

PENERAPAN DATA MINING UNTUK PEMINATAN RASA ROTI BAKAR DI CV. SARIRAOS MENGGUNAKAN METODE NAÏVE BAYES

Nurhadi Surojudin¹⁾, Enuh²⁾

Program Studi Teknik Informatika Fakultas Teknik
Universitas Pelita Bangsa
nurhadi@pelitabangsa.ac.id

Disetujui, 25 September 2019

Abstraksi

Perubahan perilaku pembeli terhadap perkembangan zaman hal ini menyebabkan para pelaku usaha kuliner berinovasi supaya bisnis tetap berjalan dan bertahan hal ini akan berpengaruh besar pada penghasilan dan omset yang didapat untuk mengembangkan minat rasa pada roti bakar CV.Sariraos dengan menganalisa *data mining* dan metode *naïve bayes* kegunaannya untuk menghitung peluang dari satu kelas dan masing-masing atribut serta menentukan kelas mana yang paling optimal kemudian analisa tersebut dijelaskan menggunakan aplikasi *rapidminer* kegunaannya untuk memprediksi *accuracy data* dan menggunakan metode *naïve bayes* diharapkan dapat memberikan keputusan yang akurat dalam menentukan rekomendasi minat rasa, metode klasifikasi dan prediksi yang mewakili aturan-aturan tersebut kemudian dikembangkan dengan perhitungan rumus *confusion matrix* berdasarkan hasil penelitian maka dapat disimpulkan dengan metode *naïve bayes* minat rasa yang paling disukai adalah rasa coklat dengan pengujian menggunakan aplikasi *rapidminer* maka menunjukkan tingkat akurasi yang didapat 92.64 %.

Kata kunci : Minat rasa, *Data mining*, *Naïve bayes*, *Rapidminer* , *Confusion matrix*

Abstract

Changes in buyer behavior towards the times this causes culinary business actors to innovate so that businesses continue to run and survive this will have a major effect on the income and turnover obtained to develop a taste interest in CV Sariraos toast by analyzing data mining and its use of naïve Bayes methods. To calculate the odds of one class and each attribute and determine which class is the most optimal, then the analysis is explained using the rapidminer application, its usefulness is to predict accuracy data and using the naïve Bayes method is expected to provide accurate decisions in determining recommendation of taste interest, classification method and predictions that represent these rules are then developed by calculating the confusion matrix formula based on the results of the research, so it can be concluded that with the naïve Bayes method, the most preferred taste is chocolate by testing using an applica. the rapidminer shows the accuracy rate obtained is 92.64%.

Keywords: *taste interest, data mining, Naïve Bayes, Rapidminer, Confusion matrix*

1. Pendahuluan

1.1 Latar Belakang

Dewasa ini dengan kemajuan zaman teknologi yang ada sekarang ini sangat dibutuhkan sekali sehingga bisa dimanfaatkan menjadi sebuah data yang lebih bermanfaat pada data mining penggalian informasi ini berguna sekali untuk meningkatkan keuntungan dan penyusunan dalam strategi pemasaran. Yang salah satunya menggunakan sebuah metode yaitu *klasifikasi naïve bayes* merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan prekuensi dan kombinasi dari nilai dataset yang diberikan (Saleh, 2015) kegunaannya untuk mengetahui minat dan ketertarikan produk yang tersedia dengan memprediksi probabilitas suatu anggota class. Dengan demikian perusahaan tersebut dapat mengetahui tingkat ketertarikan pembeli terhadap suatu produk yang kita tawarkan sehingga perusahaan dapat menentukan target dan sasaran pasar lebih rinci dan efisien. Oleh karena itu metode *naïve bayes data mining* sendiri memiliki arti menganalisa data dari

