

# KLASIFIKASI DATA PELANGGAN POTENSIAL *PUT IN SERVICE* INDIHOME MENGGUNAKAN ALGORITMA NAÏVE BAYES

Muhammad Fauzan Al-Ghifary Dwi Wibowo<sup>1\*)</sup>, Reni Rahmadewi<sup>2</sup>

<sup>1</sup>Jurusan Teknik Elektro, Fakultas Teknik, Universitas Singaperbangsa Karawang, Karawang

<sup>2</sup>Jln. HS Ronggo Waluyo, Kabupaten Karawang, Jawa Barat 41361, Indonesia

email: <sup>1</sup>[muhammad.fauzan18171@student.unsika.ac.id](mailto:muhammad.fauzan18171@student.unsika.ac.id), <sup>2</sup>[reni.rahmadewi@ft.unsika.ac.id](mailto:reni.rahmadewi@ft.unsika.ac.id)

**Abstract** – In the digital era like now, the invitation to get a job is very tight, companies are also increasingly demanding prospective employees to have work experience. This experience allows them to know and understand the atmosphere and apply the knowledge gained during college to the world of work. Knowing and learning how to classify customer data, an opportunity for students to gain experience. Data is an observation result in the form of characteristics of the representation that represents an object. While the understanding of information is the result of structured input processing [1]. Classification is an information mining functionality that wants to create a model to predict part of the objects in the information line [2]. Naïve Bayes algorithm is a grouping of data that is needed when predicting the possibility of a class. The customer classification of test data and training data required for grouping is 25 test data and 75 training data, where training information is then processed using the naïve Bayes procedure [6]. After the grouping process using the naïve Bayes procedure ended, after that the test information experiment amounted to 25 customer information [7]. Suggestions for further research, the method used needs to be redeveloped using other data mining methods, and it is better if the data used needs to be integrated with the database in order to avoid duplication or damage to data and data security is better maintained.

**Keywords** – Naïve Bayes, Data Mining, Training Data, Database, Classification

**Abstrak** – Di era digital seperti sekarang invitasi untuk memperoleh lapangan pekerjaan sangatlah ketat, perusahaan juga semakin menuntut calon karyawannya untuk memiliki pengalaman bekerja. Pengalamannya ini membuat mereka dapat mengetahui dan memahami suasana serta menerapkan ilmu yang diperoleh pada saat berkuliah pada dunia kerja. Mengetahui dan mempelajari bagaimana proses klasifikasi data pelanggan, kesempatan bagi mahasiswa untuk memperoleh pengalaman. Data adalah sebuah hasil observasi yang berupa karakteristik dari representasi yang mewakili sebuah objek. Sedangkan pengertian informasi adalah hasil dari pemrosesan input yang terstruktur [1]. Klasifikasi merupakan sebuah fungsionalitas informasi mining yang hendak menciptakan model guna memprediksi bagian dari objek- objek pada baris informasi [2]. Algoritma Naïve Bayes merupakan pengelompokan data yang diperlukan pada saat memprediksi kemungkinan suatu class. Klasifikasi pelanggan data uji serta data latih yang diperlukan untuk pengelompokan yang berjumlah 25 data uji serta 75 data latih, dimana informasi latih setelah itu di proses dengan memakai prosedur naïve bayes [6]. Setelah proses pengelompokan yang memakai prosedur naïve bayes berakhir, setelah itu percobaan informasi uji yang berjumlah 25 informasi pelanggan [7]. Saran untuk penelitian selanjutnya metode yang digunakan perlu dikembangkan kembali dengan menggunakan metode data mining lainnya, dan sebaiknya data-data yang

digunakan perlu diintegrasikan dengan database agar menghindari duplikasi maupun kerusakan pada data dan keamanan data lebih terjaga.

**Kata Kunci** – Naïve Bayes, Data Mining, Data Latih, Database, Klasifikasi

## I. PENDAHULUAN

Indihome (Indonesia Digital Home) ialah salah satu layanan komunikasi serta informasi dari PT. Telekomunikasi Indonesia. Layanannya itu berbentuk semacam telepon rumah, internet, serta layanan tv interaktif. Karna penunjukan inilah Telkom membagikan bukti diri kepada Indihome selaku 3 layanan dalam satu paket( 3 in 1) tidak hanya internet, pelanggan pula memperoleh siaran Televisi berbayar serta saluran telepon.

