



Penerapan Algoritma *Naïve Bayes* Dalam Prediksi Tingkat Kepuasan Konsumen Di Hotel Novotel Karawang

Ryan Septianto

Prodi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa
ryanseptianto50@gmail.com

Abstract

Novotel Hotel is a hotel that has been established since 2018 with more than 40 employees. At the Novotel hotel, there are work divisions such as kitchen,cooks, security, waiters, managers who already have their respective divisions working regularly in the hotel. At Novotel hotels, the existing facilities such as registration of visitor data, rooms, food presentation, service, to the costs incurred for the facilities used do not yet have an assessment of customer satisfaction with Novotel hotels. This provides a lack of material for evaluating the Novotel Karawang hotel in improving the quality of the hotel to provide comfortable facilities to consumers. Hotel Novotel Karawang to improve hotel quality with data obtained from consumers is an evaluation strategy for improving the quality of existing facilities at the hotel, the method used to obtain data is by distributing questionnaires to consumers who are given several questions related to the hotel, from the data obtained it will be processed using the Naïve Bayes algorithm to find out data on customer satisfaction with Novotel hotels as evaluation material.

Keywords: *Novotel Hotel, Satisfaction, Consumers, Nave Bayes.*

Abastrak

Hotel Novotel merupakan hotel yang suda berdiri sejak tahun 2018 dengan jumlah pegawai lebih dari 40 orang. Pada hotel Novotel memiliki divisi kerja seperti kitchen, juru masak, keamanan, pramusaji, manajer yang sudah memiliki bagian kerja masing-masing divisi yang sudah teratur di hotel. Pada hotel Novotel dalam fasilitas yang ada seperti pendaftaran data pengunjung, kamar, penyajian makanan, pelayanan, hingga biaya yang dikeluarkan untuk fasilitas yang digunakan belum memiliki penilaian kepuasan konsumen terhadap hotel Novotel. Hal ini memberikan kekurangan untuk bahan evaluasi hotel Novotel Karawang dalam meningkatkan kualitas hotel untuk memberikan kenyamanan fasilitas terhadap konsumen. Hotel Novotel Karawang untuk meningkatkan kualitas hotel dengan bahan data yang diperoleh dari konsumen merupakan strategi evaluasi peningkatan kualitas fasilitas yang ada di hotel, cara yang dilakukan untuk mendapatkan data yaitu dengan

penyebaran kuisisioner terhadap konsumen yang diberikan beberapa pertanyaan terkait hotel, dari data yang didapatkan maka akan diolah menggunakan algoritma *Naïve Bayes* untuk mengetahui data kepuasan konsumen terhadap hotel Novotel sebagai bahan evaluasi.

Kata kunci:Hotel Novotel,Keputusan, Konsumen, *Naïve Bayes.*

1. Pendahuluan

1.1. Latar Belakang

Kebutuhan akan informasi yang akurat dan cepat sangat dibutuhkan dalam kehidupan sehari-hari seiring dengan berkembangnya teknologi informasi. Namun, kebutuhan informasi yang meningkat untuk memperoleh pengetahuan baru terkadang tidak menyajikan informasi yang tepat, biasanya informasi tersebut masih perlu digali dari data dengan populasi yang cukup besar. Analisis data diperlukan untuk mengetahui potensi informasi yang ada. Hal ini mendorong munculnya bidang pengetahuan baru

untuk memecahkan masalah penggalian informasi atau pola penting dari data dalam jumlah besar yang disebut data mining.

Kepuasan pelanggan merupakan suatu tanggapan dari konsumen atas kinerja yang telah diberikan sesuai dengan harapan pelanggan. Kepuasan pelanggan merupakan suatu tingkatan dimana kebutuhan, keinginan dan harapan dari pelanggan dapat terpenuhi yang akan mengakibatkan terjadinya pembelian ulang atau kesetiaan yang berlanjut. Hotel Novotel Karawang untuk meningkatkan kualitas hotel dengan bahan data yang diperoleh dari konsumen merupakan strategi evaluasi peningkatan kualitas fasilitas yang ada di hotel, cara yang dilakukan untuk mendapatkan data yaitu dengan penyebaran kuisioner terhadap konsumen yang diberikan beberapa pertanyaan terkait hotel, dari data yang didapatkan maka akan diolah menggunakan algoritma *Naïve Bayes* untuk mengetahui prediksi tingkat kepuasan konsumen terhadap hotel Novotel sebagai bahan evaluasi. Prediksi jumlah tingkat kepuasan merupakan faktor penting yang menentukan kelancaran usaha suatu perusahaan. Prediksi ini sangat berguna untuk menentukan berapa banyak tingkat kepuasan pada bulan selanjutnya. Permasalahan yang umum dihadapi oleh suatu perusahaan adalah bagaimana tingkat akurasi dalam memprediksi atau meramalakan tingkat kepuasan dimasa mendatang. Prediksi tersebut sangat berpengaruh untuk menentukan tingkat kepuasan mana yang sangat dominan. Untuk menghasikan prediksi yang tepat tentu saja dibutuhkan kecermatan dan ketelitian. Tujuan penelitian ini adalah penulis mencoba menerapkan data mining metode decision tree Algoritma *Naïve Bayes* pada Hotel Novotel Karawang dan diharapkan dapat memberikan informasi berupa prediksi kepuasan konsumen. Sehingga kedepannya pemilik bisnis ini dapat melakukan analisa tingkat kepuasan mengikuti trend dan kebutuhan konsumen.

