

MODEL DATA MINING DALAM PENENTUAN KELAYAKAN PEMILIHAN TEMPAT TINGGAL MENGGUNAKAN METODE NAIVE BAYES

Diana Laily Fithri

Fakultas Teknik, Prodi Sistem Informasi
Universitas Muria Kudus
Email: diana.laily@umk.ac.id

ABSTRAK

Berkembangnya teknologi dan kemajuan pertumbuhan penduduk di Indonesia, mengakibatkan semakin meningkatnya daya beli untuk kebutuhan tempat tinggal. Hal ini terbukti dengan semakin banyaknya, perumahan-perumahan yang sudah disiapkan dengan menggunakan cara angsuran atau cicilan yang murah. Oleh karena itu, banyak konsumen yang masih merasa kesulitan dalam menentukan letak atau lokasi tempat tinggal. Tempat tinggal yang baik adalah tempat tinggal yang dapat memberikan kenyamanan bagi penghuninya. Penelitian ini membahas tentang pemilihan tempat tinggal yang layak sesuai dengan 7 kriteria, yaitu Keamanan, fasilitas Umum, bebas banjir, harga, air bersih, model, sejuk dan nyaman. Dalam penentuan kelayakan pemilihan tempat tinggal dengan metode *Naive Bayes* dan pemodelan *Data mining* dihasilkan 3 kriteria utama dalam penentuan kelayakan pemilihan tempat tinggal yaitu keamanan, bebas banjir dan harga tempat tinggal. Model *Data mining* untuk penentuan tempat tinggal menggunakan metode *Naive Bayes* memberikan manfaat yang besar untuk konsumen dalam mempertimbangkan penentuan pencarian tempat tinggal, yang diantaranya menghasilkan probabilitas antara layak (sesuai) atau tidak sesuai, yang nantinya dapat digunakan oleh masyarakat umum untuk mempermudah dalam penentuan pencarian tempat tinggal.

Kata kunci: *naive bayes*, masyarakat, tempat tinggal.

ABSTRACT

Development of technology and the progress of population growth in Indonesia, resulting in increasing the purchasing power for housing needs. This is evidenced by the increasing number, estates that had been prepared by using installment or installments cheap. Therefore, many consumers still find it difficult to determine the position or location of residence. Good place to live is a residence that can provide comfort for occupants. This study discusses the selection of adequate housing according to seven criteria, namely security, public facilities, flood free, price, clean air, model, cool and comfortable. In determining the feasibility of the election of a residence with Naive Bayes methods of Data mining and modeling produced three main criteria in determining the feasibility of selecting a place to stay that is security, flood free and priced place to stay. Data mining models to the place of residence using the naïve Bayes provide great benefits to consumers in considering the determination finding a place to stay, which include generating a probability between decent (as appropriate) or do not conform, which can later be used by the general public to facilitate the determination of the search space stay.

Keywords: *naive bayes, community, place of residence.*

1. PENDAHULUAN

Pertumbuhan penduduk di Indonesia, mengalami peningkatan yang cukup signifikan, dan ini mengakibatkan banyak para penduduk yang mencari tempat tinggal juga mengalami peningkatan. Di semua daerah mulai dari desa sampai kota melakukan lahan kosongnya untuk pembangunan perumahan-perumahan. Perumahan yang idelal dan sehat harus tidak boleh lepas dari air bersih, keamanan, fasilitas, dan transportasi. Tempat tinggal idaman adalah tempat tinggal yang sehat dimana bangunan tempat berlindung dan beristirahat serta sebagai sarana pembinaan keluarga yang dapat menumbuhkan kehidupan sehat secara fisik, mental dan social, sehingga seluruh keluarga dapat bekerja secara produktif. Teknologi informasi yang berkembang semakin cepat di kehidupan manusia dapat memberikan dampak positif dalam kehidupan sehari-hari dan nantinya dapat membantu dan mempermudah dalam pencarian dan penentuan tempat tinggal. Dalam penelitian ini, kriteria yang digunakan diantaranya sanitasi air bersih, keamanan lokasi, harga dan fasilitas rumah untuk tempat tinggal, pembuangan sampah, transportasi, model rumah, dan tentunya harus bebas banjir. Metode *Naive Bayes* untuk penentuan dalam

pemilihan tempat tinggal digunakan untuk memudahkan konsumen dalam memilih tempat tinggal sesuai dengan keinginan konsumen, yang dapat menghasilkan tempat tinggal layak huni atau tidak sesuai.

