

IMPLEMENTASI *RULE-BASED DOCUMENT SUBJECTIVITY* PADA SISTEM *OPINION MINING*

Imam Fahrur Rozi³

Permasalahan yang pertama dihadapi dalam mengembangkan sistem opinion mining adalah menentukan apakah suatu teks tergolong kalimat opini atau bukan (*document subjectivity*). Pada penelitian ini dikembangkan sistem *rule-based document subjectivity*. Kalimat pertama kali akan diproses menggunakan Hidden Markov Model Part-of-speech (POS) Tagging. Hasil dari proses tersebut berupa teks yang masing-masing kata didalamnya sudah memiliki tag. Pada hasil proses POS Tagging kemudian diterapkan rule. Dari pengujian didapatkan nilai precision dan recall untuk proses document subjectivity terbaik adalah 0.99 dan 0.88.

Kata-kata kunci: opinion mining, POS tagging, hidden markov model, rule-based

Abstract

The first problem that should be solved in developing opinion mining is how to detect the document subjectivity or how to classify the text document into opinion or non opinion. This research is aimed at developing a system that could detect document subjectivity of text document. First, text will be processed using Hidden markov Model POS Tagging. It will produce a sentence or text with tag/word class assigned in each word. Then, that result will be processed by using rule-based method in order to detect the subjectivity of the document (opinion or not). This system could perform very well. It is proved by value of precision and recall which are produced by testing process is high (0.99 in precision and 0.88 in recall).

Keywords: *opinion mining, POS Tagging, hidden markov model, rule-based*

1. PENDAHULUAN

Informasi tekstual yang tersebar di dunia ini tergolong ke dalam 2 kategori utama yaitu, fakta dan opini. Fakta merupakan

³ *Imam Fahrur Rozi. Dosen Program Studi Manajemen Informatika, Jurusan Teknik Elektro, Politeknik Negeri Malang.*

ekspresi yang bersifat obyektif atas suatu objek atau entitas, maupun peristiwa. Berbeda dengan fakta, opini sifatnya lebih subyektif dan ia mengekspresikan tentang perasaan, penilaian, pendapat seseorang terhadap suatu entitas maupun peristiwa.

Mayoritas penelitian yang telah dilakukan sampai saat ini yang berkenaan dengan pengolahan informasi teks masih sebatas pada pengolahan data faktual/fakta, seperti *information retrieval*, *web search*, *text clustering* dan beberapa jenis lainnya yang tergolong dalam *text mining* ataupun *natural language processing* lainnya. Masih sedikit penelitian yang masuk ke bidang *opinion mining*. Padahal pengolahan data opini juga penting. Salah satu alasan sedikitnya penelitian dalam *opinion mining* adalah terbatasnya jumlah informasi tekstual yang didalamnya terdapat ungkapan opini. Akan tetapi, kemajuan dalam dunia *World Wide Web* beberapa tahun terakhir ini telah merubah kondisi tersebut. WWW telah memfasilitasi manusia untuk dapat mem-publish ungkapan opini mereka, baik dalam bentuk blog, forum di internet, grup-grup diskusi dan berbagai sosial media/networking yang ada di internet. Sumber-sumber informasi dari internet sangat beragam dan banyak. Hal ini menyebabkan kesulitan untuk menganalisis opini yang tersebar di WWW secara manual. Disinilah letak relevansi penelitian *opinion mining*, yaitu untuk mengembangkan sistem untuk analisis opini yang terotomatisasi. Penelitian ini ditujukan untuk mengembangkan sistem yang dapat mendeteksi *document subjectivity* dari suatu dokument teks.

2. KAJIAN PUSTAKA

Penelitian tentang *opinion mining* mulai marak setelah Turney melakukan penelitian tentang *opinion mining* yang diterapkan pada data review konsumen suatu produk menggunakan *Semantic Orientation-Pointwise Mutual Information (SO-PMI)*, dengan nilai akurasi terbaik yang didapat adalah 84% untuk review kendaraan dan 66% untuk data review film. Pada tahun yang sama, Pang dkk. mengembangkan sistem *opinion mining* menggunakan metode *machine learning*, dengan nilai akurasi berkisar 72% sampai 83%. Hal serupa juga dilakukan oleh Kesler dan Nikolov pada tahun 2009, dengan nilai uji F sebesar 0.689. Penelitian ini difokuskan pada pengembangan subsistem

document subjectivity menggunakan metode rule-based rule-based setelah sebelumnya dokument eks diproses menggunakan Hidden Markov Model POS Tagging.

