

**ANALISIS SENTIMEN DALAM PENILAIAN KUALITAS HOTEL  
MENGUNAKAN PYTHON**

**UAS PENGOLAHAN ARTIFICIAL INTELLIGENCE**

*Disusun untuk memenuhi salah satu syarat menyelesaikan mata kuliah Artificial Intelligence  
di Semester 7*

**Disusun Oleh:**

Mutakin / 432007006190133

Rais Zainuri / 432007006190146



**TEKNIK INFORMATIKA (S-1)**

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER**

**STMIK TASIKMALAYA**

**2023**

# ANALISIS SENTIMEN DALAM PENILAIAN KUALITAS HOTEL MENGUNAKAN PYTHON

Mutakin<sup>1</sup>, Rais Zainuri<sup>2</sup>

Jl.R.E. Martadinata No.272 A Tasikmalaya, Telp. (0265) 310830

Jurusan Teknik Informatika, STMIK Tasikmalaya

e-mail: [mutakin.email@gmail.com](mailto:mutakin.email@gmail.com), [raiszainuri09@gmail.com](mailto:raiszainuri09@gmail.com)

## Abstrak

Hotel merupakan hunian sementara yang biasa dilakukan oleh warga untuk beristirahat di perjalanan jauh atau liburannya. Salah satu yang dilakukan dalam pemilihan hotel adalah dengan melihat review atau ulasan yang tertera di media online terkait hotel yang akan dipilih. Penelitian ini bertujuan untuk melakukan analisis sentimen pada penilaian kualitas hotel yang diberikan oleh pelanggan melalui media online. Analisis sentimen dilakukan dengan menggunakan library Python, seperti Natural Language Processing Toolkit (NLTK) dan TextBlob. Data yang digunakan adalah review dan rating yang diberikan pelanggan pada google map terkait pengalamannya di hotel tersebut. Data mentah yang diperoleh kemudian diolah menggunakan algoritma SVM atau Support Vector Machine, sehingga menjadi model yang dapat diimplementasikan menjadi situs. Dalam penelitian ini menghasilkan sebuah situs untuk mengetahui suatu komentar yang berkaitan dengan hotel itu positif atau negatif. Hal Ini diharapkan dapat membantu hotel dalam memahami kepuasan pelanggan dan memperbaiki layanannya untuk memenuhi harapan pelanggan.

**Kata kunci:** *hotel, textblob, svm, algoritma, review.*

## Abstract

Hotels are temporary shelters that are usually used by residents to rest on long trips or holidays. One of the things that is done in selecting a hotel is to look at the reviews or reviews listed in online media regarding the hotel to be selected. This study aims to carry out sentiment analysis on hotel quality assessments provided by customers through online media. Sentiment analysis is performed using Python libraries, such as the Natural Language Processing Toolkit (NLTK) and TextBlob. The data used are reviews and ratings given by customers on the Google Map regarding their experience at the hotel. The raw data obtained is then processed using the SVM algorithm or Support Vector Machine, so that it becomes a model that can be implemented into a website. In this research a site to find out whether a comment related to the hotel is positive or negative. This is expected to help hotels understand customer satisfaction and improve their services to meet customer expectations.

**Keyword:** *hotel, textblob, svm, algoritma, review.*

## Pendahuluan

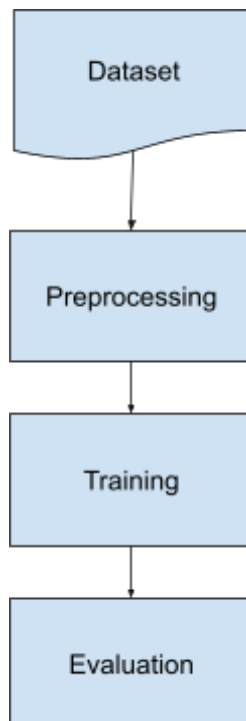
Pada era digital saat ini, kepuasan pelanggan menjadi salah satu faktor penting dalam industri perhotelan. Salah satu cara untuk mengetahui tingkat kepuasan pelanggan adalah melalui review dan rating yang diberikan pada media online seperti Google Maps. Melalui review dan rating ini, hotel dapat mengetahui kelebihan dan kekurangan yang dimiliki, sekaligus memperbaiki layanannya untuk memuaskan pelanggan. Dari sisi pelanggan, merupakan kemudahan dalam mencari informasi mengenai hotel yang akan mereka singgahi sebagai pertimbangan sebelum melakukan transaksi.

