

**ANALYSIS OF STORE CUSTOMER SENTIMENT USING THE NAIVE  
BAYES CLASSIFIER METHOD DHENDSUL STORE  
(Case Study: Dhedsul Store)**

**Antonius L B Hallan<sup>1</sup>, Gergorius Kopong Pati<sup>2</sup>, Karolus Wulla Rato<sup>3</sup>**

Teknik Informatika STIMIKOM Stella Maris Sumba

*hallan83@gmail.com<sup>\*1</sup>, gregkopong80@gmail.com<sup>2</sup>*

**ABSTRACT**

*Sentiment Analysis is a technique for extracting text data to obtain information about positive, neutral or negative sentiment. Sentiment analysis is provided by internet users on social media to provide a personal assessment or opinion. The Dhedsul shop that often gets user sentiment through social media is the Dhedsul shop. The sentiment of opinions from consumers about the Dhedsul Shop can be analyzed and used to obtain useful information for other customers and the Dhedsul Shop. By using the Text Mining technique, the classification method will determine whether a sentiment is positive, neutral or negative. One algorithm that is widely used in sentiment analysis is the Naïve Bayes classification method. This research uses the Naïve Bayes Classifier (NBC) method with tf-idf weighting accompanied by the addition of an emotional icon (emoticon) conversion feature to determine the existing sentiment classes from tweets about the Dhedsul Store. The research results show that the Naïve Bayes method without adding features is able to classify sentiment with an accuracy value of 80.85%, while if the tf-idf weighting feature is added along with emotional icon conversion it is able to increase the accuracy value to 81%.*

**Keywords:** *Sentiment, Naïve Bayes Classifier*

**ABSTRAK**

*Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Analisis sentimen diberikan oleh pengguna internet pada media sosial untuk memberikan suatu penilaian atau opini pribadi. Toko Dhedsul yang sering mendapatkan sentimen pengguna melalui media sosial adalah Toko Dhedsul. Adanya sentimen opini dari konsumen tentang Toko Dhedsul dapat dianalisis dan dimanfaatkan untuk mendapatkan informasi yang berguna bagi pelanggan lain maupun pihak Toko Dhedsul. Dengan menggunakan teknik Text Mining metode klasifikasi, akan diketahui suatu sentimen bernilai positif, netral atau negatif. Salah satu algoritme yang banyak digunakan dalam analisis sentimen adalah metode klasifikasi Naïve Bayes. Penelitian ini menggunakan metode Naïve Bayes Classifier (NBC) dengan pembobotan tf-idf disertai penambahan fitur konversi ikon emosi (emoticon) untuk mengetahui kelas sentimen yang ada dari tweet tentang toko Toko Dhedsul. Hasil penelitian menunjukkan bahwa metode Naïve Bayes tanpa penambahan fitur mampu mengklasifikasi sentimen dengan nilai akurasi sebesar 80,85%, sementara jika ditambahkan fitur pembobotan tf-idf disertai konversi ikon emosi mampu meningkatkan nilai akurasi menjadi 81%*

**Kata Kunci:** *Sentiment, Naïve Bayes Classifier.*

## PENDAHULUAN

Belakangan ini dunia usaha berkembang dengan sangat pesat. Ditandai dengan semakin banyaknya perusahaan-perusahaan, serta ide-ide untuk pengembangan suatu usaha itu sendiri agar meningkatkan laba usaha yang dikehendakinya terus berkembang serta menjadikan kepuasan tersendiri bagi konsumen yang hendak berbelanja. Perkembangan kota Tambolaka sebagai ibu kota kabupaten Sumba Barat Daya menimbulkan dampak besar dalam perubahan pusat bisnis. Perubahan yang paling jelas terlihat adalah timbulnya berbagai bisnis yang semakin banyak. Dengan meningkatnya usaha jual beli barang yang semakin pesat, maka semakin banyak pula muncul usaha jual beli barang yang ada di kota Tambolaka khususnya usaha toko. Dengan demikian meningkat pula persaingan dalam usaha tersebut. Selain timbulnya persaingan yang tinggi, pola pikir dan perilaku pelanggan juga mengalami kemajuan. Perkembangan pola pikir tersebut, misalnya pelanggan lebih cenderung dengan menilai kualitas produk, harga yang ditawarkan serta tingkat pelayanan yang diberikan customer yang ada pada toko tersebut.

