

Analisis Sentimen Pada Review Pengguna Aplikasi *Mobile* Untuk Evaluasi Faktor *Usability*

Septiyawan Rosetya Wardhana, Diana Purwitasari, Siti Rochimah

Jurusan Teknik Informatika dan Institut Teknologi Sepuluh Nopember

Kampus ITS, Jl. Raya ITS, Sukolilo, Surabaya 60111,

Telp. + 62 31 5939214, Fax. + 62 31 5939363

e-mail: rossywardhana@gmail.com¹⁾, siti@if.its.ac.id²⁾, diana.purwitasari@gmail.com³⁾

Abstrak

Faktor *usability* merupakan aspek yang paling diperhatikan dalam pembuatan maupun pengujian aplikasi *mobile*. Evaluasi *usability* dapat dilakukan dengan melakukan analisis sentimen pada review pengguna aplikasi *mobile*. Orientasi sentimen inilah yang umumnya dijadikan sebagai acuan dalam proses evaluasi. Selain itu, setiap review pasti memiliki tingkat sentimen yang mencerminkan tinggi rendahnya orientasi sentimen, sehingga akan lebih efektif apabila tingkat sentimen juga dipertimbangkan dalam proses evaluasi. Berbeda dengan jenis perangkat lunak lainnya, aplikasi *mobile* memiliki batasan dan permasalahan sendiri yang tidak dimiliki perangkat lunak lain. Model PACMAD (*People At The Centre of Mobile Application Development*) merupakan model *usability* yang karakteristiknya disesuaikan dengan batasan dan permasalahan yang dimiliki oleh aplikasi *mobile*. Oleh karena itu, dalam penelitian ini diusulkan suatu metode analisis sentimen dengan mempertimbangkan tingkat sentimen pada opini pengguna aplikasi *mobile* untuk evaluasi faktor *usability* berdasarkan model PACMAD. Data review pengguna akan diklasifikasikan ke dalam faktor *usability* PACMAD dengan menggunakan metode TF.ICF, kemudian dianalisis orientasi dan tingkat sentimennya dengan menggunakan metode SentiWordNet Interpretation. Berdasarkan hasil uji coba menggunakan data review dari 5 aplikasi *mobile* diperoleh rata-rata nilai akurasi klasifikasi faktor *usability* sebesar 82% dan akurasi sentimen sebesar 79%.

Kata kunci : analisis sentimen, evaluasi faktor *usability*, model PACMAD, review pengguna.

Abstract

Usability factors are the most noted aspects in *mobile application development or evaluation process*. *Usability factors evaluation process* can be done by doing *sentiment analysis* in *user review* at that *mobile application*. This *sentiment orientation* which is generally used as *reference* in *evaluation process*. Moreover, every review must have *sentiment level* that reflect *high and low sentiment orientation*, so that it will be more effective if it consider *sentiment level* in *evaluation process*. Different with other software type, *mobile application* has some limitations and issues itself which another software doesn't have. PACMAD (*People At the Centre of Mobile Application Development*) model is *usability model* which its characteristic is adapted with *mobile application limitation and issues*. Therefore, this research propose *sentiment analysis method* which consider *sentiment level* in *mobile application review* to evaluate *usability factor* based on PACMAD model. In this research, *user review* will be classified into PACMAD *usability factors* using TF.ICF method. Then, that review will be analyzed based on *sentiment orientation and level* using SentiWordNet Interpretation method. Based on test results using review data from 5 *mobile application*, it is gained 82% average of *usability factors evaluation accuracy* and 79% *sentiment accuracy*.

Keywords: *sentiment analysis, usability factors evaluation, PACMAD model, user review.*

1. Pendahuluan

Aplikasi *mobile* atau *apps* merupakan perangkat lunak yang khusus dirancang untuk bisa dijalankan pada perangkat *mobile* seperti *smartphone, tablet* maupun *smartwatch*. Dalam proses pembuatan maupun pengujian aplikasi *mobile*, faktor *usability* aplikasi menjadi aspek yang paling diperhatikan. *Usability* merupakan suatu atribut kualitas yang mencerminkan bagaimana tingkat kemudahan suatu perangkat lunak ketika digunakan oleh penggunaannya. Pengujian *usability* secara umum biasanya melibatkan evaluasi berbagai macam aplikasi *mobile* berdasarkan pertimbangan faktor *usability*

dari aplikasi yang digunakan. Bagaimanapun juga dalam evaluasi *usability* pasti melibatkan pengguna sebagai acuannya. Salah satu bentuk komentar pengguna terhadap aplikasi yang digunakan adalah review pengguna terhadap aplikasi tersebut. Opini pengguna merupakan bagian dari teks yang mengandung nilai informasi tentang pengalaman pengguna dalam menggunakan perangkat lunak [1]. Opini pengguna mengandung suatu komentar maupun persepsi pengguna yang dapat dijadikan salah satu aspek untuk mengukur kualitas suatu perangkat lunak. Permasalahan untuk mengevaluasi faktor *usability* pada aplikasi *mobile* memang merupakan permasalahan yang cukup kompleks.

