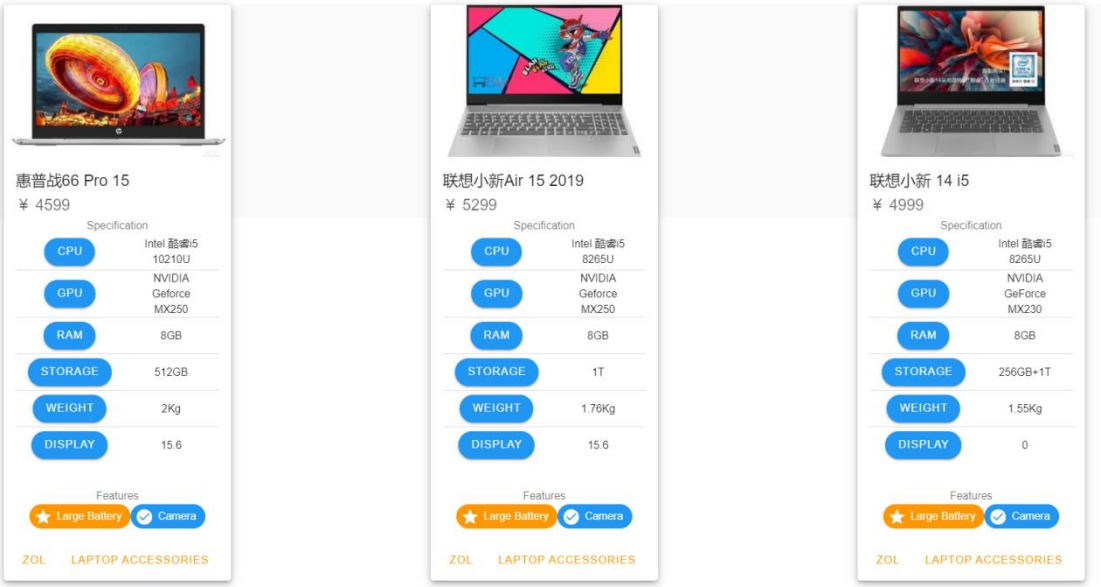
- [x] OFFICE 3 sets
- [ ] Small Programming Projects, Data Processing Projects
- [ ] Two-dimensional drawing software, Graphic design software
- [ ] Other

Q11: she care most about the price and the battery for daily use.

What aspects of the notebook computer do you value very much:

- [x] Price
- [ ] Brand
- [ ] Weight
- [x] Battery
- [ ] Screen Size

Click “Submit”, then she get these three best fit result .in less than 2 minutes



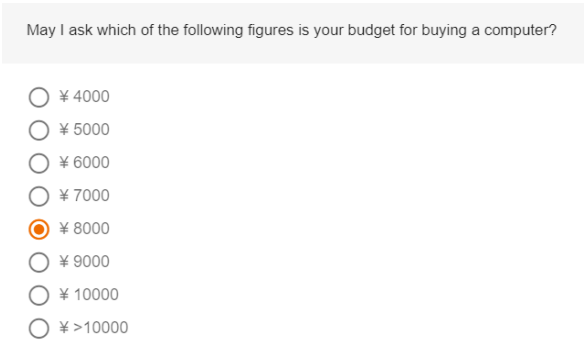
Explain:

Firstly, the prices are about ￥5000 which satisfies the requirement. Secondly, all there three have a powerful battery. Lily thinks this recommendation system is useful, and she clicks the ‘ZOL’ button to get further details.

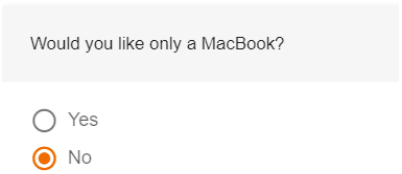
CASE 2:

My friend Pony is a freshman at a university. He likes to play various games. And he wants to pick a laptop for use on the campus. But he does not know which to pick. So I tell him to try our recommendation system.

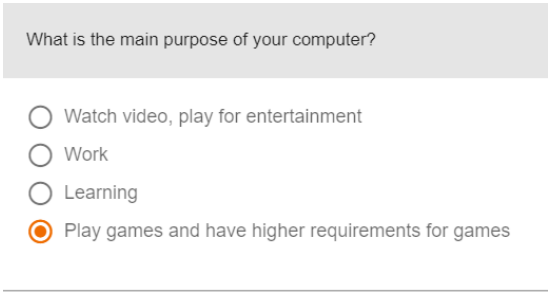
Q1: his parents gave he ￥8000 to buy a laptop.



Q2:

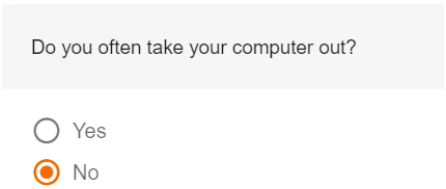


Q3: for games



Q4:

Q5: usually use in dormitory



Q6: large screen for games

May I ask which of the following categories do you satisfy with screen size and weight?

- (●) Screen should be large, heavy does not matter
- ( ) Medium, 15.6 inches is standard
- ( ) Smaller, I like small and light
- ( ) Doesn't matter

Q7:

The following is the evaluation of notebook brand effect in 2020, the first-line brand often has the design ability of notebook mold and has its own factory, after-sale network spread out all over the country; Third-line brands often provide higher configuration at lower prices.

First line brands:

☐ Lenovo	☐ DELL	☐ ACER
☐ ASUS	☑ HP	☑ ROG
☐ MSI	☐ ThinkPad	☐ Alienware
☐ GIGABYTE	☐ HUAWEI	☐ Microsoft

Q8:

Are you very interested in large stand-alone games?

- (●) Yes
- ( ) No

Q9:

Usually play games attention will be more inclined to special effects?