Penelitian ini bertujuan untuk mengetahui sentimen pelanggan pada toko dhedsul, untuk mengetahui sentimen pelanggan tersebut bersifat negatif ataupun positif dapat dilihat dari opini pelanggan pada tingkat pelayanan, kualitas produk dan harga yang ditawarkan oleh toko dhedsul. Analisis Sentimen adalah jenis natural language yang dapat dilakukan untuk mendapatkan Opini Pelanggan. Dengan adanya analisis sentimen dapat dilakukan pengolahan kata untuk melacak mood pelanggan toko dari komentar pelanggan yang didapatkan dari hasil wawancara dan kuesioner. Opini atau informasi yang terkandung dari hasil wawancara dan kuesioner ini sangat berharga sebagai alat penentu kebijakan dalam mengklasifikasi dan ini bisa dilakukan dengan Text Mining. Text Mining adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dokumen dimana Text Mining merupakan variasi dari Data Mining yang berusaha menemukan pola menarik dari sekumpulan data tekstual yang berjumlah besar.

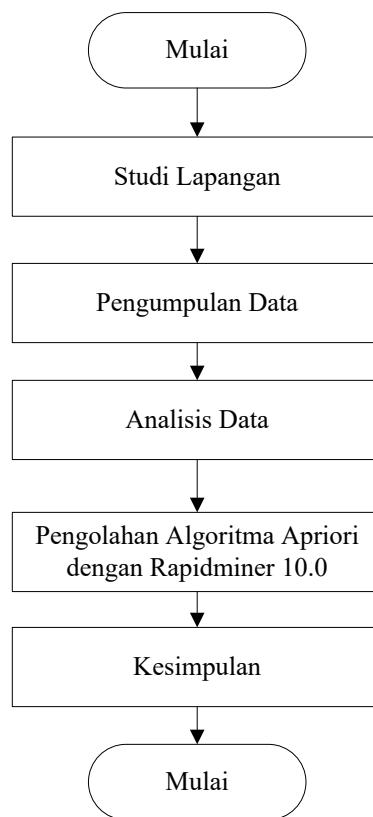
Analisa sentimen terhadap pelanggan toko telah banyak dilakukan, beberapa peneliti terdahulu diantaranya adalah penelitian yang dilakukan oleh: Kautsar & Kemas menganalisis sentimen terhadap toko online (e-commerce) lazada dan tokopedia dengan metode klasifikasi yang digunakan adalah naive bayes classifier (NBC) untuk melakukan perbandingan seberapa besar tingkat akurasi yang didapat dari hasil prediksi klasifikasi antara toko online lazada dan tokopedia, kemudian hasil yang didapatkan dari analisis sentimen tersebut ialah akurasi yang didapatkan oleh lazada sebesar 75.26% dan tokopedia sebesar 83.97%. Kemudian pada penelitian Siti Rahayu, et al melakukan analisa sentimen terhadap review marketplace online (e-commerce) shopee dengan menggunakan metode klasifikasi naive bayes classifier untuk mengetahui sentimen marketplace online (e-commerce) shopee menghasilkan tingkat akurasi sebesar 87%. Naive Bayes Classifier adalah algoritma yang sering digunakan dalam pengkategorian teks, dengan menggabungkan probabilitas kata-kata dan kategori untuk memprediksi probabilitas kategori dari sebuah dokumen.

Oleh karena itu, penelitian ini akan menawarkan solusi dalam melakukan analisis sentimen terhadap komentar pelanggan toko dhedsul di Kabupaten sumba barat daya.

Analisis sentimen ini digunakan untuk mengetahui performa algoritma Naive Bayes Classifier dalam melakukan klasifikasi berdasarkan komentar pelanggan toko dhedshul yang ada di Kabupaten sumba barat daya, sehingga mempermudah pemilik usaha dalam memantau perkembangan toko dan membantu dalam pengambilan keputusan untuk langkah apa selanjutnya yang tepat dalam mengembangkan usaha yang ada. Untuk membuat sebuah analisis sentimen perlu banyak hal dilakukan, di antaranya memilih classifier dan sumber data yang akan digunakan untuk analisis. Untuk menerapkan metode ini, peneliti menggunakan alat bantu yaitu rapidminer yang merupakan salah satu tools yang digunakan untuk pengolahan data seperti data mining ( Qisthiano et al., 2021 ) Dari alasan diatas maka penulis akan menyusun tugas akhir, yang dituangkan dalam judul: “Analisis Sentimen Pelanggan Toko Menggunakan Metode Naive Bayes Classifier” ( studi kasus : toko dhedsul