Berkaitan dengan faktor *usability*, El Hales dalam penelitiannya pernah menggunakan *effectiveness*, *efficiency* dan *satisfaction* sebagai faktor untuk mengevaluasi *usability* perangkat lunak dengan menggunakan *opinion mining* atau penambangan opini [2]. Faktor *usability* yang digunakan meliputi *effectiveness*, *efficiency* dan *satisfaction*. Kemudian dengan menggunakan deteksi orientasi sentimen pada metode *opinion mining*, review pengguna akan diklasifikasikan ke dalam kategori positif atau negatif. Hasil dari penelitian tersebut adalah rata-rata jumlah review positif dan negatif yang telah dikelompokkan berdasarkan faktor *usability*. Proses deteksi sentimen dalam penelitian tersebut belum mempertimbangkan tingkat sentimen serta nilai atau arti kata pada kalimat terhadap review pengguna. Selain itu, faktor *usability* yang digunakan juga belum memperhatikan berbagai faktor dan batasan yang ada dalam aplikasi *mobile* secara keseluruhan.

Berbeda dengan aplikasi *desktop*, aplikasi *mobile* memiliki arsitektur dan model yang berbeda. Aplikasi *mobile* memiliki batasan lain yang berbeda dengan jenis perangkat lunak lain seperti konteks *mobile*, koneksi, layar kecil, resolusi yang berbeda, kemampuan daya yang terbatas serta metode entri data yang berbeda. Sedangkan beberapa faktor yang harus dipertimbangkan dalam evaluasi *usability* adalah pengguna, tujuan dan konteks penggunaan. Berdasarkan batasan dan faktor yang harus dipertimbangkan tersebut maka Harrison dkk mengusulkan model *usability* PACMAD sebagai *guideline* dalam mendesain maupun mengevaluasi aplikasi *mobile* yang terdiri dari *effectiveness*, *efficiency*, *satisfaction*, *learnability*, *memorability*, *errors* dan *cognitive load* [3]. Berdasarkan permasalahan tersebut, maka dalam penelitian ini diusulkan suatu metode analisis sentimen review pengguna aplikasi *mobile* pada evaluasi faktor kebergunaan PACMAD dengan mempertimbangkan tingkat sentimen review. Penambahan tingkat sentimen dan penggunaan model PACMAD pada evaluasi *usability* aplikasi *mobile* ini diharapkan dapat menghasilkan hasil evaluasi *usability* yang lebih akurat dan sesuai dengan konteks aplikasi *mobile* yang ada saat ini.

2. Tinjauan Pustaka/ State of the Art

Bagian ini memuat berbagai literatur berupa penelitian-penelitian sebelumnya yang terkait dengan penelitian yang sedang dibahas. Pada bagian ini juga diharapkan dapat menunjukkan keterhubungan hingga perbedaan mendasar antara penelitian yang sedang dilakukan dengan penelitian-penelitian sebelumnya.

Usability merupakan sesuatu yang kritis dalam sebuah sistem atau aplikasi. Menurut Alaa, *usability* merupakan faktor penting dalam suatu sistem [2]. *Usability* merupakan suatu ukuran atau tingkat dimana suatu produk dapat digunakan oleh pengguna untuk mencapai tujuan pembuatan produk yang sebenarnya dengan efektif, efisien dan penuh kepuasan dari segi konteks penggunaan. Evaluasi *usability* menjadi salah satu hal yang cukup penting dalam pembuatan dan pengembangan perangkat lunak. Tujuan dari evaluasi ini adalah untuk mendapatkan timbal balik dari pengguna mengenai pengalamannya dalam menggunakan perangkat lunak yang bersangkutan [4].

Berbeda dengan aplikasi *desktop* pada umumnya, aplikasi *mobile* memiliki arsitektur dan kebutuhan yang berbeda [5]. Hal tersebut dikarenakan aplikasi *mobile* memiliki batasan-batasan khusus yang tidak bisa disamakan dengan aplikasi *desktop*. Beberapa batasan aplikasi *mobile* yang diajukan oleh Zhang dan Adipat antara lain adalah konteks *mobile*, koneksi, ukuran layar, tampilan resolusi, sumber daya yang terbatas dan metode entri data yang berbeda [6].

Harrison dkk memperkenalkan model *usability* baru yang mempertimbangkan isu-isu perangkat *mobile* [3]. Model *usability* tersebut adalah PACMAD (*People At the Centre of Mobile Application Development*). Dalam PACMAD terdapat 7 atribut yang merupakan penggabungan atribut *usability* ISO dengan Nielsen serta menambahkan atribut *cognitive load* didalamnya. Atribut-atribut *usability* tersebut antara lain adalah *effectiveness*, *efficiency*, *satisfaction*, *learnability*, *memorability*, *errors* dan *cognitive load*. Penjelasan mengenai masing-masing atribut dapat dilihat pada Tabel 1.

