

# Sistem Informasi Prediksi Tamu Homestay Mengulangi Kedatangannya

(Study Kasus Homestay di Tawangmangu Kab Karanganyar)

Maria Atik Sunarti Ekowati, Elvies Umbulolo,

Universitas Kristen Surakarta, Surakarta

Email korespondensi : maria.atik@gmail.com

## ABSTRAK

Homestay adalah salah satu usaha yang bergerak dibidang jasa sewa, baik kamar atau rumah untuk penginapan. Tawangmangu, kabupaten karanganyar adalah salah satu destinasi wisata di Eks Karisidenan Surakarta, yang cukup terkenal, oleh sebab itu banyak masyarakat di sekitar tempat wisata tersebut menyediakan penginapan bagi wisatawan dan salah satunya adalah Homestay. Hal tersebut tentu menimbulkan persaingan usaha antar Homestay dan penginapanlain. Data mining dimaksudkan untuk memberikan solusi bagi para pengambil keputusan di dunia bisnis untuk mengembangkan bisnis mereka. Untuk itu penulis tertarik mengangkat permasalahan ini kedalam penelitian dengan judul “Perancangan Sistem Informasi Aplikasi Data Mining Untuk Memprediksi Kemungkinan Tamu Homestay Mengulangi Kedatangannya “, Tujuan dari penelitian ini adalah untuk melihat dimana homestay yang paling banyak diminati konsumen untuk disewa, sehingga pemilik homestaypun dapat memprediksi kemungkinan tamu mereka mengulangi kedatangannya pada homestay tersebut untuk menginap. Manfaatnya mempermudah analisis data yang besar dan membantu memberikan informasi pengolahan data tamu homestay yang mengulangi datang menginap. Dan salah satu metode yang terdapat dalam data mining yang digunakan dalam penelitian ini adalah metode Clustering (Pengelompokkan). Hasil dari penelitian ini adalah sebuah rancangan aplikasi yang pada tahap penelitian selanjutnya dibangun guna membantu usaha bisnis homestay ditawangmangu sebagai gambaran dalam pengambilan keputusan peningkatan infrastruktur, industrialisasi, dan inovasi yang berkelanjutan di homestay guna mendapatkan pola kedatangan tamu yang mengulang menginap.

Kata kunci: Data mining, Prediksi, Aplikasi, Pohon keputusan, Homestay

## ABSTRACT

*Homestay is one of the businesses engaged in rental services, either rooms or houses for lodging. Tawangmangu, Karanganyar district is one of the tourist destinations in the Ex Karisidenan Surakarta, which is quite well known, therefore many people in the tourist area provide accommodation for tourists and one of them is Homestay. This certainly creates competition between homestays and other lodgings. Data mining is intended to provide solutions for decision makers in the business world to grow their business. For this reason, the authors are interested in raising this problem into the study with the title "Designing a Data Mining Application Information System to Predict the Possibility of Homestay Guests to Repeat Their Arrival". The purpose of this study was to see which homestays are most desirable for consumers to rent, so that homestay owners can predict the possibility their guests repeated their arrival at the homestay to stay overnight. The benefits make it easy to analyze large data and help provide data processing information for repeat homestay guests coming to stay. And one of the methods contained in the data mining used in this study is the Clustering method. The results of this study are an application design which is then developed in the research phase to help the homestay business be seen as an illustration of decision making in improving infrastructure, industrialization, and continuous innovation in the homestay in order to get the pattern of arrival of guests who repeat the stay.*

*Keywords: Data mining, Prediction, Application, Decision Tree, Homestay*

## 1. PENDAHULUAN

Era perkembangan teknologi informasi yang begitu maju saat ini, menyebabkan tingkat akurasi suatu data sangat dibutuhkan dalam kehidupan sehari-hari. Setiap informasi yang ada menjadi suatu hal penting untuk menentukan setiap keputusan dalam situasi tertentu. Hal ini menyebabkan penyediaan informasi menjadi sarana untuk dianalisa dan diringkas menjadi suatu pengetahuan dari data yang bermanfaat ketika pengambilan suatu keputusan dilakukan [1].

