

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Beberapa penelitian yang telah dilakukan seperti Mohammad Taufan AZ, Surnayo dan Wijoyo (2014) dalam penelitian “Sistem Pendukung Keputusan untuk Investasi Perumahan Area Malang Menggunakan Algoritma Bayesian” melakukan pengklasifikasian rumah kedalam 4 klasifikasi yaitu mewah, sedang, sederhana dan minimalis menggunakan algoritma bayesian atau metode *Naïve Bayes Classifier*. Hasil dari klasifikasi tersebut akan dibobot untuk mendapatkan ranking yang mendekati kriteria persyaratan dalam pengambilan keputusan.

Tia Imandasari Eka Irawan, Agus Perdana Windarto dan Anjar Wanto (2019) melakukan penelitian “Algoritma *Naive Bayes* dalam Klasifikasi Lokasi Pembangunan Sumber Air” dengan memanfaatkan metode *Naïve Bayes Classifier* untuk membantu pihak PDAM Tirta Lihou mendapatkan alternatif lokasi sumber air bersih yang layak digunakan.

Paulus Harsadi, Dwi Remawati, Kriswantoro Budi Satrio (2020) dengan penelitian “Pemetaan Usaha Café di Manahan dengan Menggunakan Metode *Naïve Bayes* (Studi Kasus Dinas Perdagangan)” untuk mengklasifikasikan data café dengan hasil berupa presentasi nilai kemungkinan mikro, kecil atau menengah berdasarkan uji akurasi

Diana Laily Fithri (2016) dalam penelitian “Model Data Mining dalam Penentuan Kelayakan Pemilihan Tempat Tinggal Menggunakan Metode *Naïve Bayes*” dengan menggunakan model data *mining* untuk penentuan tempat tinggal menggunakan metode *naïve bayes* dengan membantu mempertimbangkan penentuan pencarian tempat tinggal yang menghasilkan probabilitas antara layak atau tidak dalam penentuan pencarian tempat tinggal.

Ajeng Widya Ayuningtyas (2020) dalam penelitian “Analisis Sentimen terhadap STMIK AKAKOM Yogyakarta Menggunakan Metode *Naïve Bayes* Berdasarkan Komentar Facebook dan Ulasan di Google Maps” yang membahas tentang penggunaan metode *Naïve Bayes* yang digunakan untuk mengklasifikasikan teks komentar ke dalam kelas positif, netral atau negatif sehingga dapat diketahui kecenderungan sentimen masyarakat terhadap institusi.

Perbedaan penelitian ini dengan kelima penelitian yang telah dijelaskan diatas terletak pada pemanfaatan *Naïve Bayes* dalam menyelesaikan beberapa masalah atau beberapa objek. Pada penelitian ini, metode *Naïve Bayes* digunakan untuk membantu dalam pemilihan perumahan sedangkan dalam beberapa penelitian diatas menggunakan *naïve bayes* untuk mengklasifikasi perumahan ke beberapa kelas, menentukan lokasi pembangunan sumber air, pengklasifikasian jenis café, dan mengklasifikasikan beberapa bentuk komentar dan ulasan ke beberapa kelas.

Tabel 2.1 Acuan Tinjauan Pustaka

Penulis	Objek	Metode	Hasil
Mohammad Taufan AZ, Surnayo Dan Wijoyo (2014)	Sistem Pendukung Keputusan Untuk Investasi Perumahan Area Malang Menggunakan Algoritma Bayesian.	<i>Naïve Bayes Classifier</i>	Pengklasifikasian rumah kedalam 4 klasifikasi mewah, sedang, sederhana dan minimalis. Kemudian hasil klasifikasi tersebut dibobot untuk mendapatkan ranking yang mendekati kriteria.
Tia Imandasari, Eka Irawan, Agus Perdana Windarto, Anjar Wanto (2019)	Algoritma Naive Bayes Dalam Klasifikasi Lokasi Pembangunan Sumber Air	<i>Naïve Bayes Classifier</i>	Menggunakan metode <i>Naïve Bayes Classifier</i> pihak pdam tirta lihou berencana mencari alternatif solusi sumber mata air untuk beberapa unit produksi sehingga dapat memenuhi kebutuhan masyarakat.
Paulus Harsadi, Dwi Remawati, Kriswantoro Budi Satrio (2020)	Pemetaan Usaha Cafe Di Manahan Dengan Menggunakan Metode Naive Bayes Studi Kasus Dinas Perdagangan	<i>Naïve Bayes Classifier</i>	Mengklasifikasikan data cafe baru dengan hasil akhir berupa prosentasi nilai kemungkinan mikro, kecil, atau menengah, berdasarkan uji akurasi penggunaan metode <i>Naïve Bayes</i> sebesar 80%.

