

**ANALISIS SENTIMEN PADA MEDIA SOSIAL *TWITTER*
TERHADAP PENGGUNAAN KENDARAAN LISTRIK
DENGAN MENGGUNAKAN METODE *NAÏVE BAYES*
CLASSIFIER DAN *INFORMATION GAIN***

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-I Pada Jurusan Teknik Informatika



Oleh:

Ivando Sibarani

NIM: 09021282025046

**JURUSAN TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024**

LEMBAR PENGESAHAN SKRIPSI

**ANALISIS SENTIMEN PADA MEDIA SOSIAL *TWITTER*
TERHADAP PENGGUNAAN KENDARAAN LISTRIK
DENGAN MENGGUNAKAN METODE *NAÏVE BAYES*
CLASSIFIER DAN *INFORMATION GAIN***

Oleh :

Ivando Sibarani

NIM: 09021282025046

Inderalaya, 8 Oktober 2024

Pembimbing I

Pembimbing II

Rizki Kurniati, S.Kom, M.T.

NIP 199107122019032016

Desty Rodiah, S.Kom, M.T.

NIP 198912212020122011

Mengetahui

Dekan Fakultas Ilmu Komputer



Hadipurnawan Satria, Ph.D.

NIP 198004182020121001

TANDA LULUS UJIAN SIDANG SKRIPSI

Pada hari Jum'at tanggal 13 September 2024 telah dilaksanakan ujian sidang skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Ivando Sibarani
NIM : 09021282025046
Judul : Analisis Sentimen pada Media Sosial Twitter terhadap Penggunaan Kendaraan Listrik dengan menggunakan Metode *Naïve Bayes Classifier* dan *Information gain*
dan dinyatakan LULUS

1. Ketua Penguji

Yunita, S.Si., M.Cs.
NIP 198306062015042002

2. Pembimbing 1

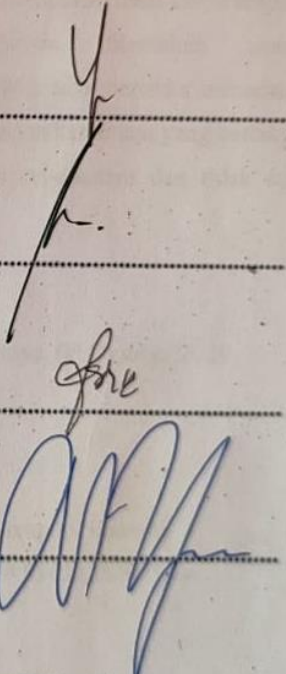
Rizki Kurniati, S.Kom., M.T.
NIP 199107122019032016

3. Pembimbing 2

Desty Rodiah, S.Kom., M.T.
NIP 198912212020122011

4. Penguji 1

Novi Yusliani, S.Kom., M.T.
NIP 198211082012122001



Mengetahui,

Ketua Jurusan Teknik Informatika



Hadipurnawan Satria, Ph.D.

NIP 198004182020121001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Ivando Sbarani
NIM : 09021282025046
Program Studi : Teknik Informatika
Judul : Analisis Sentimen pada Media Sosial Twitter terhadap Penggunaan
Kendaraan Listrik dengan menggunakan Metode *Naïve Bayes Classifier* dan *Information gain*

Hasil Pengecekan Turnitin : 9%

Menyatakan bahwa Laporan Penelitian saya merupakan hasil karya sendiri dan bukan hasil penjiplakkan/plagiat. Apabila ditemukan unsur penjiplakkan/plagiat dalam laporan penelitian ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Indralaya, 08 Oktober 2024



Ivando Sibarani

NIM 09021282025046

Motto dan Persembahan

“Hidup Memang Tidak Adil, Jadi Biasakan Dirimu”

Kupersembahkan Karya Tulis ini Kepada:

- Orang Tua
- Keluarga Besar
- Teman-teman TI REG 2020
- Para Dosen dan staff Fakultas Ilmu
Komputer Universitas Sriwijaya

