

# Analisis Sentimen Terhadap Gerai Minuman Xing Fu Tang Indonesia Menggunakan Data Ulasan Google

Arifah Jihan Fadhillah

Informatika, Fakultas Industri Kreatif, Institut Teknologi dan Bisnis Kalbis  
Jalan Pulomas Selatan Kav. 22, Jakarta Timur, 13210  
Email: 2016102811@student.kalbis.ac.id

**Abstract:** *This study aims to obtain sentiment analysis results from Xing Fu Tang beverage outlets in Indonesia. The data used in this study comes from customer-written reviews of all Xing Fu Tang outlets in Indonesia on the Google Review site. Xing Fu Tang outlets are scattered across various cities, including Jakarta, Tangerang, Bekasi, Bandung, and Surabaya. The amount of data taken manually on June 4, 2020 is 500 reviews. The study uses the Naïve Bayes Multinomial classification method through the application of the Scikit-learn module. The overall precision of the sentiment analysis model developed during the study was 78%. Result of this study is only reliable during the time of study because amount of review will be updated and the customer opinions will change over time.*

**Keywords:** *Xing Fu Tang, Google Review, Sentiment Analysis, Multinomial Naïve Bayes, Scikit-Learn*

**Abstrak:** *Penelitian ini bertujuan untuk mendapatkan hasil analisis sentimen dari gerai minuman Xing Fu Tang di Indonesia. Data yang digunakan dalam penelitian ini berasal dari ulasan-ulasan yang ditulis oleh pelanggan dari seluruh gerai Xing Fu Tang di Indonesia dalam situs Google Review. Gerai Xing Fu Tang tersebar dalam beberapa kota yaitu Jakarta, Tangerang, Bekasi, Bandung, dan Surabaya. Jumlah data yang diambil secara manual pada tanggal 4 Juni 2020 adalah 500 ulasan. Penelitian menggunakan metode klasifikasi Multinomial Naïve Bayes dengan menerapkan modul Scikit-learn. Hasil akurasi secara keseluruhan dari model analisis sentimen yang dikembangkan saat penelitian berlangsung sebesar 78%. Hasil penelitian hanya valid saat waktu penelitian berlangsung karena data ulasan yang akan selalu bertambah dan opini pelanggan akan selalu berubah seiring dengan berjalannya waktu.*

**Kata kunci:** *Xing Fu Tang, Google Review, Analisis Sentimen, Multinomial Naïve Bayes, Scikit-Learn*

## I. PENDAHULUAN

Xing Fu Tang merupakan gerai minuman yang berasal dari Taiwan dengan jenis minuman yang terkenal saat ini yaitu Brown Sugar Boba Milk (perpaduan antara susu, bola-bola tapioka dan gula yang telah terkaramelisasi) [1]. Xing Fu Tang sendiri memiliki berbagai gerai minuman yang tersebar di beberapa kota di Indonesia seperti Indonesia, Tangerang, dan

Surabaya. Xing Fu Tang terkenal ramai dengan banyaknya jumlah pelanggan yang ingin mencicipi rasa minuman yang ditawarkan oleh gerai-gerai tersebut [2].

Salah satu faktor yang mempengaruhi popularitas Xing Fu Tang adalah opini pelanggan. Faktor ini dirujuk dalam penelitian dari Tee dan kawan-kawan yang mengatakan bahwa word of mouth dan brand awareness memberikan pengaruh

yang positif dan signifikan terhadap Xing Fu Tang [3]. Hal ini menunjukkan bahwa keputusan pelanggan membeli produk Xing Fu Tang berdasarkan rekomendasi yang didapat untuk membeli produk tersebut dan pengetahuan mengenai jenis produk dan harga yang ditawarkan.

Rekomendasi yang didapat umumnya dapat berasal dari keluarga atau teman maupun dari pihak lain seperti opini pelanggan mengenai Xing Fu Tang yang tersebar melalui sosial media, portal berita dan situs ulasan restoran. Informasi yang banyak dan tersebar luas di internet akan membutuhkan banyak waktu bagi perusahaan untuk mengamati dan menganalisis opini pelanggan yang akan terus bertambah mengenai produk Xing Fu Tang. Analisis opini pelanggan ini sangat penting untuk mengetahui kualitas yang diinginkan pelanggan terhadap produk Xing Fu Tang dan membantu perusahaan menyusun strategi yang tepat agar meningkatkan minat pembelian produk.