perspektif yang berbeda dan meringkasnya menjadi informasi yang berguna (Medkhar et al, 2013) ini dapat juga diterapkan dalam memprediksi pola pembelian dari analisis kumpulan data penjualan dimasa lampau. Metode naïve bayes merupakan bagian dari teknik yang mampu menangani masalah ketidak pastian dengan menekankan pada konsep probabilitas kegunaanya untuk pengukuran tingkat keakurasian dan optimal algoritma ini menerapkan probabilitas yaitu dimana dengan mencari probabilitas yang tertinggi dengan memasukan beberapa atribut. Pada penelitian ini adalah bagaimana penerapan metode *naïve bayes* untuk menciptakan pengetahuan baru yang kemudian digunakan dalam optimasi strategi pemasaran produk yang baik untuk menghadapi persaingan bisnis kuliner. Permasalahan yang dialami didalam pemasaran saat ini adalah masih menggunakan cara tradisional dari mulut ke mulut sehingga belum tahu bagaimana optimasi yang baik untuk pemasaran produk kuliner tersebut, serta kurangnya menguasai teknik pengolahan data penjualan dalam mengatasi solusi terbaik dari permasalahan tersebut maka dari itu peneliti memberikan kesimpulan dengan mengambil judul “ **Penerapan Data Mining Untuk Peminatan Rasa Roti Bakar Di CV. Sariraos Menggunakan Metode Naïve Bayes** ”.

1.2 Identifikasi Masalah

Berdasarkan latar belakang masalah diatas maka peneliti mengidentifikasi permasalahan sebagai berikut:

1. Belum adanya metode dalam analisa data yang dapat membantu didalam pemasaran CV.Sariraos.
2. Belum adanya metode yang digunakan untuk menentukan jumlah penjualan.
3. Kurang maksimalnya CV.Sariraos dalam mengolah data pemasukan atau pengeluaran.

1.1 Batasan Masalah

Sebagaimana diuraikan latar belakang diatas ada beberapa permasalahan yang dianalisa peneliti, maka perlu adanya suatu batasan masalah untuk memberikan kemudahan dalam penelitian sebagai berikut:

1. Metode yang digunakan adalah dengan algoritma naïve bayes untuk menentukan rasa paling disukai pada CV.Sariraos saat ini.
2. Parameter yang dapat digunakan adalah kepuasan pelanggan, sehingga dapat melihat perkembangan daya pembeli pada CV.Sariraos.

1.2 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah kami uraikan diatas, maka penelitian ini dengan menggunakan penerapan algoritma klasifikasi *naïve bayes* kegunaannya untuk melihat rasa paling disukai semua kalangan mulai dari anak-anak sampai orang dewasa.

1. Untuk membantu meningkatkan sistem pelayanan dan penjualan pada CV. Sariraos
2. Untuk menambah ilmu pengetahuan, pengalaman, dan pengamatan dengan menggunakan metode *naïve bayes* untuk menyelesaikan laporan penelitian

2. Tinjauan Studi

2.1 Tinjauan Pustaka

Tinjauan pustaka merupakan bahan yang tertulis berupa buku, makalah maupun jurnal yang merupakan bahasan topic yang akan diteliti.

2.1 Roti

Roti adalah makanan yang terbuat dari tepung, dengan tahap diadonkan fermentasi dipanggang dengan oven roti dapat dibedakan menjadi tiga macam roti kukus, roti goreng dan roti panggang (bakar) bahan pembuatan roti Antara lain tepung terigu, gula, garam, ragi, mentega, air, telur, dan susu secara umum roti biasanya dijadikan sebagai makanan hidangan penutup.

2.2.1 Data Mining

Data mining merupakan salah satu cara untuk mendapatkan pengetahuan baru dengan memanfaatkan jumlah data yang sangat besar beberapa teknik telah dikembangkan dan diimplementasikan untuk mengumpulkan pengetahuan yang mungkin berguna untuk pengambilan keputusan teknik-teknik yang digunakan untuk pengextrakan pengetahuan dalam data mining adalah pengenalan pola clustering, asosiasi, prediksi dan klasifikasi (defiyanti, 2015)

2.3.7 Teknik Pembelajaran Data Mining

Teknik yang digunakan data mining kaitannya erat dengan penemuan dan pengajaran serta pembelajaran yang dibagi menjadi tiga metode pembelajaran yaitu:

1. Pengawasan belajar (*Supervised learning*)

Teknik yang sering digunakan teknik ini sama dengan *programming by example* ini melibatkan fase pelatihan dimana data pelatihan historis yang karakternya dikotak-kotakan sehingga kita mengetsahui hasil yang diolah dalam algoritma data mining, proses ini melatih algoritma untuk mengenali variabel-variabel dan nilai kunci yang nantinya digunakan sebagai dasar dalam membuat perkiraan-perkiraan ketika ada data baru yang masuk.