*Dalam pelaksanaannya, Telkom bekerjasama dengan developer teknologi telekomunikasi guna membangun rumah yang mempunyai konsep digital. Pelayanan Indihome cuma bisa diterapkan pada rumah yang di wilayahnya ada jaringan serat optik dari Telkom( FTTH) serta zona yang masih memakai kabel tembaga.*

Penelitian ini mengklasifikasi, serta menganalisis data pelanggan potensial *Put In Service* Indihome dengan menggunakan Algoritma Naïve Bayes.

---

\*) **penulis korespondensi:** M. Fauzan Al-Ghifary D. W  
Email: [muhammad.fauzan18171@unsika.ac.id](mailto:muhammad.fauzan18171@unsika.ac.id)

## II. PENELITIAN YANG TERKAIT

### A. Data dan Informasi

Data adalah sebuah fakta atau hasil dari observasi yang mentah dengan berupa ukuran atribut (karakteristik) dari representasi atau penyajian fakta yang mewakili sebuah objek. Sedangkan pengertian informasi adalah hasil dari pemrosesan input yang terstruktur, memiliki arti dan dapat bermanfaat bagi orang yang mengetahuinya. Informasi memiliki sifat yang relevan, dapat diandalkan (reliable), lengkap, dapat dipahami, diverifikasi dan memiliki keterkaitan dengan waktu. [1]

### B. Klasifikasi

Klasifikasi merupakan suatu kegunaan penggalian informasi yang akan menghasilkan model untuk meramalkan klasifikasi item dalam baris informasi. Tatanan sendiri memiliki 2 fase siklus yang terdiri dari tahap karakterisasi dan tahap

pembelajaran [2]. Setiap perhitungan karakterisasi memiliki aturan yang sama yaitu melakukan pembelajaran dalam model sehingga dapat meramalkan kontribusi nama kelas yang tepat sebagai hasil. Sedangkan pada tahap urutan, tahap ini mengambil suatu informasi yang tanda kelasnya tidak jelas ketika diantisipasi dan kemudian dicoba menggunakan model yang telah dibuat dalam pengalaman yang berkembang [3].

### C. Algoritma Naïve Bayes

Algoritma *Naïve Bayes* merupakan pengumpulan informasi yang diperlukan saat meramalkan peluang suatu kelas. Strategi Bayes adalah rencana terukur untuk melakukan deduksi induktif pada masalah pengumpulan. Kapan pun kami pertama kali memeriksa ide dan definisi mendasar dalam hipotesis Bayes, maka kami menggunakan hipotesis ini untuk dikelompokkan dalam Penambangan Informasi. [2]. Kemudian, pada titik itu, Karakterisasi Bayesian bergantung pada hipotesis Bayes yang memiliki kemampuan serupa untuk berkumpul sebagai pohon pilihan dan organisasi otak. Kemudian Bayesian Order juga didemonstrasikan memiliki akurasi dan kecepatan tinggi ketika diterapkan pada koleksi informasi dengan ukuran data yang cukup besar [4].

Teorema bayes memiliki bentuk sebagai berikut :

$$P(H|E) = \frac{P(H|E) \times P(H)}{P(E)}$$

Keterangan :

H = Teori informasi X adalah kelas tertentu

X = Informasi dengan kelas yang tidak jelas

P(H) = Probabilitas teori H (prob. sebelumnya) P(X|H) = Probabilitas X dalam keadaan ini

P(X) = Probabilitas dari X

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob).

Ada 4 fase proses Naïve Bayes, lebih spesifiknya:

1. Gandakan semua faktor kelas,
2. Menghitung jumlah kasus per kelas,
3. Menghitung jumlah kelas per tabel,
4. Analisis hasil per kelas.

## III. METODE PENELITIAN

### A. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data Put In Service, yaitu 100 dataset dari sto CKP. Ada sekitar 75 Data latih yang digunakan sedangkan ada sekitar 25 data uji yang telah ditandai dengan mempertimbangkan faktor-faktor saat ini. Masalah yang harus diselesaikan dalam pemeriksaan ini adalah cara agar hasil nilai ketepatan yang jauh lebih ideal.

### B. Pengolahan Data

Pada tahap pengolahan data memiliki 3 proses, yaitu :

- a) Memeriksa data yang inkonsisten,
- b) Menghapus duplikasi data,
- c) Memperbaiki kesalahan pada data.

