**ANALISA SENTIMEN MENGGUNAKAN NAÏVE BAYES UNTUK MENGETAHUI PRESENTASE KOMENTAR PADA APLIKASI GO-JEK**

*Putri Rizqiyah*

*Program Studi Informatika, Institut Teknologi Telkom Purwokerto*

*Jl. D.I Panjaitan No. 128 Purokwerto*

*E-mail :* [*15102031@ittelkom-pwt.ac.id/*](mailto:15102031@ittelkom-pwt.ac.id/) *Hp. Tlp. +628974255772*

# ABSTRACT

perkembangan smartphone di Indonesia yang sangat pesat, menjadi cikal bakal pertumbuhan aplikasi mobile. penduduk Indonesia dengan jumlah penduduk sebanyak 250 juta jiwa adalah pasar yang besar, sehingga banyak peluang yang dimanfaatkan oleh para kalangan, salah satunya adalah kebutuhan transportasi online yang dapat memudahkan masyarakat dalam pemesanan transportasi secara online, salah satunya adalah aplikasi Go-jek, Antusiasme masyarakat dalam menyambut platform ini begitu besar, terbukti dengan kedudukan Go-jek pada rating tertinggi yang dipilih konsumen sebanyak 72,6 persen. Banyaknya komentar masyarakat yang diberikan kepada aplikasi ini merupakan hal yang menarik untuk diolah menjadi suatu informasi. Pemanfaatan informasi tersebut membutuhkan teknik analisis sehingga informasi yang dihasilkan dapat membantu banyak pihak.

Metode yang digunakan pada penelitian ini adalah metode klasifikasi Naïve Bayes. Naïve Bayes Classifier merupakan teknik pembelajaran algoritma data mining yang memanfaatkan metode probabilitas dan statistic. Metode ini gunakan untuk mengklasifikasikan komentar yang diberikan masyarkat terhadap aplikasi Go-jek. Komentar yang diberikan pada apilkasi ini nantinya akan di klasifikasikan menjadi komentar postifi, negatif, dan netral.

Tujuan dilakukannya penelitian ini adalah untuk mengetahui seberapa besar komentar positif, negative dan netral yang diberikan masyarakat terhadap aplikasi ini, sehingga nantinya akan memudahkan pihak Go-jek dalam memberikan kebijakan atau pengembangan aplikasi di masa depan.

*Keywords : Analisa Sentimen, Go-jek, Klasifikasi, Komentar, Naïve Bayes*

# ABSTRAK

# PENDAHULUAN

1. **Latar Belakang**

Perkembangan *smartphone* di Indonesia, merupakan cikal bakal pertumbuhan aplikasi *mobile* di Indonesia, Menristekdikti menyebutkan angka pengguna smartphone di Indonesia kini mencapai sekitar 25% dari total penduduk atau sekitar 65 juta orang [1]. Angka ini merupakan angka yang sangat menarik untuk peluang pasar, sehingga dimanfaatkan banyak kalangan berlomba-lomba untuk menciptakan aplikasi mobile yang dapat membantu kehidupan dan aktifitas sehari-hari masyarakat. Salah satunya perusahaan transportasi online, semenjak kemunculannya pada tahun 2015, aplikasi transportasi online ini disambut baik oleh masyarakat. Berdasarkan survey yang dilakukan oleh Yayasan Lembaga Konsumen Indonesia (YLKI) pada tanggal 5-16 April 2017 dan melibatkan sebanyak 4.668 responden, yang dilakukan untuk men-survey pelayanan transportasi online. Dari survey didapatkan bahwa alasan utama responden memilih transportasi online adalah, murah yakni 84.1%, cepat 81.9%, nyaman 78.8% dan aman 61.4% . Adapun rating tertinggi yang dipilih konsumen dalam pelayanan transportasi online diduduki oleh aplikasi Go-Jek sebanyak 72.6 %[2]. Meskipun menduduki ranking 1 yang dipilih oleh konsumen, sebagian masyarakat tentu tidak semuanya memberikan komentar positif terhadap aplikasi go-jek sebagian masyarakat mungkin ada yang memberikan komentar negative mengenai aplikasi ini, makadari itu dibutuhkan suatu teknik dan metode untuk mengetahui presentase klasifikasi komentar-komentar yang diberikan masyarakat pada aplikasi go-jek. Klasifikasi dibagi menjadi komentar negatif, positif, dan netral. Presentase banyaknya komentar negatif, positif, dan netral dibutuhkan oleh berbagai kalangan seperti masyarakat, pemerintah, bahkan pihak perusahaan go-jek itu sendiri. Presentase ini bisa saja dibutuhkan untuk membuat kebijakan bagi pemerintah, atau rencana pengembangan aplikasi dimasa depan untuk pihak perusahaan.