Memahami sebuah pengetahuan dari data pada suatu informasi saja tidak cukup untuk mengambil suatu keputusan. Diperlukan juga suatu analisa dari setiap data yang ada untuk mendapatkan bahan

pertimbangan dari informasi yang tersedia. Dengan menggunakan data mining, setiap kumpulan atau gudang data dapat memberikan pengetahuan penting yang menjadi informasi yang sangat berharga bagi suatu organisasi, seperti pada organisasi pendidikan. Pada organisasi pariwisata, suatu sistem informasi dapat digunakan untuk memperoleh informasi yang menunjang setiap kegiatan pada pengambilan suatu keputusan. Data mining dapat digunakan untuk menyusun sistem yang mempunyai kemampuan memprediksi pola kedatangan tamu menginap di homestay datang kembali, untuk selanjutnya bisa menjadi strategi dalam proses prediksi kedatangan tamu homestay kembali datang menginap [3].

Hotel, Penginapan dan homestay merupakan salah satu organisasi yang bergerak di bidang pariwisata. Informasi kedatangan tamu yang berwisata ke tawangmangu dan menginap sangat penting guna meningkatkan pelayanan yang dapat membuat tamu nyaman dan bisa menikmati pelayanan homestay dan wisata di tawangmangu kabupaten karanganyar.

Data training yang akan digunakan oleh peneliti adalah data tamu yang datang dan menginap di salah satu homestay yang berada di tawangmangu kabupaten karanganyar. Dari kumpulan data mining dan data testing, dapat diketahui informasi kelulusan yang dapat mempengaruhi beberapa keputusan program studi menggunakan data mining [2].

Adapun masalah penelitian dirumuskan sebagai berikut :

1. Bagaimana model rancangan sistem informasi aplikasi data mining guna memprediksi kedatangan tamu kembali menginap di homestay yang sama.
2. Bagaimana hasil rancangan sistem informasi aplikasi data mining tersebut.

Adapun batasan masalah dalam penelitian ini adalah :

1. Data yang akan dianalisa merupakan data tamu datang menginap di homestay kawasan wisata tawangmangu.
2. Algoritma yang digunakan dalam metode clustering yang dilakukan adalah Agglomerative Hierarchical Clustering Algorithm (AHC).
3. Hasil penelitian adalah sebuah rancangan aplikasi yang pada tahap penelitian selanjutnya dibangun guna membantu usaha bisnis homestay sebagai gambaran dalam pengambilan keputusan peningkatan infrastruktur, industrialisasi, dan inovasi yang berkelanjutan guna mendapatkan pola guna memprediksi kedatangan tamu yang mengulang menginap di homestay wilayah kawasan wisata tawangmangu karanganyar.

Adapun studi sebelumnya dari penelitian ini adalah sebagai berikut :

1. Sandy Kurniawan (2007:16), "Penerapan Data Mining dengan metode interpolasi untuk memprediksi minat konsumen asuransi".
2. Enur Irdiansyah (2010:2), "Penerapan Data Mining Pada Penjualan Produk Minuman di PT. Pepsi Cola Indobeverage menggunakan metode clustering".

Tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk menerapkan Data Mining pada prediksi kedatangan tamu menginap di homestay menggunakan metode clustering.
2. Untuk melihat homestay yang paling banyak diminati konsumen.
3. Hasil akhir dari penerapan data mining ini adalah adanya sebuah grafik yang menunjukkan tingkat minat konsumen homestay yang tinggi untuk datang menginap berulang.

## 2. METODE PENELITIAN

1. Metode Pengumpulan Data : Metode pengumpulan data yang digunakan untuk mendapatkan data dari suatu informasi, maka metode yang digunakan dalam proses pengumpulan data sebagai berikut : Metode Observasi, Metode Studi Pustaka
2. Alat dan Bahan Penelitian  
Alat yaitu : Processor Intel core i3, RAM 2 GB, Hardisk 320 GB, Printer  
Bahan Yaitu : Windows 7, Microsoft Word 2007
3. Metode Pengembangan Perangkat Lunak :  
Menurut [15] Teknik analisis data dalam pembuatan perangkat lunak menggunakan pemodelan perangkat lunak dengan paradigma waterfall, yang meliputi beberapa proses diantaranya : System / Information Engineering, Analisis, Design, Coding, Pengujian, Maintenance
4. Tahapan Data Mining  
Menurut [10] Data mining adalah proses yang menggunakan teknik statistik, perhitungan, kecerdasan buatan dan machine learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar. Data mining bukanlah suatu bidang yang sama sekali baru. Dalam aplikasinya, data mining sebenarnya merupakan bagian dari proses Knowledge Discovery in Database (KDD), bukan sebagai teknologi yang utuh dan berdiri sendiri. Data mining

merupakan salah satu bagian langkah yang penting dalam proses KDD terutama berkaitan dengan ekstraksi dan penghitungan pola-pola dari data yang ditelaah, adapun tahapan data mining sebagai berikut :

