SINGAPORE INTELLIGENT INSURANCE RECOMMENDER SYSTEM

PROJECT REPORT

Group Member Names
Zhang Yu
Li Sheng
Prashant Chaudhary
Veda Yogeesh

DOCUMENT VERSION <1.0>

21-April-2020

AUTHOR

Name
Veda Yogeesha

APPROVALS

Approval Date	Approved Version	Approver
02-May	1.0	Zhang
02-May	1.0	Prashant
02-May	1.0	Li Sheng

Table of Contents

1. Introduction

To make good decisions, it is necessary to possess information in a easily understandable manner which can be easily compared to make necessary judgements based on the individual needs. However, with the growing amount of information and features, it would be also challenging to the user to be able to apply his preferences to filter the relevant applicable details. Specifically, in the area of Insurance policies that have complex terminologies and multitude of features, it is not easily comprehensible to the user.

Business Value Created:

1. Intelligent and yet simple user interface to determine the most suitable policy by considering the user demographic details and preferences. And also recommend based upon the historical policy transactions.

2. Insurance Providers can also utilize the information captured to understand and evaluate the trend of user preferences and comparison of other Insurance provider policies to be more competitive in the market. This will assist them to drive business and operational decisions from data to maximize profitability and efficiency metrics.

1.1 Purpose of the document

Insurance is a means of protection from financial loss. An insurance policy is a contract used to indemnify individuals for covered losses and in managing cash flow uncertainty. Insurance industry has already several well-established players and in Singapore particularly there are several new entrants as well. Competition in insurance is important in providing competitive premiums and quality service. However, a lot of competition means that insurers have to differentiate themselves via quality, innovation or price. With the overwhelming number of plans, individuals will have more options and determining the right plan can be time consuming and might not result in choosing the best plan.

Traditional recommender systems usually provide a comparison of the different policies and do not take in to account the user preferences and past preferences. This recommender system will help to look through all the current available plans, past preferences, new policies in the market and customer demographics to identify the most suitable policy.

Market Research:

Lot of information on the insurance policies are available to educate users on the different kinds of insurance. However, all these websites require users to read through several pages of information, hence for a common man it would be time consuming and might not be easily comprehensible to digest the information. Also going through an insurance agent to

purchase the insurance policy would also mean paying commission, paying higher premium and might bind the user with a company which is not having a good portfolio.

Here are certain websites which provide guidance to the user in choosing the right policy and understanding of the various features, calculating the amount an individual will be required to shell out to insure his health and for his dependents.

1. **Life Insurance Association Singapore (LIA)** is the not-for-profit trade association of life insurance product providers and life reinsurance providers based in Singapore and registered by the Monetary Authority of Singapore (MAS). It provides information on the different kinds of policies, provides tools to determine the amount an individual should get insured.

2. **compareFirst** is a joint effort by Consumers Association of Singapore, MAS, LIA and money sense to enable consumers to compare the various life insurance products. It provides general product information on policies.

3. **MoneySense** is Singapore's national financial education programme, started in 2003. They help Singaporeans to manage their money well and make sound financial decisions on their own.

Though the above websites are impartial and share detailed information. Users can be coming from different backgrounds and of different generations. They might not be having the liberty or the ability to understand the different policies from different companies. Also, they might not have access to a company performance or insurance claim processing time which would also help to be a deciding factor for determination of the right policy. We have performed market research by searching several websites, meeting several customers who are unable to easily determine the best insurance policy which is recommended based on the user needs. Having considered all the above challenging issues which a user will require to go through, this application will determine the most optimal policy without going through any such hassle described above.

We have also performed knowledge elicitation by interviewing the domain experts to obtain a deeper understanding of the policies, current pain points and operations in evaluating the right policy.

1.2 Project Scope

A system to provide intelligent recommendation on the insurance policy the user can opt for,

1. Considering his preferences and evaluating the current policies to choose the optimal policy.
2. Historical data of other users who have previously purchased different policies to choose the most preferred one.

1.3 Scope of the document

This document is structured to provide you an understanding of how the different functionalities, tools are synchronized to operate in providing the solution to the business problem. Also there is information provided upon how the health insurance system in Singapore functions and how the different tools are utilized in the system.

Here is the list of core modules of the system which will briefed in detail during the later sections.

Sl no	Functionality
1	User Demographics and Preference capture
2	Normalization of the policies
3	Determination of optimal policy by OptaPlanner
4	User feedback collection
5	Data mining based on historic transactions by Orange
6	Dialog flow

1.4 Terms/Acronyms and Definitions

Stated here are some of the basic terms and the corresponding definitions relating to health insurance.

Term	Description
Deductible	A deductible is what you pay annually for health services before your insurance company pays its share. For instance, if you have a deductible of $1,000, your insurance plan might not start covering its share of your bills until you've paid $1,000 for healthcare in a given year. However, plans often cover the cost of things like preventive care doctor's visits even before you've paid your full deductible amount.
Premium	Your premium is what you'll pay the insurance company for the privilege of having an active insurance plan. Most people pay theirs every month, but your payments might be due once a quarter or once a year.
Co-payment	The co-payment (or co-pay) is the amount you owe each time you receive certain types of medical care. Co-pays can vary depending on the kind of service you're getting. For example, you may have to pay a $30 co-pay for each visit to your GP and $60 for each visit to a specialist.

Coinsurance	After you've met your deductible for the year you're not off the hook when it comes to medical bills. You'll generally face some amount of coinsurance. That's the percentage you'll pay of medical expenses. For example, you might meet your $2,500 deductible in May and from then on your coinsurance would be 20%. That means you would pay $20 of a $100 bill and the insurance company would pay the other $80.
Claim limits	There are limits to what you can claim under a policy. For example, limits may be included for all claims as well as for each illness, disability, per month, year, or for a lifetime.
Age limit	There is no age limit for MediShield Life. It covers you for life, and there is no age limit for entry into the scheme. Private insurance plans may have an age limit, and may not be available to you once you reach a certain age. Some health insurance policies provide cover for your whole life.
Policy riders	A policy rider may be added to an existing policy, which provides additional coverage.
Claim	a request by a plan member, or a plan member's health care provider, for the insurance company to pay for medical services.
Underwriting	the process by which health insurance companies determine whether to extend coverage to an applicant and/or set the policy's premium.
Pre-existing condition	a health problem that has been diagnosed, or for which you have been treated, before buying a health insurance plan.

1.5 Risks and Assumptions

Here are the assumptions undertaken for the current design of the system.

- Currently this application search is based on the private ward policies alone
- Data mining is performed based on 1000 previous transactions.
- Cognitive process of modelling and solving the decision problem involving the multiple criteria decision-making model using the Grey relational analysis
- Orange is utilized to perform data mining
- Dialogflow is implemented to support answer user queries and view the supporting FAQ from Insurance providers.

2. System/ Solution Overview

This section will provide the necessary details which has supported in obtaining the necessary understanding of the different concepts required for the design of this application.