ABSTRACT

The use of electric vehicles is increasing worldwide as a solution to reduce carbon emissions and dependence on fossil fuels. However, the adoption of this technology faces various challenges, including public acceptance. Given the large number of opinions on Twitter, it is important to classify sentiment to facilitate the identification of opinions on electric vehicles. This study aims to analyze public sentiment towards the use of electric vehicles through data from Twitter. This study uses 4174 tweet data with 1903 positive sentiments, 1603 negative sentiments, and 668 neutral sentiments. The method used in this study is the Naïve Bayes Classifier and Information gain feature selection. The results of the study with the Naïve Bayes Classifier model applied with Information gain obtained the highest accuracy, precision, recall, and f1-score at a ratio of 80:20 of 61.77%, 40.1%, 47.8%, and 43.13%, with a training time of 0.0054 seconds. Meanwhile, by only using the naive bayes method, the values obtained were 60.2%, 40.75%, 46.7% and 42.55%, with a training time of 0.0061 seconds.

Keywords: sentiment analysis, electric vehicles, Twitter, Naïve Bayes Classifier, Information gain

ABSTRAK

Penggunaan kendaraan listrik semakin meningkat di seluruh dunia sebagai solusi mengurangi emisi karbon dan ketergantungan pada bahan bakar fosil. Namun, adopsi teknologi ini menghadapi berbagai tantangan, termasuk penerimaan dari masyarakat. Melihat banyaknya opini yang ada di Twitter, penting untuk melakukan klasifikasi sentimen untuk memudahkan identifikasi opini terhadap kendaraan listrik. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap penggunaan kendaraan listrik melalui data dari Twitter. Penelitian ini menggunakan data *tweet* sebanyak 4174 dengan 1903 sentimen positif, 1603 sentimen negatif, dan 668 sentimen netral. Metode yang digunakan dalam penelitian ini adalah *Naïve Bayes Classifier* dan seleksi fitur *Information gain*. Hasil penelitian dengan model *Naïve Bayes Classifier* yang diterapkan dengan *Information gain* mendapatkan *accuracy*, *precision*, *recall*, dan *f1-score* tertinggi pada rasio 80:20 sebesar 61.77%, 40.1%, 47.8%, dan 43.13%, dengan waktu latih selama 0.0054 detik. Sedangkan dengan hanya menggunakan metode *naive bayes* menfapatkan nilai sebesar 60.2%, 40.75%, 46.7% dan 42.55%, dengan waktu latih selama 0.0061 detik.

Kata kunci : analisis sentimen, kendaraan listrik, Twitter, *Naïve Bayes Classifier*, *Information gain*

KATA PENGANTAR

Puji dan Syukur kehadirat Allah SWT atas rahmat, nikmat, dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul **“Analisis Sentimen Pada Media Sosial Twitter Terhadap Penggunaan Kendaraan Listrik Dengan Menggunakan Metode *Naïve Bayes Classifier* Dan *Information gain*”**. Tugas Akhir ini disusun dalam rangka memenuhi salah satu syarat dalam menyelesaikan pendidikan program Strata-1 Program Studi Teknik Informatika Fakultas Ilmu Komputer, Universitas Sriwijaya.

Penulis ingin mengucapkan terima kasih kepada seluruh pihak yang telah memberikan dukungan dan bantuan baik dalam segi materil maupun moril selama proses pembuatan tugas akhir ini. Adapun penulis merujuk secara khusus pihak yang telah membantu sebagai berikut :

1. Kedua orang tua saya dan kedua kakak saya, yang selalu memberikan dukungan, bantuan serta kritik dan saran.
2. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Hadipurnawan Satria, Ph.D. selaku Ketua Jurusan Teknik Informatika.
4. Ibu Rizki Kurniati S.Kom., M.T. selaku pembimbing akademik yang telah memberikan bimbingan serta arahan dalam urusan akademik.
5. Ibu Rizki Kurniati S.Kom., M.T. dan ibu Desty Rodiah, S.Kom., M.T. selaku dosen pembimbing I dan dosen pembimbing II yang telah memberikan arahan, bimbingan, kritik, dan saran kepada penulis sehingga dapat menyelesaikan tugas akhir.
6. Seluruh Bapak dan Ibu Dosen Program Studi Teknik Informatika yang telah memberikan ilmu kepada penulis selama perkuliahan.
7. Staff Jurusan Teknik Infomatika beserta staff Fakultas Ilmu Komputer yang telah membantu dalam urusan administrasi dan akademik selama masa perkuliahan.
8. Teman - temanku “ERSIZ” yang yang selalu memberikan semangat, masukan, serta hiburan selama perkuliahan.

