

Pengembangan Analisis Sentimen Pada Twitter Untuk Layanan Ojek Online Menggunakan Rule Based

Dwi Puspitasari¹, Dian Hanifuddin Subhi², Beby Novia Putri Lovenida³

^{1,2,3} Teknologi Informasi, Teknik Informatika, Politeknik Negeri Malang

¹dwi.puspitasari@polinema.ac.id, ²dhanifudin@gmail.com, ³Bebyrpl@gmail.com

Abstrak— Ojek online terus berupaya memberikan pelayanan terbaik untuk memuaskan pelanggan mereka seperti memberikan sejumlah diskon di layanan yang ada ataupun memberikan jenis pelayanan yang baru. Dengan perkembangan di setiap layanan, membuat minat masyarakat terhadap Gojek dan Grab semakin tinggi. Apresiasi juga sering diberikan terhadap dua perusahaan ojek online tersebut. Media social merupakan wadah masyarakat untuk mengaspresiasi Gojek dan Grab, baik berupa pujian, saran maupun kritik. Salah satu media social yang menjadi wadah apresiasi adalah twitter. Opini yang disampaikan pelanggan melalui akun resmi twitter Gojek yakni @gojekindonesia dan Grab yakni @GrabId bisa dijadikan alat ukur kepuasan pelanggan terhadap layanan ojek online di kedua penyedia layanan ojek online tersebut dengan menganalisis twitter mereka. Opini yang berupa komentar pelanggan akan di Analisa, apakah opini tersebut bisa di klasifikasikan menjadi positif, negative ataupun netral. Analisis sentiment ini menggunakan rule based untuk mengukur klasifikasi opini pelanggan. Data yang digunakan yakni 1000 tweet. Pengujian dilakukan menggunakan *k-fold cross validation* dengan hasil akurasi klasifikasi rule based ini mendapatkan akurasi terbaik 79%, *Precision* terbaik 100%, *Recall* terbaik 80% dan *f-measure* terbaik 79%

Kata kunci— Ojek Online, Sentimen Analisis, Rule Based

I. PENDAHULUAN

Di Indonesia ojek online mulai banyak bermunculan, dari perusahaan yang besar seperti Gojek dan Grab, hingga perusahaan yang sedang berkembang di beberapa kota yang belum terjamah kedua perusahaan ojek online tersebut. Kedua perusahaan tersebut terus berupaya memberikan pelayanan terbaik mereka seperti memberikan sejumlah diskon di layanan yang ada. Dengan perkembangan di setiap layanan, membuat minat masyarakat terhadap Gojek dan Grab semakin tinggi. Apresiasi juga sering diberikan terhadap dua perusahaan ojek online tersebut. Media social

merupakan wadah masyarakat untuk mengaspresiasi Gojek dan Grab, baik berupa pujian, saran maupun kritikan.

Banyak organisasi, perusahaan bisnis khususnya menggunakan sosial ini media untuk mendapatkan beberapa keuntungan untuk bisnis mereka. Pemasaran media sosial adalah cara baru untuk mempromosikan produk dan meningkatkan pendapatan dan menjangkau lebih banyak pelanggan dengan lebih sedikit anggaran dan sumber daya. Pemasaran media sosial memberi manfaat untuk bisnis dengan mengurangi biaya dan waktu staf, dan meningkatkan generasi pendapatan[1].

Menurut Pemasaran Media Sosial Laporan Industri 2015, manfaat Pemasaran Media Sosial peningkatan eksposur ke bisnis, peningkatan lalu lintas, mengembangkan penggemar setia, memberikan wawasan pasar, dan lead yang dihasilkan [2]. Beberapa penelitian juga menunjukkan perusahaan itu dapat mengukur kepuasan pelanggan dan wawasan pelanggan dengan menganalisis data media social[3].