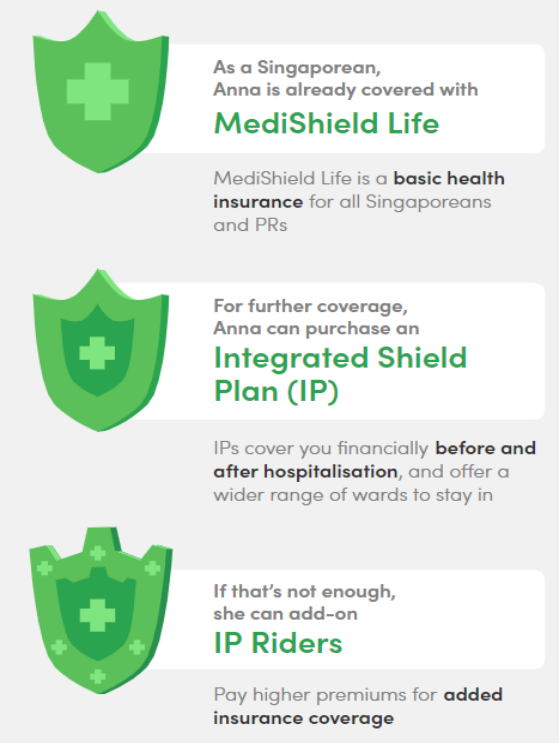
We will discuss upon the following areas:

a. Basic Health Insurance
b. Multi criteria decision making model

c. Intuitionistic Fuzzy sets
d. Dialogflow
e. OptaPlanner
f. Data Mining

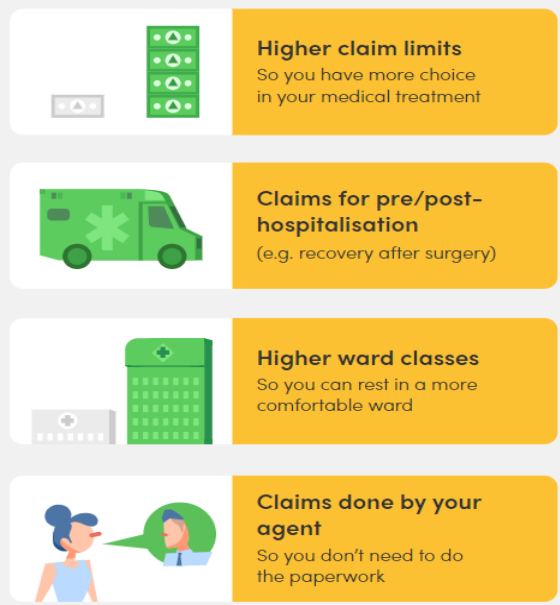
2.1 Basics of Health Insurance in Singapore

uml



MediShield Life is the most basic sort of health insurance and covers basic healthcare in public hospitals' lower-class wards. However, addition of Integrated Shield Plan (IP) gives you the option to get treated in better class wards and private hospitals, managed by the private insurer.

Fig 1: Depicts the levels of coverage by the insurance plans

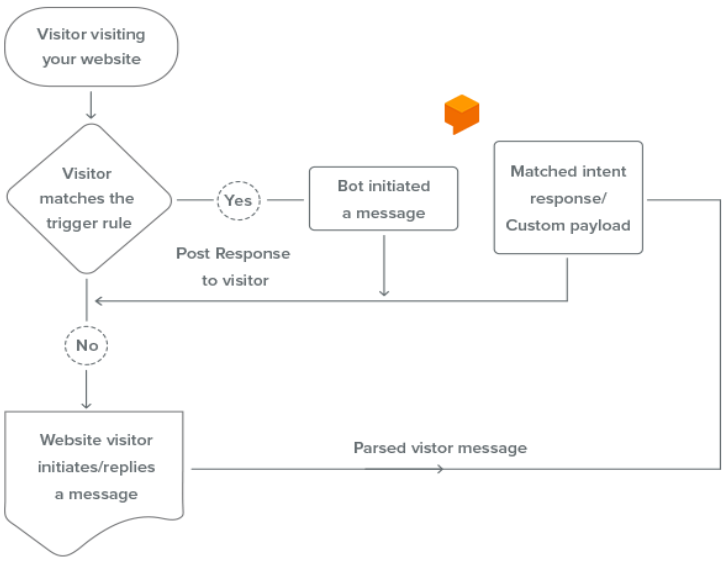


An Integrated Shield Plan (IP) provides coverage on top of your MediShield Life. Different types of IPs provide different coverage. Some cover you for a higher-type ward (A/B1) in a public hospital, while others cover admission into a private hospital.

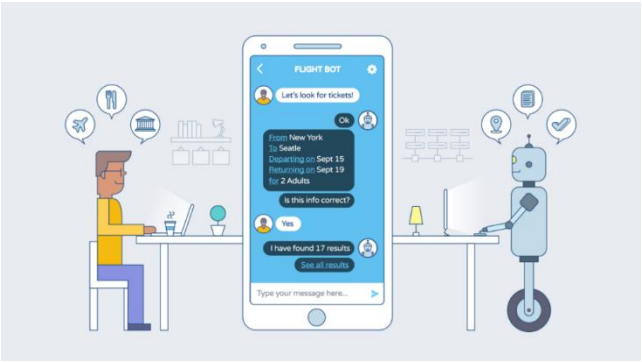
Fig 2: Depicts the benefits of having the Integrated Shield plan

2.2 Dialogflow

Dialogflow is user-friendly, intuitive natural language processing Google service tool. It is an end-to-end development suite for building conversational interfaces for websites, mobile applications, and messaging platforms. It can be used to build chatbots, voice assistants, etc., especially relevant to those that are capable of having natural and rich interactions with the users. It is also powered by machine learning to recognize the intent and context of what a user says, allowing a conversational interface to provide highly efficient and accurate responses.



The above picture represent the flow chart for the Dialogflow execution



Here the user may want to check upon certain frequently asked questions about the policies/ insurance providers. By associating the FAQ and policy references with Dialogflow, the bot can respond to user queries.

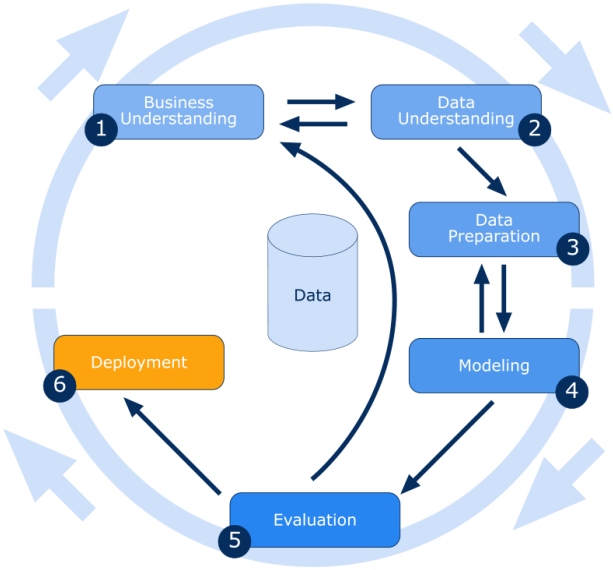
Benefits of integrating Dialogflow:

- Easy to tailor-made bots that suit your business aspects to support the FAQ.
- It is easily configurable. Once integrated with Dialogflow, the bots you have created in the platforms is all set to assist your website visitors.
- Dialogflow doesn't require any coding knowledge. The interface will guide you to build your bot as you wish without any basic coding knowledge and allows you to integrate with your application in minutes.

