

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/270794878>

Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation

Conference Paper · October 2014

CITATIONS

7

READS

6,403

3 authors, including:



Ghulam Asrofi Buntoro

Universitas Muhammadiyah Ponorogo

38 PUBLICATIONS 97 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Solidification Process of Metal [View project](#)

Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation

Ghulam Asrofi Buntoro¹, Teguh Bharata Adji², Adhistya Erna Purnamasari³

Jurusan Teknik Elektro FT UGM

Jln. Grafika 2 Yogyakarta 55281 INDONESIA

¹ghulamasrofibuntoro@gmail.com, ²adji.tba@gmail.com, ³adystya@te.ugm.ac.id

ABSTRACT

Twitter is now becoming a very popular communication media on the Internet, there are a lot of positive comments and negative comments, one way to identify them is by Sentiment Analysis. Today many news are accessed from social media especially Twitter; from Twitter public comments, we can determine the level of public interest sentiment to the problems that exist. With more sentiment parameters that can be obtained, information that is produced will be more accurate and reliably. In this paper will be generated 7 parameters of Sentiment Analysis, which are very positive, positive, somewhat positive, neutral, somewhat negative, negative, very negative. In studies improve by 7 parameters Sentiment Analysis with 23,43% accuracy.

ABSTRACT

Twitter sekarang ini menjadi media komunikasi yang sangat populer di Internet, di dalamnya banyak komentar positif dan negatif, salah satu cara untuk mengidentifikasi komentar tersebut adalah dengan Analisis sentimen. Saat ini banyak berita yang diambil dari media sosial, terutama Twitter, karena dari komentar publik, kita dapat menentukan tingkat sentimen masyarakat terhadap suatu permasalahan. Parameter sentimen yang lebih banyak, akan menghasilkan informasi yang lebih detail, akurat dan terpercaya. Dalam makalah ini dihasilkan 7 parameters analisis sentimen, yaitu sangat positif, positif, agak positif, netral, agak negatif, negatif, sangat negatif. Penelitian ini mampu meningkatkan parameters Sentiment Analysis menjadi 7 parameters dengan akurasi 23,43%.

Keywords—Sentiment Analysis, Lexicon Based, Double Propagation

I. PENDAHULUAN

Perkembangan di bidang Teknologi Informasi Komputer dari masa ke masa sangat pesat dan peranannya dalam kehidupan manusia dapat dirasakan dalam berbagai bidang kegiatan kehidupan manusia, baik secara individu ataupun kelompok (organisasi atau perusahaan).

Secara umum terdapat dua tipe informasi tekstual di *web* yaitu fakta dan opini. Fakta adalah pernyataan objektif mengenai entitas dan kejadian di dunia sedangkan opini adalah pernyataan subjektif yang merefleksikan sentimen atau persepsi orang mengenai entitas ataupun kejadian di dunia. Ketika suatu organisasi/perusahaan/perorangan ingin memperoleh opini publik mengenai produk, citra dan layanannya maka mereka tidak perlu melakukan survei konvensional dan fokus group yang mahal biayanya. *Web* melalui situs *review online*, *blog* pribadi, situs jejaring sosial menyediakan sumber-sumber opini yang besar jumlahnya bagi kebutuhan individu maupun organisasi. Melalui *web* orang dapat mengekspresikan apa saja, termasuk pendapatnya akan suatu hal tanpa adanya keterpaksaan.

Mikroblog seperti Twitter dan Facebook sekarang menjadi perangkat komunikasi yang sangat populer di kalangan pengguna internet. Pada konferensi resmi pengembangan Twitter Chirp 2010, perusahaan tersebut menyampaikan statistik mengenai situs dan pengguna Twitter. Statistik tersebut menyebutkan bahwa pada bulan April 2010, Twitter memiliki 106 juta akun dan sebanyak 180 juta pengunjung unik setiap bulannya. Jumlah pengguna Twitter disebutkan terus meningkat 300.000 *user* setiap harinya (Yarrow, 2010). Digital Buzz blog suatu situs yang menyediakan statistik infografik menyebutkan data statistik yang sama.

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 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 dan aplikasi berbasis analisis sentimen berkembang pesat. Bahkan di Amerika terdapat sekitar 20-30 perusahaan yang memfokuskan pada layanan analisis sentimen [1].