- a. Data cleaning  
Untuk menghilangkan data noise (data yang tidak relevan/ berhubungan langsung dengan tujuan akhir proses datamining, misal: data mining yang bertujuan untuk menganalisa hasil prediksi, maka data-data dalam kumpulan seperti "nama pegawai", "umur", dan sebagainya dapat di-ignore) dan tidak konsisten.
  - b. Data integration  
Untuk menggabungkan multiple data source.
  - c. Data selection  
Untuk mengambil sebuah data yang sesuai untuk keperluan analisa.
  - d. Data transformation  
Untuk mentransformasikan data ke dalam bentuk yang lebih sesuai untuk di mining. Data mining Proses terpenting dimana metode tertentu diterapkan untuk menghasilkan data pattern.
  - e. Pattern evaluation  
Untuk mengidentifikasi apakah benar interenting patterns yang didapatkan sudah cukup mewakili knowledge berdasarkan perhitungan tertentu.
  - f. Knowledge presentation  
Untuk mempresentasikan knowledge yang sudah didapat dari user.
5. Metode Penerapan Data Mining : Metode yang digunakan untuk penerapan data mining adalah metode clustering (pengelompokan).
- a. Teknik Clustering  
Secara garis besar metode clustering dibagi dalam 2 tipe yaitu : hierarchical dan non hierarchical. Hierarchical menggunakan  $N \times N$  similarity matrix, sedangkan non hierarchical membagi dataset menjadi sebuah level single partisi, dengan atau tanpa pencocokan antara clusters. Selain itu hal mendasar yang membedakan kedua metode ini adalah : metode pengelompokan hirarki digunakan apabila belum ada informasi jumlah kelompok, sedangkan metode pengelompokan non hirarki bertujuan mengelompokan  $n$  objek ke dalam  $k$  kelompok ( $k < n$ ). [9].
  - b. Agglomerative Hierarchical Clustering Algorithm (AHC algorithm)  
Proses Clustering yang akan dilakukan menggunakan Agglomerative Hierarchical Clustering Algorithm (AHC algorithm). Walaupun lebih lambat, hasil akhir pengelompokan agglomerative hierarchical clustering lebih baik dibandingkan kmeans yang sering digunakan, karena lebih efisien . Proses dasar dari hierarchical clustering adalah:
    1. Mulailah dengan mendefinisikan setiap item sebagai cluster sebagai data yang ada adalah sejumlah  $N$  items, maka didapat sekarang cluster yang ada sejumlah  $N$  items tersebut. Jarak antara clusters dan items sama.
    2. Selanjutnya, mencari pasangan cluster yang paling dekat dan menggabungkan keduanya dalam satu cluster.
    3. Hitung jarak antara cluster yang baru dibentuk dengan cluster yang sudah ada sebelumnya. Ulangi tahap b dan c sampai cluster terbentuk  $N$  cluster yang diinginkan. Sumber : Clemenccon (2014), Pengantar Data Mining.
6. Analisis Basis Data  
Struktur logika dari basis data dapat digambarkan dalam sebuah grafik dengan menggunakan Entity Relationship Diagram (ERD). ERD merupakan hubungan antara entitas yang digunakan dalam sistem untuk menggambarkan hubungan antara entitas atau struktur data dan relasi antar file. Komponen utama pembentukan ERD adalah Entity (entitas) dan Relation (relasi) sehingga dalam hal ini ERD merupakan komponen-komponen himpunan entitas dan himpunan relasi yang dideskripsikan lebih jauh melalui sejumlah atribut-atribut (property) yang menggambarkan seluruh fakta dari sistem yang ditinjau, (Kamber, 2012).
7. Diagram Konteks  
Diagram konteks atau disebut juga dengan model sistem fundamental merepresentasikan seluruh elemen sistem sebagai sebuah bubble tunggal dengan data input dan output yang ditunjukkan oleh anak panah yang masuk dan keluar secara bertahap, (wang, 2014). Berdasarkan hasil analisis, maka diagram konteks dari simulasi aplikasi penerapan data mining penjualan dapat dilihat pada gambar 1 dibawah ini :



