

Optimization of house purchase recommendation system (HPRS) using genetic algorithm

K.A.F.A. Samah, I.M. Badarudin, E.E. Odzaly, K.N. Ismail, N.I.S. Nasarudin, N.F. Tahar, M.H. Khairuddin

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (Kampus Jasin), Malaysia

Article Info

Article history:

Received Mar 1, 2019

Revised Jun 28, 2019

Accepted Jul 22, 2019

Keywords:

Genetic algorithm optimization

House budget

House preferences

House purchase

recommendation system

(HPRS)

ABSTRACT

This paper presents the optimization of house purchase recommendation system (HRPS) using Genetic Algorithm. Everyone in this world has their own dream house and plan to purchase it depending on their budget and based on the house preferences. Homebuyers face problem in comparing the house property websites according to the factors during the house survey. Subsequently, it is time-consuming in making the decision and they need to bear the transportation cost as they will need to travel to one house developer office to another. In addition, some of the homebuyers felt disappointed when their expectations were not met. Thus, in order to optimize the preferences, in this paper, we present a web-based house purchase recommendation system (HPRS) using a genetic algorithm. Then, it follows by test the functionality and usability of the system. As a result, the system found to function accordingly and obtain more than the average score of system usability scale testing. For further research, it is recommended to add more data to the database and compare with other algorithms.

*Copyright © 2019 Institute of Advanced Engineering and Science.
All rights reserved.*

Corresponding Author:

K.A.F.A. Samah,

Faculty of Computer and Mathematical Sciences,

Universiti Teknologi MARA (Kampus Jasin), 77300 Merlimau, Melaka, Malaysia.

Email: khyrina783@melaka.uitm.edu.my

1. INTRODUCTION

A house is a home, structure or building that functions as a shelter from the rain and heat for humans [1] and acts as a basic need for every human in the world where most daily activities of a human occur at home [2]. There are two types of houses which are “Landed” and “Non-Landed” property [3]. Example of “Non-Landed” houses are condominium, flat, and apartments. Each homebuyer have their own house preferences when they want to purchase a house and they strongly prefer to purchase it if truly meets their housing preferences. It is considered as a tough decision as the process involved long duration for the commitment from the first down payment until the monthly payment for the housing loan [4]. In terms of location, homebuyers preferred to live in urban area so as to be close to their workplace and services [5]. Housing purchasing trend in Malaysia is determined mostly by demographic, price and location [6-7]. Locational qualities like accessibility to city centre and distance from work place to the house are one of the main concerns for house buyers in urban areas [8]. Furthermore, facilities like public transport, supermarket, hypermarket, hospital, and clinic are the factors that homebuyer will consider when purchasing a house [9].

However, there were a few problems encountered in the process of buying a house. Firstly, the internet has become a platform for homebuyer to search for the house. The National Association of Realtors (NAR) stated that 88% of homebuyers choose the internet when searching for house information [10]. They do not know the house property websites when doing the survey, therefore they will search all over the internet to find the website based on the desired house location. Homebuyer would have to spend more time when doing a house survey from different house developer until it meets their house preferences [11].

Through a survey conducted in Melaka involving 52 participants, 81.1% of respondents use the website as a platform to do their house survey and 73.6% of respondents claimed difficulties in doing the survey for the house. The increasing of house price too high has affected the ability of the homebuyer to buy a house [1]. Through the survey that had been conducted, it shows that 48.1% respondent had a salary between RM1,000-RM5,000, it claimed that high price houses are not affordable for a homebuyer with a low-income salary and they need to find other alternatives. Housing affordability refers to a condition where people save their income to buy a house, as well as to pay for other expenses. The middle-income people are concern because affordability not only affects their ability to become a homeowner but also the types of home that they are able to purchase [12].

Each homebuyer has their own expectations on the house that they purchase yet there was a contradiction with the attributes provided by the housing developers and causes dissatisfaction [13]. One of the most important factor, when a homebuyer wants to buy a house is house preferences. In the same survey, it shows that 86.5% of respondents will refer to the provided facilities as one of the factors in decision making to purchase a house. Besides, other commitments such as car loan, bank loan, personal saving and others have affected the housing affordability for the homebuyers [14]. This was proven whereby 71.2% of them agreed that they have difficulties to calculate the budget due to own commitments. Other contributing factor like good environment of residential neighbourhood [15] preferably with gated neighbourhood, is belief can help to reduce crimes in residential areas [16]. In addition to that, homebuyer will also consider house features such as building's shape, the appearance and the dimensions with different size of lots, the built-up area and scaling of the housing [17].