Analisis sentimen sendiri atau juga biasa disebut dengan opinion mining adalah salah satu bagian dari text mining. Bidang ini melakukan studi mengenai opini orang-orang, sentimen, evaluasi, tingkah laku dan emosi terhadap suatu entitas seperti produk, layanan, organisasi, individu, permasalahan, topik, acara dan atribut-atributnya [3].

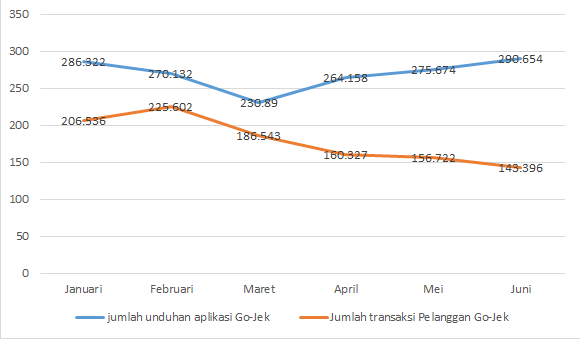
Fauziah Afshah dalam penelitiannya yang berjudul Analisa sentiment menggunakan Naïve Bayes untuk melihat persepsi masyarakat terhadap kenaikan harga jual rokok pada media sosial Twitter, melibatkan 350 dataset dengan klasifikasi 150 positif, 150 negatif dan 50 netral presentase akurasi terbesar sebanyak 85% berada pada opini netral [4]. Sedangkan Ahmad Fathan Hidayatullah dalam penelitiannya yang berjudul analisis sentiment dan klasifikasi kategori terhadap tokoh publik pada twitter, perhitungan akurasi klasifikasi menggunakan metode Naïve Bayes dengan fitur *term frequency* diperoleh sebesar 73.81% [5], dari sini bisa disimpulkan bahwa analisa sentimen menggunaka metode klasifikasi naïve bayes mendapatkan akurasi yang tinggi, makadari itu penulis memutuskan untuk memilih metode ini pada penelitian yang sedang dilakukan.

Tabel 1.1 Data pemberian rating *user* terhadap aplikasi go-jek per tahun 2018

Sumber: Google Play store

|  |  |  |
| --- | --- | --- |
| No | Kategori Rating | Jumlah Rating  (pengguna) |
| 1 | Rating Bintang 5 | 696.680 |
| 2 | Rating Bintang 4 | 178.904 |
| 3 | Rating Bintang 3 | 90.677 |
| 4 | Rating Bintang 2 | 31.610 |
| 5 | Rating Bintang 1 | 62.544 |

Dari tabel 1.1 bisa dilihat kategori pemberian rating oleh *user* untuk aplikasi go-jek, yang menjadi permasalahan adalah tidak semua jumlah pengguna yang ada pada setiap rating menggambarkan klasifikasi komentar pengguna, makadari itu diperlukan suatu metode dan analisis untuk mengklasifikasikan komentar pengguna kedalam beberapa kategori, dimana dalam penelitian ini berupa kategori positif, negatif, dan netral.



Gambar 1.1 Kurva unduhan aplikasi go-jek dan Transaksi pelanggan Go-Jek selama Januari-juni 2016

Sumber : PT Go-Jek Indonesia

Selain itu dilihat dari gambar 1.1 yang menunjukkan kurva unduhan aplikasi dan transaksi pelanggan Go-jek mengalami gap, dimana unduhan aplikasi Go-jek ini meningkat sedangkan transaksi yang dilakukan pada aplikasi ini justru menurun. Sehingga dibutuhkan sebuah penelitian faktor apa yang membuat transaksi pelanggan menurun, hal ini yang membuat penulis melakukan penelitian untuk mengetahui penyebab berkurangnya transaksi pada aplikasi ini melalui komentar-komentar yang diberikan pengguna pada aplikasi ini di play store.