Gambar 1. Konteks diagram Analisa dan perancangan Aplikasi Data Mining

### 8. Analisis Sistem

Analisis Sistem dapat didefinisikan sebagai penguraian dari suatu sistem informasi yang utuh ke dalam bagian-bagian komponennya dengan maksud untuk mengidentifikasi dan mengevaluasi permasalahan, kesempatan serta hambatan-hambatan yang terjadi dan kebutuhan-kebutuhan yang diharapkan sehingga dapat diusulkan perbaikan [13]. Atau secara lebih mudah analisis sistem adalah penelitian atas sistem yang telah ada dengan tujuan untuk merancang sistem yang baru atau diperbarui. Tahapan analisis sistem ini merupakan tahapan yang sangat kritis dan sangat penting, karena kesalahan dalam tahap ini akan mengakibatkan juga kesalahan di tahap selanjutnya. Tugas utama analisis sistem dalam tahap ini adalah menemukan kelemahan-kelemahan dari sistem yang berjalan sehingga dapat diusulkan perbaikannya. Data tamu homestay yang sudah ada akan diolah atau dianalisis untuk mengetahui tingkat kecenderungan konsumen di homestay pada faktor ketertarikannya menginap kembali di homestay yang sama. Dari pengolahan data tersebut akan diperoleh suatu pola konsumsi masyarakat terhadap homestay tersebut.

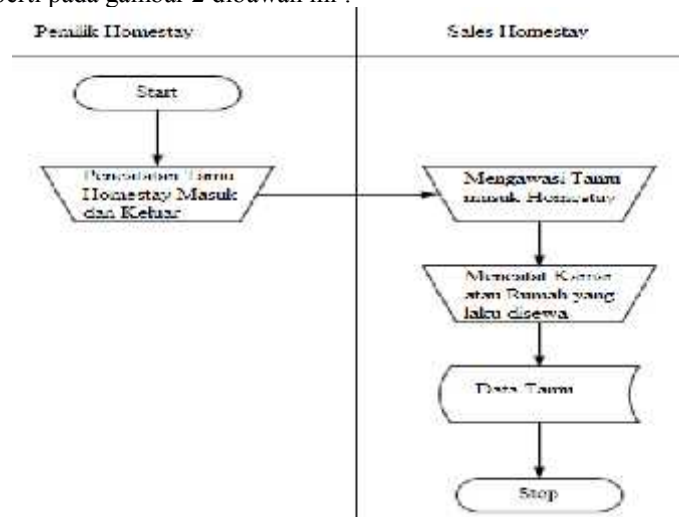
### 9. Analisis Sistem yang akan diusulkan

Berdasarkan permasalahan yang diatas, penulis tertarik untuk membuat data mining prediksi kedatangan tamu menginap kembali di homestay yang sama dengan menggunakan sebuah perancangan sistem informasi aplikasi untuk diselesaikan dengan simulasi pemrograman PHP pada yahap penelitian selanjutnya, [14].

## 3. HASIL DAN PEMBAHASAN

### 3.1. Analisis Sistem yang Sedang Berjalan

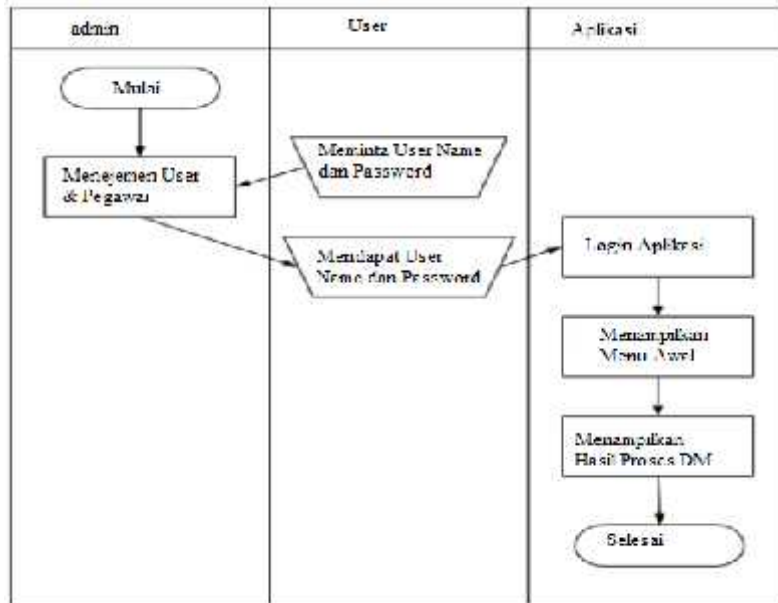
Analisis sistem yang sedang berjalan didefinisikan sebagai penguraian dari suatu sistem untuk mengidentifikasi dan mengevaluasi permasalahan yang terjadi, (Komar, 2014). Adapun sistem yang telah berjalan sekarang seperti pada gambar 2 dibawah ini :