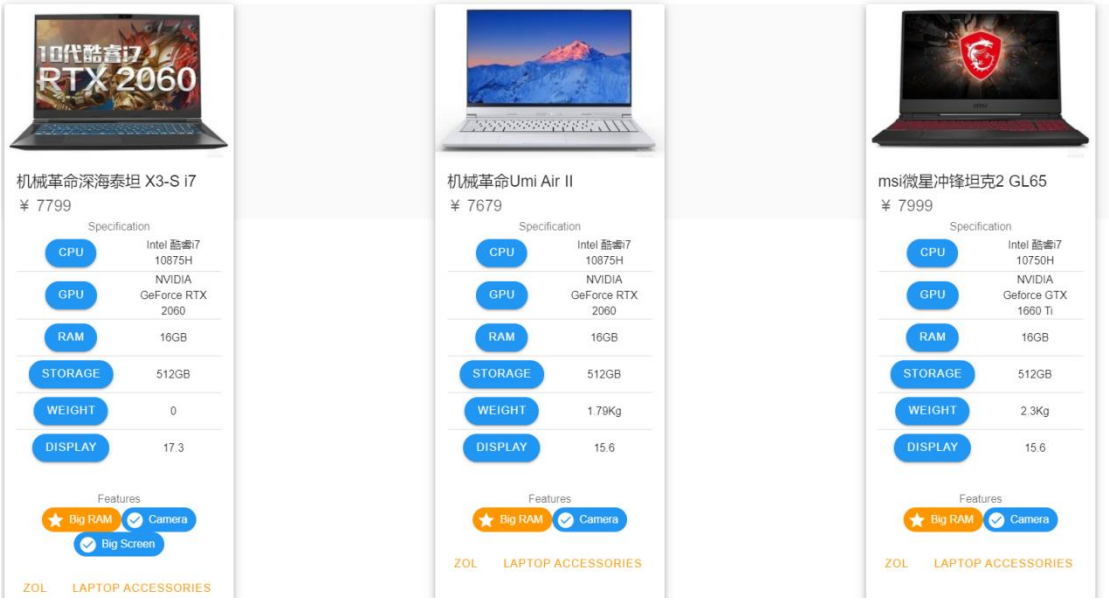
- (●) Yes
- ( ) No

Q10: budget limited

What aspects of the notebook computer do you value very much:

- ☑ Price
- ☐ Brand
- ☐ Weight
- ☐ Battery
- ☐ Screen Size

Then submit, he gets these three:



Explain:

The GPUs are all suit for 3d games, Big screen is the priority. Gamebook application scenarios require high computing power and high-quality requirements, such as playing 3 A or training complex neural network models, so notebook computers need high hardware to provide the corresponding capabilities, such as high memory, powerful graphics cards, high refresh rate screen.

CASE 3:

My new friend Jackson, who has been admitted to a major in artificial intelligence. Then he wants to get a new laptop for study. He wants to try this recommendation system and see if the system can help him choose a laptop.

Q1 : he choose ￥8000

May I ask which of the following figures is your budget for buying a computer?

- ○ ¥ 4000
- ○ ¥ 5000
- ○ ¥ 6000
- ○ ¥ 7000
- ◉ ¥ 8000
- ○ ¥ 9000
- ○ ¥ 10000
- ○ ¥ >10000

Q2: NO

Q3:

What is the main purpose of your computer?

- ○ Watch video, play for entertainment
- ○ Work
- ◉ Learning
- ○ Play games and have higher requirements for games

Q4:

Do you want cool color RGB lights with keyboards?

- ◉ Yes
- ○ No

Q5:

Do you often take your computer out?

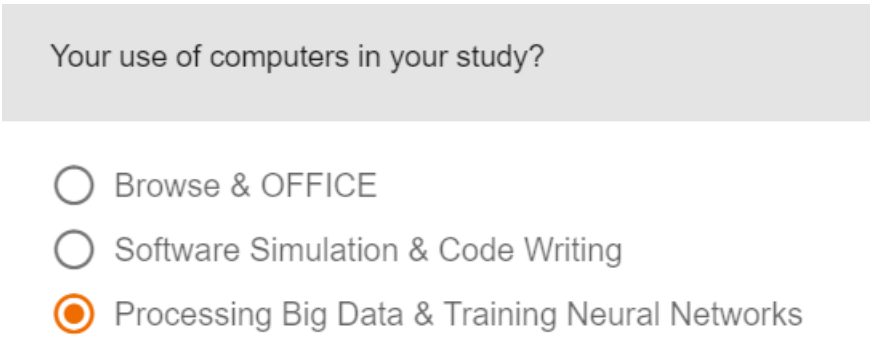
- ◉ Yes
- ○ No

Q6:

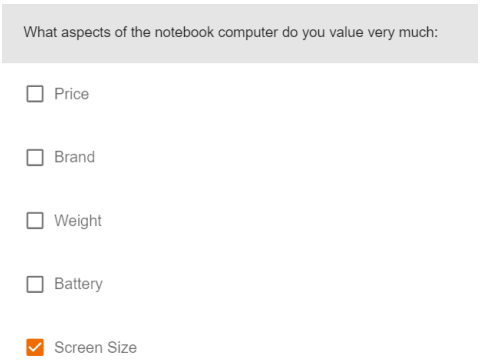
May I ask which of the following categories do you satisfy with screen size and weight?

- ◉ Screen should be large, heavy does not matter
- ○ Medium, 15.6 inches is standard
- ○ Smaller, I like small and light
- ○ Doesn't matter

