Proceeding of International Conference on Multidisciplinary Research for Sustainable Innovation, Vol. 1 No. 1 (2024)
DOI to be assigned soon

Research Paper

Design of A Property Product Recommendation System Using Association Rule Method Based on User Interaction Patterns

Susi Wagiyati Purtiningrum[1], Yudi Irawan Chandra[2]*, Dian Gustina[1], Nafisah Yuliani[1], Fahrul Nurzaman[1], Jhonny Z.A[1], Agus Wismo[1]
[1] Universitas Persada Indonesia Y.A.I, Indonesia
[2] Sekolah Tinggi Manajemen Informatika dan Komputer Jakarta STI&K, Indonesia

Received : January 26, 2024	Revised : February 15, 2024	Accepted : February 17, 2024	Online : March 13, 2024

Abstract

This work creates an association rule-based real estate product recommendation system. Personalizing property suggestions based on user behaviour optimises property searches. Data-driven insights enhance dynamic property market user experience. Association rules alter property advice. Data-driven insights and adaptability improve property search by proposing homes depending on user engagement patterns. Strong algorithms establish location, budget, and property associations, and association rule technology and user interaction patterns increase property recommendations. Personalized property discovery uses accurate and adaptive suggestions from continuous learning. Results reveal that user interaction pattern-based association rule techniques improve property suggestion accuracy and personalization. The system's tailored advice improves property market decisions, confirming its usefulness and adaptability. Insufficient user data might distort suggestions, especially for specific interests. Not enough user diversity can lower system accuracy. User data and privacy must be secured to optimize the recommendation system. Association rules and user engagement patterns can transform property recommendations. This innovative technique can improve property searches, provide personalized ideas, and help consumers make informed decisions in a competitive market.

Keywords *Property Recommendation System, Association Rule Method, User Interaction Patterns, Personalized Suggestions, Real Estate Decision-making*

INTRODUCTION

Some property websites group properties by kind (houses, shop houses, offices, boarding houses, etc.), size, status (for sale or rent), ownership letter status, facilities, contents, provinces, and cities for easy searching and property information. There are two categories of information seekers. The first type knows their needs. Type 2 is unsure about their needs. For the second sort of user, provide help to assess their property needs (Owen, 2019).

This study suggests employing Apriori and Association Rule algorithms to suggest products based on user profiles or product features. The system uses the Apriori Association Rule to find product relationships in a dataset. The system collects user-viewed advertising for analysis. The user's search process weights data attributes by frequency. System usage analysis proposes advertisements.

Fitriah et al. (2023) used the Apriori Algorithm for BM Warehouse Store Sales Data, while Wijaya et al. (2023) proposed a hybrid content-based and collaborative method for the Implementation of Data Mining Purchasing Patterns at Santoso Tiga Sumenep Store. Since it uses user datasets instead of training data or expert input, the Apriori Association Rule is more accurate than Naive Bayes. With more data, the Apriori Association Rule method takes longer.

PT Brighton Property provided data for this study. Processing Apriori Association Rule data

Corresponding author's email: yirawanc@gmail.com

can boost the company's property product promotion and client transaction interest. How do we recommend property products using system user activity? Evaluation of user activity pattern data using Apriori association rules?

Learning from the Brighton Indonesia website, property type, price, area, and address are limited. By trying blackbox testing, apriori association rule-analyzed system utilisation will develop a feature recommendation system and help buyers find their dream home and showcase the company's assets with a simple property search and information system. Research suggests that the apriori association rule technique can increase system usage-based property marketing, which may increase consumer interest in the company's properties.

LITERATURE REVIEW

Recommendation System

A recommendation system is software designed to assist users in making decisions when faced with large amounts of information. The system provides personalized recommendations, such as what items to buy, what books to read, or what music to listen to (Jogiyanto, 2005). Personalized recommender systems must first recognize each existing user.

Recommendation systems must create and maintain a user profile that includes user interests. For instance, Amazon's e-commerce recommendation system stores all customer purchase transactions, comments, and reviews or ratings (Pradana, 2016; Sun et al., 2023). Two approaches can be used to form a user profile: implicit and explicit. The implicit approach involves storing and learning user behaviour to build user profiles. This behaviour can take the form of likes/dislikes, ratings, and other actions related to various items. In contrast, the explicit approach directly asks the user to describe the items they like.

