

***SENTIMENT ANALYSIS PADA MASYARAKAT TERHADAP LIGHT
RAIL TRANSIT (LRT) KOTA PALEMBANG MENGGUNAKAN METODE
IMPROVED K-NEAREST NEIGHBOR***

SKRIPSI

Program Studi Sistem Informasi

Jenjang Sarjana



Oleh:

SITI NUR ARAFAH

NIM 09031281823145

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2022

HALAMAN PENGESAHAN

SKRIPSI

***SENTIMENT ANALYSIS PADA MASYARAKAT TERHADAP LIGHT
RAIL TRANSIT (LRT) KOTA PALEMBANG MENGGUNAKAN METODE
IMPROVED K-NEAREST NEIGHBOR***

Sebagai salah satu syarat untuk penyelesaian
studi di Program Studi Sistem Informasi SI

Oleh

SITI NUR ARAFAH

09031281823145

Palembang, 21 Juli 2022

**Menyetujui,
Ketua Jurusan Sistem Informasi,**

Pembimbing,


Endang Lestari Ruskan, S.Kom., M.T.
NIP 197811172006042001


Fathoni, MMSI.
NIP 197210182008121001

HALAMAN PERNYATAAN

Yang Bertanda Tangan di Bawah Ini :

Nama : Siti Nur Arafah
NM : 09031281823145
Prodi : Sistem Informasi Reguler
Judul Skripsi : *Sentiment Analysis* pada Masyarakat terhadap *Light Rail Transit*
(LRT) Kota Palembang menggunakan Metode *Improved K-Nearest Neighbor*

Hasil Pengecekan *Software iThenticate/Turnitin* : 16%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/*plagiat*. Apabila ditemukan unsur penjiplakan/*plagiat* dalam laporan skripsi 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.



Palembang, 27 Mei 2022



Siti Nur Arafah
NIM. 09031281823145

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Senin

Tanggal : 27 Juni 2022

Nama : Siti Nur Arafah

NIM : 0903128123145

Judul : *Sentiment Analysis* Pada Masyarakat terhadap *Light Rail Transit (LRT)* Kota Palembang Menggunakan Metode *Improved K-Nearest Neighbor*

Komisi Penguji:

1. **Ketua** : Dr. Ermatita, M.Kom

2. **Pembimbing** : Fathoni, MMSI.

3. **Penguji 1** : Rizka Dhini Kurnia, M.Sc

4. **Penguji 2** : Bayu Wijaya Putra, M.Kom



Mengetahui,
Ketua Jurusan Sistem Informasi




Endang Lestari Ruskan, M.T
NIP 197811172006042001

HALAMAN PERSEMBAHAN

MOTTO

“Jangan pergi mengikuti kemana jalan akan berujung.

Buat jalanmu sendiri dan tinggalkanlah jejak”

-Ralph Waldo Emerson-

Skripsi ini kupersembahkan kepada:

- ❖ *Allah SWT*
- ❖ *Ibu dan Bapak*
- ❖ *Keluarga Besarku*
- ❖ *Dosen Pembimbing Akademik*
- ❖ *Dosen Pembimbing dan Dosen Penguji*
- ❖ *Sistem Informasi 2018*
- ❖ *Almamater Universitas Sriwijaya*

KATA PENGANTAR



Alhamdulillah Puji Syukur kepada kehadiran Allah SWT yang telah memberikan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan Skripsi ini dengan baik dengan Judul “***Sentiment Analysis Pada Masyarakat Terhadap Light Rail Transit (LRT) Kota Palembang Menggunakan Metode Improved K-Nearest Neighbor***”. Skripsi ini dibuat untuk memenuhi syarat dalam menyelesaikan Program Sarjana (S1) Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.

Penulis menyadari bahwa skripsi ini tidak mungkin terselesaikan tanpa adanya dukungan, bantuan, bimbingan dan nasehat dari berbagai pihak selama penyusunan skripsi ini. Karena itu, pada kesempatan ini penulis mengucapkan rasa terima kasih kepada:

1. Prof. Dr. Ir. H. Anis Saggaff, MSCE, IPU. selaku Rektor Universitas Sriwijaya
2. Bapak Jaidan Jauhari, S.Pd, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Ibu Endang Lestari Ruskan, M.T. selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Fathoni, S.T., MMSI. selaku Dosen Pembimbing Skripsi yang telah dengan sabar memberikan bimbingan, arahan serta saran kepada penulis sehingga skripsi ini dapat diselesaikan dengan baik.
5. Seluruh Dosen pengajar Fakultas Ilmu Komputer Universitas Sriwijaya yang telah memberikan ilmu pengetahuan yang tak ternilai selama penulis menempuh pendidikan di Fakultas Ilmu Komputer Universitas Sriwijaya.
6. Kedua Orang tua tercinta, Bapak Amat Jamili dan Ibu Sri Mulyani yang selalu memberikan dukungan, kasih sayang, nasehat, semangat serta do'a yang tiada hentinya untuk penulis agar dapat mencapai hasil yang terbaik dan menjadi anak yang dapat dibanggakan.

