
**SISTEM APLIKASI UNTUK MENENTUKAN KELAYAKAN AREA
DISTRIBUSI MAJALAH DENGAN METODE NAIVE BAYES
CLASSIFIER PADA PT. MEDIA KOMUNITAS SERPONG**

***APPLICATION SYSTEM FOR DETERMINING THE FEASIBILITY OF
THE MAGAZINE DISTRIBUTION AREA WITH THE NAIVE BAYES
CLASSIFIER METHOD IN SERPONG COMMUNITY MEDIA INC.***

Nurfiqih¹, Shandi Noris²

^{1,2}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pamulang
Jl. Surya Kencana No. 1, Pamulang, Tangerang Selatan-Indonesia
E-mail : ²dosen00354@unpam.ac.id

ABSTRAK

Data mining merupakan teknologi yang sangat berguna untuk membantu perusahaan untuk menemukan informasi yang sangat penting dari gudang data mereka, data mining meramalkan tren dan sifat-sifat perilaku bisnis yang sangat berguna untuk mendukung pengambilan keputusan penting. Data mining dapat menjawab pertanyaan-pertanyaan bisnis yang dengan cara tradisional memerlukan banyak waktu untuk menjawabnya. Algoritma *Naive Bayes Classifier* adalah salah satu bagian dari *data mining* yang memakai teknik prediksi berbasis probabilistik sederhana yang berdasar pada aturan bayes, atau juga pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu *class*.

Penentuan keputusan yang dilakukan PT. Media Komunitas Serpong sebagai upaya untuk meningkatkan area baru diperlukan keputusan yang perlu diperhatikan penuh, untuk itu metode *Naive Bayes Classifier* adalah cara yang lebih efektif dilakukan untuk menentukan kelayakan area baru secara cepat dan akurat.

Kata kunci : Area Distribusi, *Data Mining*, Penerapan metode *Naive Bayes Classifier*

ABSTRACT

Data mining is a very useful technology to help companies to find the most important information from their data warehouse, data mining forecast trends and traits business behavior is very useful to support critical decision making. Data mining can answer business questions that the traditional way takes a lot of time to answer. Naive Bayes classifier algorithm is one part of the data mining technique that uses a simple probabilistic-based prediction based on the Bayes rule, or statistical classification also can be used to predict the probability of membership of a class.

Determination of decisions made by PT. Media Komunitas Serpong in an effort to improve the new area necessary decisions that need to be filled, for the Naive Bayes classifier method is a more effective way is done to determine the feasibility of new areas quickly and accurately.

Keywords : Area Distribution, *Data Mining*, Application of *Naive Bayes Classifier* method

1. PENDAHULUAN

Dalam dunia bisnis yang selalu dinamis dan penuh persaingan, para pelakunya harus senantiasa memikirkan cara-cara untuk terus survive dan jika mungkin mengembangkan skala bisnis mereka. Untuk mencapai hal itu, dapat diringkaskan empat kebutuhan bisnis, yaitu penambahan jenis maupun peningkatan kapasitas produk, pengurangan biaya operasi perusahaan, dan peningkatan efektifitas pemasaran serta memperoleh keuntungan. Untuk memenuhi kebutuhan-kebutuhan diatas, banyak cara yang dapat ditempuh. Salah satunya adalah dengan melakukan analisis data perusahaan baik data internal maupun data eksternal (Enur Irdiansyah,2010).

Setiap organisasi baik organisasi profit maupun organisasi non profit mempunyai standar operasional prosedur di mana setiap kegiatan operasinya selalu dicatat dan didokumentasikan. Data-data tersimpan atau didokumentasikan dalam basis data. Dokumentasi data operasional dalam basis data tersebut, merupakan salah satu faktor yang dibutuhkan oleh perusahaan yang bergerak dalam bidang percetakan dan pendistribusian majalah(Sumber berdasarkan wawancara dengan Syarief Hidayat, Spv Sirkulasi dan Distribusi PT. Media Komunitas Serpong,2014). Data operasional yang terdokumentasi bisa digunakan untuk proses penentuan area distribusi baru. Hal ini dikarenakan para pengambilan keputusan (decision maker) bisa mengambil keputusan dengan benar.

Masalah yang dihadapi perusahaan adalah menentukan kelayakan area distribusi yang digunakan pada saat ini, masih dianggap kurang begitu efektif, efektif dalam hal keakuratan dalam menghitung kelayakan area baru dan efisien dalam hal cepat mengambil keputusan dan tidak memakan waktu lama. Pada Penelitian ini digunakan Metode Naïve Bayes untuk penyelesaian permasalahan yang ada.