Apriori

Another technique used in data mining is the Apriori algorithm, a type of association rule. This rule explains the association between attributes, also known as affinity analysis or market basket analysis (Laksito & Kusrini, 2019). Apriori finds comparable things based on user ratings. The algorithm develops a user profile from item properties. In a document, the forming attribute is its words. These user profile parameters are weighted based on criteria. These are the algorithm steps:

1. The product is divided based on a vector of its forming components.
2. The system creates a user profile based on the weight of the component vector that forms an item. User profiling can use the TF-IDF (term frequency-inverse document frequency) algorithm. TF represents the number of terms in a document (Dwi, 2016), while the IDF value can be calculated using the formula:

$$idf_i = \log\left(sin\frac{n}{df_i}\right) \quad (1)$$

n is the number of all documents, while df is the number of documents that have the term i.

Based on their profile's resemblance to the item's vector of components, the system will determine a user's likes or dislikes. If the algorithm thinks the user will like it, it will recommend it. Benefits of content-based recommendation systems:

1. Content-based recommendation systems can explain results generation.
2. Content-based recommendation engines may suggest unrated products. Drawbacks of content-based recommendation algorithms include:

3. Content-based recommendation systems cannot produce unexpected results (Serendipity Problem).
4. Content-based recommendation systems require user profiles with interests and preferences. Poorly profiled and inactive new users cannot obtain recommendation system recommendations (Cold Start Problem).

Association Rule Mining

Associative rules between objects are found via association analysis or association rule mining (Toivonen, 2017). Many data mining approaches include association analysis. Frequent pattern mining, a stage of association analysis, helps academics create efficient algorithms. Support, the percentage of database combinations of these items, and confidence, the relationship strength, and value associative rules. Association analysis has two steps:

1. High-frequency pattern analysis
   This stage finds item combinations that match database support value minimums. The formula for item support value:

   $$\text{Support (A)} = \text{Number of Transactions containing A / Total Transactions} \quad (2)$$

   while the support value of 2 items is obtained from the following formula:

   $$\text{Support (A} \cap \text{B)} = \text{Number of Transactions containing A and B / Total Transactions} \quad (3)$$

2. Associative rule formation
   After all the high-frequency patterns are found, the associative rule that fulfils the minimum requirement for confidence is found by calculating the confidence of the associative rule A B. The confidence value of the rule A B is obtained from the following formula:

$$\text{P(B | A)} = \text{Number of Transactions containing A and B / Number of Transactions containing A} \quad (4)$$

Minimum Support

The minimum support value distinguishes frequent items or criteria from infrequent ones. It is determined based on the analysis observations and affects the analysis results (Sinthuja et al., 2019). A higher support value results in fewer selected items or criteria and no attachment relationship. Similarly, a smaller support value allows more items or criteria to pass the selection, resulting in less specific analysis. Therefore, testing and determining the appropriate support value is crucial to produce the most accurate analysis.

Minimum Confidence

The minimum confidence value is used to measure the relationship between items in an association rule. It is a unit of measurement that needs to be determined by the relevant parties. The greater the minimum confidence value, the weaker the relationship between the appropriate criteria, but it leads to maximum analysis results. Similarly, decreasing the minimum confidence value strengthens the correlation between the relevant criteria, but it does not maximize the analysis results.

Blackbox Testing

Black Box testing, as defined by Bonifácio and Moura (2017), verifies that all software functions were executed correctly according to functional requirements. With this test, software engineers can get input conditions that meet all program functional criteria. This test can detect erroneous or missing functions, interface errors, data structure or external database access errors, initialization and termination mistakes, functional validity, system sensitivity to input values, and data restrictions.

RESEARCH METHOD

Development Model

The system development model used in this research is the Waterfall Model (Chandra et al., 2022). The stages of the software development process are shown in Figure 1.

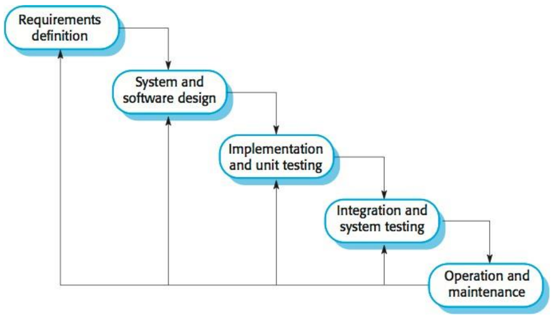

Figure 1. Waterfall Model Activity Stage

1. Definition of Requirements
   PT Brighton's Head of Marketing is interviewed to discover the problem's emphasis, which is analysed using property product data.
2. System and Software Design
   Problems will be solved, and a system will be designed at this level. System design employs a data flow diagram, conceptual data model, physical data model, database table structure, and presentation.
3. Implementation and Unit Testing
   PHP-coded system creation using Codeigneiter Framework. This research manages databases with MYSQL. The manufacturing process includes system unit or sub-program testing.
4. Integration and System Testing
   System testing after completion. Testing is done using Black Box Testing. Trials are meant to ensure the system meets design and research goals.
5. Operations and Maintenance
   This research is system design, so this stage is skipped.