Menurut Siti Agustin (2012) dalam penelitiannya yang berjudul Analisis penunjang keputusan pemilihan tempat tinggal dengan Metode *Analytical Hierarchy Process* (AHP) secara online berdasarkan jenis kelamin menjelaskan tentang tempat tinggal adalah tempat tinggal yang dapat memberikan keamanan dan nyaman bagi penghuninya, dengan menggunakan kriteria-kriteria dalam pemilihan, yaitu Air bersih, keamanan, bebas banjir, dan fasilitas. Tahapan dari penelitian tersebut, setelah membuat analisa maka membuat perankingan berdasarkan kriteria-kriteria tersebut.

Menurut I wayan Supriyana (2012) dalam penelitiannya yang berjudul Sistem Pendukung Keputusan dalam Pemilihan Tempat Kost dengan Metode Pembobotan (Studi Kasus Sleman Yogyakarta) menjelaskan bahwa Penentuan tempat tinggal di daerah yang baru dikenal dipengaruhi oleh banyak factor, diantaranya lokasi, fasilitas, sistem kontrak dan harga. Makalah ini menentukan kriteria-kriteria yang digunakan didalam pemilihan tempat tinggal. kriteria-kriteria tersebut dianalisis menggunakan metode pembobotan.

Menurut Angga Rifa Nur Rijati dalam penelitiannya yang berjudul Sistem Pendukung Keputusan untuk pemilihan Perumahan dengan menggunakan Metode *Fuzzy Tahani* (Studi Kasus PT.Bukit Semarang Jaya Metro) menjelaskan bahwa Metode yang digunakan dalam menentukan kriteria perumahan adalah *Fuzzy Tahani* yang dapat menghasilkan informasi mengenai profil perumahan berdasarkan tujuh kriteria yakni: Harga, Desain rumah, Luas tanah , Luas bangunan , Fasilitas pendukung , Lokasi dan Waktu tempuh ke pusat kota. Kriteria tersebut dapat digunakan konsumen dalam membantu untuk pemilihan perumahan.

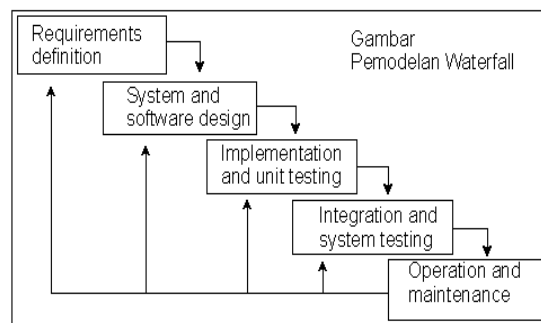
Menurut Ahmad Efendi dalam penelitiannya yang berjudul Rancang Bangun Sistem Pendukung Keputusan Pemilihan tempat kost dengan menggunakan *Weight Product* menjelaskan tentang menentukan pemilihan tempat kost dipengaruhi oleh banyak faktor, diantaranya lokasi, fasilitas, sistem kontrak, dan harga. Skripsi ini menentukan kriteria-kriteria yang akan digunakan dalam pemilihan tempat tinggal. Untuk membantu dalam penentuan dalam menetapkan pilihan tempat kost maka dibutuhkan sistem pendukung keputusan.

Oleh Karena itu, untuk mengembangkan penelitian yang sudah ada, penelitian ini menambah kriteria-kriteria yang sudah ada, diantaranya Sanitasi air bersih, pembuangan sampah, transportasi, Fasilitas, Keamanan, Bebas banjir, Model Rumah, Harga dan Fasilitas Rumah.