Tabel 1. Model *Usability* PACMAD

No	Atribut	Penjelasan
1	<i>Effectiveness</i>	Kemampuan pengguna untuk melengkapi tujuan mereka dalam konteks yang spesifik

No	Atribut	Penjelasan
2	<i>Efficiency</i>	Kemampuan pengguna untuk melengkapi tujuan berdasarkan kecepatan dan akurasi
3	<i>Satisfaction</i>	Tingkat kesenangan dan kepuasan yang diusahakan untuk pengguna melalui penggunaan aplikasi
4	<i>Learnability</i>	Kemudahan pengguna dalam memperoleh keahlian dalam menggunakan aplikasi <i>mobile</i>
5	<i>Memorability</i>	Kemampuan pengguna untuk menguasai penggunaan aplikasi secara efektif
6	<i>Errors</i>	Kesalahan yang dibuat oleh sistem
7	<i>Cognitive Load</i>	Beban kognitif yang diterima pengguna saat menggunakan aplikasi <i>mobile</i>

2.1. Data

Data yang akan digunakan dalam penelitian ini adalah data review pengguna aplikasi *mobile* berbahasa Inggris yang diperoleh dari Google *Playstore*. Google *Playstore* dipilih karena merupakan pusat aplikasi *mobile* terbesar di dunia dan digunakan oleh kebanyakan orang khususnya pengguna *smartphone* android. Data berupa dokumen review pengguna diperoleh menggunakan metode *crawling* dengan memanfaatkan *HTML Parser* pada website atau situs Google *Playstore* versi Bahasa Inggris. Dari proses *crawling* tersebut maka dapat diperoleh berbagai informasi mengenai opini pengguna terhadap aplikasi *mobile* seperti nama *reviewer*, tanggal review, *rating* serta konten review yang memuat opini atau pendapat pengguna terhadap aplikasi *mobile*.

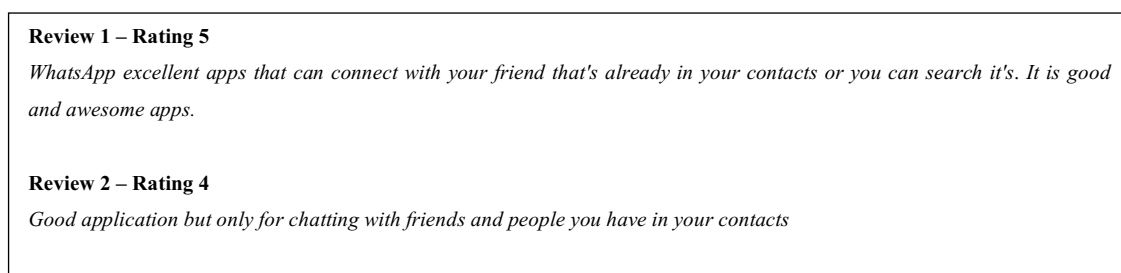
Pada tahap penelitian, jumlah aplikasi *mobile* yang akan digunakan sebagai data dan diambil reviewnya adalah 5 aplikasi yaitu *Whatsapp Messenger*, *Google Street View*, *Instagram*, *Waze* dan *Line*. Jumlah data yang diambil pada tahap uji coba adalah 620 data review, dengan rincian yaitu 520 untuk data *training* dan 100 untuk data *testing*. Tabel 2 dibawah ini menunjukkan rincian pengambilan data pada penelitian.

Tabel 2. Rincian Aplikasi *Mobile* dan URL Pengambilan Data Review Pengguna

No	Aplikasi <i>Mobile</i>	URL
1	<i>Whatsapp Messenger</i>	https://play.google.com/store/apps/details?id=com.whatsapp&hl=en
2	<i>Google Street View</i>	https://play.google.com/store/apps/details?id=com.google.android.street&hl=en
3	<i>Instagram</i>	https://play.google.com/store/apps/details?id=com.instagram.android&hl=en
4	<i>Waze</i>	https://play.google.com/store/apps/details?id=jp.naver.line.android&hl=en
5	<i>Line</i>	https://play.google.com/store/apps/details?id=jp.naver.line.android&hl=en

Kemudian data review yang diperoleh tersebut akan dibagi menjadi 2 yaitu *data training* dan *data testing*. Data review pengguna yang digunakan pada setiap aplikasi kemudian akan diberi label faktor *usability* oleh para ahli. Data akan dianggap valid dan diterima sebagai data *training* apabila keseluruhan ahli yang melabeli opini memiliki kesepakatan yang sama mengenai faktor *usability* yang dilabelkan.

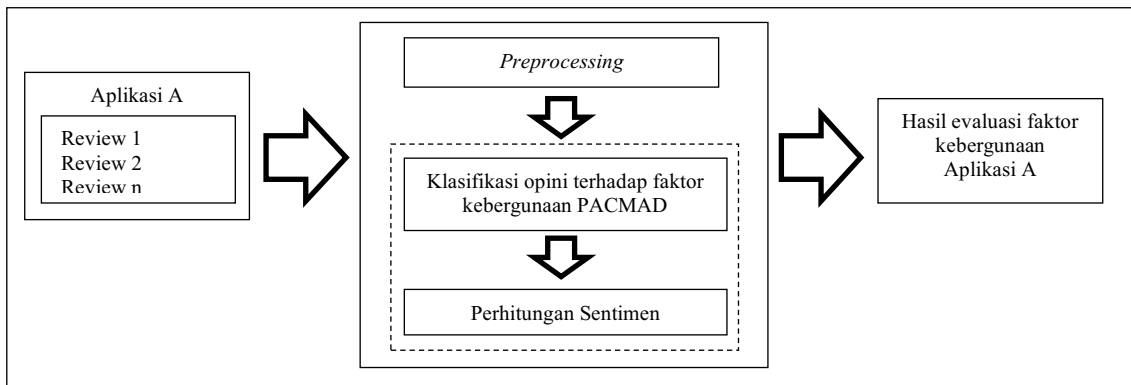
Faktor *usability* dari hasil pelabelan para ahli tersebut akan menjadi *groundtruth* atau data kebenaran untuk proses pengujian faktor *usability* pada data uji. Sedangkan untuk pengujian sentimen dan opini pengguna akan menggunakan *rating* pada data opini pengguna sebagai data kebenarannya, dimana *rating* 1 dan 2 akan bernilai negatif, *rating* 3 bernilai netral dan *rating* 4 dan 5 akan bernilai positif. Untuk contoh data review pengguna aplikasi *mobile* ditunjukkan pada Gambar 1.