Pembuatan model pembelajaran mesin diusulkan dalam penelitian ini untuk melakukan analisis opini atau sentimen untuk mempermudah perusahaan menarik data ulasan pelanggan dan melakukan analisis dalam waktu yang lebih singkat. Analisis sentimen sendiri adalah salah satu bidang studi yang menganalisa opini, sentimen, penilaian, pendapat dan emosi dari pelanggan terhadap suatu entitas yang diekspresikan dalam bentuk teks [4].

Tahapan yang dilakukan peneliti untuk pembuatan aplikasi dimulai dari penarikan data ulasan dari situs Google Review dan diberikan label secara manual. Data yang telah ditarik harus melalui rangkaian tahap pra-pemrosesan terlebih dahulu untuk

menghilangkan simbol dan kata-kata yang tidak mengandung sentimen ataupun kata-kata yang tidak merujuk sentimen itu berasal. Data lalu dibagi menjadi dua yaitu data latih dan data uji. Data latih dan data uji digunakan untuk pengembangan model pembelajaran mesin analisis sentimen.

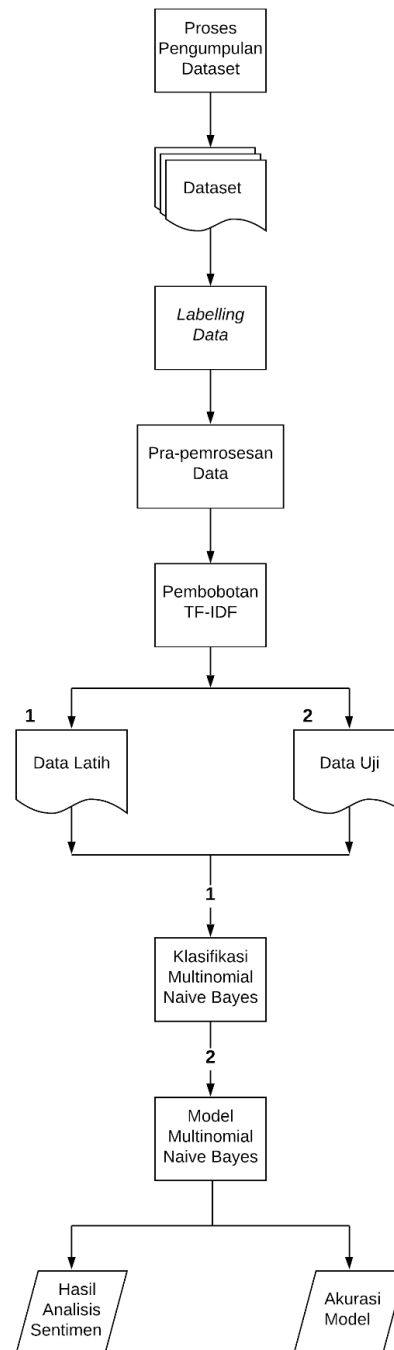
Penelitian menggunakan pendekatan pembelajaran mesin dengan menerapkan metode klasifikasi Multinomial Naïve Bayes. Pendekatan tersebut digunakan karena cocok dengan jumlah dan tipe data yang digunakan serta didukung dari penelitian-penelitian sebelumnya yang menggunakan metode klasifikasi yang sama dengan data yang satu jenis tetapi pada kasus yang berbeda menghasilkan akurasi yang tinggi.

Berbagai penelitian dengan metode Naïve Bayes serta variasinya diacu dalam penelitian ini. Salah satunya adalah penelitian dari Parasati dkk. yang melakukan analisis sentimen aspek terhadap Bakso Presiden Malang menggunakan data Google Review dan Tripadvisor [5]. Yusnitasari dkk. melakukan analisis sentimen terhadap ulasan restoran Fish Street di situs Zomato menghasilkan akurasi sebesar 87,5% [6]. Hal serupa dilakukan oleh Laksono dkk. dengan data situs TripAdvisor pada ulasan restoran terpopuler di Surabaya dengan hasil akurasi sebesar 72,06% [7]. Selain mengacu penelitian yang dilakukan terhadap ulasan restoran berbahasa Indonesia, penelitian berbahasa lain yang menggunakan metode yang sama dirujuk sebagai perbandingan untuk melihat efektivitas dari metode tersebut yaitu Sharif dkk. pada ulasan restoran berbahasa Bengali dengan akurasi sebesar 80,48% [8].