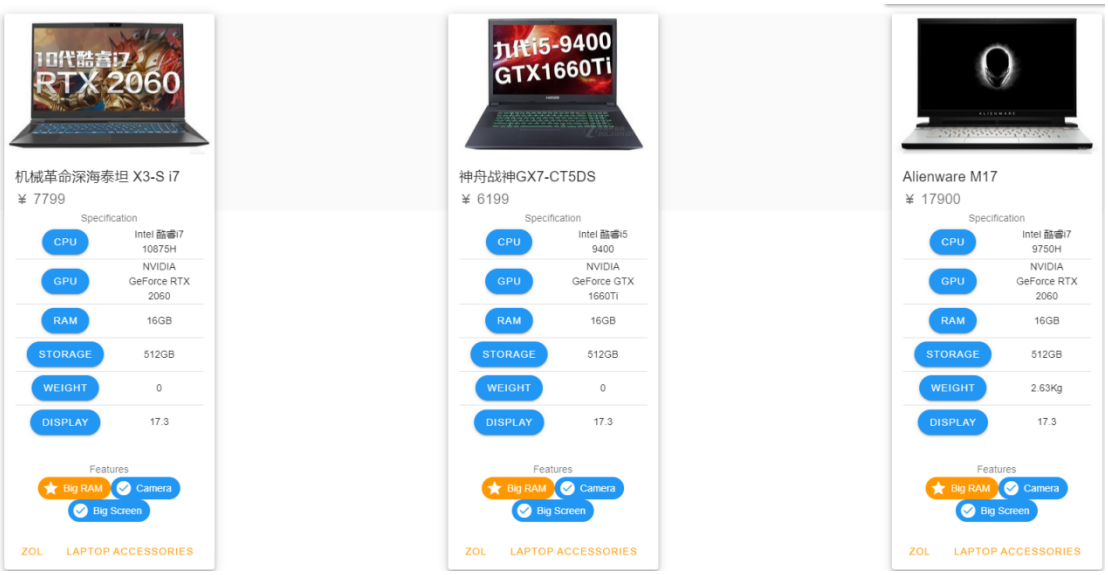
Q7:



Q8: he only wants a big screen laptop, the price does not matter



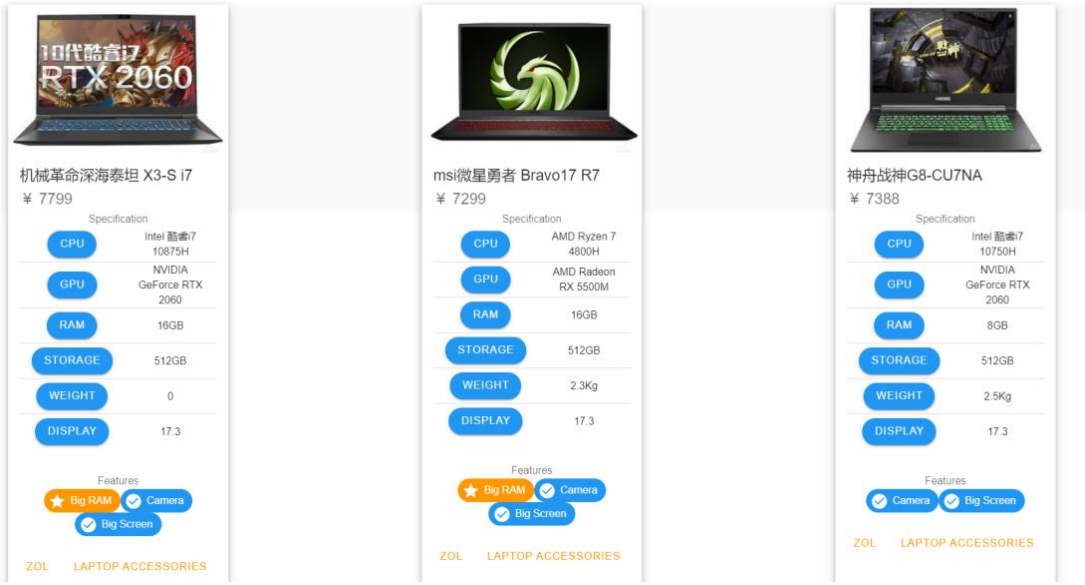
Then submit



Explain:

There is an Alienware M17 which price ￥17900 much more than 8000. Because in the last question, he did not choose the ‘price’, so the system will not consider price as a very important factor, then the other price range will be considered. But they all have a big screen which Jackson required. The RTX 2060 is good for deep neural network training.

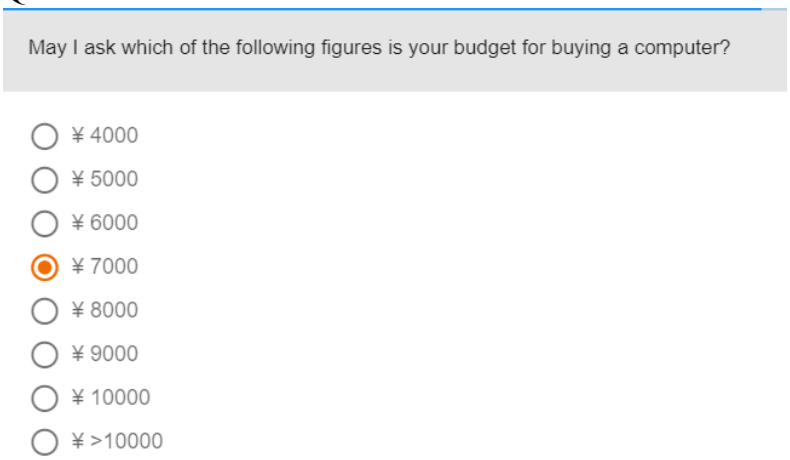
Let us see what happened if we select "price" at the same time. Here is the result:



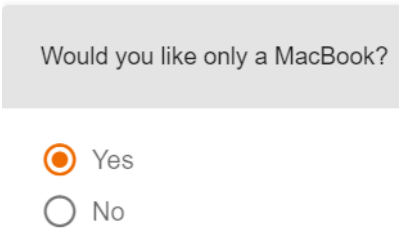
CASE 4:

In an invalid case, Jane wants to buy a MacBook, but she just has ¥ 7000 budget. When she uses this system to get a laptop, something different happened.