1.2. Tinjauan Literatur

1.2.1. Hotel

Usaha penyediaan akomodasi berupa kamar-kamar di dalam suatu bangunan yang dapat dilengkapi dengan jasa pelayanan makan dan minum, kegiatan hiburan dan/atau fasilitas lainnya secara harian dengan tujuan bangunan berkamar banyak yang disewakan sebagai tempat untuk menginap. Hotel merupakan tempat makan, minum, beristirahat bagi orang-orang yang sedang atau bentuk akomodasi yang dikelola secara komersial, disediakan bagi setiap orang untuk memperoleh pelayanan, penginapan, makan dan minum [1].

1.2.2. Rapid Miner

Rapid Miner adalah alat komputasi statistik dikembangkan dan berhasil diterapkan pada berbagai data untuk dianalisis dan memantau prosesnya. RapidMiner proyek dimulai pada tahun 2001 oleh Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer di Artificial Kelompok Intelijen dari Katharina Morik di Universitas Teknologi Dortmund. Rapid Miner adalah salah satu alat penambangan data yang digunakan untuk menganalisis informasi yang diakses web. Ini digunakan untuk penelitian, pendidikan, pembuatan prototipe cepat, aplikasi pengembangan, dan aplikasi industri. Aplikasi ini berlisensi *Open Source*, yang termasuk pembersihan data, transformasi data, optimasi, validasi dan visualisasi. Itu visualisasi berisi melihat data yang dianalisis dalam bentuk sebar plot, Bar, Pie chart. Berbagai pengelompokan dan klasifikasi algoritma untuk melakukan proses analitik. Salah satunya Fitur utama alat ini yaitu, ia akan menganalisis data tanpa coding program, namun jika ada orang ingin menganalisis data dengan pengkodean mereka sendiri itu juga bisa dimasukkan dalam aplikasi. Berbagai jenis dataset dapat diimpor oleh RapidMiner seperti, excel, csv, xml, arff, akses dll . Sejak 2007, *Rapid Miner* telah sangat diperluas dan menjadi salah satu yang paling penambangan data penting dan alat analisis data [2]

1.2.3. Data Mining

Data mining merupakan suatu langkah dalam melakukan *Knowledge Discovery in Databases (KDD)*. *Knowledge discovery* sebagai suatu proses terdiri atas pembersihan data (*data cleaning*), integrasi data (*data integration*), pemilihan data (*data selection*), transformasi data (*data transformation*), data mining, evaluasi pola (*pattern evaluation*) dan penyajian pengetahuan (*knowledge presentation*).[3]

1.2.4. Naïve Bayes

Naïve Bayes merupakan salah satu metode yang dapat digunakan untuk mengklasifikasikan data. Bayesian classification merupakan pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Teorema ini dikemukakan oleh ilmuwan Inggris, Thomas Bayes, dengan memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. Teorema Bayes dikombinasikan dengan "*Naïve*" yang berarti setiap atribut/variable bersifat bebas (*independent*).[4]

1.2.5. Knowledge Discovery Database (KDD)

Knowledge Discovery In Databases (KDD) adalah proses non trivial untuk mencari dan mengidentifikasi pola (*pattern*) dalam data.

Knowledge Discovery In Database (KDD) merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar serta hubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dai pola-pola sejumlah kumpulan data .