**1.2 Rumusan Masalah**

Rumusan masalah dari penelitian ini adalah sebagai berikut:

1. Rating yang diberikan oleh *user* terhadap aplikasi Go-jek belum menggambarkan kondisi penerimaan pengguna terhadap aplikasi tersebut.
2. Pengguna masih kesulitan faktor apa yang mengharuskan pengguna menggunakan aplikasi Go-jek.
3. Penggunaan metode klasifikasi Naïve Bayes untuk menentukan presentase komentar pada aplikasi Go-jek

**1.3 Pertanyaan Penelitian**

Berdasarkan rumusan masalah maka pertanyaan penelitian yang dikemukakan adalah sebagai berikut:

1. Bagaimana menggambarkan rating yang diberikan oleh pengguna ke dalam suatu klasifikasi berdasarkan komentar pengguna?
2. Apa faktor yang mengharuskan pengguna menggunakan aplikasi tersebut?
3. Bagaimana penerapan metode klasifikasi Naïve Bayes dalam menentukan presentase komentar pada aplikasi Go-jek?

**1.4 Tujuan Penelitian**

Tujuan dari penelitian yang dilakukan ini adalah, menghitung presentase komentar positif, negatif dan netral yang diberikan pengguna terhadap aplikasi go-jek, sehingga memudahkan pengguna dalam memutuskan menggunakan aplikasi ini atau tidak berdasarkan presentase komentar yang diberikan.

**1.5 Manfaat Penelitian**

Manfaat dari penelitian ini adalah:

1. Mengetahui presentase komentar positif, negatif, dan netral berdasarkan komentar pengguna
2. Memudahkan pengguna dalam keputusan penggunaan aplikasi ini.
3. Mengetahui seberapa besar tingkat akurasi metode naïve bayes terhadap klasifikasi komentar aplikasi Go-jek.

# LANDASAN TEORI

* 1. **Dasar Teori**

Berikut ini merupakan informasi mengenai jurnal-jurnal yang penulis baca sebagai bahan referensi dan acuan dalam melaksanakan penelitian ini.