II. PENELITIAN TERKAIT

Metode *Double Propagation* (DP) pertama kali diusulkan oleh Qiu, dkk [2]. Metode ini merupakan

metode *semi unsupervised*. Disebut demikian karena metode ini masih membutuhkan kamus kata yang berisi sedikit kata sifat. Metode ini dapat melengkapi kamus katanya secara otomatis saat proses ekstraksi fitur produk berlangsung.

Caranya adalah dengan menemukan fitur menggunakan kata sifat yang terdapat dalam kamus, kemudian memanfaatkan fitur yang telah terekstrak tadi untuk menemukan kata sifat lain yang terdapat dalam teks opini. Kata sifat baru tersebut secara otomatis akan ditambahkan dalam kamus kata. Proses ini berlangsung terus-menerus hingga tidak ada fitur dan kata sifat baru yang ditemukan. Metode ini juga mampu mengekstrak fitur yang tidak terkomentari secara langsung oleh suatu kata opini. Misalnya dalam kalimat “*Smartphone ini mengagumkan layar dan kameranya*”, metode *Double Propagation* akan mampu mengekstrak dua buah fitur yakni *layar* dan *kamera*. Sementara metode ekstraksi yang menggunakan bantuan kamus kata hanya akan mampu mengekstrak fitur *layar* saja, karena fitur itulah yang berelasi secara langsung dengan kata opini *mengagumkan*. Proses lebih jelas tentang bagaimana cara kerja metode ini dapat dibaca pada paper yang ditulis oleh Qiu. Untuk proses penentuan sentimen kata opini, metode ini dirasa cukup efektif dalam membantu mengenali lebih banyak fitur produk yang terdapat dalam teks opini. Sehingga diharapkan hasil yang lebih baik[3].

Penelitian yang dilakukan oleh Qiu, dkk [2]. Adalah Memperluas *Domain Sentimen Lexicon Based* dengan *Double Propagation* dengan 2 parameters *Sentiment Analysis* yaitu: positif dan negatif. Selanjutnya Penelitian yang dilakukan Azhar, dkk [3]. Yaitu mengenai Otomatisasi Perbandingan Produk Berdasarkan Bobot Fitur pada Teks Opini dengan 3 parameters *Sentiment Analysis* yaitu: positif, netral dan negatif. Kemudian Penelitian Liu dkk [4] membandingkan dua buah produk dengan mengamati tiap fitur dari kedua produk, sistem tersebut dinamakan dengan *opinion observer* dengan 2 parameters *Sentiment Analysis* yaitu: positif dan negatif. Sedangkan penelitian yang dilakukan ini adalah Analisis sentimen *tweet* di media sosial *Twitter* dengan kombinasi *Lexicon Based* dan *Double Propagation* dengan 7 parameters *Sentiment Analysis* yaitu: sangat positif, positif, agak positif, netral, agak negatif, negatif, sangat negatif.

III. METODOLOGI

Metodologi yang digunakan dalam penelitian ini adalah sebagai berikut.

A. Pengumpulan Data

Data minimal 128 *tweet*. Data yang diambil hanya *tweet* dalam berbahasa Indonesia, diambil secara acak baik dari user biasa ataupun media *Twitter* dengan kata kunci jokowi dan prabowo. Metode pencarian dilakukan dengan melakukan filter untuk kata-kata positif dan negatif maupun tanpa menggunakan filter, sehingga kata-kata yang didapat mengandung unsur positif, negatif maupun netral.

Pada tahap *preprocessing*, dilakukan 3 langkah sebagai berikut.

1. Seleksi komentar

Pada tahap ini, dilakukan seleksi komentar yang mengandung kata kunci pemilu, karena *Twitter* terdapat fungsi retweet, yaitu memberikan komentar terhadap *tweet* komentar seseorang, karena komentar *tweet* akan mengganggu dalam proses *Sentiment Analysis tweet*. Jadi dalam *preprocessing* ini komentar *tweet* dihapus.

2. Cleansing

Kalimat yang didapat biasanya masih terdapat noise, yaitu kesalahan acak atau varian dalam variable terukur [5], untuk itu, kita harus menghilangkan noise tersebut. Kata yang dihilangkan adalah karakter HTML, kata kunci, ikon emosi, *hashtag* (#), *username* (@username), *url* (http://website.com), dan *email* (nama@website.com) [5].

