BAB 2
TINJAUAN PUSTAKA

2.1 Profil E-kosan.com

E-kosan.com adalah situs yang berfungsi untuk menyediakan informasi tentang kosan. E-kosan.com bisa menjadi salah satu solusi mahasiswa sebagai pencari kosan karena dengan adanya e-kosan.com pencari kosan tidak perlu repot-repot untuk mencari kosan dengan cara manual adalah mencari kosan dengan cara mendatangi kosan satu persatu secara langsung. Tetapi dengan adanya e-kosan.com pencari kosan hanya perlu membuka situs e-kosan.com pencari kosan hanya perlu membuka situs e-kosan.com dan memilih kosan yang sesuai dengan keinginan kemudian mendatangi kosan tersebut.

Pengguna e-kosan terdiri dari pemilik kosan dan penghuni atau pencari kosan. Pendiri kosan bisa mendaftarkan kosannya agar dipubliksaikan oleh pegawai e-kosan.com. Sedangkan pencari kosan hanya bisa untuk mencari kosan yang diinginkan secara gratis.

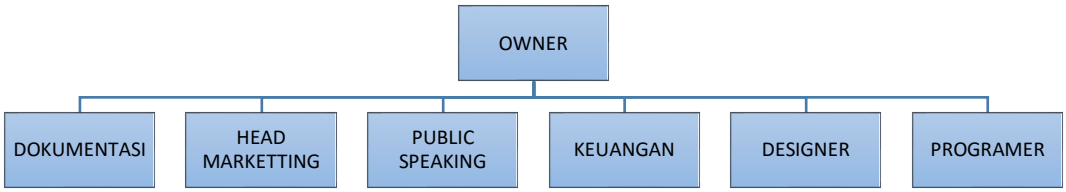
2.1.1 Sejarah

E-kosan.com adalah sebuah aplikasi berbasis situs yang bediri pada tanggal 22 November 2012 kemudian dipublikasikan pada tanggal 12 Desember 2012. Awalnya dibangun untuk memenuhi tugas besar mata kuliah analisi desain sistem informasi pada program studi Teknik Informatika Universitas Komputer Indonesia. Berkat keseriusan pembuat e-kosan sehingga membuat situs ini bertahan hingga kini.

Sejak Mei 2012 sampai sekarang (April 2015) E-kosan telah dikunjungi sebanyak 84.943 orang (berdasarkan data *Google Analytic*). Selain data dari *Google Analytic* e-kosan juga memiliki cukup banyak penggemar di *Facebook* da *Twitter*. Pada *Fans Page Facebook* terdapat 2.745 orang (April 2015) dan 3.517 orang yang mengikuti akun *Twitter* E-kosan (April2015).

2.1.2 Struktur Organisasi

Berikut adalah Struktur Organisasi yang berlaku pada E-Kosan:



Gambar 2. 1 Struktur Organisasi E-kosan.com

2.1.3 Logo

Berikut adalah logo yang digunakan E-Kosan.com:



Gambar 2. 2 Logo E-kosan.com

2.1.4 Job Description

Dari Struktur organisasi diatas, masing-masing jabatan mempunyai tugas untuk menjadikan situs e-kosan.com lebih berkembang. Adapun uraian tugas dari masing-masing jabatan tersebut adalah:

a. Owner

Owner mempunyai kewajiban dalam menyediakan modal, yang harus dikeluarkan untuk kebutuhan perusahaan dalam mengembangkan e-kosan.com.

b. Dokumntasi

Mempunyai tugas pokok melaksanakan pengumpulan pengelolahan data dalam rangka pelaksanaan kegiatan evaluasi pendokumentasian dan pelaporan kegiatan unsur-unsur organisasi di e-kosan.com

c. Head Marketing

Bertanggung jawab terhadap kelancaran dan terlaksananya tujuan pemasaran sesuai strategi dan sasaran pemasaran yang ditetapkan oleh e-kosan.com.