7. Kak Angga Pratama. selaku Admin Jurusan Sistem Informasi yang telah membantu penulis dalam proses administrasi selama penyusunan skripsi.
8. Sahabat dan teman terbaik yang tidak bisa disebutkan satu persatu, yang selalu memberikan dukungannya serta selalu setia untuk mendengarkan keluh kesah dari penulis selama melakukan penyusunan skripsi.
9. Para pejuang skripsi satu bimbingan yang selalu memberikan dukungan dan semangat kepada penulis untuk selalu tetap semangat dalam menyusun skripsi ini.
10. Seluruh anggota Sinus A 2018 yang telah mengisi hari-hari penulis selama penulis menempuh pendidikan di Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.

Akhir kata, penulis menyadari bahwa masih terdapat kekurangan dalam Skripsi ini, untuk itu saran dan kritik yang bersifat membangun akan penulis terima dengan senang hati.

Palembang, 27 Mei 2022

Penulis,



SITI NUR ARAFAH
NIM 09031281823145

**SENTIMENT ANALYSIS PADA MASYARAKAT TERHADAP LIGHT
RAIL TRANSIT (LRT) KOTA PALEMBANG MENGGUNAKAN METODE
IMPROVED K-NEAREST NEIGHBOR**

Oleh

Siti Nur Arafah 09031281823145

ABSTRAK

Light Rail Transit (LRT) merupakan sistem transportasi cepat yang berkelanjutan, yang dibangun untuk mengatasi masalah kemacetan di kota Palembang. Agar dapat menarik minat masyarakat untuk beralih menggunakan transportasi umum dibandingkan dengan transportasi pribadi salah satunya dengan meningkatkan kualitas pelayanan yang diberikan. Analisis sentimen digunakan untuk mengelompokkan opini positif dan negatif pada pengguna jasa transportasi LRT Kota Palembang. Selain melakukan penarikan data melalui *crawling* data pada data *tweet*, peneliti juga melakukan penyebaran kuesioner. Dalam melakukan proses klasifikasi analisis sentimen, penelitian ini menggunakan metode *Improved K-Nearest Neighbor* yang merupakan modifikasi dari metode *K-Nearest Neighbor*. Hasil penelitian ini merupakan data latih dan uji terhadap 1617 *record* data yang menghasilkan akurasi tertinggi sebesar 74.07% pada 90% data latih dan 10% data uji, dengan nilai presisi 70%, *recall* 56% dan *f-1 score* 59%, sedangkan akurasi terendah dengan nilai akurasi 63.04% pada 50% data latih dan 50% data uji, dengan presisi 44%, *recall* 42% dan *f-1 score* 42%.

Kata Kunci: Analisis Sentimen, *K-Nearest Neighbor*, *Light Rail Transit (LRT)*.

**SENTIMENT ANALYSIS ON THE COMMUNITY TO LIGHT RAIL
TRANSIT (LRT) OF PALEMBANG CITY USING IMPROVED K-NEAREST
NEIGHBOR METHOD**

By

Siti Nur Arafah 09031281823145

ABSTRACT

Light Rail Transit (LRT) is a sustainable fast transportation system, which was built to overcome the congestion problem in the city of Palembang. In order to attract people's interest to switch to using public transportation compared to private transportation, one of them is by improving the quality of services provided. Sentiment analysis is used to classify positive and negative opinions on users of Palembang City LRT transportation services. In addition to retrieving data through crawling data on tweet data, the researchers also distributed questionnaires. In conducting the classification process of sentiment analysis, this study uses the Improved K-Nearest Neighbor method which is a modification of the K-Nearest Neighbor method. The results of this research are testing and training data on 1617 data records and the highest accuracy of 74.07% on 90% training data and 10% testing data, with 70% precision, 56% recall and 59% f-1 score, while the lowest accuracy with an accuracy of 63.04% on 50% training data and 50% testing data, with 44% precision, 42% recall and 42% f-1 score.