3. Parsing

yaitu proses memecah dokumen menjadi sebuah kata dengan melakukan analisa terhadap kumpulan kata dengan memisahkan kata tersebut dan menentukan struktur sintaksis dari tiap kata tersebut.[6]

4. Normalisasi Kalimat

Bertujuan untuk menormalkan kalimat sehingga kalimat gaul menjadi normal [7], sehingga bahasa gaul tersebut dapat dikenali sebagai bahasa yang sesuai dengan KBBI.

Yang harus dilakukan untuk normalisasi kalimat adalah:

- Meregangkan tanda baca (*punctuation*) dan *symbol* selain *alphabet*

Meregangkan tanda baca adalah memberikan jarak terhadap tanda baca dari kata-kata sesudah atau sebelumnya, tujuannya agar tanda baca dan *symbol* selain *alphabet* tidak menjadi satu dengan kata-kata pada saat proses tokenisasi.

- Mengubah menjadi huruf kecil semua
- Normalisasi kata

Aturan-aturan dalam proses normalisasi dapat dilihat pada Tabel 1.

Tabel 1. Aturan normalisasi kata [7]

Tidak Normal / gaul	Normal
Akhiran -ny	Akhiran -nya
Akhiran -nk	Akhiran -ng
Akhiran -x	Akhiran -nya
Akhiran -z	Akhiran -s
Akhiran -dh	Akhiran -t
Kata berulang: sama2	Kata berulang: sama-sama
Ejaan: oe	Huruf: u
Ejaan: dj	Huruf: j

- Menghilangkan huruf yang berulang

Ketika sedang senang atau kesal, seseorang bebas menuliskan opini berdasarkan emosinya, biasanya seseorang menuliskan dengan mengulang huruf yang sama. Contohnya: “kereeen” untuk mengekspresikan kesenangan. Kata berulang seperti “kereeen” akan di normalisasi menjadi “keren”.

- Menghilangkan *emoticon*

Ketika sedang menulis status (*tweet*) seseorang kadang salah atau kurang tepat dalam penggunaan emoticon, entah disengaja atau tidak banyak yang melakukannya. Contohnya: Mereka hanya bisa memfitnah karena tidak bisa ketemu fakta buruk :), kata opini fitnah tapi emoticonnya senyum :), dengan begitu emoticon akan mengganggu dalam proses *Sentiment Analysis tweet*, jadi dalam proses ini emoticon dihapus atau diabaikan. Beberapa *emoticon*, *feeling and sentiment* dapat dilihat pada Gambar 1.

Emoticon	Feeling	Sentiment
:) :-)	Happy	Positive
:(:-(Sad	Negative
:D :-D	Very Happy!	Positive
D: D=	Very Sad	Negative
* * * * *	Fascinated	Positive
D:< D: D8	Horror, disgust, sadness	Negative
xD XD	Laughing, big grin	Positive
: = :-	Straight face no expression	Neutral

Gambar 1 *Emoticon, Feeling and Sentiment*

B. Tokenisasi

Setelah normalisasi kalimat, selanjutnya kalimat tersebut dipecah kedalam token-token menggunakan pembatas / delimiter spasi. Token yang digunakan dalam penelitian ini yaitu [5]:

- *unigram*: token yang terdiri dari hanya satu kata, contohnya: Pemilu.

Berikut ilustrasinya:

Opini : Pemilu tahun ini berlangsung secara jujur dan adil.
Unigram : Pemilu, tahun, ini, berlangsung, secara, jujur, dan, adil.

C. Part of Speech (POS) Tagger

POS tagger adalah sebuah proses untuk memberikan kelas pada sebuah kata. Dalam proses *POS tagger* dilakukan dengan cara *parsing*, kemudian ditentukan kelas tiap kata dengan menggunakan bantuan kamus yang di buat sendiri berdasarkan Kamus Besar Bahasa Indonesia (KBBI) menggunakan metode *Maximum Entropy*. Proses *POS tagging* terbagi ke dalam tiga proses yaitu pemisahan setiap token dalam dokumen dengan pengecekan setiap kata dalam dokumen, mengidentifikasi setiap kata dalam

dokumen dengan pemberian jenis kata, pengecekan kata yang belum teridentifikasi terhadap bentuk imbuhan dan akhiran sehingga diperoleh kata dasar.