Data Collection Methods

In collecting the data needed in conducting this research are as follows:

1. Interview
   The PT Brighton Property marketing department was interviewed. The purpose of this interview was to determine customer property needs. Interviews included requests for research data.
2. Data and Documentation Study
   Collecting data needed for research activities obtained from product marketing staff, such as a list of property data sold and detailed property data.

Current System Analysis

In the current property marketing system, there is no product recommendation system based on system usage patterns. In general, the current property product marketing system can only search for property products based on property categories and property areas.

Problem Analysis

System visitors do not get property product recommendations that match usage patterns. System visitors must independently search for property products based on the search parameters, such as property categories and property area locations.

Problem Solving Method Analysis

The Apriori method is used to provide property recommendations that match user behaviour. The following is the flow of the apriori method work process as shown in Figure 2.

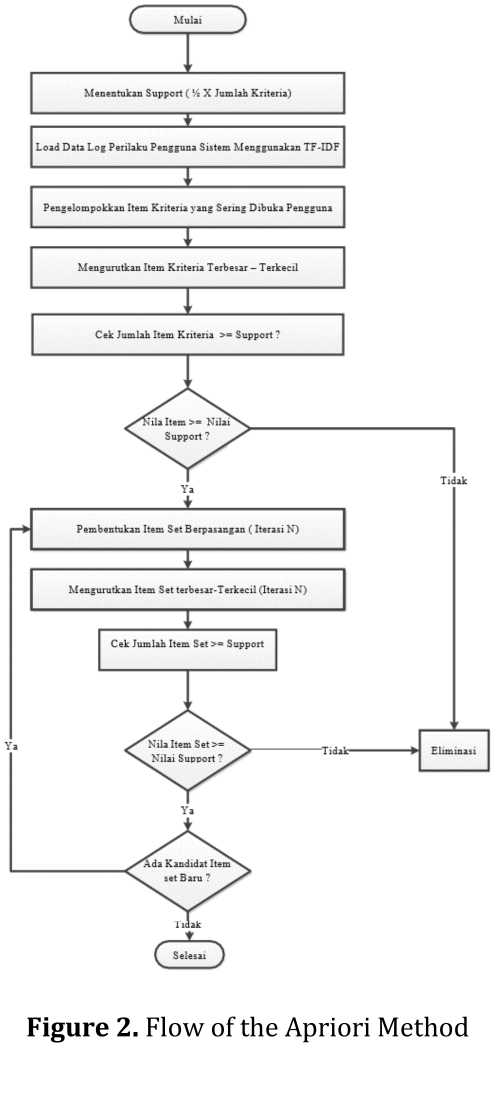

Figure 2. Flow of the Apriori Method

FINDINGS AND DISCUSSION

Application of Association Rule Method with Support Value 1

At this stage, the application of the association rule method with a support value of 1 is tested; the example data used can be seen in Table 1.

Table 1. Sample Ad Details Opened

User ID	Advertising ID	Criteria Name
26	1	Building Area > 500 m2, bathroom 2, Bedroom 3, PDAM, Electricity 2.200 Watt
	2	Building Size < 300 m2, bathroom 1, Bedroom 2, PDAM, Electricity 2.200 Watt
	3	Building Area > 500 m2, bathroom 2, Bedroom 5, PDAM, Electricity 3.500 Watt
	23	Building Size > 300 m2 < 500 m2, room bathroom 2, bedroom 3, PDAM, Electricity 2.200 Watt
	24	Building Area > 500 m2, bathroom 2, Bedroom 5, Well Pump, Electricity 3,500 Watt
	25	Building Area > 500 m2, bathroom 3, Bedroom 6, PDAM, Electricity 4.400 Watt

Based on the ad search pattern data above, the frequency value for the first stage iteration (C1) is shown in Table 2.