Keyword: *Sentiment Analysis, K-Nearest Neighbor, Light Rail Transit (LRT).*

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN BEBAS PLAGIAT	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan	6
1.4 Manfaat	6
1.5 Batasan Masalah	7
BAB II TINJAUAN PUSTAKA	8
2.1 Penelitian Terkait	8
2.2 Masyarakat	10
2.3 <i>Light Rail Transit (LRT)</i>	10
2.3.1 Fasilitas dan Layanan LRT	11
2.3.2 Jadwal Keberangkatan LRT	13
2.4 Analisis Sentimen	14
2.4.1 Sentimen Positif	15
2.4.2 Sentimen Negatif	15
2.4.3 Sentimen Netral	15
2.5 <i>Text Mining</i>	16
2.5.1 <i>Preprocessing Text</i>	16
2.6 Pembobotan Kata TF-IDF	18
2.7 <i>Improved K-Nearest Neighbor</i>	19
2.8 Akurasi, Presisi, dan <i>Recall</i>	20

2.8.1 Akurasi	20
2.8.2 Presisi	21
2.8.3 Recall	21
2.9 Twitter	22
2.9.1 Twitter <i>Application Programming Interface</i> (API)	24
2.10 Python	25
BAB III METODOLOGI PENELITIAN	26
3.1 Tahap Penelitian	26
3.2 Analisis Kebutuhan Data	28
3.2.1 Alat Penelitian	28
3.2.2 Bahan Penelitian	29
3.3 Metode Pengumpulan Data	30
3.3.1 Studi Literatur	30
3.3.2 Survey Kuesioner	30
3.3.3 Observasi (<i>Crawling Data</i>)	31
3.4 Metode Pengolahan Data	33
3.4.1 Proses <i>Pre-processing text</i>	33
3.4.2 Pembobotan Kata TF-IDF	36
3.4.3 Penerapan Metode <i>Improved K-NN</i>	37
3.5 Implementasi	38
3.6 Kesimpulan dan Saran	38
BAB IV HASIL DAN PEMBAHASAN	39
4.1 Pengumpulan Data	39
4.1.1 Proses <i>Crawling Data</i>	39
4.1.2 Proses Pengumpulan Data Kuesioner	40
4.1.3 Hasil Pengumpulan Data	41
4.2 Hasil <i>Preprocessing</i>	42
4.2.1 <i>Case Folding Data</i>	42
4.2.2 <i>Tokenizing Data</i>	44
4.2.3 <i>Filtering (Stopword Removal) Data</i>	45
4.2.4 <i>Stemming Data</i>	47
4.2.5 Hasil Proses TF-IDF	49
4.3 Hasil Sentimen Analisis Data Data	50
4.3.1 <i>Library NLTK</i>	50
4.3.2 <i>Library Sastrawi</i>	50
4.3.3 <i>Library Polarity</i>	51

4.3.4 Hasil Pembagian Data.....	52
4.4 Pengujian Data dengan <i>Python 3</i>	53
4.4.1 Pemanggilan <i>Library</i> Awal	53
4.4.2 <i>Library Sklearn</i>	54
4.4.3 <i>Library Split Data</i>	55
4.4.4 <i>Library Model Improved K-Nearest Neighbor</i>	55
4.5 Hasil Pengujian Validasi	55
4.6 Hasil Pembahasan	58
BAB V KESIMPULAN DAN SARAN.....	69
5.1 Kesimpulan	69
5.2 Saran.....	70
DAFTAR PUSTAKA	71

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Jadwal Keberangkatan.....	14
Gambar 3.1 Tahapan Penelitian	26
Gambar 3.2 Proses <i>Crawling data sentiment</i>	31
Gambar 3.3 Sampel data <i>sentiment</i>	32
Gambar 3.4 Alur Proses Pengolahan Data	33
Gambar 3.5 Alur Proses <i>Preprocessing</i>	34
Gambar 3.6 Alur Proses Pembobotan Kata	37
Gambar 3.7 Alur Proses <i>Improved KNN</i>	38
Gambar 4.1 Operator <i>Crawling Data</i>	39
Gambar 4.2 Proses Pengumpulan Data dengan Kuesioner	40
Gambar 4.3 Proses Pemanggilan Data	42
Gambar 4.4 <i>Library Case Folding</i>	43
Gambar 4.5 Data Setelah Proses <i>Case Folding</i>	44
Gambar 4.6 <i>Library Tokenizing</i>	44
Gambar 4.7 Data Setelah Proses <i>Tokenizing</i>	45
Gambar 4.8 <i>Library Stopword Removal</i>	46
Gambar 4.9 Hasil Setelah <i>Filtering Data dan Stopword Removal</i>	47
Gambar 4.10 <i>Library Stemming</i>	48
Gambar 4.11 Data Setelah Proses <i>Stemming</i>	49
Gambar 4.12 Proses <i>Library TF-IDF</i>	49
Gambar 4.13 Hasil Proses <i>Library TF-IDF</i>	50
Gambar 4.14 <i>Library NLTK</i>	50
Gambar 4.15 <i>Library Sastrawi</i>	51
Gambar 4.16 <i>Library Polarity</i>	51
Gambar 4.17 Grafik Sentimen Analisis.....	53
Gambar 4.18 Library Awal Proses Klasifikasi	53

