

## BAB II

### LANDASAN TEORI

#### 2.1 Kepala Negara

Negara Indonesia merupakan negara berbentuk republik dan menganut sistem presidensiil, maka Presiden Indonesia adalah seorang kepala negara sekaligus kepala pemerintahan Indonesia. Hal ini ditegaskan dalam UUD 1945 Pasal 1 Ayat 1 “Negara Indonesia ialah Negara Kesatuan, yang berbentuk Republik” dan UUD 1945 Pasal 4 Ayat 1 “ Presiden Republik Indonesia memegang kekuasaan pemerintahan menurut Undang-Undang Dasar”.

Sebagai kepala negara, Presiden adalah simbol resmi negara Indonesia di dunia. Sebagai kepala pemerintahan, Presiden dibantu oleh wakil presiden dan menteri-menteri dalam kabinet, memegang kekuasaan eksekutif untuk melaksanakan tugas-tugas pemerintah sehari-hari.

##### 2.1.1 Tugas Presiden Sebagai Kepala Negara

Sebagai Kepala Negara, Presiden memiliki tugas-tugas khusus yang harus dilakukan selaku Kepala Negara. Tugas Presiden sebagai Kepala Negara tercantum dalam peraturan Undang-Undang Dasar 1945 (UUD '45) adalah:

1. Presiden memegang kekuasaan yang tertinggi atas Angkatan Darat, Angkatan Laut dan Angkatan Udara (Pasal 10).
2. Presiden mengangkat duta dan konsul (Pasal 13 ayat 1).
3. Presiden menerima penempatan duta Negara lain dengan memperhatikan pertimbangan Dewan Perwakilan Rakyat (Pasal 13 ayat 3).
4. Negara menjamin kemerdekaan tiap-tiap penduduk untuk memeluk agamanya masing-masing dan untuk beribadat menurut agamanya dan kepercayaannya itu (Pasal 29 Ayat 2).
5. Negara memprioritaskan anggaran pendidikan sekurang-kurangnya dua puluh persen dari anggaran pendapatan dan belanja negara serta dari

**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.

anggaran pendapatan dan belanja daerah untuk memenuhi kebutuhan penyelenggaraan pendidikan nasional (Pasal 31 Ayat 4).

6. Negara memajukan kebudayaan nasional Indonesia di tengah peradaban dunia dengan menjamin kebebasan masyarakat dalam memelihara dan mengembangkan nilai-nilai budayanya (Pasal 32 Ayat 1).
7. Negara menghormati dan memelihara bahasa daerah sebagai kekayaan budaya nasional (Pasal 32 Ayat 2).
8. Fakir miskin dan anak-anak terlantar dipelihara oleh negara (Pasal 34 Ayat 1).
9. Negara mengembangkan sistem jaminan sosial bagi seluruh rakyat dan memberdayakan masyarakat yang lemah dan tidak mampu sesuai dengan martabat kemanusiaan (Pasal 34 Ayat 2).
10. Negara bertanggung jawab atas penyediaan fasilitas pelayanan kesehatan dan fasilitas pelayanan umum yang layak (Pasal 34 Ayat 3).

### 2.1.2 Tugas Presiden Sebagai Kepala Pemerintahan

Tugas Presiden sebagai Kepala Pemerintahan berdasarkan UUD 1945 adalah:

1. Presiden menetapkan peraturan pemerintah untuk menjalankan undang-undang sebagaimana mestinya (Pasal 5 ayat 2).
2. Menteri-menteri itu diangkat dan diberhentikan oleh Presiden (Pasal 17 ayat 2).
3. Hubungan wewenang antara pemerintah pusat dan Pemerintahan Daerah, Provinsi, Kabupaten, dan Kota, diatur dengan undang-undang dengan memperhatikan kekhususan dan keragaman daerah (Pasal 18B Ayat 1).
4. Hubungan keuangan, pelayanan umum, pemanfaatan sumber daya alam dan sumber daya lainnya antara pemerintah pusat dan pemerintah daerah diatur dan dilaksanakan secara adil dan selaras berdasarkan undang-undang (Pasal 18B Ayat 2).