d. Public Speaking

Mempunyai tugas pokok yaitu menyampaikan pesan atau ide kepada public dengan metode yang sesuai sehingga publik bisa memahami pesan atau ide, dan kemudian memperoleh manfaat dari pesan tersebut.

e. Keuangan

Keuangan mempunyai tugas pokok melaksanakan penyusunan rencana kegiatan, pengkoordinasian, pembinaan dan pengandalian terhadap anggaran yang terjadi di e-kosan.com

f. Designer

Mempunyai tugas pokok membuat desain website e-kosan.com, desain ini dimaksudkan dalam mengambangkan website e-kosan.com yang lebih baik dari yang sebelumnya. Sehingga dapat dibuka atau dinikmati pada sebuah layar seperti layar komputer, layar handphone,layar tablet.

g. Admin *Twitter*

Mempunyai tugas pokok yaitu pengawasan pengoperasian kegiatan dalam mempromosikan e-kosan.com dimedia sosial *Twitter*.

h. Admin *Facebook*

Mempunyai tugas pokok yaitu pengawasan pengoperasian kegiatan dalam mempromosikan e-kosan.com dimedia sosial *Facebook*.

i. Programer

Membangun/mengembangkan website e-kosan.com terutama pada tahap *construction* dengan melakukan *coding* dengan bahasa pemrograman, mengimplementasikan *requiretment* dan desain proses bisnis ke komputer dengan menggunakan algoritm atau logika dan bahasa pemrograman dan melakukan testing terhadap website e-kosan.com.

2.2 Landasan Teori

2.2.1 Data Mining

Menurut David Hand, Heikki Mannila dan Padhraic Smyth dalam bukunya *Princples of Data mining* [3], data mining merupakan proses analisa dari sekumpulan (terkadang sangat besar) data pengamatan untuk menemukan adanya hubungan-hubungan yang tidak terduga sebelumnya dan untuk merangkum data yang menjadi bentuk yang mudah dimengerti dan berguna bagi pemilik data. Dari pengertian tersebut dapat ditarik kesimpulan bahwa konsep data mining berhubungan dengan data jumlah yang sangat besar. Tujuan dari data mining adalah berusaha mencari manfaat dari sekumpulan data tersebut.

Dilihat dari disiplin ilmu yang digunakan, data mining merupakan ilmu multi disiplin. Data mining menyangkut berbagai disiplin ilmu seperti database, kecerdasan buatan(*artificial intellegence*), *information science*(ilmu informasi), *high performace computing*, visualisasi, *machine learning*, statistik, *neural networks*(jaringan syaraf tiruan), pemodelan matematika, *information retrieval* dan *information extraction* serta pengenalan pola.

Data mining juga dikenal dengan nama *Knowledge Discovery(mining) in database* (KDD), ektrasi pengetahuan *(knowledge extaction)*. Secara umum, definisi data mining dapat diartikan sebagai berikut[4]:

1. Proses penemuan pola yang menarik dari data yang tersimpandalam jumlah besar.
2. Ekstraksi dari suatu informasi yang berguna atau menarik(*non-trivial*, implisit, sebelumnya belum diketahui potensian kegunaannya) pola atau pengatahuan dari data yang disimpan dalam jumlah besar.

3. Eksplorasi dari analisa secara otomatis atau semiotomatis terhadap data-data dalam jumlah besar untuk mencari pola dan aturan yang berarti.

