

## APLIKASI R-SHINY UNTUK *SENTIMENT ANALYSIS* TERHADAP ULASAN RESTORAN DI SINGAPURA MENGGUNAKAN METODE NAÏVE BAYES

Noviyanti Santoso<sup>1</sup>, Mashuri<sup>2</sup>, Anggie Calista<sup>1</sup>, Fahrila Annasiah<sup>1</sup>, Fifi Dwi Haryanti<sup>1</sup>, Kharin Octavian<sup>1</sup>, Ranto<sup>1</sup>, Yofani Kurnia Putri<sup>1</sup>

<sup>1</sup>Departemen Statistika Bisnis, Fakultas Vokasi, Institut Teknologi Sepuluh Nopember

<sup>2</sup>Departemen Teknik Mesin Industri, Fakultas Vokasi, Institut Teknologi Sepuluh Nopember

e-mail: [noviyantisantoso30@gmail.com](mailto:noviyantisantoso30@gmail.com)

**Abstrak** – Perkembangan teknologi yang pesat memberikan banyak kemudahan bagi penggunanya saat ini. Kehadirannya memberikan kemudahan akses pada banyak bidang. Seperti kemudahan dalam melakukan pemilihan restoran yang dapat dilakukan dalam genggam tangan dan dapat diakses dimana saja secara cepat dan praktis. Pemilihan sebuah restoran menjadi hal yang sangat mudah dilakukan, karena tersedia banyak rekomendasi pilihan dan ulasan dari pengunjung sebelumnya terhadap restoran tersebut, yang dapat digunakan oleh orang lain dalam menentukan restoran yang akan dikunjungi. Setiap individu pasti memiliki pandangan atau pendapat yang berbeda mengenai restoran-restoran tersebut. Pendapat dan opini dari pengguna dapat dijadikan masukan bagi restoran untuk meningkatkan kualitasnya. *Sentiment analysis* digunakan untuk memperoleh pendapat atau opini dari para pengguna suatu platform internet dari penyedia restoran tersebut. Maka dari itu akan dibuat sebuah aplikasi sentiment analysis berbasis website dengan menggunakan R-Shiny melalui program RStudio. Project ini dapat mempermudah penggunanya untuk melihat ulasan beberapa restoran secara keseluruhan, sehingga dapat menjadi dasar pertimbangan dalam pemilihan restoran. Website ini akan menunjukkan sentiment analysis terkait teks, yang berada pada kalimat maupun opini yang ditulis. Sehingga bisa melihat apakah puas atau tidak puas dengan menggunakan aplikasi ini secara instan dan mudah.

**Kata Kunci:** Sentiment Analysis, Restoran, Ulasan, R-Shiny.

### PENDAHULUAN

Kegiatan kuliner menjadi salah satu kegiatan wajib yang biasanya dilakukan ketika bepergian ke luar negeri bagi mayoritas masyarakat Indonesia. Founder JavaMifi, Andintya Maris, mengatakan bahwa berdasarkan data pengguna JavaMifi pada tahun 2019, salah satu negara Asia yang sering dikunjungi orang Indonesia adalah Singapura, setelah Jepang. Singapura dipilih dikarenakan dekat dengan Indonesia dan memiliki fasilitas yang baik, serta memiliki berbagai macam kuliner. Restoran di Singapura tersebut menyajikan kuliner tertentu, entah itu makanan secara umum atau kuliner khas pada daerah atau negara tersebut. Menurut Peraturan Menteri Perdagangan Republik Indonesia, restoran adalah kegiatan usaha yang menyediakan berbagai jenis makanan, dan atau makanan ringan, dan atau minuman dengan proses pelayanan dan penyajian kepada pelanggan di dalam satu tempat secara tetap (Prasetyo, 2017). Restoran yang berbeda-beda tersebut, bahkan meskipun dengan menu yang sama tentunya bisa memiliki perbedaan. Sehingga, bagi calon pelanggan yang mana belum pernah mengonsumsi kuliner di suatu restoran tentu kebingungan dalam hal memilih restoran, karena belum memiliki gambaran mengenai pelayanan dan rasa makanan yang ada pada setiap restoran. Maka dari itu para pengunjung membutuhkan sebuah rekomendasi terkait beberapa restoran yang ada. Di sisi lain, melalui

TripAdvisor menyediakan ulasan dari pelanggan di berbagai restoran, yang mana ulasan ini belum semua dimanfaatkan dengan baik oleh pemilik usaha tersebut. Sehingga dengan adanya perkembangan teknologi diperlukan analisis yang dapat merangkum dan mengklasifikasikan berbagai ulasan-ulasan mengenai restoran-restoran itu menjadi apakah restoran tersebut memuaskan atau tidak yaitu dengan *sentiment analysis*.