1.2.6. Microsoft Word

Microsoft Word merupakan perangkat lunak pengolah kata terbaik dan terpopuler di dunia. Selain itu, Microsoft Word juga tersedia di berbagai jenis sistem operasi. Aplikasi ini telah didistribusikan di sistem operasi Microsoft Windows, Mac OS, IOS, dan Android. Hal ini menyebabkan popularitas Microsoft Word bertahan, walaupun banyak terdapat aplikasi pengolah kata yang terus bermunculan. Selain kualitas informasi, kegunaan juga merupakan salah satu aspek penting yang harus diperhatikan agar sebuah aplikasi dapat terus digunakan oleh banyak pengguna.[5]

1.2.7. Microsoft Excel

Microsoft Excel adalah Program aplikasi pada Microsoft Office yang digunakan dalam pengolahan angka (Aritmatika). Microsoft Excel sangat membantu pegawai perkantoran dalam menyelesaikan permasalahan yang mudah sampai dengan yang rumit dalam bidang administratif perkantoran khususnya. Microsoft Excel memang dikenal dengan penggunaan rumus-rumus atau formula dalam lembar kerjanya. Penggunaan rumus yang efektif akan memudahkan pegawai perkantoran dalam membuat laporan pekerjaan dengan menggunakan Microsoft Excel.[6]

2. Metode Penelitian

Penelitian ini menggunakan metode Knowledge Discovery in Database (KDD). Metode yang digunakan untuk mendapatkan pengetahuan asal dari sebuah database merupakan definisi dari KDD. Proses kerja KDD terlihat pada gambar di bawah ini:[8]



Gambar 1. Proses Knowledge Discovery in Database

2.1. Data Selection (Data Seleksi)

Tahapan ini merupakan tahapan dimana data-data dikumpulkan dan kemudian diolah dengan metode Decision Tree algoritma *Navie bayes*. Berikut ini adalah data kepuasan tamu hotel Novotel Karawang yang akan digunakan untuk pemrosesan data yang diolah dengan metode decision tree. Jumlah dataset asli pada data kepuasan tamu sebanyak 152 record.

Dari semua atribut yang ada terdapat 5 atribut yang akan digunakan dalam proses KDD. Atribut tersebut yaitu:

Tabel 1. Atribut Yang Digunakan

No	Atribut	Tipe
1	Nama	text
2	Fasilitas	text
3	Pelayanan	text
4	Kebersihan	text
5	Status	text

2.2. Processing (Cleaning)

Setelah dilakukan data selection, atribut atau variabel yang digunakan terdiri dari 3 atribut yang terdapat dalam penentuan minat. Dengan 3 atribut. Pada atribut fasilitas,pelayanan,dan kebersihan memiliki value yang dapat dilihat pada gambar dibawah ini:

No	Atribut	Value
1	Fasilitas	Cukup
		Baik
		Sangat Baik
2	Pelayanan	Cukup
		Baik
		Sangat Baik
3	Kebersihan	Cukup
		Baik
		Sangat Baik

Gambar 2. Atribut dan Value Yang Digunakan

2.3. Data Transformation

Dalam tahap data transformation tidak ada data yang dibersihkan, pembagian data set dengan menggunakan process split data untuk mnghasilkan data training dan data testing dengan rasio 75:25. Data training berjumlah 112 data dan data testing 40 data dari total keseluruhan yaitu 152 data. Tabel data tersebut dapat dilihat pada gambar berikut:

Fasilitas	Pelayanan	Kebersihan	Status
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Sangat Baik	Baik	Cukup	Puas
Cukup	Cukup	Cukup	Kurang Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Baik	Baik	Cukup	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Cukup	Cukup	Cukup	Kurang Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Baik	Baik	Baik	Puas
Cukup	Cukup	Cukup	Kurang Puas
Baik	Baik	Cukup	Puas
Baik	Baik	Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Cukup	Cukup	Cukup	Kurang Puas
Baik	Baik	Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Cukup	Cukup	Cukup	Kurang Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas
Baik	Baik	Sangat Baik	Puas
Sangat Baik	Sangat Baik	Sangat Baik	Sangat Puas

Gambar 3. Data Testing

2.4. Data Mining

Metode yang penulis gunakan dalam mencari pola menarik adalah metode Decision Tree yaitu metode klasifikasi yang paling sering digunakan. Dalam pengerjaannya tidak memerlukan waktu yang lama dan hasilnya pun mudah untuk dipahami. Algoritma *Navie bayes* memberikan hasil akhir berupa pohon keputusan (decision tree) dari hasil perhitungan probabilitas.

2.5. Interpretation / Evalaution

Pada tahap evaluasi melibatkan hasil perhitungan accuracy dari data testing dengan bantuan Confusion Matrix. Sedangkan, dalam tahap interpretasi merupakan visualisasi dari hasil evaluasi yang sudah dilakukan.