Setelah melakukan observasi mengenai pengaruh opini pelanggan terhadap penjualan produk Xing Fu Tang dan mempertimbangkan solusi yang dibutuhkan dari penelitian-penelitian sebelumnya, peneliti memutuskan untuk melakukan pengembangan model pembelajaran analisis sentimen gerai minuman Xing Fu Tang menggunakan algoritma Multinomial Naïve Bayes dengan modul Scikit-learn. Peneliti memilih menggunakan ulasan Xing Fu Tang khususnya di Indonesia pada situs Google Review karena jumlah data ulasan yang banyak dan terfokus pada gerai minuman Xing Fu Tang di Indonesia.

Manfaat utama dari penelitian ini adalah untuk mengetahui opini pelanggan Xing Fu Tang secara keseluruhan pada waktu pengambilan data. Selain itu, dari sisi akademis, penelitian ini berguna sebagai sarana edukasi untuk menunjukkan cara mengolah data teks terutama teks ulasan dimulai dari cara pengambilan data, pra-pemrosesan data, pembobotan kata, pembuatan model dan evaluasi. Sementara untuk manfaat dari segi praktis, penelitian ini dapat menghasilkan model sederhana untuk mendapatkan hasil analisis sentimen dari gerai minuman Xing Fu Tang di Indonesia.

## II. METODE PENELITIAN



Gambar 1 Skema Metode Penelitian

### A. Pengambilan Data

Data ulasan diambil dari gerai-gerai minuman Xing Fu Tang dari situs Google Review. Google Review adalah layanan gratis yang digunakan Google untuk mengumpulkan masukan berharga dari pelanggan yang pernah melakukan pembelian pada bisnis yang terdaftar di Google [9]. Google Review sendiri terdapat dalam dua layanan Google yaitu Google Maps yang dapat diakses oleh pelanggan dan Google My Business yang dapat diakses oleh pemilik bisnis. Pelanggan tidak dibayar oleh Google untuk memberikan ulasan. Google juga dapat menghapus ulasan yang dianggap spam. Data ulasan diperoleh melalui akses Google Maps [10].



Gambar 2 Contoh Ulasan Pelanggan Xing Fu Tang

Pengambilan data ulasan dilakukan secara manual. Hal ini dikarenakan keterbatasan API Google yang hanya mengizinkan pengambilan hanya lima ulasan teratas setiap gerai. Hasil pengambilan data secara manual dikumpulkan dalam bentuk CSV. Data ulasan yang telah diambil akan diberikan label terlebih dahulu secara manual. Ulasan yang mengandung opini positif akan diberi label ‘1’. Ulasan yang mengandung opini negatif akan diberi label ‘-1’. Kemudian data yang telah diberikan label tersebut akan melalui pra-pemrosesan data terlebih dahulu.

Ulasan	Sentimen
Pantesan antri, paling dekat dari daerah selatan atau depok	1
Nih bagi yg suka Ngantri minuman hits 🍷 Dulu antrinya sampe kaya uler tangan panjang bener g geboa u 🍷	1
akhirnyaaa xin fu tang buka di jakarta, tepatnya di lippon mal	1
Milk Tea Bubble yang ini beda dari yang lain, rasa nya cukup	1
Xing Fu Tang, minuman boba asal Taiwan. Brown sugar boba	1
Untuk sekarang antriannya ga terlalu rame seperti opening.	1
Ini heboh banget antrenya sampe 3 zona dari depan ke belakang	1
Ngantrinya cuma 5 menit ! Haha soalnya antrinya 21.25 WIB	1
Wawwwww bagus ada pertunjukan change Mask (ganti topi)	1
Finally, perjuangan ngantri selama berjam-jam ga sia-sia. Av	1
Antriiii bangeeeeeet luuaarbiasaaaaaaa.Tetali setelah makan	1
Bener bener enak apalagi Brown Sugar Fresh Milk nya. Bobb	1