Berdasarkan aturan linguistik pada kata diperoleh sentimen sementara. Penentuan sentimen dilakukan dengan melihat adanya kata yang mengandung opini baik yang memiliki *polarity* positif maupun negatif dari *tweet* yang sudah dilabeli kelas katanya. Kelas kata yang dipilih adalah kata sifat (*adjective*), kata keterangan (*adverb*), kata benda (*noun*) dan kata kerja (*verb*), sesuai dengan penelitian [3] bahwa keempat jenis kata di atas merupakan jenis kata yang paling banyak mengandung sentimen. Dalam sistem ini jika suatu *tweet* terdapat kata benda (NN) pada sebelum atau setelah kata sifat (JJ) atau kata keterangan (RB) dan kata benda (memiliki *polarity* berlawanan dengan kata sifat atau kata keterangan maka *polarity* yang diperoleh berdasarkan kata sifat atau kata keterangan, karena kata sifat atau kata keterangan memberikan penegasan terhadap kata benda.[6]

D. Penentuan Class Attribute

Data Twitter yang sudah dilakukan Preprocessing kemudian akan ditentukan class attribute, class attribute yang dimunculkan dalam penelitian ini ada 7, diantaranya sangat positif, positif, agak positif, netral, agak negatif, negatif dan sangat negatif. Dengan 7 class attribute ini diharapkan mampu memberi penilaian masyarakat secara akurat terhadap objek tertentu.

E. Load Dictionary

Setelah dilakukan tokenisasi dan ditentukan class attribute, langkah selanjutnya adalah load dictionary. Banyak jenis kamus yang dapat digunakan, contohnya: kamus kata kunci sentimen positif (positif keywords), kamus kata kunci sentimen negatif (negatif keywords), kamus kata negasi (negation keywords), dan kamus normalisasi bahasa gaul/alay.

Berikut adalah contoh kamus dan isinya [5]:

- Positif keywords: baik, hebat, jujur, cerdas, keren.
- Negatif keywords: bohong, korupsi, jahat, jelek.
- Negation keywords: nggak, tidak, bukan, jauh.
- Kamus konversi bahasa gaul ke KBBA: sp = siapa, spt = seperti, brp = berapa, hrg = harga, ciyus = serius.

F. Extract dan Pembobotan Fitur

Untuk mengekstrak fitur produk dari teks opini yang terdapat pada dataset, digunakan metode *Double Propagation* (DP)[2]. Seperti dijelaskan pada bagian sebelumnya, metode ini tidak membutuhkan kamus kata opini yang lengkap untuk dapat bekerja. Karena dalam prosesnya, metode ini tidak hanya akan mengekstrak kata opini melainkan juga mengekstrak kata sifat yang dicurigai sebagai kata opini.

Hasil akhir dari proses ini adalah daftar kata opini beserta bobot untuk masing-masing kata kunci (fitur). Bobot untuk masing-masing kata kunci (fitur) didapatkan melalui Persamaan 1.

$$W_f = \sum_{i=1}^n p_i \quad (1)$$

dimana W_f adalah bobot dari kata kunci fitur f yang didapatkan melalui penjumlahan n skor polaritas kata opini p yang mengomentari fitur f . Skor polaritas suatu kata opini p akan bernilai 1 jika kata tersebut adalah kata opini positif, dan bernilai -1 jika kata tersebut adalah kata opini negatif[3].

Kata kunci diekstrak dari kalimat menggunakan token yang dicocokkan dengan kamus untuk mendapatkan kata kunci positif dan negatif. Prosesnya bisa dilihat sebagai berikut.

- Ekstraksi kata kunci positif dan kata kunci negatif
Token-token dicocokkan dengan kata kunci yang ada dalam kamus yang berisikan kata kunci positif ataupun kata kunci negatif.
- Ekstraksi kata sifat yang dicurigai sebagai kata opini
Kata sifat yang dicurigai sebagai kata opini diekstrak menjadi kata opini kemudian ditambahkan dalam kamus opini. Contohnya: “banyak” adalah kata sifat, akan tetapi juga sebagai kata opini dan terdapat dalam kamus positif maka akan menjadi kata opini positif.
- Evaluasi negasi
Dengan menggunakan kata negasi, kata yang dideteksi positif dapat berubah menjadi negatif maupun sebaliknya. Contohnya “korupsi” adalah kata kunci negatif, akan tetapi jika sebelumnya terdapat kata “tidak” dan menghasilkan token “tidak korupsi” maka nilainya akan berubah menjadi positif. Kata-kata negasi tersebut, dapat dilihat dari kamus yang berisikan kata negasi.

