**KATA PENGANTAR**

Dengan menyebut nama Allah SWT yang Maha Pengasih lagi Maha Panyayang, Kami panjatkan puja dan puji syukur atas kehadirat-Nya, yang telah melimpahkan rahmat, hidayah, dan inayah-Nya kepada kami, sehingga kami dapat menyelesaikan makalah tentang “Analisis Sentimen Pada Sosial Media”.

Makalah ilmiah ini telah kami susun dengan maksimal dan mendapatkan bantuan dari berbagai pihak sehingga dapat memperlancar pembuatan makalah ini. Untuk itu kami menyampaikan banyak terima kasih kepada semua pihak yang telah berkontribusi dalam pembuatan makalah ini.

Terlepas dari semua itu, Kami menyadari sepenuhnya bahwa masih ada kekurangan baik dari segi susunan kalimat maupun tata bahasanya. Oleh karena itu dengan tangan terbuka kami menerima segala saran dan kritik dari pembaca agar kami dapat memperbaiki makalah ilmiah ini.  
 Akhir kata kami berharap semoga makalah ilmiah tentang limbah dan manfaatnya untuk masyarakan ini dapat memberikan manfaat maupun inpirasi terhadap pembaca.

      
                                                                                        Jakarta,  Maret 2015

                                                                                              Penyusun

# DAFTAR ISI

KATA PENGANTAR…………………………………………………………………………………..1

DAFTAR ISI…………………………………………………………………………………….2

BAB I PENDAHULUAN

1. Latar Belakang……………………………………………………………………………3
2. Rumusan Masalah………………………………………………………………………...4
3. Tujuan Penulisan Makalah………………………………………………………………..4

BAB II PEMBAHASAN

1. Definisi Analisis sentiment……………………………………………………………….5
2. Pembahasan Penyedia Layanan Telekomunikasi selluler pada Twitter………………….5
3. Pengaruh parameter support vector machine dan Lexicon Based

Features dan Akurasi Sistem……………………………………………………………..6

BAB III PENUTUP

1. Kesimpulan………………………………………………………………………………10

DAFTAR PUSTAKA……………………………………………………………………………11

# BAB I

# PENDAHULUAN

1. Latar belakang

Teknologi komunikasi semakin berkembang dari tahun ke tahun terutama teknologi telepon seluler. Untuk mendukung kelancaran komunikasi dengan orang lain, masyarakat membutuhkan penyedia layanan telekomunikasi seluler yang dapat menunjang jalanya telepon seluler mereka dan sesuai dengan kebutuhan. Perusahan –perusahaan penyedia layanan telekomunikasi seluler berlomba-lomba untuk meningkatkan inovasi dan mutu pada layanan produk mereka,misalnya seperti biaya bertelepon yang semakin murah, bonus layanan pesan singkat yang semakin banyak paket layanan data internet dengan kuota yang banyak dan kecepatan koneksi yang tinggi, serta sinyal yang dapat menjangkau ke semua wilayah. Tercatat tiga penyedia layanan telekomunikasi yang paling popular di Indonesia yaitu XL Axiata, Telkomsel, dan Indosat

Setiap penyedia layanan telekomunikasi seluler mempunyai keunggulan tersendiri yang mampu memuaskan kebutuhan penggunanya sekaligus mempunyai kekurangan yang membuat masyarakat mengeluh dan berpindah ke telekomunikasi seluller yang lainya sehingga setiap perusahaan penyedia layanan telekomunikasi seluler pun perlu memperhatikan apa yang menjadi keunggulan dan kelemahan produk layanan sendiri atau perushaan lain agar tidak terjadi penurunan layanan dan berpindahanya konsumen ke penyedia jasa yang lain

Agar dapat mengetahui bagaimana tingkat kepuasan dari pengguna terhadap suatu penyedia layanan telekomunikasi seluler, maka diperlukan analisis sentimen pada Twitter pengguna yang menyangkut penyedia layanan telekomunikasi seluler tersebut. Analisis sentimen atau opinion mining adalah studi komputasional dari opini-opini orang, sentimen, dan emosi melalui entitas dan atribut yang dimiliki yang diekspresikan dalam bentuk teks Analisis sentimen ini dapat mengelompokkan polaritas dari teks dalam kalimat atau dokumen untuk mengetahui apakah opini pada kalimat atau dokumen tersebut apakah termasuk positif atau negative metode yang digunakan adalah vector machine klasifikasi dengan support machine proposes teks terdiri dari tokenisasi, data cleaning, case floding ,filterisasi dan stemming. Hasil dari praproses teks akan dilakukan pembobotan. Pertama,pembobotan kata yang terdiri dari perhitungan TF,WTF, DF, IDF, dan TF-IDF. Kemudian pembobotan Lexicon Based Features dengan cara mengecek kata bersentimen positif atau negatif berdasarkan kamus atau lexiconyang ada di sentiment dictionaries(data yang berisi kata-kata bersentimen) dan menghitung frekuensi kemunculannya pada dokumen teks