Gambar 4.19 <i>Library Sklearn</i>	54
Gambar 4.20 Hasil Perubahan Data	54
Gambar 4.21 <i>Library Split Data</i>	55
Gambar 4.22 <i>Library Improved K-Nearest Neighbor</i>	55
Gambar 4.23 Hasil <i>Recall</i> dan <i>Precision</i>	57
Gambar 4.24 <i>Library Proses Case Folding</i>	58
Gambar 4.25 <i>Library Tokenizing Proses Cleaning</i>	60
Gambar 4.26 Hasil Frekuensi Token.....	63
Gambar 4.27 Hasil Ranking Kata TF-IDF	66
Gambar 4.28 <i>Library Polarity</i> (Pelabelan Data).....	66
Gambar 4.29 Grafik Persentase Data Sentimen.....	67

DAFTAR TABEL

	Halaman
Tabel 2.1 Penelitian Terkait.....	8
Tabel 3.1 Sampel data yang sudah dikumpulkan.....	29
Tabel 3.2 Proses <i>Preprocessing</i>	34
Tabel 3.3 Proses <i>Case Folding</i>	35
Tabel 3.4 Proses <i>Tokenizing</i>	35
Tabel 3.5 Proses <i>Filtering</i>	36
Tabel 3.6 Proses <i>Stemming</i>	36
Tabel 4.1 Hasil Pengumpulan Data.....	41
Tabel 4.2 Data Sebelum <i>Case Folding</i>	43
Tabel 4.3 Contoh Text atau Karakter yang dihilangkan	44
Tabel 4.4 Data Sebelum Proses <i>Tokenizing</i>	45
Tabel 4.5 Sampel Kata yang dihilangkan.....	46
Tabel 4.6 Hasil Proses <i>Stopword Removal</i>	47
Tabel 4.7 Data Sebelum <i>Stemming</i>	48
Tabel 4.8 Sampel Hasil Label Sentimen Hasil	52
Tabel 4.9 Pembagian Data Sentimen	52
Tabel 4.10 Hasil Data Uji <i>Confusion Matrix</i>	55
Tabel 4.11 Hasil Data Tertinggi Dataset (90:10).....	56
Tabel 4.12 Hasil Akurasi Tertinggi Dataset (90:10).....	56
Tabel 4.13 Hasil Data Terendah Dataset (50:50).....	56
Tabel 4.14 Hasil Akurasi Terendah Dataset (50:50).....	56
Tabel 4.15 Data Sebelum dan Setelah Proses <i>Case Folding</i>	59
Tabel 4.16 Data Sebelum dan Setelah Proses <i>Cleaning</i>	61
Tabel 4.17 Data Sebelum dan Setelah Proses <i>Tokenizing</i>	62
Tabel 4.18 Data Sebelum dan Setelah Proses <i>Filtering</i>	63
Tabel 4.19 Data Sebelum dan Setelah Proses <i>Stemming</i>	64
Tabel 4.20 Pembagian Data Sentimen	67

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Dataset Awal.....	A-1
Lampiran 2 Dataset Hasil Preprocessing	B-1
Lampiran 3 Turnitin.....	C-1
Lampiran 4 Fasilitas LRT Palembang.....	D-1
Lampiran 5 Form Perbaikan	E-1
Lampiran 6 Kartu Konsultasi TA	F-1
Lampiran 7 SK Pembimbing TA	G-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sistem transportasi di Indonesia saat ini telah berkembang pesat dan memberikan mobilitas yang meningkat kepada masyarakat. Dalam era perkembangan dan pertumbuhan penduduk yang begitu pesat, Kota Palembang merupakan ibu Kota dari provinsi Sumatera Selatan yang termasuk Kota terbesar kedua di Pulau Sumatera setelah Kota Medan. Berdasarkan Sensus penduduk tahun 2020 Sumatera Selatan, Kepadatan Penduduk di Kota Palembang mencapai hingga 1.7 juta jiwa, yang dimana mengakibatkan kemacetan lalu lintas di Kota Palembang semakin tahunnya semakin parah. Menurut hasil dari studi Pemerintah Provinsi (Pemprov) Sumatera Selatan, mulai 2012 lalu lintas di 10 jalan protokol Kota Palembang akan mengalami kemacetan total pada tahun 2019 karena kapasitas jalan sudah tidak sanggup untuk menampung rata-rata lalu lintas pada tahun 2019 yang akan meningkat sebesar 6 kali lipat (Noerdin, 2016).