3. Hasil dan Pembahasan

Pada penelitian ini dilakukan perhitungan secara manual dan dengan bantuan tools RapidMiner untuk mendapatkan sebuah pohon keputusan (decision tree) yang dapat memprediksi tingkat kepuasan tamu Hotel Novotel Karawang

3.1. Perhitungan Secara Manual

Sesuai dengan langkah pembuatan pohon keputusan (decision tree) setelah melakukan analisis data yang menghasilkan data training dan data testing maka dilanjutkan dengan menentukan nilai probabilitas pada data training. Hasil perhitungan akan digunakan untuk menentukan root atribut atau atribut yang akan menjadi the best classifier.

Rumus menghitung nilai entropy:

$$P(C) = \frac{X_{C1}}{X_C + X_C + X_C}$$

Keterangan :

P(C) = Class yang di hitung

X_{C1} = Jumlah nilai Class

X_C = Jumlah keseluruhan Class

Didapatkan hasil perhitungan seperti pada gambar di bawah ini:

Class			Probabilitas		
Sangat Puas	Puas	Kurang Puas	Sangat Puas	Puas	Kurang Puas
62	68	22	0.407	0.447	0.144

Gambar 4. Hasil Perhitungan Probabilitas

No	Atribut	Value	Sangat Puas	Puas	Kurang Puas	Sangat Puas	Puas	Kurang Puas
1	Fasilitas	Cukup	0	5	21	0,015	0,073	0,88
		Baik	0	58	1	0,015	0,85	0,08
		Sangat Baik	62	5	0	0,97	0,073	0,04
		Total	62	68	22	2,996		
2	Pelayanan	Cukup	0	4	22	0,015	0,088	0,92
		Baik	3	58	0	0,061	0,85	0,04
		Sangat Baik	59	6	0	0,923	0,058	0,04
		Total	62	68	22	2,995		
3	Kebersihan	Cukup	0	9	21	0,015	0,132	0,88
		Baik	3	53	1	0,061	0,779	0,08
		Sangat Baik	59	6	0	0,920	0,088	0,04
		Total	62	68	22	2,995		

Gambar 5. Hasil Perhitungan keseluruhan Probabilitas

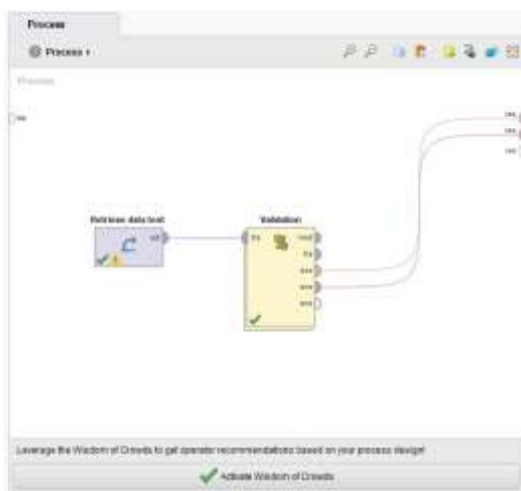
3.2. Pengujian Menggunakan RapidMiner

Dengan menggunakan algoritma *Navie bayes*, akan dibentuk pohon keputusan (decision tree) menggunakan bantuan tools RapidMiner. Langkah pertama adalah import data

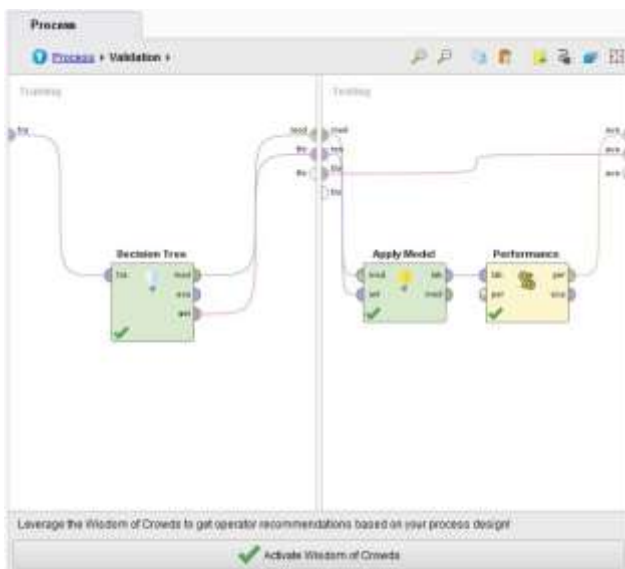


Gambar 6. Import Data

Dilanjutkan dengan menambahkan operators yang dibutuhkan, yaitu operator *Split Validation* dan menggunakan perintah *deccision tree*, *apply model* dan *performance*.