Recommendation techniques help to provide suggestions to the user. The aim of the recommendation techniques is to generate a result to the user when they making a decision while faced with different choices [18]. The three techniques that suitable for this project is the Rule-Based Expert System, Fuzzy Logic, and Genetic Algorithm. A rule-based system is a type of expert system. This type of expert system uses rules to display the information into the system. A rule-based system can be used for the purpose of knowledge discovery, which means to extract information or knowledge discovered from data [19]. Although rule-based system can be applied easily, the number of rules can become complicated if involve in a larger problem [20]. A nonlinear mapping of an input data is set to a scalar output data can be defined as Fuzzy Logic (FL). It is a suitable tool that can deal with a problem that has dynamic behavior and helps to provide the best solutions to problems that involve the manipulation of several variables [18]. FL contains 4 main parts which are Fuzzification, Fuzzy Inference Process, and Defuzzification. GA is the search algorithm basis on the developmental ideas of natural selection and genetic. Besides providing an alternative method to solve a problem, GA also performs better than other traditional methods in most of the problem [21]. To calculate the similarity between two contexts from the multidimensional perspective, GA is used to achieve the best optimization value [22]. Three specifications have been used as a comparison between the recommendation techniques which are understanding, receive many inputs at one time and run in a single run, and modularity. Rule-based easy to understand and modular but not able to receive many inputs at one time and not able to run in a single run. FL has more disadvantages since it only easy to understand. As a comparison result between three techniques, GA is the most suitable for this project as it easy to understand, able to receive many inputs at one time and run in a single run and modular.

There was two existing buying house recommendation system that has been developed to help homebuyer to do the house survey; iProperty Malaysia dan Propertyguru. These applications are available on the website and we have compared between both system with our HPRS. Two types of comparison specifications been used which are the input from homebuyers and the output. It found that iProperty Malaysia able to receive six inputs such as any states, type of property, min price, max price, built-up size and no of bedrooms and the result prompted the list of houses that meet the specification input from homebuyers. Even the Propertyguru able to receive twelve inputs such as areas in Melaka, type of house, min price, max price, no of bedrooms, floor area, PSF, tenure, furnishing, floor level, listed and constructed, yet the output is the same as iProperty Malaysia. Nevertheless, the HPRS able to receive many inputs such as areas in Melaka Tengah, input the price, types (landed or non-landed), no of bedrooms, input includes the facilities around the housing area (hospital, transportation, school, etc) and the unique of the system is allowed user to provide input the budget that been approved by bank and the output following the preferences that within the budget.

Thus, in this paper, we designed and developed a web-based house purchase recommendation system (HPRS) using Genetic Algorithm (GA) in order to overcome the problem, a house purchase recommendation system developed with the hope it able to help the homebuyers to do house survey as per their budget and house preferences. This paper is organized as follows. In section 1, we have reviewed some background and literature of the developed system. The methodology is explained in section 2. The results and discussion are in section 3 and finally section 4 for the conclusion of the paper.

2. RESEARCH METHOD

In this study, the research methodology adopted Agile methodology. It is a slight methodology for software development. Agile model believes that every project needs to be handled differently and the existing methods need to be accommodated to best suit the project requirements. Rather than a single large process model implemented in SDLC, the development life cycle is divided into smaller parts in agile development [23]. Agile development is based on the idea of incremental and iterative development, in which the phases within a development life cycle are revisited over and over again. It iteratively improves software by using user feedback to converge on solutions. Data collected from four (4) house developers which are Metacorp Properties Sdn.Bhd (Taman Tasik Utama, Ayer Keroh), PDG Property Sdn.Bhd (Taman Saujana Indah), Country Villas Resort Sdn. Bhd (Country Villas Resort), and GJH Holdings Sdn.Bhd (Vista Kirana). Generally, for the system development, it is divided into three parts which are use case diagram, HPRS flowchart diagram and the implementation of the GA.

2.1. Use Case Diagram

Figure 1 shows the use-case diagram that involved five activities for the admin and four activities for the homebuyer. Use case description is important to clearly explain with details each use case involved so the developer understands exactly what is in the requirement to satisfy.

