



Analisis dan Implementasi Metode Naive Bayes Pada Classification Sentimen Komentar Pelanggan

Rusmin Saragih

STMIK Kaputama, Indonesia, email: evitha12014@gmail.com

Info Artikel

Kata Kunci:
Analisis; Implementasi; Naive Bayes; Sentimen; Pelanggan

Keywords:
Analysis; Implementation; Naive Bayes; Sentiment; Customer



Lisensi: cc-by-sa

Copyright © 2023 Rusmin Saragih

Abstrak

Hotel Sambas mengalami kesulitan untuk mendapatkan makna atau kesimpulan dari keseluruhan komentar yang diberikan pelanggan terhadap produk dan layanan hotel dikarenakan banyaknya komentar yang ada, pertahun mencapai 675 komentar. Sistem analisis sentiment analysis system bertujuan untuk membantu pihak hotel dalam mendapatkan makna dari komentar yang banyak dengan menggunakan metode Naive Bayes Classifier (NBC). Metode ini mengelompokkan komentar berdasarkan kategori-kategori yang ditinjau oleh hotel. Komentar dibagi berdasarkan sentimen positif dan negatif, sehingga dapat dievaluasi kepuasan pelanggan terhadap produk dan jasa yang disediakan secara terkomputerisasi dan spesifik. Hasil dari penelitian yang dilakukan mendapatkan enam kategori yang ditinjau dengan 55 keyword kata benda, terdapat 120 keyword sentimen dengan 66 kata sentimen positif dan 54 kata sentimen negatif. Hasil pengolahan terhadap 175 data latih disimpulkan bahwa hasil klasifikasi sentimen yang didapat adalah sentimen positif sebanyak 155 komentar dan sentimen negatif sebanyak 20 komentar. Kategori sentimen positif terbesar adalah kategori kamar sebesar 73 komentar dan kategori dengan sentimen negatif terbesar adalah kategori kamar sebesar 17 komentar. Tingkat akurasi dalam penentuan kategori adalah sebesar 77.14% dan 75.42% dalam penentuan sentimen memiliki tingkat precision 99.12% dan recall 72.9%.

Abstract

Hotel Sambas has difficulty getting meaning or conclusions from all the comments given by customers regarding hotel products and services due to the large number of comments, reaching 675 comments per year. The sentiment analysis system aims to help hotels obtain meaning from many comments using the Naive Bayes Classifier (NBC) method. This method groups comments based on the categories reviewed by the hotel. Comments are divided based on positive and negative sentiments, so that customer satisfaction with the products and services provided can be evaluated in a computerized and specific manner. The results of the research carried out obtained six categories that were reviewed with 55 noun keywords, there were 120 sentiment keywords with 66 positive sentiment words and 54 negative sentiment words. The results of the processing of 175 training data concluded that the sentiment classification results obtained were 155 positive sentiments and 20 negative sentiments. The category with the largest positive sentiment is the room category with 73 comments and the category with the largest negative sentiment is the room category with 17 comments. The level of accuracy in determining categories is 77.14% and 75.42% in determining sentiment with a precision level of 99.12% and recall of 72.9%.

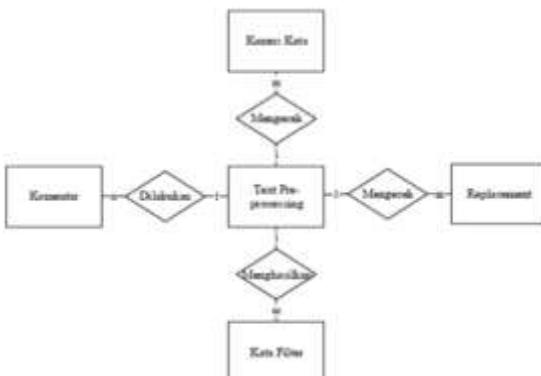
1. PENDAHULUAN

Perkembangan bisnis perhotelan kini menjadi lebih mudah dengan adanya pemesanan kamar hotel melalui perusahaan penyedia jasa pemesanan kamar hotel via online [1], [2]. Pada saat ini, hotel sambas sudah bekerja sama dengan perusahaan pemesanan kamar via online, salah satunya adalah agoda.com [3], [4]. Pengguna hotel cenderung memesan melalui agoda.com dikarenakan lebih cepat dan lebih mudah. Pengguna yang memesan melalui agoda.com diundang untuk memberikan feedback dan komentar terhadap hotel tersebut. Pelanggan yang memberikan komentar kadang tidak sesuai dengan penilaian yang diberikan [5].