**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.

5. Presiden mengesahkan rancangan undang-undang yang telah disetujui bersama untuk menjadi undang-undang (Pasal 20 Ayat 4).
6. Rancangan undang-undang anggaran pendapatan dan belanja negara diajukan oleh Presiden untuk dibahas bersama Dewan Perwakilan Rakyat dengan memperhatikan pertimbangan Dewan Perwakilan Daerah (Pasal 23 Ayat 2).
7. Anggota Badan Pemeriksa Keuangan dipilih oleh Dewan Perwakilan Rakyat dengan memperhatikan pertimbangan Dewan Perwakilan Daerah dan diresmikan oleh Presiden (Pasal 23F Ayat 1).
8. Calon Hakim Agung diusulkan Komisi Yudisial kepada Dewan Perwakilan Rakyat untuk mendapatkan persetujuan dan selanjutnya ditetapkan sebagai hakim agung oleh Presiden (Pasal 24A Ayat 3).
9. Anggota Yudisial diangkat dan diberhentikan oleh Presiden dengan persetujuan Dewan Perwakilan Rakyat (Pasal 24B Ayat 3).
10. Mahkamah Konstitusi mempunyai sembilan orang anggota hakim konstitusi yang ditetapkan oleh Presiden, yang diajukan masing-masing tiga orang oleh Mahkamah Agung, tiga orang oleh Dewan Perwakilan Rakyat, dan tiga orang oleh Presiden (Pasal 24C Ayat 3).
11. Perlindungan, pemajuan, penegakan, dan pemenuhan hak asasi manusia adalah tanggung jawab negara, terutama pemerintah (Pasal 28I Ayat 4).
12. Setiap warga negara wajib mengikuti pendidikan dasar dan pemerintah wajib membiayainya (Pasal 31 Ayat 2).
13. Pemerintah mengusahakan dan menyelenggarakan satu sistem pendidikan nasional, yang meningkatkan keimanan dan ketakwaan serta ahlak mulia dalam rangka mencerdaskan kehidupan bangsa, yang diatur dengan undang-undang (Pasal 31 Ayat 3).
14. Pemerintah memajukan ilmu pengetahuan dan teknologi dengan menunjang tinggi nilai-nilai agama dan persatuan bangsa untuk kemajuan peradaban serta kesejahteraan umat manusia (Pasal 31 Ayat 5).

**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.

### 2.1.3 Wewenang Presiden

Wewenang Presiden sebagaimana tercantum dalam UUD 1945 adalah:

1. Presiden berhak mengajukan rancangan undang-undang kepada Dewan Perwakilan Rakyat (Pasal 5 Ayat 1).
2. Presiden dengan persetujuan Dewan Perwakilan Rakyat menyatakan perang, membuat perdamaian dan perjanjian dengan negara lain (Pasal 11 Ayat 1).
3. Presiden dalam membuat perjanjian internasional lainnya yang menimbulkan akibat yang luas dan mendasar bagi kehidupan rakyat yang terkait dengan beban keuangan negara, dan/atau mengharuskan perubahan atau pembentukan undang-undang harus dengan persetujuan Dewan Perwakilan Rakyat (Pasal 11 Ayat 2).
4. Presiden menyatakan keadaan bahaya. Syarat-syarat dan akibatnya keadaan bahaya ditetapkan dengan undang-undang (Pasal 12).
5. Presiden memberi grasi dan rehabilitasi dengan memperhatikan pertimbangan Mahkamah Agung (Pasal 14 Ayat 1).
6. Presiden memberi amnesti dan abolisi dengan memperhatikan pertimbangan Dewan Perwakilan Rakyat (Pasal 14 Ayat 2).
7. Presiden memberi gelar, tanda jasa, dan lain-lain tanda kehormatan yang diatur dengan undang-undang (Pasal 15).
8. Presiden membentuk suatu dewan pertimbangan yang bertugas memberikan nasihat dan pertimbangan kepada Presiden, yang selanjutnya diatur dalam undang-undang (Pasal 16).
9. Dalam hal ihwal kegentingan yang memaksa, Presiden berhak menetapkan peraturan pemerintah sebagai pengganti undang-undang (Pasal 22 Ayat 1).
10. Cabang-cabang produksi yang penting bagi negara dan yang menguasai hajat hidup orang banyak dikuasai oleh negara (Pasal 33 Ayat 2).
11. Bumi dan air dan kekayaan alam yang terkandung di dalamnya dikuasai oleh negara dan dipergunakan untuk sebesar-besar kemakmuran rakyat (Pasal 33 Ayat 3).