Gambar 3 Pemberian Label Ulasan

### B. Pra-pemrosesan Data

Pra-pemrosesan data merupakan proses yang dilakukan untuk menyiapkan data mentah dan tidak terstruktur menjadi lebih mudah untuk diolah [11]. Proses ini terdiri dari gabungan beberapa tahapan agar dapat menjadikan data tersebut lebih terstruktur dan dapat dikoleksi. Berikut merupakan tahapan-tahapan yang dilakukan oleh peneliti dalam proses ini:

1. Semua huruf kapital diubah menjadi huruf kecil dan proses ini disebut *case folding* [12].
2. Pembersihan data dilakukan terutama untuk menghapus angka dan simbol yang dapat mengganggu pembelajaran mesin model. Pembersihan data adalah proses mengubah data mentah, yang diterima dari satu atau lebih penyedia, menjadi dataset yang dapat digunakan dalam visualisasi, pemodelan, dan pengambilan keputusan [13].
3. Proses *stopwords removal* dilakukan yaitu membuang kata-kata yang tidak terlalu penting untuk digunakan dalam analisis sentimen [14]. Contoh *Stopword* adalah “yang”, “ini”, “dari”. Modul python NLTK digunakan sebagai panduan *stopwords* [15].

### C. Pembobotan Kata

Tahapan berikutnya adalah proses *Term Frequency – Inverse Document Frequency* atau TF-IDF untuk data latih dan data uji yaitu pembobotan kata berdasarkan kehadiran kata rendah pada keseluruhan dokumen tetapi tinggi di dokumen tertentu. Berikut merupakan rumus dari TF-IDF:

$$TFIDF = TF \times IDF = TF \times \log_e \frac{|D|}{DF} \quad (1)$$

- TF = *term frequency*, banyaknya sebuah kata/istilah muncul.
- DF = *document frequency*, jumlah dokumen di mana kata/istilah tersebut muncul. Minimal 1 dokumen [14].

### D. Pembuatan Model

Setelah melalui data melalui tahap pembobotan kata dan telah berubah menjadi matriks bobot. Peneliti memulai pembuatan model analisis sentimen. Analisis sentimen merupakan salah satu bidang studi yang menganalisa opini, sentimen, penilaian, pendapat dan emosi dari pelanggan terhadap suatu entitas yang diekspresikan dalam bentuk teks. Entitas tersebut dapat berupa produk, jasa, organisasi, individu, peristiwa, permasalahan atau topik. Analisis sentimen sendiri dilakukan berdasarkan data yang tersebar di situs atau sosial media yang dituliskan oleh pelanggan. Kehadiran analisis sentimen mempermudah untuk pengambilan data dan jumlah data yang didapat lebih banyak karena diperoleh dari berbagai sumber [4].

Pendekatan yang digunakan dalam penelitian ini adalah metode *machine learning*. *Machine learning* merupakan salah satu pencabangan ilmu dari kecerdasan buatan dimana program komputer belajar dari pengalaman yang berhubungan dengan beberapa tugas dan ukuran performansi sehingga ketika tugas-tugas tersebut diukur dengan ukuran performansi akan meningkatkan pengalaman yang didapat [16]. Hal ini menunjukkan semakin banyak pengalaman atau pembelajaran yang dilakukan oleh komputer tersebut, maka akan meningkatkan kecerdasan dari komputer tersebut.

Machine learning terdiri dari dua kategori yaitu *supervised learning* dan *unsupervised learning*. *Supervised learning* sejak awal mengetahui *input* yang diberikan dan *output* yang diharapkan sementara *unsupervised learning* akan melakukan pembelajaran tanpa ada pengetahuan terdahulu. Dalam penelitian, peneliti menggunakan metode supervised learning karena peneliti mengetahui output yang diharapkan berupa sentimen positif, negatif dan netral. Salah satu bentuk supervised learning adalah metode klasifikasi Multinomial Naïve Bayes.