Analisis sentiment merupakan klasifikasi teks yang tujuannya untuk mengelompokkan teks (dokumen) yang mengandung opini sebagai positive sentiment, negative sentiment, atau netral [4]. Klasifikasi teks dapat dilakukan dengan dua cara berbeda, yakni klasifikasi manual dan otomatis. Fitur dan algoritma merupakan dua hal utama dalam text classification [5]. Secara garis besar ada dua algoritma di sentiment analysis, yakni rule based method dan statistical based method [5]. Pada pendekatan statistik proses pemisahan data sesuai kelas masing-masing menggunakan perhitungan matematis atau dikenal sebagai machine learning [6]. Sedangkan pada pendekatan rule based method menggunakan bantuan human expert berupa aturan (rule) yang digunakan untuk memisahkan data sesuai kelas masing-masing [6]

II. TINJAUAN PUSTAKA

A. Ojek Online

Ojek online merupakan hasil perkembangan dari teknologi yang memanfaatkan internet dan smartphone

Letakkan informasi sumber dana atau sponsor di sini. Jika tidak ada, maka Anda dapat menghapus kotak teks ini.

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.

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 layanan 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 [4]

B. Analisis Sentimen

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 manusia. 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 [7].

C. Rule Based

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 [6]

Rule dibuat berdasarkan karakter dari masing-masing kelas dan dinotasikan dalam bentuk “IF ...(kondisi)... THEN ...(solusi)...”. Dimana “IF” merupakan kondisi prasyarat (rule antecedent) 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 [6]:

- 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. [6]

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.

TABEL 1 PREPOSISI DAN ADJEKTIF

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

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.

TABEL 2 PREPOSI DAN VERB

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

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.

TABEL 3 VERB DAN ADJEKTIF

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

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 adjektif 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 dioperasikan 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 dibanakan.

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 diabaikan dan akan dianggap sebagai tidak dikenal.

12) Adverb + Adjektif

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

TABEL 4 ADVERB DAN ADJEKTIF

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

13) Adverb + Verb

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

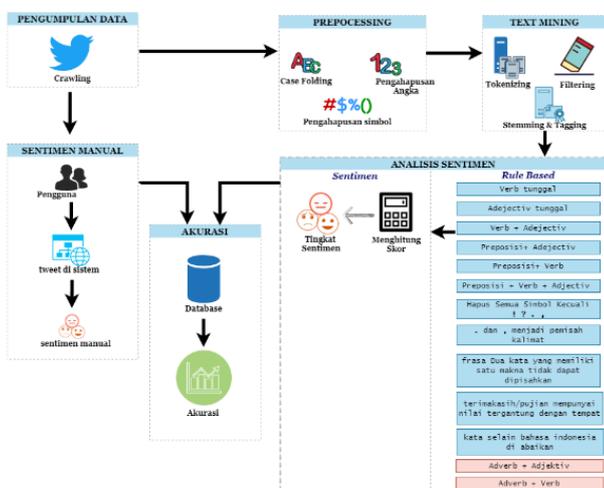
TABEL 5 ADVERB DAN ADJEKTIF

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

III. PERANCANGAN

A. Deskripsi Sistem

Sistem mempunyai beberapa proses yaitu pengumpulan data, preprocessing, text mining dan analisis sentimen. Pada proses preprocessing terdapat case folding, penghapusan angka dan penghapusan simbol. Lalu pada proses text mining terdapat tokenizing, filtering, stemming dan tagging. Dan pada proses analisis sentimen terdapat penggunaan rule based, perhitungan skor dan menentukan tingkat skor. Pada sistem juga memerlukan inputan sentiment manual untuk menentukan akurasi pada pengujian.



GAMBAR 1 DESIGN SISTEM

B. Pengambilan Data

1. 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[8].