Komentar yang diberikan oleh salah satu pelanggan ini jika dibaca dan dimengerti makna komentarnya merupakan komentar yang positif, tidak ada satu pun kata yang negatif dalam komentar ini, namun penilaian yang diberikan oleh pelanggan adalah 4,7 dari skala 10,0 yang terbilang kurang baik, sehingga dapat disimpulkan bahwa komentar dengan penilaian yang diberikan tidaklah sesuai atau berlawanan, dari sini dapat dilihat bahwa komentar yang positif tidak mendapat nilai yang baik atau sebaliknya. Namun jika menggunakan bantuan sentiment analysis pihak hotel dapat menganalisis dengan mengandalkan komentar yang diberikan dari pelanggan dan kesimpulan berdasarkan komentar yang diberikan berdasarkan makna komentar tidak dari nilai yang diberikan.

2. METODE PENELITIAN

ERD (*Entity Relationship Diagram*) sistem analisis sentimen yang dikembangkan terdiri dari 5 entitas yaitu: Komentar (atribut terdiri dari: id_komentar, nama_pengirim, tanggal_waktu, isi_komentar), Text Pre-processing (atribut terdiri dari: id_filter, case_folding, tokenize, replacement, filtering, kata_filter), Replacement (atribut terdiri dari: id_replacement, kata_tidak_baku, kata_baku), Kamus Kata (atribut terdiri dari: id_kamus, keyword, kata_sifat, kata_negasi), Kata Filter (atribut terdiri dari: id_filter, kata_filter) [2], [6].



Gambar 1. Entity Relationship Diagram Sistem

2.1 Analisis Sistem

Komentar-komentar yang ditujukan untuk hotel sebagai bentuk feedback atau penilaian dari masyarakat mengenai hotel [7], [8]. Komentarkomentar dari pelanggan terhadap produk dan layanan Hotel Sambas secara garis besar dapat dibedakan menjadi 2, yaitu komentar positif dan negative [9], [10]. Pihak hotel belum melakukan tindakan lebih lanjut mengenai komentar- komentar yang tersebar di situs-situs valid [11].

2.2 Analisis Sentimen

Sentiment Analysis adalah merupakan perpaduan dari data mining dan text mining, atau sebuah cara yang digunakan untuk mengolah berbagai opini yang diberikan oleh konsumen atau para pakar melalui berbagai media, mengenai sebuah produk, jasa ataupun sebuah instansi (Kustinah & Indriawati, 2017). Sentiment analysis merupakan sebuah metode yang digunakan untuk memahami, mengekstrak data opini, dan mengolah data tekstual secara otomatis untuk mendapatkan sebuah sentimen yang terkandung dalam sebuah opini (Sukawirasa et al., 2008). Pada Sentiment analysis terdiri dari 3 jenis opini, yaitu opini positif, opini negatif dan opini netral, sehingga dengan sentimen analysis perusahaan atau instansi yang terkait dapat mengetahui respon masyarakat terhadap suatu pelayanan atau produk, melalui feedback masyarakat atau pun para ahli (Hafidz, 2021), (V. A. Safitri et al., 2019).

Sentiment mengacu pada fokus topik tertentu, pernyataan pada suatu topik mungkin akan berbeda makna dengan pernyataan yang sama pada subjek yang berbeda, oleh karena itu pada beberapa penelitian, terutama pada review produk, perkerjaan didahului dengan menentukan elemen dari sebuah produk yang sedang dibicarakan sebelum memulai proses Sentiment Analysis (Celarier, n.d.), (Pinem, 2018).