Gambar 1. Contoh data review pengguna.

3. Metode Penelitian

Pada penelitian yang diusulkan mengenai evaluasi *usability* aplikasi *mobile* ini, data masukan yang digunakan adalah data review aplikasi *mobile*. Beberapa proses yang dilakukan dalam penelitian ini antara lain adalah *preprocessing*, klasifikasi faktor *usability*, dan perhitungan sentimen. Sedangkan keluaran yang dihasilkan dalam penelitian ini adalah nilai hasil evaluasi dari setiap faktor *usability* yang diujikan. Gambar 2 dibawah ini merupakan gambaran alur kerja metode yang diusulkan secara umum.

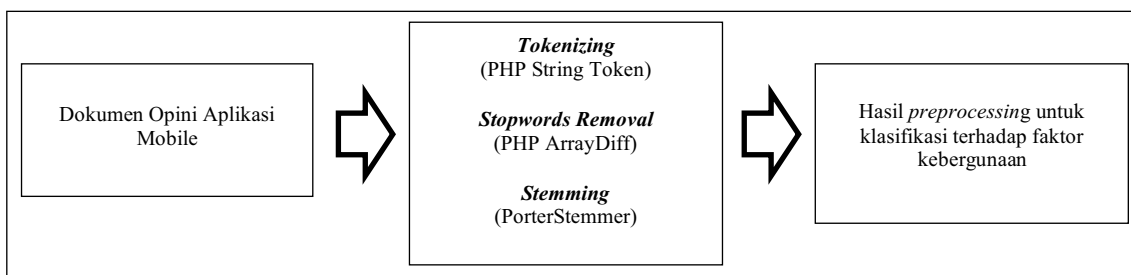


Gambar 2. Contoh data review pengguna.

Pada Gambar 2. dapat dilihat bahwa untuk mendapatkan keluaran berupa hasil evaluasi *usability* aplikasi *mobile* terdapat 3 tahapan yang harus dilakukan yaitu *preprocessing*, klasifikasi faktor *usability*, dan perhitungan sentimen. Setiap proses memiliki fungsi tersendiri yang memanfaatkan *output* dari proses ataupun tahapan sebelumnya.

3.1. Preprocessing

Tahapan awal yang harus dilakukan dalam penelitian ini adalah *preprocessing*. Gambar 3. dibawah ini menunjukkan alur dalam *preprocessing*.



Gambar 3. Alur *preprocessing*

Tahapan *preprocessing* diatas merupakan tahapan *preprocessing* yang digunakan untuk mempersiapkan data yang nantinya digunakan untuk proses klasifikasi opini terhadap faktor *usability*. Pada *preprocessing* ini terdapat beberapa sub proses yang antara lain adalah *tokenizing*, *stopwords removal*, dan *stemming*. *Tokenizing* merupakan proses memecah dokumen maupun kalimat menjadi kata atau *term* [7]. Dalam proses ini, spasi dan tanda baca lainnya akan dihilangkan sehingga menghasilkan output berupa daftar kata atau *term* dalam bentuk *list* atau *array*. Proses *tokenizing* dalam penelitian menggunakan fungsi PHP String Token atau *strtok()* yang telah disediakan oleh *library* PHP.

Stopwords removal dilakukan untuk menghilangkan kata-kata yang sering muncul dan tidak diperlukan seperti *and*, *or*, *in*, *at* dan *stopword* lainnya. Untuk proses menghilangkan *stopword* pada penelitian ini menggunakan fungsi PHP Array Different atau *arraydiff()* yang telah disediakan oleh *library* array pada PHP. Sedangkan dalam proses *stemming*, kata atau *term* yang dihasilkan dari proses *stopwords removal* akan diubah menjadi *root* kata atau kata dasar. Untuk proses *stemming* sendiri bisa memanfaatkan *library* yaitu dari *PorterStemmer*.