9. Teman - teman kelas Reguler angkatan 2020 yang selalu membagikan informasi dan menciptakan banyak momen kebersamaan selama perkuliahan yang tentu saja tidak akan pernah dilupakan.
10. Semua pihak lainnya yang tidak dapat disebutkan satu persatu yang telah berperan bagi penulis dalam menyelesaikan tugas akhir, penulis ucapkan Terima kasih banyak atas semuanya.

Semoga tugas akhir ini dapat menjadi manfaat bagi penulis mau pun pembaca pada umumnya, serta dapat menjadi referensi dan rujukan untuk peneliti-peneliti selanjutnya. Penulis menyadari bahwa dalam proses penyelesaian tugas akhir ini, masih terdapat beberapa kekurangan, Oleh karena itu, penulis berharap akan kritik dan saran yang membangun agar tugas akhir ini menjadi karya tulis yang lebih baik agar dapat menjadi bekal pengetahuan yang bermanfaat bagi peneliti di masa depan.

Indralaya, 08 Oktober 2024

Ivando Sibarani

DAFTAR ISI

Motto dan Persembahan	i
<i>ABSTRACT</i>	ii
ABSTRAK	iii
KATA PENGANTAR	iv
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR	ix
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah	I-3
1.4 Tujuan Penelitian	I-4
1.5 Manfaat penelitian	I-4
1.6 Batasan Masalah	I-5
1.7 Sistematika penulisan	I-5
1.8 Kesimpulan	I-6
BAB II LANDASAN TEORI	II-Error! Bookmark not defined.
2.1 Pendahuluan	II-Error! Bookmark not defined.
2.2 Landasan Teori	II-Error! Bookmark not defined.
2.2.1 Text Mining	II-Error! Bookmark not defined.
2.2.2 Text Preprocessing	II-Error! Bookmark not defined.
2.2.3 Analisis Sentimen	II-Error! Bookmark not defined.
2.2.4 TF-IDF	II-Error! Bookmark not defined.
2.2.5 Seleksi fitur	II-Error! Bookmark not defined.
2.2.6 Information gain	II-Error! Bookmark not defined.
2.2.7 Naive bayes Classifier	II-Error! Bookmark not defined.
2.2.8 Evaluasi	II-Error! Bookmark not defined.
2.2.9 Metode Waterfall	II-Error! Bookmark not defined.
2.3 Penelitian yang relevan	II-Error! Bookmark not defined.
2.4 Kesimpulan	II-Error! Bookmark not defined.

BAB III METODOLOGI PENELITIAN	III-Error! Bookmark not defined.
3.1 Pendahuluan	III-Error! Bookmark not defined.
3.2 Deskripsi Umum	III-Error! Bookmark not defined.
3.3 Dataset	III-Error! Bookmark not defined.
3.4 Pengumpulan data	III-Error! Bookmark not defined.
3.5 Preprocessing Data	III-Error! Bookmark not defined.
3.6 TF-IDF	III-Error! Bookmark not defined.
3.7 Information gain	III-Error! Bookmark not defined.
3.8 Naïve Bayes Classifier	III-Error! Bookmark not defined.
3.9 Pengembangan Perangkat Lunak...	III-Error! Bookmark not defined.
3.10 Menejemen Proyek Penelitian.....	III-Error! Bookmark not defined.
3.11 Kesimpulan	III-Error! Bookmark not defined.
BAB IV PENGEMBANGAN PERANGKAT LUNAK	IV-Error! Bookmark not defined.
4.1 Pendahuluan	IV-Error! Bookmark not defined.
4.2 Metode Pengembangan Perangkat Lunak...	IV-Error! Bookmark not defined.
4.2.1 Requirement	IV-Error! Bookmark not defined.
4.2.2 System Design	IV-Error! Bookmark not defined.
4.2.3 Coding and Testing	IV-Error! Bookmark not defined.
4.2.4 Implementasi	IV-Error! Bookmark not defined.
4.2.5 Maintenance.....	IV-Error! Bookmark not defined.
4.3 Kesimpulan	IV-Error! Bookmark not defined.
BAB V HASIL DAN PEMBAHASAN	V-Error! Bookmark not defined.
5.1 Pendahuluan	V-Error! Bookmark not defined.
5.2 Data Hasil Pengujian.....	V-Error! Bookmark not defined.
5.3 Analisis Data Hasil Pengujian	V-Error! Bookmark not defined.
5.4 Kesimpulan	V-Error! Bookmark not defined.
BAB VI KESIMPULAN DAN SARAN.....	VI-Error! Bookmark not defined.
6.1 Pendahuluan	VI-Error! Bookmark not defined.
6.2 Kesimpulan	VI-Error! Bookmark not defined.
6.3 Saran.....	VI-Error! Bookmark not defined.