2. Pengambilan Data untuk Kamus Jenis Kata

Pengambilan data ini berupa Multi Word Expression Lists. Multi Word Expression Lists ini berfungsi untuk membantu dalam pengenalan jenis kata dengan kata lain kamus yang akan di gunakan untuk mendeteksi jenis kata. Multi Word Expression Lists diambil dari penelitian yusuf syaifudin tentang Named Entity Recognition for Bahasa Indonesia, yang bisa di akses pada laman <https://github.com/yusufsyaifudin/indonesia-ner>

C. Pengujian K-Fold Cross Validation

K-fold adalah salah satu metode Cross Validation yang populer dengan melipat data sebanyak K dan mengulangi (iterasi) eksperimennya sebanyak K juga. Pada pengujian ini data yang akan digunakan adalah 1000 tweet termasuk data uji yang nantinya akan dibagi menjadi 10 bagian atau k=10 sehingga data yang diperoleh adalah 1000 data dibagi menjadi 10 lipatan. Selain itu akan ditentukan mana yang termasuk data training dan mana yang termasuk data testing dengan perbandingan 8:2. Pengujian menggunakan data yang sudah dipartisi akan diulang sebanyak 10 kali (K=10) dengan posisi data testing berbeda disetiap iterasinya. Misalkan iterasi pertama data tes pada posisi awal, iterasi kedua data testing di posisi kedua begitu seterusnya[9]

	Total Sampel = 1000									
iterasi 1/10	Test set									
iterasi 2/10		Test set								
iterasi 3/10			Test set							
iterasi 4/10				Test set						
iterasi 5/10					Test set					
iterasi 6/10						Test set				
iterasi 7/10							Test set			
iterasi 8/10								Test set		
iterasi 9/10									Test set	
iterasi 10/10										Test set

GAMBAR 2 ILUSTRASI 10-FOLD CROSS VALIDATION

IV. IMPLEMENTASI

A. Analisis Sentimen

Pada penelitian ini menggunakan metode Rule Based Sentiment Analysis, dengan tahapan – tahapan sebagai berikut :

- 1) Tahapan pertama yakni pengumpulan data. Data diambil dari data komentar di akun resmi twitter gojek @gojekindonesia dan grab @GrabID. Pengambilan data dilakukan dengan cara crawling. Berikut contoh dari Data yang ada

TABEL 6 CONTOH HASIL CRAWLING

2020-06-24	admin @gojekindonesia, tolong cek dm doongggg :(
2020-06-24	Eh @gojekindonesia dr sore kagak bs login ini, gak bs balik Knp gak ada opsi OTP via call sih ? Nyebelin sumpah
2020-06-24	Kaget bgt @gojekindonesia phk 400+

	karyawannya :(Padahal aku baru aja browsing buat apply job di sana :(
--	--

- 2) Hasil *crawling* disimpan di dalam database, untuk selanjutnya data akan otomatis menampilkan sentiment hasil dari sistem.
- 3) Untuk tahap selanjutnya yaitu *text preprocessing*. Dimana proses klasifikasi pada tahapan ini akan dilakukan pengecekan apakah data mengandung Case Folding, tanda baca dan penghapusan nomor. Hal tersebut untuk mengurangi bagian tidak penting pada proses analisis sentiment. Contoh proses *text preprocessing* adalah

TABLE 1 HASIL TEXT PREPROCESSING

tweet	preprocessing
admin @gojekindonesia, tolong cek dm doongggg :(admin tolong cek dm doongggg
Eh @gojekindonesia dr sore kagak bs login ini, gak bs balik Knp gak ada opsi OTP via call sih ? Nyebelin sumpah	Eh dr sore kagak login ini, gak bs balik knp gak ada opsi otp via call sih? nyebelin sumpah
Kaget bgt @gojekindonesia phk 400+ karyawannya :(Padahal aku baru aja browsing buat apply job di sana :(kaget bgt phk karyawannya padahal aku baru aja browsing buat apply job di sana :(

- 4) Setelah proses *text preprocessing* selesai, proses selanjutnya yakni *text mining*. Tahapan *text mining* adalah
 - Tokenizing merupakan proses untuk memecah urutan karakter menjadi kata per kata.
 - Filetering merupakan proses menghapus kata tidak penting
 - Stemming and tagging merupakan mencari kata dasara dan menghilangkan imbuhan seperti: membayar menjadi bayar.