*Sentiment analysis* merupakan studi pengklasifikasian yang dilakukan pada opini beberapa orang, serta emosi yang terkandung di dalamnya. *Sentiment analysis* mengelompokkan data teks menjadi kategori positif, negatif, atau netral (Sipayung et al., 2016). Selanjutnya pengujian *sentiment analysis* juga bisa dilakukan dengan metode lainnya, yakni seperti yang telah dilakukan oleh (Solecha, 2019) yaitu *sentiment analysis* pada ulasan restoran dilakukan menggunakan metode Naïve Bayes berbasis *Particle Swarm Optimization* yang dengan ketepatan klasifikasi yang dihasilkan sebesar 88,50%. *Sentiment analysis* dilakukan menggunakan metode *Naïve Bayes Classifier* karena metode tersebut hanya memerlukan penggunaan total *data training* yang sedikit sebagai parameter dalam proses pengklasifikasian (Imron, 2019).

Project ini bertujuan untuk melakukan *sentiment analysis* menggunakan metode *Naïve Bayes Classifier* berbasis website menggunakan

aplikasi *R-shiny* terhadap ulasan dari pelanggan restoran di Singapura dengan mengklasifikasikannya ke dalam puas dan tidak puas. Sehingga diharapkan nantinya, calon pelanggan restoran dapat dengan mudah dan cepat (tanpa membaca satu per satu ulasan) dalam mengetahui bagaimana penilaian atau kepuasan orang-orang yang sebelumnya sudah pernah mengunjungi restoran tersebut, selain itu juga mereka dapat membandingkan ulasan antar beberapa restoran sehingga hal ini dapat dijadikan pertimbangan dalam memilih restoran, dan akhirnya didapatkan keputusan sesuai harapan. Selain itu juga informasi yang diberikan dapat membantu pelaku usaha restoran untuk mengembangkan usahanya dengan meningkatkan kualitas berdasarkan kemunculan kata kunci terbanyak mengenai restorannya.

## METODOLOGI

Proses penentuan rangka memperuntukkan dengan teknik *data mining* yaitu *Supervised Learning* yang dapat menjadikan susunan belajar dimana tersedia tutor (*data testing*) dan tutorial (*data training*) yang nantinya akan bekerja sebagai luaran dari proses *training* yang telah dilakukan. Selanjutnya hasil proses *data training* dan *data testing* akan menghasilkan sebuah *sentiment analysis* menggunakan metode *Naive Bayes Classifier*. *Sentiment analysis* bertujuan untuk mengklasifikasikan ulasan restoran menjadi positif dan negatif.

### 1. Melakukan Pengambilan Data

*Data testing* yang digunakan diambil dari website <https://www.tripadvisor.com>, yakni pada data restoran di Singapura. *Scraping* dilakukan pada nama restoran dan link ulasan dari restoran tersebut, dari hasil *scraping* ulasan restoran yang tercantum akan digunakan untuk menentukan kategori ulasan positif dan negatif.

### 2. Melakukan Pre-processing

*Pre-processing* data dilakukan dengan proses *cleaning* atau pembersihan data agar dapat menyederhanakan proses analisis. *Pre-processing* data berupa kegiatan menghapus penggunaan tanda baca yang tidak digunakan seperti hastag, underscore dan lain-lain. Selain itu perlu dilakukan penggantian penggantian keseluruhan huruf yang masih terdapat huruf besar menjadi huruf kecil (to-

lower) dan menghilangkan kata-kata yang tidak penting (*stopword*). Dilanjutkan dengan mengganti kata yang memiliki awalan dan akhiran menjadi kata dasar (*stemming*), melakukan lematisasi dan melakukan penggantian kata yang masih memiliki kesalahan (*typo*) menjadi kata baku (*slang word*). Setelah semua proses dilakukan, maka data akan lebih mudah untuk diolah.

### 3. Melakukan Proses Klasifikasi

*Sentiment analysis* dari model *supervised learning* akan menghasilkan klasifikasi ulasan positif dan negatif. Algoritma *Naive Bayes* merupakan teknik yang diterapkan untuk menentukan kelas dari setiap masalah atau obyek yang sudah dibagi berdasarkan setiap masalah, *naive bayes* dipilih menjadi salah satu algoritma yang banyak digunakan dalam proses *sentiment analysis* karena sederhana, cepat dan juga memiliki tingkat akurasi yang tinggi.