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.

## 2.2 Twitter

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*).

Semua pengguna dapat mengirim dan menerima *Tweets* melalui situs Twitter, aplikasi eksternal yang kompatibel (telepon seluler), atau dengan pesan singkat (SMS) yang tersedia di negara-negara tertentu (Twitter, 2013). Pengguna dapat menulis pesan berdasarkan topik dengan menggunakan tanda # (*hashtag*). Sedangkan untuk menyebutkan atau membalas pesan dari pengguna lain bisa menggunakan tanda @.

Pesan pada awalnya diatur hanya mempunyai batasan sampai 140 karakter disesuaikan dengan kompatibilitas dengan pesan SMS. Batas karakter 140 juga meningkatkan penggunaan layanan memperpendek URL seperti bit.ly, goo.gl, dan tr.im, dan jasa hosting konten, seperti Twitpic, Tweepphoto, memozu.com dan NotePub untuk mengakomodasi multimedia isi dan teks yang lebih panjang dari pada 140 karakter (Twitter, 2013). Twitter menggunakan bit.ly untuk memperpendek otomatis semua *URL* yang dikirim-tampil. Fitur yang terdapat dalam Twitter, antara lain:

### 1. Laman Utama (*Home*)

Pada halaman utama kita bisa melihat *Tweets* yang dikirimkan oleh orang-orang yang menjadi teman kita atau yang kita ikuti (*following*).

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.

## 2. Profil (*Profile*)

Pada halaman ini yang akan dilihat oleh seluruh orang mengenai profil atau data diri serta *Tweets* yang sudah pernah kita buat.

## 3. *Followers*

Pengikut adalah pengguna lain yang ingin menjadikan kita sebagai teman. Bila pengguna lain menjadi pengikut akun seseorang, maka *Tweets* seseorang yang ia ikuti tersebut akan masuk ke dalam halaman utama.

## 4. *Following*

Kebalikan dari pengikut, *following* adalah akun seseorang yang mengikuti akun pengguna lain agar *Tweets* yang dikirim oleh orang yang diikuti tersebut masuk ke dalam halaman utama.

## 5. *Mentions*

Biasanya konten ini merupakan balasan dari percakapan agar sesama pengguna bisa langsung menandai orang yang akan diajak bicara.

## 6. *Favorite*

*Tweets* ditandai sebagai favorit agar tidak hilang oleh halaman sebelumnya.

## 7. Pesan Langsung (*Direct Message*)

Fungsi pesan langsung lebih bisa disebut SMS karena pengiriman pesan langsung di antara pengguna.

## 8. *Hashtag*

*Hashtag* “#” yang ditulis di depan topik tertentu agar pengguna lain bisa mencari topik yang sejenis yang ditulis oleh orang lain juga

## 9. *List*

Pengguna Twitter dapat mengelompokkan ikutan mereka ke dalam satu grup sehingga memudahkan untuk dapat melihat secara keseluruhan para nama pengguna (*username*) yang mereka ikuti (*follow*).

## 10. Topik Terkini (*Trending Topic*)

Topik yang sedang banyak dibicarakan banyak pengguna dalam suatu waktu yang bersamaan.

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.

## 2.2.1 Twitter API

*Application Programming Interface* (API) merupakan fungsi-fungsi untuk menggantikan bahasa yang digunakan dalam *system calls* dengan bahasa yang lebih terstruktur dan mudah dimengerti oleh programmer. Fungsi yang dibuat dengan menggunakan API tersebut kemudian akan memanggil *system calls* sesuai dengan sistem operasinya. Tidak tertutup kemungkinan nama dari *system calls* sama dengan nama di API (Twitter, 2013).