Q1:

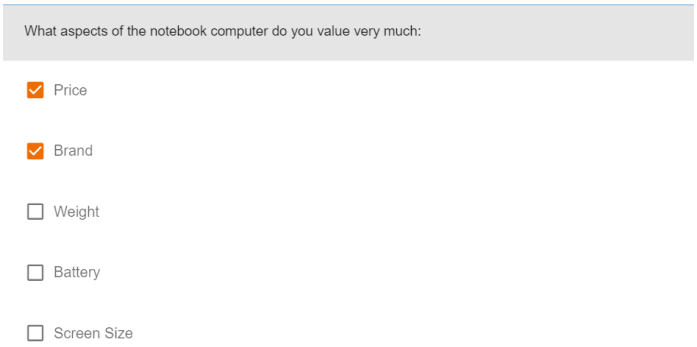


Q2:

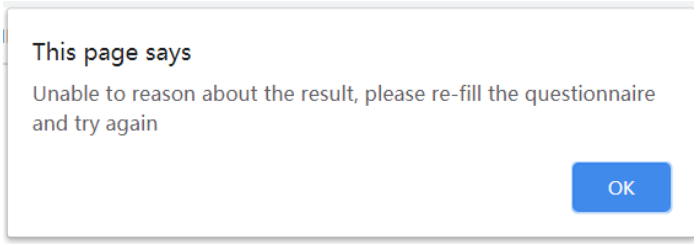


Q3-Q7 does not influence, skip them here.

Q8: She only wants the MacBook , but her budget is limited at the same time

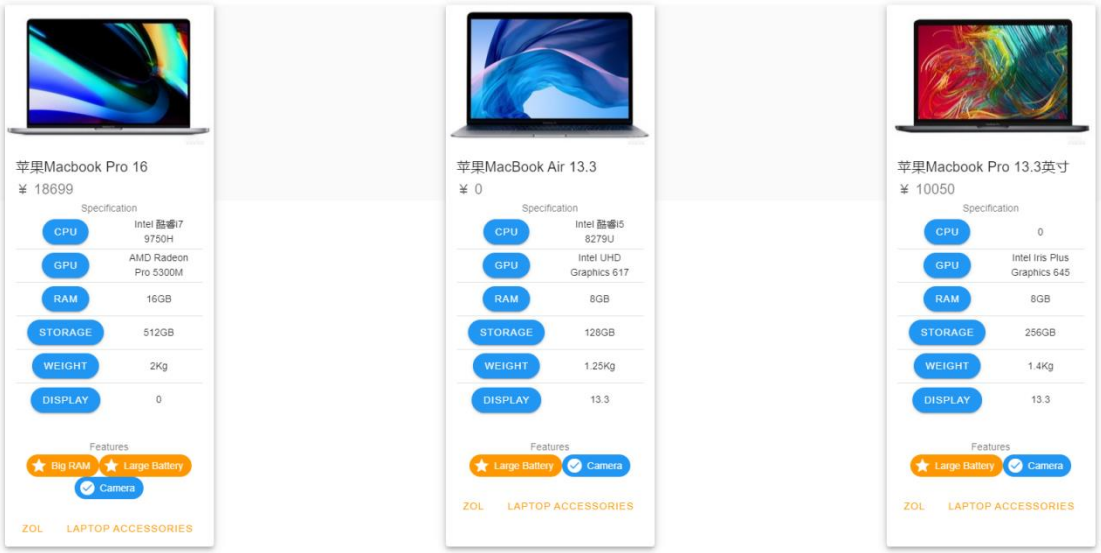


Then submit, we get:



Explain:

￥7000 is not enough to buy any Macbook, so no match result here. Maybe Jane can add more budget or cancel the Price in the last question. Then there will be some results. Let us cancel the option ‘price’ and try. We get this:



The middle one in the database has a missing value on price. The reason is the source page does not have the relative data.

Appendix 5: Mapped System Functionalities against the knowledge

Business rule OR Business process OR Knowledge-based reasoning techniques:
(1) Crawl Data - We Crawl data from the online shopping Web and do data preprocess such as data clean and restructure.
(2) Data Clean
(3) Rule Engine - Decision tree/ Rule making

Business resource optimization techniques Search OR Constraint satisfaction OR Evolutionary computing:
(1) Rule Engine - There are hard/soft constraints to scale the selection space. Details can refer to Appendix - rule engine

Knowledge Discovery using suitable data mining techniques:
(1) Database - we apply K-means to classify PC clusters and categorize them as several types.

Appendix 6: Rule Engine

What role does the rule engine play in this project?

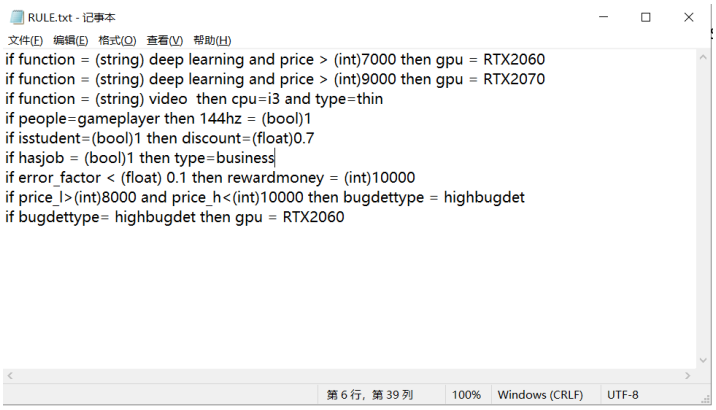
Let's imagine there is a guy who has no idea about the laptop hardware thing. So he doesn't know how to choose a laptop which can cover his requirements but not too expensive. One day this guy has to choose a laptop for some reason. And he would be asked the questions like "what do you want to do with your laptop? " by the shop assistant. Then, he will recommend some laptops to you which can cover your requirement using his knowledge about the link between user preference and computer hardware configuration. Here our project plays the role of the shop assistant, and our rule engine represent the domain knowledge of the "configuration"

In this project, we used the self-made Rule Engine to build the knowledge base. How to deploy this engine and the fundamental of this RULE ENGINE will be explained in the following scope.

How to use it?

a) Function Description

Customize the rule file in a certain format, save it in .txt format, use this rule engine to digest the rules in this format, and build a knowledge table in memory.



Pic.12 – A rule file

Currently, this rule engine only supports up to two conditions per rule, and logical operations between two conditions only support "and" and "or".

The condition value and the result value only accept a string, bool, int, float types, and the output currently only supports the equal (=) operator.