Lanjutan Tabel 2.1

Penulis	Objek	Metode	Hasil
Diana Laily Fithri (2016)	Model Data <i>Mining</i> Dalam Penentuan Kelayakan Pemilihan Tempat Tinggal Menggunakan Metode <i>Naive Bayes</i>	<i>Naive Bayes</i> <i>Classifier</i>	Menghasilkan probabilitas antara layak (sesuai) atau tidak sesuai, yang nantinya dapat digunakan oleh masyarakat umum untuk mempermudah dalam pencarian tempat tinggal
Ajeng Widya Ayuningtyas (2020)	Analisis Sentimen Terhadap STMIK AKAKOM Yogyakarta Menggunakan Metode <i>Naive Bayes</i> Berdasarkan Komentar <i>Facebook</i> Dan Ulasan Di <i>Google Maps</i>	<i>Naive Bayes</i> <i>Classifier</i>	Penelitian ini membahas tentang metode <i>Naive</i> <i>Bayes</i> yang digunakan untuk mengklasifikasikan teks komentar ke dalam kelas positif, netral atau negatif sehingga dapat diketahui kecenderungan sentimen masyarakat terhadap institusi.

2.2 Dasar Teori

2.2.1 Sistem Pendukung Keputusan

Sistem pendukung keputusan (*Decision Support System*) mulai dikembangkan pada tahun 1960-an, tetapi istilah sistem pendukung keputusan itu sendiri baru muncul pada tahun 1971, yang diciptakan oleh G. Anthony Gorry dan Michael S. Scott Morton. Sistem pendukung keputusan didefinisikan sebagai sistem yang digunakan untuk mendukung dan membantu pihak manajemen melakukan pengambilan keputusan pada semi terstruktur dan tidak terstruktur.

Pada dasarnya konsep sistem pendukung keputusan hanyalah sebatas pada kegiatan membantu para manager melakukan penelitian serta menggantikan posisi serta peran manajer.

Sistem pendukung keputusan merupakan sistem informasi interaktif yang menyediakan informasi, pemodelan dan manipulasi data. Sistem ini digunakan untuk membantu mengambil keputusan dalam situasi yang semi terstruktur dan tidak terstruktur dimana tak seorangpun tahu secara pasti bagaimana keputusan seharusnya dibuat. Tahap ini merupakan tahap pelaksanaan dari keputusan yang telah diambil. Pada tahap ini perlu disusun tindakan yang terencana, sehingga hasil keputusan dapat dipantau dan disesuaikan apabila diperlukan perbaikan.

2.2.2 Naïve Bayes Classifier

Naïve Bayes merupakan suatu algoritma yang dapat mengklasifikasikan suatu variabel tertentu dengan menggunakan metode probabilitas dan statistik. Klasifikasi-klasifikasi Bayes dapat memprediksi kelas anggota probabilitas.

Klasifikasi Bayes berdasarkan atas teorema Bayes. *Naive Bayes* merupakan sebuah pengklasifikasian probalistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengansumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. *Naive Bayes* juga didefinisikan sebagai pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya.

Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu.

Keuntungan penggunaan *Naive Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. *Naive Bayes* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan. Langkah-langkah metode *Naive Bayes* sebagai berikut.

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(H)}$$

Dimana:

X : data dengan class yang belum diketahui.

C : hipotesis data menggunakan suatu class spesifik

$P(C|X)$: probabilitas hipotesis H berdasar kondisi X (partiori probabilitas)

$P(C)$: probabilitas hipotesis H (prior probabilitas)

$P(X|C)$: probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: probabilitas

Penerapan metode *naïve bayes* ini dapat dilihat pada kasus berikut, dimana terdapat beberapa data latih dan sebuah data uji. Berikut contoh kasus yang akan dihitung menggunakan metode *naïve bayes*.