### C. Evaluasi dan Validasi Data

Evaluasi dan validasi data harus dimungkinkan dengan menggunakan persetujuan silang 10 lipatan. Jika Anda menggunakan strategi ini dengan memisahkan secara serampangan menjadi setiap bagian yang terdiri dari 10 bagian, untuk setiap bagian itu sendiri, siklus pengumpulan akan dilakukan terlebih dahulu. Sedangkan estimasi presisi diperkirakan menggunakan disarray grid dan Collector Working Qualities (ROC) bend untuk mengukur harga AUC. AUC digunakan sebagai proporsi eksekusi prasangka dengan menilai kemungkinan hasil yang diperoleh dari tes sembarangan memilih dari populasi positif atau negatif, semakin terlihat harga AUC, semakin banyak membumi karakterisasi berikutnya. Karena AUC adalah subset dari daerah satuan kuadrat, nilai selanjutnya akan terus menjadi sesuatu yang serupa yaitu antara 0,0 dan 1,0.

## IV. HASIL DAN PEMBAHASAN

### A. Klasifikasi Pelanggan

Klasifikasi Pelanggan Data uji dan Data latih yang digunakan untuk klasifikasi berjumlah 25 Data uji dan 75 Data latih, dimana informasi persiapan kemudian akan ditangani menggunakan teknik naïve bayes. Informasi persiapan yang digunakan harus terlihat pada Tabel 4.1. [6]

**Tabel 4.1 Data Latih**

Kode Pelanggan	Jumlah Pembelian	Interval Waktu	Lokasi	Target
P1	Sangat banyak	Bulanan	Dekat	Berpotensi
P2	Sangat banyak	Bulanan	Jauh	Berpotensi
P10	Banyak	Bulanan	Jauh	Berpotensi
P61	Sedikit	Bulanan	Dekat	Berpotensi
P75	Sangat banyak	Bulanan	Dekat	Berpotensi

Informasi yang digunakan sebagai informasi pengujian dan persiapan melibatkan informasi pada Februari 2021 yang seharusnya dapat dilihat pada Tabel 4.2.

**Tabel 4.2 Komposisi Data Latih & Data Uji**

Jenis Data	Jumlah	Keterangan
Data Latih	75	Data Put In Service pada bulan Februari 2021
Data Uji	25	Data Put In Service pada bulan Februari 2021

Nilai kemungkinan setiap basis tidak sepenuhnya ditentukan dari nilai kemungkinan setiap aturan ke dalam klasifikasi, khususnya :

- a. Kemungkinan jumlah pembelian di setiap kelas, seperti yang ditampilkan pada Tabel 4.3

**Tabel 4.3 Kemungkinan Jumlah Pembelian**

Jumlah Pembelian	Jumlah Kejadian Jumlah Pembelian		Probabilitas	
	Berpotensi	Tidak Berpotensi	Berpotensi	Tidak Berpotensi
Sangat banyak	19/55	0/20	0,34	0
Banyak	18/55	0/20	0,33	0
Sedikit	17/55	2/20	0,31	0,1
Sangat banyak	1/55	18/20	0,002	0,9

- pengujian, yang menambahkan hingga 25 informasi klien, akan terlihat pada Tabel 4.7. [7]

**Tabel 4.7 Data Uji**

Kode pelanggan	Jumlah Pembelian	Interval Waktu	Lokasi	Prediksi target
P76	Banyak	Bulanan	Dekat	Berpotensi
P80	Banyak	Bulanan	Dekat	Berpotensi
P85	Sangat sedikit	Bulanan	Dekat	Berpotensi
P88	Banyak	Bulanan	Jauh	Tidak Berpotensi
P90	Banyak	Bulanan	Dekat	Berpotensi
P95	Sangat sedikit	Bulanan	Dekat	Berpotensi
P100	Banyak	Bulanan	Dekat	Berpotensi

- b. Kemungkinan rentang waktu untuk setiap kelas seperti yang ditampilkan pada Tabel 4.4

**Tabel 4.4 Probabilitas Interval Waktu**

Interval Waktu	Jumlah Kejadian Jumlah Pembelian		Probabilitas	
	Berpotensi	Tidak Berpotensi	Berpotensi	Tidak Berpotensi
Bulanan	53/55	3/20	0,95	0,15

Hasil urutan yang diperkenalkan pada tabel 4.8 kemudian dapat diubah dengan menggunakan teknik grid yang tidak teratur seperti di bawah ini :

**Tabel 4.8 Pengujian Confusion Matrix**

Confusion Matrix	Data Sebenarnya		
Data Prediksi	TRUE	21	0
	FALSE	2	0

- c. Kemungkinan daerah di setiap kelas seperti yang ditampilkan pada Tabel 4.5

**Tabel 4.5 Probabilitas Lokasi**

Jumlah Pembelian	Jumlah Kejadian Lokasi		Probabilitas	
	Berpotensi	Tidak Berpotensi	Berpotensi	Tidak Berpotensi
Dekat	39/56	8/20	0,72	0,34
Jauh	16/56	14/20	0,31	0,65

Dari tabel di atas, cenderung terlihat bahwa penyajian pemanfaatan teknik naïve bayes dapat diperkirakan dengan menghitung keuntungan dari akurasi, review, dan ketepatan. Dalam estimasi ini hasil yang di dapat langsung dikonversikan ke dalam bentuk presentase [8].