2.2.2 Web Mining

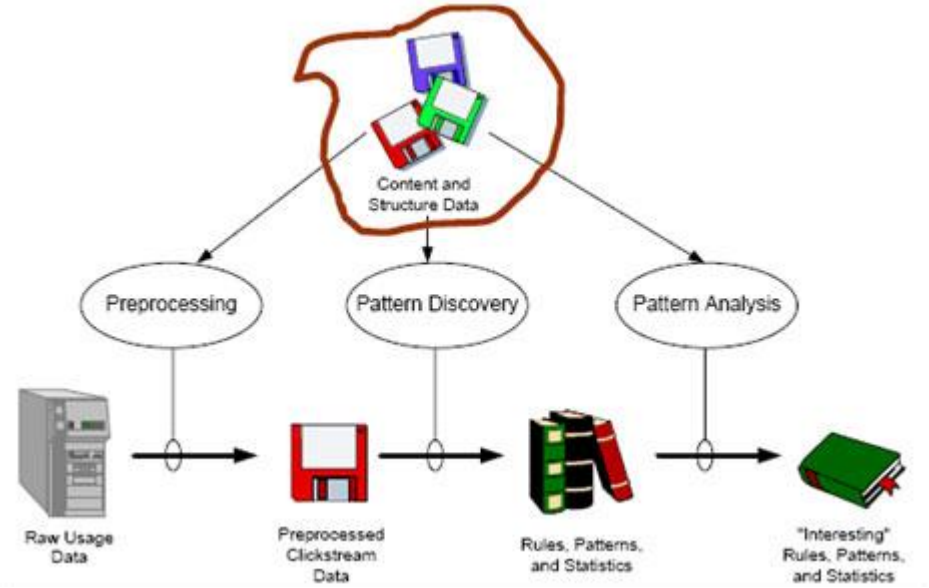
Web Mining merupakan suatu aplikasi bagian dari data mining yang berusaha menggali pola-pola yang tersedia didalam web itu sendiri. Jadi antara *data mining* dan *web mining* hanya berbeda dalam hal target data yang dianalisa. Data mining umumnya menganalisa data yang berasal dari OLTP*(Online Transactional Process)* dan data transaksi lainnya. Sedangkan *web mining* target analisanya adalah dari data web, seperti data akses pengunjung, struktur halaman web, format halaman web dan sebagainya[5].

2.2.2.1 Web Usage Mininig

Menurut Srivastava, *web usage mining* merupakan teknik data mininig yang berusaha mengungkap pola penggunaan dari halaman web, dalam rangka coba untuk memahami dan meningkatkan pelayanan kebutuhan dari aplikasi berbasis web[3]. Pada *web usage mining* yang dianalisanya adalah pengguna atau pengunjung dari halaman web. Sehingga karena yang coba dianalisa adalah tingkah laku dari pengunjung (pengguna) dari web maka hasil dari *web usage mining* banyak digunakan dalam *e-marketing* dan *e-commerce*. Hasil analisa dapat digunakan untuk meningkatkan layanan dari aplikasi web.

Hasil web usage mining antara lain informasi mengenai segmentasi pengunjung dari situs (aplikasi web). Segmentasi dapat dilihat berdasrkan lokasi (negara, kota atau wilayah), waktu akses (pagi, siang, sore, atau malam), penggunaan browser dan sebagainya. Dalam situs *ecommerce* misalnya dapat digunaan untuk melihat pola pengunjung dalam pembeliaan produk seperti produk apa saja yang paling banyak dibeli(diakses), pengunjung dari mana saja yang banyak melakukan pembelian, dan sebagainya.

Secara garis besar, proses web usage mining terbagi menjadi 3 (tiga) fase, yaitu *preprocessing*, *pattern discovery* dan *pattern analysis*, digambarkan menganai proses web usage mining seperti terlihat pda gambar berikut:



Gambar 2. 3 Proses *Web Usage Mining*

a. *Fase Pra-processing*

Tujuannya adalah untuk mengubah data web usage mining menjadi data yang dapat diandalkan. Proses data *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data seperti kesalahan cetak(*tipografi*). Pada proses persiapan ini data asli dari data log kunjungan disaring sehingga hanya data yang valid saja yang akan diolah.

b. *Fase Pattern Discovery*

Tujuan dari fase untuk menemukan pola dengan menggunakan teknik seperti: *Path analysis, Associoation rule, Classification, Clustering.*

c. *Fase Analysis*

Fase ini biasanya menggunakan teknik seperti: OLAP, *Knowledge Query Management, Intelligent Agents, Visulization.*

2.2.3 Association Rules

Association rules merupakan salah satu teknik data mining yang berfungsi untuk menemukan asosiasi antara variabel, korelasi atau suatu struktur diantara item atau objek-objek didalam data base transaksi, database relasional, maupun pada penyimpanan informasi lainnya. *Association Rules Discovery* seringkali

digunakan dalam *Market Basket Analysis* dan juga sebagai *Affinity Grouping*. *Association rules mining* digunakan untuk mencari pola-pola yang sering muncul, asosiasi-asosiasi, korelasi atau hubungan sebab akibat diantara himpunan dari item-item atau objek-objek dalam database[6].

