

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB II

### LANDASAN TEORI

#### 2.1 Transportasi *Online*

Transportasi *online* merupakan transportasi yang penggunaannya memanfaatkan teknologi *smartphone*. Contoh salah satu transportasi *online* adalah ojek *online*. Dimana ojek *online* merupakan *trend* terbaru dari kalangan masyarakat. Ojek *online* memberikan kemudahan yaitu dalam pemesanannya dengan menggunakan teknologi *smartphone*. Layanan utama ojek *online* adalah mengantar penumpang ke tempat tujuan. Namun selain itu juga melayani antar jemput barang, pemesanan makanan/minuman, belanja, tiket dan sebagainya (Marikxon, 2015). Cara kerja ojek *online* yaitu dengan membuka aplikasi yang terpasang di *smartphone*, lalu memilih layanan yang diinginkan dan mengkonfirmasi pesanan.

Beberapa perusahaan jasa transportasi berbasis aplikasi *online* di Indonesia yaitu :

##### 1) *GO-JEK*

*GO-JEK* merupakan perusahaan berjiwa sosial yang memimpin revolusi industri transportasi Ojek. *GO-JEK* bekerjasama dengan para pengemudi ojek berpengalaman di Jakarta, Bandung, Bali & Surabaya dan menjadi solusi utama dalam pengiriman barang, pesan antar makanan, berbelanja dan berpergian ditengah kemacetan.

Dengan menggunakan *GO-JEK APP*, anda dapat memesan *GO-JEK Driver* untuk mengakses semua layanan ini. Dengan cara masukkan alamat anda untuk mengetahui biaya penggunaan layanan. Gunakan layanan '*Use My Location*' untuk mengarahkan *Driver* ke tempat anda berada.

Setelah anda mengkonfirmasi pesanan, teknologi *location-based* kami akan mencari *Driver* yang posisinya bedekatan dengan anda. Setelah seorang *Driver*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ditugaskan, anda dapat melihat foto *Driver*, mengirimkan sms dan juga dapat menelponnya. (Rizkyawan, Aziz,.dkk. 2017)

2) *Grab*

*Grab*(sebelumnya dikenal sebagai *GrabTaxi*) adalah sebuah perusahaan asal Singapura yang melayani aplikasi penyedia transportasi dan tersedia di enam negara di Asia Tenggara, yakni Malaysia, Singapura, Thailand, Vietnam, Indonesia,dan Filipina.*Grab* memiliki visi untuk merevolusi industri pertaksian di Asia Tenggara, sehingga dapat memberikan keamanan dan kenyamanan bagi pengguna kendaraan seantero Asia Tenggara.Hingga bulan Maret 2015, jumlah pengguna *Grab* mencapai 3,8 juta pengguna. *Grab* tersedia untuk sistem operasi *Android*, *iOS*, dan *BlackBerry*. Di Indonesia, *Grab* melayani pemesanan kendaraan seperti ojek, mobil, dan taksi. Saat ini *Grab* tersedia di seluruh Jakarta dan sekitarnya.

Jenis-jenis Layanan *Grab* merupakan aplikasi layanan yang menyediakan layanan transportasi untuk menghubungkan lebih dari 10 juta penumpang dan 185.000 pengemudi di seluruh wilayah Asia Tenggara.

Aplikasi *Grab* menawarkan 5 pilihan layanan transportasi mulai dari taksi, mobil pribadi, sepeda motor hingga pengiriman paket untuk memenuhi kebutuhan penumpang di Singapura, Malaysia, Thailand, Filipina, Vietnam, dan Indonesia. Layanan yang ditawarkan antara Lain :

- a. *GrabTaxi* Layanan taksi premium yang menghubungkan antara pengemudi taksi dan calon penumpang taksi.
- b. *GrabCar* Layanan transportasi untuk mereka yang memilih kenyamanan berkendara layaknya menggunakan mobil pribadi, menghubungkan antara pemilik/pengemudi mobil pribadi dan calon penumpang/penyewa.
- c. *GrabBike* Sebuah alternatif layanan transportasi untuk mereka yang ingin lebih cepat dan aman sampai ke tujuan dengan menggunakan sepeda motor. Menghubungkan antara pemilik/pengemudi sepeda motor dengan calon penumpang.
- d. *GrabExpress* Layanan pengiriman paket yang cepat, aman dan terpercaya.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Sejak diluncurkan pada 2012, aplikasi *Grab* telah diunduh di lebih dari 13 juta perangkat di seluruh Asia Tenggara. Lebih dari 250.000 pengemudi di jaringan *Grab*, termasuk seluruh layanan di bawah aplikasi *Grab*, telah menggunakan aplikasi ini untuk menerima pemesanan layanan transportasi di Singapura, Indonesia, Filipina, Malaysia, Thailand dan Vietnam. Aplikasi *Grab* dan seluruh layanannya saat ini tersedia di 30 kota di enam negara di Asia Tenggara (Satria dan Yudhi, 2016).

