

BAB I. LANDASAN TEORI

1.1 Penelitian Terdahulu

Beberapa penelitian yang dijadikan rujukan dalam penelitian ini adalah sebagai berikut:

Pada penelitian yang dilakukan oleh Sonia Anastasia and Indra Budi yang berjudul *Twitter Sentiment Analysis of Online Transportation Service Providers* menunjukkan bahwa GO-JEK dan Grab adalah dua online paling populer penyedia layanan transportasi di Indonesia. Sebuah kompetisi antara keduanya ketat untuk mendapatkan pelanggan baru dan menyediakan layanan terbaik. Tidak ada survei resmi yang dilakukan untuk mengukur kepuasan pelanggan di kedua perusahaan untuk memberikan wawasan perusahaan mana yang menawarkan kualitas layanan yang lebih baik. Penelitian ini bertujuan untuk mengukur GO-JEK dan Grab kepuasan pelanggan melalui analisis sentimen data twitter. kedua perusahaan menggunakan twitter untuk menjangkau pelanggan mereka dan mempromosikan layanan mereka. Kami mengumpulkan 126.405 tweet dari Februari hingga Maret 2016 mengandung kata kunci GOJEK dan Grab. Kemudian, kami memproses pra-tweet dan beri label secara manual sebelum diklasifikasikan menggunakan tiga algoritma: Mendukung Mesin Vektor, Naïve Bayes, dan Decision Tree. Kita membuat dan membandingkan beberapa skema klasifikasi dengan berbeda kata kunci, dataset, dan teknik validasi silang k-fold. Akhirnya, kami menghitung Net Sentiment Score yang berkorelasi dengan pelanggan kepuasan menggunakan hasil klasifikasi. Percobaan menunjukkan kepuasan pelanggan grab lebih tinggi dari Gojek. Itu Studi juga menunjukkan bahwa pelanggan cenderung menyebut kedua perusahaan Akun Twitter untuk pengalaman buruk dan tidak menyebutkan akun perusahaan untuk komentar positif (Anashtasia, 2016).

Penelitian yang dilakukan oleh Dwi Lidiana Wati yang Berjudul Analisis Sentimen Pada Twitter Untuk Layanan Ojek Online Menggunakan Metode Multinomial Naïve Bayes Dengan Pembobotan Tf Idf mendapatkan hasil yang lebih optimal, maka akan ditambahkan pembobotan TF (Term Frequency) dan IDF (Invers Document Frequency). Dataset berjumlah 900 tweet yang didapatkan dari mention para pengguna ojek online kepada akun official masing-masing ojek

online. Akun official yang digunakan adalah @gojekindonesia, GrabID, dan Uber_IDN. Berdasarkan hasil penelitian, algoritma Multinomial Naïve Bayes dan pembobotan TF IDF didapatkan hasil rata-rata akurasi sebesar 92,467% pada stopword dan 92,667% pada laplace. Sedangkan pada pengujian Naive Bayes Classifier mendapatkan hasil rata-rata sebesar 92,222% pada stopword dan 92,667% pada laplace.

Pada penilitan terdahulu dengan judul *Temporal Sentiment Analysis for Opinion Mining of ASEAN Free Trade Area on Social Media* oleh Mustofa Kamal menerapkan pendekatan baru untuk menganalisis pendapat tentang AFTA oleh penambang pendapat berdasarkan analisis sentimen temporal. Pendekatan ini akan mengekstraksi dan memproses data tekstual secara otomatis untuk mendapatkan sentiment. Informasi yang terkandung dalam kalimat opini. Hasil penelitian menunjukkan efektivitas pendekatan sederhana untuk mendapatkan informasi sentimen dari pendapat orang dan beberapa pemangku kepentingan dan siswa yang memiliki kepedulian terhadap AFTA menilai solusi ini.