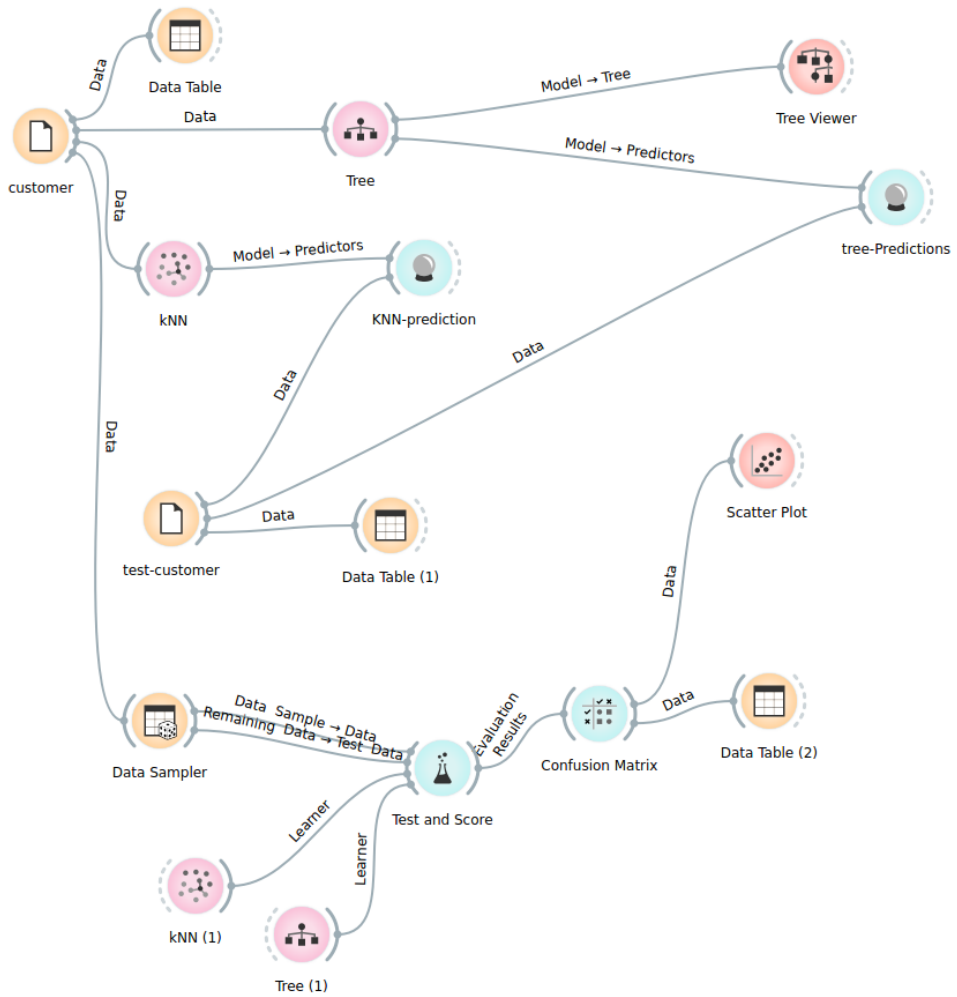
2.3 Data mining

Based on the recommendation paradigm of the supervised learning, we determine the most optimal policy based on the past transactions. In the pursuit of extracting useful and relevant information from data, process of data exploration, preprocessing, modeling, evaluation and knowledge extraction is performed. Exploratory visualization using Orange tool helped to comprehend various patterns in the data set.

CRISP DM, framework was used to perform the data mining process.



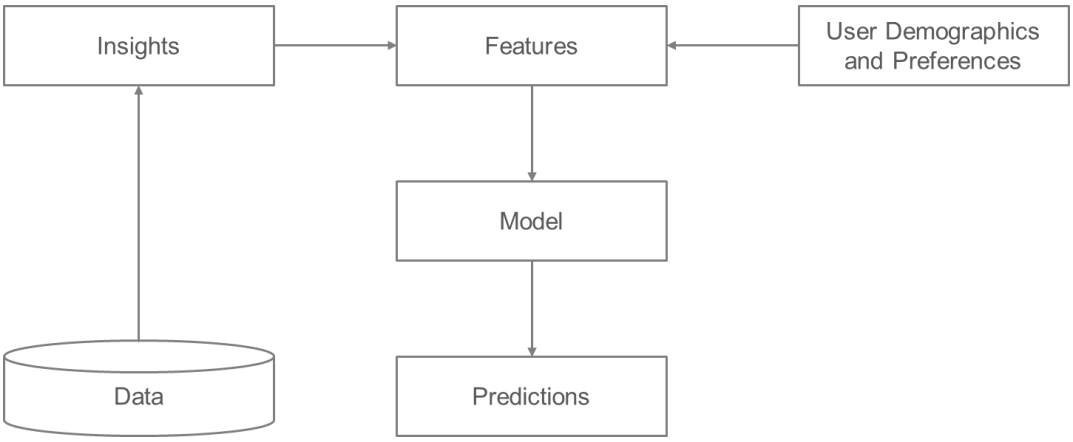
- Business Understanding : The business objective is to determine the most suitable policy based on the historical transactions performed by different users previously. This result will be utilized by the subject user to make informed decisions.
- Data Understanding : For the data, we have the historical transactions of the several users and their policies purchased. For training and test sets, we have taken in to account several factors such as user's characteristic data (age, sex, net worth..), preferences and policy details. It's a lot of information, some of the details are missing, partially captured and in different formats. There are certain key details which are required for our analysis, which would help to determine the policy.
- Data Preparation: In this stage, as the name suggests the data requires to be prepared for further analysis. This acquired data had to cleansed, formatted to achieve a proper organized data structure. Certain records which were missing critical data had to be omitted since this would not give much information and would become outlier. Data had to be partitioned in to 2 sets for training and testing.
- Data Analysis and Modelling: Orange tool was used for data analysis. We had narrowed upon the KNN model based on the performance and accuracy observed in the Orange tool.



Orange implements functions for construction of classification models, their evaluation and scoring. Both Decision Trees and KNN model was utilized for the study. Neighbors-based classification is a type of *instance-based learning* or *non-generalizing learning*: it does not attempt to construct a general internal model, but simply stores instances of the training data. Classification is computed from a simple majority vote of the nearest neighbors of each point: a query point is assigned the data class which has the most representatives within the nearest neighbors of the point. The k-neighbors classification in KNeighborsClassifier is the most commonly used technique. The basic nearest neighbors regression uses uniform weights: that is, each point in the local neighborhood contributes uniformly to the classification of a query point. The default value, weights = 'uniform', assigns equal weights to all points. weights = 'distance' assigns weights proportional to the inverse of the distance from the query point. Alternatively, a user-defined function of the distance can be supplied, which will be used to compute the weights.

Evaluation: In this phase, the model results were evaluated in the context of the business objectives defined in the first phase. Based on the 2 models used, the classification accuracy and certainty was found to be better in KNN model. Hence this model was used for determining the best suitable policy.

Finally after the evaluation was successful, the code was integrated and deployed.



The above diagram provides a high level pictorial presentation on how prediction was performed.

2.4 Multi criteria decision making model

It is a cognitive process of modelling and solving the decision problem involving the multiple criteria using the Grey relational analysis model. The objective of this model is to support the user in providing the best alternatives when multiple criteria is involved. The raw inputs might be in different ranges and their values in different units. While the distribution can be tamed as described above, the values might still need to be brought into comparable ranges. For this, the features might need to be transformed to some min-max range (so min is always 0 and max is always 1) or z-scored values could be used (so the mean of each feature is zero and standard deviation is 1). Such

transformations then let the model do the actual job of learning the relative importance of these features instead of forcing them to also compensate for these feature differences.

Here is the explanation on the steps to be followed:

• Defining the object of decision and perform normalization:

The object of the decision is to list all the set of all candidate alternatives for which the decision has to be made.

(Grey relational analysis)

Normalize the each response considered using

$$x_i(k) = \frac{\max y_i(k) - y_i(k)}{\max y_i(k) - \min y_i(k)}$$

• Defining a consistent family of criteria and apply user weightage :

The performance of alternatives is determined based upon a set of user input preference captured values. Following which, the user weightage is applied to all the applicable policies which will be used to evaluate the most optimal policy. IFS logic is used to calculate the user weightage value.