## METODE PENELITIAN

Dalam tahapan penelitian ini digunakan tahapan penelitian yang sistematis untuk membantu penelitian agar terarah dengan baik



Gambar 1. Tahapan Penelitian

Metode Pengumpulan Data Pengumpulan data dilakukan untuk memperoleh informasi yang diperlukan untuk mencapai tujuan penelitian. Pengumpulan data akan

dilakukan dengan langkah-langkah berikut:

- a. Data primer, penulis mendapatkan data yang diperoleh langsung dari lapangan yaitu melalui wawancara dan observasi.
  - 1) Wawancara, penulis mengumpulkan data dan informasi melalui wawancara dan diskusi langsung.
  - 2) observasi, penulis mengumpulkan data dengan cara merekam langsung pada objek penelitian untuk memperoleh data yang diperlukan dalam penelitian. Untuk mengetahui informasi yang dibutuhkan, peneliti mengumpulkan data laporan penjualan.
- b. Sampel Penelitian Sampel adalah bagian dari jumlah dan karakteristik populasi. Pengambilan sampel dilakukan jika populasi besar dan peneliti tidak mungkin mempelajari semua yang ada dalam populasi (Taufiqurrahman et al., 2017).

## **HASIL DAN PEMBAHASAN**

Proses data pada penelitian ini membagi data-data yang telah dilakukan proses preprocessing menjadi dua bagian yaitu data latih dan data uji. Pembuatan data latih ini bertujuan sebagai pedoman dalam menentukan hasil data uji. Data komentar yang disimpan dalam excel bisa dilihat pada gambar dibawah ini:

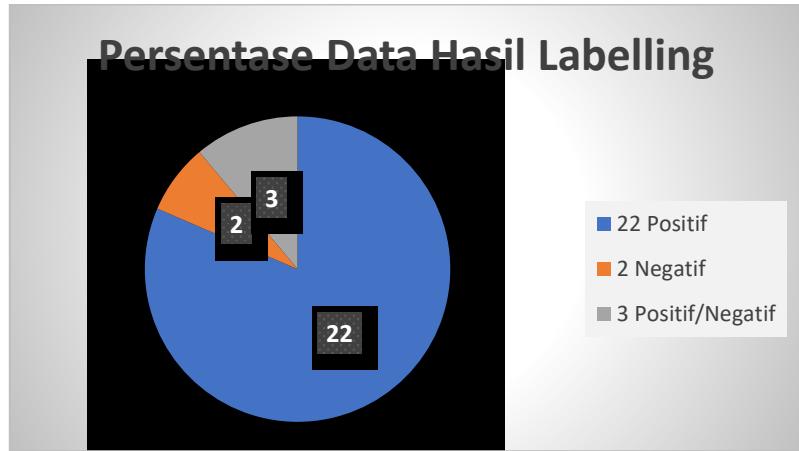
	text	sentimen
1		
2	saya puas dengan kualitas barang	positif
3	Harga terjangkau	positif
4	Pelayanan ramah	positif
5	Adanya potongan harga	positif
6	Tempatnya bisa dijangkau	positif
7	Pelayanan tidak baik	negatif
8	Barang banyak sekennya	negatif
9	saya puas dengan kualitas barang	positif
10	kualitas barangnya bagus	positif
11	Pelayanan ramah	positif
12	saya puas dengan kualitas barang	positif
13	tidak baik pelayanan	negatif
14	Tempatnya bisa dijangkau	positif
15	saya puas dengan kualitas barang	positif
16	kualitas barangnya bagus	positif
17	Adanya potongan harga	positif
18	Tempatnya bisa dijangkau	positif

Gambar 1 Contoh Data Latih

## **Percentase Data Hasil Labelling**

Pada proses pengambilan data yang dilakukan berdasarkan koesioner yang dibagikan pelanggan toko dengan menghasilkan data dengan jumlah 27 komentar. Setelah seluruh komentar berhasil disimpan selanjutnya dilakukan proses pelabelan. Dalam proses pelabelan yang dilakukan seharusnya untuk menentukan komentar yang didapatkan apakah bernilai positif, negative. Namun pada proses pelabelan pada penelitian ini terdapat beberapa komentar yang tidak sesuai dengan sentimen komentar yang sebenarnya. Dan hasil akhir dari proses pelabelan data menghasilkan data sebanyak 22 komentar bernilai positif, 2 komentar bernilai negatif dan 3 komentar bernilai positif

negatif. Persentase data bernilai positif, negatif dan positif negatif dan dapat dilihat pada gambar 2.