Gambar2. Flow Map Sistem yang sedang berjalan di homestay

### 3.2. Analisis Sistem yang Diusulkan

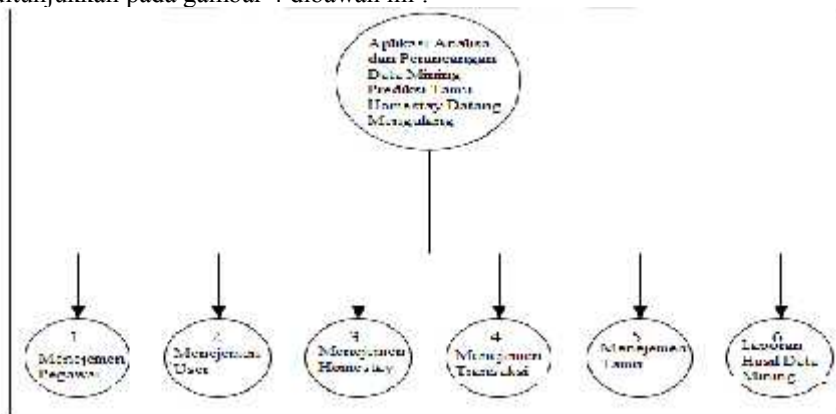
Analisis sistem yang diusulkan sekarang seperti pada gambar 3 dibawah ini :



Gambar 3. Flow Map Sistem yang diusulkan pada pemilik homestay

**3.3. Perancangan Diagram Berjenjang**

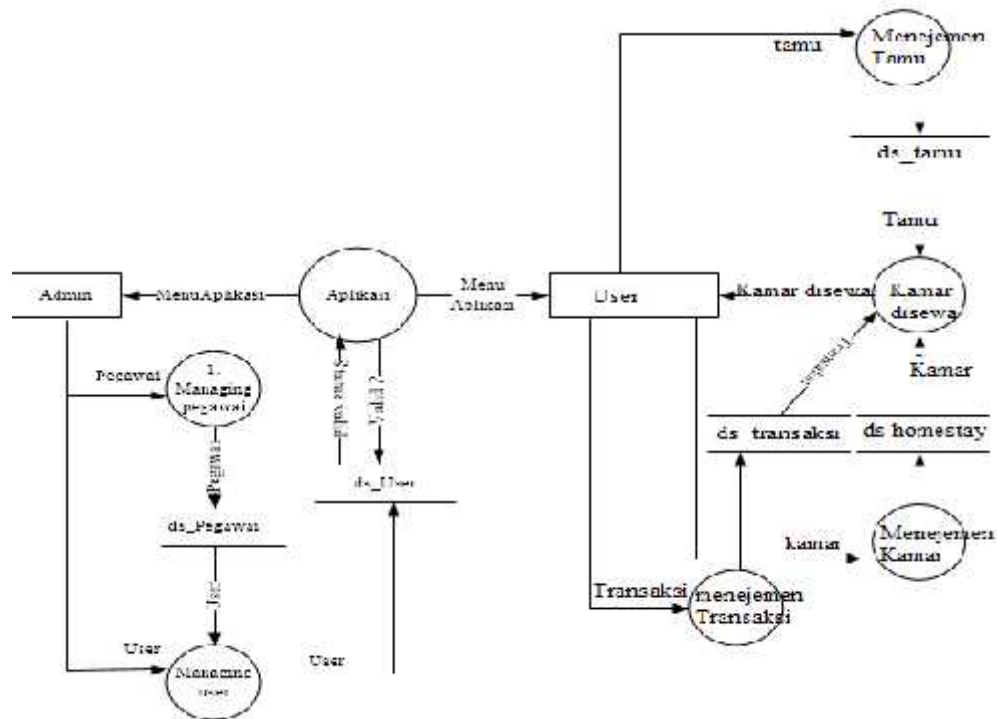
Menggambarkan struktur dari sistem. Berupa suatu bagan berjenjang yang menggambarkan semua proses yang ada pada sistem [4]. Dipergunakan untuk mempersiapkan penggambaran DFD ke level lebih bawah lagi, yang ditunjukkan pada gambar 4 dibawah ini :