DAFTAR PUSTAKAi

DAFTAR TABEL

Tabel II- 1 Tabel Confusion Matrix	II-Error! Bookmark not defined.
Tabel III- 1 Dataset	III-Error! Bookmark not defined.
Tabel III- 2 Case Folding	III-Error! Bookmark not defined.
Tabel III- 3 Cleaning	III-Error! Bookmark not defined.
Tabel III- 4 Tokenizing	III-Error! Bookmark not defined.
Tabel III- 5 Normalisasi	III-Error! Bookmark not defined.
Tabel III- 6 Stopword Removal	III-Error! Bookmark not defined.
Tabel III- 7 Stemming	III-Error! Bookmark not defined.
Tabel III- 8 Tabel TF.....	III-Error! Bookmark not defined.
Tabel III- 9 Tabel DF	III-Error! Bookmark not defined.
Tabel III- 10 TF-IDF.....	III-Error! Bookmark not defined.
Tabel III- 11 Data Training	III-Error! Bookmark not defined.
Tabel III- 12 Perhitungan Entropy data	III-Error! Bookmark not defined.
Tabel III- 13 Perhitungan Entropy Kata.....	III-Error! Bookmark not defined.
Tabel III- 14 Perhitungan Entropy negasi Kata	III-Error! Bookmark not defined.
Tabel III- 15 Nilai Gain.....	III-Error! Bookmark not defined.
Tabel III- 16 Nilai n	III-Error! Bookmark not defined.
Tabel III- 17 Nilai n1	III-Error! Bookmark not defined.
Tabel III- 18 Nilai $P(a_i v_j)$	III-Error! Bookmark not defined.
Tabel III- 19 Data Testing Naïve Bayes	III-Error! Bookmark not defined.
Tabel III- 20 Fitur setelah Information gain	III-Error! Bookmark not defined.
Tabel III- 21 Hasil Klasifikasi Naïve Bayes	III-Error! Bookmark not defined.
Tabel III- 22 Jadwal Penelitian dalam <i>Work Breakdown Structure</i> (WBS).....	III-Error! Bookmark not defined.
Tabel IV- 1 Rincian Dataset.....	IV-Error! Bookmark not defined.
Tabel IV- 2 Skenario preprocess data	IV-Error! Bookmark not defined.
Tabel IV- 3 Skenario Analisis sentiment teks..	IV-Error! Bookmark not defined.
Tabel IV- 4 function dari preprocess_ignb.py .	IV-Error! Bookmark not defined.
Tabel IV- 5 function dari classify.py	IV-Error! Bookmark not defined.
Tabel IV- 6 function dari app.py	IV-Error! Bookmark not defined.
Tabel IV- 7 Rencana Pengujian Preprocess Data dan Analisis Sentimen	IV-Error! Bookmark not defined.
Tabel IV- 8 Rencana Pengujian Klasifikasi teks.....	IV-Error! Bookmark not defined.
Tabel IV- 9 Pengujian Preprocess Data dan analisis sentimen	IV-Error! Bookmark not defined.
Tabel IV- 10 Pengujian Klasifikasi teks	IV-Error! Bookmark not defined.
Tabel V- 1 Rincian data Pengujian	V-Error! Bookmark not defined.
Tabel V- 2 Hasil pengujian	V-Error! Bookmark not defined.