3.2. Klasifikasi Faktor Usability

Metode yang digunakan untuk melakukan klasifikasi faktor *usability* dalam penelitian ini adalah metode TF.ICF (Term Frequency, Inverse Class Frequency). TF.ICF merupakan metode yang mengacu pada pembobotan kata terhadap class atau kategori tertentu [8]. Dimana untuk formula TF.ICF dapat dilihat pada persamaan 1 dan persamaan 2.

$$W_{TF}(t_i, c_j) = f(t_i, c_j) \tag{1}$$

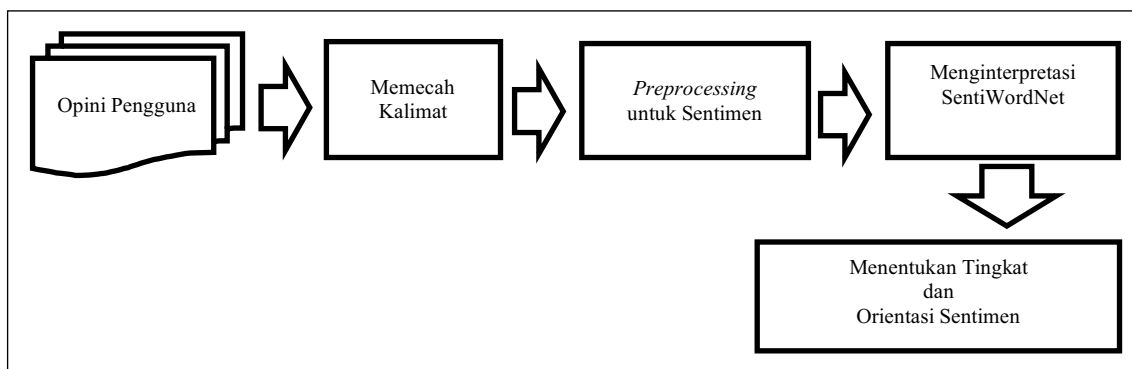
Dari persamaan 1. tersebut maka dapat disimpulkan bahwa nilai TF term i pada $class$ j merupakan jumlah kemunculan kata i pada kategori atau $class$ j . Sedangkan ICF digunakan untuk menghitung seberapa signifikan pengaruh kata atau term dari suatu $class$ terhadap $class$ atau kategori lainnya. Untuk formula ICF dapat dilihat pada persamaan 2.

$$W_{ICF}(t_i, c_j) = 1 + \log \frac{c}{c(t_i)} \quad (2)$$

Persamaan 2 diatas menunjukkan bahwa perhitungan ICF dilakukan dengan menghitung log dari jumlah keseluruhan $class$ (c) dibagi dengan jumlah $class$ yang mengandung $term$ i (t_i). Hasil dari TF kemudian dikalikan dengan ICF sehingga menjadi TF.ICF.

3.2. Perhitungan Sentimen

Perhitungan sentimen merupakan suatu proses yang digunakan untuk menghitung tingkat sentimen dan menentukan apakah suatu dokumen review pengguna bersifat positif, negatif, atau netral. Proses ini menjadi sangat penting, dikarenakan untuk mengetahui hasil evaluasi *usability* pada suatu opini pengguna maka sistem harus bisa memutuskan apakah review tersebut bersifat positif, negatif atau bahkan netral terlebih dahulu. Tahapan perhitungan sentimen pada penelitian ini dapat dilihat pada Gambar 4 dibawah ini.



Gambar 4. Tahapan perhitungan sentimen.

Pada Gambar 4 diatas maka dapat disimpulkan bahwa tahapan-tahapan yang dilakukan untuk menghitung sentimen review pengguna antara lain adalah memecah kalimat, *preprocessing* untuk sentimnen, menginterpretasi SentiWordNet, menentukan tingkat dan orientasi sentimen. Pada tahap memecah kalimat, review pengguna yang berupa dokumen akan dipecah menjadi kumpulan kalimat. Selanjutnya, setiap kalimat pada dokumen opini akan dimasukkan dalam *preprocessing* untuk sentimen meliputi *tokenizing*, *stopwords removal* dan POS Tagging yang telah dijelaskan pada proses sebelumnya. Setelah melalui *preprocessing*, opini pengguna akan diolah kembali pada proses interpretasi SentiWordNet. Dalam interpretasi SentiWordNet, review pengguna akan dicocokkan dengan SentiWordNet untuk mencari nilai sentimen yang sesuai dengan kata yang dicocokkan. Tujuan dari tahapan ini adalah untuk mencari nilai sentimen yang mewakili setiap kata yang ada pada opini pengguna sesuai dengan arti dan kategori kata tersebut.

Proses interpretasi SentiWordNet dilakukan dengan menghitung rata-rata nilai positif dan negatif dari setiap kata dalam opini pengguna yang diujikan berdasarkan pengelompokkan kategori kata atau tag yang meliputi kata benda, kata sifat, kata kerja maupun keterangan pada SentiWordNet. Jadi hasil dari tahap ini adalah interpretasi nilai positif dan negatif dari kata berdasarkan kelompok tag nya. Sebagai contoh, kata “good” dalam SentiWordNet memiliki beberapa *sense* yang berbeda meliputi 4 *sense* dalam kata benda, 21 *sense* dalam kata sifat dan 2 *sense* dalam keterangan. Sehingga untuk perhitungan interpretasi SentiWordNetnya dilakukan dengan menghitung rata-rata nilai positif dan negatif dari kata “good” pada setiap kategori kata. Hasil dari perhitungan tersebut dapat dilihat pada Tabel 3.