2.6 Intuitionistic Fuzzy sets (IFS)

Is an extension of fuzzy sets. Unlike the Fuzzy set, wherein there the membership element is a single value between 0 and1. Here it incorporates 3 elements namely Membership degree, non-membership degree and Hesitancy margin.

Linguistic Scale/ Decision Maker	Membership degree	Non-membership degree	Hesitancy degree
Very Important	0.9	0.1	0.05
Important	0.75	0.2	0.05
Medium	0.5	0.4	0.1
Unimportant	0.35	0.6	0.15
Very Unimportant	0.1	0.8	0.1

User expresses his preferences for the policy preferences, these are transformed into a measurable values (weights) using the linguistic variables defined in the above table. Assume that decision group contains ‘n’ decision makers. The importance of the decision makers are considered as linguistic terms expressed in intuitionistic fuzzy numbers. Let Dj be an intuitionistic fuzzy number for rating of jth decision maker. Then the weight of ‘j’th decision maker can be obtained as:

$$\lambda_j = \frac{\left(\mu_j + \pi_j\left(\frac{\mu_j}{\mu_j + v_j}\right)\right)}{\sum_{j=1}^{n}\left(\mu_j + \pi_j\left(\frac{\mu_j}{\mu_j + v_j}\right)\right)} \quad for\ all\ j = 1, 2, \ldots, n$$

2.5 OptaPlanner

OptaPlanner is a lightweight, embeddable constraint satisfaction engine which optimizes planning problems. It is also known as *Constraint Satisfaction Programming* (which is part of the *Operations Research* discipline).

Usually, a planning problem has at least two levels of constraints:

- A hard constraint must not be broken.
- A soft constraint should not be broken if it can be avoided. For example: Teacher A does not like to teach on Friday afternoon.

These constraints define the score calculation of a planning problem. Each solution of a planning problem can be graded with a score.

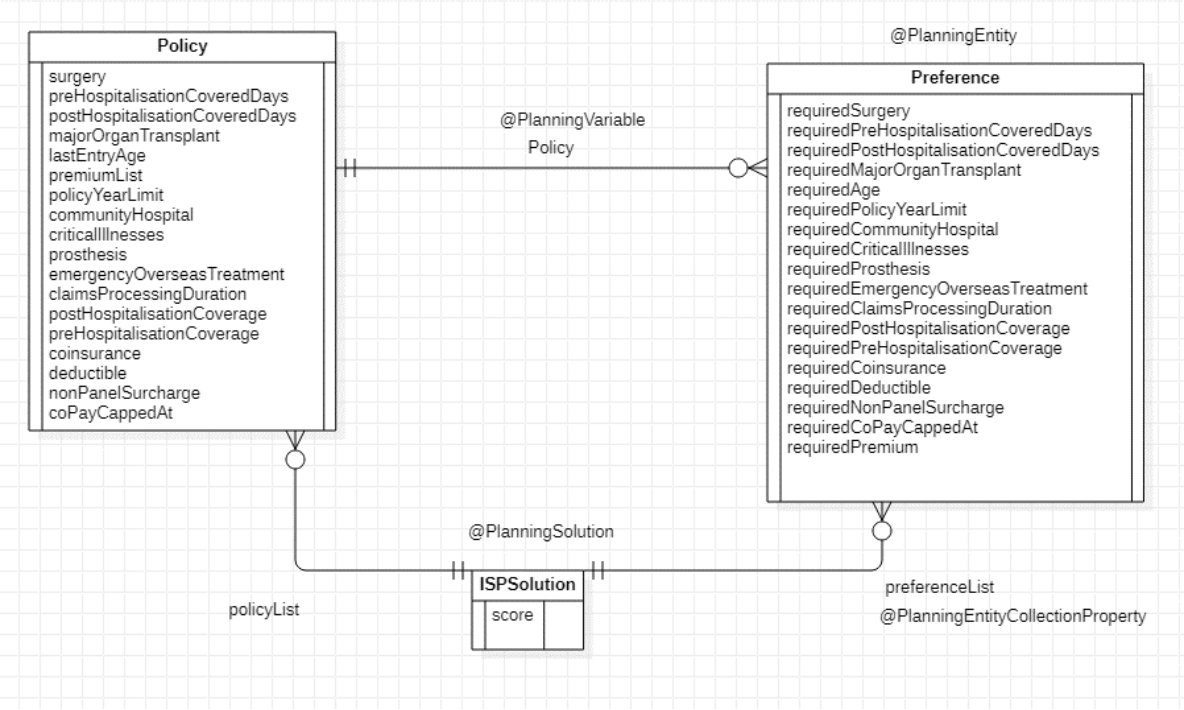


Fig: Opta Class Diagram

The picture above depicts the design in OptaPlanner,

- Planning entity: the class (or classes) that changes during planning [Preference].
- Planning variable: the property (or properties) of a planning entity class that changes during planning. In this examples, that's the property [Policy].
- Solution: the class that represents a data set and contains all planning entities. [ISPSolution].

3. Functional Specifications

This section will explain the overall functionality of the system.

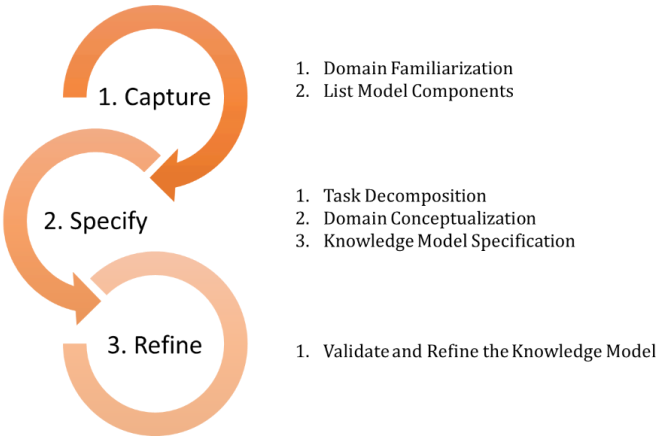
We begin the section by providing an overview of the knowledge modelling process and followed by use cases, flow chart, activity diagram, field list, rules and followed by detailed functional explanation.

3.1 Knowledge Modelling

The diagram mentioned here, capture the differences in the Data, Information and Knowledge*. How the data is synthesized to represent useful knowledge is the core objective of the knowledge modelling process.

	characteristic	*example*
Data	uninterpreted raw	...---...
Information	meaning attached to data	SOS
Knowledge	* attach purpose and competence to information * potential to generate action	emergency alert ->start rescue operation

For an synthetic tasks the system does not yet exist: the purpose of the task is to construct a system description. Knowledge modelling helps to construct an abstract description of a system of how the different knowledge components play in problem-solving, in a way that is understandable for humans. Knowledge modelling is 3 stage process as pictorially presented here,



1. Knowledge Capture: In this initial phase we began with the exercise of encompassing the data. And this data can be categorized primarily into the following categories:

Sl No	Source	Description	Approach to obtain
1	Insurance Provider Website	Provides information on the Insurance Provider, Policies, policy features	Performing web search on the Insurance Provider websites
2	Consumer	Provides information on their policy preferences and pain points of choosing an agent and determining the optimal policy	Elicitation of this information by interview with the users and forums
3	Subject Matter Expert	SME was able to provide us with a better understanding of the insurance policies and how the user demographics and preferences play a important part for an agent to determine the most suitable policy.	Elicitation of this information by interview with the SME

2. Knowledge Specification: goal of this stage is to get a complete specification of the knowledge model.

   The following activities was carried out to build such a specification:

   a. choose a task template;
   b. construct an initial domain schema;
   c. specify the three knowledge categories.