Transportasi merupakan salah satu sarana pembangunan yang strategis dan penting yang mempercepat perekonomian dan dapat mempengaruhi seluruh aspek kehidupan masyarakat. Transportasi berfungsi sebagai sarana penghubung antar tempat maupun wilayah, sehingga kualitas dari pelayanan dan fasilitas akan menjadi dasar yang dapat mempengaruhi konsumen dalam memilih jenis jasa yang berkembang di era saat ini. Masyarakat akan menuntut pelayanan dan fasilitas yang lebih baik sesuai dengan kebutuhannya untuk beralih menggunakan transportasi umum. Sebagai salah satu contoh transportasi umum di Kota Palembang yaitu *Light Rail Transit* (LRT).

LRT merupakan sistem transportasi cepat yang berkelanjutan, yang dibangun untuk mengatasi masalah kemacetan di kota Palembang. LRT merupakan kereta api ringan yang dibangun menggunakan aluminium campuran yang ramah lingkungan dan memiliki kecepatan sekitar 40 km/jam hingga 80 km/jam. LRT dioperasikan oleh PT. Kereta Api Indonesia Divisi Regional III Palembang, yang memiliki beberapa jalur layanan, yang dimulai dari stasiun Bandara Internasional Sultan Mahmud Badaruddin II sampai Jakabaring *Sport City*(Widiyanti, 2019)

Dilihat dari angka kemacetan Kota Palembang, masih banyak masyarakat di yang menggunakan kendaraan pribadi dan belum beralih menggunakan LRT sebagai alat transportasi umum. Hal ini perlu dipertimbangkan lebih lanjut, untuk mengetahui pendapat masyarakat mengenai apa yang dirasakan ketika menggunakan LRT Palembang, sehingga dapat menjadi standar dan solusi dalam pengambilan keputusan pemerintah untuk meningkatkan pelayanan dan kualitas pada LRT kota Palembang.

Di era sekarang ini, Teknologi telah berkembang sangat pesat, sehingga, dapat memudahkan untuk akses dan pengumpulan informasi, salah satunya adalah akses informasi melalui jejaring pengguna sosial media. Pada penelitian tahun 2015, *Twitter* merupakan salah satu jejaring media sosial yang paling cepat berkembang dan dapat dilihat bahwa jumlah pengguna aktif internet di Indonesia telah mencapai angka 72,7 juta pengguna, dimana setiap harinya 4.1 juta *tweets* berasal dari Indonesia (CNN Indonesia, 2016). *Twitter* merupakan layanan dari jejaring media sosial, yang memungkinkan penggunanya untuk dapat berinteraksi dengan pengguna lainnya dalam sebuah pesan yang berbasis teks hingga 280 karakter (*tweets*) (Baihaqi et al., 2020). LRT merupakan salah satu transportasi

yang banyak dibahas di media sosial Twitter. *Twitter* menyediakan data yang tersedia melalui *Twitter API*, sehingga memudahkan proses pengumpulan *tweets* dalam jumlah besar (Oktinas, 2017). Pada pengguna *twitter API*, *tweets* akan diolah menjadi data untuk pencarian sesuai *keywords*, data tersebut akan diproses menggunakan analisis sentimen, yang apabila diproses dengan baik dan benar dapat menghasilkan informasi yang dapat berguna dalam pengambilan keputusan dan meningkatkan persepsi seorang akan sesuatu hal yang baru diketahui.

Selain melakukan penarikan data melalui *crawling data* pada data *tweet*, peneliti juga melakukan penyebaran kuesioner. Alasan peneliti melakukan penyebaran kuesioner, peneliti ingin melihat respon masyarakat secara langsung terhadap penggunaan fasilitas dan layanan LRT Palembang, dengan menggunakan kuesioner peneliti bisa mendapatkan reaksi dan opini masyarakat yang lebih beragam, serta mendapatkan *record data* yang lebih banyak.