Tabel 3. Perhitungan *Sense* Kata

No	Kata	Tag	Rata-rata Positif	Rata-rata Negatif
1	Good	Noun	0.531	0
2	Good	Adjective	0.5	0

3	Good	Adverb	0.188	0
---	------	--------	-------	---

Tahap yang dilakukan selanjutnya adalah menentukan orientasi sentimen. Pada tahap ini akan dihitung nilai polaritas dari review pengguna yang nantinya digunakan sebagai acuan untuk menentukan tingkat dan orientasi sentimen dari review pengguna tersebut. Tujuan tahap ini adalah menghitung tingkat sentimen dan menentukan apakah review pengguna bersifat positif, negatif atau netral. Metode yang digunakan pada tahap ini adalah metode rata-rata kalimat dan rata-rata review. Dimana metode ini mempertimbangkan nilai atau arti kata pada suatu kalimat serta nilai positif dan negatif dari kata yang memiliki pengaruh terhadap tingkat sentimen kata.

Review pengguna yang sudah dipecah menjadi kalimat dan melalui tahap *preprocessing* akan dicari nilai interpretasi positif dan negatifnya menggunakan metode *SentiWordNet interpretation* yang telah dijelaskan sebelumnya. Setelah itu, rata-rata nilai positif dan nilai negatif akan dihitung berdasarkan jumlah interpretasi kata positif dan negatif yang ada pada kalimat tersebut. Kemudian, setelah semua rata-rata nilai positif dan negatif pada setiap kalimat diperoleh maka akan dihitung rata-rata nilai positif dan negatif pada review. Rata-rata nilai positif dan negatif pada review inilah yang dijadikan sebagai tingkat sentimen, dimana tingkat sentimen yang paling tinggi dalam suatu review akan dijadikan sebagai orientasi sentimen. Apabila tingkat sentimen positif lebih tinggi daripada tingkat sentimen negatif maka review akan memiliki orientasi sentimen positif, dan apabila tingkat sentimen negatif lebih tinggi maka review akan berorientasi negatif. Namun apabila tingkat sentimen antara positif dan negatif memiliki nilai yang sama maka review akan dianggap objektif atau netral.

4. Hasil dan Pembahasan

Dalam penelitian yang mengutamakan proses klasifikasi dalam penerapan metodenya, maka pengukuran yang cocok digunakan adalah akurasi. Akurasi merupakan perbandingan jumlah data relevan yang diuji oleh sistem dengan keseluruhan data yang diuji oleh sistem. Untuk penjelasan lebih detail maka dapat dilihat *confusion matrix* pada Tabel 4.

Tabel 4. *Confusion Matrix*

Sistem	Relevan	Tidak Relevan
Ditampilkan Sistem	tp (<i>true positive</i>)	fp (<i>false positive</i>)
Tidak Ditampilkan Sistem	fn (<i>false negative</i>)	tn (<i>true negative</i>)

Untuk formula akurasi dapat dilihat pada persamaan 3.

$$Akurasi = \frac{tp}{tp+fp} \tag{3}$$

Dimana *tp (true positive)* merupakan jumlah data relevan yang diuji oleh sistem, *fn (false negative)* adalah jumlah data relevan yang belum diuji oleh sistem, *fp (false positive)* adalah jumlah data yang belum diuji sistem dan tidak relevan. Sedangkan *tn (true negative)* merupakan data yang tidak relevan dan tidak ditampilkan oleh sistem. Pengujian dilakukan memasukkan data review pengguna ke dalam sistem yang telah dibangun dalam penelitian ini untuk menemukan faktor *usability* serta orientasi dan tingkat setimen dari review yang diujikan. Tabel 5 berikut merupakan hasil ujicoba dengan menggunakan 100 data uji dan 500 data *training*. Sedangkan untuk diagram hasil evaluasi *usability* berdasarkan jumlah orientasi sentimen dapat dilihat pada Gambar 5.