TABLE 8 HASIL TEXT MINING

preprocessing	hasil text mining
admin tolong cek dm doongggg	tolong cek dm doongggg
Eh dr sore kagak login ini, gak bs balik knp gak ada opsi otp via call sih? nyebelin sumpah	Eh dr sore kagak login ini, gak bs balik knp gak ada opsi otp via call sih? nyebelin sumpah
kaget bgt phk karyawannya padahal aku baru aja browsing buat apply job di sana	kaget bgt phk karyawannya padahal browsing apply job

- 5) Tahap selanjutnya yaitu analisis sentiment dengan rule based yang sudah di jelaskan di atas. Kata yang sudah melalui tahapan *preprocessing* dan *text mining* akan di klasifikasikan berdasarkan rule based yang ada dan tahapan *word degree* ini, kata akan di klasifikasikan apakah kata tersebut bersentimen positif, negative ataupun netral. Setiap kata memiliki nilai yang berbeda. Maka proses pemberian nilai sebagai berikut :

TABLE 2 HASIL ANALISIS SENTIMEN

tweet	sentiment
tolong cek dm doongggg	tolong Verba 1 cek Nomina 0 = 1
Eh dr sore kagak login ini, gak bs balik knp gak ada opsi otp via call sih? nyebelin sumpah	sore Nomina 0 kagak Adverbia 0 opsi Nomina 0 via Preposisi 0 sumpah Nomina 0 = 0
kaget bgt phk karyawannya padahal browsing apply job	kaget Adjektiva -1 karyawan Nomina 0 = -1

B. Pengujian K-fold Cross Validation

Untuk pengujian fold ke-1 sampai fold ke-10 jumlah data tes yang dimasukkan adalah 100 tweet dengan posisi data tes seperti digambarkan pada Gambar 2 dan hasilnya dapat dilihat pada tabel berikut:

TABLE 9 PENGUJIAN K-FOLDING

	Precision	Recall	Akurasi	F-measure
iterasi 1	78	80	79	79
iterasi 2	83	10	57	18
iterasi 3	67	50	57	57
iterasi 4	21	54	43	31
iterasi 5	67	10	28	17
iterasi 6	31	42	38	36
iterasi 7	67	61	62	64
iterasi 8	100	59	71	74
iterasi 9	46	50	48	48
iterasi 10	33	68	64	45
rata-rata	59	48	55	47

Dari tabel 9 dapat dilihat bahwa akurasi yang didapat dari pengujian data tes yang tertinggi pada saat pengujian *fold* ke-1 yakni 79% dan akurasi terendah pada saat pengujian *fold* ke-5 yaitu 28% dengan rata rata akurasi 55%. Presisi tertinggi pada saat pengujian *fold* ke-8 yaitu 100% dan presisi terendah pada saat pengujian *fold* ke-4 yaitu 21% dengan rata rata presisi 59%. *Recall* tertinggi pada saat pengujian *fold* ke-1 yaitu 80% dan *Recall* terendah pada saat pengujian *fold* ke-2 dan ke-5 yaitu 10% dengan rata rata *Recall* yaitu 48%. Hasil pengujian *f-measure* tertinggi pada saat pengujian *fold* ke-1 yakni 79% dan *f-measure* terendah pada saat pengujian *fold* ke-5 yaitu 17% dengan rata rata *f-measure* 47%.

V. HASIL DAN PEMBAHASAN

A. Pengembangan rule based

Pengembangan rule based yang dilakukan mengacu pada penelitian sebelumnya, dimana di penelitian sebelumnya mendeteksi bagian yaitu (1) Verb, (2) Adjective, (3) Preposition, (4) Noun, (5) Adverb, (6) Symbol, (7) Phrase, and (8) Complimentary. Pada pengembangan kali ini hal yang di lakukan adalah menambah rule berupa menspesifikasi perhitungan pada deteksi adverb. Pada penelitian sebelumnya deteksi adverb di lakukan dengan

mendeteksi preposisi + verb + adjektif dihitung dengan Nand lalu And. Untuk pengembangan rule baru adverb di lakukan dengan mendeteksi 2 bagian yakni : adverb + verb dan adverb+ adjektif dimana keduanya di hitung dengan perhitungan Nand.