Tabel 2.1 Informasi Jurnal

|  |  |  |  |
| --- | --- | --- | --- |
| No | Judul, Nama peneliti, Tahun Penelitian | Masalah, Metode, Hasil | Perbedaan dengan penelitian yang kalian lakukan |
| 1 | Analisis Sentimen Terhadap Wacana Politik Pada Media Masa Online Menggunakan Algoritma Support Vector Machine dan Naïve Bayes, Andi Nurul Hidayat, 2015 | Mengetahui metode mana yang memiliki performansi yang lebih baik untuk di implementasikan dalam sentiment analysis terhadap wacana politik, Support Vector Machine dan Naïve Bayes, hasilnya adalah berupa efektifitas kalimat opini positif, netral, negatif terhadap wacana politik tersebut | Pada penelitian ini data yang digunakan adalah data berita pada media massa online sedangkan penelitian yang penulis lakukan adalah menganalisis suatu komentar dari sebuah aplikasi *smartphone* |
| 2 | Penerapan Analisis Sentimen pada Twitter Berbahasa Indonesia Sebagai pemberi Rating, Support Vector Machine, Nuvirta Monarizqa, Lukito Edi Nugroho, Bimo Sunarfri Hantono, 2015 | Menganalisis komentar-komentar di twitter untuk diterjemahkan menjadi sebuah ranking, Support Vector Machine,  Nilai akurasi dari konfigurasi algoritme SVM  dengan preprocessing yang diterapkan penelitian tersebut untuk analisis sentimen pada teks berupa twit berbahasa Indonesia adalah sebesar 73.43%. Konfigurasi ini kemudian menjadi core aplikasi analisis sentimenNilai akurasi dari konfigurasi algoritme SVM  dengan preprocessing yang diterapkan penelitian ini untuk analisis sentimen pada teks berupa twit berbahasa Indonesia adalah sebesar 73.43%. Konfigurasi ini kemudian menjadi core aplikasi analisis sentimen | Pada penelitian ini data yang digunakan adalah twit pada media sosial twitter dan mengklasifikasikan komentar tersebut berdasarkan emotikon pada twit yang di analisis sedangkan penelitian yang penulis lakukan klasifikasi komentar berdasarkan adjective yang telah di kategorikan oleh penulis kedalam kategori positif, negatif dan netral |
| 3 | K-Nearest Neighbors Sebagai Analisis Sentimen Review Produk Appstore For Android, K-Neares Neighbors, Sucitra Sahara Mochammad Wahyudi, 2015 | Bagaimana menerapkan algoritma K-NN untuk menghasilkan nilai akurasi pada klasifikasi review appstore for android, K-Nearest Neighbors, hasil pengujian yang peneliti lakukan, pengukuran akurasi menggunakan confusion matrix dan kurva ROC membuktikan bahwa hasil pengujian algoritma K-NN sebesar 74.50% | Perhitungsn akurasi pada penelitian ini menggunakan confusion matrix sedangkan perhitungan akurasi yang dilakukan oleh penulis menggunakan metode naïve bayes dan menggunakan tools *rapid miner* |
| 4 | Analisis Sentimen dan Klasifikasi Kategori terhadap Tokoh publik pada Twitter, Ahmad Fathan Hidayatullah, Azhari SN, 2014 | Menganalisis *tweet* berbahasa Indonesia yang membicarakan tentang tokoh publik menjelang pemilihan umum 2014 di Indonesia, Support Vector Machine, Naïve Bayes, pengujian klasifikasi dengan fitur *term frequency* diperoleh sebesar 79.91% dan pengujian klasifikasi dengan fitur TF-IDF didapatkan sebesar 79.68%. berdasarkan data tsb *term frequency* menghasilkan performansi dengan akurasi lebih baik sebesar 0.23% | Penelitian yang dilakukan melakukan klasifikasi tweet berdasarkan kombinasi antara kelas sentimen dan kelas kategori, sedangkan penelitian yang penulis lakukan hanya menggunakan klasifikasi kelas sentimen nya saja |
| 5 | Analisis Sentimen Data Presiden Jokowi Dengan Preprosesing Normalisasi Dan Stemming Menggunakan Metode Naïve Bayes Dan SVM, Nurirwan Saputra, Teguh Bharata Adji, Adhistya Erna Permanasari, 2015 | Pemilihan kombinasi yang terbaik dari proses normalisasi data dan *stemming* terhadap akurasi analisis sentiment, Naïve Bayes dan SVM, akurasi yang dihasilkan ketika data dilakukan *stemming* terdapat peningkatan rata-rata sebesar 0.85% untuk metode naïve bayes dan 0.85% untuk metode SVM | Penelitian ini menggunakan *search techniques* dalam pengambilan data dan metode yang digunakan adalah Boolean *searching* sedangkan pengambilan data yang dilakukan penulis mengambil data satu bulan terakhir sebanyak jumlah yang dibutuhkan |
| 6 | Analisis Sentimen Terhadap Acara Televisi Berdasarkan Opini publik, Aditia Rakhmat Sentiaji, Adam Mukaharil Bachtiar, 2014 | Menjadikan sentimen publik sebagai indikator penilaian kualitas suatu acara di televise, Naïve Bayes Classifier, Dari hasil penelitian yang telah dilakukan terlihat  bahwa algoritma naïve bayes dapat mengklasifikasikan suatu opini yang berupa tweet ke dalam dua kelas yaitu positif dan negatif dengan akurat. Tingkat keakurasian dari pengklasifikasian tersebut sangat dipengaruhi oleh proses training. | Basis data yang digunakan dalam penelitian ini adalah basis data NoSql sedangkan basis data yan digunakan penulis adalah mySql |
| 7 | Pengklasifikasian Topik dan Analisis Sentimen dalam Media Sosial, YegarHeru Susilo, Tegar Heru Susilo, Siti Rochimah, 2013 | Mengklasifikasikan dokumen untuk membedakan status bertopik akademis dari yang non akademis dari status yang diunggah mahasiswa, Support Vector Machine, Kolaborasi metode SVM dengan pembobotan fitur Term Frequency – Inverse Document Frequency, dapat dipakai untuk melakukan klasifikasi topic dalam bahasa Indonesia dengan tingkat akurasi 93% | Penelitian ini dilakukan dengan mengklasifikasikan status dinding mahasiswa ke dalam kategori akademis dan non-akademis, sedangkan penelitian yang penulis lakukan mengklasifikasikan komentar ke dalam kategori positif, negatif dan netral |
| 8 | Analisa Sentimen Menggunakan Naïve Bayes Untuk Melihat Persepsi Masyarakat Terhadap Kenaikan Harga Jual Rokok pada Media Sosial Twitter, Fauziah Afshoh, 2017 | Menganalisa persepsi masyarakat terhadap kenaikan harga rokok pada media sosial twitter, naïve Bayes Classifier, hasil dari penelitian menunjukan bahwa dengan menggunakan data opini sebanyak 350 buah diketahui bahwa nilai sentimen positif yang paling banyak terbentuk dalam menanggapi topik kenaikan harga jual rokok dengan persentase tweets sebesar 53% | Pada penelitian ini membandingkan kinerja dari Naïve Bayes dan klasifikasi Lexicon Based, sedangkan penelitian yang penulis lakukan hanya menggunakan metode naïve bayes saja. |
| 9 | Penerapan Character N-Gram Untuk Sentiment Analysis Review Hotel Menggunakan Algoritma Naïve Bayes, Elly Indrayuni, Mochamad Wahyudi, 2015 | Bagaimana merangkum sekumpulan aplikasi dan website penyedia layanan wisata dalam satu review aplikasi dan mengelompokan opini-opini penggunanya ke dalam opini positif dan negatif secara otomatis, Naïve Bayes, N-Gram, penerepan algoritma naïve bayes terbukti menghasilkan nilai akurasi hingga 82.67% | Pada penelitian ini data yang digunakan untuk review sebanyak 300 data review hotel yang terdiri dari 150 review positif, 150 review negatif, sedangkan penelitian yang penulis lakukan mengelompokkan komentar kedalam 3 klasifikasi yaitu positif, negatif dan netral |
| 10 | Sentiment Analysis and Opinion Mining - A Facebook Posts and Comments Analyzer,  S. M. Junaid , S. W. Jaffry , M. M. Yousaf , L. Aslam , S. Sarwar, 2017 | Mendesain sebuah framework melalui data dan opini yang ada pada postingan dan komentar media sosial facebook untuk membantu pengguna untuk memahami sentimen diskusi di sebuah postingan atau komentar, binary classifier, menciptakan sebuah framework untuk menunjukan sentimen analisis dan opinion mining melalui postingan pengguna pada media sosial facebook | Pada penelitian ini penulis meciptakan sebuah framework yang memudahkan pengguna untuk memahami kategori postingan yang ada di facebook sehingga pengguna bisa dengan mudah bergabung dengan diskusi yang ada pada media sosial tersebut, sedangkan penelitian yang penulis lakukan adalah mengklasifikasikan kategori komentar untuk pihak perusahaan dan membantu pengguna untuk memberikan gambaran tentang kualitas aplikasi tersebut |