2. Tanpa pengawasan belajar (*Unsupervised learning*)

Teknik ini tidak melibatkan fase pelatihan seperti yang terdapat pada *supervised learning* maka teknik ini tergantung pada penggunaan algoiritma yang dapat mendeteksi semua pola seperti *associations* dan *sequences*, yang muncul kriteria penting yang spesifik dalam data masukan pendekatan ini mengarah pada pembuatan banyak aturan (*rules*) yang mengkarakterisasikan penemuan *associations*, *clussters*, dan *segments*. Atutan ini kemudian dianalisis untuk menemukan hal-hal yang penting.

3. Bantuan dalam belajar (*Reinforcement learning*)

Tehnik ini jarang digunakan sehingga dibandingkan dengan dua tehnik lainnya, namun tekni ini memiliki penerapan-penerapan yang terus dioptimalkan dari waktu ke waktu dan memiliki kontrol adaptif. Tehnik ini sangat menyerupai kehidupan nyata yaitu seperti "on-job-training", dimana seorang pekerja diberikan sekumpulan tugas yang sulit sehingga membutuhkan keputusan-keputusan. Pada beberapa titik waktu kelak diberikan penilaian atas *performance* pekerja tersebut kemudian pekerja tersebut diminta mengevaluasi keputusan-keputusan yang telah dibuatnya sehubungan dengan hasil *performance* pekerja tersebut. *Reinforcement learning* sangat tepat digunakan untuk menyelesaikan masalah-masalah yang sulit yang bergantung pada waktu.

3. Kerangka Konsep

Metode penelitian adalah tahapan yang harus dilalui seorang peneliti dimulai dari rumusan masalah sampai diperoleh hasil penelitian sehingga bisa membentuk suatu alat sistematis, adapun tujuan dari penelitian adalah untuk menemukan atau memperoleh suatu hasil untuk mengisi, kekurangan dan kekosongan atau menguji suatu pengetahuan agar hasil yang dicapai tidak menyimpang dari tujuan yang telah ditentukan. Metode penelitian untuk menyelesaikan penelitian ini jenis penelitian pengelompokan dengan melakukan tingkat akurasi rasa paling disukai dengan menggunakan metode klasifikasi *naïve bayes* merupakan sebuah metode klasifikasi menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan inggris Thomas bayes *Cris J Hinde(2009)* mengatakan memiliki tingkat akurasi yang lebih baik dibandingkan model clasifikasi lainnya. CV.Sariraos adalah produksi yang bergerak dibidang kuliner yang berdiri semenjak 2006 yang terletak didaerah cikarang tepatnya dipasir gombang jababeka 1 pintu 7.

1. Pelanggan datang ke Toko Berkah untuk melakukan transaksi pembelian.
2. Pelanggan melakukan pemilihan barang dan menentukan jumlah barang yang akan dibeli.
3. Pelanggan memesan barang yang diinginkan untuk dibeli.

4. Admin mengeluarkan dua bukti pembayaran (*invoice*), satu untuk diserahkan ke pelanggan dan satu lagi disimpan sebagai arsip pembayaran untuk membuat pencacatan penjualan yang nantinya diserahkan ke pemilik.
5. Pelanggan melakukan pembayaran ditempat.
6. Pelanggan menerima bukti pembayaran dan barang yang telah dibeli.

4. Metodologi Penelitian

Alat bantu yang digunakan peneliti untuk mengumpulkan data penelitian berdasarkan permasalahan yang telah diuraikan pada subbab sebelumnya maka diperlukan peralatan yang diperlukan untuk meneliti :

Dalam penelitian ini yang dibutuhkan dalam data pembuatan sebagai bahan penelitian. roti panggang atau roti bakar menunjukkan pada dengan proses pembuatan radiasi pemanasan sehingga menjadi kecoklatan dan garing roti bakar pada umumnya dengan olesan margarin dengan rasa manis seperti selai coklat, strawberry dan lain sebagainya. Biasanya yang digunakan adalah roti iris yang nantinya didalam roti tersebut dioleskan selai banyak berbagai macam aneka rasa dan tergantung selera yang kita mau roti tawar berbagai macam warna mulai dari warna putih, coklat, multigrain dan lain sebagainya sehingga banyak disukai berbagai kalangan mulai dari anak-anak sampai orang dewasa.