Sentiment analysis atau analisis sentimen adalah teknik untuk mengklasifikasikan dan menganalisis perasaan dan pandangan yang terkandung dalam teks. Dalam hal ini, teks yang dianalisis adalah review dan rating yang diberikan pelanggan pada hotel. Analisis sentimen dapat membantu hotel untuk mengetahui apakah review dan rating yang diterima positif, negatif, atau netral. Namun untuk analisis sentimen yang kami lakukan, klasifikasi hanya dibagi menjadi dua, yaitu positif dan negatif.

Penelitian ini bertujuan untuk melakukan analisis sentimen pada review dan rating hotel melalui sentiment analysis. Hasil dari penelitian ini diharapkan dapat membantu hotel dalam memahami kepuasan pelanggan dan memperbaiki layanannya agar sesuai dengan harapan pelanggan.

## Metode Penelitian

Penelitian ini menggunakan data yang bersumber dari Google Maps Review berbahasa Indonesia. Untuk penelitian, penulis meneliti dari Objek City Hotel. Dalam penelitian ini menggunakan kerangka sebagai berikut,



**Gambar 1.** Kerangka berpikir dalam penelitian.

Langkah-langkah kerangka adalah sebagai berikut:

1. *Business Understanding*

Proses memahami objek penelitian yang akan dilakukan untuk menghasilkan hasil yang sesuai dengan harapan. Pemahaman terhadap objek adalah suatu yang ditekankan dalam penelitian, dapat dicari dengan menggali informasi dari sumber objek.

## 2. *Data Understanding*

Proses yang dilakukan pada tahap ini adalah proses pengambilan data mentah yang diperoleh dari Google Maps Review, dan mengambil objek City Hotel pada tanggal 09 Februari 2022 dari hampir sehingga mendapatkan 1406 data yang berhasil diperoleh dengan menggunakan Scraping.

## 3. *Data Preparation (Preprocessing)*

Dilakukan preparation sebelum masuk ke tahap modeling,. Proses tersebut meliputi beberapa tahapan dari mulai mengubah case menjadi kecil, menghapus karakter dan sebagainya.

## 4. *Modeling and Training*

Algoritma yang digunakan dalam penelitian ini adalah SVM, SVM bekerja dengan membangun model yang memisahkan data menjadi kelas yang berbeda dengan menggunakan garis terpendek antara dua kelas.

## 5. *Evaluation*

Setelah model dilatih, model dapat diuji dengan menggunakan data testing. Beberapa metrik yang digunakan untuk mengevaluasi model SVM antara lain akurasi, presisi, dan recall.

# Hasil dan Pembahasan

Analisis sentimen terhadap review atau ulasan yang digunakan untuk menggali penyebab pemberian suatu opini tertentu oleh konsumen, dilakukan dalam beberapa tahapan. Dalam hal ini terdapat empat tahapan utama yang harus dilakukan, yaitu berupa tahap pengambilan data, preprocessing data awal, analisis data, pengembangan model dan *Deployment*.

## 1. Pengambilan Data

Tahap pengambilan data ini merupakan tahap inisiasi pada analisis sentimen. Secara teknis penggunaan pemrograman Python digunakan untuk mempermudah proses ini dengan cara *web scraping* menggunakan *library Selenium*. Jenis data yang akan didapatkan berupa kumpulan kalimat ulasan yang masih memiliki beragam jenis tipe. Misalkan masih terdapat tanda baca, angka, dan emoticon. Data yang akan diambil berupa nama, rating, tanggal dibuat, dan ulasan. Berikut contoh data awal yang didapatkan seperti pada Tabel 1.