- d. Kemungkinan objektif untuk setiap kelas adalah seperti yang ditampilkan pada Tabel 4.6

**Tabel 4.6 Probabilitas Target**

Jumlah kejadian target		Probabilitas	
Berpotensi	Tidak berpotensi	Berpotensi	Tidak berpotensi
56/75	20/75	0,73	0,27

- a) Jadi untuk menghitung nilai review yang melibatkan rumus perhitungan sebagai berikut :

$$\text{Nilai recall} = 21/(21+2) = 91\%$$

- b) Jadi untuk menghitung nilai akurasi menggunakan rumus (2.1), hasil perhitungannya adalah sebagai berikut :

$$\text{Nilai ketepatan} = 21+2/(21+2+0+2) = 92\%$$

- c) Untuk menghitung nilai precision menggunakan rumus hasil perhitungan sebagai berikut :

$$\text{Nilai Precision} = 21/(21+0) = 100\%$$

## V. KESIMPULAN DAN SARAN

Penelitian ini memiliki beberapa kesimpulan meliputi:

- Eksplorasi ini dipimpin menggunakan peralatan Rapidminer pada dataset menggunakan perhitungan Naïve Bayes.
- Perhitungan Naïve Bayes adalah urutan terukur yang dapat digunakan untuk meramalkan kemungkinan pendaftaran kelas. Metode ini juga mempunyai sistem klasifikasi yaitu, Bayesian Classification.
- Pengujian pada informasi pernyataan ulang Bantuan Put In dengan proses penambangan perhitungan Naïve Bayes menciptakan tingkat presisi yang benar-benar layak, di mana dalam menguji model informasi, seluruh kumpulan informasi digunakan sebagai pengujian informasi.

## B. Menguji Teknik Gullible Bayes Memanfaatkan Informasi Pengujian

Setelah sistem pengelompokan yang menggunakan teknik naïve bayes selesai, maka pada saat itu, pengujian informasi

Saran untuk penelitian selanjutnya metode yang digunakan perlu dikembangkan kembali dengan menggunakan metode data mining lainnya, dan sebaiknya data-data yang digunakan perlu diintegrasikan dengan database agar menghindari duplikasi maupun kerusakan pada data dan keamanan data lebih terjaga.

#### UCAPAN TERIMA KASIH

Terima kasih kepada pembimbing, anggota kelompok tugas akhir, dan rekan-rekan Teknik Elektro Universitas Singaperbangsa Karawang yang telah mendukung dan membantu dalam proses pembuatan penelitian ini.

#### DAFTAR PUSTAKA

- [1] Regulation of Transcription Elongation in Response to Osmostress, <https://www.yeastgenome.org/reference/S000207095>
- [2] Kusnawi., (2007) Pengantar Solusi Data Mining
- [3] Ninki Hermaduanty., (2008) Bidang Kesehatan Sistem Pendukung Keputusan Berbasis SMS untuk Menentukan Status Gizi dengan Metode K-Nearest Neighbor
- [4] Muhamad et al., (2013) Implementasi Algoritma Naïve Bayes Berbasis Particle Swarm Optimization Untuk Memprediksi Penyakit Hepatitis
- [5] Harisnawan, E., (2016) Laporan Kegiatan Magang PT Telekomunikasi Indonesia Witel Yogyakarta (Doctoral dissertation, STIE YKPN)
- [6] Putro, H. F., Vulandari, R. T., & Saptomo, W. L. Y. (2020). Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan. *Jurnal TIKOMSIN (Teknologi Informasi dan Komunikasi Sinar Nusantara)*, 8(2).
- [7] Manalu, E., Sianturi, F. A., & Manalu, M. R. (2017). Penerapan Algoritma Naïve Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada Cv. Papadan Mama Pastries. *Jurnal Mantik Penusa*, 1(2).
- [8] Webb, G. I., Keogh, E., & Miikkulainen, R. (2010). Naïve Bayes. *Encyclopedia of machine learning*, 15, 713-714.