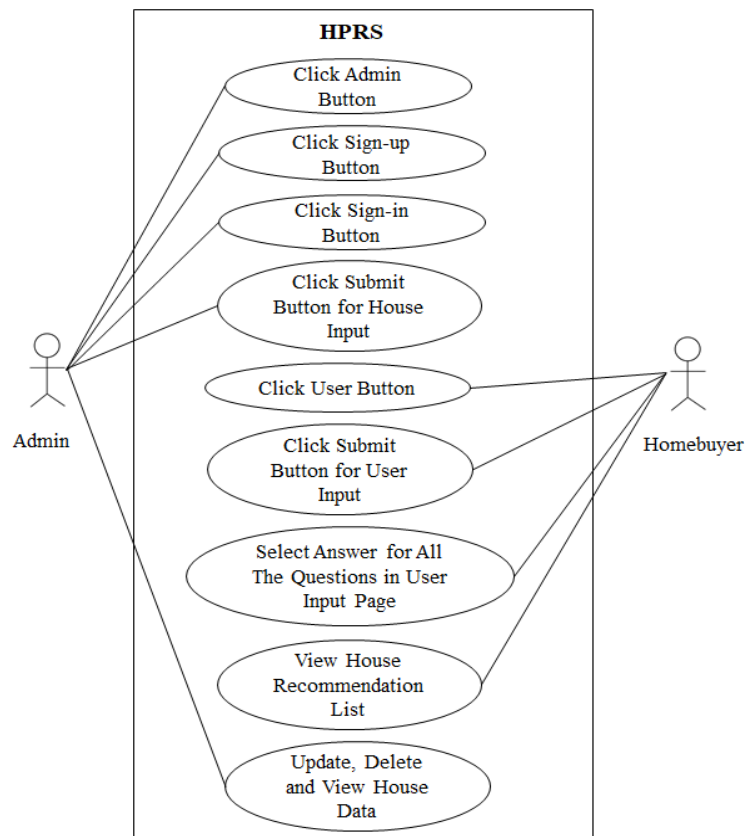


Figure 1. Overall HPRS Use Case Diagram

2.2. System Flowchart Diagram

A system flowchart shows the overall flow of HPRS and is shown in Figure 2. It starts with the input from the homebuyer and follows by the GA process. Finally, the result is the list of recommendation details of the house and the house price within the budget.

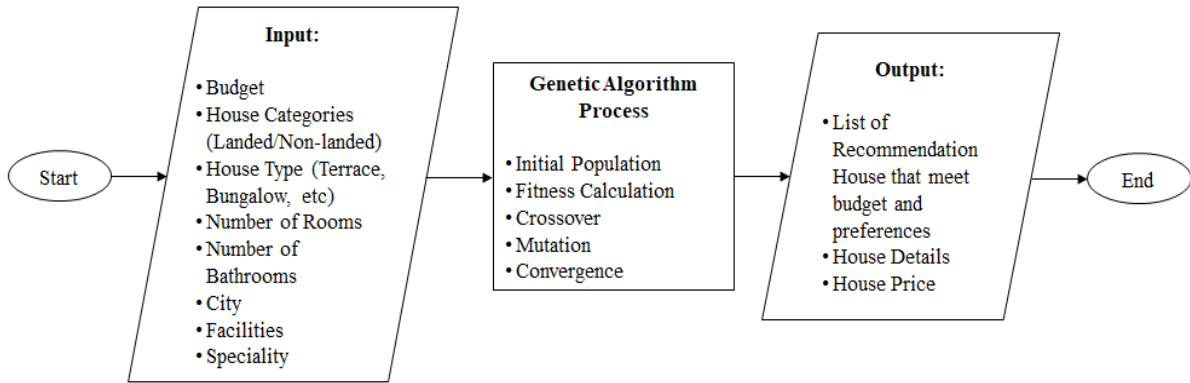


Figure 2. Flowchart Diagram for HPRS

2.3. Implementation of Genetic Algorithm

GA involved 5 processes (as illustrated in Figure 2) start with generate population, fitness calculation, crossover, mutation, and convergence. All the input given by the homebuyer will be saved into the system.