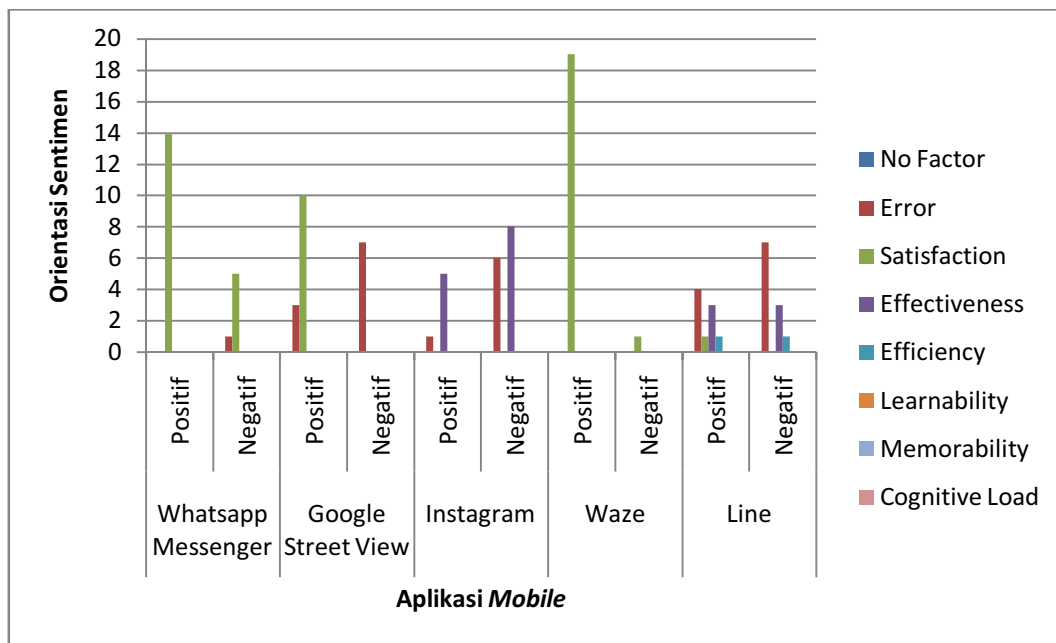
Tabel 4. Analisa Faktor *Usability*

<i>Usability</i>	<i>Whatsapp Messenger</i>				<i>Google Street View</i>				<i>Instagram</i>			
	<i>Pos</i>	<i>Lvl</i>	<i>Neg</i>	<i>Lvl</i>	<i>Pos</i>	<i>Lvl</i>	<i>Neg</i>	<i>Lvl</i>	<i>Pos</i>	<i>Lvl</i>	<i>Neg</i>	<i>Lvl</i>
<i>Error</i>	0	0	1	0.21	3	0.186	7	0.222	1	0.114	6	0.168
<i>Satisfaction</i>	14	0.345	5	0.1	10	0.227	0	0	0	0	0	0
<i>Effectiveness</i>	0	0	0	0	0	0	0	0	5	0.104	8	0.156
<i>Efficiency</i>	0	0	0	0	0	0	0	0	0	0	0	0
<i>Learnability</i>	0	0	0	0	0	0	0	0	0	0	0	0

Usability	Whatsapp Messenger				Google Street View				Instagram			
	Pos	Lvl	Neg	Lvl	Pos	Lvl	Neg	Lvl	Pos	Lvl	Neg	Lvl
Memorability	0	0	0	0	0	0	0	0	0	0	0	0
Cognitive Load	0	0	0	0	0	0	0	0	0	0	0	0

Tabel 5. Analisa Faktor Usability

Usability	Waze				Line			
	Pos	Lvl	Neg	Lvl	Pos	Lvl	Neg	Lvl
Error	0	0	0	0	4	0.122	7	0.125
Satisfaction	19	0.172	1	0.135	1	0.177	0	0
Effectiveness	0	0	0	0	3	0.162	3	0.126
Efficiency	0	0	0	0	1	0.211	1	0.165
Learnability	0	0	0	0	0	0	0	0
Memorability	0	0	0	0	0	0	0	0
Cognitive Load	0	0	0	0	0	0	0	0



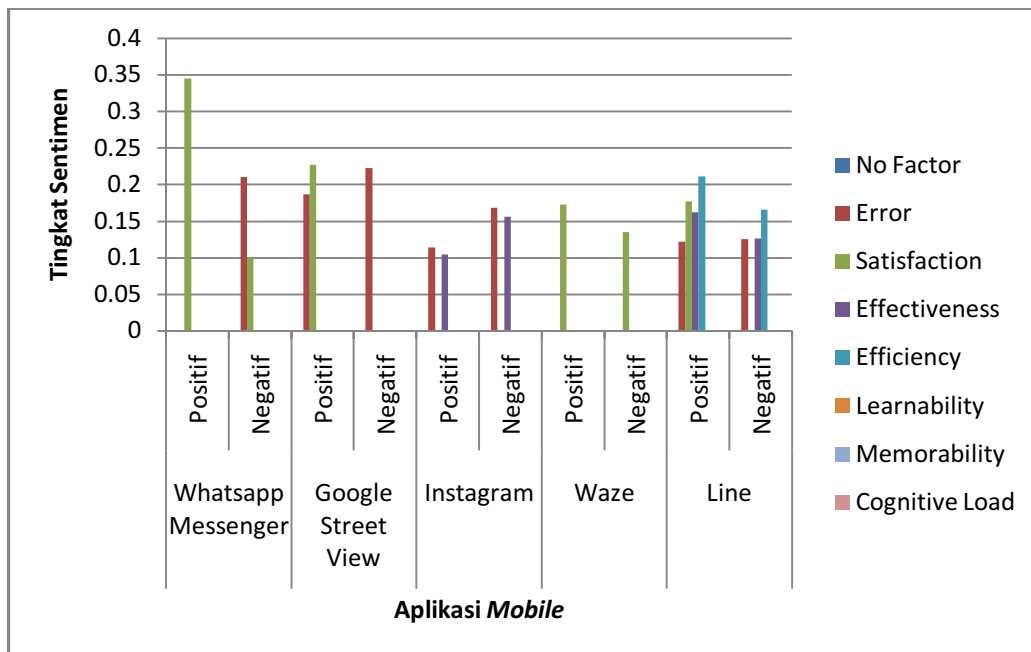
Gambar 5. Diagram Hasil Evaluasi Usability Berdasarkan Jumlah Orientasi Sentimen.

Berdasarkan Tabel 4 dan Tabel 5 diatas, maka dapat disimpulkan bahwa faktor *usability error*, *satisfaction*, *effectiveness* dan *efficiency* merupakan faktor *usability* yang sering diulas oleh para pengguna aplikasi *mobile*. Kebanyakan review pengguna memang mengomentari masalah *error* dan kepuasan pengguna dalam menggunakan aplikasi. Namun banyak juga yang mengomentari masalah fitur serta keakuratan dan kecepatan aplikasi ketika digunakan. Untuk ulasan yang lebih detail maka dapat dilihat diagram pada Gambar 5.