Pada awalnya perusahaan Summize yang menyediakan fasilitas mencari data di twitter. Kemudian perusahaan Summize ini diakuisisi dan diganti merek menjadi *Twitter Search* sehingga *Search API* terpisah sebagai entitas sendiri. API Twitter terdiri dari 3 (tiga) bagian yaitu (Twitter, 2013) :

### a. *Search API*

*Search API* dirancang untuk memudahkan user dalam mengelola *query search* di konten Twitter. User dapat menggunakannya untuk mencari *tweet* berdasarkan *keyword* khusus atau mencari *tweet* lebih spesifik berdasarkan username Twitter. *Search API* juga menyediakan akses pada data *Trending Topic*.

### b. *Representational State Transfer (REST) API*

*REST API* memperbolehkan *developer* untuk mengakses inti dari Twitter seperti *timeline*, *status update* dan informasi user. *REST API* digunakan dalam membangun sebuah aplikasi Twitter yang kompleks yang memerlukan inti dari Twitter.

### c. *Streaming API*

*Streaming API* digunakan *developer* untuk kebutuhan yang lebih intensif seperti melakukan penelitian dan analisis data. *Streaming API* dapat menghasilkan aplikasi yang dapat mengetahui statistik *status update*, *follower* dan lain sebagainya.

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.

## 2.3 Analisis Sentimen

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

Menurut Pang dkk (2008) yang dikutip Rosdiansyah (2014), yang mengatakan bahwa pada suatu kalimat analisis sentimen menggambarkan bagian pertimbangan penilaian terhadap entitas atau kejadian tertentu. Yang termasuk entitas adalah produk, layanan, topik, isu, orang, organisasi atau peristiwa yang menjadi objek target pada kalimat sentimen menurut Liu (2012) dalam Rosdiansyah (2014).

Tugas dasar dalam analisis sentimen adalah mengelompokkan teks yang ada dalam sebuah kalimat atau dokumen kemudian menentukan pendapat yang dikemukakan dalam kalimat atau dokumen tersebut apakah bersifat positif, negatif atau netral (Dehaff, M., 2010) yang dikutip (Saraswati, 2011). *Sentiment analysis* juga dapat menyatakan perasaan emosional sedih, gembira, atau marah.

## 2.4 Text Mining

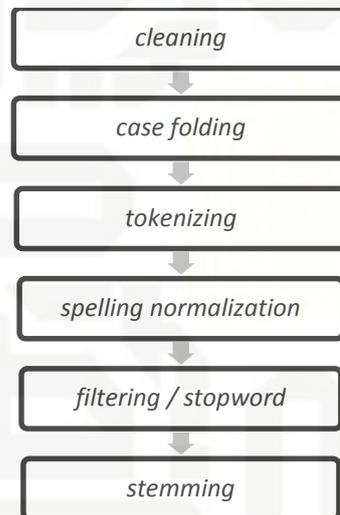
*Text mining* (penambangan teks) adalah penambangan yang dilakukan oleh komputer untuk mendapatkan sesuatu yang baru, sesuatu yang tidak diketahui sebelumnya atau menemukan kembali informasi yang tersirat secara implisit, yang berasal dari informasi yang diekstrak secara otomatis dari sumber-sumber data teks yang berbeda-beda (Feldman & Sanger, 2007). *Text mining* merupakan teknik yang digunakan untuk menangani masalah klasifikasi, *clustering*, *information extraction* dan *information retrieval* (Berry & Kogan, 2010).

Pada dasarnya proses kerja dari *text mining* banyak mengadopsi dari penelitian *data mining* namun yang menjadi perbedaan adalah pola yang digunakan oleh *text mining* diambil dari sekumpulan bahasa alami yang tidak terstruktur sedangkan dalam *data mining* pola yang diambil dari *database* yang terstruktur (Han & Kamber, 2006). Tahap-tahap *text mining* secara umum adalah *text*

*preprocessing* dan *feature selection* (Feldman & Sanger, 2007), (Berry & Kogan, 2010). Dimana penjelasan dari tahap-tahap tersebut adalah sebagai berikut:

### 2.4.1 Text Preprocessing

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 2.1



**Gambar 2.1** Proses text preprocessing

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

1. *Cleaning* adalah proses untuk membersihkan *tweet* dari kata-kata yang tidak diperlukan untuk mengurangi *noise* pada proses klasifikasi. Kata yang dihilangkan adalah karakter HTML, hashtag (#), *username* atau *mention* (@username), *link url* (<http://situs.com>), emoticon, dan RT (tanda *retweet*).
2. *Case folding* adalah 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.