2.3.1. Process 1: Generate Population

Before we start generating the population, it started with chromosome encoding. The chromosome encoding is created randomly. Next, the population is generated, and it will create chromosomes as much as possible as of example shown in Figure 3. Every house contains the house details such as the house price, the category and type of the house, the availability of facilities and specialty, the agent phone number, the number of rooms and bathrooms, the built-up size and the address house.

House33	House23	House35	House46	House4
Budget	Budget	Budget	Budget	Budget
Category of House	Category of House	Category of House	Category of House	Category of House
House Type	House Type	House Type	House Type	House Type
House Price	House Price	House Price	House Price	House Price
House Address	House Address	House Address	House Address	House Address
No of Room	No of Room	No of Room	No of Room	No of Room
No of Bathroom	No of Bathroom	No of Bathroom	No of Bathroom	No of Bathroom
Build up Size	Build up Size	Build up Size	Build up Size	Build up Size
Facilities	Facilities	Facilities	Facilities	Facilities
Speciality	Speciality	Speciality	Speciality	Speciality
Agent Phone Number	Agent Phone Number	Agent Phone Number	Agent Phone Number	Agent Phone Number

Figure 3. Example of Chromosome Encoding for Population Generation

2.3.2. Process 2: Fitness Calculation

At this phase, the fitness value of each chromosome been calculated by comparing the chromosome from the user and from chromosome from the database using the computer coding. The fitness calculation been total and measured by comparing the user input with the available information.

2.3.3. Process 3 and 4: Crossover and Mutation.

The best fitness value has been chosen, and it will be sorted from the highest fitness value until the lowest fitness value. The first 3 chromosomes with the highest fitness value will undergo the crossover and mutation process, then sort it again using the fitness value. The new chromosome created been called as a child. Figure 4 shows the example of the process of crossover happens on the chromosome. Two value from chromosome X will be the crossover with two value from chromosome Y to create a new child. Mutation process with continues with the child.

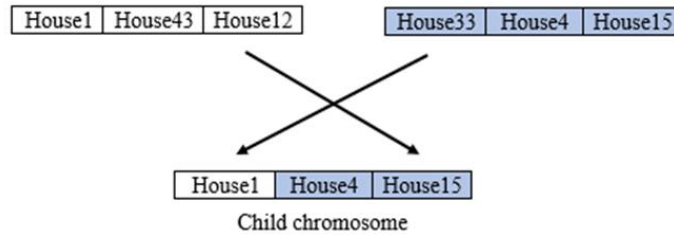


Figure 4. Example of Crossover Process on Chromosome

Therefore, the 1st generation of a chromosome is being produced. Then, the other half of the chromosome with low fitness value will be repopulated. After the chromosome has been repopulated, the steps will be repeated until the 50th generation.

2.3.4. Process 5: Convergence

The convergence process started when the system received the same chromosome with the highest fitness value. Else not, it will loop again until the result is optimized. Therefore, the data from the first 3 chromosomes are considered as a best result, the GA will stop and display it to the homebuyer. The data have been chosen at this stage because the system needs to provide the optimized result.

3. RESULTS AND ANALYSIS

This section discusses the results and findings gained from the HPRS. In this study, there were two types of testing conducted which is functionality testing and usability testing.

3.1. Functionality Testing

The functionality testing is used in order to validate the result that has been obtained from each module of the use case. It checked every module and examine whether it can produce the expected output. Figure 5 shows the snapshot of the HPRS where the homebuyer needs to answer the given questions started with the city area, house preferences, and followed by the details of the specific question. The questions are vary depending on the selection such as if homebuyer select landed house (as in Question 2), the non-landed question (as in Question 2c) will not appear.

Figure 5. Snapshot of HPRS

The functionality test has been conducted followed the HPRS use case and Table 1 shows the HPRS result of the five times testing on the functionality to ensure it functioning as per plan.

Table 1. HPRS Functionality Test Result

Component	Test1	Test2	Test3	Test4	Test5
Click Admin Button	Ok	Ok	Ok	Ok	Ok
Click Sign-up Button	Ok	Ok	Ok	Ok	Ok
Click Sign-in Button	Ok	Ok	Ok	Ok	Ok
Click Submit Button for House Input	Ok	Ok	Ok	Ok	Ok
Click User Button	Ok	Ok	Ok	Ok	Ok
Select Answer for All The Questions in User Input Page	Ok	Ok	Ok	Ok	Ok
View House Recommendation List	Ok	Ok	Ok	Ok	Ok
Update, Delete and View House Data	Ok	Ok	Ok	Ok	Ok

Next, we test the HPRS result by providing one of the examples of user input for specific house preferences with the budget of RM900,000 as in Table 2. As we know that different houses will have different house preferences, thus, the homebuyer needs to provide the input of the house preferences as much as possible.