Table 2. Frequency Value Iteration C1

User ID	Number Criteria	Criteria	Total Frequency
26	1	Building area < 300 m2	1
	2	Building area > 300 m2 < 500 m2	1
	3	Building area > 500 m2	4
	4	Bathroom 2	4
	5	Bathroom 3	1
	6	Bedroom 2	1
	7	Bedroom 3	2
	8	Bedroom 5	2
	9	Bedroom 6	1
	10	PDAM	5
	11	Well Pump	1
	12	Electricity 2200 Watt	3
	13	Electricity 3500 Watt	2
	14	Electricity 4400 Watt	1

With a support value of 2, the final result of iteration stage 1 must have a minimum frequency value of 2. Here are the results of iteration stage 1 as in Table 3 below.

Table 3. Results of Iteration 1 (L1)

User ID	Number Criteria	Criteria	Total Frequency
26	3	Building area > 500 m2	4
	4	Bathroom 2	4
	7	Bedroom 3	2
	8	Bedroom 5	2
	10	PDAM	5

User ID	Number Criteria	Criteria	Total Frequency
	12	Electricity 2200 Watt	3
	13	Electricity 3500 Watt	2

The next step is to form a new item set by crossing the product of each item set that passes stage 1. The following results of the formation of the item set of iteration stage 2 (C2) are shown in Table 4.

Table 4. Item Set Formation Iteration C2

User ID	Criteria Number
26	3, 4
	3, 7
	3, 8
	3, 10
	3, 12
	3, 13
	4, 7
	4, 8
	4, 10
	4, 12
	4, 13
	7, 8
	7, 10
	7, 12
	7, 13
	8, 10
	8, 12
	8, 13
	10, 12
	10, 13
	12, 13

Based on the results of forming new item sets in the second iteration stage, the system calculates the frequency value again for each set. The following is the result of the frequency value of iteration stage 2 as in Table 5 below.

Table 5. C2 Iteration Frequency Value

User ID	Criteria Number	Frequency
26	3, 4	3
	3, 7	1
	3, 8	2
	3, 10	5
	3, 12	1
	3, 13	2
	4, 7	2

User ID	Criteria Number	Frequency
	4, 8	2
	4, 10	4
	4, 12	2
	4, 13	2
	7, 8	0
	7, 10	2
	7, 12	2
	7, 13	0
	8, 10	1
	8, 12	0
	8, 13	2
	10, 12	3
	10, 13	1
	12, 13	0

With a support value of 2, the final result of iteration stage 2 must have a minimum frequency value of 2. The following results of iteration stage 2 are shown in Table 6:

Table 6. Results of Iteration 2 (L2)

User ID	Criteria Number	Frequency
26	3, 4	3
	3, 8	2
	3, 10	5
	3, 13	2
	4, 7	2
	4, 8	2
	4, 10	4
	4, 12	2
	4, 13	2
	7, 10	2
	7, 12	2
	8, 13	2
	10, 12	3

The next stage is to form the 3rd iteration item set (C3). Formation of item sets in stage 3 by looking at the frequency value of each item set, as in table 7 below:

Table 7. C3 Iteration Formation

User ID	Criteria Number
26	3, 4, 7
	3, 4, 8
	3, 4, 10
	3, 4, 12

User ID	Criteria Number
	3, 4, 13
	4, 7, 8
	4, 7, 10
	4, 7, 12
	4, 7, 13
	7, 8, 10
	7, 8, 12
	7, 8, 13
	8, 10, 12
	8, 10, 13
	10, 12, 13

The frequency values of the item pairs in the third iteration are shown in Table 8:

Table 8. Item Pair Value Iteration 3 (L3)

User ID	Criteria Number	Frequency
26	3, 4, 7	1
	3, 4, 8	2
	3, 4, 10	2
	3, 4, 12	1
	3, 4, 13	2
	4, 7, 8	0
	4, 7,10	1
	4, 7, 12	2
	4, 7, 13	0
	7, 8, 10	0
	7, 8, 12	0
	7, 8, 13	0
	8, 10, 12	0
	8, 10, 13	1
	10, 12, 13	0

With a support value of 2, the final result of iteration stage 3 must have a minimum frequency value of 2. The following results of iteration stage 3 are shown in Table 9 below.