Gambar 7. Operator Split Validation



Gambar 8. Operator Decision tree, apply model, performance

Lakukan running proses untuk mendapatkan hasil berupa pohon keputusan dari data testing dan mendapatkan keakuratan dari data samples. Berikut adalah pohon keputusan yang didapatkan dengan bantuan tools RapidMiner.



Gambar 9. Pohon Keputusan

3.3. Evaluasi Hasil Manual

Setelah selesai menghitung probabilitas dan gain, selanjutnya yang akan dilakukan adalah menghitung *accuracy* pada data dengan bantuan *confusion matrix*. *Confusion matrix* memiliki empat istilah hasil proses klasifikasi, yaitu *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, *False Negative (FN)*

	Actual Puas	Actual Sangat Puas	Actual Kurang Puas	class precision
pred Puas	5	0	0	100.0%
pred Sangat Puas	1	0	0	75.0%
pred Kurang Puas	0	0	0	100.0%
class recall	66.7%	100.0%	100.0%	

Gambar 10. Confusion Matrix Akurasi

A. Accuracy

Accuracy adalah presentase dari total data uji coba yang benar diidentifikasi. Dalam pembahasan *accuracy*, *decision tree* dan *kappa* maka dapat dilihat untuk masing-masing presentase yang didapat yaitu sebesar *accuracy* 90.91%.

B. Kappa statistic

analisis statistical yang didasarkan pada kecocokan penafsiran atau derajat kesepakatan (*degree of agreement*) untuk data kualitatif, pada dasarnya *kappa statistic* menunjukkan analisis di antara kelas-kelas yang berbeda. Semakin tinggi *kappa statistic*, maka akan di pertimbangkan sebagai performa yang mempunyai kriteria atau kinerja yang bagus. *Kappa* yang dihasilkan dari percobaan data uji adalah sebesar 0.833.

	Actual Puas	Actual Sangat Puas	Actual Kurang Puas	class precision
pred Puas	5	0	0	100.0%
pred Sangat Puas	1	0	0	75.0%
pred Kurang Puas	0	0	0	100.0%
class recall	66.7%	100.0%	100.0%	

Gambar 11. Tampilan Menu Data kappa

4. kesimpulan

Berdasarkan penelitian yang dilakukan dengan latar belakang yang sedang dialami pada Hotel Novotel adalah mengenai pengeolahan perhitungan menggunakan microsoft excel dan aplikasi *Rapid Miner* dalam pengolahan data dengan hasil kuesioner yang telah diisi oleh koresponden yang memberikan perangkangan/penilaian pada pertanyaan-pertanyaan yang sudah dibuat.

Dari hasil pembahasan yang telah diuraikan oleh penulis, maka dapat ditarik kesimpulan Menerapkan metode data mining algoritma *naïve bayes* dengan memanfaatkan training data guna menghasilkan probabilitas setiap atribut dengan kelas yang berbeda, sehingga perhitungan tersebut menghasilkan nilai yang dapat dioptimalkan untuk menentukan kepuasan konsumen hotel dan membantu mengevaluasi penilaian menjadi lebih akurat.

Referensi

- [1] A. Susepti, "KEPUASAN DAN LOYALITAS TAMU HOTEL (Studi tentang Persepsi Tamu Hotel Mahkota Plengkung Kabupaten Banyuwangi)," vol. 50, no. 5, pp. 27–36.
- [2] P. Stmik and J. Sti, "PENGUNAAN MODEL KLUSTRISASI DENGAN METODE K-MEANS UNTUK MENDETEKSI AKTIVITAS PENGGUNA WEB MENGGUNAKAN RAPIDMINER BERDASARKAN USER-AGENT-BASED STUDI KASUS PADA APLIKASI-APLIKASI," vol. 4, no. September, 2020.
- [3] H. D. Wijaya and S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," vol. 7, no. 1, pp. 1–7, 2020.
- [4] J. Teknologi, I. Jtsi, M. R. Handoko, F. Teknik, and U. T. Indonesia, "SISTEM PAKAR DIAGNOSA PENYAKIT SELAMA KEHAMILAN MENGGUNAKAN METODE NAIVE BAYES BERBASIS WEB," vol. 2, no. 1, pp. 50–58, 2021.
- [5] S. D. Puspitasari, "Analisis Usability Pada Aplikasi Microso Word Berdasarkan Model Analisis Usability Pada Aplikasi Microsoft Word Berdasarkan Model Nielsen," *Academia*, 2017.
- [6] S. Informasi, "Pelatihan pengolahan data administrasi perkantoran menggunakan aplikasi microsoft excel pada balai desa pinanggripan kecamatan air batu kabupaten asahan," vol. 2, no. 2, pp. 107–114, 2019.