G. Determine Sentiment

Proses yang terakhir adalah menentukan sentimen suatu kalimat opini, penentuan dilakukan dengan penjumlahan n skor polaritas kata opini p yang mengomentari fitur f . Skor polaritas suatu kata opini p akan bernilai 1 jika kata tersebut adalah kata opini positif, dan bernilai -1 jika kata tersebut adalah kata opini negatif[4].

Setelah diketahui kata yang mengandung positif, negatif dan netral di dalam sebuah kalimat, selanjutnya dihitung bobot nilai yang terkandung dalam kalimat tersebut yang dilakukan dengan menjumlahkan nilai kata opini. Jika jumlah nilai opini positif dalam kalimat tersebut ≥ 3 , maka nilai sentimen dari kalimat tersebut adalah sangat positif, apabila nilai opini positif dalam kalimat tersebut = 2, maka nilai sentimen dari kalimat tersebut adalah positif, jika nilai opini positif dalam kalimat tersebut = 1, maka nilai sentimen dari kalimat tersebut adalah agak positif, jika nilai opini dalam kalimat tersebut = 0, maka nilai sentimen dari kalimat tersebut adalah

netral, jika nilai opini negatif dalam kalimat tersebut = -1, maka nilai sentimen dari kalimat tersebut adalah agak negatif, jika nilai opini negatif dalam kalimat tersebut = -2, maka nilai sentimen dari kalimat tersebut adalah negatif, dan jika nilai opini negatif dalam kalimat tersebut = -3, maka nilai sentimen dari kalimat tersebut adalah sangat negatif. Determine Sentimen dapat dilihat pada Tabel 2.

Tabel 2 Determine Sentimennya

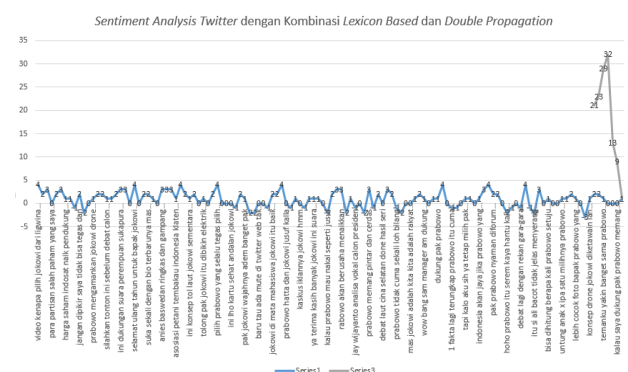
Sentimen	Nilai
Sangat Positif	≥ 3
Positif	2
Agak Positif	1
Netral	0
Agak negatif	-1
Negatif	-2
Sangat Negatif	≤ -3

IV. UJI COBA DAN PEMBAHASAN

Ujicoba dilakukan dengan data yang berjumlah 128 *tweet*. Proses *preprocessing* dilakukan dengan seleksi komentar, *cleansing*, *parsing*, normalisasi kalimat, *tokenisasi*, dan konversi bahasa gaul ke KBBA. selanjutnya menghapus kata yang berulang dari *tweet* tersebut, terakhir menghapus *emoticon* yang ada dalam *tweet*. Selanjutnya data mengalami proses *POS tagger*, dari semua proses di atas menghasilkan data pelatihan untuk membangun model klasifikasi.

Skenario ujicoba untuk melihat seberapa baik sistem yang telah dibuat, dengan mengambil data dari *tweet* dengan kata kunci Jokowi dan Prabowo sebanyak 128 *tweet*, kemudian dilakukan proses *preprocessing* dan *POS Tagging* selanjutnya sistem melakukan proses klasifikasi dengan menghitung nilai sentimen tiap *tweet* dan mencocokkannya dengan 7 parameters sentimen yaitu sangat positif, positif, agak positif, netral, agak negatif, negatif, dan sangat negatif, dengan hasilnya adalah nilai sentimen tiap data *tweetnya*.

Hasil ujicoba *Sentiment Analysis tweet* dengan 7 parameters sentimen bisa dilihat pada Gambar 2.



Gambar 2. Hasil ujicoba *Sentiment Analysis* dengan 7 parameters

Kombinasi *Lexicon based* dan *double propagation* mampu menghasilkan 7 parameters *Sentiment Analysis* untuk lebih jelasnya bisa dilihat hasil dari metode ini pada Tabel 3.