Table 9. Results of Iteration 3 (L3)

User ID	Criteria Number	Frequency
26	3, 4, 8	2
	3, 4, 10	2
	3, 4, 13	2
	4, 7, 12	2

The next step is to form the 4th iteration item set (C4) and count the number of item pair frequencies that have been successfully formed. The following results of the formation of the 4th iteration and the frequency value are shown in Table 10.

Table 10. Iteration C4

User ID	Criteria Number	Frequency
26	3, 4, 8, 10	1
	3, 4, 8, 12	0
	3, 4, 8, 13	2
	4, 8, 10, 12	0
	4, 8, 10, 13	1
	8, 10, 12, 13	0

With a support value of 2, the final result of iteration stage 4 must have a minimum frequency value of 2. The following results of iteration stage 4 that have frequency values above or equal to the support value are shown in Table 11 below.

Table 11. Iteration 3 (L3) result

User ID	Criteria Number	Frequency
26	3, 4, 8, 13	2

Iteration 4 produced 1 connected criteria relationship. Therefore, advertising guidelines are Building Area > 500 m2, 2 baths, 5 bedrooms, 3,500 Watt Electricity. The association rule technique was evaluated with support values of 2, 3, and 4, yielding the following results:

1. No pair/criteria set item in stage 3 meets the support value, preventing the system from making the correct recommendation.
2. No frequency pair passes in the next iteration. The system cannot make the proper recommendation in stage 3 since no pair/item set of criteria passes the support value.

Data Flow Diagram (DFD)

Data flow diagrams show how operations and data flow in the proposed system. The data flow diagram has these subdiagrams:

1. Level 0 Context Diagram

   Data and system user flow are shown in context diagrams. See Figure 3 for the system context diagram design:

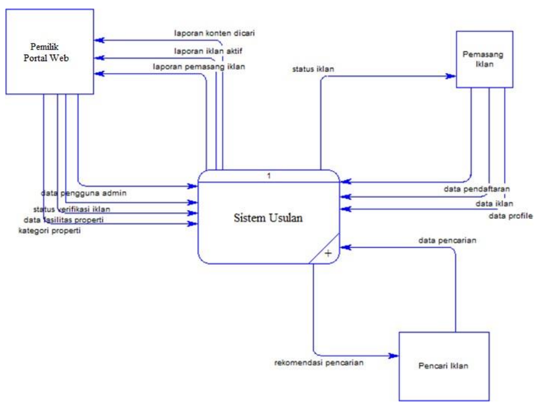

Figure 3. Context Diagram (Level 0)

2. Data Flow Diagram Level 1

   Based on the context diagram, a derivative level of the process is made to be able to see the process flow in the system in detail. The following level 1 diagram design is shown in Figure 4.

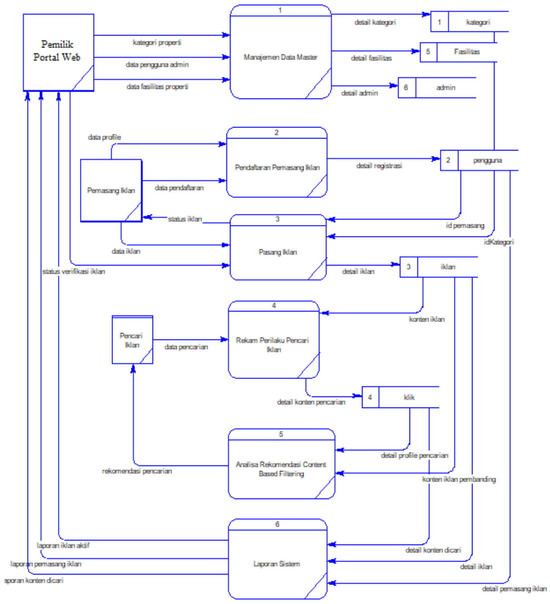

Figure 4. Data Flow Diagram Level 1

System Database Design

Designing system databases is data modelling. System builders learn data flow and database table column contents from data modelling. The Physical Data Model database system design is shown in Figure 5.