Tabel 2.2 Data Latih Perumahan

No	Nama	Luas Rumah	Luas Tanah	Jarak Dari Pusat Kota	Tipe Rumah	Harga	Lokasi	Pilih Perumahan
1	Syafi Arkam	21	50-60 m ²	0-10 km	Tipe 21	300-500 juta	Pusat Kota	Tidak
2	Leni Syamsiah	36	60-72 m ²	0-10 km	Tipe 36	600-900 juta	Pusat Kota	Tidak
3	Tutri Wulan	21	50-60 m ²	11-20 km	Tipe 21	100-300 juta	Pinggir Kota	Iya
4	M. Irfan	54	120-150 m ²	0-10 km	Tipe 54	600-900 juta	Pusat Kota	Tidak
5	M. Sahril	60	120-150 m ²	0-10 km	Tipe 60	> 1000 juta	Pusat Kota	Iya
6	Rahmat	120	>150 m ²	0-10 km	Tipe 120	> 2000 juta	Pusat Kota	Iya
7	Yusnidar	36	60-72 m ²	11-20 km	Tipe 36	300-500 juta	Pinggir Kota	Iya
8	Nilam Sari	60	120-150 m ²	11-20 km	Tipe 60	> 1000 juta	Pinggir Kota	Tidak
9	Zulfahmi	21	50-60 m ²	11-20 km	Tipe 21	300-500 juta	Pinggir Kota	Tidak
10	Fauziah	45	72 - 90 m ²	21-30 km	Tipe 45	600-900 juta	Pinggir Kota	Iya
11	Ade Gunawan	60	120-150 m ²	11-20 km	Tipe 60	> 1000 juta	Pusat Kota	Iya

No	Nama	Luas Rumah	Luas Tanah	Jarak Dari Pusat Kota	Tipe Rumah	Harga	Lokasi	Pilih Perumahan
12	Irwanto	54	120-150 m ²	0-10 km	Tipe 54	600-900 juta	Pusat Kota	Iya
13	Rizwan Hadi	36	60-72 m ²	11-20 km	Tipe 36	300-500 juta	Pinggir Kota	Tidak
14	Hidayatullah	21	50-60 m ²	0-10 km	Tipe 21	300-500 juta	Pusat Kota	Iya
15	Wilsa Rizki	36	60-72 m ²	0-10 km	Tipe 36	300-500 juta	Pusat Kota	?

Lanjutan Tabel 2.2 Data Latih Perumahan

Untuk data latih diatas (baris 15), maka :

1. Menghitung nilai probabilitas hipotesis $P(C_i)$.

$$P(\text{pilih_perumahan} = \text{"Iya"}) = 8/14 = 0,57$$

$$P(\text{pilih_perumahan} = \text{"Tidak"}) = 6/14 = 0,43$$

2. Menghitung nilai probabilitas berdasarkan kondisi pada hipotesis

$$P(X|C_i) \quad P(\text{luas_rumah} = \text{"36"} | \text{pilih_perumahan} = \text{"Iya"}) = 1/8 = 0,125$$

$$P(\text{luas_rumah} = \text{"36"} | \text{pilih_perumahan} = \text{"Tidak"}) = 2/6 = 0,333$$

$$P(\text{luas_tanah} = \text{"60-72"} | \text{pilih_perumahan} = \text{"Iya"}) = 1/8 = 0,125$$

$$P(\text{luas_tanah} = \text{"60-72"} | \text{pilih_perumahan} = \text{Tidak}) = 2/6 = 0,333$$

$$P(\text{jarak_dari_kota} = \text{"0-10"} | \text{pilih_perumahan} = \text{"Iya"}) = 4/8 = 0,5$$

$$P(\text{jarak_dari_kota} = \text{"0-10"} | \text{pilih_perumahan} = \text{Tidak}) = 3/6 = 0,5$$

$$P(\text{tipe_rumah} = \text{"Tipe 36"} | \text{pilih_perumahan} = \text{"Iya"}) = 1/8 = 0,125$$

$$P(\text{tipe_rumah} = \text{"Tipe 36"} | \text{pilih_perumahan} = \text{Tidak}) = 2/6 = 0,333$$

$$P(\text{harga} = \text{"300-500 juta"} | \text{pilih_perumahan} = \text{"Iya"}) = 2/8 = 0,25$$

$$P(\text{harga} = \text{"300-500 juta"} | \text{pilih_perumahan} = \text{Tidak}) = 3/6 = 0,5$$

$$P(\text{lokasi} = \text{"Pusat Kota"} \mid \text{pilih_perumahan} = \text{"Iya"}) = 5/8 = 0,625$$

$$P(\text{lokasi} = \text{"Pusat Kota"} \mid \text{pilih_perumahan} = \text{"Tidak"}) = 3/6 = 0,5$$

Sehingga,

$$\begin{aligned} P(X \mid \text{pilih_perumahan} = \text{"Iya"}) &= 0,125 \times 0,125 \times 0,50 \times 0,125 \times 0,25 \times 0,62 \\ &= \mathbf{0,0001526} \end{aligned}$$

$$\begin{aligned} P(X \mid \text{pilih_perumahan} = \text{"Tidak"}) &= 0,33 \times 0,33 \times 0,50 \times 0,33 \times 0,50 \times 0,50 \\ &= \mathbf{0,0046158} \end{aligned}$$