DAFTAR GAMBAR

Gambar II- 1 Metode Waterfall.....	II-Error! Bookmark not defined.
Gambar III- 1 Alur Program	III-Error! Bookmark not defined.
Gambar III- 2 Information gain.....	III-Error! Bookmark not defined.
Gambar III- 3 Naïve Bayes	III-Error! Bookmark not defined.
Gambar IV- 1 Use case Diagram	IV-Error! Bookmark not defined.
Gambar IV- 2 Activity diagram Preprocess Data dan Sentiment Analisis	IV-Error! Bookmark not defined.
Gambar IV- 3 Activity Diagram Analisis sentimen teks	IV-Error! Bookmark not defined.
Gambar IV- 4 Analisis Sentimen Naïve Bayes	IV-Error! Bookmark not defined.
Gambar IV- 6 Sequence Diagram sentimen analisis teks	IV-Error! Bookmark not defined.
Gambar IV- 7 Class Diagram.....	IV-Error! Bookmark not defined.
Gambar IV- 32 Interface halaman awal	IV-Error! Bookmark not defined.
Gambar IV- 33 Interface setelah input file	IV-Error! Bookmark not defined.
Gambar IV- 34 Interface setelah tombol ‘Preprocess Data’ ditekan	IV-Error! Bookmark not defined.
Gambar IV- 35 Hasil Sentimen Analisis Naive bayes	IV-Error! Bookmark not defined.
Gambar IV- 36 Sentimen Analisis Information gain dan Naive bayes.....	IV-Error! Bookmark not defined.
Gambar IV- 37 Interface awal klasifikasi teks.	IV-Error! Bookmark not defined.
Gambar IV- 38 Interface setelah input teks dan menekan ‘analisis’	IV-Error! Bookmark not defined.
Gambar V- 1 Diagram Perbandingan Hasil Evaluasi	V-Error! Bookmark not defined.

BAB I PENDAHULUAN

1.1 Pendahuluan

Pada bab pendahuluan ini membahas latar belakang masalah, rumusan masalah, tujuan, manfaat, batasan masalah, sistematika penulisan serta kesimpulan dari skripsi. Bab ini akan memberikan penjelasan umum mengenai keseluruhan isi laporan. Pendahuluan dimulai dengan penjelasan mengenai masalah yang ada dan bagaimana penyelesaian suatu masalah.

1.2 Latar Belakang

Penggunaan kendaraan listrik semakin mendapatkan perhatian di seluruh dunia sebagai solusi untuk mengurangi emisi karbon dan ketergantungan pada bahan bakar fosil (Alfarizi dan Fitriani. 2023). Namun, tantangan yang dihadapi dalam mengadopsi teknologi kendaraan listrik tidak hanya sebatas pada aspek tersebut, tetapi juga mencakup pada penerimaan dari masyarakat (Ernawati et al. 2023).

Twitter sebagai platform media sosial yang populer di Indonesia, dengan 14.75 juta pengguna aktif pada April 2023, menjadi sumber data yang potensial untuk memahami opini publik¹. Oleh karena itu, pemahaman mendalam tentang sentimen, dan sikap masyarakat terhadap kendaraan listrik sangatlah penting untuk melihat respon masyarakat, pengembangan kebijakan publik, merancang strategi yang efektif dalam pemasaran dan mendorong adopsi teknologi ini.

¹ Databoks “Jumlah Pengguna Twitter di Indonesia Capai 14,75 Juta per April 2023”(<https://databoks.katadata.co.id/datapublish/2023/05/31/jumlah-pengguna-twitter-di-indonesia-capai-1475-juta-per-april-2023-peringkat-keenam-dunia>, diakses pada 15 Maret 2024)

Analisis sentimen menjadi hal yang efektif dalam menggali data opini publik. Dengan mengumpulkan dan mengolah opini yang tersebar di *Twitter*, analisis sentimen dapat membantu mengklasifikasikan respons masyarakat ke dalam kategori positif, negatif, atau netral (Alfarizi dan Fitriani, 2023).