Fungsi ini paling banyak digunakan untuk manganalisa data dalam rangka keperluan strategi pemasaran, desain katalog, dan proses pembuatan keputusan bisnis.Tipe *association rules* bisa dinyatakan sebagai misal:"70% dari orang-orang yang membeli mie, jus dan saus akan membeli juga roti tawar".

Aturan asosiasi mangcapture item atau kejadian dalam data berukuran besar yang berisi data transaksi. Dengan kemajuan teknologi, data penjualan dapat disimpan dalam jumlah besar yang disebut dengan "basket data". Aturan asosiasi yang didefinisikan pada basket data, digunakan untuk keperluan promosi, desain katalog, segmentasi costumer, dan target pemasaran. Secara tradisional, aturan asosiasi digunakan untuk menemukan trend bisnis dengan menganalisa transaksi costumer. Dan dapat digunakan secara efektif pada bidang web mining yang diilustrasikan sebagai berikut: pada log pengunjung, kita menemukan bahwa aturan asosiasi:"A dan B implies C", memiliki nilai *confidence* 80%, dimana A, B, dan C adalah halaman web yang bisa diakses. Jika seorang user mengunjungi halaman A dan B, maka terdapat 80% kemungkinan dia akan mengunjungi halaman C juga pada sesion yang sama, sehingga halaman C perlu diberi direct link dari A atau B. Informasi ini dapat digunakan untuk membuat link secara dinamik ke halaman C dari halaman A atau B sehingga user dapat melakukan direct link ke halaman C. Informasi semacam ini digunakan untuk melakukan link ke halaman produk yang berbeda secara dinamik berdasarkan interaksi web visitors.

Adapun rumus untuk perhitungan *support* dan *confidence* adalah[7]:

a. *Support*

   Support adalah prosentase kombinasi item tersebut dalam database,

   Rumus untuk mencari nilai *support* dari 1 item adalah:

$$Support(A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ transaksi} \qquad (2.1)$$

Sedangkan jika mencari nilai support 2 item maka rumusnya adalah:

$$Support(A \cup B) = \frac{Jumlah\ transaksi\ mengandung\ A\ dan\ B}{Total\ transaksi} \quad (2.2)$$

b. *Confidence*

Confidence adalah kuatnya hubungan antar item dalam aturan asosiatif. Rumus *confidence* adalah:

$$Confidence(A|B) = \frac{Jumlah\ transaksi\ mengandung\ A\ dan\ B}{Total\ transaksi\ mengandung\ A} \quad (2.3)$$

2.2.3.1 Algoritma CT-Pro

Algoritma CT-Pro merupakan algoritma yang dimodifikasi dari algoritma *FP-Growth* dimana modifikasi yang dilakukan adalah pada struktur data yang digunakan. Struktur data yang digunakan adalah *Compressed FP-Tree* (*CFPTree*) dimana informasi dari sebuah *FP-Tree* diringkas dengan struktur yang lebih kecil, sehingga baik pembentukan *tree* maupun *frequent itemset mining* dapat dilakukan lebih cepat[9].