Pada penelitian terdahulu, judul *Twitter Rule-based Sentiment Degree Measurement of Opinion Mining of Community Participatory in the Government of Surabaya* oleh Berlian Juliartha. Karena kompleksitas tugas untuk mengatasi tingkat partisipatif dan untuk mengkategorikan kelompok partisipatif, mereka menghadapi kesulitan untuk memenuhi target. Dalam penelitian ini, kami menyajikan sistem baru untuk mengukur tingkat sentimen masyarakat partisipatif. Kami menyediakan 5 fungsi dalam sistem kami, yaitu: (1) Pengumpulan Data, (2) Pengolah Data, (3) Penambangan Teks, (4) Analisis Sentimen dan (5) Validasi. Kami mengusulkan berdasarkan aturan kami teknik untuk analisis sentimen penambang pendapat dengan deteksi 8 bagian penting, yaitu (1) Verb, (2) Adjective, (3) Preposisi, (4) Kata benda, (5) Adverbia, (6) Simbol, (7) Frasa, dan (8) Complimentary. Untuk penerapan sistem yang kami usulkan Untuk penerapan sistem yang kami usulkan, kami membuat serangkaian percobaan dengan 408 data partisipasi masyarakat di Twitter untuk Diskominfo Surabaya. Setelah dilakukan pengujian sistem menggunakan validasi LOO, bahwasannya metode berbasis rule bases memberikan akurasi 81,33%, metode SVM menggunakan fungsi SMO memberi 69,36% dan metode Naive Bayes (Tan Li Im, 2015).

Pada penelitian dengan judul *Sentiment Analysis Menggunakan Rule Based Method Pada Data Pengaduan Publik Berbasis Lexical Resources* oleh Masfulatul Lailiyah bahwa Sentiwordnet banyak digunakan dalam opinion mining maupun sentiment analysis teks dalam bahasa Inggris. Pada penelitian ini, kami mencoba memanfaatkan Sentiwordnet untuk menganalisa sentiment dari pengaduan masyarakat berbahasa Indonesia serta membandingkannya dengan sentimen leksikon Indonesia. Diperoleh nilai akurasi sebesar 47% untuk data pengaduan pada media twitter dan 56.85% untuk data pengaduan pada media center ketika menggunakan Sentiwordnet. Sedangkan pada penggunaan sentimen leksikon Indonesia diperoleh nilai akurasi sebesar 65.4% untuk data pengaduan pada media twitter dan 81.4% untuk data pengaduan pada media center (Zhou Jin, 2016).

Pada penelitian yang dilakukan oleh Sujata Rani yang berjudul *Rule based Sentiment Analysis System for Analyzing Tweets* sistem analisis sentimen telah diusulkan yang menghitung sentimen dan skor setiap tweet yang diekstraksi dari twitter tentang parameter apa pun yang diberikan oleh pengguna. Itu sistem yang diusulkan telah mencapai akurasi 86% dalam hal menguji pada 500 tweet dan 94% sementara menguji pada 200 bahasa Inggris kalimat. Hasil yang diberikan oleh sistem sangat merangsang dan di masa depan, memiliki potensi untuk digunakan untuk manfaat masyarakat (Rani, 2017).

1.2 Ojek Online

Kesibukan setiap orang membuat mereka ingin mendapatkan sesuatu yang instan, apalagi menyangkut kegiatan sehari-hari. Kegiatan seperti berpergian, pemesanan sesuatu, sampai melakukan bersih – bersih ingin mereka lakukan secara instan. Hal ini memunculkan alasan untuk memunculkan suatu teknologi yang baru yakni berupa ojek online. Ojek online merupakan hasil perkembangan dari teknologi yang memanfaatkan internet dan smartphone untuk menjalankan suatu kegiatan sehari-hari. Ojek online di masa kini tidak hanya soal mengantarkan orang, namun dengan perkembangan teknologi pula menjadikan ojek online mempunyai banyak layanan, seperti pemesanan makanan, pengiriman barang, pembelian tiket, pembelian pulsa, melakukan cleaning service hingga melakukan belanja. System ojek online kini bergantung pada internet dan smarphone. Aplikasi sudah disediakan