Dalam analisis sentimen, salah satu metode klasifikasi yang umum digunakan adalah *Naïve Bayes Classifier*. Algoritma ini didasarkan pada teorema Bayes, yang dikemukakan oleh ilmuwan Inggris Thomas Bayes. Metode ini menggunakan konsep probabilitas dan statistik, dimana *Naïve Bayes Classifier* mencari nilai probabilitas tertinggi untuk mengelompokkan data uji ke dalam kategori yang tepat (Ramadhani dan Suryono. 2024). Namun, ada masalah umum dalam klasifikasi teks yang disebabkan oleh banyaknya dimensi ruang fitur. Pemilihan fitur dapat menjadi solusi untuk masalah ini dengan *Information gain*. Dengan menggunakan *Information gain*, fitur yang kurang atau tidak relevan akan dihilangkan, sehingga dimensi data berkurang dan dapat meningkatkan kinerja (Anggita dan Abdulloh. 2023). Beberapa penelitian telah menunjukkan bahwa penerapan *Information gain* pada algoritma klasifikasi mampu meningkatkan akurasi yang dihasilkan.

Pada penelitian yang berjudul “Analisis Sentimen Kendaraan Listrik Menggunakan Algoritma *naive bayes* dengan Seleksi Fitur *Information gain* dan *Particle Swarm Optimization* ” oleh Salman Alfarizi dan Eka Fitriani pada 2023. Hasil yang diperoleh dengan metode *Naïve Bayes* sebesar 79.43%. Hasil yang didapatkan setelah menggunakan algoritma *Naïve Bayes*, *Information gain* dan *PSO* meningkat menjadi 84.54%.

Penelitian selanjutnya dilakakukan oleh Amelia Isnanda, Yuyun Umaidah, dan Jajam Haerul Jaman pada 2021, dengan judul “Implementasi *Naïve Bayes Classifier* Dan *Information gain* Pada Analisis Sentimen Penggunaan *E-Wallet* Saat Pandemi”. Hasil akurasi yang didapat dengan metode *naïve bayes* sebesar 84%. Sedangkan hasil akurasi dengan menggunakan *Naïve Bayes* dan *Information gain* sebesar 92%.

Penelitian lainnya dengan judul “Perbandingan Metode Klasifikasi *Support Vector Machine* Dan *Naïve Bayes* Pada Analisis Sentimen Kendaraan Listrik” oleh Ni Wayan Ernawati, I Nyoman Satya Kumara, dan Widyadi Setiawan pada 2023. Pada penelitian ini hasil *accuracy* yang didapatkan dengan metode *Naïve Bayes* sebesar 82% dan hasil dari SVM sebesar 81%. Dengan waktu latih yang dibutuhkan oleh metode SVM yaitu 37.42 detik sedangkan metode *naive bayes* hanya membutuhkan 0.10 detik. Hal ini membuktikan *Naïve Bayes* menjadi metode penelitian yang baik karena memiliki waktu latih yang cepat dan akurasi yang tinggi.

Tujuan dari penelitian ini adalah untuk mengetahui sentimen yang diberikan oleh masyarakat terhadap kendaraan listrik dengan menggunakan metode *naive bayes* dan seleksi fitur *information gain*.

1.3 Rumusan Masalah

1. Bagaimana membangun sebuah sistem analisis sentimen masyarakat terkait penggunaan kendaraan listrik dapat dianalisis melalui data yang terdapat pada media sosial Twitter dengan menggunakan metode *Naïve Bayes* dan *Informaiton Gain*?

2. Bagaimana akurasi model *Naïve Bayes* dalam menganalisis sentiment terhadap penggunaan kendaraan Listrik di twitter setelah diterapkan metode *Information gain*?

1.4 Tujuan Penelitian

1. Tujuan dari penelitian ini adalah untuk membangun sebuah sistem yang mampu menganalisis sentimen masyarakat terkait penggunaan kendaraan listrik berdasarkan data yang diambil dari media sosial Twitter. Sistem ini akan dibangun dengan menggunakan metode *Naïve Bayes Classifier* dan teknik seleksi fitur *Information gain*, sehingga dapat meningkatkan akurasi analisis sentimen terhadap penggunaan kendaraan listrik.
2. Mengetahui model *Naïve Bayes Classifier* dalam analisis sentimen setelah diterapkan metode *Information gain*.