Naive Bayes bekerja berdasarkan teori probabilitas yang memandang semua fitur dari data sebagai bukti probabilitas dan merupakan salah satu bagian dari Data Mining, Metode Naive Bayes teguh (robust) terhadap data-data yang terisolasi yang biasanya merupakan data dengan karakteristik berbeda (outlier). Naive Bayes juga bisa menangani nilai atribut yang salah dengan mengabaikan data latih selama proses pembangunan model dan prediksi, Tangguh menghadapi atribut yang tidak relevan. Serta Atribut yang mempunyai korelasi bisa mendegradasi kinerja klasifikasi Naive Bayes, karena asumsi independensi atribut tersebut sudah tidak ada(Eko Prasetyo,2012).

Tujuan Penelitian ini Agar mempermudah dalam penentuan kelayakan area distribusi secara efektif dalam hal keakuratan perhitungan dan efisien dalam hal kecepatan dalam menentukan keputusan.

2. PERANCANGAN

2.1 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data yang didapatkan dari hasil survei kurir riset pemasaran. Data ini merupakan data pemilihan area kelayakan distribusi majalah, penelitian ini cenderung memilih area yang didalamnya ditempati oleh kalangan menengah ke atas. Data tersebut terdiri dari enam kelas diantaranya adalah:

C1 : Strategis

C2 : Biaya izin

C3 : Hunian

C4 : Minat Pembaca

C5 : Layak Distribusi

Tahap awal cara kerja dari proses perhitungan *Naive Bayes* adalah dengan melakukan pengambilan data distribusi dari data perumahan dibebberapa cluster. Adapun variabel penentu yang digunakan dalam mengklasifikasikan data cluster yaitu :

1. Strategis

Dimana posisi cluster yang ada dekat dengan jalan utama yang dikelompokkan dalam dua kategori yaitu ya dan tidak.

2. Biaya Izin

Merupakan variabel biaya izin sebagai syarat masuknya majalah kedalam cluster yang dikelompokkan dalam dua kategori yaitu ya dan tidak.

3. Hunian

variabel hunian tipe 45 ke atas, sasaran utama distribusi

4. Tingkat Pembaca

Merupakan variabel tingkat pembaca yang dikelompokkan dalam tiga kategori yaitu tinggi, menengah dan rendah, berdasarkan presentase hunian A & B tingkat pembaca $\geq 60\%$ dari total hunian A & B dikatakan tinggi, $>40\% - <60\%$ dikatakan menengah, $\leq 40\%$ dikatakan rendah.

Tabel 2.1 Data Cluster

(Sumber : Divisi Sirkulasi Distribusi, PT Media Komunitas Serpong)

Data Ke-	Cluster	Strategis	Biaya Izin	Hunian	Minat Pembaca	Layak Distribusi
1	The Green	Ya	Ya	713	Tinggi	Layak
2	De Latinos	Ya	Ya	170	Tinggi	Layak
3	Ingenia	Tidak	Tidak	89	Rendah	Tidak Layak
4	Giri Loka 3	Ya	Tidak	165	Tinggi	Layak
5	Neo Catalina	Tidak	Ya	7	Rendah	Tidak Layak
6	Pavilion Residence	Tidak	Ya	108	Menengah	Layak
7	Taman Provence	Tidak	Tidak	70	Tinggi	Layak
8	Simplicity	Ya	Tidak	0	Rendah	Tidak Layak
9	Teras Bukit Golf	Ya	Tidak	35	Tinggi	Layak
10	Victorian River Park	Ya	Tidak	153	Menengah	Layak
11	Taman Tirta Golf	Tidak	Tidak	44	Tinggi	Layak
12	The Avani	Tidak	Tidak	456	Rendah	Tidak Layak
13	Foresta	Tidak	Ya	875	Tinggi	Layak
14	The Icon	Tidak	Tidak	534	Menengah	Layak
15	Green Cove	Ya	Tidak	288	Tinggi	Layak
16	Castilla	Tidak	Ya	453	Rendah	Tidak Layak
17	Kencana Loka	Ya	Tidak	325	Tinggi	Layak
18	British	Tidak	Ya	182	Menengah	Layak
19	Serenade Lake	Tidak	Ya	30	Tinggi	Layak
20	Anggrek Loka	Tidak	Tidak	50	Tinggi	Layak
21	Anggrek Loka I (Anggrek Residence) 2 -3	Ya	Tidak	135	Menengah	Layak
22	Taman Giri Loka	Tidak	Ya	280	Tinggi	Layak
23	Fontainbleau	Ya	Ya	0	Tinggi	Layak
24	Versailles	Tidak	Tidak	100	Tinggi	Layak
25	Golden Vienna 1	Tidak	Ya	86	Tinggi	Layak
26	Prestigia	Ya	Ya	45	Tinggi	Tidak Layak