B. Kendala sistem dalam memahami sentimen

Kendala yang sering ditemui dalam memahami kalimat untuk diolah dalam proses klasifikasi sentimen yaitu Ketika dalam melabeli data secara manual menggunakan opini beberapa orang, namun Ketika sistem yang akan melabeli data maka sistem melalui setiap proses sehingga menghasilkan sentiment, dimana hal tersebut terkadang mengakibatkan output yang berbeda. Contoh : “ Ini fiturnya yang nganterin makanan barengan beberapa customer bikin pingsan. Udah lama nambah lama. ??? “. Dalam contoh tersebut pelabelan manual akan memberikan sentiment negative, namun sistem akan melabeli sentiment netral di karenakan sistem melalui beberapa proses untuk bisa melabeli kalimat tersebut.

Kendala yang menyebabkan sentiment manual dan sentimen sistem berbeda adalah sistem tidak bisa membaca kata yang tidak terdapat di kamus. Kebanyakan kata tersebut berupa kata yang tidak baku seperti kata “gemes” kata tersebut tidak ada di kamus dikarenakan tidak baku, kata bakunya yaitu “gemas”. Jika sistem menemui sebuah kata yang mempunyai 2 makna atau jenis kata, maka sistem akan membaca kata berdasarkan kamus jenis kata yang lebih dahulu contoh mental (Keadaan kejiwaan seseorang) yang masuk dalam kamus adjektiv sedangkan mental (terpelanting atau terlempar karena berbenturan dengan benda lain) yang masuk dalam kamus verb.

Hal-hal diatas merupakan beberapa kendala sistem dalam memeberikan sentiment suatu kalimat. Kendala tersebut juga dapat mempengaruhi tingkat akurasi sistem, sehingga mengakibatkan kinerja sistem tidak berjalan secara maksimal.

KESIMPULAN

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan dapat ditarik kesimpulan sebagai berikut:

1. Metode *Rule Based* dapat di gunakan untuk melakukan klasifikasi sentimen tentang ojek online pada twitter
2. Hasil uji akurasi klasifikasi yang dilakukan oleh aplikasi menggunakan *10-fold cross validation* dengan rata rata adalah 55%, *Precision* 59%, *Recall* 48% dan *f-measure* 47% dan akurasi terbaik 79%,*Precision* terbaik 100%, *Recall* terbaik 80% dan *f-measure* terbaik 79%.

DAFTAR PUSTAKA

- [1] Vidya, N.A., Fanany, M.I., & Budi, I., 2015. Twitter Sentiment to Analyze Net Brand Reputation of Mobile Phone
- [2] N. A. Vidya, M. I. Fanany and I. Budi, "Twitter Sentiment to Analyze Net Brand Reputation," *Procedia Computer Science* 72, pp. 519-526, 2015.
- [3] L. Zhang, K. Huac, H. Wang, G. Qiane and L. Zhang, "Sentiment Analysis on Reviews of Mobile Users," *Proceedings of The 11th International Conference on Mobile Systems and Pervasive Computing*, pp. 458-465, 2014.

- [4] Liu, Bing. 2010. *Sentiment Analysis and Subjectivity*. Chicago: University of Illinois.
- [5] Atmadja, A. R. & Purwarianti, A., 2015. Comparison on The Rule Based Method and Statistical Based Method on Emotion Classification for Indonesian Twitter Text. Bali, s.n.)
- [6] Lailiyah, M. (2017). *Sentiment Analysis Menggunakan Rule Based Method pada Data Pengaduan Publik Berbasis Lexical Resource*. Surabaya: Masters thesis Institut Teknologi Sepuluh Nopember
- [7] Benamara, F., Cesarano, C., & Reforgiato, D., (2006). *Sentiment Analysis: Adjectives and Adverbs are better then Adjectives Alone*. ICWSM'2006 Boulder, CO USA.
- [8] Saputra, Pramana Yoga, "Implementasi Teknik Crawling Untuk Pengumpulan Data Dari Media Sosial Twitter", *Jurnal Dinamika Dotcom* Vol 8 Nomor 2, Juli 2017
- [9] Ratnawati F. *Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter*. 2018;