The rule engine can output two results when one rule was fired.

b) Data Structure

Input: Use dictionary structure to record each fact as input
Output: Use dictionary structure to record each result as output

The reason we choose a dictionary is that it's easy to convert the dictionary into Json format.

c) Rule Format

There is something that needs to be noticed when we write the Rule.txt file.
Every rule should be write in the following format:

if [fact1] = [fact_value] and [fact2] = [fact2_value] then [result_name]=[result_value]

The 'and' can be changed with 'or' , or you can just have one conditional. The underlined operator can be replaced with any of the following :
{ "!=",">=","<=","=","<",">"}.

The content in the '[ ]' can be customize as your like. But in this part you should avoid using the following keywords:" ", "!=", ">=", "<=", "=", "<", ">", "if", "then", " and", " or", "(", ")" .

An Invalid Example:

if personality = kind and sex=male then type= good man
('good man' here should be changed as 'good_man')

d) Choose variable type

All [fact_value], [result_value] only support four types: bool, string, int, float.
Using a pair of '( )' before the variable to specify the type.

Example:

if error_factor < (float) 0.1 then reward = (int)10000

When it comes to comparing numbers, you need to use parentheses before the value to specify the type of value. If it is bool type, then you should use '(bool)1' , '(bool)0' to represent 'True' , 'False' respectively.

Example:

if is_student = (bool)1 then discount=float(0.7)

All types are default to the 'string' type if there are no parentheses specified the types.

Example:

if income > 10000 then is_rich = (bool)1.
This rule may not work because the fact_value 10000 here will be explained as a string type, you can't compare a string type with a numerical type. Here should be:

'if income > (int)10000 then is_rich = (bool)1'

or

'if income > (float)10000 then is_rich = (bool)1'

e) Conflict Strategy

Sometimes there would be some conflict in the rule file. Our strategy is the rules written later have higher priority.

Watch an example with two rules

Example:

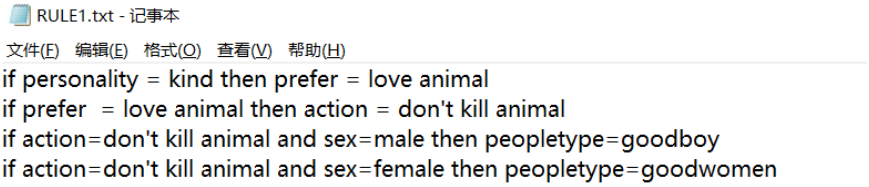
if a > (int)3 then b = (int)2
if a > (int)7 then b = (int)4

When we input {'a' :8}, the result will be {'b' :4}. Because the second rule has a higher priority. But if you change the position of the two rules, the result will be {'b': 2}. In other words, the latter result will overwrite the previous result.

f) reasoning more than one time

The result obtained after a round of inference will be used as input to enter the rule engine again for matching until no new result is generated, and the inference ends.to show you this, watch the following example

Example:

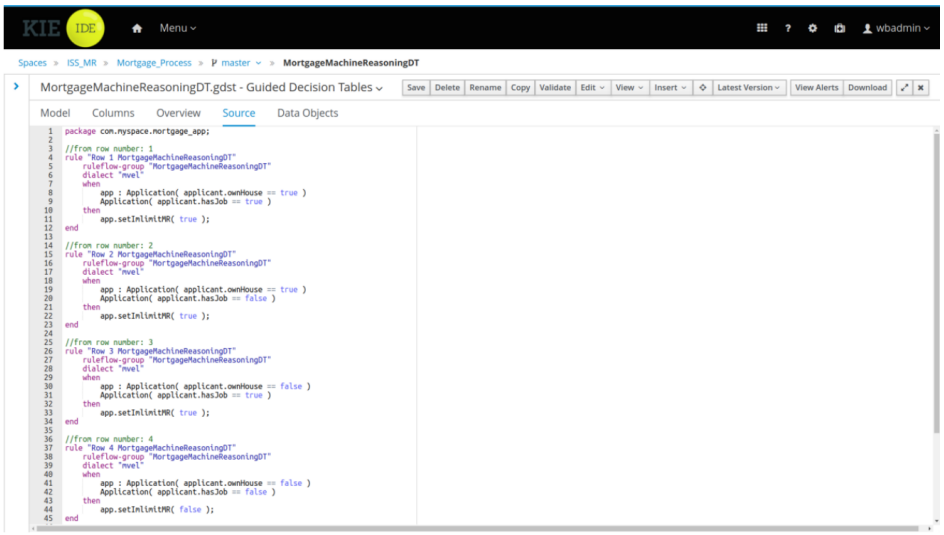


Input: case ={'personality': 'kind', 'sex': 'male'}
Output: {'prefer': 'love animal', 'action': "don't kill animal", 'peopletype': 'goodboy'}

It obvious that the action in the third rule is not the item we input but is derived from the 'personality: kind' after the first rule and the second rule, and finally we get the result of 'people type: good boy'

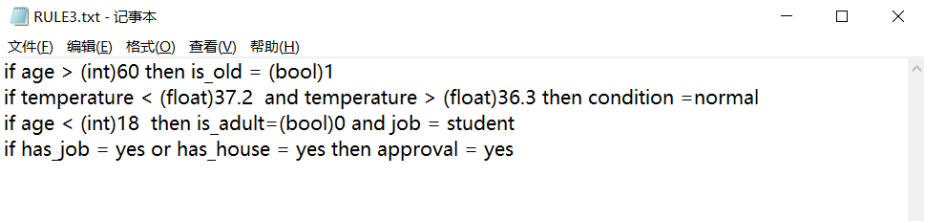
Fundamental of the self-made rule engine:

KIE with a rule engine can define rules easily. Let's view the rule in source code form in the KIE before we explain how our self-made rule engine work.



Apparently, all the rules are constructed in when-then form. Users need to name their variables and set their conditional values. In the self-made rule engine, people should write down their rules in a .TXT file in the specified format. Then, the rule engine will extract out the variable name, variable type, the relationship between two conditions, and value automatically. All this information will be stored in a table.