Gambar 4 Diagram Berjenjang

**3.4. Perancangan Data Flow Diagram Level 1 (DFD)**

Data flow diagram adalah suatu diagram yang menggunakan notasi-notasi yang menggambarkan arus data dari sistem, yang penggunaannya sangat membantu untuk memahami sistem secara logika, terstruktur dan jelas (Liu, 2012), seperti pada gambar 5 dibawah ini :

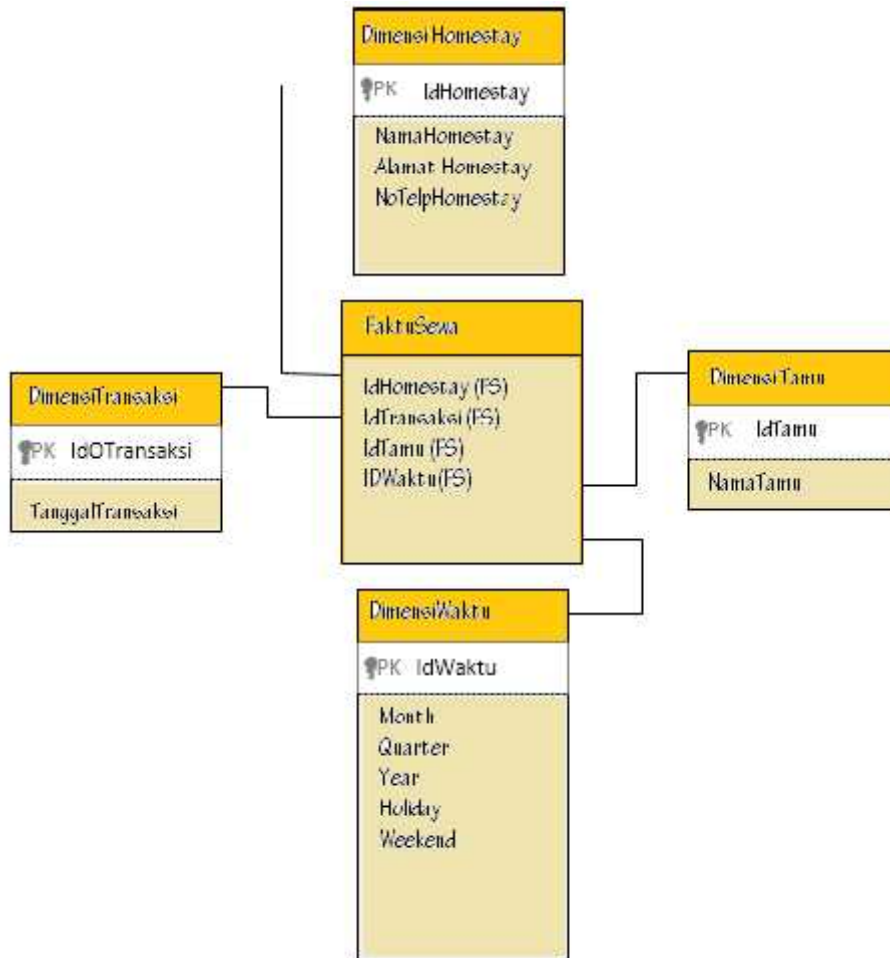


Gambar 5 Data Flow Diagram Level 1 (DFD)

### 3.5. Perancangan Data Warehouse

#### a. Perancangan Star Schema

Menurut (Connolly, 2013) star schema adalah model data dimensional yang mempunyai fact table di bagian tengah, dikelilingi oleh tabel dimensi yang terdiri dari data reference (yang bisa di-denormalized). Star schema mengambil karakteristik dari factual data yang di-generate oleh event yang terjadi di masa lampau, seperti ditunjukkan pada gambar 6 dibawah ini :



Gambar 6 Star Schema

3.6. Perancangan Tabel

1. Tabel Dimensi

a) Tabel DimensiTransaksi (ds\_Transaksi), dapat dilihat pada tabel 1 dibawah ini :