.

frekuensi kata tersebut dinormalisasi dengan metode Min-max dan menjadi bobot untuk Lexicon Based Features. Setelah fitur pembobotan kata dan Lexicon Based Featuresdiperoleh, maka data-

data tersebut akan diklasifikasi dengan metode Support Vector Machine. Data latih yang mengandung fitur bobot TF-IDF dan Lexicon Based Featuresakan diolah dalam proses pelatihan hingga terbentuk suatu hyperplane pemisah antar kelas data latih. Setelah terbentuk hyperplane tersebut, maka data uji dapat dimasukkan ke dalam fungsi sign(h(x)) untuk mendapatkan hasil klasifikasi analisis sentimen berupa sentimen positif atau negatif.

Dengan penelitian ini diharapkan agar para perusahaan penyedia layanan telekomunikasi seluler dapat melihat bagaimana tingkat kepuasan masyarakat terhadap produk layanan mereka sehingga mereka akan terpacu untuk meningkatkan layanan dan mutu pada produk layanan mereka.

1. Rumusan Masalah

Berdasarkan latar belakang di atas maka rumusan makalah pada makalah ini adalah sebagai berikut :

1. Definisi Analisis Sentimen
2. Pembahasan Penyedia Layanan Telekomunikasi selluler pada Twitter
3. Pengaruh parameter support vector machine dan Lexicon Based Features dan Akurasi Sistem
4. Hasil Pembahasan
5. Tujuan Penulisan Makalah
6. Untuk menambah ilmu pengetahuan mengenai masalah yang di angkat dalam makalah
7. Untuk mengetahui pengguna penyedia layanan telekomunikasi selluler di Twitter

# BAB 2

# PEMBAHASAN

1. Definisi Analisis Sentimen

Analisis sentimen atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini.Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif. Salah satu contoh penggunaan analisis sentimen dalam dunia nyata adalah identifikasi kecenderungan pasar dan opini pasar terhadap suatu objek barang. Besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian

1. Pembahasan Penyedia Layanan Telekomunikasi selluler pada Twitter

Tahapan proses pada sistem analisis sentimen terhadap tingkat kepuasan pengguna penyedia layanan telekomunikasi seluler pada Twitter dengan metode Support Vector Machine dan Lexicon Based Featuresantara lain praproses teks, pembobotan kata, pembobotan Lexicon Based Features Pembobotan kata bertujuan untuk memberikan bobot pada fitur kata berdasarkan frekuensi kemunculan kata.Fitur kata yang telah diberi bobot dapat digunakan untuk proses klasifikasi. Tahapan pembobotan kata yaitu:

1. Term Frekuency (TF)

Term frequency atau tf merupakan jumlah kemunculan atau frekuensi kata pada suatu dokumen. Sementara Wtf adalah jumlah bobot dari tf yang telah dihitung dengan logaritma.

1. Documen Frekuency (DF)

Document Frequency(df) merupakan frekuensi atau jumlah dokumen yang mengandung suatu kata

1. Inverse Documen Frekuency (IDF)

adalah bobot kebalikan dari bobot document frequency. Kata yang jarang muncul di banyak dokumen mempunyai bobot Inverse Document Frequency

yang tinggi. Persamaan dari Inverse Document Frequency(idf)

Keterangan :

N : jumlah dokumen text

Dft : jumlah dokumen yang mengandung suatu jata t.