Data dilatih menggunakan metode klasifikasi Multinomial Naïve Bayes yaitu salah satu jenis Naïve Bayes berdasarkan distribusi fitur yang multinomial. Multinomial Naïve Bayes umumnya digunakan pada klasifikasi dokumen. Dokumen dapat dilihat sebagai suatu tas yang berisikan berbagai kata dengan kemungkinan kehadiran kata-kata tersebut dapat berulang. Frekuensi kata-kata ini dapat difasilitasi oleh Multinomial Naïve

Bayes. Rumus Multinomial Naïve Bayes dapat dipaparkan pada (2).

$$P(E|H) = N! \times \prod_{i=1}^k \frac{P_i^{n_i}}{n_i!} \quad (2)$$

Di mana probabilitas dokumen E terdapat dalam kelas H adalah jumlah kata dalam dokumen (N!) dikali dengan pembagian dari probabilitas kata i dalam dokumen (Pi) dan jumlah kata i dalam dokumen (ni) [17].

Sebelum pelatihan data, data dibagi menjadi data latih dan data uji terlebih dahulu. Teknik validasi yang digunakan adalah Hold-Out dengan porsi pembagian data 80/20 [18]. Peneliti menerapkan metode Naïve Bayes menggunakan Scikit-learn yaitu modul Python untuk pembelajaran mesin yang dibangun di atas SciPy dan didistribusikan di bawah lisensi 3-Clause BSD [19].

Hasil pengembangan model tersebut ditampilkan dalam bentuk aplikasi. Peneliti menggunakan bantuan modul PyQt untuk membuat tampilan aplikasi [20]. Grafik yang ditampilkan dalam aplikasi menggunakan bantuan modul Matplotlib [21].

**E. Evaluasi**

Hasil dari klasifikasi analisis sentimen dinyatakan dalam bentuk ukuran evaluasi dan ditampilkan dalam bentuk tabel Confusion Matrix sebagai berikut [18]:

- TP adalah *true positives* yaitu jumlah ulasan positif yang diklasifikasikan dengan benar.
- TN adalah *true negatives* yaitu jumlah ulasan negatif yang diklasifikasikan dengan benar.
- FP adalah *false positives* yaitu jumlah ulasan positif yang salah

diklasifikasikan oleh model analisis sentimen.

- FN adalah *false negatives* yaitu jumlah ulasan negatif yang salah diklasifikasikan oleh model analisis sentimen.

Tabel 1 *Confusion Matrix*

	Ya Prediksi	Tidak Prediksi	Jumlah
Ya Aktual	TP	TN	P
Tidak Aktual	FP	FN	N
Jumlah	P'	N'	P+N

Dari tabel Confusion Matrix, peneliti dapat menghitung nilai akurasi yaitu presentase dari jumlah ulasan yang diklasifikasikan dengan benar dalam data uji oleh model analisis sentimen. Berikut merupakan rumus dari akurasi:

$$\frac{TP+TN}{P+N} \quad (3)$$

**III. HASIL DAN PEMBAHASAN**

Penelitian menghasilkan 500 data ulasan dari berbagai ulasan gerai Xing Fu Tang di Google Maps. Ulasan tersebut kemudian diberikan label secara manual. Hasil ulasan yang telah diberikan label tersebut disimpan dalam bentuk CSV. Ketika data diunggah, peneliti mendapatkan informasi berupa jumlah ulasan yang positif sebesar 318 dan negatif sebesar 182.

```
Jumlah Sentimen
1      318
-1     182
Name: Sentimen, dtype: int64
```

Gambar 4 Junlah Sentimen

Tahap pra-pemrosesan data yang dilakukan selanjutnya adalah mengubah huruf kapital menjadi huruf kecil serta menghapus angka dan simbol yang tidak penting. Gambar 5 merupakan hasil pembersihan data.

```
Hasil Cleaning
0      pantesan antri paling deket dari daerah selata...
1      nih bagi yg suka ngantri minuman hits dulu ant...
2      akhirnyaaa xin fu tang buka di jakarta tepatny...
3      milk tea bubble yang ini beda dari yang lain r...
4      xing fu tang minuman boba asal taiwan brown su...
...
495     akhirnya masuk indonesia nih bubble tea asli t...
496     akhirnya dapatkan kesempatan untuk mencoba tok...
497     rasa si lumayan b aja bs jd yg antri cuman kr...
498     boba nya biasa ajae
499     boba ini sangat terkenal di kalangan remaja te...
Name: Ulasan, Length: 500, dtype: object
```