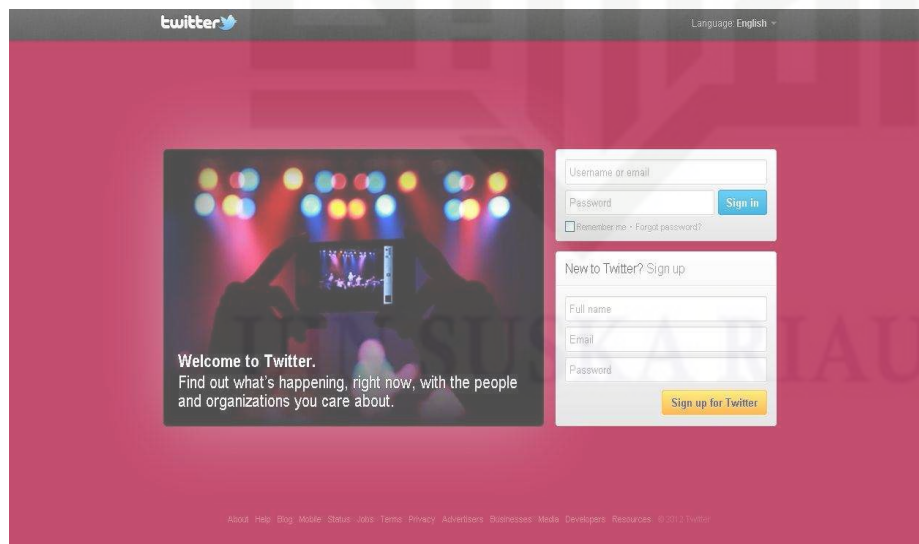
Ada beberapa tujuan dan manfaat lahirnya jasa transportasi berbasis aplikasi *online*, diantaranya sebagai berikut:

1. Praktis dan mudah digunakan, layanan jasa transportasi berbasis aplikasi *online* mencakup menggunakan telpon pintar yang sudah menggunakan internet dan aplikasi jasa transportasi *online* yang ada didalamnya, kita dapat melakukan pemesanan layanan jasa transportasi.
2. Transparan, dengan jasa transportasi berbasis *online* ini juga memungkinkan pelanggan mengetahui dengan pasti setiap informasi jasa transportasi *online* secara detail seperti nama *driver*, nomor kendaraan, posisi kendaraan yang akan dipakai, waktu perjalanan, lisensi perusahaan dan lain sebagainya.
3. Lebih terpercaya, maksudnya disini lebih terpercaya adalah parapengemudi atau *driver* sudah terdaftar didalam perusahaan jasa transportasi berbasis aplikasi *online* ini berupa identitas lengkap dan perlengkapan berkendara yang sesuai Standar Nasional Indonesia (SNI) sehingga dapat meminimalisir resiko kerugian terhadap pengguna jasa transportasi lain.
4. Adanya asuransi kecelakaan bagi pengguna dan pengemudi, salah satu perusahaan jasa transportasi berbasis *online* yaitu *GO-JEK* telah melakukan kerja sama dengan perusahaan asuransi Allianz dalam memberikan perlindungan asuransi kecelakaan bagi para pengguna jasa transportasi *GO-JEK*.

## 2.2 Twitter

### 2.2.1 Definisi Twitter

Twitter merupakan sosial media yang menjadi salah satu alat berkomunikasi melalui jaringan internet yang sedang digandrungi oleh pengguna internet dan sosial media pada saat ini. Twitter adalah sebuah situs web yang dimiliki dan dioperasikan oleh Twitter Inc, yang menawarkan jaringan sosial berupa *mikroblog* sehingga memungkinkan penggunanya untuk mengirim dan membaca pesan *tweets* (Twitter, 2013). *Mikroblog* adalah salah satu jenis alat komunikasi *online* dimana pengguna dapat memperbarui status tentang mereka yang sedang memikirkan dan melakukan sesuatu, apa pendapat mereka tentang suatu objek atau fenomena tertentu. *Tweets* adalah teks tulisan hingga 140 karakter yang ditampilkan pada halaman profil pengguna. *Tweets* bisa dilihat secara publik, namun pengirim dapat membatasi pengiriman pesan ke daftar teman-teman mereka saja. Pengguna dapat melihat *Tweets* pengguna lain yang dikenal dengan sebutan pengikut (*follower*). Twitter didirikan pada Maret 2006 oleh perusahaan rintisan *Obvios Corp* (Hadi, 2010)



**Gambar 2.1 Tampilan Homepage Twitter (Sumber: [www.Twitter.com](http://www.Twitter.com))**

Semua pengguna dapat mengirim dan menerima *Tweets* melalui situs Twitter, aplikasi eksternal yang kompatibel, atau dengan pesan singkat (SMS) yang tersedia

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

di negara-negara tertentu (Twitter, 2013) .pengguna dapat menulis pesan berdasarkan topik dengan menggunakan *hashtag* (#). Sedangkan untuk membalas atau menyebutkan pesan dari pengguna lain bisa menggunakan tanda @.