5. Hasil Dan Pembahasan

5.1 Hasil

5.1.1 Penentuan Kriteria

Dalam menganalisa ada beberapa atribut untuk hasil dan pembahasan pada CV.Sariraos sebagai berikut:

Tabel 5.1 Atribut

Atribut	Dipenelitian
Nama	Ciri peminat rasa
Jenis kelamin	Laki-laki dan Perempuan
Status perkawinan	Yes dan No
Status pekerjaan	Pelajar dan Pekerja
Minat rasa	Coklat dan strawberry

5.2 Penerapan Metode *Naïve Bayes*

Metode *naïve bayes* bisa disebut *classifacatoin* sebuah metode statistik yang kegunaannya untuk memproses didalam menentukan probabilitas dari suatu kelas teori *naïve bayaes* mempunyai kemampuan mengklasifikasi yang serupa dengan metode *dicision tree*, *naïve bayes* juga bisa mengatasi *datasheet* besar dengan *variable* dan kontinyu berdasarkan asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output dengan kata lain probabilitas mengamati secara bersama yang diambil dari produk probabilitas secara individu.

Kelebihan menggunakan metode *naïve bayes* kita hanya membutuhkan data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian data string bersifat konstan dibedakan dengan data numeric yang bersifat kontinyu, perbedaan ini akan terlihat pada saat menentukan nilai probabilitas setiap kriteria baik data numeric maupun data string adapun penerapan metode *naïve bayes* adalah sebagai berikut.

5.2.1 Data Training Menggunakan Metode Naïve Bayes Aplikasi Rapidminer

Berikut ini terlampir tabel data training terdiri dari 408 data yang nanti digunakan sebagai penghitungan dengan menggunakan metode *naïve bayes* berikut ini adalah data training dari minat rasa yang berjumlah 408 datasheet.

5.2.2 Data Training Menggunakan Aplikasi Rapidminer

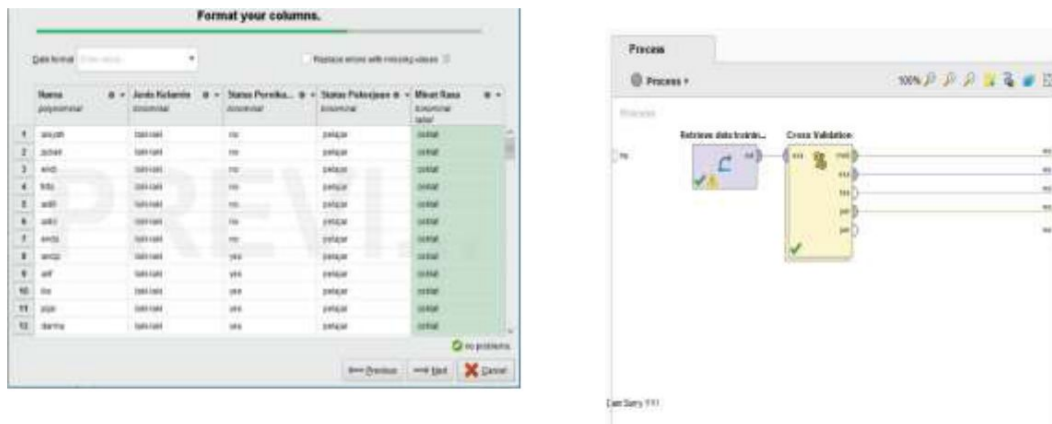
Untuk menentukan data yang akan dianalisis maka langkah utama yang harus dihitung dengan menggunakan aplikasi *rapidminer*, data yang digunakan dalam penelitian ini yang diperoleh dari lapangan berjumlah 408 data dimulai dari Jenis Kelamin, Status Pernikahan Status Pekerjaan Minat Rasa data training yang akan diprediksi pada atribut label, klasifikasi *naïve bayes* dipakai untuk memprediksi minat rasa pada CV.Sariraos.

5.3 Langkah-Langkah Menggunakan Aplikasi Rapidminer

Dalam proses pengumpulan data merupakan proses paling utama untuk diterapkan sehingga menemukan pengetahuan tersembunyi dari data yaitu dengan menggunakan metode klasifikasi metode *naïve bayes*

5.3.1 Proses Import Data

Pada tahapan ini data training yang *di-import* ke dalam tool aplikasi *rapidminer* 9.1.000 dari data 408 kami jadikan dua data yaitu data *training* 408 sementara data training yang akan diprediksi untuk menentukan penyajian data yang berupa tabel *confusion metrix* setelah tabel diketahui.