Tabel 3 Hasil ujicoba *Sentiment Analysis*

Sentiment	Jumlah	Nilai
Sangat positif	21	67
Positif	23	46
Agak positif	29	29
Netral	32	0
Agak negatif	13	-13
Negatif	9	-18
Sangat negatif	1	-3
Jumlah	128	112

Untuk mengetahui akurasi sistem yang telah dibuat yaitu dengan membandingkan hasil sistem dengan hasil dari 3 mahasiswa yang telah memberikan penilaian *sentiment analysis* terhadap data *tweet*, ada berapa banyak nilai sentimen yang sama dari data *tweet*, kemudian dihitung prosentase akurasi dengan Persamaan 2.

$$\text{Akurasi} = \frac{t}{n} \times 100 \quad (2)$$

dimana t adalah jumlah nilai sentimen yang sama pada data *tweet* antara sistem dengan 3 mahasiswa dan n adalah jumlah seluruh data *tweet*.

Hasil nilai sentiment dari 3 mahasiswa dapat dilihat pada Tabel 4.

Tabel 4 Perhitungan sentiment 3 mahasiswa

Mahasiswa	Persamaan Sentimen	Prosentase
1	26	20,31
2	40	31,25
3	24	18,75

Setelah didapatkan nilai sentimen dari tiap mahasiswa, kemudian dihitung dengan Persamaan 2 seperti dibawah ini:

$$\text{Akurasi} : \frac{26 + 40 + 24}{3} = \frac{90}{3} \times 100 = 30,00\%$$

Hasil yang didapatkan dari perhitungan di atas adalah akurasi dari sistem ini yaitu 30,00 %.

V. KESIMPULAN

Metode *Double Propagation* (DP) mampu untuk mengekstrak dan memberi bobot nilai sentimen sebuah kata opini, juga mampu mengenali kata sifat yang termasuk kata opini. Kombinasi Metode *Lexicon based* dan Metode *Double Propagation* (DP) mampu untuk meningkatkan parameters sentimen menjadi 7 parameters sentimen yaitu sangat positif, positif, agak positif, netral, agak negatif, negatif, dan sangat negatif. Ujicoba sistem yang telah

dibuat dilakukan dengan menggunakan 128 data *tweet* dengan 3 mahasiswa sebagai perbandingan akurasi, hasil akurasi mencapai 30,00 %. Selanjutnya metode ini masih bisa ditingkatkan nilai akurasi dan digunakan untuk landasan penelitian lain yang memanfaatkan opini publik atau *Sentiment Analysis* pada bidang lain.

REFERENCES

- [1] Go, A., Huang, L., & Bhayani, R. (2009). Twitter Sentiment Analysis. Final Project Report, Stanford University, Department of Computer Science.
- [2] Qiu, Guang., Bing, Liu., Jiajun Bu and Chun Chen. 2009. "Expanding Domain Sentimen Lexicon through Double Propagation". In Proceedings of IJCAI
- [3] Yufis Azhar, Agus Zainal Arifin, Diana Purwitasari. 2013. "Otomatisasi Perbandingan Produk Berdasarkan Bobot Fitur pada Teks Opini".
- [4] Liu, B., Hu, M., dan Cheng, J. (2005). Opinion observer: analyzing and comparing opinions on the Web. In Proceedings of the 14th international conference on World Wide Web. ACM, hal. 342-351
- [5] J. Ariawan, "Data Preprocessing." [Online]. Available: <https://www.google.com/search?q=apa+itu+noise+dalam+data+mining&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a&channel=sb>. [Accessed: 10-Mar-2014].
- [6] Noviah Dwi Putranti, Edi Winarko (2014). "Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan *Maximum Entropy* dan *Support Vector Machine*"
- [7] M. Yusuf Nur Sumarno Putro. 2011 "Analisis Sentimen pada Dokumen berbahasa Indonesia dengan Pendekatan Support Vector Machine". Masters, Binus.
- [8] N. Adiyasa, "Analisis Sentimen Pada Opini Berbahasa Indonesia Menggunakan Pendekatan Lexicon-Based," Catatan Kecil, 2011. [Online]. Available: <http://adiyasan.wordpress.com/2013/02/08/sentimen-analysis-menggunakan-pendekatan-lexicon-based/>. [Accessed: 10-Mar-2014].