* 1. **Landasan Teori**
     1. **Smartphone**

*Smartphone* adalah telepon genggam yang mempunyai kemampuan tinggi, kadang-kadang dengan fungsi menyerupai komputer.  Bagi beberapa orang, *smartphone* merupakan telepon yang bekerja menggunakan seluruh perangkat lunak sistem operasi yang menyediakan hubungan standar dan mendasar bagi pengembang aplikasi [6].

1. **Macam-macam Sistem Operasi Smartphone**

Berikut adalah macam-macam sistem operasi pada *smartphone*: Android OS, BlackBerry OS, IOS, Windows Phone, Bada OS, Firefox OS, MeeGo OS, Palm OS, Symbian OS, Tizen OS, Ubuntu OS [6].

* + 1. **Sejarah awal Go-jek**

Bermula di Tahun 2010 sebagai perusahaan transportasi roda dua melalui panggilan telepon, Go-Jek kini telah tumbuh menjadi perusahaan layanan transportasi dan gaya hidup berbasis aplikasi. Go-jek adalah perusahaan teknologi dengan misi sosial untuk meningkatkan kesejahteraan dan mata pecaharian pekerja di berbagai sector informal di Indonesia. 3 esensi nilai yang di juarai oleh Go-Jek yaitu: kecepatan, inovasi, dan dampak sosial [7].

1. **Fitur aplikasi Go-jek**

Fitur yang ada pada aplikasi ini adalah: Go-Jek, Go Life, dan Go-Pay, dengan spesifikasi fitur sebagai berikut:

**Go-Jek.**

Go-Ride: Layanan transportasi online untuk sepeda motor.

Go-Car: Layanan transportasi online untuk mobil.

Go-Food: Layanan pemesanan makanan pada restoran-restoren tertentu.

Go-Mart: Layanan pemesanan kebutuhan rumah pada mini market.

Go-Send: Layanan pengiriman barang.

Go-Box: Layanan penyewaan taksi barang untuk pindah rumah.

Go-Tix: Layanan pemesanan Tiket untuk bioskop.

Go-Med: Layanan pemesanan kebutuhan medis.

**Go-Life**

Go-Messege: Layanan pemesanan pijat di rumah.

Go-Clean: Layanan kebersihan rumah.

Go-Glam: Layanan salon yang bisa datang ke lokasi pengguna.

Go-Auto: Layanan reparasi mesin otomotif.

**Go-Pay**

Go-Pay: Layanan pembayaran untuk transaksi secara non tunai.

Go Points: fitur untuk pengumpulan point yang didapatkan setiap kali menggunakan aplikasi go-jek point ini nantinya bisa digunakan untuk transaksi go-jek atau ditukar dengan sebuah hadiah.

Go-Pulsa: Layanan pembelian pulsa secara online untuk semua operator[7].