Gambar 5 Hasil Pembersihan Data

Proses *stopword removal* merupakan salah satu bagian dalam parameter pembobotan kata. Gambar 6 merupakan hasil pembobotan kata.

```
Hasil TF-IDF
[[0.22202756 0.05369857 0. ... 0.21309613 0.13546712 0.
 [0. ... 0.31387022 ... 0. ... 0. ... 0.27357802
 [0. ... 0.1773134 ... 0. ... 0. ... 0.1545513
 ...
 [0. ... 0.44428925 ... 0. ... 0. ... 0.38725487
 [0. ... 0. ... 0. ... 0. ... 0.
 [0. ... 0. ... 0. ... 0. ... 0.]
```

Gambar 6 Hasil Pembobotan Kata

Pelatihan dan pengujian data dilakukan menggunakan data yang telah dibobotkan. Hasil pengujian ditampilkan dalam bentuk *confusion matrix* dengan ulasan positif yang diklasifikasikan dengan benar berjumlah 57 dan ulasan negatif yang diklasifikasikan dengan benar berjumlah 21.

```
Hasil Confusion Matrix
[[21 18]
 [ 4 57]]
```

Gambar 7 Hasil Pengujian

Penelitian juga mendapatkan hasil yang lebih rinci mengenai akurasi model yang ditampilkan dalam bentuk *classification report*.

```
Hasil Klasifikasi
precision    recall  f1-score   support

-1         0.84         0.54         0.66         39
 1         0.76         0.93         0.84         61

accuracy          0.78
macro avg         0.80         0.74         0.75         100
weighted avg     0.79         0.78         0.77         100
```

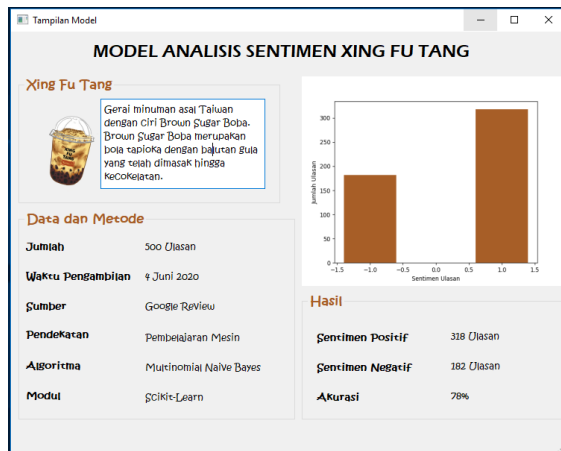
Gambar 8 Rincian Klasifikasi

Setelah melakukan pelatihan dan pengujian, peneliti mendapatkan akurasi keseluruhan dari model yang telah dibuat yaitu sebesar 78%.

```
Hasil Akurasi
0.78
```

Gambar 9 Hasil Akurasi

Hasil pembuatan model tersebut diaplikasikan dalam bentuk tampilan. Hal ini untuk mempermudah bagi pengguna untuk mendapatkan informasi mengenai penelitian secara singkat. Peneliti menampilkan penjelasan singkat mengenai Xing Fu Tang, data dan metode yang digunakan, grafik sentimen serta hasil penelitian yang berupa jumlah sentimen dan akurasi model.



Gambar 10 Tampilan Aplikasi

#### IV. SIMPULAN

Berdasarkan hasil analisis, percobaan, dan pengembangan model analisis sentimen yang telah dilakukan, maka peneliti merumuskan kesimpulan sebagai berikut:

1. Jumlah ulasan sentimen positif yang didapatkan saat penelitian berlangsung lebih banyak yaitu 318 ulasan dibandingkan sentimen negatif yaitu 182 ulasan.
2. Analisis sentimen berhasil dilakukan dengan metode klasifikasi Multinomial Naïve Bayes yang dibantu oleh modul Scikit-learn dengan akurasi 78% dimana dari 100 ulasan yang diujikan 78 berhasil diklasifikasikan dengan benar (57 ulasan positif dan 21 ulasan negatif).
3. Hasil penelitian hanya valid saat waktu penelitian berlangsung karena data ulasan yang akan selalu bertambah dan opini pelanggan akan selalu berubah seiring dengan berjalannya waktu.