Analisis sentimen atau *opinion mining* adalah proses otomatis untuk memahami, mengekstraksi, dan memproses atau mengolah data berbentuk tekstual untuk memperoleh informasi sentimen yang terkandung pada kalimat opini (Buntoro et al., 2015). Pada Penelitian ini, Analisis sentimen digunakan untuk mengelompokkan opini positif dan negatif pada pengguna jasa transportasi LRT Kota Palembang, sehingga lebih cepat dan mudah bagi pengelola untuk mengatasi dan meninjau kembali kekurangan yang terdapat pada LRT Kota Palembang, baik dari segi pelayanan maupun segi fasilitas yang ada.

Dalam melakukan proses klasifikasi analisis sentimen, dilakukan menggunakan metode *Improved K-Nearest Neighbor* yang merupakan modifikasi dari metode *K-Nearest Neighbor*. Modifikasi dilakukan karena ada perubahan

penentuan nilai- k , dimana penentuan nilai- k tetap dilakukan hanya saja setiap kategori memiliki nilai- k yang berbeda. Metode *Improved KNN* dipilih, karena performanya lebih baik pada saat menggunakan *k-values* yang beragam dibandingkan metode KNN tanpa modifikasi. Metode ini bersifat fleksibel dan tangguh terhadap data yang *noisy* serta efektif digunakan dengan jumlah data latihan yang besar (Li & Han, 2013).

Salah satu penelitian yang pernah dilakukan oleh (Okfalisa et al., 2018) melakukan perbandingan antara metode KNN dan Modifikasi KNN pada *data Conditional Cash Transfer Implementation Unit* (Unit Pelaksana Program Keluarga Harapan) dengan *7395 record data*. Hasilnya metode K-NN mendapatkan akurasi tertinggi 94,95% dengan rata-rata akurasi selama pengujian 93.94% dan akurasi tertinggi pada MK-NN mendapatkan akurasi tertinggi 99.51% dengan dengan rata-rata akurasi 99,20%, Sehingga dapat dianalisis bahwa kemampuan algoritma MKNN dalam akurasi lebih baik daripada algoritma KNN.

Selanjutnya pada penelitian yang berjudul “*Sentiment Analysis on Movie Reviews Using Information Gain and K-Nearest Neighbor*” oleh (Octaviani Faomasi Daeli, 2020). Pada penelitian ini digunakan *polarity* dataset v2.0 dari *Cornell movie review*. Hasilnya metode K-NN mendapatkan akurasi tertinggi yaitu 60% pada $K=3$. Namun, setelah menggunakan *Information Gain*, akurasinya meningkat menjadi 96.8% pada $K=3$. Bahkan akurasinya tertinggi dibanding metode lainnya yaitu *Naive Bayes*, *Support Vector Machine*, dan *Random Forest*. Sehingga dapat disimpulkan dalam penelitian tersebut metode KNN memiliki nilai akurasi yang lebih baik daripada metode lainnya.

Berdasarkan uraian latar belakang diatas, maka dilakukan penelitian yang berjudul “*Sentiment Analysis Pada Masyarakat terhadap Light Rail Transit (LRT) Kota Palembang Menggunakan Metode Improved K-Nearest Neighbor*” yang membahas mengenai analisis sentimen dari data *twitter* dan pendapat langsung pengguna LRT Palembang. Pemilihan objek penelitian ini, yaitu LRT Palembang. Karena moda transportasi ini merupakan salah satu moda transportasi umum berkelanjutan yang memiliki banyak ulasan di jejaring media sosial *twitter* yang dibutuhkan untuk penelitian ini.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan permasalahannya sebagai berikut:

1. Bagaimana hasil dari analisis sentimen pada masyarakat terhadap LRT Kota Palembang?
2. Bagaimana penerapan metode *Improved K-Nearest Neighbor* pada analisis sentimen masyarakat terhadap LRT Kota Palembang?
3. Bagaimana hasil akurasi dari menerapkan metode *Improved k-Nearest Neighbor* pada analisis sentimen masyarakat terhadap LRT Kota Palembang?

1.3 Tujuan Penelitian

Berikut merupakan tujuan dari penelitian ini antara lain sebagai berikut:

1. untuk mengetahui sentimen masyarakat terhadap LRT Kota Palembang,
2. untuk menerapkan metode *Improved K-Nearest Neighbor* pada analisis sentimen masyarakat terhadap LRT Kota Palembang.
3. untuk mengetahui nilai akurasi dari menerapkan metode *Improved K-Nearest Neighbor* pada analisis sentimen masyarakat terhadap LRT Kota Palembang

1.4 Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian ini:

1. Manfaat Teoritis

Hasil penelitian ini diharapkan dapat dijadikan sebagai bahan bacaan dan referensi untuk penelitian-penelitian selanjutnya yang ingin meneliti mengenai analisis sentimen pada LRT Kota Palembang.