1. **Kelebihan aplikasi Go-jek diantara aplikasi transportasi online lainnya**

Kelebihan aplikasi Go-Jek dibandingkan aplikasi transpostasi online lainnya adalah, adanya layanan untuk gaya hidup sehari seperti yang sudah dijelaskan dalam point sebelumnya.

* + 1. **Analisis sentimen**

Analisis sentimen sendiri atau juga biasa disebut dengan opinion mining adalah salah satu bagian dari text mining. Bidang ini melakukan studi mengenai opini orang-orang, sentimen, evaluasi, tingkah laku dan emosi terhadap suatu entitas seperti produk, layanan, organisasi, individu, permasalahan, topik, acara dan atribut-atributnya [3].

* + 1. **Text mining**

Text mining adalah proses mengambil informasi berkualitas tinggi dari teks infor masi berkualitas tinggi biasanya diperoleh melalui per amalan pola dan kecenderungan melalui sarana seperti pembelajaran polastatistik [8].

* + 1. **Klasifikasi**

*Klasifikasi* adalah suatu proses pengelom-pokan data dengan didasarkan pada ciri-ciri tertentu ke dalam kelas-kelas yang telah ditentukan pula

* + 1. **Analisa Sentimen dengan Naïve Bayes Classifier.**

Naïve Bayes Classifier merupakan teknik pembelajaran algoritma data mining yang memanfaatkan metode probabilitas dan statistik. Naïve Bayes Classifier dalam melakukan klasifikasi terdapat dua proses penting yaitu learning (training) dan testing.

Proses klasifikasi Naïve Bayes Classifier terhadap dokumen yaitu dengan

mempresentasikan setiap dokumen dengan atribut “*X*1,*X2,X3, …., Xn*” yang mempunyai makna bahwa *X*1 untuk kata pertama, *X2* adalah kata kedua, dan seterusnya. Untuk himpunan kategori komentar dipresentasikan dengan *V*. Saat melakukan proses klasifikasi dokumen, Naïve Bayes Classifier akan mencari nilai probabilitas tertinggi dari [4]:



Keterangan:

VMAP: semua kategori yang diujikan

Vj: Kategori komentar, dengan:

J1: Sentimen positif

J2: Sentimen Negatif

J3: Sentimen Netral

*P(XI |VJ)*: probabilitas Xi pada kategori Vj

*P(Vj):* probabilitas dari Vj

* + 1. **Data *preprocessing***

Tahap preprocessing merupakan tahap awal untuk mengolah data teks menjadi analisa sentimen. Terdapat beberapa metode yang digunakan pada tahap preprocessing seperti:

1. *Parsing / Tokenizer* merupakan proses untuk memisahkan data teks menjadi beberapa token.

2. *Cleansing* adalah proses pembersihan data teks yaitu dengan menghilangkan data yang tidak konsisten atau tidak relevan.

3. *Normalization* digunakan untuk menormalisasikan kalimat. Banyak pengguna dalam menuliskan kata menggunakan singkatan sehingga akan menghambat proses pemahaman teks.

4. *Case folding* yaitu merubah bentuk kata menjadi bentuk dasarnya agar sebuah karakter dapat seragam (lower case)[4].

* + 1. **Rapid Minner**

Rapid Miner adalah aplikasi data mining yang tidak perlu dipertanyakan lagi dan berbasis sistem open-source. Tersedia sebagai aplikasi yang berdiri sendiri untuk analisis data dan sebagai mesin data mining untuk integrasi ke dalam produk sendiri. Ribuan aplikasi RapidMiner di lebih dari 40 negara memberikan pengguna mereka keunggulan yang kompetitif. Solusi yang di usung antara lain :Integrasi data, Analitis ETL, Data Analisis, dan Pelaporan dalam satu suite tunggal.

1. **Kerangka Berfikir**

MASALAH

1. Rating yang diberikan oleh *user* terhadap aplikasi Go-jek belum menggambarkan kondisi penerimaan pengguna terhadap aplikasi tersebut.

2. Pengguna masih kesulitan faktor apa yang mengharuskan pengguna menggunakan aplikasi Go-jek.

TUJUAN

Tujuan dari penelitian yang dilakukan ini adalah, menghitung presentase komentar positif, negatif dan netral yang diberikan pengguna terhadap aplikasi go-jek, sehingga memudahkan pengguna dalam memutuskan menggunakan aplikasi ini atau tidak berdasarkan presentase komentar yang diberikan.