   Task template can be captured as below:

Task Features	Explanation
Nature of Output	How to recommend a policy
Nature of Input	Policy features, User demographics, User Preferences, Insurance Provider details
System Nature	Analyzing the data and constructing a recommender system
Constraints	Matching user preferences to find the suitable policy

   Based on the above task template, inference can be analyzed as consisting of the problem-solving factors arranged in a hierarchical structure, it provides a abstraction mechanism over the details of the reasoning process.

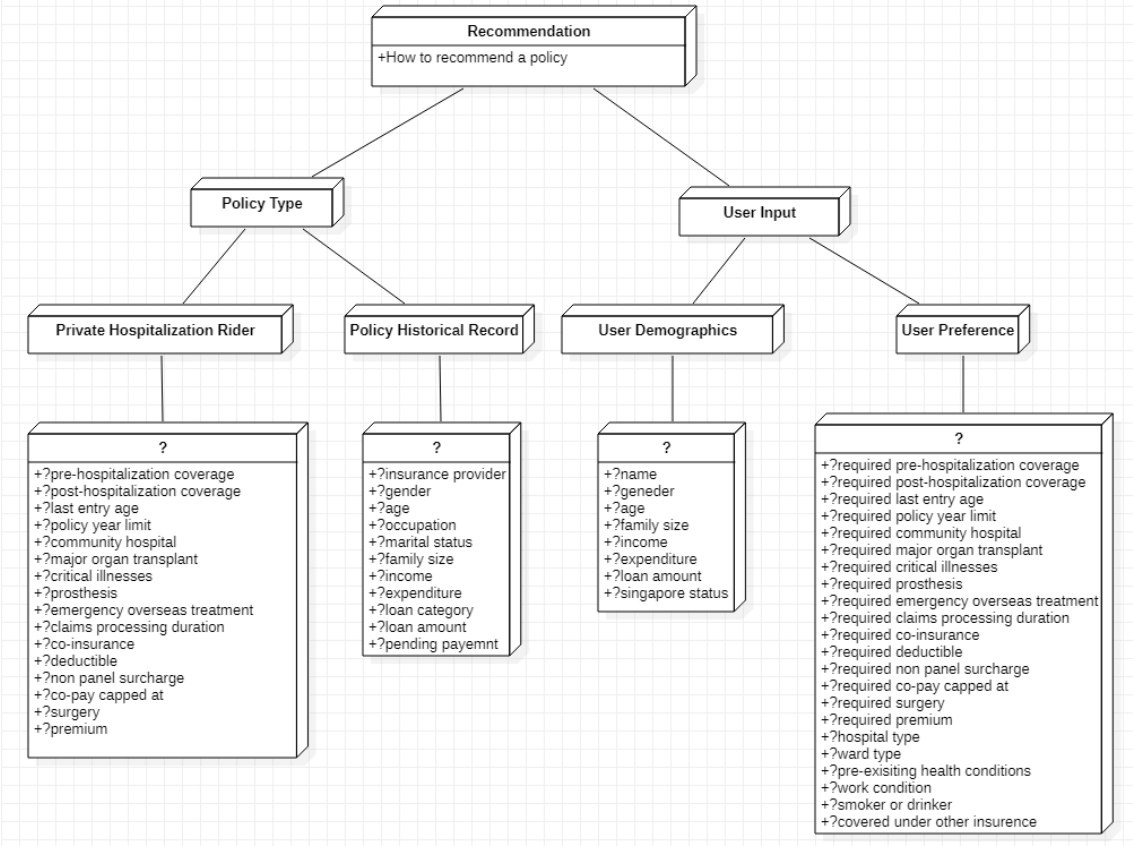
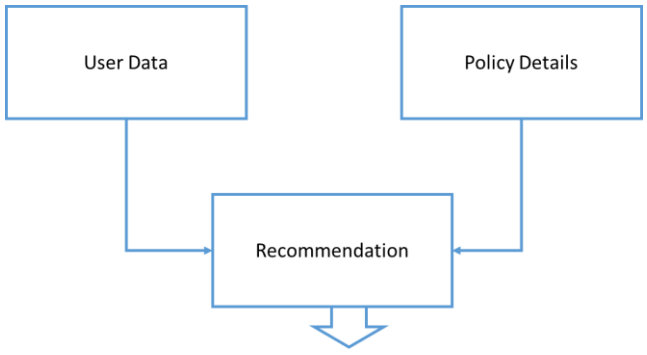


Fig. Inference Diagram of the Recommendation policy

Based on the inference, we have constructed the initial domain schema.



3. Knowledge Refinement:

   Is a process of validating the knowledge model to verify whether the model is right? This was performed by testing the system built based on the knowledge model. After testing the various domain information with the respective knowledge model action for the expected behavior. Necessary fine tuning and fix was performed on the model built to ensure the expectation is achieved.

3.2 Use cases

Description of how a person who actually uses that process or system will accomplish a goal.

Here we provide the details upon the different possible use cases:

1. User only provides the demographics to determine the optimal policy
2. User provides the demographics and the preferences to determine the suitable policy
3. User to use Dialogflow
4. User would like to check upon the previous transactions history to determine the suitable policy
5. User can choose to restart the exercise by re-choosing the preferences again.

UC-1	Current Optimal Policy Determination
Primary Actor(s)	User A
Initiation	User will capture the below details 1. Demographic details 2. Policy preference information
Conditions	User chooses to provide the policy preference details
Post-conditions	System will consume the user provided preference details to perform the intelligent evaluation of all the policies to determine the most optimal policy.

UC-2	Historically Preferred Policy Determination
Primary Actor(s)	User A
Initiation	User will capture the below details 1. Demographic details
Conditions	User choose to look upon the historically preferred policy based upon his basic information details.
Post-conditions	System will consume the basic user details to perform the intelligent evaluation of all the historical data to determine the most optimal policy.

3.3 Field level specifications

Buttons, Links and Icons:

Button, Link, Icon Label	OnClick Event	Visible	Enabled Vs Disabled	Navigate To	Validation	Dependencies
Start Survey	**User will initiate the operation to fill the form details.**	**Yes, always**	**Always Enabled, after selection a new form will be displayed to capture the user details**	**User Demographic page**	**N/A**	**N/A**
Next	**User will be navigated to the next form page**	**Yes, always**	**Always Enabled, after selection a new form will be displayed**	**User Preferences Page**	**N/A**	**User should have completed capturing the mandatory information in the existing page to proceed.**

Form Elements:

Field Label	UI Control	Mand?	Editable	Data Type	Value Set	Data Example	Data Source
Name	Textbox	Mandatory	Yes	Alpha-numeric	none	Li Sheng	User entry
Age	Dropdown	Mandatory	Yes	Numeric	Numbers ranging from 20 to 75	25	User selection
Gender	Dropdown	Mandatory	Yes	Characters	Male/ Female	Male	User selection
Singapore Status	Dropdown	Mandatory	Yes	Characters	Singaporean/ PR/ Foreigner	Singaporean	User selection
Expenditure	Selection	Optional	Yes	Numeric	Any amount	10000	User selection
Income	Selection	Optional	Yes	Numeric	Any amount	100000	User selection
Loan Amount	Selection	Optional	Yes	Numeric	Any amount	1000000	User selection
Family Size	Selection	Optional	Yes	Numeric	Number of members including self	5	User selection
Hospital Type	Selection	Mandatory	Yes	Characters	Hospital Type List	Govt and Restructured/ Private	User selection
Health Condition	Selection	Mandatory	Yes	Characters	Yes/ No	Yes	User selection
Hazard Profession Type	Selection	Mandatory	Yes	Characters	Yes/ No	Yes	User selection