2. Manfaat Praktis

Hasil penelitian diharapkan dapat dijadikan sebagai informasi mengenai kualitas dari pelayanan dan fasilitas LRT Palembang, dimana informasi tersebut dapat dijadikan sebagai tolak ukur bahan pertimbangan Pemerintah dalam meningkatkan kualitas pelayanan dan fasilitas dari LRT Kota Palembang dan juga mengetahui kemampuan dari metode *Improved K-Nearest Neighbor* dalam bidang analisis sentimen.

1.5 Batasan Masalah

Untuk menghindari adanya pembahasan yang menyimpang, maka penelitian ini dibatasi:

1. Data yang digunakan dari *twitter* diambil melalui *twitter* API serta menggunakan Kuesioner kepada pengguna layanan dan fasilitas dari LRT Palembang.
2. *tweets* dan pendapat yang digunakan mengandung kata kunci #LRTPalembang
3. Proses analisis sentimen menggunakan perangkat lunak *Python 3*
4. Data *tweets* dikelompokkan menjadi 3 sentimen yaitu positif, negative dan netral.
5. Metode yang digunakan dalam penelitian menggunakan metode *Improved K-Nearest Neighbor*.

DAFTAR PUSTAKA

- Anthon Fathanudien, S. A. &. (2015). Partisipasi Masyarakat dalam Mewujudkan Kuningan sebagai Kabupaten Konservasi (Studi di Kabupaten Kuningan). *UNIFIKASI: Jurnal Ilmu Hukum*, 2(1), 67–90.
<https://doi.org/10.25134/unifikasi.v2i1.26>
- Aziz, A., Fauziah, F., & Fitri, I. (2021). Analisis Sentimen Terhadap Kebijakan Pemerintah Tentang Larangan Mudik Hari Raya Idulfitri di Indonesia Tahun 2021 Menggunakan Metode Naïve Bayes. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 5(2), 842–851.
<http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/381>
- Baihaqi, W. M., Pinilih, M., Rohmah, M., Informasi, T., Purwokerto, U. A., Informasi, S., Purwokerto, U. A., & Korespondensi, P. (2020). *Kombinasi K-Means Dan Support Vector Machine (SVM) Untuk K-Means And Support Vector Machine (SVM) Combination To Predict Sara Elements On Tweet*. 7(3), 501–510. <https://doi.org/10.25126/jtiik.202072126>
- Budi, S. (2017). Text Mining Untuk Analisis Sentimen Review Film. *Techno.COM*, 16(1), 1–8.
- Buntoro, G. A., Adji, T. B., & Purnamasari, A. E. (2015). *Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation*. *CITEE*, January, pp. 7-8.
- Divva, G., Zulma, M., & Chamidah, N. (2021). *Perbandingan Metode Klasifikasi Naive Bayes , Decision Tree Dan K- Nearest Neighbor Pada Data Log Firewall*. *April*, 679–688.
- Falahah, & Nur, D. D. A. (2015). Pengembangan Aplikasi Sentiment Analysis Menggunakan Metode Naïve Bayes (Studi Kasus Sentiment Analysis dari media Twitter). *Seminar Nasional Sistem Informasi Indonesia*, November, 335–339.
- Ghofari, K. AI. (2021). Pembuatan Sistem Pencarian Hadis dengan menggunakan Metode Pembobotan TF-IDF. *Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika*. 207–212.
- Haida, N., Pratama, G., Sukarnoto, T., & Widiawati. (2021). Ecobankers : Journal of Economy Banking Pengaruh Pemahaman Masyarakat Tentang Riba Terhadap Minat. *Ecobankers: Journal of Economy Banking*, 2(2), 131–139.
- Juditha, C. (2017). Sentimen Dan Imparsialitas Isi Berita Tentang Ahok Di Portal Berita Online. *Jurnal Penelitian Komunikasi Dan Pembangunan*, 18(1), 57.
<https://doi.org/10.31346/jpkp.v18i1.839>