1. Term Frequency-Inverse Document Frequency (tf-idf)

Pembobotan ini adalah hasil perkalian dari pembobotan term frequency dan inverse document frequency dari suatu term

Ketarangan :

Wtd : Term Frekuenci

Idft : Inverse Document Frequency

1. Pengaruh parameter support vector machine dan Lexicon Based Features dan Akurasi Sistem
2. Support Vector Machine

adalah suatu teknik yang relatif baru untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. Support Vector Machine masuk kelas supervised learning, dimana dalam implementasinya perlu adanya tahap pelatihan menggunakan sequential training SVM dan disusul tahap pengujian Konsep klasifikasi dengan Support Vector Machine adalah mencari hyperplaneterbaik yang berfungsi sebagai pemisah dua kelas data. Konsep klasifikasi dengan Support Vector Machine adalah mencari hyperplaneterbaik yang berfungsi sebagai pemisah dua kelas data. Support Vector Machine mampu bekerja pada dataset yang berdimensi tinggi dengan menggunakan kernel trik. Support Vector Machine hanya menggunakan beberapa titik data terpilih yang berkontribusi (support vector) untuk membentuk model yang akan digunakan dalam proses klasifikasi. Ilustrasi metode Support Vector Machine

persamaan support vector machine

f(x)= w.x +b

atau

f(x)=

Keterangan :

W : parameter hyperplane yang dicari (garis yang tegak lurus antara garis Hyperplane dan titik support vector)

X : titik data masukan Support Vector Machine

Ai : nilai bobot setiap titik data

K(x,xi) : fungsi kernel

B : parameter hyperlane yang dicari (nilai bias)

Untuk penelitian ini menggunakan kernel polynominal. Persamaanya :

K(x,y)=(x.y+c)d

Keterangan :

K(x,y) :Nila kernel dari data x dan data y

x : fiur data 1

y : fitur data 2

c : nilai konsta

1. Pembobotan Lexicon Based Features

Lexicon Based Features merupakan fitur kata yang memiliki sentimen positif atau negatif berdasarkan kamus atau lexicon. Lexicon adalah kumpulan kata sentimen yang telah diketahui dan terhimpun Untuk proses pembobotan pada fitur ini, dibutuhkan kamus atau lexicon yang berisi kata-kata yang mengandung sentimen yang disebut dengan sentiment dictionaries Sentimen yang digunakan berupa positif dan negatif. Terdapat dua macam fitur yang digunakan pada penelitian ini yaitu fitur jumlah kata positif dan fitur jumlah kata negatif dalam dokumen. Agar bobot Lexicon Based Features seimbang dengan bobot tf-idf, maka fitur jumlah kata positif dan kata negatif perlu dinormalisasikan dengan metode Min-max

1. Normalisasi Min-Max

Normalisasi Min-max merupakan proses transformasi yang mana merubah nilai suatu variabel diskalakan menjadi berukuran lebih kecil dengan jangkauan antara -1 sampai1 atau 0,0 sampai 0,1 normalisasi digunaka untuk mengurangi kesalahan pada proses data mining Metode normalisasi ini merupakan metode yang paling sederhana dengan melakukan transformasi linier terhadap data asli dan memiliki kelebihan yaitu terdapat keseimbangan nilai perbandingan antara nilai data sebelum dinormalisasi dengan nilai data yang telah dinormalisasi.

Keterangan :

V’i : Nilai data baru hasil normalisasi min-max

Vi : Nilai data yang akan di normalisasi

Min :Nilai minimum data

Max : Nilai maximum data

New\_min :Nilai minimum yang di harapkan dari proses normalisasi

New\_max :Nilai maxsimum yang di harapkan dari proses normalisasi

1. Pengujian Pengaruh Lexicon Based Featuresdan Akurasi Sistem

Pada pengujian pengaruh Lexicon Based Featuresterhadap analisis sentimen menggunakan Support Vector Machine ini akan dilakukan perbandingan hasil akurasi sistem yang menggunakan Lexicon Based Features dan tanpa menggunakan Lexicon Based Features. Nilai-nilai parameter Support Vector Machine yang digunakan dalam pengujian ini adalah nilai parameter Support Vector Machine optimal berdasarkan hasil pengujian sebelumnya yaitu nilai degree= 2, nilai konstanta learning rate= 0,0001, dan jumlah iterasi maksimum = 50 kali. Sistem analisis sentimen menggunakan metode Support Vector Machine dengan menggunakan Lexicon Based Features atau tanpa menggunakan Lexicon Based Features. Hasil pengujian pengaruh Lexicon Based Features dan akurasi sistem ditunjukkan pada Tabel berikut

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | Dengan Lexicon  Based Features | | Tanpa Lexcion  Based Features | |
|  | | Hasil Sebenarnya | | Hasil Sebenarnya | |
| Hasil  prediksi |  | Positif | Negative | Positif | Negative |
| Positif | 33 | 18 | 26 | 7 |
| Negatif | 1 | 38 | 8 | 48 |
| Accuracy | | 79% | | 84% | |
| Precision | | 69% | | 76% | |
| Reccal | | 97% | | 86% | |
| F-Measure | | 78% | | 80% | |