Gambar 2 Persentase Data Hasil Labelling

### Ekstraksi Fitur

Setelah terbentuknya file yang akan dijadikan dataset, maka selanjutnya data tersebut akan dibentuk menjadi sebuah model klasifikasi. Namun sebelum membentuk model, ada beberapa tahapan yang harus dilakukan agar terbentuknya suatu model yang baik. Yang pertama dilakukan adalah membaca file xlsx dan kemudian dilakukan tokenisasi terhadap seluruh dokumen dalam file tersebut. Berdasarkan hasil tokenisasi yang dilakukan, maka penulis juga ingin mengetahui frekuensi kata yang banyak diperbincangkan oleh pengunjung toko, untuk itu penulis memvisualisasikannya dalam bentuk wordcloud pada gambar



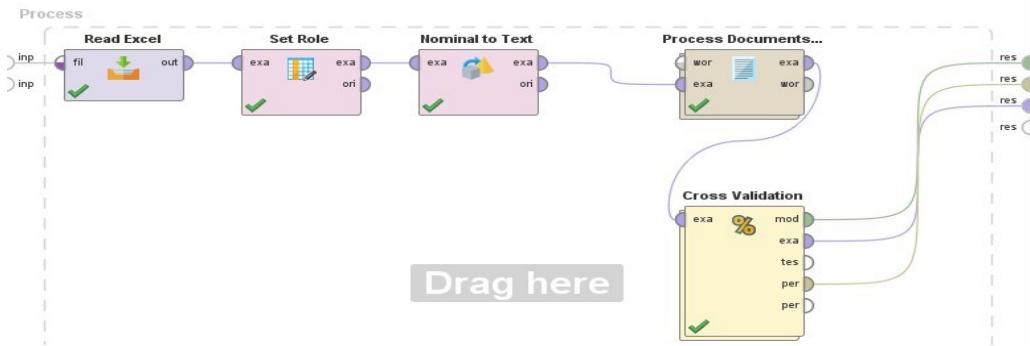
Gambar 3 Proses Document From Data

### Perhitungan Akurasi Dengan RapidMiner

Penelitian ini menggunakan tools RapidMiner versi 10.1. perhitungan akurasinya akan proses dengan menggunakan metode Naive Bayes Classifier dan Support Vektor Machine. Untuk mengetahui tingkat akurasi dari kedua metode tersebut.

### Proses Dokument dengan metode Naive Bayes Classifier

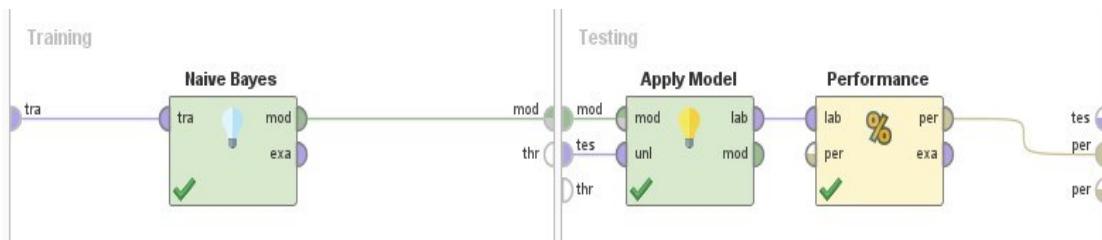
Pada proses ini diuji dengan data latih dari excel. Seperti yang terlihat pada gambar dibawah ini



Gambar 4 Proses Dokumen Data Latih

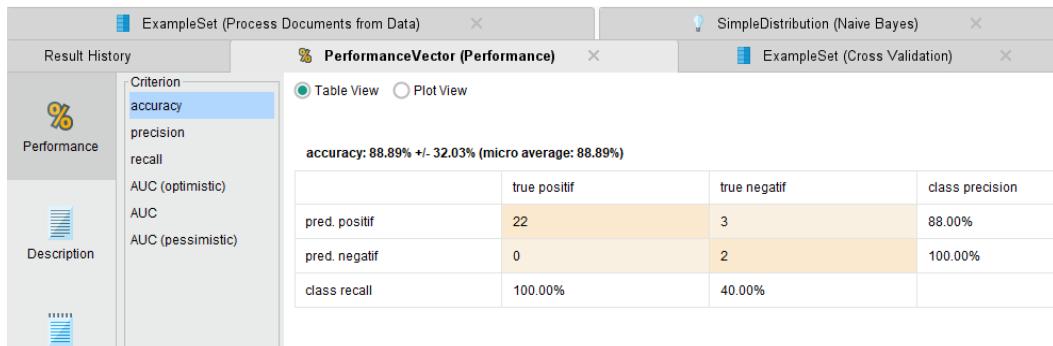
### Perhitungan Akurasi dengan Naive Bayes Classifier pada RapidMiner