2. METODOLOGI PENELITIAN

2.1 Metode Pengembangan Sistem

Metode pengembangan sistem yang digunakan untuk Model *Data mining* dalam penentuan Kelayakan pemilihan tempat tinggal adalah dengan menggunakan Metode Rekayasa Perangkat Lunak Air Terjun (*Waterfall*) menurut Summerfile [6]. Tahapan Metode pengembangan sistem menurut Summerfile dapat dilihat di Gambar 1.



Gambar 1: Tahapan *Waterfall* [6]

Tahap – tahap dalam model Air Terjun tersebut antara lain :

- 1) Pendefinisian Masalah dan Analisis Kebutuhan
Pada tahap ini, meliputi penentuan pokok-pokok permasalahan yang terjadi dan kebutuhan dalam perancangan penentuan pemilihan kelayakan tempat tinggal.
- 2) Perancangan Sistem dan Perangkat Lunak
Pada tahap ini, meliputi perancangan sistem informasi dan perangkat.

- 3) Implementasi dan Unit Testing
Pada tahap ini, meliputi penerapan dan pengujian hasil perancangan.
- 4) Integrasi dan Pengujian Sistem
Pada tahap ini, meliputi penyatuan dan pengujian Sistem secara menyeluruh.
- 5) Pengoperasian dan Perawatan
Pada tahap ini, meliputi pengoperasian Sistem dan pemeliharaan fakta yang ada di lapangan.

2.2 Metode Naive Bayes

Teorema keputusan *bayes* adalah pendekatan statistik yang fundamental dalam pengenalan pola (*pattern recognition*). Pendekatan ini didasarkan pada kuantifikasi *trade-off* antara berbagai keputusan klasifikasi dengan menggunakan probabilitas dan ongkos yang ditimbulkan dalam keputusan keputusan tersebut. Ide dasar dari *bayes* adalah menangani masalah yang bersifat hipotesis yakni mendesain suatu klasifikasi untuk memisahkan objek. Misalkan terdapat dua jenis objek dengan kemungkinan kemunculan random, selanjutnya ingin diprediksi objek apa yang akan lewat selanjutnya.

Objek pertama diwakili oleh h_1 dan objek kedua diwakili oleh h_2 . karena apa yang akan muncul bersifat probablistik maka h adalah suatu variabel yang harus di deskripsikan secara probablistik. Selanjutnya probabilitas a priori, $(P h_1)$ dan $(P h_2)$ masing-masing melihat peluang munculnya objek 1 dan objek 2. Walaupun probabilitas kemunculan kedua objek tersebut tidak diketahui dengan pasti tapi setidaknya dapat diestimasi dari data yang tersedia. Misalkan N adalah jumlah total kedua objek, kemudian N_1 dan N_2 masing masing menyatakan jumlah objek 1 dan objek 2, selanjutnya

$$P(h_1) \approx \frac{N_1}{N} \text{ dan } P(h_2) \approx \frac{N_2}{N} \quad (1)$$

Untuk mengestimasi probabilitas digunakan rumus, yang sering disebut m-estimate :

$$P(a_i | v_j) = \frac{n_c + mp}{n + m} \quad (2)$$

2.3 Data mining

Data mining sering didefinisikan sebagai proses penggalian yang valid, dipahami sebelumnya dan tidak diketahui, informasi dari database yang besar dalam rangka meningkatkan dan mengoptimalkan keputusan bisnis. Secara garis besar *Data mining* dapat dikelompokkan menjadi 2 kategori utama:

- 1) *Descriptive mining*, yaitu proses untuk menemukan karakteristik penting dari data dalam suatu basis data. Teknik *Data mining* yang termasuk dalam *Descriptive mining* adalah *clustering*, *association*, dan *sequential mining*.
- 2) *Predictive*, yaitu proses untuk menemukan pola dari data dengan menggunakan beberapa variable lain di masa depan. Salah satu teknik yang terdapat dalam *predictive mining* adalah klasifikasi.

Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data *warehouse*, atau penyimpanan informasi lainnya. *Data mining* berkaitan dengan bidang ilmu – ilmu lain, seperti database system, data warehousing, statistik, machine learning, information retrieval, dan komputasi tingkat tinggi. Selain itu, *Data mining* didukung oleh ilmu lain seperti neural network, pengenalan pola, spatial data *analysis*, *image database*, *signal processing*. *Data mining* didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringkali semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Data yang dibutuhkan dalam jumlah besar. *Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu:

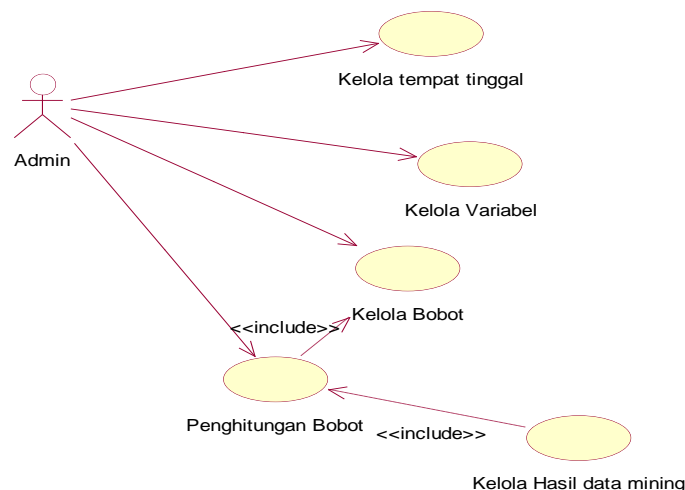
- 1) Deskripsi
Deskripsi adalah menggambarkan pola dan kecenderungan yang terdapat dalam data secara sederhana. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.
- 2) Klasifikasi
Suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang ada yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Klasifikasi menggunakan *supervised learning*.

- 3) Estimasi
Estimasi hampir sama dengan klasifikasi, perbedaannya adalah variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun dengan menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi.
- 4) Prediksi
Prediksi memiliki kesamaan dengan klasifikasi dan estimasi, dalam prediksi nilai dari hasil prediksi akan ada dimasa mendatang. Beberapa teknik yang digunakan dalam klasifikasi dan estimasi dapat juga digunakan (untuk keadaan yang tepat) untuk prediksi.
- 5) Klastering
Klastering merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas objek - objek yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan record-record dalam kluster lain. Klastering menggunakan *unsupervised learning*.
- 6) Asosiasi
Tugas asosiasi atau sering disebut juga sebagai "market basket analysis" dalam *Data mining* adalah menemukan relasi atau korelasi diantara himpunan item-item dan menemukan atribut yang muncul dalam satu waktu. Asosiasi menggunakan *unsupervised learning*. Penting tidaknya suatu aturan assosiatif dapat diketahui dengan dua parameter, *support* dan *confidence*.

3. HASIL DAN PEMBAHASAN

3.1 Analisa Use Case Sistem Kelayakan Pemilihan Tempat Tinggal

Use case diagram adalah diagram yang menggambarkan seorang actor dalam mengelola proses-proses yang ada dalam sistem tersebut. *Use Case* diagram penentuan pemilihan kelayakan tempat tinggal dapat dilihat dalam Gambar 2.



Gambar 2. Use Case Diagram Kelayakan Pemilihan Tempat Tinggal

3.2 Proses Penghitungan Pembobotan Penentuan Pemilihan Kelayakan Tempat Tinggal

Penghitungan pembobotan untuk penentuan pemilihan kelayakan tempat tinggal memiliki kriteria diantaranya :

- 1) Sanitasi air bersih
- 2) Pembuangan Sampah
- 3) Transportasi
- 4) Fasilitas
- 5) Keamanan
- 6) Bebas rumah
- 7) Model Rumah