Name	Rating	Created	Review
alfira candika	2/5	5 bulan lalu	Situasi kamar cukup lembab, di tembok ada bekas jamur dan bagian kasur cukup basah (lembab) terma
Nur Bayani	2/5	2 bulan lalu	Baru sekali ini ke hotel yang bebas merokok dimana2. Sampe lift nya aja bau rokok. Kamar very poor. Di
Anton Jovan	2/5	sebulan lalu	Seminggu yang lalu saya pesan kamar yang Deluxe 500.000/malam, Astaga kagetnya bukan Maen harg
widawati Dadang	1/5	3 bulan lalu	ruangan standar tidak ada yang non smoking room, bau rokok sekali.. tulisannya bintang 3, tapi tidak bi
Lionel Dessi	1/5	sebulan lalu	Service sangat buruk mengecewakan sekali bikin kapok dan gamau balik lagi fitnes di city hotel . Saya l
Khansa Ghaisani	3/5	4 bulan lalu	Kamarnya bener2 ga sesuai dgn digambar. Temboknya lembab, sprej banyak noda jd terlihat kusam. Lai
Sidharta Prabowo	2/5	sebulan lalu	Di kira fitnes sama kolam renang merupakan fasilitas untuk tamu hotel ternyata bayar, terus gak ada te
Erna Ek	4/5	2 minggu lalu	Sering ke city hotel karena member gym-nya. Gym-nya komplit ,peralatannya bagus dan terawat .setela
angga praditya	2/5	2 bulan lalu	Fasilitas kamar kurang,kamar mandi plafonnya bocor,air hangat hidup sebentar,tv byk semut ny
achmad audy	4/5	5 bulan lalu	Enak sebenarnya ini hotel, pelayannya polite, tapi pas mandi bau solokan, kirain dr belerang, taunya kai
Heroe Herlambang	2/5	5 bulan lalu	Kayaknya ni hotel dalam tahap pembenahan. Saya stay di 338. Untuk kopi dan teh tidak tersedia. Lift ta
rachmat saleh	4/5	5 bulan lalu	Over all okay tapi ada catatan utk tembok dan furniture krg diperhatikan dan 3 tahun sy pernah nginap
Dewi Moz	3/5	4 bulan lalu	Lumayan udah ngga spooky lagi nginep disini, rasa makanannya standar saja. Sarapannya ga banyak var
De Riska Nirvana	5/5	5 bulan lalu	makin bagus skrg diperluas sma dperbanyak kamarnya ya apa gmna? Uda lama bgt swimming pool nya
bunda azka	3/5	5 bulan lalu	Untuk hotel keren masi dalam penambahan bangunan jd makin keren klo dh jadi.. Makanan murah... Ak

**Tabel 1.** Data hasil scraping

Pada proses pengumpulan data awal didapatkan data kotor berupa ulasan pada objek City Hotel berjumlah 1050 ulasan. Data tersebut berasal dari hasil *scraping* website google maps menggunakan Python.

## 2. Tahap *Preprocessing* Data

Selanjutnya pada tahapan ini dilakukan pembersihan untuk menghilangkan noise, bentuk kata dan mengurangi jumlah kata. *Preprocessing* sendiri memiliki beberapa tahap.

### a. *Data Cleansing*

Pada tahap ini data kotor yang telah dikumpulkan dibersihkan dari karakter karakter yang tidak diinginkan. Dan Pemilihan Kolom yang diperlukan saja untuk tahap modeling.