oleh penyedia layanan ojek online, dan siap di gunakan oleh para pelanggan. Setiap pelanggan yang mempunyai aplikasi ojek online tinggal duduk manis sambil menanti pesanan mereka. Ojek online kini hadir untuk memudahkan para pelanggan dalam melakukan kegiatan sehari-hari mereka dengan memanfaatkan teknologi yang ada. Penyedia layanan ojek online pun berkembang pesat seperti Gojek dan Grab. Tidak menutup kemungkinan banyak penyedia layanan ojek online yang berada di kota – kota kecil yang juga berkembang.

Namun seiring persaingan yang semakin ketat, kedua penyedia layanan ojek online yakni Gojek dan Grab mampu berada di titik teratas dalam hal penyedia layanan ojek online. Gojek dan Grab terus berusaha menyediakan layanan terbaik mereka untuk para pelanggan. Dalam menyediakan layanan untuk pelanggan , penyedia layanan ojek online juga memanfaatkan media social untuk wadah pelanggan untuk menyalurkan kritik saran nya terhadap penyedia layanan tersebut, sebagai contoh pada akun twitter resmi gojek yakni @gojekindonesia , tertulis “@gojekindonesia min mau nanya nih, kenapa akun saya tiba-tiba logout sendiri. Ini udah ke 2 kalinya akun saya keluar sendiri dan saya gk bisa pakek akun saya yg dulu. Saya jadi kecewa sebagai pengguna setia gojek”. Contoh pada akun twitter resmi Grab yakni @GrabId , “@GrabId tolong DMnya dilihat lagi ya, sudah delivered tapi makanan gak diantar driver ,CSnya salah baca...” . Dari contoh data yang ada tersebut, pihak penyedia layana ojek online pasti akan merespon kritik dan saran dari para pelanggan yang di salurkan melalui media social. Kedua penyedia pelayanan ojek online tersebut juga memanfaatkan akun social media mereka untuk mengukur kepuasan pelanggan dan wawasan pelanggan dengan menganalisis data media sosial (L. Zhang et al & Desouza, 2014).

1.3 *Analysis Sentiment*

Sentiment analysis atau dikenal juga dengan sebutan opinion mining merupakan cabang dari klasifikasi teks. Tujuan dari sentiment analysis adalah mengelompokkan teks (dokumen) yang mengandung opini sebagai positive sentiment, negative sentiment, atau netral (Liu, 2010). Dalam sentiment analysis fitur yang dihasilkan dan algoritma klasifikasi menjadi poin utama (Atmadja, 2015). Tugas dasar dalam analisis sentimen adalah mengelompokkan teks yang ada

dalam sebuah kalimat atau dokumen kemudia menentukan pendapat yang dikemukakan dalam kalimat atau dokumen tersebut apakah bersifat positif, negatif atau netral (Dehaff, M., 2010). Analisis sentimen bertujuan menganalisis pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenaan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu (E.Boy, 2007).