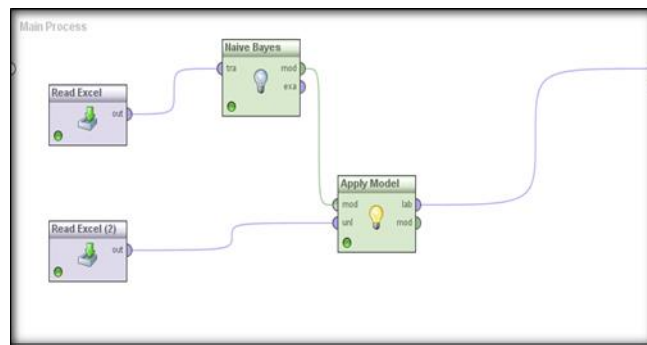
Berdasarkan kriteria yang digunakan untuk menentukan pemilihan kelayakan tempat tinggal, ada 3 kriteria dengan nilai yang paling tinggi yaitu Keamanan, bebas banjir dan model Rumah. Penghitungan pembobotan penentuan pemilihan kelayakan tempat tinggal dapat dilihat dalam Tabel 1.

Tabel 1. Proses penghitungan penentuan pemilihan kelayakan tempat tinggal

<i>Kriteria</i>	<i>Pembobotan</i>	<i>Hasil</i>
P (Sanitasi Air bersih / Layak)	= $(5+(7*0,5))/(4+3)$	= 0,35
P (Sanitasi Air bersih / Tidak Layak)	= $(5+(3*0,5))/(266+3)$	= 0,99
P (Pembuangan sampah / Layak)	= $(6+(7*0,5))/(4+3)$	= 1,35
P (Pembuangan sampah / Tidak Layak)	= $(4+(3*0,5))/(266+3)$	= 3,02
P (Transportasi / Layak)	= $(4+(7*0,5))/(4+3)$	= 1,07
P (Transportasi / Tidak Layak)	= $(6+(3*0,5))/(266+3)$	= 3,02
P (Fasilitas / Layak)	= $(5+(7*0,5))/(4+3)$	= 1,21
P (Fasilitas / Tidak Layak)	= $(5+(3*0,5))/(266+3)$	= 3,02
P (Keamanan / Layak)	= $(7+(7*0,5))/(4+3)$	= 1,5
P (Keamanan / Tidak Layak)	= $(3+(3*0,5))/(266+3)$	= 3,01
P (Bebas Banjir / Layak)	= $(7+(7*0,5))/(4+3)$	= 1,5
P (Bebas Banjir / Tidak layak)	= $(3+(7*0,5))/(266+3)$	= 3,02
P (Model Rumah /layak)	= $(7+(7*0,5))/(4+3)$	= 1,5
P (Model Rumah/Tidak layak)	= $(3+(7*0,5))/(266+3)$	= 3,02

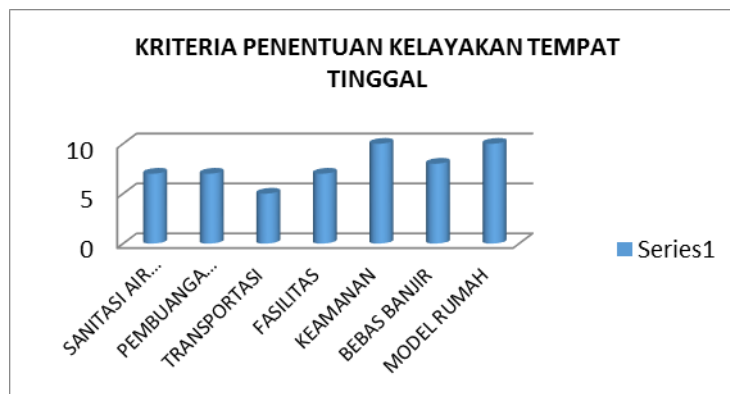
3.3 Model Data Mining Untuk Penentuan Kelayakan Tempat Tinggal

Model *Data mining* dengan menggunakan metode *Naive Bayes* dapat digunakan untuk mempermudah pengguna dalam menentukan kelayakan pemilihan tempat tinggal. Karena dalam metode tersebut, harus menggunakan data training dan data testing untuk menentukan hasil tentang penentuan kelayakan tempat tinggal. Model *Data mining* dengan menggunakan software *Rapid Miner* dapat dilihat di Gambar 3.