Tabel 1. Tabel DimensiTransaksi (ds\_Transaksi)

No	Nama Field	Type Data	Keterangan
1	IdTransaksi	Varchar(2)	Primary Key
2	TglTransaksi	date	

b) Tabel DimensiHomestay (ds\_Homestay), dapat dilihat pada tabel 2 dibawah ini :

Tabel 2. Tabel DimensiHomestay (ds\_Homestay)

No	Nama Field	Type Data	Keterangan
1	IdHomestay	Varchar(3)	PrimayKey
2	NamaHomestay	Varchar(20)	
3	AlamatHomestay	Varchar(32)	
4	NoTlpHomestay	Varchar(12)	

c) Tabel DimensiTamU (ds\_TamU), dapat dilihat pada tabel 3 dibawah ini :

Tabel 3. Tabel DimensiTamU (ds\_TamU)

No	Nama Field	Type Data	Keterangan
1	IdTamU	Varchar(1)	Primary Key
2	NamaTamU	Varchar(30)	

## 2. Tabel Fakta

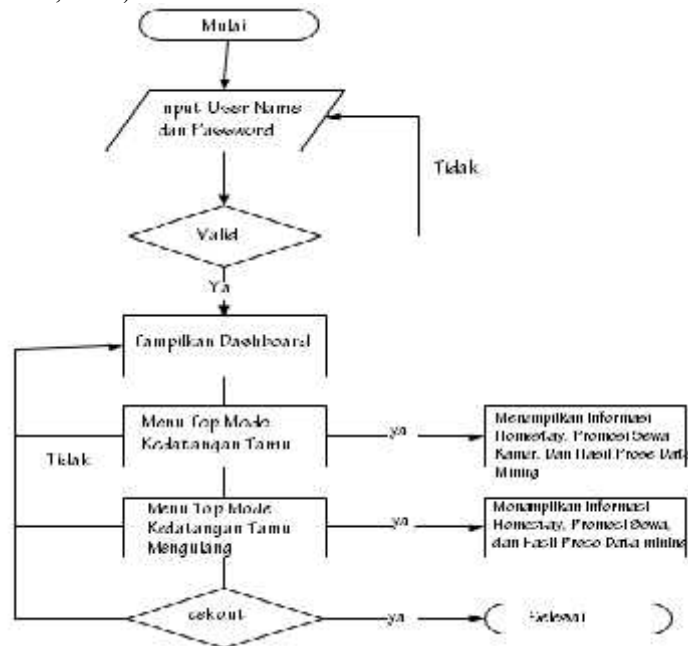
a) Tabel FaktaSewa (fS\_ Sewa), dapat dilihat pada tabel 4 dibawah ini :

Tabel 4. Tabel FaktaSewa (fS\_ Sewa)

No	Nama Field	Type Data	Keterangan
1	IdHomestay	Varchar(3)	Foreign Key
2	IdTransaksi	Varchar(2)	Foreign Key
3	IdTamu	Varchar(1)	Foreign Key
4	IdWaktu	Date	Foreign Key

### 3.7. Perancangan Flowchart

Flowchart atau Bagan alir adalah bagan (chart) yang menunjukkan alir (flow) di dalam program atau prosedur sistem secara logika. Bagan alir (flowchart) digunakan terutama untuk alat bantu komunikasi dan untuk dokumentasi, (Simon, 2014). Gambar 7. berikut adalah flowchart dari sistem:



Gambar 7. Flowchart

### 3.8. Perancangan Interface Aplikasi

Perancangan antarmuka (interface) merupakan bagian penting dalam perancangan aplikasi, karena berhubungan dengan tampilan dan interaksi pengguna dengan aplikasi.

## 4. KESIMPULAN DAN SARAN

### Kesimpulan

Dari hasil penelitian dapat disimpulkan bahwa :

1. Analisis dan perancangan teknik data mining dengan menggunakan metode association rule mining untuk memprediksikan tamu berulang datang menginap di homestay
2. Menghasilkan suatu rancangan aplikasi guna membangun sebuah aplikasi untuk menarik pelanggan, mempertahankan kepercayaan dan loyalitas pelanggan terhadap layanan sewa kamar homestay sehingga tamu berulang datang menginap di homestay.