Hasil Pengujian Pengaruh Lexicon Based Features dan Akurasi Sistem

Dari hasil pengujian tersebut dapat dilihat pada grafik pengujian pengaruh Lexicon Based Features dan akurasi sistem yang ditunjukkan dari hasil pengujian pengaruh Lexicon Based Featuresterhadap sistem analisis sentimen menunjukkan bahwa tingkat akurasi sistem analisis sentimen dengan menggunakan Lexicon Based featureslebih rendah daripada sistem analisis sentimen yang tanpa menggunakan Lexicon Based Features. Hal ini dikarenakan terdapat kata-kata bersentimen positif di dalam data uji yang seharusnya bersentimen negatif dan sebaliknya.

Hasil Grafik Pengujian

Untuk hasil akurasi sistem dengan menggunakan Lexicon Based Features menghasilkan accuracysebesar 79%, precision sebesar 65%, recall sebesar 97%, dan f-Measure sebesar 78%. Hal ini menunjukkan bahwa sistem yang menggunakan Lexicon Based Features baik digunakan dalam analisis sentimen, walaupun tingkat akurasinya lebih rendah daripada sistem analisis sentimen yang tidak menggunakan Lexicon Based Features yaitu nilai accuracysebesar 86%, precision sebesar 84%, recall sebesar 76%, dan f-measure sebesar 80%. Namun perbandingan tingkat akurasi tersebut tidak menunjukkan perbedaan yang terlalu jauh atau signifikan. Untuk nilai recall sistem analisis sentimen dengan menggunakan Lexicon Based Features lebih besar yaitu 97% dibandingkan dengan nilai recall sistem analisis sentimen yang tanpa menggunakan Lexicon Based Featuressebesar 76%. Hal ini dikarenakan sistem analisis sentimen dengan menggunakan Lexicon Based Features lebih dapat mengenali sentimen data uji melalui pengecekan kata bersentimen pada saat pembobotan Lexicon Based Features apabila digabungkan dengan sistem analisis sentimen tanpa menggunakan Lexicon

# BAB 3

# PENUTUP

1. Kesimpulan

Berdasarkan hasil pengujian yang telah dilakukan sebelumnya, maka dapat diambil kesimpulan mengenai analisis sentimen terhadap tingkat kepuasan pengguna penyedia layanan telekomunikasi seluler pada Twitter dengan metode Support Vector Machine dan Lexicon Based Features yaitu sebagai berikut.

1. Metode Support Vector Machine dapat diimplementasikan pada analisis sentimen terhadap tingkat kepuasan pengguna penyedia layanan telekomunikasi seluler dengan hasil klasifikasi berupa sentimen positif atau negatif.
2. Untuk penelitian analisis sentimen terhadap tingkat kepuasan pengguna penyedia layanan telekomunikasi seluler pada Twitter dengan metode Support Vector Machine dan Lexicon Based Features, didapatkan nilai-nilai parameter Support Vector Machine yang optimal yaitu nilai degree untuk perhitungan kernel polynomial adalah 2 serta nilai konstanta learning rate dan jumlah iterasi maksimum untuk perhitungan sequential training SVM (support machine)
3. Penambahan Lexicon Based Features memiliki pengaruh untuk analisis sentimen terhadap tingkat kepuasan pengguna penyedia layanan telekomunikasi seluler pada Twitter, namun tingkat akurasinya lebih rendah dibandingkan dengan analisis sentimen yang tidak menggunakan Lexicon Based Features. Hal ini dikarenakan terdapat kalimat sindiran atau muncul kata-kata bersentimen positif pada data uji yang seharusnya bersentimen negatif dan sebaliknya.

# DAFTAR PUSTAKA

* 1. <https://www.researchgate.net/publication/320234928_Analisis_Sentimen_Tingkat_Kepuasan_Pengguna_Penyedia_Layanan_Telekomunikasi_Seluler_Indonesia_Pada_Twitter_dengan_Metode_Support_Vector_Machine_dan_Lexicon_Based_Features?enrichId=rgreq-76130ba60790d20550b42c59c4b1bca0-XXX&enrichSource=Y292ZXJQYWdlOzMyMDIzNDkyODtBUzo1NDYxMjI4ODE2NzExNjhAMTUwNzIxNzI0OTYxMQ%3D%3D&el=1_x_3&_esc=publicationCoverPdf>
  2. <https://www.slideshare.net/ikekurniati/sentiment-analysis-71333347>