Habit Type	Selection	Mandatory	Yes	Characters	Yes/ No	Yes	User selection
Other Insurance	Selection	Mandatory	Yes	Characters	Yes/ No	Yes	User selection
Co-Payment Percentage	Selection	Mandatory	Yes	Numeric	5%, 10%, 20%, 25 % and above	10%	User selection
Co-Insurance Payment	Selection	Mandatory	Yes	Numeric	3000 SGD, 5000 SGD, 12000 SGD, Full Amount	10%	User selection
Deductible	Selection	Mandatory	Yes	Numeric	0 SGD,1000 SGD,2500 SGD,3500 SGD	10%	User selection
Non-Panel Surcharge	Selection	Mandatory	Yes	Numeric	3000 SGD, 5000 SGD, 12000 SGD, Full Amount	10%	User selection
Preferences	Selection	Mandatory	Yes	Characters	Yes/ No	Yes	User selection
Ward Type	Selection	Optional	Yes	Characters	Ward Types List	Yes	User selection
Premium Amount Payment	Selection	Optional	Yes	Characters	Premium Amount Range List	Yes	User selection
Pre-Hospitalization Coverage	Selection	Optional	Yes	Numeric	Pre-Hospitalization Coverage Days Range List 3m, 4m, 6m, 1Y	Yes	User selection
Post-Hospitalization Coverage	Selection	Optional	Yes	Numeric	Post-Hospitalization Coverage Days Range List 3m, 4m, 6m, 1Y, 2Y	Yes	User selection
Annual Coverage limit	Selection	Optional	Yes	Characters	Annual Coverage limit Range List >400 K, 400 K-800 K, 800 K to 1.2 M, >1.2 M	Yes	User selection

3.4 Form Business Rules and Dependencies

Drools Rules:

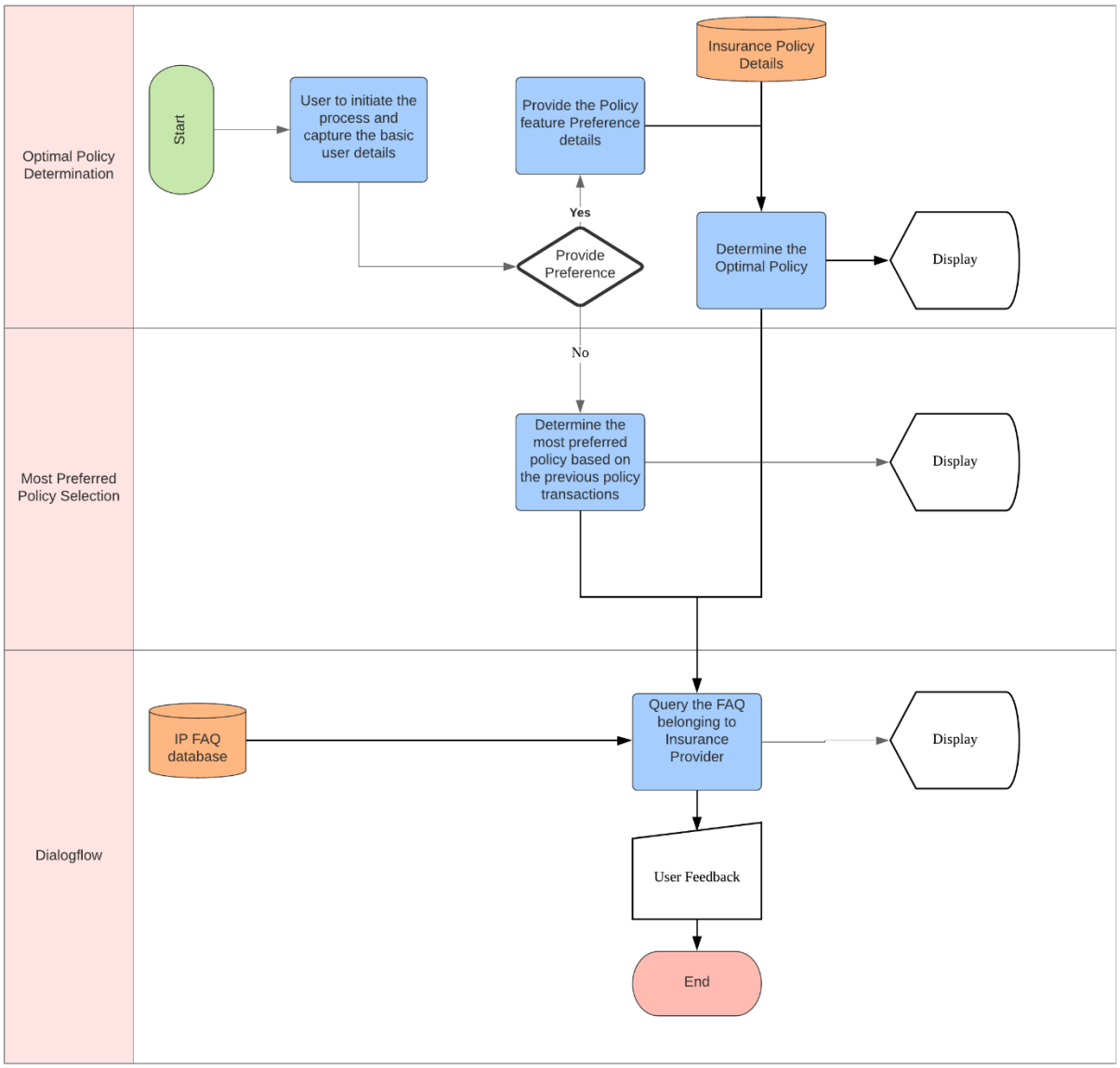
Rule Name	Validation / Business Rules
Invalid Age	To ensure the age of the user is not greater than the applicable last entry age for the policy which is >75.
Status	To validate whether the Individual is a Foreigner or Singaporean or Permanent Resident

Opta Rules:

Rule Name	Rule Type
Hard last entry age	Hard
Post-Hospitalization Covered Days	Soft
Pre-Hospitalization Covered Days	Soft
Policy Year Limit	Soft
Critical Illnesses	Soft
Major Organ Transplant	Soft
Community Hospital	Soft
Premium Amount	Soft
Surgery	Soft
Co-Pay Capped At	Soft
Non-Panel Surcharge	Soft
Deductible	Soft
Co-insurance	Soft
Pre-Hospitalization Coverage	Soft
Post-Hospitalization Coverage	Soft
Claims Processing Duration	Soft
Emergency Overseas Treatment	Soft
Prosthesis	Soft