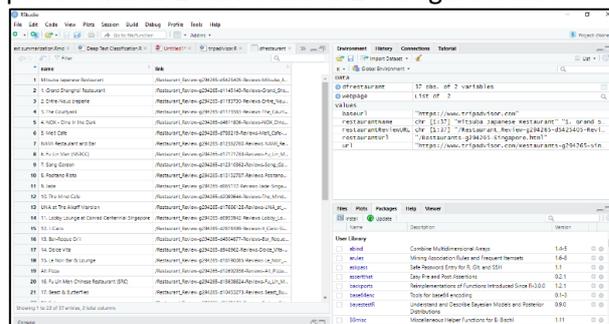
## HASIL DAN PEMBAHASAN

Algoritma pada project ini menggunakan *Naive Bayes Classifier*. Langkah-langkah dan hasil analisisnya adalah sebagai berikut

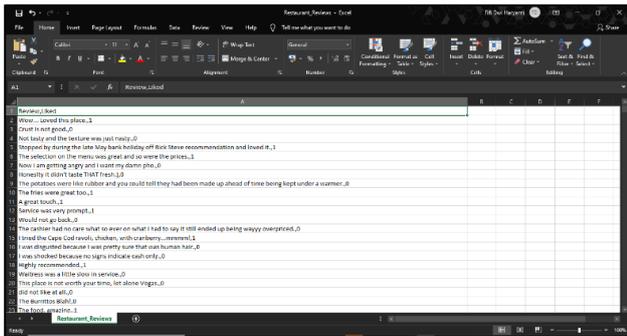
### 1. Pengumpulan dan Pengolahan Data

#### a. Pengumpulan Data dengan Metode *Scraping Website*

Dari *page* pertama data yang dihasilkan mempunyai nama restoran sejumlah 33 dengan *link ulasan* yang terdapat pada setiap nama restoran dan kemudian terdapat 1000 data pada *data training* untuk ulasan restoran. Tampilan dari *data scraping* dan juga *data training* divisualisasikan pada Gambar 1 dan Gambar 2 sebagai berikut.



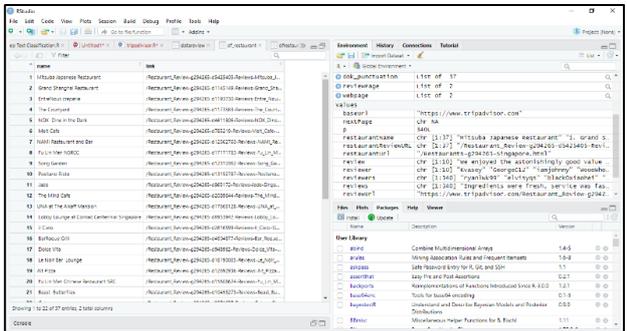
Gambar 1. Data Scraping



Gambar . Data Training

b. *Cleaning Data Scraping*

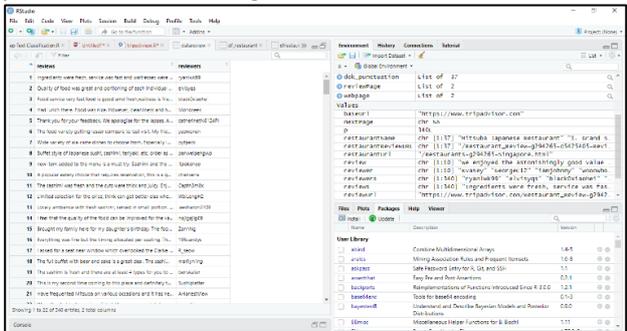
Data dari hasil *scraping* nama restoran mempunyai karakter yang bisa mempersulit proses analisis nantinya. Sehingga perlu dilakukan tindakan *cleaning*(pembersihan) data dengan menghapus karakter yang tidak diperlukan. Data hasil *scraping* nama restoran yang sudah di *cleaning* ditunjukkan pada Gambar 3.



Gambar 3. *Cleaning Data Scraping*

c. Data Ulasan dan Pemberi Ulasan Restoran

Data yang telah dicleaning kemudian akan dilakukan pengambilan data ulasan dan pemberi ulasan dari setiap restoran. Tampilan\_hasil data ulasan dan pemberi ulasan restoran ditunjukkan pada Gambar 4 sebagai berikut.