### 2.2.2 Manfaat Twitter

Mengacu pada buku *Welcome To Twitterland* oleh Ali Akbar (Akbar, Ali, 2012), berikut adalah 8 manfaat dari Twitter:

a. Media Silaturahmi

Melalui Twitter, pemakai dapat bersilaturahmi dengan saudara, teman ataupun kerabat yang jauh jaraknya. Komunikasi bisa menjadi akrab dan tidak terbatas pada ruang dan waktu.

b. Mendapatkan Berita Terkini

Banyak berita terkini yang dapat diperoleh melalui Twitter. Misalnya saja melalui akun Twitter @detik.com, para *follower* dapat mengakses berita terkini yang di-*update* oleh akun tersebut secara langsung setiap hari.

c. Motivasi

Melalui akun Twitter, *follower* pun dapat mendapatkan motivasi dan semangat hidup. Misalnya saja dengan mem-*follow* @MotivasiTweet, dimana para *follower*-nya bisa mendapatkan berbagai nasehat dan motivasi hidup yang dapat membangkitkan semangat dan pikiran positif.

d. Informasi

Informasi yang bermanfaat pun dapat diperoleh melalui Twitter. Misalnya ketika ingin mencari informasi produk halal, maka pengguna Twitter dapat mem-*follow* @halalcorner.

e. Pemecahan Masalah / Konseling

Pengguna Twitter bisa saja mendapatkan pemecahan / konseling atas masalah yang dihadapinya dengan ketika pengguna meng-*update* status yang berhubungan dengan masalahnya dan mendapatkan beragam *feedback* /jawaban dari *follower*-nya yang dapat berupa masukan dan saran yang dapat dijadikan sebagai alternatif atas masalahnya.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

f. Media Hiburan

Manfaat hiburan pun merupakan salah satu manfaat Twitter yang sangat penulis rasakan sebagai salah satu pengguna Twitter. Berbagai lelucon yang memberikan penghiburan dapat didapatkan melalui @pepatahgombal, @rajagombal ataupun @radityadika (penulis buku humor ternama “Kambing Jantan”).

g. Cari Jodoh

Biasanya pengguna Twitter yang merasa tertarik dengan pengguna akun Twitter lainnya akan mem-follow akun tersebut dan berusaha berkenalan dengan pemakai akun Twitter tersebut.

h. Sumber Rejeki

Sampai saat ini, telah banyak sekali perusahaan dan pebisnis (konvensional maupun online) yang memanfaatkan Twitter sebagai salah satu media untuk mencari keuntungan, dengan melakukan berbagai tindakan promosi baik melalui pemasangan iklan, update status yang berbaur “bisnis” dan penawaran lainnya.

Pada buku Twitter API: *Up and Running* (Makice, 2009:27), dituliskan bahwa terdapat studi yang menyimpulkan bahwa terdapat empat fungsi pemakaian Twitter oleh para membernya yaitu:

- a) Daily Chatter  
Membicarakan mengenai aktivitas dan rutinitas harian.
- b) Bercengkrama  
Menggunakan “@” untuk berbicara secara spesifik kepada pengguna Twitter lainnya.
- c) Berbagi Informasi  
Membagikan pesan informasi kepada pengguna lainnya.
- d) Melaporkan Berita  
Melaporkan berita dan perkembangan informasi kepada pengguna lain.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 2.3 Sentimen

### 2.3.1. Pengertian Sentimen

Menurut *Merriam-Webster Online Dictionary* (2016), sentimen memiliki dua pengertian. Pertama, sentimen adalah sebuah sikap, pemikiran atau penilaian yang berdasarkan pada perasaan yang disebut dengan opini. Kedua, sentimen adalah perasaan yang muncul ketika mengalami nostalgia atau peristiwa yang sentimental (Rahmad, 2017). Penelitian ini menitik beratkan pada pengertian sentimen yang pertama, yaitu sentimen sebagai sikap, pemikiran atau penilaian yang berdasarkan pada perasaan yang disebut dengan opini.

Sentimen yang terdapat dalam teks merupakan emosi, penilaian atau opini terhadap sebuah subjek yang diekspresikan oleh penulis teks. Proses menentukan sentimen secara keseluruhan dari sebuah teks, baik teks yang berupa dokumen atau kalimat disebut dengan klasifikasi sentimen (Akkaya & Zhang, 2001).

### 2.3.2 Analisis Sentimen

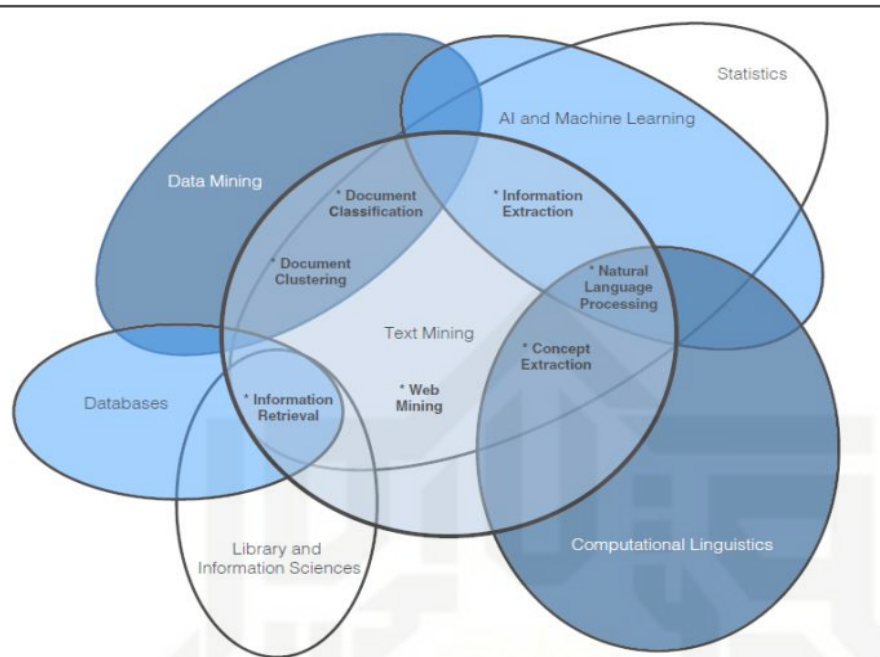
Sentimen analisis atau opinion mining mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi *linguistic* dan *text* mining yang bertujuan menganalisa pendapat, sentiment, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenaan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu (Saraswati, 2011).