Table 2. Example of User Input

Budget	House Preferences
RM900,000	City: Ayer Keroh House Categories: Landed House Type: Semi-D House Position: Corner Lot Number of Rooms: 4 Number of Bathrooms: 3 Facilities: Hypermarket, Mall, Hospital Speciality: Guarded, Unfurnished, Car Porch

Using the same input, HPRS has been tested for three times to measure the consistency of the result. Figure 6 until Figure 8 shows the output of HPRS which provides the recommended list of 3 houses that match with the user input as in previous Table 3. The house price shown is the house price, that is lower than the homebuyer budget. Figure 7 shows the lowest house price at Taman Tiara Villa, but the built-up size is only 2,800 sq. compared to Figure 8 that provides the bigger built-up size which is 3,200 sq. Despite Figure 9 shows the HPRS result 3 with different specification 5 number of rooms instead of 4, yet the house price still within the budget RM900,000 which is RM802,018.



Figure 6. HPRS Result 1



Figure 7. HPRS Result 2

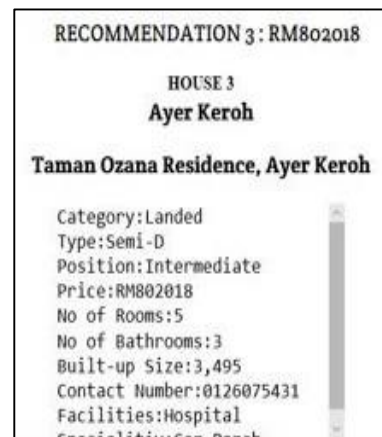


Figure 8. HPRS Result 3

3.2. Usability Testing

Usability testing is a method to test the HPRS by letting the users of the system perform a certain task and how they understand the usage of the system. For this project, the testing was done using the system usability scale. The System Usability Scale (SUS) is a survey consist of ten-item scale giving assessments of usability. The statement of the items is made, and the respondents then indicate the agreement or

disagreement with the statement on a 5-point scale. The statements covered the aspects of usability like the need for support, training, and complexity [24]. Thus, it has a great level of validation for measuring the usability of the system. Respondents of the SUS are asked to immediately record their response to the survey after they had tried using HPRS. All items are required to be filled and they must mark the center point of the scale if they cannot respond to the specific item. Hence, the survey can be done without any biases. Table 3 shows all the questions of the system usability scale questionnaire.

Table 3. System Usability Scale Questionnaire

Scale: 1-Strongly Disagree, 2-Disagree, 3-Neutral, 4-Agree, 5-Strongly Agree						
No	Question	Scale				
1	I think that I would like to use this system frequently	1	2	3	4	5
2	I found the system unnecessarily complex	1	2	3	4	5
3	I thought the system was easy to use	1	2	3	4	5
4	I think that I would need the support of a technical person to be able to use this system	1	2	3	4	5
5	I found the various functions in this system were well integrated	1	2	3	4	5
6	I thought there was too much inconsistency in this system	1	2	3	4	5
7	I would imagine that most people would learn to use this system very quickly	1	2	3	4	5
8	I found the system very cumbersome to use	1	2	3	4	5
9	I felt very confident using the system	1	2	3	4	5
10	I needed to learn a lot of things before I could get going with this system	1	2	3	4	5

Respondents selected are from Melaka Tengah area and varied in term of age and gender. The respondents have been asked to use the HPRS and directly answered the questionnaire given. Figure 9 summarized the SUS questionnaire result in 30 respondents for each of the question given.