Berdasarkan diagram orientasi sentimen pada Gambar 5, maka dapat disimpulkan bahwa pada aplikasi *Whatsapp Messenger*, *Google Street View* dan *Waze*, faktor positif *satisfaction* merupakan faktor tertinggi yang berhasil diperoleh berdasarkan data uji yang digunakan. Hal tersebut menandakan bahwa aplikasi tersebut sudah stabil dan disukai oleh banyak pengguna. Namun perlu diperhatikan juga faktor *error* pada aplikasi *Google Street View*. Berdasarkan diagram diatas, jumlah faktor negatif *error* yang dimiliki oleh aplikasi *Google Street View* cukup tinggi.

Untuk aplikasi *instagram*, faktor *effectiveness* dan *error* menjadi faktor yang paling dominan

berdasarkan hasil evaluasi *usability*. Namun nilai *effectiveness* dan *error* negatif jauh lebih tinggi daripada *effectiveness* dan *error* positif, sehingga dapat disimpulkan bahwa aplikasi *instagram* terbaru kurang sesuai dengan tujuan pembuatan sebenarnya. Mengingat pada *update* terbaru aplikasi *instagram*, *developer* banyak mengubah fungsi aplikasi ini menjadi mirip dengan aplikasi *snapchat*. Sedangkan pada aplikasi *line*, faktor *error* negatif menjadi sangat dominan dikarenakan pada *update* terbaru banyak terjadi *crash* pada saat membuka *timeline* aplikasi ini. Untuk analisa berdasarkan tingkat sentimen dapat dilihat pada Gambar 6.



Gambar 6. Diagram Tingkat Sentimen Aplikasi Mobile.

Berdasarkan diagram tingkat sentimen pada Gambar 6, maka dapat disimpulkan bahwa tingkat sentimen yang dihasilkan melalui evaluasi *usability* yang diusulkan dalam penelitian ini memiliki nilai yang hampir sama dengan orientasi sentimen yang ditunjukkan pada Gambar 5. Sehingga analisa setiap aplikasi *mobile* yang diujikan tidak jauh berbeda dengan analisa pada orientasi sentimen. Berdasarkan hasil ujicoba maka dapat diperoleh nilai akurasi klasifikasi *usability* sebesar 82% dan nilai akurasi sentimen sebesar 79%.

4. Simpulan

Review pengguna memang sangat efektif digunakan sebagai acuan dalam evaluasi faktor *usability* aplikasi *mobile*. Hal tersebut dikarenakan, review pengguna merupakan gambaran pengalaman pengguna dalam menggunakan aplikasi. Analisis sentimen untuk evaluasi faktor *usability* dalam penelitian ini merupakan metode evaluasi faktor *usability* yang mempertimbangkan orientasi dan tingkat sentimen pada review pengguna aplikasi *mobile*. Dengan mempertimbangkan tingkat sentimen pada review pengguna maka nilai evaluasi yang dihasilkan menjadi lebih efektif serta sesuai dengan faktor *usability* yang diulas oleh pengguna. Berdasarkan hasil pengujian yang dilakukan, nilai akurasi klasifikasi yang diperoleh adalah 82% dan nilai akurasi sentimen yang diperoleh adalah 79%.

Untuk penelitian selanjutnya dapat dikembangkan metode yang dapat melakukan klasifikasi multi *usability* pada setiap review pengguna serta dapat diujicobakan pada review pengguna dengan bahasa lain seperti bahasa Indonesia.

Daftar Pustaka

- [1] I. Atoum and A. Ootom, "Mining Software Quality from Software Reviews: Research Trends and Open Issues," *Int. J. Comput. Trends Technol.*, vol. 31, no. 2, pp. 74–83, 2016.
- [2] A. M. El-Halees, "Software usability evaluation using opinion mining," *J. Softw.*, vol. 9, no. 2, pp. 343–349, 2014.
- [3] R. Harrison, D. Flood, and D. Duce, "Usability of mobile applications: literature review and

- rationale for a new usability model,” pp. 1–16, 2013.
- [4] F. Alshehri and M. Freeman, “Methods of usability evaluations of mobile devices,” *Australas. Conf. Inf. Syst.*, no. 2007, pp. 1–10, 2012.
- [5] K. Lapin, “Deriving Usability Goals for Mobile Applications,” *Proc. 2014 Multimedia, Interact. Des. Innov. Int. Conf. Multimedia, Interact. Des. Innov. - MIDI '14*, pp. 1–6, 2014.
- [6] Zhang, “Challenges , Methodologies , and Issues in the Usability Testing of Mobile Applications,” *IJHCI*.
- [7] Z. Kelemen and R. Kusters, “Towards Applying Text Mining Techniques on Software Quality Standards and Models,” *arXiv Prepr. arXiv ...*, pp. 1–8, 2013.
- [8] D. Wang and H. Zhang, “Inverse-category-frequency based supervised term weighting schemes for text categorization,” *J. Inf. Sci. Eng.*, vol. 29, no. 2, pp. 209–225, 2013.