1.5 Manfaat penelitian

1. Dapat mengetahui dan mempelajari hasil kinerja dari metode *naive bayes classifier* dalam melakukan analisis sentimen setelah ditambahkan seleksi fitur *information gain*.
2. Hasil penelitian dapat digunakan oleh industri kendaraan listrik untuk mengetahui persepsi publik mengenai kendaraan listrik.
3. Hasil penelitian dapat digunakan sebagai dasar untuk penelitian lanjutan dalam bidang yang sama dan berhubungan.

1.6 Batasan Masalah

1. Data yang dianalisis adalah data dari media sosial *Twitter*.
2. Kategori sentimen analisis dibagi menjadi 3, yaitu positif, negatif, dan netral.

1.7 Sistematika penulisan

Sistematika penulisan skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini dijelaskan mengenai latar belakang, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah, dan sistematika penulisan

BAB II KAJIAN LITERATUR

Pada bab ini membahas mengenai dasar-dasar teori yang digunakan dalam penelitian, seperti definisi dari *Text Mining*, *Text Preprocessing*, analisis sentimen, *Information gain*, *naive bayes*, evaluasi, dan metode *waterfall*.

BAB III METODOLOGI PENELITIAN

Pada bab ini membahas mengenai tahapan yang dilakukan pada penelitian. Setiap rencana tahapan penelitian dideskripsikan dengan rinci dengan berfokus pada satu kerangka kerja. Di akhir bab berisi perancangan manajemen proyek pada pelaksanaan penelitian.

BAB IV PENGEMBANGAN PERANGKAT LUNAK

Bab ini menjelaskan proses pengembangan perangkat lunak yang digunakan dalam penelitian ini. Pembahasannya meliputi metode pengembangan perangkat lunak, tahap-tahap yang dilalui mulai dari analisis kebutuhan, perancangan sistem, implementasi, pengujian, hingga pemeliharaan. Setiap tahap dijelaskan secara

terperinci untuk memberikan gambaran lengkap tentang bagaimana perangkat lunak dikembangkan untuk keperluan analisis sentimen pada media sosial Twitter terhadap kendaraan listrik.

BAB V HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil dari pengujian perangkat lunak yang telah dikembangkan. Analisis terhadap hasil pengujian dijelaskan secara mendalam, termasuk data yang diperoleh dari pengujian *naïve bayes classifier* dengan dan tanpa seleksi fitur menggunakan *information gain*. Bab ini juga membahas perbandingan hasil akurasi, presisi, recall, dan F1-score, serta interpretasi dari hasil tersebut dalam konteks sentimen masyarakat terhadap kendaraan listrik.

BAB VI KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan, yaitu mengenai efektivitas penggunaan *naïve bayes classifier* dan *information gain* dalam analisis sentimen terhadap kendaraan listrik. Selain itu, bab ini juga menyajikan saran yang dapat dijadikan acuan untuk penelitian lebih lanjut, termasuk pengembangan metode lain yang mungkin dapat meningkatkan performa klasifikasi atau penerapan dataset yang lebih besar untuk hasil yang lebih representatif.

1.8 Kesimpulan

Dalam penelitian ini, analisis sentimen pada media sosial *Twitter* terhadap penggunaan kendaraan listrik dilakukan menggunakan metode *naïve bayes classifier* dan *information gain*. Tujuan utama penelitian adalah untuk mengidentifikasi sentimen pengguna *Twitter* terhadap kendaraan listrik,

mengimplementasikan metode *naïve bayes classifier*, dan menganalisis efektivitas *information gain* dalam meningkatkan akurasi klasifikasi sentimen.

Dengan demikian, penelitian ini tidak hanya memberikan pemahaman yang lebih baik mengenai sentimen pengguna terhadap kendaraan listrik, tetapi juga menyediakan wawasan yang berguna bagi industri kendaraan listrik dalam merancang strategi pemasaran dan pengembangan produk. Selain itu, penelitian ini juga memberikan kontribusi terhadap pengembangan teknologi analisis sentimen, yang dapat bermanfaat bagi berbagai aplikasi lainnya dalam pemrosesan bahasa alami.