#### DAFTAR RUJUKAN

- [1] PT Pelepas Dahaga, "About Us." [Daring]. Tersedia pada: <https://www.xingfutang.co.id/about>. [Diakses: 16-Mar-2020].
- [2] L. Mustinda, "Xing Fu Tang: Antre 2 Jam Untuk Cicipi Brown Sugar Boba Milk," *detikcom*, 19-Jun-2019.
- [3] J. Tee, R. Erdiansyah, dan S. Aulia, "Pengaruh Word Of Mouth dan Brand Awareness terhadap Keputusan Pembelian Xing Fu Tang di Indonesia," *Prologia*, vol. 4, no. 1, hlm. 18–24, 2020.
- [4] B. Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*. 2015.
- [5] W. Parasati, F. A. Bachtiar, dan N. Y. Setiawan, "Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 4, 2020.
- [6] T. Yusnitasari, D. Ikasari, E. E. S. Pratiwi, dan N. S. Ramdani, "Analisis Sentimen Terhadap Review Restoran Fish Streat pada Aplikasi Zomato Menggunakan Stemming Nazief Adriani dan Naive Bayes Classifier," *Pros. Setrinov*, vol. 3, hlm. 163–174, 2017.
- [7] R. A. Laksono, "Sentiment Analysis of Restaurant Customer Reviews on TripAdvisor using Naive Bayes," *2019 12th Int. Conf. Inf. Commun. Technol. Syst.*, hlm. 49–54, 2019.
- [8] O. Sharif, M. M. Hoque, dan E. Hossain, "Sentiment Analysis of Bengali Texts on Online Restaurant Reviews Using Multinomial Naïve Bayes," *2019 1st Int. Conf. Adv. Sci. Eng. Robot. Technol.*, vol. 2019, no. Icasert, hlm. 1–6, 2019.
- [9] Google, "Mengumpulkan ulasan melalui Google Customer Review," 2020. [Daring]. Tersedia pada: <https://support.google.com/merchants/answer/7548675?hl=id>. [Diakses: 06-Jun-2020].
- [10] Google, "Xing Fu Tang Maps," 2020. [Daring]. Tersedia pada: <https://www.google.com/maps/search/Xing+Fu+Tang/@->



- 2.6552272,101.1842874,6z/data=!3m1!4b1 . [Diakses: 04-Jun-2020].
- [11] R. Feldman dan J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. New York: Cambridge University Press, 2009.
- [12] C. D. Manning, P. Raghavan, dan H. Schütze, *Introduction to Information Retrieval*. 2008.
- [13] S. E. Buttrey dan L. R. Whitaker, *A Data Scientist's Guide to Acquiring, Cleaning, and Managing Data in R*. .
- [14] O. W. Purbo, *Text Mining Analisis Medsos, Kekuatan Brand & Intelejen di Internet*, 1 ed. Yogyakarta: Penerbit ANDI, 2019.
- [15] S. Bird, E. Loper, dan E. Klein, *Natural Language ToolKit (NLTK) Book*. 2009.
- [16] B. Santoso, A. Azis, dan Zohrahayaty, *Machine Learning & Reasoning Fuzzy Logic Algoritma, Manual, Matlab, & Rapid Miner*. Deepublish Publisher, 2020.
- [17] I. H. Witten, E. Frank, dan M. A. Hall, *Data Mining (Fourth Edition)*. 2017.
- [18] Suyanto, *Machine Learning Tingkat Dasar dan Lanjut*. Bandung: Penerbit Informatika, 2018.
- [19] Pedregosa et al, "Scikit-Learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, hlm. 2825–2830, 2011.
- [20] Riverbank Computing Limited, "What is PyQt?" [Daring]. Tersedia pada: <https://riverbankcomputing.com/software/pyqt/intro>. [Diakses: 13-Apr-2020].
- [21] J. D. Hunter, "Matplotlib: A 2D graphics environment," *Comput. Sci. Eng.*, vol. 9, no. 3, hlm. 90–95, 2007, doi: 10.1109/MCSE.2007.55.