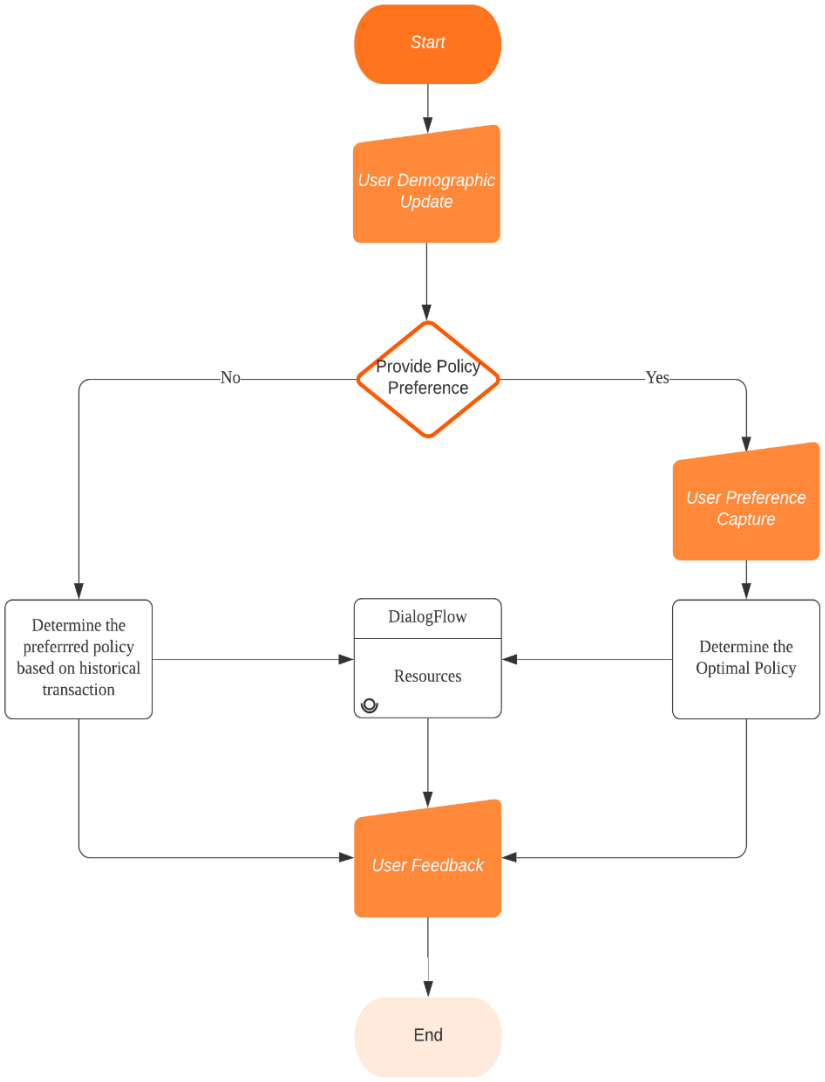
3.5 UML Diagrams

3.5.1 Activity Diagram



The above activity diagram helps to describe how the different activities are coordinated to provide a service. The actor/ user will choose to perform the desired activity for the recommendation of the policy, could also tap on to Dialogflow for any queries and eventually provide feedback.

3.5.2 Flowchart



The flowchart diagram helps to understand the workflow designed in the system.

1. To capture the basic user information details.
2. User will make a decision on whether to determine the recommendation based on the current insurance providers policy information or to check upon the historical information.
3. Upon deciding to provide the preferences, system will intelligently determine the most optimal policy based on the preferences provided.
4. User can also opt to determine the preferred policy based on the historical transactions performed.
5. User can choose to provide the feedback and restart the process again or close the application.

4 System Model

SDLC (Software Development Life Cycle) – Waterfall model is followed to build this recommender system. With the well-structured flow of phases, it has helped to build a high quality software ready to use.

Initially we begin with the clear definition of the problem statement for Insurance policy determination. Followed by the knowledge modelling study which has helped to guide us upon the constructive problem solving methodology. During the design phase, we discussed on the various possible approaches which can be utilized for the model, also how to provide the best experience to the user and drive profitability for Insurance Providers. Logical data flow diagram and a draft requirement blue print of design specification is documented based on the knowledge model.

The picture below presents logical data flow diagram, this describes how data flows and how it is utilized for the recommendation.

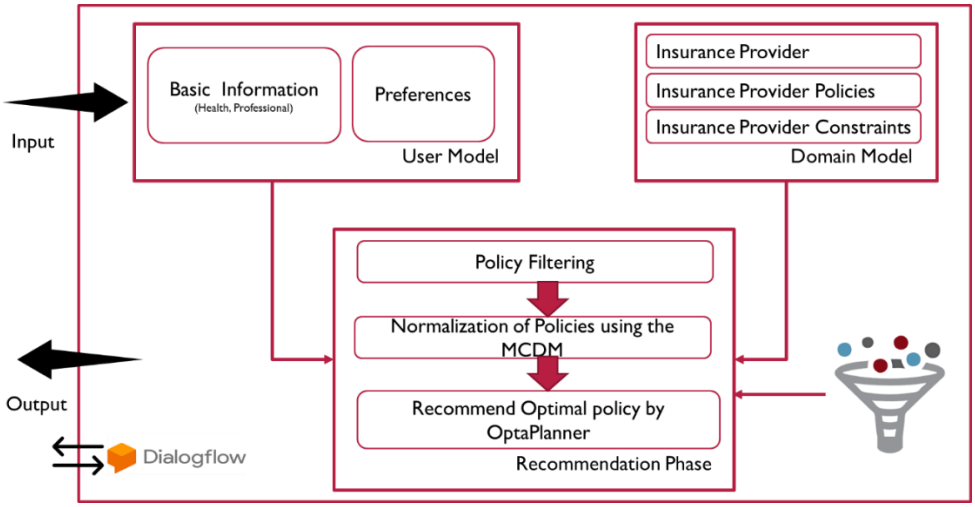
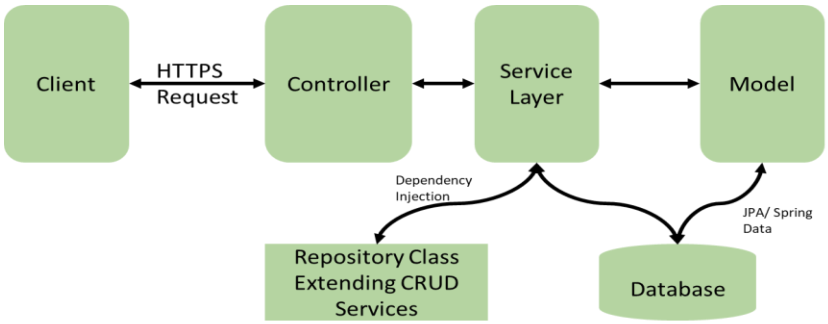


Fig: Local Data Flow Diagram

During the Development phase, actual development started. Development is performed using the Spring boot model. Spring boot a module of Spring Framework is used to create a standalone, production-based application. It follows a layered architecture in which each layer communicates with the layer directly below or above. The picture below represents the Spring Boot flow architecture.



4.1 System Architecture

The system architecture diagram shown below is used to show relationship between the different components of the system. It is built upon the main principles of scalability, user friendliness, simplicity, high intelligence, robustness and modularity.

In addition, we have used several external systems as mentioned here,

1. KIE (to invoke JBPM, Drools and OptaPlanner) – helps with the workflow, reasoning and intelligent constraint solving.
2. Orange - for data exploration and analysis
3. Dialogflow – for answering the user queries associated with the Insurance providers.