Watch this rule file:



After being explained by the rule engine, the rule table will be built.

input_type1	input_value1	input_type2	input_value2	relationship1	relationship2	input_relationship	result_type1	result_value1	result_type2	result_value2	result_relationship1	result_relationship2
age	60	0	0	5	0	0	is_old	True	0	0	3	0
temperature	37.2	temperature	36.3	4	5	1	condition	normal	0	0	3	0
age	18	0	0	4	0	0	is_adult	False	job	student	3	3
has_job	yes	has_house	yes	3	3	2	approval	yes	0	0	3	0

Rule table

Explain:

(1) input_type1, input_type2, result_value1, result_value2 record the variables' names.

(2) input_value1, input_value2, result_value1, result_value2, record the variables values.

(3)"relationship1" record the relationship between 'input_type1' and 'input_value1', 'result_relationship1' record the relationship between 'result_type1' and 'result_value1'. number 0,1,2,3,4,5 represent "!=",">=","<=","=","<",">" respectively.

(4) 'input_relationship' is 0 means there is just one condition in the rule,'1' means the 'and' logic operator, '2' means 'or' logic operator between the two input conditions.

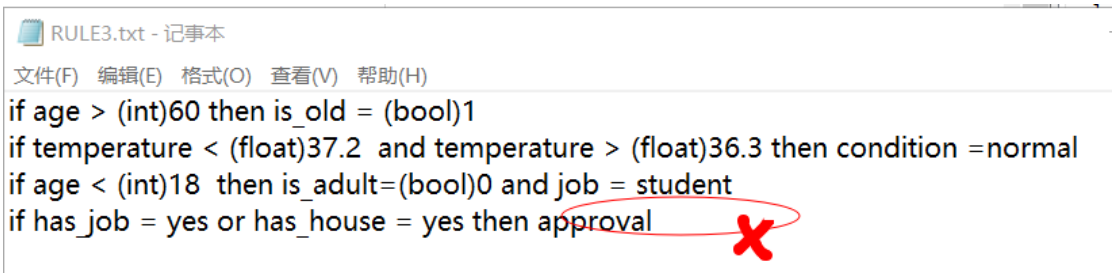
In this way, the rule table can represent all the information in the rule.txt.

During the inference, every input variable name will be compared with the 'input_type'. If they are matched, then compare their values to decide if this rule will be fired or not. If the rule was fired, the result will be appended to the result dictionary. That' how the rule engine working.

Future Improvements:
The rule engine is used to customize your own rule in a very simple and convenient way. But like the first version viable engine, it can be improved in the following aspect.

(1) the exception control can be various
Now the rule engine didn't have many reminders when there exist some exceptions in the rule.txt file. But We do design the basic reminder in this engine.

Example 1: to detect the incomplete rule



Console:

```
RULE NO. 1 has been bulit
RULE NO. 2 has been bulit
RULE NO. 3 has been bulit
条件错误, 无关系符号(=,>,<..)
```

Example 2: to detect the wrong type

```
if __name__=='__main__':
    rt=ruletable("RULE3.txt")
    rule_table= rt.rule_table
    case={'age':'none'}
    result=rt.reason(case)
```

Console:

```
RULE NO. 1 has been bulit
RULE NO. 2 has been bulit
RULE NO. 3 has been bulit
RULE NO. 4 has been bulit
jugde type error ,string cant compare to int please check the
rule.txt or input
```

Example 2 tells about the error type, but it doesn't tell about the exact error location, if there are huge amount rules in rule.txt, it's troublesome to locate the error one.

Here, we think the rule engine can be improved in the exception catching aspect:
(1) have more detect type in rule.txt
(2) specify the error location.

2. more function and more powerful

As we mentioned before, the output of the rule engine can only like this:
Result {A: 'something', B:'3'}
Which means A equal to 'something' and B equal to 3. But there is no way to do it when we want to express B>3 and A is not 'something'. This limitation was from the output format of the engine. We output key-value pairs, in this format, we have no idea to express like 'B>3'. So this is the point about how to make it more powerful.

Another point is the input condition of each rule. In the rule engine, we can accept only two rules at most. If input with more than two conditions in one rule, we have to split it. But in most cases, two conditions connect with 'and' and 'or' is enough in practice.

Appendix 7: Individual Report – Chen Jiajun

ISS_MR_PM Laptop Recommender Individual Report
Chen Jiajun A0215519E

1. Personal contribution

This project experience is rewarding for me. I work together with Wang Ding, Zuo Zhen, and Li Yuan. It is a great journey to finish a project with those fantastic guys who bring me lots of joy and sharing knowledge selflessly.

In actual development, I was mainly responsible for front-end design and development. Also, I was always actively involved in the weekly meeting. I would like to suggest some solutions and give my opinions to my teammates. It was a great joy to discuss with my classmates and apply the knowledge I learnt from the course in the project.

2. What is learnt?

During this project, I gain valuable experience from these three parts - system design, front-end development, teamwork.

(1) System Design
From the discussion with my teammates, I have learnt the current situation of the laptop market. Wang Ding, who was responsible for market research, gave me some methods for market research. I used to focus on the coding part. However, in this project, I found that discovering the business values and feasibility of the system is also a necessary part.

In the system design part, I got a whole picture of a machine learning system including front-end system, back-end system, algorithm core, database, etc. Although I did not need to develop the entire system, it is also helpful to understand the data flow and system design.

(2) Front-end Development
I develop the whole front-end application in the project. It is a challenge for me because I did not have prior experience with web development. But I want to expand my skills to keep myself updated with new techniques so that I took over this part.

In the beginning, to find a short path to complete the front-end development, I did some quick learning about basic coding language (HTML + CSS + JavaScript). Then, I did some surveys about the latest trend in web development. I found that Vue.js and Vuetify.js UI Framework have some friendly user manuals and it is widely used in the industry. So I choose Vue.js and Vuetify.js UI to develop the website. Usually, I developed the application and checked the framework manual at the same time. I tried

to apply some new things I learned in the front-end application. Sometimes I stuck in some fancy features or modern developing rules. It has stricter rules compared to python or c++. I would check the manual again and start with the fundamental part I missed. It took a few days to learn how to build a web application demo.