Sentimen Analysis atau opinion merupakan proses memahami, mengekstrak dan mengolah data terstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Sentiment analysis dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif (Bopang, 2002).

Sentimen Analisis dapat dibedakan berdasarkan sumber datanya, beberapa level yang sering digunakan dalam penelitian Sentimen Analisis adalah Sentimen Analisis pada level dokumen dan Sentimen Analisis pada level Kalimat (Fink Clayton, 2011)







**Gambar 2.2 Diagram Venn 6 bidang terkait dan 7 area praktek text mining (Sumber: Megawati, 2015)**

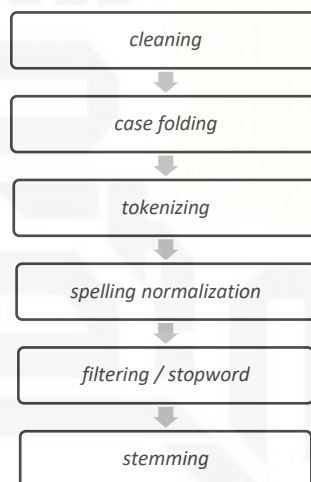
Berdasarkan Gambar 2.2 dapat dijelaskan bahwa ;

1. Pencarian dan perolehan informasi (*search and information retrieval*), yaitu penyimpanan dan penggalian dokumen teks misalnya dalam mesin pencarian (*search engine*) dan pencarian kata kunci (*keywords*).
2. Pengelompokan dokumen, yaitu pengelompokan dan pengkategorian kata, istilah, paragraf, atau dokumen dengan menggunakan metode *cluster (clustering)* Data mining.
3. Klasifikasi dokumen, yaitu pengelompokan dan pengkategorian kata, istilah, paragraf atau dokumen dengan menggunakan metode klasifikasi (*classification*) data mining berdasarkan model terlatih yang sudah memiliki label.
4. *Web mining*, yaitu penggalian informasi dari internet dengan skala fokus yang spesifik.
5. Ekstraksi informasi (*information extraction*), yaitu mengidentifikasi dan mengekstraksi informasi dari data yang sifatnya semi-terstruktur atau tidak terstruktur dan mengubahnya menjadi data yang terstruktur.

6. *Natural language processing* (NLP), yaitu pembuatan program yang memiliki kemampuan untuk memahami bahasa manusia.
7. Ekstraksi konsep, yaitu pengelompokan kata atau frase ke dalam kelompok yang mirip secara semantik.

#### 2.4.1. *Preprocessing* (Pra-proses)

Menurut (Feldman & Sanger, 2007) yang dikutip oleh (Indranandita, Amelia, & dkk, 2008) *Text preprocessing* merupakan tahapan proses awal terhadap teks untuk mempersiapkan teks menjadi data yang akan diolah selanjutnya. Sekumpulan karakter yang bersambungan (teks) harus dipecah-pecah menjadi unsur yang lebih berarti, yang dapat dilakukan dalam tingkatan yang berbeda. Tahapan yang dilakukan dari *text preprocessing* dapat dilihat pada gambar 2.3



**Gambar 2.3** Proses *text preprocessing*

Tahapan *preprocessing* terdiri dari proses *cleaning*, *case folding*, *tokenizing*, *spelling normalization*, *filtering* dan *stemming* (Indranandita, Amelia, & dkk, 2008).

##### 1. *Cleaning*

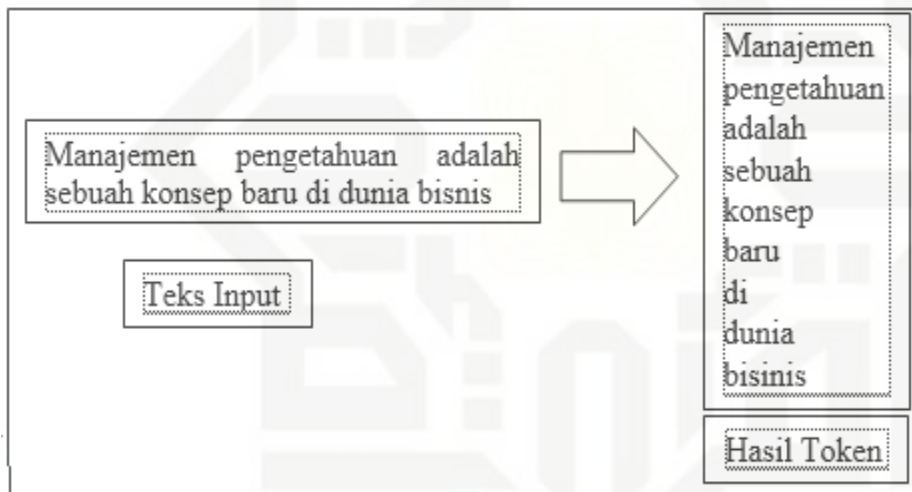
*Cleaning* adalah proses untuk membersihkan dokumen dari kata-kata yang tidak diperlukan untuk mengurangi *noise* pada proses klasifikasi. Kata yang dihilangkan adalah karakter HTML, *hashtag* (#), *username* (@username), *url* (<http://situs.com>), dan *emoticon*.