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. *Tokenizing* yaitu proses memecah tweet atau kalimat menjadi sebuah kata dengan melakukan analisa terhadap kumpulan kata dengan memisahkan kata tersebut dan menentukan struktur sintaksis dari tiap kata tersebut.

4. *Spelling normalization*

Normalisasi merupakan tahapan identifikasi kata silang dan penulisan kata berlebihan kemudian diganti dengan kata kamus KBBI (Rosdiansyah, 2014). Pada tahap ini setiap dijumpai kata yang penggunaan huruf berlebihan dan kata yang tidak baku akan diubah. Adapun algoritma normalisasi yang dilakukan pada penelitian ini sebagai berikut :

1. Cari kata yang akan dinormalisasi dalam kamus. Jika ditemukan maka di asumsikan bahwa kata tersebut adalah *root word*, maka algoritma berhenti.
2. Jika tidak ditemukan hapus huruf berlebihan dimulai untuk setiap huruf pada kata, periksa huruf pertama kata tersebut, kemudian *recoding*. Periksa huruf selanjutnya jika huruf sama dengan huruf sebelumnya maka hapus huruf tersebut, jika tidak simpan huruf lakukan hal yang sama pada huruf selanjutnya.
3. Melakukan *recoding*
4. Jika telah diperiksa untuk setiap huruf periksa kata hasil proses sebelumnya pada kamus.
5. Jika ditemukan maka algoritma berhenti. Jika tidak ditemukan algoritma ini mengembalikan kata yang asli sebelum dilakukan penghapusan huruf berlebihan.
6. Selanjutnya periksa kata pada kamus silang.
7. Jika ditemukan lakukan perubahan kata menjadi kata baku. Jika tidak ditemukan maka kembalikan kata pada *root word*.

6. *Filtering* adalah proses untuk 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). Selain itu juga melakukan pembuangan kata yang tidak berpengaruh dalam klasifikasi sentimen suatu *tweet* yang disebut *stopword*. Contoh *stopword* bahasa

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.

Indonesia yaitu kata gaul seperti “cc”, “gw”, atau “ente”, kata hubung seperti “di”, “dengan”, “dari”, “yang”, “ke”, “dan”, penunjuk waktu, dan kata Tanya.

## 6. *Stemming*

*Stemming* adalah tahapan mencari kata *root* / kata dasar dari setiap kata hasil dari proses normalisasi. 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 ECS (Solihin, Andy, & DKK, 2013). Algoritma ini mengacu pada aturan KBBI yang mengelompokkan imbuhan yang diperbolehkan atau imbuhan yang tidak diperbolehkan.

Berikut merupakan langkah-langkah yang dilakukan oleh algoritma ECS :

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”).
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.

### 2.4.3 Penggalan Informasi Pada Text Mining

Tahap akhir penggalan informasi pada *text mining* yaitu ekstraksi ilmu pengetahuan (*knowledge discovery*), dimana terdapat beberapa jenis kategori utama yang bisa dilakukan sebagai berikut (Miner, dkk, 2012 dikutip oleh Chyntia, 2015).