Analisis sentimen terbagi menjadi metode klasifikasi dan rule-based. Metode berbasis klasifikasi menggunakan metode machine learning. Machine Learning memerlukan dataset untuk digunakan sebagai data training. Pendekatan machine learning umumnya menggunakan supervised classifier. Machine learning biasanya dimulai dengan membentuk dataset pelatihan yang kemudian digunakan untuk melatih classifier. Akhirnya, classifier diuji pada dataset uji. Oleh karena itu, dibutuhkan effort untuk mengumpulkan dan melakukan class tag pada sampel dataset tersebut, selain itu proses training juga membutuhkan waktu. Akurasi dari pendekatan klasifikasi machine learning sangat baik, akan tetapi performa klasifikasinya domain dependent terhadap dataset yang digunakan pada saat training. Sedangkan metode rule-based menggunakan lexicon and rule yang dibuat secara manual oleh manu-sia. Lexicon-Based menggunakan dictionary atau kamus lexicon untuk melakukan penilaian terhadap kata. Pada dictionary, kata-kata dipasangkan dengan nilai polaritasnya. Yang harus dilakukan sebelum melakukan analisis menggunakan lexicon adalah menentukan kata yang akan dianalisis dari corpus. Pemilihan kata tersebut dapat dilakukan dengan melakukan Part-Of-Speech Tagging lalu mencari kata-kata dengan tipe yang ingin diambil. Dalam pendekatan leksikon, polaritas ditentukan dalam proses pencocokan teks ulasan dengan kelas positif dan negatif dari leksikon sentimen. Kinerja analisis sentimen dalam pendekatan leksikon dapat ditingkatkan dengan pengklasifikasi berbasis aturan, yang secara khusus diamati untuk domain yang agak sempit. Dengan demikian, algoritma berbasis aturan untuk analisis sentimen digunakan untuk mengklasifikasikan komentar pengaduan layanan. Rasio positif / negatif dari setiap kalimat ditentukan dan nilai-nilai sentimen dari seluruh teks dihitung (Benamara, 2006).

1.4 *Rule Based*

Pendekatan berbasis rule based adalah pendekatan yang mengklasifikasikan teks ke dalam kelompok terorganisir dengan menggunakan aturan linguistic yang di rancang sendiri. Aturan – aturan ini memerintahkan system untuk menggunakan elemen teks yang relevan secara semantik untuk mengidentifikasi kategori yang relevan berdasarkan isinya. Setiap aturan terdiri dari pola dan kategori yang diprediksi (Poria).

Rule dibuat berdasarkan karakter dari masing-masing kelas dan dinotasikan dalam bentuk “IF ...(kondisi)... THEN ...(solusi)...”. Dimana “IF” merupakan kondisi prasyarat (rule antecedant) yang terdiri dari satu atau lebih atribut tes, dimana tesnya bersifat logika. Sedangkan “THEN” merupakan konsekuen (rule consequent) yang berisi hasil prediksi kelas. Misalnya untuk memisahkan data ke dalam kelas positif, negatif dan netral rule yang bisa digunakan adalah :

- Jika jumlah frekuensi kemunculan kata bersentimen positif lebih banyak dari kata bersentimen negatif, maka data digolongkan sebagai kelas positif.
- Jika jumlah frekuensi kemunculan kata bersentimen negatif lebih banyak dari kata bersentimen positif, maka data digolongkan sebagai kelas negatif.
- Jika jumlah frekuensi kemunculan kata bersentimen positif sama dengan kata bersentimen negatif, maka data digolongkan sebagai kelas netral.
- Jika tidak ditemukan kata bersentimen positif maupun kata bersentimen negatif pada data, maka digolongkan sebagai kelas netral.

Aturan (rule) yang digunakan bisa bersifat mutually exclusive atau exhaustive. Mutually exclusive berarti classifier mengandung aturan-aturan yang bersifat independen satu sama lain, sedangkan exhaustive berarti classifier mengandung aturan-aturan yang mencatat setiap kemungkinan kombinasi nilai atribut. Dimana setiap record hanya boleh dilingkupi paling banyak satu aturan saja (Lailiyah M. , 2017).

Contoh penerapan Rule Based yakni pada penelitian dengan judul Twitter Rule-based Sentiment Degree Measurement of Opinion Mining of Community Participatory in the Government of Surabaya oleh Berlian Juliartha , dengan permasalahan menantang bagi Diskominfo Surabaya untuk menetapkan target dengan memberikan respon balik dalam 24 jam. Karena kompleksitas tugas untuk

mengatasi tingkat partisipatif dan untuk mengkategorikan kelompok partisipatif, mereka menghadapi kesulitan untuk memenuhi target. Dalam penelitian ini, kami menyajikan sistem baru untuk mengukur tingkat sentimen masyarakat. Mereka menyediakan 5 fungsi dalam system yakni :