## 2. Case Folding

Proses penyeragaman bentuk huruf dengan mengubah semua huruf menjadi huruf kecil, dan juga menghilangkan tanda baca dan angka, dalam hal ini hanya menggunakan huruf antara a sampai z.

## 3. Tokenizing

*Tokenizing* adalah proses pemotongan *string* input berdasarkan kata yang menyusunnya, dan *tokenizing* juga dapat diartikan sebuah proses memecah dokumen atau kalimat menjadi sebuah kata dengan melakukan analisa terhadap kumpulan kata dengan memisahkan kata tersebut dan menentukan struktur sintaksis dari tiap kata tersebut. Contoh proses *tokenizing* pada Gambar 2.4 berikut :



Gambar 2.4 Proses Tokenisasi

## 2.5 Stopword

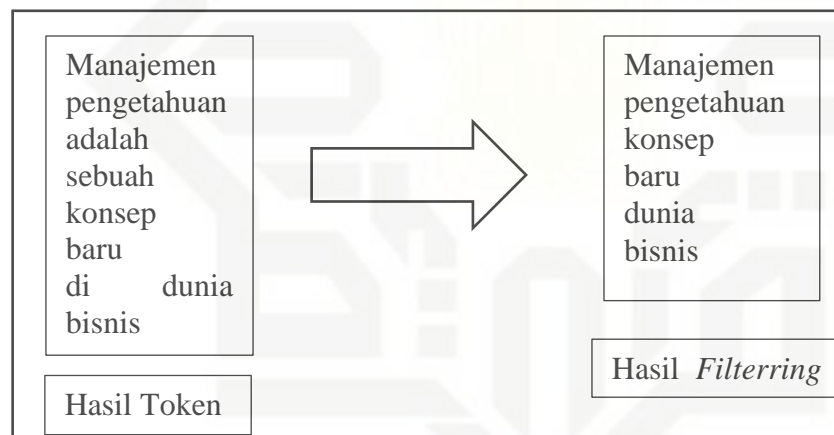
*Stopword* merupakan kumpulan kata-kata yang sering muncul dalam suatu dokumen. *Stopword* pada umumnya adalah sebuah kata penghubung yang tidak begitu penting, maka *stopword* dapat diabaikan dan tidak ikut dalam proses pengindeksan *stopword* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words*. Contoh *stopword* adalah "yang", "dan", "di", dan sebagainya.

#### 4. *Spelling Normalization*

Merupakan perbaikan dan substitusi kata-kata yang salah eja ataupun disingkat dengan bentuk tertentu. Substitusi kata dilakukan untuk menghindari jumlah perhitungan dimensi kata yang melebar. Perhitungan dimensi kata akan melebar jika kata yang salah eja atau disingkat tidak diubah karena kata tersebut sebenarnya memiliki kontribusi dalam merepresentasikan dokumen tetapi akan dianggap sebagai entitas yang berbeda proses penyusunan matriks.

#### 5. *Filtering*

*Filtering* adalah tahap mengambil kata-kata penting dari hasil token. Biasanya tahap ini menggunakan algoritma *stop-list* (membuang kata-kata kurang penting) atau *word-list* (menyimpan kata penting).



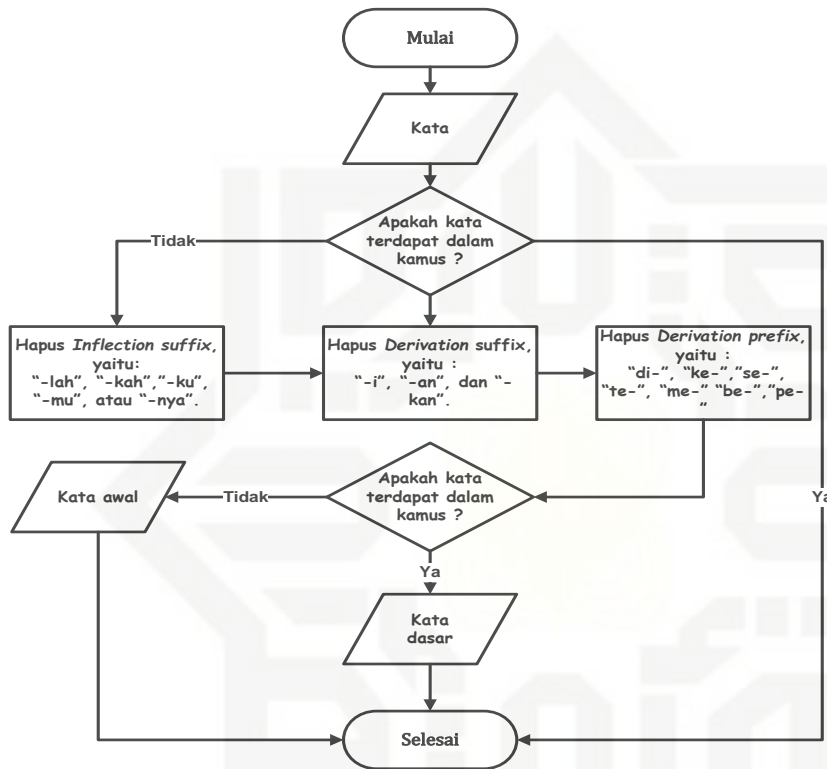
**Gambar 2.5** Proses *Filtering* atau Penyaringan Kata Penting