#### 1. Klasifikasi/ prediksi,

Klasifikasi adalah bentuk analisis data yang mengekstrak model untuk menggambarkan kelas data (Jiawei, Kamber, & Pei, 2012 dikutip oleh Chyntia, 2015). Model yang dibangun meliputi pengklasifikasian dan prediksi kategori label kelas. Klasifikasi data mempunyai dua tahapan proses, yaitu tahap pembelajaran (*learning step*) dimana model klasifikasi dibangun berdasarkan label yang sudah diketahui sebelumnya dan tahapan klasifikasi (*classification step*) dimana model digunakan untuk memprediksi label kelas dari data yang diberikan (Miner, dkk, 2012 dikutip oleh Chyntia, 2015). Klasifikasi memiliki berbagai aplikasi, termasuk deteksi penipuan, penargetan marketing, prediksi kinerja, manufaktur, diagnosis medis, dan banyak lainnya. Sebagai contoh, kita dapat membangun sebuah model klasifikasi untuk mengkategorikan apakah suatu aplikasi pinjaman bank termasuk aman atau berisiko. Karena pada awal pembangunan model label kelas dari data telah diketahui, klasifikasi juga disebut sebagai metode *supervised learning*.

#### 2. Pengelompokan (*Clustering*)

Tidak seperti klasifikasi, pada model *clustering* pengelompokan data dilakukan dengan menggunakan algoritma yang sudah ditentukan dan selanjutnya data akan

**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.

diproses oleh algoritma untuk dikelompokkan menurut karakteristik alaminya. Tidak ada unsur pembimbingan (dengan pemberian label kelas), melainkan algoritma akan berjalan dengan sendirinya untuk mengelompokkan data tersebut. Data yang lebih dekat (mirip) dengan data lain akan berkelompok dalam satu *cluster*, sedangkan data yang lebih jauh (berbeda) dari data yang lain akan berpisah dalam kelompok yang berbeda. Untuk masalah pengelompokkan data berdasarkan kemiripan / ketidakmiripan antar data tanpa ada label kelas yang diketahui sebelumnya disebut dengan pembelajaran tidak terbimbing atau *unsupervised learning*. Dalam konteks yang lain, pembelajaran tidak terbimbing disebut juga pengelompokan atau *clustering*. Menurut struktur, *clustering* terbagi menjadi dua, yaitu *hirarki* dan *partisi*. Dalam pengelompokan berbasis hierarki, satu data tunggal bisa dianggap sebuah *cluster*, dua atau lebih *cluster* kecil dapat bergabung menjadi sebuah *cluster* besar, begitu seterusnya hingga semua data dapat bergabung menjadi sebuah *cluster*. Di sisi lain, pengelompokan berbasis partisi membagi set data ke dalam sejumlah *cluster* yang tidak bertumpang-tindih antara satu *cluster* dengan *cluster* yang lain, artinya setiap data hanya menjadi anggota satu *cluster* saja.

### 3. Asosiasi

Asosiasi merupakan proses pencarian hubungan antar elemen data. Dalam dunia industri retail, analisis asosiasi biasanya disebut *market Basket Analysis* (Miner, dkk, 2012 dikutip oleh Chyntia, 2015). Asosiasi tersebut dihitung berdasarkan ukuran *Support* (presentase dokumen yang memuat seluruh konsep suatu produk A dan B) dan *confidence* (presentase dokumen yang memuat seluruh konsep produk B yang berada dalam subset yang sama dengan dokumen yang memuat seluruh konsep produk A).

### 4. Analisis Tren

Tujuan dari analisis tren yaitu untuk mencari perubahan suatu objek atau kejadian oleh waktu (Miner, dkk, 2012 dikutip oleh Chyntia, 2015). Salah satu aplikasi analisis tren yaitu kegiatan identifikasi evolusi topik penelitian pada artikel akademis.



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.

$IVI$  = Total kata unik pada keseluruhan kelas  $c$

Selanjutnya,

$$P(c | dn) = P(c) * \prod p(w|c) \tag{2.5}$$

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

$Pc$  = Priors

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

## 2.7 Blackbox

*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.

*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?

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. 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 *K-fold Cross Validation*

*K-fold cross validation* adalah teknik yang dapat digunakan jika memiliki jumlah data yang terbatas. Cara kerja *K-Fold Cross Validation* adalah sebagai berikut ;

1. Seluruh data dibagi menjadi K bagian.
2. *Fold* ke -1 adalah ketika bagian ke-1 menjadi data uji (*testing data*) dan sisanya menjadi data latih (*training data*). Selanjutnya, hitung akurasi berdasarkan porsi data tersebut.
3. *Fold* ke-2 adalah ketika bagian ke-2 menjadi data uji (*testing data*) dan sisanya menjadi data latih (*training data*). Selanjutnya, hitung akurasi berdasarkan porsi data tersebut.