- a. Pengumpulan data
- b. Pengolahan data
- c. Pengembangan teks
- d. Analisis sentiment
- e. Validasi

mereka juga mengusulkan teknik untuk analisis sentiment pengembangan pendapat dengan 8 deteksi penting, yakni :

- a. Verb
- b. Adjective
- c. Preposisi
- d. Kata benda
- e. Adverbia
- f. Simbol
- g. Frasa
- h. Yang bersifat pujian

Pada percobaan dengan 408 data partisipatif masyarakat ditwitter untuk Diskominfo Surabaya dan dibandingkan dengan sentimen lain algoritma klasifikasi yang SVM dan Naive Bayes Classifier. Sistem kami yakni rule based melakukan tingkat akurasi 81,33% dan mengungguli algoritma pembanding lainnya.

Pada pengembangan rule baru yang mengacu pada penelitian dengan judul Twitter Rule-based Sentiment Degree Measurement of Opinion Mining of Community Participatory in the Government of Surabaya, maka pengembangan rule based sebagai berikut :

Aturan-aturan part-of-speech yang digunakan dalam proses analisa sentiment ini akan dijelaskan seperti dibawah ini :

1) Adjektif Tunggal

Jika sebuah kalimat mengandung adjektif tanpa diikuti verb dan preposisi, maka nilai kalimat sama dengan nilai adjektif. Sebagai contoh:

“pelanggan marah”

Maka akan diambil kata marah karena sesuai dengan rule. Kata marah memiliki nilai sentiment -1 dan akan menjadi representasi sentiment dari kalimat tersebut.

2) Verb Tunggal

Jika sebuah kalimat mengandung verba tanpa diikuti adjektif dan preposisi, maka nilai kalimat sama dengan nilai adjektif. Sebagai contoh kalimat “pesanan grab bike saya di batalkan”. Kata batalkan memiliki kelas kata verba dengan nilai sentiment -1. Sedangkan kata pesanan dan grab bike tidak terdapat dalam kamus sentiment dan kata saya memiliki kelas kata nomina, sehingga representasi sentiment kalimat tersebut adalah -1 (negatif)..

3) Preposisi + Adjektif

Jika terdapat adjektif setelah preposisi maka nilai kalimat dihitung dengan rule AND operator.

Preposisi	Adjektif	Nilai
1	1	1
1	-1	-1
-1	1	-1
-1	-1	1

Tabel 2. 1 Preposisi dan Adjektif

Sebagai contoh pada kalimat “pesanan makanan saya kenapa sangat lama sih?”. kata sangat merupakan kelas kata preposisi yang memiliki nilai sentiment 1, sedangkan kata lama merupakan kelas kata adjektif yang memiliki sentiment -1. Cara menghitung representasi sentiment pada contoh kalimat diatas adalah menggunakan and operator seperti pada tabel diatas sehingga nilai sentiment kalimat diatas adalah -1 (negatif).

4) Preposisi + Verb

Jika terdapat verb setelah preposisi maka nilai kalimat dihitung dengan AND operator.

Preposisi	Verb	Nilai
1	1	1
1	-1	-1
-1	1	-1
-1	-1	1

Tabel 2. 2 Preposisi dan Verb

Sebagai contoh kalimat “saya sudah complain terus kenapa tidak di respon dengan cepat”. Kata tidak merupakan kelas kata preposisi yang memiliki nilai sentiment -1, sedangkan kata respon merupakan kelas kata verba yang memiliki nilai sentiment 1. Cara menghitung representasi sentiment pada contoh kalimat diatas adalah menggunakan and operator seperti pada tabel di atas sehingga nilai sentiment kalimat diatas adalah -1 (negatif).

5) Verb + Adjektif

Rule NAND operator dapat merepresentasikan sentiment value kalimat yang mengandung verb diikuti dengan adjektif.