- Li, B., & Han, L. (2013). Distance weighted cosine similarity measure for text classification. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8206 LNCS, 611–618. https://doi.org/10.1007/978-3-642-41278-3_74
- Lusiana, U. (2012). Application of Calibration Curve, Accuracy and Precision Chart as Internal Quality Control at COD Testing in Wastewater. *Biopropal Industri*, 3(1), 1–8.
- Naf'an, M. Z., Burhanuddin, A., & Riyani, A. (2019). Penerapan Cosine Similarity dan Pembobotan TF-IDF untuk Mendeteksi Kemiripan Dokumen. *Jurnal Linguistik Komputasional*, 2(1), 23–27.
- Noviar, A., & Mukti, S. (2021). Klasifikasi Arritmia pada Sinyal EKG menggunakan Deep Neural Network. *Jupiter*, 13, 29–38.
- Nurudin, M., Mara, M. N., & Kusnandar, D. (2014). *Ukuran sampel dan distribusi sampling dari beberapa variabel random kontinu*. 03(1), 1–6.
- Nurwahyuni, S., Bina, U., & Informatika, S. (2019). *Online Krl Access Menggunakan Metode*. 7(1), 31–38.
- Nurzahputra, A., Muslim, M. A., & Khusniati, M. (2017). Penerapan Algoritma K-Means Untuk Clustering Penilaian Dosen Berdasarkan Indeks Kepuasan Mahasiswa. *Techno.Com*, 16(1), 17–24. <https://doi.org/10.33633/tc.v16i1.1284>
- Octaviani Faomasi Daeli, N. (2020). Sentiment Analysis on Movie Reviews Using Information Gain and K-Nearest Neighbor. *Open Access J Data Sci Appl*, 3(1), 1–007. <https://doi.org/10.34818/JDSA.2020.3.22>.
- Okfalisa, Gazalba, I., Mustakim, & Reza, N. G. I. (2018). Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification. *Proceedings - 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2017, 2018-Januari*, 294–298. <https://doi.org/10.1109/ICITISEE.2017.8285514>
- Oktinas, W. (2017). Analisis Sentimen Pada Acara Televisi Menggunakan Improved K-Nearest Neighbor. *Universitas Sumatera Utara*, 1(2), 6-38.
- Pratama, A. Y., Umidah, Y., & Voutama, A. (2021). Analisis Sentimen Media Sosial Twitter Dengan Algoritma K-Nearest Neighbor dan Seleksi Fitur Chi-Square (Kasus Omnibus Law Cipta Kerja). *Sains Komputer & Informatika*, 5(2), 897–910. <https://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/view/386/365>

- Qisthiano, M. R., Kurniawan, T. B., Negara, E. S., & Akbar, M. (2021). *Pengembangan Model Untuk Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu dengan Metode Naïve Bayes*. 5, 987–994. <https://doi.org/10.30865/mib.v5i3.3030>
- Que, V. K. S., Iriani, A., & Purnomo, H. D. (2020). Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 9(2), 162–170. <https://doi.org/10.22146/jnteti.v9i2.102>
- Ratniasih, N. L., Sudarma, M., & Gunantara, N. (2017). Penerapan Text Mining Dalam Spam Filtering Untuk Aplikasi Chat. *Majalah Ilmiah Teknologi Elektro*, 16(3), 13. <https://doi.org/10.24843/mite.2017.v16i03p03>
- Rokhman, K. A., Berlilana, B., & Arsi, P. (2021). Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online. *Journal of Information System Management (JOISM)*, 3(1), 1–7. <https://doi.org/10.24076/joism.2021v3i1.341>
- Rustiana, D., & Rahayu, N. (2017). Analisis Sentimen Pasar Otomotif Mobil: Tweet Twitter Menggunakan Naïve Bayes. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 8(1), 113–120. <https://doi.org/10.24176/simet.v8i1.841>
- Sari, I. P., Purnama, I., & Ritonga, A. A. (2021). *Implementasi API pada Aplikasi Al-Qur ' an Berbasis Android dengan Metode UCD*. 5(April), 615–623. <https://doi.org/10.30865/mib.v5i2.2913>
- Siregar, R. R. A., Sinaga, F. A., & Arianto, R. (2017). Aplikasi Penentuan Dosen Penguji Skripsi Menggunakan Metode TF-IDF dan Vector Space Model. *Computatio : Journal of Computer Science and Information Systems*, 1(2), 171. <https://doi.org/10.24912/computatio.v1i2.1014>
- Sodik, F., Dwi, B., & Kharisudin, I. (2020). *Perbandingan Metode Klasifikasi Supervised Learning pada Data Bank Customers Menggunakan Python*. 3, 689–694.
- Telaumbanua, K., Sudarto, S., Butar-Butar, F., & Bilqis, P. S. (2021). Identifikasi Sampah Berdasarkan Tekstur Dengan Metode GLCM dan GLRLM Menggunakan Improved KNN. *Explorer*, 1(2), 45–52. <https://doi.org/10.47065/explorer.v1i2.94>
- Widiyanti, D. (2019). Pengembangan Park and Ride untuk meningkatkan pelayanan Angkutan LRT Kota Palembang. *Jurnal Penelitian Transportasi Darat*. 21, 103–116. doi: 10.25104/jptd.v21i2.1562.