### b. *Data Labeling*

Selanjutnya setelah data dibersihkan, kemudian data ulasan akan dilakukan *labeling* menggunakan *library* python yaitu TextBlob untuk mendapatkan nilai polarity. Jika nilai polarity nya lebih dari 0.0 maka sentiment ulasan tersebut positif sebaliknya jika nilai polarity nya dibawah 0.0 maka sentiment ulasan tersebut negatif. Berikut ini kode untuk *labeling*

```

crimen > data_labelling.py
import os
from textblob import TextBlob
import pandas as pd

path = './dataset/clean/review-2023-02-07.csv'

result = pd.read_csv(path)

review = result["review"].to_list()

data = {}
data['review'] = review
sentiments = []

test = 1

for dataHasil in review:
    hasilReview = TextBlob(dataHasil)

    test = test+1
    print(test)

    try:
        hasilReview = hasilReview.translate(from_lang="id", to="en")
    except Exception as e:
        print(e)

    if hasilReview.sentiment.polarity > 0.0:
        sentiments.append("positif")
    else:
        sentiments.append("negatif")

# print(sentiments)

data["sentiments"] = sentiments

df = pd.DataFrame(data)
df.to_csv("./dataset_labeled.csv")

```

**Gambar 2** Tahap *Labeling* menggunakan TextBlob

Contoh hasil labeling

	Unnamed: 0	review	sentiments
0	0	situasi kamar lembab tembok bekas jamur kasur ...	negatif
1	1	hotel bebas merokok dimana2 sampe lift nya aja...	positif
2	2	seminggu pesan kamar deluxe 500 000 malam asta...	positif
3	3	ruangan standar non smoking room bau rokok tul...	positif
4	4	service buruk mengecewakan bikin kapok gamau f...	negatif

**Gambar 3** Contoh Data Hasil Labeling

### 3. Analisis Data

Data yang telah diperoleh dari tahap labeling memiliki beberapa kolom yakni Review dan Sentiment.

Pada dataset tersebut diperoleh 2 kategori fitur yakni :

```
Jumlah Feature
positif      667
negatif      383
Name: sentiments, dtype: int64
```

**Gambar 4** Jumlah Feature masing masing kategori

#### 4. Modeling

Selanjutnya data yang telah diberi label kemudian akan diolah untuk pembuatan model yang akan digunakan untuk memprediksi sentiment dari ulasan hotel. Modeling memiliki beberapa tahapan yaitu :

1. Menghapus karakter yang tidak diperlukan  
Dengan menghapus karakter yang tidak diperlukan, maka hasil dari modeling diharapkan menjadi lebih bersih.
2. Menghapus kata kata yg tidak memiliki makna (stopwords)  
Dengan menghapus kata-kata ini, menghindari informasi tingkat rendah dari teks untuk memberi lebih banyak fokus pada informasi penting.
3. Pemisahan data train dan test
4. Tahap training.  
Tahap yang dilakukan untuk mentrain dari data menjadi sebuah model.
5. scoring akurasi  
Tahap yang dilakukan untuk meneliti hasil dari scoring training.

## Kesimpulan Dan Saran

### a. Kesimpulan

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan, dapat disimpulkan bahwa analisis sentimen menggunakan metode Algoritma SVM memiliki performa yang baik dan tinggi. Hasil tersebut deploy melalui link <https://sentiment-review-hotel.up.railway.app/>, yang menghasilkan sebuah aplikasi yang dapat mendeteksi sentimen positif atau negatif dari kalimat yang dimasukan.

### b. Saran

Pada penelitian ini masih terdapat beberapa kekurangan sehingga perlu adanya perbaikan pada penelitian selanjutnya, diantaranya adalah:

1. Memperbanyak data yang digunakan dalam pelatihan, sehingga sistem dapat mengenali lebih banyak pola data.
2. Memperbaiki kualitas sistem sehingga *friendly* bagi pengunjungnya.

## DAFTAR PUSTAKA

Angelina Puput Giovani, Ardiansyah, Tuti Haryanti, Laela Kurniawati, Windu Gata, 2020, Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi.

Didik Garbian Nugroho, Yulison Herry Chrisnanto, Agung Wahana. Analisis Sentimen Pada Jasa Ojek Online Menggunakan Metode Naïve Bayes