DESIGN PENELITIAN

SOLUSI

Naïve Bayes Classifier

LANDASAN TEORI

PENDEKATAN

TOOLS

DATA

1. Aplikasi Go-jek
2. Naïve Bayes Classifier
3. Rapid Minner
4. Analisa Sentimen
5. Klasifikasi

Komentar pengguna pada aplikasi Go-Jek di play store

Rappid Minner

HASIL

.

menghitung presentase komentar positif, negatif dan netral yang diberikan pengguna terhadap aplikasi go-jek, sehingga memudahkan pengguna dalam memutuskan menggunakan aplikasi ini atau tidak berdasarkan presentase komentar yang diberikan.

1. Mengetahui presentase komentar positif, negatif, dan netral berdasarkan komentar pengguna
2. Memudahkan pengguna dalam keputusan penggunaan aplikasi ini.
3. Mengetahui seberapa besar tingkat akurasi metode naïve bayes terhadap klasifikasi komentar aplikasi Go-jek.

MANFAAT

Gambar 2.1 Kerangka Berfikir

# METODE PENELITIAN

**Metode penelitian yang digunakan adalah :**

## Deskriptif Khusus

Penelitian ini bertujuan untuk menghitung presentase komentar yang diberikan pengguna pada aplikasi Go-jek, sehingga pengguna dapat mengetahui faktor apa yang mengharuskan pengguna menggunakan aplikasi Go-jek dan juga membantu pihak perusahaan dalam merancang pengembangan aplikasi di masa depan. Berikut tahapan-tahapan yang dilakukan pada penelitian ini:

Studi Literatur

pengumpulan data

Input data

Data preprocessing

Klasifikasi Naive Bayes

Menghitung presentase komentar

Menghitung akurasi dari pengujian data

Gambar 3.1 Tahapan penelitian

Gambar diatas merupakan tahapan penelitian yang penulis lakukan dalam melakukan penelitian ini adapun rincian jelasnya adalah sebagai berikut:

1. Studi Literatur

Studi literatur ini dilakukan untuk mengetahui trend penelitian yag terjadi pada bidang analisa sentimen, studi literatur ini dilakukan untuk meriview jurnal-jurnal yang telah ada dan digunakan untuk menentukan studi kasus serta metode yang akan digunakan oleh penulis saat melakukan penelitian.

1. Pengumpulan Data

Pengumpulan data ini dilakukan untuk memperkuat alasan mengapa penelitian ini harus dilakukan, pengumpulan data diambil dari sumbernya langsung yaitu *website* resmi Go-jek serta komentar pengguna aplikasi Go-jek pada Google Play Strore.

1. Input Data

Data yang di input pada tahapan ini adalah data komentar pengguna pada aplikasi Go-jek selama satu bulan terakhir sebanyak 200 komentar dengan klasifikasi komentar negatif, positif dan netral

1. Data *preprocessing*

Tahap *preprocessing* adalah tahap awal dalam pengolahan data teks menjadi analisa sentimen. Terdapat beberapa metode yang digunakan pada tahap *preprocessing* diantaranya adalah:

1. *Parsing* merupakan proses untuk memisahkan data teks menjadi beberapa token.
2. *Cleansing* merupakan proses pembersihan data teks yaitu dengan menghilangkan data yang tidak konsisten.
3. *Normalization* digunakan untuk menormalisasi kalimat. Seperti kata singkatan, contohhnya “tp” jika dinormalisasi menjadi “tapi” lalu kata “gara2” dinormalisasi menjadi “gara-gara”.
4. *Case Folding* yaitu merubah bentuk kata menjadi bentuk dasarnya agar sebuah karakter dapat seragam (*lower case)*. Contohnya “menyebalkan” dirubah mejadi kata dasarnya yaitu “sebal”
5. Klasifikasi Naïve Bayes

Tahapan selanjutnya yaitu klasifikasi menggunakan naïve bayes, pada tahapan ini akan melalui dua tahap lagi yaitu data *training* dan data *testing.* Data training berjumlah 150 data komentar dengan klasifikasi positif, negatif, dan netral yang diklasifikasikan secara manual lalu data testing sebanyak 200 data komentar dengan klasifikasi yang sama.

1. Menghitung presentase komentar

Pada tahap ini akan dihitung seberapa besar presentase komentar positif, negatif dan netral yang didapatkan dari data *testing.*

1. Menghitung akurasi dari pengujian data

Pada tahapan ini akan dilakukan menggunakan *tools Rapid Minner* untuk menguji seberapa besar akurat algoritma naïve bayes untuk pada dataset penelitian ini.