#### 6. *Stemming*

*Stemming* adalah tahapan mencari kata *root* / kata dasar dari setiap kata hasil dari proses *filtering*. Karena data komentar yang akan diklasifikasi menggunakan bahasa Indonesia maka algoritma *stemming* untuk berbahasa Indonesia yang mempunyai tingkat keakuratan yang lebih baik dibanding algoritma lainnya adalah algoritma Nazief & Andriani (Agusta, 2009).

Proses *stemming* menggunakan Algoritma Nazief dan Adriani dapat dilihat pada gambar 2.6 dibawah ini. Proses *Stemming* pada teks bahasa Indonesia lebih

rumit karena terdapat variasi imbuhan yang harus dibuang untuk mendapat *root word* (kata dasar) dari sebuah kata. Algoritma ini mengacu pada aturan KBBI (Kamus besar bahasa Indonesia) yang mengelompokkan imbuhan yang diperbolehkan atau imbuhan yang tidak diperbolehkan.



**Gambar 2.6 Flowchart Algoritma Nazief dan Adriani**

Berikut merupakan langkah-langkah yang dilakukan oleh algoritma Nazief dan Adriani (Agusta, 2009).

1. Kata yang belum di *stemming* dicari pada KBBI. Apabila kata langsung ditemukan, berarti kata tersebut adalah kata dasar, kata dikembalikan dan algoritma dihentikan.
2. Hilangkan *inflectional suffixes* terlebih dahulu, jika ini berhasil dan *suffix* adalah pertikel (“lah” atau “kah”), langkah ini dilakukan lagi untuk menghilangkan *inflectional possessive pronoun suffixes* (“ku”, “mu” atau “nya”).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. Partikel *Derivational suffix* (“i”, “-an”, “-kan”) kemudian dihilangkan, langkah dilanjutkan lagi untuk mengecek apakah masih ada *derivational suffix* yang tersisa, jika ada maka akan dihilangkan. Apabila tidak ada lagi maka lakukan langkah selanjutnya.
4. *Derivational prefix* (“di-”, “ke-”, “se-”, “te-”, “me-”, “be-”, “pe-”) dihilangkan, kemudian langkah dilanjutkan lagi untuk mengecek apakah masih ada *derivational prefix* yang tersisa, jika ada maka akan dihilangkan. Apabila tidak ada lagi maka lakukan langkah selanjutnya.
5. Setelah tidak ada lagi imbuhan yang tersisa, maka algoritma ini dihentikan kemudian kata dasar tersebut di cari pada KBBI, jika kata ditemukan berarti algoritma ini berhasil tapi jika kata dasar tidak ditemukan maka dilakukan *recoding*.
6. Jika semua langkah telah dilakukan tetapi kata dasar tidak ditemukan pada kamus, maka algoritma ini mengembalikan kata yang asli sebelum dilakukan *stemming*.

#### 2.4.2. Transformation

Pada tahapan ini pemrosesan teks dilanjutkan dengan proses transformasi teks menjadi data numerik sebagai representasi dari setiap dokumen. Pada *text transformation* ini kita hanya menentukan (TF) saja, yaitu jumlah frekuensi kemunculan kata dalam dokumen tersebut.

### 2.6 Naïve Bayes Classifier

Algoritma *Naive Bayes Classifier* merupakan algoritma yang digunakan untuk mencari nilai probabilitas tertinggi untuk mengklasifikasi data uji pada kategori yang paling tepat (Feldman & Sanger, 2007). Dalam penelitian ini yang menjadi data uji adalah dokumen *tweets*. Ada dua tahap pada klasifikasi dokumen. Tahap pertama adalah pelatihan terhadap dokumen yang sudah diketahui kategorinya. Sedangkan tahap kedua adalah proses klasifikasi dokumen yang belum diketahui kategorinya.

Sebuah keuntungan dari *naive bayes classifier* adalah bahwa ia memerlukan sejumlah kecil data pelatihan untuk mengestimasi parameter (rata-rata dan varian

dari variabel) yang diperlukan untuk klasifikasi. Karena variabel diasumsikan independen, hanya varian dari variabel-variabel untuk setiap kelas yang perlu ditentukan dan bukan keseluruhan *covariance matrix*. Penghitungan nilai probabilitas tersebut menggunakan persamaan, (Jurafsky, 2011) :

$$P(c) = \frac{N_c}{N} \quad (2.1)$$

$N_c$  = Banyak dokument dalam suatu kelas (n)

$N$  = Jumlah keseluruhan dokumen data latih

dan

$$P(w|c) = \frac{\text{count}(w,c+1)}{\text{count}(c)+|V|} \quad (2.2)$$

Count (w,c) = Frekuensi kata **w** pada kelas **c**

Count (c) = Total frekuensi kata pada masing-masing kelas **c**

$|V|$  = Total kata unik pada keseluruhan kelas **c**

Selanjutnya,

$$P(c | dn) = P(c) * \prod p(w|c) \quad (2.3)$$

$P(c|dn)$  = *Choosing a class*

$P_c$  = *Priors*

$\prod p(w|c)$  = *Total Conditional Probabilities*

## 2.7 Black Box

*Black box testing* adalah pengujian yang dilakukan hanya mengamati hasil eksekusi melalui data uji dan memeriksa fungsional dari perangkat lunak. Jadi dianalogikan seperti kita melihat suatu kotak hitam, kita hanya bisa melihat penampilan luarnya saja, tanpa tau ada apa dibalik bungkus hitam nya. Sama seperti pengujian *blackbox*, mengevaluasi hanya dari tampilan luarnya (*interface* nya), fungsionalitasnya tanpa mengetahui apa sesungguhnya yang terjadi dalam proses detilnya (Mustaqbal, 2015).