Gambar 5.2 Proses Data Import 408 Data

Row No.	Most-Rare	Nama	Jenis Kelamin	Status Perk.	Status Pake...
1	sehat	adip	laki-laki	no	pakai
2	sehat	zaki	laki-laki	no	pakai
3	sehat	andi	laki-laki	no	pakai
4	sehat	titu	laki-laki	no	pakai
5	sehat	andi	laki-laki	no	pakai
6	sehat	adip	laki-laki	no	pakai
7	sehat	andi	laki-laki	no	pakai
8	sehat	andi	laki-laki	no	pakai
9	sehat	adip	laki-laki	no	pakai
10	sehat	no	laki-laki	no	pakai
11	sehat	andi	laki-laki	no	pakai
12	sehat	andi	laki-laki	no	pakai
13	sehat	andi	laki-laki	no	pakai
14	sehat	andi	laki-laki	no	pakai

Gambar 5.3 Data Import

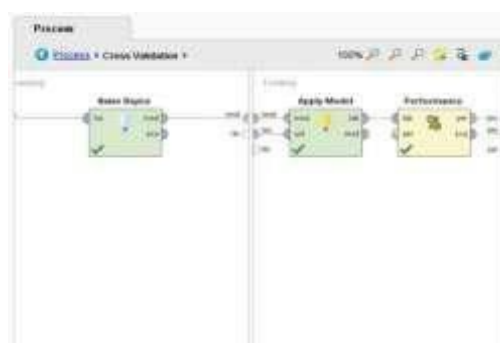
5.3.2 Proses Prediksi Data Training

Proses data training dengan metode *naïve bayes* proses pengujian data yang grafik untuk menghasilkan data testing sehingga bisa *attribute* operator lalu operator *apply model* operator *performance* training dengan *rapoidminer*.

	Real (sehat)	Real (tidak)	Class Predicted
pred (sehat)	342	14	86.27%
pred (tidak)	11	30	73.23%
Overall	85.53%	72.86%	

adalah suatu proses pelatihan sedangkan data testing adalah akan menghasilkan pola dan dan menentukan hasil prediksi dihubungkan pada operator *select* dihubungkan ke *cross validation* kemudian dihubungkan ke sehingga bisa diprediksi data menggunakan aplikasi

5.3.3 Proses Uji data training



Gambar 5.4 Proses Data Training Dengan *Naïve Bayes*

Proses performance adalah untuk mengitung seberapa akurat antara data actual dengan yang sudah diprediksi proses ini juga bisa dibagi menjadi tiga macam yaitu *accuracy*, *precision*, dan *recall* dalam proses ini data yang digunakan adalah 408 data *training*.

Gambar 5.5 Proses Data Training Dengan Hasil *Accurasy*

<i>precision</i>	$\frac{342}{342 + 14} = 0,9607$	$= 0,9607 \times 100 = 96.07 \%$
<i>Recall</i>	$\frac{342}{342 + 16} = 0,9553$	$= 0,9553 \times 100 = 95.53 \%$
<i>Accuracy</i>	$\frac{342 + 36}{342 + 14 + 36 + 16} = 0,9264$	$= 0,9264 \times 100 = 92.64 \%$

6. Kesimpulan

Berdasarkan penelitian pada prediksi minat rasa roti bakar pada CV.Sariraos pada penelitian data mining dengan menggunakan metode *naïve bayes* peneliti dapat mengambil kesimpulan sebagai berikut:

1. Dengan menggunakan metode *naïve bayes* kita dapat mengklasifikasi minat rasa pada roti bakar CV.Sariraos.
2. Minat rasa yang paling disukai adalah rasa coklat.
3. Proses data mining dengan menggunakan metode *naïve bayes* dapat menganalisa data minat rasa dengan pengujian menggunakan aplikasi *rapidminer* yaitu tingkat akurasi 92.64%

Daftar Pustaka

- A. Trimanto, F. Faqih, I. M. Irfani, and S. Timur, "Penerapan Data Mining Untuk Evaluasi Status Kelulusan Mahasiswa Fakultas Teknologi Pertanian Tahun 2015 Menggunakan Algoritma Naïve Bayes Classifier," 2015.
- D. R. Sari, D. Hartama, I. S. Damanik, and A. Wanto, "Penerapan Metode Naïve Bayes dalam Memprediksi Kepuasan Mahasiswa Terhadap Cara Pengajaran Dosen," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. November, p. 287, 2019, doi: 10.30645/senaris.v1i0.34.