Langkah-langkah algoritma CT-Pro yaitu[2]:

1. Menemukan item-item frequent.
2. Membuat *Global CFP-Tree* dimulai dari jumlah frequent terbesar ke frequent terkecil pada Global item table. Setelah terurut kemudian dibuat *Global CFP-Tree*.
3. Melakukan Mining
   a. Dimulai dari mengurutkan *Global item table* mulai dari item berfrekuensi terkecil hingga terbesar.
   b. Untuk setiap item yang terdaftar ada *Global item table* yang telah diurutkan, dilakukan pencarian node yang berkaitan dengan item

tersebut pada *Global CFP-Tree*. Dari semua node yang ditemukan untuk setiap item inilah yang disebut dengan *Local Frequent item* dan digunakan untuk membuat *Local item table* yang pembuatannya dilakukan berdasarkan jumalah minimun *support* yang telah ditentukan.

c. Setelah itu, dibuat *local CFP-Tree* berdasarkan *Local item table* yang terbentuk.
d. Dari *Local CFP-Tree* dibentuk *frequent pattern* sesuai dengan item yang di-mining. Dari *frequent pattern* dihitung dari masing-masing item yang memenuhi dihitung *confidence*-nya. Apabila memenuhi minimum *confidene*, maka masing-masing item yang bersangkutan dijadikan *knowledge*.

2.2.4 Teknik Visualisasi *Association Rule*

Visualisasi adalah proses transformasi data, informasi, dan pengetahuan dalam bentuk visual yang membuat penggunaan kemampuan visual alami manusia. Khas dari aplikasi visualisasi adalah bidang komputer grafis. Terlepas dari pentingnya visualisasi, ada banyak keterbatasan dan kesulitan yang harus diambil dalam pertimbangan seperti [11].

Memvisualisasikan aturan asosiasi bertujuan memecahkan beberapa masalah utama yang datang dengan aturan asosiasi. Pertama-tama aturan (*rule*) ditemukan oleh prosedur otomatis harus disaring. Tergantung pada apa confidence minimum dan minimum support yang ditentukan sehingga mendapatkan banyaknya jumalah aturan (*rule*) yang dihasilkan.

Menggunakan visualisasi, pengguna dapat melihat kelompok aturan terkait dan dampak dari tingkat perbedaan nilai *support* dan nilai *confidence*. Tujuan dari visualisasi aturan asosiasi adalah untuk memvisualisasikan sejumlah besar aturan asosiasi dan metadata mereka dalam dua dimensi (2D) atau tiga dimensi display (3D) dengan interaksi manusia minimum, oklusi minimum, dan tidak ada swapping layar. Ada banyak pendekatan yang dikembangkan untuk memvisualisasikan aturan asosiasi diantaranya:

1. Rule Table

2. two-dimensional matrix

3. directed graph

4. rule-item approach

5. Mosaic Plot

6. Double Decker Plot,

7. Parallel Coordinates,

2.2.4.1 Rule Table

Metode yang paling sederhana untuk visualisasi aturan asosiasi adalah dengan menggunakan tabel aturan. Berikut format tabel aturan telah digunakan :

Tabel 2. 1 Format Tabel Aturan

Items	Rule N	Confidence	Support

Contoh bentuk dari visualisasi aturan assosiasi adalah sebagai berikut :

Tabel 2. 2 Contoh Implementasi Tabel Aturan

Items	Rule N	Confidence	Support
Telur Mie	1	80%	10%
Mie Kopi Gula	2	75%	0,9%

Pada baris pertama dapat diartikan sebagai berikut :

Telur $\overset{80\%}{\longrightarrow}$ Mie dengan Support 10%

Tabel aturan adalah cara yang paling mudah untuk menunjukkan aturan asosiasi kepada pengguna.

2.3 Alat-Alat Pemodelan Sistem

Alat-alat pemodelan sistem adalah alat (*tools*) yang dibutuhkan dalam proses perancangan sistem.

2.3.1.1 Diagram Konteks

Diagram konteks adalah diagram aliran data pada tingkat paling atas yang merupakan penggambaran yang berfungsi untuk memperlihatkan interaksi atau hubungan langsung antara sistem informasi dengan lingkungannya. Diagram konteks menggambarkan sebuah sistem berupa sebuah proses yang berhubungan dengan satu atau beberapa entitas/entity.