Metode uji dapat diterapkan pada semua tingkat pengujian perangkat lunak, unit, integrasi, fungsional, sistem dan penerimaan. Ini biasanya terdiri dari kebanyakan jika tidak semua pengujian pada tingkat yang lebih tinggi, tetapi juga bisa mendominasi unit testing juga.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

*Blackbox Testing* cenderung untuk menemukan hal-hal berikut:

1. Fungsi yang tidak benar atau tidak ada.
2. Kesalahan antarmuka (*interface errors*).
3. Kesalahan pada struktur data dan akses basis data.
4. Kesalahan performansi (*performance errors*).
5. Kesalahan inisialisasi dan terminasi.

Pengujian didesain untuk menjawab pertanyaan-pertanyaan berikut:

1. Bagaimana fungsi-fungsi diuji agar dapat dinyatakan valid?
2. Input seperti apa yang dapat menjadi bahan kasus uji yang baik?
3. Apakah sistem sensitif pada input-input tertentu?
4. Bagaimana sekumpulan data dapat diisolasi?
5. Berapa banyak rata-rata data dan jumlah data yang dapat ditangani sistem?
6. Efek apa yang dapat membuat kombinasi data ditangani spesifik pada operasi sistem?

Kelebihan *Blackbox Testing* :

1. Spesifikasi program dapat ditentukan di awal
2. Dapat digunakan untuk menilai konsistensi program
3. Testing dilakukan berdasarkan spesifikasi
4. Tidak perlu melihat kode program secara detail
5. Dapat memilih subset test secara efektif dan efisien.
6. Dapat menemukan cacat.
7. Memaksimalkan testing investmen.

Kekurangan *Blackbox Testing* :

1. Bila spesifikasi program yang dibuat kurang jelas dan ringkas, maka akan sulit membuat dokumentasi setepat mungkin
2. Tester tidak pernah yakin apakah PL tersebut benar – benar lulus uji.



## 2.8 Confusion Matrix

Salah satu cara untuk menghitung akurasi adalah dengan menggunakan metode *confusion matrix*. *Confusion matrix* merupakan sebuah cara yang berguna untuk menganalisis seberapa baik *classifier* mengenali data. Metode ini menggunakan tabel matriks seperti terlihat pada tabel 2.1 berikut ini, jika data set hanya terdiri dari dua kelas, kelas yang satu dianggap sebagai positif dan yang lainnya negatif.

**Tabel 2.1 Model *confusion matrix***

Klasifikasi yang benar	Diklasifikasikan sebagai	
	+	-
+	True positive	False positives
-	False negative	True Negatives

*True positive* adalah jumlah *record* positif yang diklasifikasikan sebagai positif, *false positives* adalah jumlah *record* positif yang diklasifikasikan sebagai negatif, *false negatives* adalah jumlah *record* negatif yang diklasifikasikan sebagai positif, *true negatives* adalah jumlah *record* negatif yang diklasifikasikan sebagai negatif, kemudian masukan data uji (Mustika, 2015). Setelah data uji dimasukan ke dalam confusion matrix, hitung nilai-nilai yang telah dimasukan tersebut untuk dihitung jumlah *sensitivity (recall)*, *specifity*, *precision* dan *accuracy*. *Sensitivity* digunakan untuk membandingkan jumlah TP terhadap jumlah record yang positif sedangkan *specifity* adalah perbandingan jumlah TN terhadap jumlah *record* yang negatif. Untuk menghitung digunakan persamaan dibawah ini :

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \quad (2.4)$$

$$\text{Specifity} = \frac{TN}{(FP+TN)} \quad (2.5)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (2.6)$$

$$\text{Accuracy} = \frac{TP+TN}{(TP+TN+FP+FN)} \quad (2.7)$$

Keterangan :

- TP = jumlah *true positives*  
 TN = jumlah *true negatives*  
 FP = jumlah *false positives*  
 FN = jumlah *false negatives*

## 2.9 Kategori nilai *Tweet*

Adapun beberapa tipe data yang menjadi kriteria pada penelitian ini adalah (Effendy, 2016) :

1. Positif, kalimat disebut positif, kalau mengandung arti positif terhadap, atau akan dimungkinkan menggiring para pembaca membentuk pendapat positif terhadap suatu obyek tertentu. Seperti informasi yang berguna, kabar yang menyenangkan, pujian, dukungan, prestasi, terobosan terbaru, kejadian yang penting, keberhasilan dan keindahan.
2. Negatif, Kalimat dikategorikan negatif, kalau mengandung arti negatif terhadap, atau akan dimungkinkan menggiring para pembaca membentuk pendapat negatif terhadap suatu obyek tertentu. Berisi tentang kekecewaan masyarakat terhadap layanan ojek online.