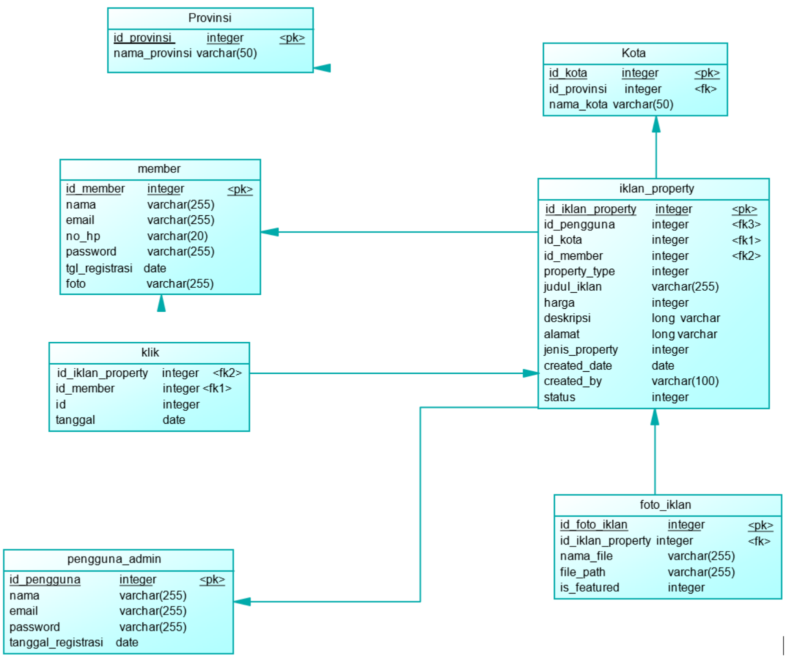

Figure 5. Physical Data Model Database System

Hardware and Software Requirements

In implementing the use of the system created, hardware and software are used, among others.

1. Hardware

   Hardware generates data, programs, and outputs. A server supporting this application has these hardware specs:

   a. Processor: Intel Core i7
   b. Memory: 16.0GB Dual-Channel
   c. HardDisk: 2794GB Seagate and 931GB Seagate
   d. Monitor: 1920x1080 pixels
   e. Keyboard 108 keys
   f. Mouse: Optic PS/2

2. Software

   In developing this application, the author uses software with the following specifications:

   a. Operating System: Windows 7 Ultimate
   b. Programme Package:
      - XAMPP (Apache, MySQL)
      - Bootstrap 5.0
      - Code Igniter 4.0
      - Notepad++
      - Google Chrome, Mozilla Firefox, Mozilla Firefox Developer.

System Implementation Results

The following are some of the implementation results in the form of page views on the web, as shown in Figure 6 below.

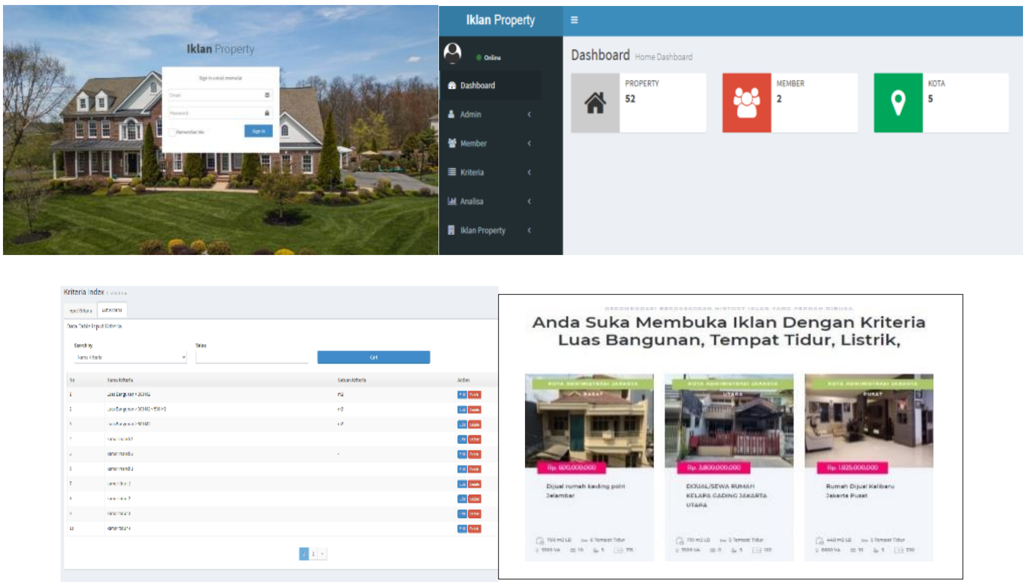

Figure 6. App Display Results in Internet Browser

The system is tested to verify appropriate operation. This research uses black box testing. Blackbox testing involves creating test scenarios. Table 12 shows the testing results.