1. **TEKNIK PENGUMPULAN DATA DAN ANALISA DATA**
2. **Teknik pengumpulan Data**

Berikut adalah tahapan pengumpulan data yang penulis lakukan dalam menyusun penelitian ini:

## Observasi

Observasi ini dilakukan dalam mengamati perkembangan aplikasi go-jek serta perkembangan pengguna dan transaksi yang dilakukan masyarakat pada aplikasi ini, observasi ini juga dilakukan untuk mengamati komentar-komentar yang diberikan masyarakat mengenai aplikasi ini, adanya banyak macam komentar yang diberikan oleh pengguna pada aplikasi ini yang membuat penulis melakukan penelitian dalam presentase klasifikasi komentar masyarakat pada aplikasi Go-jek.

1. Studi Dokumentasi

Dalam studi dokumentasi ini penulis mengumpukan data sebagai berikut:

* 1. Kurva pengguna dan transaksi apliksi go-jek
  2. Klasifikasi ranking yang diberikan pengguna pada aplikasi go-jek
  3. Ranking aplikasi go-jek diantara aplikasi transportasi online lainnya

1. **Teknik Analisa Data**

Teknik analisa data yang penulis gunakan dalam menyusun penelitian ini adalah sebagai berikut:

1. **Teknik Pengumpulan Data**

Data diambil dari kegiatan observasi dan studi dokumentasi didapatkan beberapa informasi sebagai berikut:

1. Terjadi gap antara daftar pengunduh aplikasi dan transaksi pengguna pada aplikasi go-jek
2. Mendapatkan jumlah masing-masing bintang yang diberikan oleh pengguna
3. Mengetahui ranking aplikasi go-jek diantara aplikasi transportasi online yang lainnya
4. **Teknik Pengolahan Data (Reduksi Data)**

Berdasarkan data yang diperoleh dari beberapa kegiatan seperti observasi dan studi dokumentasi, penulis kemudia melakukan reduksi data. Yaitu memisahkan data yang tidak diperlukan seperti informasi lowongan pekerjaan pada perusahaan go-jek, data driver go-jek, dan jumlah driver pada perusahaan go-jek. Dengan begitu didapatkan informasi yang dibutuhkan untuk dalam penelitian ini informasi tersebut adalah informasi tentang pengguna aplikasi go-jek serta opini masyarakat yang diberikan pada aplikasi ini, sehingga penulis dapat menyimpulkan tentang pentingnya melakukan penelitian ini.

1. **Menarik Kesimpulan**

Selanjutnya penulis menarik kesimpulan setelah melakukan observasi dan studi dokumentasi bahwasannya belum ada penelitian yang membahas tentang presentase komentar pada aplikasi go-jek dan beberapa permasalahan yang lainnya maka penulis memutuskan untuk melakukan penelitian ini.

# DAFTAR PUSTAKA

[1] kemenristekdikti, “Smartphone Rakyat Indonesia,” 2017. [Online]. Available: https://ristekdikti.go.id/smartphone-rakyat-indonesia-2/. [Accessed: 07-Jan-2018].

[2] W. Kota, “Ini Hasil Lengkap Survei Transportasi Online YLKI, 41 Persen Pengguna Pernah Dikecewakan,” 2017. [Online]. Available: http://wartakota.tribunnews.com/2017/05/12/ini-hasil-lengkap-survei-transportasi-online-ylki-41-persen-pengguna-pernah-dikecewakan. [Accessed: 07-Jan-2018].

[3] J. Prager, “Open-Domain Question–Answering,” *Found. Trends® Inf. Retr.*, vol. 1, no. 2, pp. 91–231, 2006.

[4] P. Ilmiah, F. Afshoh, P. S. Informatika, F. Komunikasi, D. A. N. Informatika, and U. M. Surakarta, “Analisa Sentimen Menggunakan Naïve Bayes,” 2017.

[5] A. F. Hidayatullah and A. Sn, “Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter,” *Semin. Nas. Inform. 2014*, vol. 2014, no. August 2013, pp. 0–8, 2014.

[6] Anonymous, “Pengertian Smartphone,” 2015. [Online]. Available: http://kliknklik.com/content/66-pengertian-smartphone. [Accessed: 06-Feb-2018].

[7] Gojek, “About Gojek,” 2010. [Online]. Available: https://www.go-jek.com/about/. [Accessed: 06-Jan-2018].

[8] andi nurul Hidayat, “Analisis Sentimen Terhadap Wacana Politik Pada Media Masa Online Menggunakan Algoritma Support Vector Machine Dan Naive Bayes,” *J. Elektron. Sistim Inf. Dan Komput.*, vol. 1, no. 1, pp. 1–7, 2015.