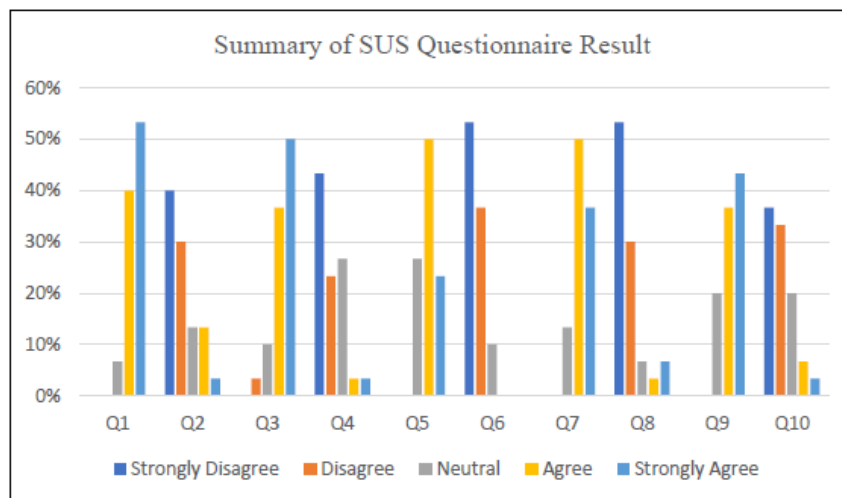


Figure 9. Summary of SUS Questionnaire Result for HPRS

SUS scoring calculated based on contributions on each item that is range from 0 to 4. For odd number items, the contribution is the scale position minus 1 while the even numbers contributed 5 minus the scale position. All of the contributions then multiplied by 2.5 to obtain the overall value of the SU. The result is in the range of 0 to 100. The average score of SUS is 68 means that the HPRS is good in term of usability [25] and Figure 10 illustrates a histogram summarized the total SUS scores obtained for the HPRS. From the histogram, the score is divided into a range of 5 starting from the lowest score obtained which is 55. The frequency of each range is shown in the vertical axis. For example, there are 4 frequencies of score 70 until 75 based on the histogram. The total score obtained from 30 respondents are 2377.5 and it is divided by the number of respondents to find the average SUS score. The average SUS score obtained for the HPRS is 76.8. Therefore, the HPRS is proven to have good usability by achieving an average score of more than 68.

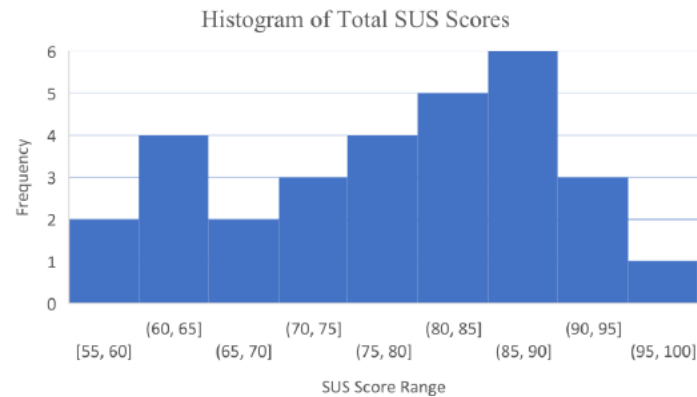


Figure 10. Histogram of Total SUS Scores for HPRS

4. CONCLUSION

As a conclusion, this House Purchase Recommendation System has achieved the objectives and able to solve the highlighted problem. Using the GA technique advantage, HPRS help homebuyer to search and buy a house based on their house preferences and budget. Indirectly, it manages to reduce time-consuming for the houses surveying through websites. The functionality testing is done by evaluating and try each use case function of the website and it is proven HPRS is successfully functioned well. However, there are a few limitations of this system that be improved. For future work, this system should be displaying the monthly installments for the house, ability for the homebuyer to view the picture of the house and state the rate of interest for the bank loan. Also to test the performance of the system by comparing with other techniques such as Naïve Bayes and Support Vector Machine.