2.3 Naive Bayes Classification (NBC)

Naive Bayes dikembangkan oleh Reverend Thomas Bayes pada abad ke 18. Klasifikasi dengan metode Naïve Bayes secara umum dilakukan dengan pendekatan peluang atau probabilitas [7], [8], [12].

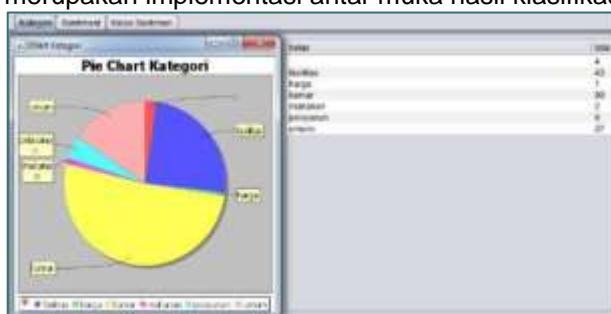
$$P(H|X) = \frac{P(X|H)}{P(X)} \times P(H) \quad (1)$$

Keterangan

X : Data dengan class yang belum diketahui
 H : Hipotesis data merupakan suatu class spesifik
 $P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)
 $P(H)$: Probabilitas hipotesis H (prior probabilitas)
 $P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H
 $P(X)$: Probabilitas X

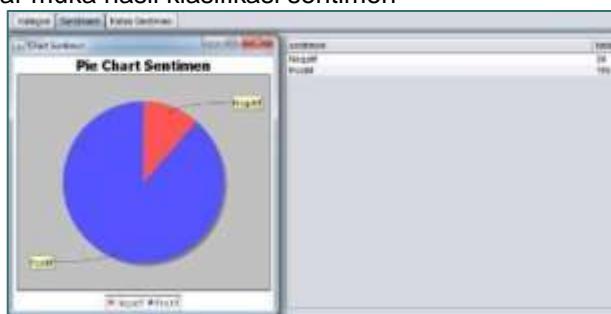
3. HASIL DAN ANALISIS

Implementasi dari perancangan yang sudah dibuat dengan bahasa pemrograman Java dan database MySql. Berikut merupakan implementasi antar muka hasil klasifikasi kategori.



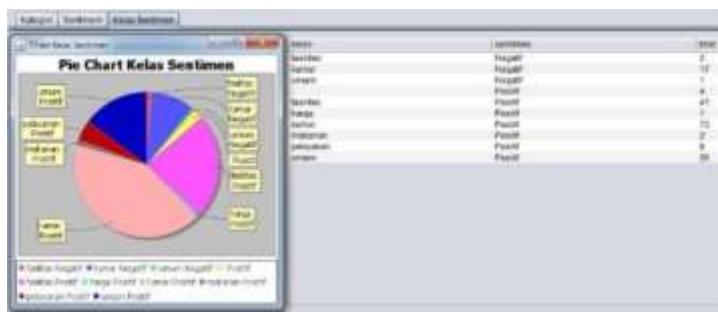
Gambar 2. Hasil Klasifikasi Kategori

Berikut implementasi antar muka hasil klasifikasi sentimen



Gambar 3. Hasil Klasifikasi Kategori

Berikut implementasi antar muka hasil klasifikasi kategori sentimen



Gambar 4. Hasil Klasifikasi Kategori Sentimen

Seluruh fungsi-fungsi yang ada pada sistem analisis sentimen ini telah melalui proses pengujian. Berdasarkan pengujian tersebut, diketahui bahwa setiap fungsi pada sistem dapat memproses data serta memberikan respon yang sesuai dengan kasus yang ada pada pengujian seperti data input yang benar, salah atau tidak lengkap. Kemampuan sistem dalam melakukan analisis dengan mengelola komentar dengan menghasilkan accuracy kategori, dan untuk sentimen precision, recall, dan accuracy.