During the development, I realized that how important the component concepts and modern developing rules are. It can shorten the time of development and make the project more friendly to read and modify. The component concepts in vue.js can let me reuse my codes and update the codes easily in the future. The developing rules can let me build a more robust system and track the abnormal part. I gained lots of coding experience from developing the front-end system.

(3) Teamwork
I used to handle all the coding work by myself in the project. In this project, I found that it is more efficient to work with others. Li Yuan, who was responsible for back-end development, taught me how to use Postman to simulate a deployment environment. He is a great guy and I enjoy working with him.

To advance the project, I need to communicate with Li Yuan to develop APIs for communication between the front-end and back-end. Also, we need to discuss the demand for the application and assign tasks properly. I have a better understanding of the responsibilities of the front-end and the back-end. It gave me some pieces of knowledge in career planning.

3. How can you apply the knowledge and the skill gained?

Firstly, I have a better understanding of machine learning knowledge learned from the courses. I feel how efficient those data cleaning skills and machine learning skills are. I may apply KNN cluster methods to get some insights from the dataset in my next project. Also, how to manipulate data and store it in the database is quite useful for a data scientist.

Secondly, I have learned how to build a website to represent some visual output from a machine learning system. Web development skills can help me to develop a more friendly and interactive application. Also, I can use these skills to build my blog and show my thoughts. It is a bonus to have a personal website when I apply for a job.

Thirdly, developing a whole project with others will give me more confidence and experience in the future workplace. I went through a whole project circle from market research to system deployment. I have learned how to play a contributing role in a team.

To sum up, this project brings me experience in developing a machine learning system and teamwork.

Appendix 8: Individual Report – Li Yuan

ISS_MR_PM Laptop Recommender Individual Report
Li Yuan A0215499R

1. Personal contribution

This project is quite interesting for me, when I proposed the project proposal, it was unanimously agreed, and everyone thought that it was a meaningful project and we had the confidence to do it well, and finally we did it.

In actual development, I was mainly responsible for Back-end , Database (Flask and SQLite3) and the Rule-based model making. Also everyone on the team helps the other team members work together to overcome each others' problems, We thought about the project presentation together, shot the project presentation materials, help each other and improve each other, pushed the project forward together.

According to our structure of the whole system, our Back-end and Database can satisfy the following requirement.
1. Be able to get the information from front end. Make sure it can connect with front end.
2. Deal with the raw user preference information, convert it into configuration information.
3. Store all the laptop information in database, and be able to select data form database.
4. Return result back to front-end.

For the Rule-based model making, it is a quiet import part embedded in the rule engine, It is constructed mainly by referring to expert suggestions, combining with the results of famous computer evaluation bloggers, analyzing and summarizing clustering effects, and classifying user groups. Finally, user problems can be converted into specific laptop configuration, and then matched in the database.

2. What is learnt?

During this project, I have a better understanding of how a reasoning system was making-up, and I gain valuable experience mainly from these three parts - Back-end, Database and Rule-based model making.

(1) Back-end
Firstly, I have troubles to choose the python web framework between Django and Flask. Because I don’t really know the pros and cons of these frameworks. But I have been exposed to flask in class once. And I thought this framework is quite convenient to build a API for front-end. So I dig into flask to learn more and tried to deploy it .

To handle the above three points. I learned much knowledge and technique. For the first

point, GET and POST methods are two HTTP request method. GET method is used to get some information from resource, and POST method is used to submit some form data need to be process and get response.

Actually, the rule engine doing the mapping job, which satisfy the second point. I just need to invoke the rule engine in back-end.

(2) Database

To satisfied the point of storing all the laptop information in database, and be able to select data form database, I learned SQLite3 and make use of it. An important feature of SQLite3 is zero configuration, which means no complicated installation or management. In another word: lightweight. But it's powerful, SQLite3 is the most widely deployed SQL database engine in the world. This tool is good for small-scale database. SQLite3 can be integrate into python program easily by using the SQLite3 library. Utilize the SQLite, I build a database for all the laptop information we get.

(3) Rule-based model making

It is a difficult problem at the beginning, because it need large professional knowledge about notebook configuration, so I search much information from the prominent blogger, and refer to their toaster scores to make a feature analysis of generally accepted notebooks, and applied what I learned in classes to integrate these characteristics to make a mapping of rules, and finally according to the experimental results to modify the weights of rules and specify the setting of the rules, these steps makes me more clear about how to set rule in a rule-based system, and how to modify and optimize the rules.

3. How can you apply the knowledge and the skill gained?

Firstly, I will apply these knowledge in the data analysis. Data analyst is a very hot job in the market. Except what I learned by my job, I also learned how to do data crawling, data cleaning, clustering, and other data process techniques from my teammates, and I will get more deep in it and apply it in the real case.

Secondly, I acknowledge that how should the front-end interact with the back-end, and I will utilize this interaction and combine it with the database, to do a better job of interacting in the future work. To do a more systematic and more standardized database, and more reliable back-end.

Thirdly, I will combine the experience of team cooperation this time and hope to play a better coordinating role in future team cooperation projects. I will make full us of my strengths both in my own work and in the promotion of the whole project. Enable the team to achieve the goals more efficiently and targeted.

Appendix 9: Individual Report – Wang Ding

ISS_MR_PM Laptop Recommender Individual Report
WANG DING A0216421U

1. Personal contribution

As an team organizer, i organize meetings regularly to discuss progress and ensure the project is carried out in an orderly manner. In case of problems, consider the opinions of the members and promote the solution of the problems. In the whole process of the project, it is necessary to firmly control the product design to meet the requirements of the course design and the business concept.

2. What is learned?

During this period, I studied expert system, crawler technology and data processing, analysis and other technologies. Because the acquisition of our data depends on crawler, the analysis of data depends on some methods such as data pre-processing and clustering, and reference of expert system framework need to depend on mature model.