Table 12. Application Trial Results

No.	Test Result
1	The admin user successfully added data, changed data, and deleted system admin user data.
2	Admin user successfully added data changes and deleted member data.
3	Admin user successfully added data changes and deleted criteria data.
4	Admin user successfully added data changes and deleted sub-criteria data.
5	Admin users have successfully managed ad data, such as adding ad data, changing ad data, deleting ad data, approving pending ads and setting ad images.
6	The admin user successfully runs and gets the results of a priori association rule analysis
7	General users successfully register as members
8	Member successfully manages ad data

CONCLUSIONS

Finally, the Property Product Recommendation System improves property suggestions using user interaction pattern-based association rules. Personal recommendations increase real estate decision-making and user experience. Adaptability and learning keep the system relevant. The study reveals that data-driven approaches increase user-centric property discovery. As the real estate market evolves, association rule techniques should be integrated into recommendation systems. This study optimises property recommendation systems with user-tailored suggestions. The Apriori Association Rule method recommends property advertising based on user behaviour. The optimal Apriori Association Rule support value is 50% of the maximum item or criteria frequency, which can affect analysis results. With a minimum confidence value of 50%, the company may propose good results. The system was 'Very Good' and user-acceptable, with a 90 acceptance test result.

Advanced machine learning and real-time user feedback could boost algorithm efficiency. Market developments and economic factors may improve the system's prediction. Explore different suggestion strategies based on user profiles and demographics to better understand individual preferences. The system's scalability and adaptability to real estate markets and cultures must be assessed. The recommendation system must be constantly improved and adapted to provide useful property suggestions.

REFERENCES

Bonifácio, A. L. & Moura, A. V. (2017). Test suite completeness and black box testing. *Software: Testing, Verification and Reliability, 27*(1-2), e1626. https://doi.org/10.1002/stvr.1626.

Chandra, Y.I., Lusita, M.D., & Ekasari, M.H. (2022). Rancang Bangun Aplikasi Informasi Puskemas Berbasis Web Mobile (Studi Kasus : Puskesmas Tanah Abang). *Jurnal Ilmiah Teknik Informatika, 23*(2), 106–115. https://doi.org/10.37817/tekinfo.v23i2.2602.

Fitriah, Riadi, I., & Herman. (2023). Analisis Data Mining Sistem Inventory Menggunakan Algoritma Apriori: Analysis Data Mining of Inventory System Using Apriori Algorithm. *Decode: Jurnal Pendidikan Teknologi Informasi, 3*(1), 118–129. https://doi.org/10.51454/decode.v3i1.132Frequent Itemset Mining Using LP-Growth

Jogiyanto, H. M. (2005). *Analisis dan desain sistem informasi*. Andi Offset.

Laksito, A.D. & Kusrini. (2019). Apriori Algorithm Optimization using Temporary Table. *2019 International Conference on Information and Communications Technology (ICOIACT)*, 560–565. https://doi.org/10.1109/ICOIACT46704.2019.8938475

Owen, J. (2019). Appraising the Property: Making the best of what you've got. In *Home Extension Design*. RIBA Publishing.

Pradana, M. (2016). Klasifikasi Bisnis E-Commerce di Indonesia. *Modus 27,* 163. https://doi.org/10.24002/modus.v27i2.554

Sinthuja, M., Puviarasan, M., & Aruna, P. (2019). Frequent Itemset Mining Using LP-Growth Algorithm Based on Multiple Minimum Support Threshold Value (Multiple Item Support Frequent Pattern Growth). *Journal of Computational and Theoretical Nanoscience, 16*(4), 1365-1372. https://doi.org/10.1166/jctn.2019.8046.

Sun, G., Xuefeng, D. & Chen, C.-C. (2023). Call for Metaverse e-commerce: future challenges and opportunities for electronic commerce. *Electronic Commerce Research*. https://doi.org/10.1007/s10660-023-09704-6

Toivonen, H. (2017). Association Rule, in: Sammut, C., Webb, G.I. (Eds.), *Encyclopedia of Machine Learning and Data Mining* (pp. 70-71). Springer US. https://doi.org/10.1007/978-1-4899-7687-1_38

Wijaya, M. I. P., Aqromi, N. L., & Afiyah, S. N. (2023). Implementasi Data Mining Pola Pembelian Pada Toko Santoso Tiga Sumenep Dengan Menerapkan Algoritma Apriori. *Jurnal Ilmiah Teknologi Informasi Asia, 17*(2), 97-108.