DAFTAR PUSTAKA

- Alfarizi. S., & Fitriani. E. (2023). Analisis Sentimen Kendaraan Listrik Menggunakan Algoritma *naive bayes* dengan Seleksi Fitur *Information gain* dan Particle Swarm Optimization. Indonesian Journal on Software Engineering (IJSE).
- Alin F. N., Totohendarto M. H., Muttaqin M. R. (2023). Analisis Sentimen Terhadap Kendaraan Listrik Pada Platform Twitter Menggunakan Metode *Naive bayes*. Informatics for Educators And Professionals : Journal of Informatics
- Anggita, S.D., & Abdulloh, F.F. (2023). Optimasi Algoritma Support Vector Machine Berbasis PSO Dan Seleksi Fitur *Information gain* Pada Analisis Sentimen. *Journal Of Applied Computer Science And Technology*.
- Darwis, D., Siskawati, N., & Abidin, Z. Penerapan Algoritma *naive bayes* untuk Analisis Sentimen Review Data Twitter BMKG Nasional. Jurnal TEKNO KOMPAK
- Ernawati. N. W., Kumara. I. N. S., & Setiawan. S. (2023). Perbandingan Metode Klasifikasi Support Vector Machine Dan Naïve Bayes Pada Analisis Sentimen Kendaraan Listrik. Jurnal SPEKTRUM.
- Firdaus, A., Istalama Firdaus, W. (2021). Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan). Jurnal JUPITER.

- Husada, H.C., & Paramita, A.S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). TEKNIKA
- Isnanda, A., Umaidah, Y., & Jaman, J.H. (2021). Implementasi Naïve Bayes Classifier Dan *Information gain* Pada Analisis Sentimen Penggunaan E-Wallet Saat Pandemi. Jurnal Teknologi Informatika dan Komputer MH. Thamrin
- Mahardika F., Mustofa K., Suseno A.T.(2023). Implementasi Metode Waterfall pada Sistem Informasi Penjualan Unit Motor Berbasis Web. JURNAL ILMU KOMPUTER
- Nofiyani, & Wulandari. (2021). Implementasi Electronic Data Processing Untuk meningkatkan Efektifitas dan Efisiensi Pada Text Mining. JURNAL MEDIA INFORMATIKA BUDIDARMA
- Noviana R. & Rasal I. (2023). PENERAPAN ALGORITMA *NAIVE BAYES* DAN SVM UNTUK ANALISIS SENTIMEN BOYBAND BTS PADA MEDIA SOSIAL TWITTER. Jurnal Teknik dan science
- Nur. F., Ahsan. M., Harianto. W. (2022). Komparasi Tingkat Akurasi *Information gain* Dan Gain Ratio Pada Metode K-Nearest Neighbor. Jurnal Mahasiswa Teknik Informatika
- Rahman Isnain, A., Indra Skti, A., Satya Marga, A., & Alita. (2021). Sentimen Analisis Publik Terhadap kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma SVM. JDMSI
- Ramadhani, B., & Suryono, R.R. (2024). Komparasi Algoritma Naïve Bayes dan

Logistic Regression Untuk Analisis Sentimen Metaverse. JURNAL
MEDIA INFORMATIKA BUDIDARMA

Ridwansyah, T. (2022). Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan *Naïve Bayes Classifier*. *Kajian Ilmiah Informatika dan Komputer*.

Tanggreani, A., I., & Sitokdana, M.N.N. (2022). Analisis Sentimen Aplikasi *E-Government* Pada *Google Play* Menggunakan Algoritma *Naïve Bayes*. *Jurnal Teknik Informatika dan Sistem Informasi*

Ulgasesa, R., Putra Negra, A.B., & Tursina. (2021). Pengaruh Stemming Terhadap Performa Klasifikasi Sentimen Masyarakat Tentang Kebijakan New Normal. *Jurnal Sistem dan Teknologi Informasi*.

Wati, R., Ernawati, S., & Rachmi, H. (2023). Pembobotan TF-IDF Menggunakan *Naïve Bayes* pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH. *Jurnal Manajemen Informatika (JAMIKA)*