27	Foglio	Ya	Tidak	671	Tinggi	Tidak Layak
28	Puspita Loka	Ya	Tidak	673	Tinggi	Layak
29	Taman Chrysan 1	Tidak	Tidak	174	Tinggi	Layak
30	Illustria	Tidak	Tidak	78	Tinggi	Tidak Layak
31	Nusa Loka	Ya	Ya	245	Rendah	Layak
32	Griya Loka	Ya	Ya	323	Rendah	Layak
33	Sheffield Greenwich	Ya	Tidak	98	Rendah	Tidak Layak
34	Magnolia	Tidak	Tidak	98	Rendah	Tidak Layak
35	Edelwais	Ya	Ya	22	Tinggi	Layak
36	Giri Loka I & II	Ya	Tidak	300	Tinggi	Layak
37	Vermount Parkland	Tidak	Ya	0	Rendah	Tidak Layak
38	Emerald Cove	Tidak	Tidak	34	Rendah	Tidak Layak
39	Golden Vienna 2	Ya	Tidak	80	Tinggi	Layak
40	Taman Chrysan 2	Ya	Tidak	0	Rendah	Layak
41	Provence Parkland	Tidak	Tidak	110	Tinggi	Layak
42	Oleaster	Ya	Ya	0	Tinggi	Layak
43	Catallina	Ya	Tidak	940	Rendah	Tidak Layak
44	The Spring	Tidak	Tidak	234	Tinggi	Tidak Layak
45	Volta	Tidak	Ya	56	Rendah	Tidak Layak
46	Venetian	Tidak	Tidak	0	Rendah	Tidak Layak
47	The Eminent	Tidak	Tidak	123	Rendah	Tidak Layak
48	Naturale	Ya	Ya	890	Rendah	Tidak Layak
49	De Prangipani	Ya	Ya	567	Rendah	Tidak Layak
50	Savila	Ya	Tidak	450	Tinggi	Layak

a. Rata-rata layak distribusi untuk hunian Tipe A & B

$$\bar{x}_{\text{Layak}} = \frac{6720}{31}$$

$$\bar{x}_{\text{Layak}} = 216.774$$

$$s_{\text{Layak}}^2 = \frac{1499843.417}{30}$$

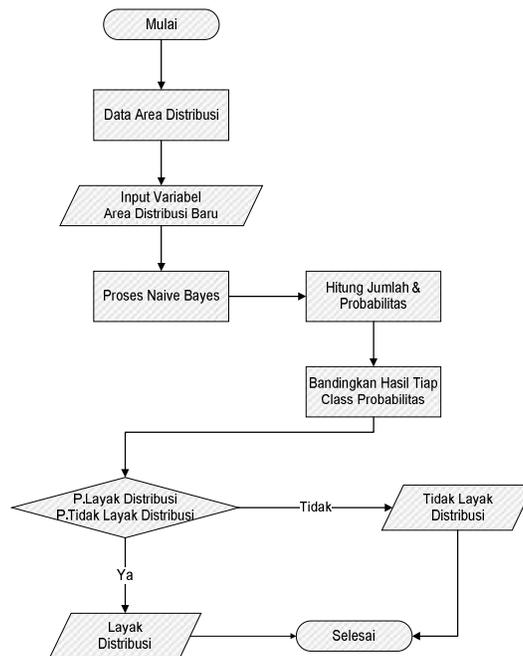
$$s_{Layak}^2 = 49994.780$$
$$s_{Layak} = \sqrt{49994.780} = 223.595$$

b. Rata-rata tidak layak distribusi untuk hunian Tipe A & B

$$\bar{x}_{Tidak Layak} = \frac{4.839}{19}$$
$$\bar{x}_{Tidak Layak} = 254.684$$
$$s_{Tidak Layak}^2 = \frac{1737622,1}{18}$$
$$s_{Tidak Layak}^2 = 96534.561$$
$$s_{Tidak Layak} = \sqrt{96534,561} = 310.700$$

2.2 Flowchart

Perancangan flowchart ini bertujuan untuk memberi gambaran bagaimana aliran proses dari aplikasi untuk prediksi minat studi ini berjalan. Mulai dari awal ketika task menu prediksi kelayakan area distribusi hingga tampil hasil prediksi kelayakan area distribusi.