## 2.10 Kategori Penilaian Ojek Online

Adapun beberapa kategori penilaian ojek *online* untuk perkembangan kedepannya adalah sebagai berikut:

1. Teknologi Aplikasi

*GO-JEK* adalah perusahaan jasa pemesanan ojek yang menggunakan teknologi aplikasi. Aplikasi *GO-JEK* dapat diunduh melalui *smartphone* baik android dan Ios (Gojek-2015). Langkah awal sebelum menggunakan layanan *GO-JEK* adalah melakukan registrasi dengan memasukkan alamat *email*, nama, nomor telepon dan *password*. Setelah registrasi berhasil, pengguna dapat memilih layanan yang tersedia di dalam aplikasi *GO-JEK*. Langkah selanjutnya, setelah memilih layanan *GO-JEK* pengguna memasukkan alamat tempat asal dan tujuan. Kemudian aplikasi *GO-JEK* akan menampilkan informasi mengenai keberadaan pengemudi, prediksi, waktu, identitas pengemudi (nama, foto, nomor telepon), dan harga. Setelah proses penggunaan pelayanan selesai, aplikasi *GO-JEK* memberikan



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Manajemen Stephen P. Robbins dan Mary Coulter (2012) manajemen mengacu pada proses mengkoordinasi dan mengintegrasikan kegiatan-kegiatan kerja agar diselesaikan secara efisien dan efektif dengan dan melalui orang lain.

Griffin (2008) manajemen adalah suatu rangkaian aktivitas (termasuk perencanaan dan pengambilan keputusan, pengorganisasian, kepemimpinan, dan pengendalian) yang diarahkan pada sumber-sumber daya organisasi (manusia, finansial, fisik, dan informasi) dengan maksud untuk mencapai tujuan organisasi secara efektif dan efisien.

Dari pengertian diatas, peneliti menyimpulkan bahwa manajemen adalah proses pengkoordinasian sekelompok orang dengan arahan-arahan untuk mencapai tujuan perusahaan, secara efektif dan efisien. Perusahaan yang memiliki manajemen yang baik adalah perusahaan yang menjalankan fungsi efektif dan efisien. Efisien berarti menggunakan berbagai sumber daya secara bijaksana dan dengan cara yang hemat biaya, sehingga produk atau jasa yang dihasilkan berkualitas tinggi namun dengan biaya yang relatif rendah, sedangkan efektif berarti membuat keputusan yang tepat dan mengimplementasikannya dengan sukses.

## 2.11 Penelitian Terkait

Pada tabel 2.2 berikut ini dapat dilihat beberapa penelitian sebelumnya yang menjadi acuan dan dasar penelitian ini :

**Tabel 2.2 Penelitian Terkait**

Peneliti	Judul	Metode	Hasil
Didik Garbian, Yulison Herry Chrisnando, Agung	Analisis sentiment pada jasa ojek <i>online</i> menggunakan <i>Naïve Bayes</i>	Naïve Bayes	Sistem mampu mengklasifikasikan sentiment menggunakan <i>naive bayes</i> dengan akurasi yang dihasilkan 80 % berdasarkan 800 data <i>tweet</i> yang terdiri atas 300 data

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Wahana (2016)			latih dan 500 sebagai data uji.
Nabila Safina dan Aris Marjuni	Analisis sentiment pada Twitter terhadap jasa transportasi online di Indonesia dengan metode <i>Support Vector Machine</i>	<i>Support Vector Machine</i> (SVM)	Data berdasar dari Twitter dengan kata kunci mention dari akun @gojekindonesia dan @grabid berjumlah 1000 data , 900 data latih dan 100 data uji. Klasifikasi menggunakan metode SVM dengan Akurasi 87 %
Deden Rustiana (2017)	Analisis sentiment pasar otomotif mobil: <i>tweet</i> Twitter menggunakan <i>naïve bayes</i>	<i>Naïve Bayes</i>	Implementasi naïve bayes pada penelitian ini menghasilkan nilai akurasi 93% seta persisi sentiment positif 90%, persisi negatif 90%, dan persisi netral sebesar 100%
Ghulam Asrofi Buntoro (2016)	Analisis sentiment calon gubernur DKI Jakarta 2017 di Twitter	<i>Naïve Bayes Classifier</i> (SVM) dan <i>Support Vector Machine</i> (SVM)	Data yang digunakan sebanyak 300 tweet. Nilai akurasi tertinggi didapat saat menggunakan metode klasifikasi <i>naïve bayes classifier</i> (NBC) Akurasi Mencapai 95%.
Amir hamzah, (2012)	Klasifikasi Teks Dengan <i>Naive Bayes Classifier</i>	<i>Naive Bayes Classifier</i> (NBC)	Akurasi mencapai 91 %

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	(NBC) untuk Pengelompokan Teks Berita dan <i>Abstract</i> Akademis		
Sandi Fajar Rodiyansyah dan Edi Winarko, 2013	Klasifikasi Posting Twitter Kemactan Lalu Lintas Kota Bandung Menggunakan <i>Naive Bayes Classification</i>	<i>Naive Bayes Classifier</i> (NBC)	Akurasi mencapai 93,58%