### Saran

1. Menetapkan waktu untuk mengimplentasikan dalam bentuk sebuah model aplikasi data mining berbasis website.
2. Menambahkan dua atau lebih metode, algoritma data mining agar dapat menjadi pembanding antar satu metode dengan metode lainnya, mengingat banyaknya metode-metode data mining yang dapat digunakan dalam hal menganalisa dan merancang data menjadi sebuah informasi guna dijadikan strategi pemasaran sehingga banyak tamu berulang datang menginap di homestay

## 5. UCAPAN TERIMA KASIH

Terimakasih pada panitia semnas yang telah memberikan kesempatan untuk kami bisa menyeminarkan hasil penelitian kami dengan judul **Perancangan Sistem Informasi Aplikasi Data Mining Untuk Memprediksi**



**Kemungkinan Tamu Homestay Mengulangi Kedatangannya (Study Kasus Homestay Di Tawangmangu Kab Karanganyar)****DAFTAR PUSTAKA**

- [1] Aggarwal, C. C. An Introduction to Data Mining. In Data Mining. [https://doi.org/10.1007/978-3-319-14142-8\\_12](https://doi.org/10.1007/978-3-319-14142-8_12); 2015a
- [2] Aggarwal, C. C. Privacy-Preserving Data Mining. Data Mining. [https://doi.org/10.1007/978-3-319-14142-8\\_20](https://doi.org/10.1007/978-3-319-14142-8_20); 2015b
- [3] Bramer, M., Data for Data Mining. Principles of Data Mining. [https://doi.org/10.1007/978-1-4471-4884-5\\_4](https://doi.org/10.1007/978-1-4471-4884-5_4); 2013
- [4] Costa, F., Verbeke, M., & De Raedt, L. Relational regularization and feature ranking. In SIAM International Conference on Data Mining 2014, SDM 2014. <https://doi.org/10.1137/1.9781611973440.755>. 2014
- [5] Dai, H., Liu, J., & Smirnov, E. Reliable Knowledge Discovery. Reliable Knowledge Discovery. <https://doi.org/10.1007/978-1-4614-1903-7>; 2012
- [6] Data, M. T. Mining Text Data. Mining Text Data. [https://doi.org/10.1007/978-1-46143223-4\\_7](https://doi.org/10.1007/978-1-46143223-4_7). 2012
- [7] Dey, S., Simon, G., Westra, B. L., Steinbach, M., & Kumar, V. Mining interpretable and predictive diagnosis codes from multi-source electronic health records. In SIAM International Conference on Data Mining 2014, SDM 2014. <https://doi.org/10.1137/1.9781611973440.120>; 2014
- [8] Dhanjal, C., Gaudel, R., & Cl  men  on, S. Online matrix completion through nuclear norm regularisation. In SIAM International Conference on Data Mining 2014, SDM 2014. <https://doi.org/10.1137/1.9781611973440.72> ; 2014
- [9] Failasufa, F. (UNIVERSITAS I. N. S. K. PENERAPAN DATA MINING UNTUK ANALISIS POLA PEMBELIAN KONSUMEN DENGAN ALGORITMA APRIORI PADA DATA TRANSAKSI PENJUALAN. Skripsi. 2014
- [10] Han, J., Kamber, M., & Pei, J. Data Mining Concepts and Techniques. Data Mining. <https://doi.org/10.1016/B978-0-12-381479-1.00001-0> 11. 2012
- [11] Han, J., Kamber, M., Pei, J., & Pei, J. Data Mining. In Data Mining: Concepts and Techniques. <https://doi.org/10.1016/B978-0-12-381479-1.00016-2>; 2012
- [12] Hu, Z. ., Wang, H. . b, Zhu, J. ., Li, M. ., Qiao, Y. ., & Deng, C.. Discovery of rare sequential topic patterns in document stream. In SIAM International Conference on Data Mining 2014, SDM 2014. <https://doi.org/10.1137/1.9781611973440.61> 13; 2014
- [13] Kantardzic, M. Data-Mining Concepts. Data Mining. <https://doi.org/10.1002/9781118029145.ch1> ; 2011
- [14] Karpatne, A., Khandelwal, A., Boriah, S., & Kumar, V., Predictive learning in the presence of heterogeneity and limited training data. In SIAM International Conference on Data Mining 2014, SDM 2014. <https://doi.org/10.1137/1.9781611973440.29> ; 2014
- [15] Liang, T.-P., & Turban, E, Introduction to the Special Issue Social Commerce: A Research Framework for Social Commerce. International Journal of Electronic; 2014