REFERENCES

- [1] H. Ramlan and E. E. Zahari, "Review the Issue of Housing Among Urban Dwellers in Malaysia with Special Reference Towards Affordability to Home Ownership," *Procedia Economics and Finance*, vol. 35, pp. 216-223, 2016.
- [2] A. I. Che-ani, et al., "Affordable Housing Performance Indicators for Landed Houses in the Central Region of Malaysia," *Modern Applied Science*, vol. 8, pp. 70-86, 2014.
- [3] A. Osmadi, et al., "Exploring The Elements Of Housing Price in Malaysia," *Asian Social Science*, vol. 11, pp. 26, 2015.
- [4] L. Abdullah, et al., "First-Time Home Buyers: Factors Influencing Decision Making," *International Conference on Innovation and Technology for Sustainable Built Environment*, pp. 249-254, 2012.
- [5] O. L. H. Leh, et al., "The Housing Preference Of Young People In Malaysian Urban Areas: A Case Study Subang Jaya, Selangor," *Geografia-Malaysian Journal of Society and Space*, vol. 12, pp. 60-74, 2017.
- [6] S. A. Gapor, et al., "Trends, Problems And Needs Of Urban Housing In Malaysia," *Malay*, vol. 248, pp. 62, 2010.
- [7] N. Zainon, et al., "Factors Affecting The Demand of Affordable Housing Among The Middle-Income Groups In Klang Valley Malaysia," *Journal of Design and Built Environment*, pp. 1-10, 2017.
- [8] M. H. Mohamad, et al., "Review of Building, Locational, Neighbourhood Qualities Affecting House Prices in Malaysia," *Procedia - Social and Behavioral Sciences*, vol. 234, pp. 452-460, 2016.
- [9] P. A. M. Khan, et al., "Housing Preference for First Time Home Buyer in Malaysia," *International Journal of Real Estate Studies*, vol. 11, pp. 1-6, 2017.
- [10] X. Yuan, et al., "Toward A User-Oriented Recommendation System for Real Estate Websites," *Information Systems*, vol. 38, pp. 231-243, 2013.
- [11] D. B. Tonara, et al., "Recommender System in Property Business a Case Study from Surabaya, Indonesia," *The Third International Congress on Interdisciplinary Research and Development*, pp. 71-75, 2013.
- [12] D. W. A. Aziz, et al., "A Study of Affordable Housing within the Middle Income Households in the Major Cities and Towns of Malaysia," *Australian Journal of Basic and Applied Sciences*, vol. 5, pp. 258-267, 2011.
- [13] J. Chia, et al., "Understanding Factors That Influence House Purchase Intention Among Consumers In Kota Kinabalu: An Application of Buyer Behavior Model Theory," *Journal of Technology Management and Business*, vol. 3, pp. 94-110, 2016.
- [14] A. A. M. Shoed and G. Subramaniam, "Housing Affordability Among Generation Y in Malaysia A Conceptual Analysis," vol. 731, pp. 16-23, 2015.
- [15] A. G. Salleh and N. Badarulzaman, "Quality of life of residents in urban neighbourhoods of Pulau Pinang, Malaysia," *Journal of Construction in Developing Countries*, vol. 17, pp. 117-123, 2012.
- [16] A. Abdullah, et al., "Fear of Crime in Gated and Non-gated Residential Areas," *Procedia-Social and Behavioral Sciences*, vol. 35, pp. 63-69, 2012.

-
- [17] S. R. Ju and S. Omar, "A Typology of Modern Housing in Malaysia," *International Journal of Human Ecology*, vol. 11, pp. 109-119, 2010.
- [18] B. A. Ojokoh, et al., "A Fuzzy Logic Based Personalized Recommender System," *International Journal of Computer Science and Information Technology & Security*, vol. 2, pp. 1008-1015, 2012.
- [19] H. Liu, et al., "Rule-Based Systems: A Granular Computing Perspective," *Granular Computing*, vol. 1, pp. 259-274, 2016.
- [20] C. Grosan and A. Abraham, "Rule-Based Expert Systems," *Intelligent Systems. Intelligent Systems Reference Library*, Springer, Berlin, Heidelberg, vol. 17, pp. 149-150, 2011.
- [21] O. S. Taiwo, et al., "Application of Genetic Algorithm to Solve Traveling Salesman Problem," *International Journal of Advance Research*, vol. 1, pp. 27-46, 2013.
- [22] M. Sadeghi and S. A. Asghari, "Recommender Systems Based on Evolutionary Computing: A Survey," *Journal of Software Engineering and Applications*, vol. 10, pp. 407, 2017.
- [23] R. P. Pawar, "A Comparative Study of Agile Software Development Methodology and Traditional Waterfall Model," *IOSR Journal of Computer Engineering*, pp. 01-08, 2015.
- [24] J. Brooke, "SUS - A Quick and Dirty Usability Scale," 1986. Available: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>.
- [25] N. Thomas, "How to Use the System Usability Scale (SUS) To Evaluate the Usability of Your Website - Usability Geek," 2015. Available: <https://usabilitygeek.com/how-to-use-the-system-usability-scale-sus-to-evaluate-the-usability-of-your-website/>.