3. Menghitung nilai $P(X|C) \times P(C)$

$$\begin{aligned} P(X \mid \text{pilih_perumahan} = \text{"Iya"}) \times P(\text{pilih_perumahan} = \text{"Iya"}) \\ &= 0,0001526 \times 0,57 = 0,00008698 \end{aligned}$$

$$\begin{aligned} P(X \mid \text{pilih_perumahan} = \text{"Tidak"}) \times P(\text{pilih_perumahan} = \text{"Tidak"}) \\ &= 0,0046158 \times 0,43 = \mathbf{0,00198479} \end{aligned}$$

Jadi untuk pemilihan rumah dengan kriteria Luas Rumah = 36, Luas Tanah = 60-70, Jarak dari kota = 0-10km, Tipe = Tipe 36, Harga = 300-500 juta dan lokasi = pusat kota lebih besar peluangnya untuk **tidak dipilih**.

2.2.3 Website

Website adalah sekumpulan halaman yang terdiri dari beberapa halaman yang berisi informasi dalam bentuk data digital baik berupa text, gambar, video, audio, dan animasi lainnya yang disediakan melalui jalur koneksi internet". Berdasarkan uraian, penulis menyimpulkan bahwa *web* adalah Sebuah *software* yang berfungsi untuk menampilkan dokumen - dokumen pada suatu *web* yang membuat pengguna dapat mengakses.

WWW (*World Wide Web*) atau biasa disebut dengan *web* merupakan salah satu sumber daya internet yang berkembang pesat. Saat ini, informasi *web* didistribusikan melalui pendekatan *hyperlink*, yang memungkinkan suatu teks, gambar, ataupun objek yang lain menjadi acuan untuk membuka halaman-halaman *web* yang lain. Dengan pendekatan *hyperlink* ini, seseorang dapat memperoleh informasi dengan meloncat dari suatu halaman ke halaman lain (Susanto, 2013)

2.2.4 PHP

PHP merupakan sebuah bahasa pemrograman yang berguna untuk menciptakan sebuah *website* yang berfungsi dengan baik antar berbagai dokumen dan halaman – halaman yang ada dan saling berhubungan satu sama lain. Teori mengenai PHP dijelaskan bahwa PHP adalah *server-side language* dimana memungkinkan untuk menyisipkan Script ke dalam kode HTML (James Clint Lumanau, Dali S Naga, 2018).

PHP Pertama kali ditemukan pada 1995 oleh seorang Software Developer bernama Rasmus Lerdorf. Ide awal PHP adalah ketika itu Radmus ingin mengetahui jumlah pengunjung yang membaca resume online-nya. Script yang dikembangkan baru dapat melakukan dua pekerjaan, yakni merekam informasi pengunjung, dan menampilkan jumlah pengunjung dari suatu *website*. Dan sampai sekarang kedua tugas tersebut masih tetap populer digunakan oleh dunia *web* saat ini. Kemudian, dari situ banyak orang mendiskusikan script buatan Rasmus Lerdorf, hingga akhirnya Rasmus mulai membuat sebuah *tool/script*, bernama Personal Home Page (PHP) (Marbun, dkk 2016). Kode PHP mempunyai ciri khusus yaitu :

- a. Hanya dapat dijalankan menggunakan *web server*, misalnya *Apache*.
- b. Kode PHP diletakkan dan dijalankan di *web server*.

Kode PHP dapat digunakan untuk mengakses database, seperti: MySQL, PostgreSQL, Oracle, dan lain-lain.

2.2.5 MySQL

MySQL adalah sebuah perangkat yang berperan sebagai server database, yang selanjutnya akan digunakan untuk mempelajari kode-kode PHP yang berkaitan atau membutuhkan akses ke server database (Hendrawati & Fatkhudin, 2015). MySQL merupakan software yang tergolong sebagai DBMS (*Database Management System*) yang bersifat *Open source*. *Open source* menyatakan bahwa software ini dilengkapi dengan *source code* (kode yang dipakai untuk membuat MySQL), selain tentu saja bentuk *executable*-nya atau kode yang dapat dijalankan secara langsung dalam sistem operasi, dan dapat diperoleh dengan cara *download* (mengunduh) di Internet secara gratis. Basisdata adalah suatu tools yang digunakan untuk menyimpan informasi, mengambil informasi kapanpun dibutuhkan, dan mengatur informasi yang tersimpan. Jika digambarkan Lemari File merupakan suatu basis data. Basis data merupakan suatu kumpulan data yang berhubungan secara logis dan deskripsi data tersebut, yang dirancang untuk memenuhi informasi yang dibutuhkan oleh suatu organisasi. Artinya, basisdata merupakan tempat penyimpanan data yang besar, dimana dapat digunakan oleh banyak pengguna (Marbun, dkk 2016).