4. Demikian seterusnya hingga mencapai *fold* ke-K.

5. Hitung rata-rata akurasi dari N buah akurasi di atas. Rata-rata akurasi ini menjadi akurasi final.

Metode *k-fold cross validation* melakukan generalisasi dengan membagi data kedalam **k** bagian berukuran sama. Selama proses berlangsung, salah satu dari partisi dipilih untuk data uji, dan sisanya digunakan untuk data latih. Langkah ini di ulangi **k** kali sehingga setiap partisi digunakan untuk data uji tepat satu kali. Metode *k-fold cross validation* menetapkan **k = N**, ukuran dari data set. Pendekatan ini memiliki keuntungan dalam penggunaan data sebanyak mungkin untuk training. *Test set* secara efektif mencakup keseluruhan data set. Kekurangan data pendekatan ini adalah banyaknya komputasi untuk mengulangi prosedur sebanyak **N** kali. *K-fold cross validation* adalah salah satu teknik untuk mengevaluasi keakuratan model (Mustika, 2015 dikutip oleh Citra, 2015).

## 2.9 Confusion Matrix

Metode ini menggunakan tabel matriks seperti yang terlihat pada 2.3 berikut ini jika data set hanya terdiri dari dua kelas, kelas yang satu dianggap sebagai positif dan yang lainnya negatif.

**Tabel 2.1 Confusion matrix**

Klasifikasi yang benar	Diklasifikasikan sebagai	
	+	-
+	True positives	False positives
-	False negatives	True negatives

*True positives* 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 masukkan data uji. Setelah data uji dimasukkan ke dalam *confusion matrix*, hitung nilai-nilai yang telah dimasukkan tersebut untuk dihitung jumlah *sensitivity (recall)*, *specificity*, *precision* dan *accuracy*. *Sensitivity* digunakan

untuk membandingkan jumlah TP terhadap jumlah record yang positif sedangkan *specificity* adalah perbandingan jumlah TN terhadap jumlah *record* yang negatif. Untuk menghitung digunakan persamaan di bawah ini:

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \quad (2.6)$$

$$\text{Specificity} = \frac{TN}{(FP+TN)} \quad (2.7)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (2.8)$$

$$\text{Accuracy} = \frac{TP+TN}{(P+N)} \quad (2.9)$$

Keterangan:

TP = jumlah *true positives*

TN = jumlah *true negatives*

FP = jumlah *false positives*

FN = jumlah *false negatives*

## 2.10 Penelitian Terkait

Pada table berikut dapat dilihat beberapa penelitian sebelumnya mengenai analisis sentimen :

**Table 2.2 Penelitian Terkait**

Peneliti	Judul	Metode	Ekstraksi fitur	Domain	Akurasi
Nur dan Santika (2011)	Analisis Sentimen Pada Dokumen Bebahasa Indonesia Dengan Pendekatan <i>Support Vector Machine</i>	SVM	TF-IDF	Twitter	73,07%
Hidayatullah dan Azhari (2014)	Analisis Sentimen dan Klasifikasi kategori terhadap tokoh public pada twitter	NBC	Unigram	Twitter	79,68%

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.

Peneliti	Judul	Metode	Ekstraksi fitur	Domain	Akurasi
Amelia Mustika (2015)	Penerapan Metode Support Vector Machine Dalam Klasifikasi Sentimen Tweet Public Figure	SVM	BI-Gram	Twitter	72%
Nugroho & dkk (2016)	Analisis Sentimen Pada Jasa Ojek Online Menggunakan Metode Naïve Bayes Classifier	NBC		Twitter pada akun @GojekIndonesia dan @GrabID	80%
Antonis Rahmad dan Yuan Lukito (2016)	Klasifikasi Sentimen Komentar Dri Fecebook Page menggnakan Naïve Bayes Classifier	NBC	TF-IDF	Facebook Page	83%