Verb	Adjektif	Nilai
1	1	1
1	-1	-1
-1	1	-1
-1	-1	-1

Tabel 2. 3 Verb dan Adjektif

Seperti contoh pada kalimat “kinerja driver nya aman dan cepat, jadi tepat waktu deh”. Kata kinerja merupakan kelas kata verba yang memiliki nilai sentiment 1, sedangkan kata aman merupakan kelas kata adjektif yang memiliki nilai sentiment 1. Kedua pola kelas kata tersebut masuk dalam rule kelima ini, sehingga cara menghitungnya menggunakan nand operator sesuai pada tabel diatas dan nilai sentiment rule kata tersebut menjadi 1. Terdapat satu kata lagi yang sesuai dengan rule adjektif tunggal yakni kata cepat yang

termasuk kelas kata adjektif dengan nilai sentiment 1. Apabila dalam kalimat terdapat lebih dari satu rule, maka cara menghitungnya cukup ditambahkan. Jadi nilai sentiment pada contoh kalimat diatas adalah $1+1 = 2$.

6) Preposisi + Verb + Adjektif

Terkadang preposisi, verb dan adjektif terlihat bersamaan dalam suatu kalimat. Untuk mengkalkulasi sentiment, pertama sistem menghitung verb dan adjektif menggunakan NAND operator seperti pada tabel verb asjektif Kedua, nilai kata preposisi dengan hasil dari perhitungan pertama dihitung menggunakan AND operator seperti pada tabel Preposisi + Verb

Sebagai contoh kalimat “driver tidak paham dengan baik, rute yang akan di lalui”. Kata tidak pada contoh kalimat diatas termasuk kelas kata preposisi dengan nilai sentiment -1. Kata paham termasuk kelas kata verba dengan nilai sentiment 1. Kata baik termasuk kelas kata adjektif dengan nilai sentiment 1. Karena ketiga kata tersebut masuk ke rule keenam, maka cara menghitungnya terlebih dahulu menghitung kelas kata verba dan adjektif menggunakan operator nand, sehingga $1 \text{ nand } 1 = 1$. Setelah itu barulah kelas kata preposisi dioperasi dengan hasil dari kelas kata verba dan adjektif menggunakan operasi and, sehingga $-1 \text{ and } 1 = -1$. Jadi kalimat diatas memiliki nilai sentiment-1 (negatif).

7) Semua symbol pada komentar dihapus kecuali dot (.), comma (,), tanda seru (!) dan tanda tanya (?).

Saat preprocessing tanda dot, comma, tanda seru dan tanda tanya tidak dihapus. Dot, tanda seru dan tanda tanya akan dirubah menjadi comma. Hal tersebut dilakukan agar pada aturan ke-8 kalimat lebih mudah dipisahkan hanya dengan parameter comma.

8) Dot and comma parameter untuk memisahkan kalimat.

Setiap tanggapan dan komentar biasanya mengandung beberapa kalimat yang dipisahkan oleh koma dan titik. Contoh “pesanan di batalkan driver, apakah tidak ada tindak lanjutnya?”. Komentar atau tanggapan tersebut akan dipisahkan menjadi dua kalimat. Nilai sentiment kalimat pertama “pesanan di batalkan driver” bernilai -1 sedangkan kalimat kedua setelah melakukan pemrosesan menjadi kata “apakah tidak ada tindak

lanjutnya” bernilai -1. Untuk menghitung nilai sentiment kedua kalimat maka dijumlahkan nilai kedua kalimat $(-1) + (-1) = -2$ (negatif).

- 9) Frase kata terdiri dari dua kata yang memiliki satu makna tidak dapat dipisah.

Dalam penelitian ini, frasa adalah dua kata yang memiliki satu makna yang biasanya bertemu pada kalimat bahasa Indonesia. Contoh-contoh itu adalah terima kasih, air mata, minta tolong dll. Itu hanya memiliki satu nilai.

- 10) Nilai sentiment “terima kasih” atau pujian bergantung pada posisi letak kata tersebut dalam kalimat.