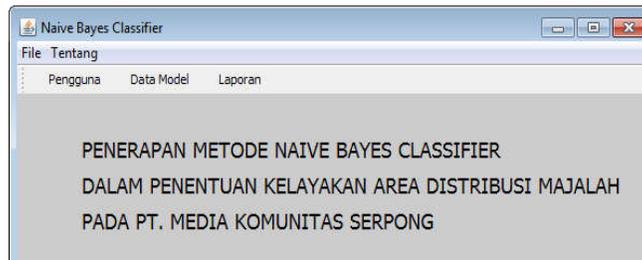


Gambar 2.1 Flowchart Sistem

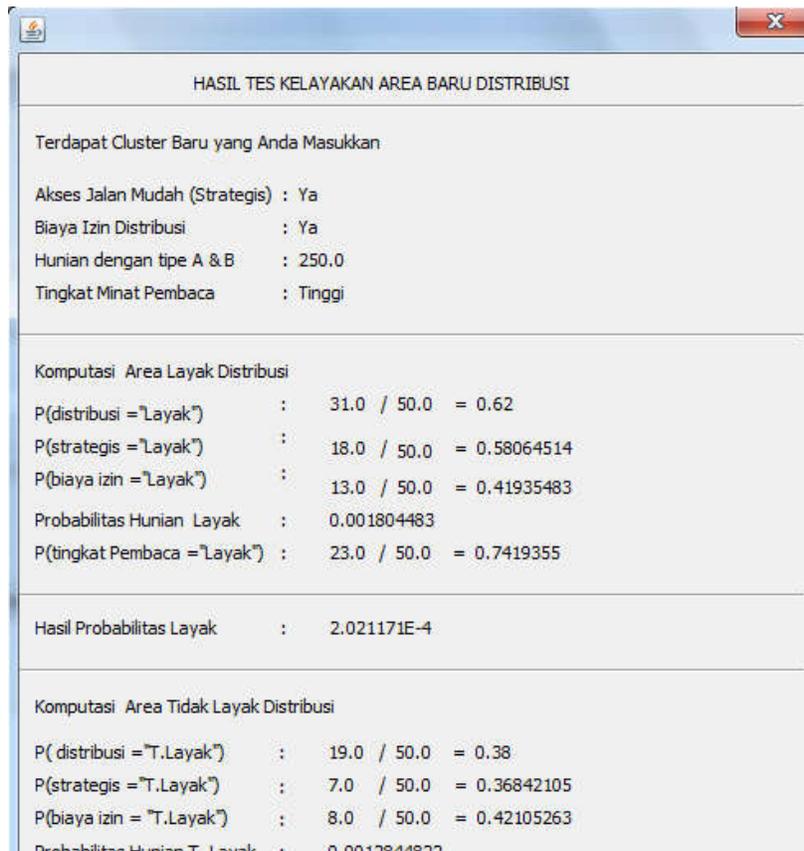
3. PEMBAHASAN

3.1 Implementasi

Teknologi yang digunakan dalam pengembangan sistem ini adalah teknologi aplikasi berbasis *desktop*, yang membentuk sebuah program yang dapat berdiri sendiri dan dapat dijalankan dalam PC sendiri. Dengan sistem *desktop based* ini, diharapkan dapat meningkatkan pengetahuan serta memberikan suatu kemudahan dalam hal menentukan kelayakan area baru di wilayah yang menyangkup kawasan distribusi.



Gambar 3.1 Form Menu Utama



Gambar 3.2 Detail Hasil Uji coba *Procedure Naive Bayes Classification*

3.2 Pengujian

$$\text{Akurasi} = \frac{\text{Jumlah data yang diprediksi secara benar}}{\text{Jumlah prediksi yang dilakukan}} \times 100 \%$$

$$\text{Akurasi} = \frac{48}{50} \times 100 \%$$

$$\text{Akurasi} = 96 \%$$

$$\text{Error} = \frac{\text{Jumlah data yang diprediksi secara salah}}{\text{Jumlah prediksi yang dilakukan}} \times 100 \%$$