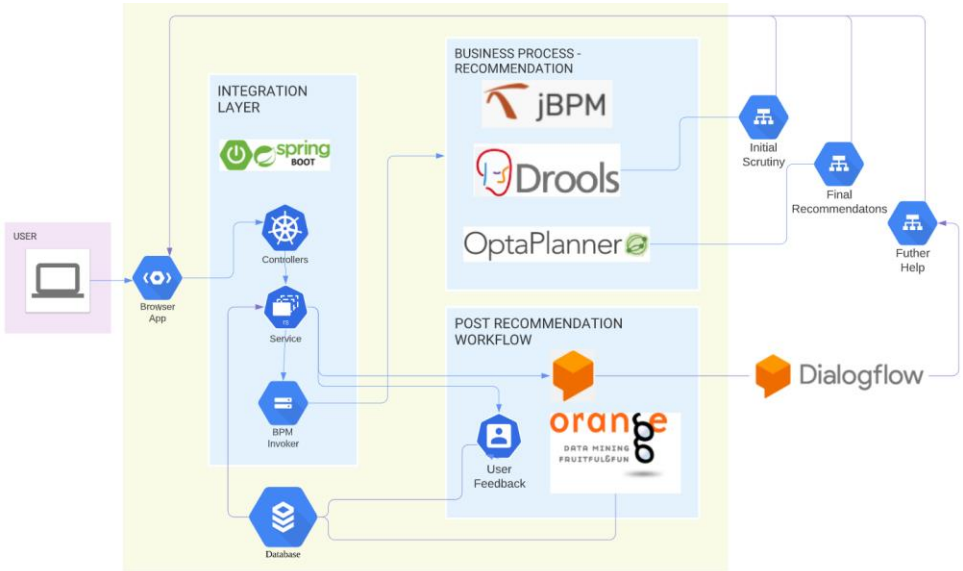


Fig: System Architecture Diagram

The key components of the application are:

1. Artifacts generated from KIE component for,
    - OptaPlanner for constraint solving.
    - jBPM, Drools for workflow managements system which executes business processes, converting business logic into assets such as cases, processes, decision tables and to execute the rules.
2. Model: Java package containing Data Objects for persistence storage. Built with Java Persistence API (JPA).
3. Service: Backend application to expose REST based API for Frontend application usage.
4. Maven for dependency management and spring boot framework for application runtime.
5. Front end design with HTML.
6. MySQL Database to store all the information using the relational model that is structured in tables. Database normalization is performed to reduce data redundancy and improve data integrity.

    Here is the snapshot of the entity relationship diagram of the recommender database,

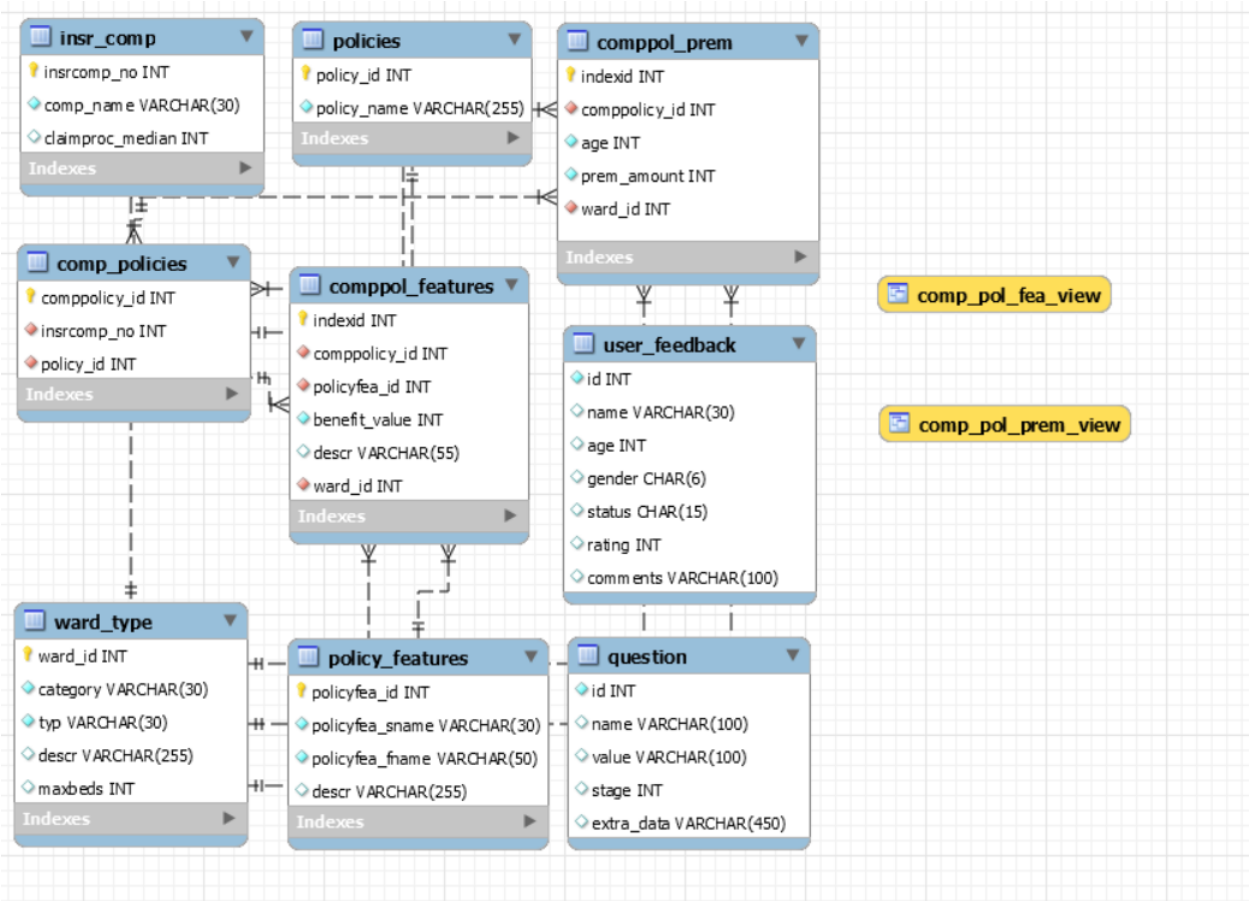


Fig: Entity Relationship Diagram

Test Model: To ensure the business objective is met, rigorous testing is performed at every stage of development and integration.

There are two types of testing, that falls under this category.

- Unit Testing: It is the testing of an individual unit or group of related units. It is often done by the programmer to test that the unit he/she has implemented is producing expected output against given input.
- Integration Testing: It is a type of testing in which a group of components are combined to produce the output. Also, the interaction between software and hardware is tested if software and hardware components have any relation

4.2 Mock-up

Mock up screens are captured in the Appendix section.

5 References

We have referenced the articles of the course materials and several papers on the MCDM, GRA and IFS. Her are the key papers which has helped us to build a better understanding of the various principles and methodology.

1. Roy B. Multicriteria methodology for decision aiding. Springer Science & Business Media.
2. Boran FE, Genc S, Kurt M, Akshay D. A multi-criteria intuitionistic fuzzy group decision making for supplier selection with TOPSIS method.
3. Kuo Y, Yang T, Huang GW. The use of grey relational analysis in solving multiple attribute decision-making problems. 2008

6 Appendix

6.1 Future enhancements

1. Real time integration with IP - By having the real time integration, we will be able to support the users the real time policy information, FAQ. Also will have the ability to provide information on the agents supporting the respective Insurance Providers

2. Feedback learning mechanism – We wanted to build a feedback model, which would help the user to identify more optimal policy based on the feedback information.

6.2 Matrix of the knowledge learnt with the system development cycle

Here is a snapshot of the different skills and knowledge learnt in this module and how it has been applied during the various stages of the lifecycle of the system development.

Machine Reasoning	Reasoning Systems	Cognitive Systems
Requirement Phase we utilized the knowledge modelling to understand the problem and provide the reasoning required to recommend the policy on the knowledge	Design Phase, based on market research and the data available we utilized the data mining methodology of CRISP framework, used the KNN model as the preferred approach.	Dialogflow, is used as a chatbot to respond user queries on the insurance policy details.
Intuitionistic Fuzzy Set, this has been used to evaluate the importance of the policy features to the user.	Orange tool for the data mining	MCDM - GRA, cognitive process of modelling and solving the decision problem involving the multiple criteria using the Grey relational analysis model

6.3 System Guides

Detailed system guide has been prepared the different type of users,

1. System end-users - users who would like to determine the most optimal policy

2. Technicians and Technology enthusiast – who would like to deploy the system for their study or future enhancement purposes.

The documents are shared in the Github folder of this project.