2.3.1.2 Data Flow Diagram

Data Flow Diagram (DFD) adalah suatu alat pemodelan yang digunakan untuk memodelkan fungsi dari sistem, menggambarkan secara rinci mengenai sistem sebagai jaringan kerja antar fungsi yang berhubungan satu sama lain dengan menunjukan dari dan kemana data mengalir serta menyimpannya. Pada umumnya dimulai dari 0, 1, 2,dst. Level ke-0 disebut dengan diagram konteks yang menggambarkan sistem secara global.

2.3.1.3 Kamus Data

Kamus data adalah suatu penjelasan tertulis mengenai data yang berada dalam database. Kamus data ikut berperan dalam perancangan dan pembangunan sistem informasi karena berfungsi untuk menjelaskan arti aliran data dan menyimpan dalam menggambarkan data flow diagram, menjelaskan spesifikasi nilai dan satuan yang relevan terhadap data yang mengalir dalam sistem tersebut, mendeskripsikan komposisi paket data yang bergerak melalui aliran. Fungsi kamus data atau *data dictionary*[10]:

1. Menjelaskan arti aliran data dan penyimpanan data dalam DFD.
2. Mendeskripsikan komposisi paket data (elemen data) yang bergerak melalui aliran data.
3. Mendeskripsikan nilai dan satuan (struktur data) yang relevan bagi penyimpanan dan aliran.
4. Mendeskripsikan hubungan detail antar penyimpanan.

2.3.1.4 Skala Pengukuran

Skala pengukuran merupakan kesepakatan yang digunakan sebagai acuan untuk menentukan panjang pendeknya interval yang ada dalam alat ukur, sehingga alat ukur tersebut bila digunakan dalam pengukuran akan menghasilkan data kuantitatif . Macam-macam skala pengukuran dapat berupa skala nominal, skala ordinal, skala interval dan skala rasio, dari skala pengukuran itu akan diperoleh data nominal, ordinal, interval dan ratio.

Skala yang dapat digunakan untuk penelitian administrasi, pendidikan dan sosila antar alain adalah:

1. Skala Likert
2. Skala Guttman
3. Rating Scale
4. Semantic Defential

Kelima jenis skala tersebut bila digunakan dalam pengukuran, akan mendapatkan data interval atau rasional.

Skala likert digunakan untuk mengukur sikap, pendapat, dan persepsi seseorang atau sekelompok orang tentang fenomena sosial. Dengan skala likert variable yang akan diukur dijabarkan menjadi indikator variabel. Kemudian indikator tersebut dijadikan sebagai titik tolak untuk menyusun item-item instrumen yang dapat berupa pernyataan atau pertanyaan. Jawaban dari setiap item instrumen mempunyai gradasi dari sangat positif sampai sangat negatif yang diantaranya yaitu :

a. Sangat setuju	a.Selalu
b. Setuju	b. Sering
c. Ragu-ragu	c. Kadang-kadang
d. Tidak Setuju	d. Tidak pernah
e. Sangat Tidak setuju	

a. Sangat positif	a. Sangat baik
b. Positif	b. Baik
c. Negatif	c. Tidak baik
d. Sangat negatif	d. Sangat tidak baik

Untuk keperluan analisis kuantitatif, maka jawaban itu dapat diberik skor misalnya:

1. Setuju/selalu/sangat positif diberi skor	5
2. Setuju/sering/positif diberi skor	4
3. Ragu-ragu/kadang-kadang/netral diberi skor	3
4. Tidak setuju/hampir tidak pernah/negatif diberi skor	2
2. Sangat tidak setuju/tidak pernah diberi skor	1

Instrumen penelitian yang menggunakan skala *liker* dapat dibuat dalam bentuk checklist ataupun pilihan ganda

Pengujian dilakukan dengan melihat hasil dari Kuesioner yang telah diisi, untuk penghitungannya dilakukan dengan menggunakan rumus dibawah ini:

$$\mathbf{Y = P/Q * 100\%} \qquad (2.3)$$

Keterangan: Y : Nilai Presentase

P : Total Skor

Q : Skor Tertinggi