$$\text{Error} = \frac{2}{50} \times 100 \%$$

$$\text{Error} = 4 \%$$

4. PENUTUP

4.1 Kesimpulan

Berdasarkan hasil pembahasan maka dapat mengambil kesimpulan bahwa Penerapan metode *Naive Bayes Classifier* sangat mudah di implementasikan dalam dunia bisnis, khususnya data dari PT. Media Komunitas Serpong yaitu data distribusi majalah, Algoritma *Naive Bayes* di dukung oleh ilmu Probabilistik dan ilmu statistika khususnya dalam penggunaan data petunjuk untuk mendukung keputusan pengklasifikasian. Pada algoritma *Naive Bayes*, semua atribut akan memberikan kontribusinya dalam pengambilan keputusan, dengan bobot atribut yang sama penting dan setiap atribut saling bebas satu sama lain. atribut penentu klasifikasi bersifat tetap yang dipakai dari hasil survey lapangan, dari kinerja yang dilakukan naive bayes didapat nilai akurasi sebesar 96%. Dimana pada data yang digunakan untuk mengklasifikasikan area distribusi tersebut terdapat atribut dengan data kontinu dibandingkan data diskrit yang hal ini berpengaruh pada perhitungan kinerja sistemnya.

Berdasarkan percobaan yang dilakukan dengan mendiskritkan data yang ada, dihasilkan bahwa kinerja dari *Naive Bayes Classifier* adalah diatas 50 %. Ini berarti semakin banyak data diskrit dibandingkan data kontinu, maka kinerja *Naive Bayes* semakin baik karena Laju error sedikit hanya 4%.

4.2 Saran

Berdasarkan kesimpulan yang diperoleh dalam penelitian ini, maka diajukan saran-saran sebagai pelengkap terhadap hasil penelitian sebagai berikut :

1. Dapat dikembangkan dengan penambahan atribut baru yang diperlukan untuk penentuan area distribusi kedepannya.
2. Dapat menggunakan metode klasifikasi yang lain sehingga dapat dikomparasi metode yang paling baik.

5. DAFTAR PUSTAKA

- 1) Gata, Windu., dan Grace Gata. 2013. *Sukses Membangun Aplikasi Penjualan dengan Java*. Jakarta: PT. Elex Media Komputindo
- 2) Prasetyo, Eko. 2012. *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: CV. ANDI OFFSET
- 3) Kusri, dan Emha Taufiq Luthfi. 2009. *Algoritma Data Mining*. Yogyakarta: CV. ANDI OFFSET
- 4) Bisri, Achmad. 2010. *Modul Praktikum Database SQL Server 2005*. Tangerang: Universitas Pamulang
- 5) Dr. Ir Harinaldi, M.Eng. 2005. *Prinsip-prinsip statistik untuk teknik dan sains*. Jakarta: Erlangga
- 6) Bustami. 2013. *Penerapan Algoritma Naive Bayes Untuk Mengklasifikasikan Data Nasabah Asuransi*. Skripsi. Releut: Universitas Malikussaleh
- 7) Alfironi Burhan M. 2013. *Implementasi Data Mining Dengan Naive Bayes Classifier Untuk Mendukung Strategi Pemasaran di Bagian*
- 8) *Humah STMIK Amikom Yogyakarta*. Skripsi. Yogyakarta: Amikom Yogyakarta
- 9) Rifqi, Naufar dkk. 2011. *Analisis dan Implementasi Klasifikasi Data Mining Menggunakan Jaringan Syaraf Tiruan Dan Evolution Strategis*. Skripsi. Bandung: IT Telkom
- 10) Yakub, Suardin. 2008. *Sistem Pakar Deteksi Penyakit Diabetes Mellitus Dengan Menggunakan Pendekatan Naive Bayesian Berbasis Web*. Skripsi. Malang: UIN
- 11) Nyura, Yusni. 2010. *Pembuatan Aplikasi Pembelajaran Bahasa Inggris Pada Handphone dengan J2ME*. Skripsi. Samarinda: Politeknik Negeri Samarinda
- 12) Nufal, Milad A. 2011. *Pengaruh Produk Persepsi Harga Promosi Dan Distribusi Terhadap Keputusan Pembelian Mebel Pada PT. Nadira Prima Di Semarang*. Skripsi. Semarang: Universitas Diponegoro

- 13) Hamakonda, Towa P dan J.N.B Tairas.2008.*Pengantar Klasifikasi Persepuluhan Dewey*.Edisi 5.Jakarta:Gunung Mulia
- 14) Lee, Finn S & Juan Santana.2010.*Data Mining: Meramalkan Bisnis Perusahaan*.Jakarta:PT. Elex Media Komputindo
- 15) Irdiansyah, Enur.2010.*Penerapan Data Mining Pada Penjualan Produk Minuman di PT. Pepsi Cola Indobeverages Menggunakan Metode Clustering*.Jurnal.Bandung:Unikom