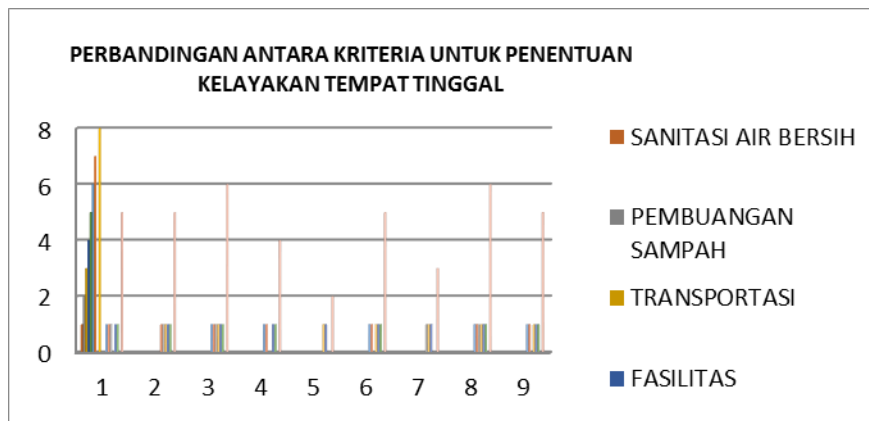


Gambar 3. Model Data Mining Untuk Penentuan Pemilihan Kelayakan Tempat Tinggal

Berdasarkan penentuan kriteria yang telah ditentukan, maka kriteria yang sangat mempunyai nilai kriteria tertinggi yaitu Keamanan, Bebas Banjir dan model rumah yang diminati oleh pemakai. Perbandingan antara variabel satu dengan variabel yang lain dapat dilihat dalam dalam Gambar 4 dan Gambar 5.



Gambar 4: Hasil Data Mining Berdasarkan Variabel



Gambar 5. Perbandingan Kriteria Untuk Penentuan Pemilihan Kelayakan Tempat Tinggal

4. KESIMPULAN

Permasalahan yang dihadapi oleh masyarakat umum diantaranya adalah dalam pemilihan penentuan kelayakan tempat tinggal. Terdapat beberapa kriteria yang dapat mempengaruhi kepada masyarakat umum dalam menentukan pemilihan kelayakan tempat tinggal, yaitu sanitasi air bersih, pembuangan sampah, transportasi, fasilitas, keamanan, bebas banjir dan model rumah. Proses penghitungan *Data mining* dengan menggunakan metode naive bayes, memiliki 3 kriteria yang paling dominan yaitu keamanan, bebas banjir dan harga tempat tinggal tersebut. *Data mining* ini juga dapat dimanfaatkan untuk memberikan penilaian kriteria yang digunakan untuk penentuan pemilihan kelayakan tempat tinggal tersebut. Pengolahan *Data mining* ini memberikan efek positif kepada masyarakat umum agar dapat diterapkan dalam kehidupan sehari-hari.

DAFTAR PUSTAKA

- [1] Agustin, Siti (2012) "Analisis penunjang keputusan pemilihan tempat tinggal dengan Metode Analytical Hierarchy Proses (AHP) secara online" Fakultas Teknik Industri.Gunadharma.
- [2] Efendi, Ahmad (2016) "Rancang bangun sistem pendukung keputusan pemilihan tempat kost dengan pemilihan Weight Product. Skripsi. Universitas Nusantara PGRI Kediri"
- [3] Keman, Soedjajadi. 2005. "Kesehatan perumahan dan lingkungan pemukiman". *Jurnal kesehatan lingkungan*, vol. 2, no. 1, juli 2005 : 29 -42. Universitas Airlangga.)
- [4] Nurrifai, Angga "Sistem Pendukung Keputusan pemilihan perumahan dengan Metode Fuzzy tahani"
- [5] Supriyana, I wayan (2012) "Sistem Pendukung Keputusan dalam Pemilihan Tempat Kost dengan Metode Pembobotan (Studi Kasus Sleman Yogyakarta).Jurnal ilmu computer Volume 5 - No 2 – September 2012.
- [6] Pressman RS, 2001, "*Software Engineering: a practitioners approach*", Mc Graw Hill Companies, Inc.