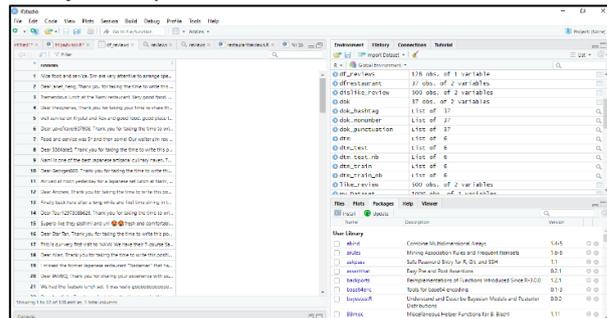


Gambar 4. Data Ulasan dan Pemberi Ulasan

d. Data Ulasan Restoran

Data ulasan dan pemberi ulasan dari setiap restoran akan diambil sejumlah 700 data yang akan digunakan sebagai data testing. Data testing ini yang akan digunakan untuk membuat *sentiment analysis* dengan menggunakan algoritma Naive Bayes

Classifier. Tampilan\_hasil data ulasan restoran ditunjukkan pada Gambar 5.

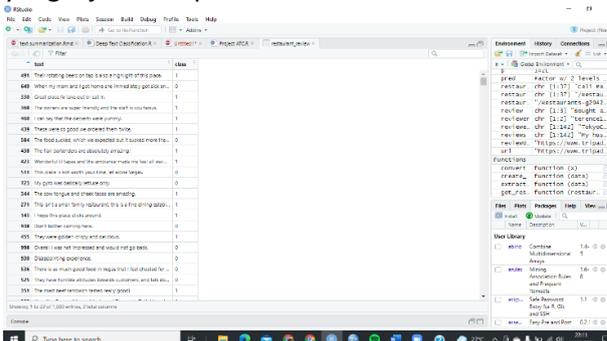


Gambar 5. Data Ulasan Restoran

2. Algoritma Naïve Bayes

a. Hasil Aktual ulasan Restoran

Pada langkah ini akan disajikan hasil dari aktual ulasan restoran, dengan menyajikan hasil aktual yang dijelaskan pada Gambar 6.



Gambar 6. Hasil Aktual

b. Hasil *Confusion Matrix* Ulasan Restoran

Langkah ini akan menunjukkan *confusion matrix* dari hasil klasifikasi ulasan restoran yang telah dilakukan sebelumnya. *Confusion matrix* ulasan restoran ditunjukkan pada Gambar 7.

```
> table(predictions, actual)
      Actual
Predictions 0  1
           0 111 41
           1  39 109
```

Gambar 7. *Confusion Matrix*

Gambar 7 menunjukkan hasil *confusion matrix* dari *ulasan* restoran, dimana 0 adalah puas dan 1 adalah tidak puas. Didapatkan puas tepat diprediksi puas sebanyak 111 dan puas salah diprediksi menjadi tidak puas sebanyak 41 Sedangkan tidak puas tepat diprediksi tidak puas sebanyak 109 dan tidak puas salah diprediksi menjadi puas sebanyak 39. Menghitung nilai sensitivitas dengan menggunakan rumus sebagai berikut.

$$\text{Sensitivitas} = \frac{TP}{TP + FN} = \frac{111}{111 + 39} = 0,74$$

Dari perhitungan diatas didapatkan nilai sensitivitas sebesar 0,74. Selanjutnya setelah



sebesar 73,3%. Nilai ketepatan dari hasil analisis dipengaruhi oleh proses pengambilan data pada *website*, dimana pada satu data pemberi ulasan memiliki dua ulasan dengan kelas yang sama.

#### **Saran**

Saran yang bisa diberikan untuk penelitian selanjutnya adalah agar memperoleh hasil klasifikasi yang lebih tinggi atau lebih maksimal, maka saat melakukan analisis sebaiknya menggunakan data yang sudah benar-benar bersih pada saat selesai proses *pre-processing* data. Saran kepada pemilik usaha restoran adalah untuk mengimplementasikan hasil *sentiment analysis* terhadap restorannya agar usaha dapat semakin berkembang dan memperbaiki kualitasnya.

#### **DAFTAR PUSTAKA**

- Imron, A. (2019). *Sentiment analysis Terhadap Tempat Wisata di Kabupaten Rembang Menggunakan Metode Naive Bayes Classifier*.
- Muthia, D. A. (2017). Sentiment analysis Pada Ulasan Restoran Dengan Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes. *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, 2(2), 39–45. <https://doi.org/10.1515/HUMOR.2006.009>.
- Prasetyo, Andjar dan Mohammad Zaenal Arifin. 2017. *Analisis Pengelolaan Limbah Makanan Restoran*. Jakarta: Indocamp.
- Sipayung, E. M., Maharani, H., & Zefanya, I. (2016). Perancangan Sistem Sentiment analysis Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Informasi*, 8(1), 958–965.
- Solecha, K. (2019). *Analisa Sentimen Dengan Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization Untuk Ulasan Restoran*. 11(1).
-