Proses *Training* dan *Testing* dengan metode Naive Bayes Classifier pada RapidMiner ditampilkan pada gambar dibawah ini:



Gambar 5 Proses *Training* dan *Testing* dengan Naive Bayes Classifier pada RapidMiner

Hasil perhitungan klasifikasi menggunakan metode Naive Bayes Classifier diperoleh sebesar 73.47% seperti yang terlihat pada gambar dibawah.



Gambar 6 Akurasi dengan Metode Naive Bayes Classifier pada RapidMiner

### Hasil Pengujian Naive Bayes

Akurasi yang diuji pada tools RapidMiner dengan menggunakan metode Naive Bayes Classifier. Hasil pengujian akurasi menggunakan metode Naive Bayes Classifier sebesar 88,9%. oleh karena itu tingkat akurasi yang diperoleh dengan menggunakan metode yang diuji pada tools RapidMiner.

Tabel 3. Hasil Perhitungan Akurasi

Metode
--------

Naive Bayes Classifier	Tingkat Predikat	True Positif	True Negatif
	Pred. Positif	22	3
	Pred. Negatif	0	2
<b>Akurasi pada tools RapidMiner</b>			
Naive Bayes Classifier		88,9%	

## KESIMPULAN

Berdasarkan Hasil pengujian akurasi menggunakan metode Naive Bayes Classifier adalah: tingkat akurasi dengan metode Naive Bayes Classifier sebesar 88,9%. memiliki hasil akurasi yang cukup baik untuk menganalisis sentimen komentar atau tweet dari pelanggan. Hasil tingkat akurasi juga menunjukan adanya kepuasaan tentang pembelian barang.

## DAFTAR PUSTAKA

- Clayton, F. (2011). Coarse-and Fine-Grained Sentiment Analysis of Social Media Text.
- Fannisa, S., Fauzi, M., & Adinugroho, S. (2018). Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. [https://www.researchgate.net/publication/265564646\\_Coarse-and\\_Fine-Grained\\_Sentiment\\_Analysis\\_of\\_Social\\_Media\\_Text](https://www.researchgate.net/publication/265564646_Coarse-and_Fine-Grained_Sentiment_Analysis_of_Social_Media_Text)
- Firmansyah, R., Fauzi, M., & Afirianto, T. (2016). Sentiment Analysis pada Review Aplikasi Mobile Menggunakan Metode Naive Bayes dan Query Expansion.
- G, V., & Chandrasekaran, D. (2015). A comparative performance evaluation of neural network based approach for sentiment classification of online reviews. Journal of King Saud University. <http://repository.ub.ac.id/id/eprint/147228/>
- Haryanto, D., Muflikhah, L., & Fauzi, M. (2018). Analisis Sentimen Review Barang Berbahasa Indonesia Dengan Metode Support Vector Machine Dan Query Expansion. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/2464>
- Hidayatullah, A. F. (2014). Analisis Sentimen Dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter. Seminar Nasional Informatika 2014 (semnasIF 2014). Yogyakarta.
- Kumar, E. (2011). NATURAL LANGUAGE PROCESSING. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/2464Hidayatullah,%20A..>
- New Delhi: LK. International Publishing House Pvt, Ltd. Muslehatin, W., Ibnu, M., & Mustakim. (2017). Penerapan Naïve Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau. Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 9, 250-256. <https://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/view/3276>

Prasetyo, E. (2012). Data Mining Konsep dan Aplikasi menggunakan MATLAB. Yogyakarta: Penerbit Andi.

Susanti, N. (2016). Uji Perbandingan Akurasi Analisis Sentimen Pariwisata menggunakan Algoritma Support Vektor Machine dan Naive Bayes. Jurnal Universitas Kristen SatyaWacana.<https://repository.uksw.edu/handle/123456789/13368>