For Web crawler and data process ,We need a lot of laptops' information to establish the laptop database, so I learned about web crawler and utilize it to grab huge amount of data.

The fundamental version of Web crawler is simple. Every time we open a new page, the browser will explain the HTML, CSS, JavaScript real-time and show them in a colorful way. If we access the web page in one program, we can get the page in source code format. It's difficult to understand the page in this format ,but it do contain nearly all the information in the original page. We all know the HTML file is structured. So we can get any thing we want in this page by explain the source code in some way.BeautifulSoup4： a powerful python library for data extraction from HTML file. Learn this library save much time while extract data from the HTML page.

Utilize pandas to process data and save them.
I use Panda library to store all the data I get from website, and output as CSV file. In this procedure, I have to deal with the raw data, make them more structured. For example, the unit of wight can be kilogram of gram, but they must to be unity. Like according to the CPU info ,I extracted the key words and put a rank on the most CPU .CSV file is more convenient when establish a form in SQLite Database. So I output them as CSV file.

After getting the data that the web page crawled, in order to classify the computer, I thought of using the k-means method in unsupervised learning to set the distance

between the non-numerical data to 1. We want to be able to map several classes to work, study, video, games and so on. So we try to get several categories by clustering the existing computer products, during which we need to use elbow method to verify the number of K values.

At the same time, I also verify the performance of the model through the following two validation methods:
1. evaluate the number and size of clusters. If the size of a class is too large, we should think about whether we need to subdivide, and if the size is too small, we should see if we need further investigation.
2. evaluation of clustering cohesion and Separation degree can be calculated separately and Separation. can also be combined calculation, silhouette method

In the construction of expert system, I refer to and take lessons from the relevant information of the existing expert system. The expert system is usually composed of six parts: human-computer interface, knowledge base, inference machine, interpreter, comprehensive database, knowledge acquisition and so on. In particular, the knowledge base and reasoning machine are separated from each other and unique. The architecture of expert system varies with the type, function and scale of expert system. According to knowledge representation technology, it can be divided into logic-based expert system, rule-based expert system, semantic network-based expert system and framework-based expert system.

Future trend: the development of expert system will be based on model reasoning, supplemented by rule reasoning, and meet the needs of commercial applications to meet the needs of real-time and large data processing. At the same time, the expert system will develop in a more specialized direction, providing targeted models and products for specific directional needs.

Knowledge acquisition is the key to the superiority of expert system knowledge base and the "bottleneck" problem of expert system design. Through knowledge acquisition, the content in the knowledge base can be expanded and modified, and the automatic learning function can also be realized.
Although our project has been formed, there are still some aspects to be improved.

3.How can you apply the knowledge and the skill gained?

I think the factory and enterprise process can be further designed as a rule-oriented, which can reduce the tedious process work and only allow manual intervention at the nodes that need to be decided by people. I will also analyze market data, product data, and find laws from the data, so that the law can be fully applied when making the next plan.

Appendix 10: Individual Report – Zuo Zhen

ISS_MR_PM Laptop Recommender Individual Report
Zuo Zhen A0215464H

1.personal contribution to group project

develop the Rule Engine And build user-guide for the engine:
We need a rule engine to establish the mapping between user preference and laptop hardware requirement. At the very beginning , we preferred to choose a mature rule engine which is stable and powerful in application. So I start to search a rule engine which is user-friendly, and try to understand the api and demos it provided. I have read the guide about JBPM, Drools for a while, but I felt they are hard to use .For example , this is one rule in Drool:

It's a little bit troubling if we want to modify maybe 30 rules in above format. Also, we need a convenient way to change the rules and deploy them. So I start to think about self-made rule engine for the above and following reason:
A) The functions of these tools are so various. But we only want to use the rule engine. Simpler one ,the better one.
B) We don't know the format of input and output in these tools, so it may cost much time in integrating it in python(We use python to build our project).
C) It seems not very hard to build a rule-engine based on python. And developing a self-made rule engine sounds interesting and excited.

The way of rule engine developing:
First, I have to draw a basic construction refer to how we build rules in KIE.

This is the conditional part("when") of a rule, there are three factors which are important for a single rule.
1.blue : conditional name or fact name, which decided if this rule should be choose. With a input ,we choose the relevant rules and see if it can be fired. Part in blue box is the keyword.
2.green: the relationship between the fact and the value.
3 red : the value of the fact . There are some requirement about this part because data has their own type, some of them are string. Some of them are number. The rule engine should be able to recognize different data type.
The structure of "then" part of a rule is quite similar with "when" part. Therefore I focused on how to process the string and extract the above three parts. I learned about the regular expression and utilize the python re library which helps to do this work.

There were some troubles in this development journey but I still felt motivated to handle them and continue. But it's really light and quite easy to use. The function is enough for our project and also able to satisfy the usual scenario.

2.what learnt is most useful for me

I think learning is not only about technique aspect ,it is also about personality.
In the development way , firstly, make sure I know how the thing works, draw a script for the ‘product’ I was going to develop. This experience enhanced my understanding of rule-based systems. Then we can move according to our timetable.

By developing a rule engine as the core of the project, not only my code ability has been improved, but also I became more patient and careful. I have been coding for several days , debug and finally completed the engine. Because The project will not be always smoothly completed. So a patient heart is also important. Make my thought clear first, and just do it.

3. how I can apply the knowledge and skills in other situations

1) The rule engine itself

I think the rule engine I developed is really light and very easy to be embedded into python program. So maybe the rule engine will be useful in my future project.

2) develop ability

AI or intelligent system is something about practice. So hands-on ability is very important for our AI people. As I mentioned before ,we did not get the idea of self-developed rule engine at very beginning. But after some analysis about the feasibility, I decided to do this. And I thought I did a good job ,because my teammates accepted the rule engine and decided to use the engine in our project.

In the process , I was exposed to some unfamiliar python library. So this is also a learning journey. But this hands-on learning journey make the learning faster ,and I enjoyed it. Which told me ,”the practice is the best way to learn”. In the future ,this rule can be applied along all my future learning career .