Jika ‘terima kasih’ berada di depan kalimat, maka nilai sentiment kata ‘terima kasih’ bernilai 1, akan tetapi jika kata ‘terima kasih’ terletak di tengah atau di akhir kalimat maka nilai sentiment kata ‘terima kasih’ adalah 0. Contoh “terima kasih sudah cepat ” maka kata tersebut memiliki nilai 1.

- 11) Penggunaan kata tidak sesuai kamus Bahasa Indonesia akan dibaikan.

Biasanya ditemukan bahasa non-formal, salah ejaan, dan bahasa Inggris pada teks komentar ojek online. Itu bisa diperbaiki pada langkah pemrosesan tetapi jika hasilnya masih tidak sesuai dengan kata dalam kamus bahasa Indonesia, itu akan dibaikan dan akan dianggap sebagai tidak dikenal.

- 12) Adverb + Adjektif

Jika terdapat adjektif setelah Adverb maka nilai kalimat dihitung dengan rule NAND operator.

Adverb	Adjektif	Nilai
1	1	1
1	-1	-1
-1	1	-1
-1	-1	-1

Tabel 2. 4 Adverb dan Adjektif

13) Adverb + Verb

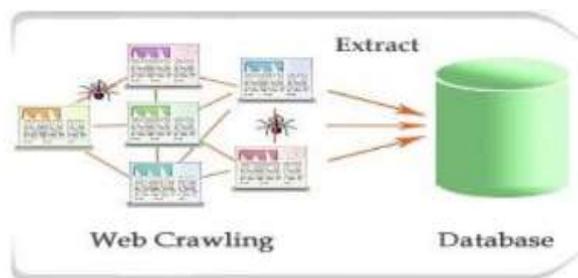
Jika terdapat Verb setelah Adverb maka nilai kalimat dihitung dengan rule NAND operator

Adverb	Adjektif	Nilai
1	1	1
1	-1	-1
-1	1	-1
-1	-1	-1

Tabel 2. 5 Adverb dan Adjektif

1.5 *Crawling*

Crawling adalah suatu Teknik yang digunakan untuk mengumpulkan informasi dalam web. Crawling bekerja secara otomatis, dimana informasi yang dikumpulkan berdasarkan atas kata kunci yang diberikan oleh pengguna (Saputra, 2017).



Gambar 2. 1 Crawling

Web crawler akan mengunjungi setiap alamat website yang diberikan kepadanya, kemudian menyerap dan menyimpan semua informasi yang terkandung didalam website tersebut. Setipa kali web crawler mengunjungi sebuah website, maka dia juga akan mendata semua link yang ada dihalaman yang dikunjunginya itu untuk kemudian dikunjungi lagi satu persatu.

Proses web crawler dalam mengunjungi setiap dokumen web disebut dengan web crawling atau spidering. Proses crawling dalam suatu website dimulai dari mendata seluruh url dari website, menelusurinya satu-persatu, kemudian memasukkannya dalam daftar halaman pada indeks search engine, sehingga setiap kali ada perubahan pada website, akan terupdate secara otomatis. Web crawling

adalah proses mengambil kumpulan halaman dari sebuah web untuk dilakukan pengindeksan sehingga mendukung kinerja mesin pencari. Salah satu contoh situs yang menerapkan web crawling adalah www.webcrawler.com . Di samping situs-situ mesin pencari terkemuka tentunya, seperti Google, Yahoo, Ask, Live, dan lain sebagainya.

1.6 *Twitter*

Twitter merupakan layanan bagi setiap orang untuk saling berinteraksi melalui suatu hal (berupa foto, video ataupun text). .Di Twitter kita bisa menuliskan sebuah status atau pesan yang nantinya akan dibaca oleh pengguna lainnya, fitur ini disebut tweet atau komentar. Jumlah pengguna harian Twitter secara global diklaim meningkat. Pada laporan finansial Twitter kuartal ke-3 tahun 2019, pengguna aktif harian di platform Twitter dicatat meningkat 17 persen, ke angka 145 juta pengguna.(Kompas,2019)