4. KESIMPULAN

Dari hasil analisis yang dilakukan didapat 6 kategori hasil tinjauan dari data latih, yaitu kamar, makanan, pelayanan, fasilitas, harga, dan umum. Dari kategori tersebut digolongkan kembali menjadi 55 keyword. Hasil dari perhitungan sistem analisis sentimen dengan metode NBC ini dari 175 data latih maka didapat 3 hasil klasifikasi yaitu Klasifikasi kategori yang tertinggi adalah kamar sebesar 90 komentar dan yang terendah adalah untuk kategori harga sebesar 1 komentar. Klasifikasi sentimen positif sebesar 155 komentar dan sentimen negatif sebesar 20 komentar. Klasifikasi kategori dengan sentimen positif tertinggi adalah kategori kamar sebesar 73 komentar dan untuk kategori dengan sentimen negatif terendah adalah kategori kamar sebesar 17 komentar.

REFERENSI

- [1] N. Hendrastuty, "Rancang Bangun Aplikasi Monitoring Santri Berbasis Android (Studi Kasus: Pesantren Nurul Ikhwan Maros)," *Jurnal Data Mining Dan Sistem Informasi*, vol. 2, no. 2, pp. 21–34, 2021.
- [2] S. Styawati, A. R. Isnain, N. Hendrastuty, and L. Andraini, "Comparison of Support Vector Machine and Naive Bayes on Twitter Data Sentiment Analysis," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. 1, pp. 56–60, 2021.
- [3] F. Dharma, S. Shabrina, A. Noviana, M. Tahir, N. Hendrastuty, and W. Wahyono, "Prediction of Indonesian inflation rate using regression model based on genetic algorithms," *Jurnal Online Informatika*, vol. 5, no. 1, pp. 45–52, 2020.
- [4] V. A. D. Safitri and B. Anggara, "Factors that affect the company innovation," in *II. InTradersUluslararası Ticaret Kongresi Kongre Kitab The Second InTraders International Conference on International Trade Conference Book*, 2019.
- [5] M. A. Handayani, E. Suwarni, Y. Fernando, F. Fitri, F. E. Saputra, and A. Candra, "PENGELOLAAN KEUANGAN BISNIS DAN UMKM DI DESA BALAIREJO," *Suluh Abdi*, vol. 4, no. 1, pp. 1–7, 2022.
- [6] A. Heaverly and E. N. EWK, "Jane Austen's View on the Industrial Revolution in Pride and Prejudice," *Linguistics and Literature Journal*, vol. 1, no. 1, pp. 1–6, 2020.
- [7] S. A. CS, "Analisis Pengaruh Intellectual Capital Terhadap Kinerja Keuangan Perusahaan (Studi Kasus Perusahaan Sektor Keuangan Yang Terdaftar Di Bursa Efek Indonesia (BEI) Pada Tahun 2008-2017)," Universitas Gadjah Mada, 2019.
- [8] S. A. Cindiyasari and M. Aisyah, "Kinerja Keuangan Perbankan: Corporate Social Responsibility, Intellectual Capital, dan Rasio Likuiditas," *Jurnal Analisis Bisnis Ekonomi*, vol. 15, no. 2, pp. 165–175, 2017.
- [9] R. Amin and others, "IMPLEMENTASI RESTFULL API MENGGUNAKAN ARSITEKTUR MICROSERVICE UNTUK MANAJEMEN TUGAS KULIAH (STUDI KASUS: MAHASISWA STMIK AKAKOM)," STMIK AKAKOM Yogyakarta, 2020.
- [10] D. A. Hafidz, "Pengembangan Sistem Informasi Edukasi dan Pemasaran Hasil Pertanian di Tulang Bawang," 2021.
- [11] A. H. Kurniawan, "Konsep Altmetrics dalam Mengukur Faktor Dampak Artikel Melalui Academic Social Media dan Non-academic Social Media," *UNILIB: Jurnal Perpustakaan*, pp. 43–49, 2020.
- [12] R. Aditomo Mahardika Putra, "Underground Support System Determination: A Literature Review," *International Journal of Research Publications*, vol. 83, no. 1, Aug. 2021, doi: 10.47119/IJRP100831820212185.