2.1 Part of Speech (POS) Tagging

Part-of-speech Tagging atau yang sering disebut Tagging atau *POS Tagging*, merupakan proses pemberian atau penentuan sebuah *part-of-speech tag* terhadap suatu kata dalam suatu kalimat. Sedangkan *part-of-speech* merupakan kategori kata ditinjau dari sudut pandang kebahasaan (gramatikal), seperti kata benda, kata kerja, kata keterangan, kata sifat dan sebagainya. *POS Tagging* merupakan komponen proses yang sangat penting dan sering digunakan dalam rangkaian pemrosesan bahasa alamiah (*Natural Language Processing*).

Ada beberapa pendekatan yang bisa digunakan untuk melakukan *POS Tagging*, yaitu pendekatan berdasar aturan (*rule based*), pendekatan probabilistik, dan pendekatan berbasis transformasi (*transformational based*).

Untuk *POS Tagging* yang menggunakan metode probabilistik, teknik *Hidden Markov Model* bisa digunakan. Hal ini dikarenakan proses *POS Tagging* bisa dipandang sebagai proses klasifikasi suatu rangkaian atau urutan tag untuk tiap kata dalam suatu kalimat.

2.2 Kelas Kata

Label atau tag yang diberikan ke suatu kata dalam suatu kalimat menunjukkan kelas kata (*word class*) dari kata yang bersangkutan, dalam konteks kalimat tersebut. Kelas kata ini juga disebut sebagai *part-of-speech*. Kumpulan atau koleksi label atau *tag part-of-speech* atau kelas kata disebut sebagai tagset. Dalam proses *POS Tagging* bahasa Indonesia tagset yang digunakan seperti ditunjukkan dalam Tabel 1.

Tabel 1. Tag atau *Label Part-Of-Speech*

POS	Arti	Contoh	POS	Arti	Contoh
OP	Kurung Buka	{{[MD	Modal	Bisa
CP	Kurung Tutup)]}	CC	Kata Sambung Setara	Dan, Atau,

					tetapi
GM	Garis Miring	/	SC	Kata Sambung Tidak Setara	Jika, Ketika
;	Titik Koma	;	DT	Determiner	Para, Ini, Itu
:	Titik Dua	:	UH	Interjection	Wah, Aduh, Oi
“	Tanda Kutip	" dan '	CDO	Kata Bilangan Berurut	Pertama, Kedua
.	Tanda Titik	.	CDC	Kata Bilangan Kolektif	Berdua
,	Tanda Koma	,	CDP	Kata Bilangan Pokok	Satu, Dua, Tiga
-	Garis	-	CDI	Kata Bilangan Tidak Biasa	Beberapa
...	Tanda Pengganti	...	PRP	Kata Ganti Orang	Saya, Mereka
JJ	Kata Sifat	Baik, Bagus	WP	Kata tanya	Apa, Siapa, Dimana
RB	Kata Keterangan	Sementara, Nanti	PRN	Kata Ganti Bilangan	Kedua-duanya
NN	Kata Benda	Kursi, Kulkas	PRL	Kata Ganti Lokasi	Sini, Situ
NNP	Benda Bernama	Toyota, Sony	NEG	Negasi	Bukan, Tidak
NNG	Benda Berpemilik	Motornya	SYM	Simbol	#,%,^,&,*
VBI	Kata Kerja Intransitif	Pergi	RP	Particle	Pun, Kah
VBT	Kata Kerja Transitif	Membeli	FW	Kata Asing	Word
IN	Preposisi	Di, Dari, Ke			

2.3 Hidden Markov Model POS Tagging

Dalam proses POS Tagging ini, data yang akan diobservasi adalah kumpulan kata atau kalimat, dan dari kalimat tersebut akan ditentukan tiap kata penyusunnya akan masuk ke dalam klas kata atau tag apa yang tepat.

Untuk ilustrasi proses HMM dalam POS Tagging, diberikan contoh kalimat “Secretariat is expected to race tomorrow”. Dari kalimat tersebut akan ditentukan rangkaian tag yang paling tepat. Dengan menggunakan teori Bayessian Interpretation, proses dimulai dengan mempertimbangkan semua urutan tag yang mungkin untuk kalimat tersebut. Dari semua kemungkinan urutan tersebut, kemudian akan dipilih urutan tag yang paling mungkin berdasar data observasi yang diberikan, dalam hal ini data observasi adalah kata (*words*) sejumlah n (w_1^n). Dengan kata lain, dari semua kemungkinan urutan tag sejumlah n (t_1^n), dipilih sebuah urutan tag yang menghasilkan $p(t_1^n | w_1^n)$ yang tertinggi. Seperti ditunjukkan dalam persamaan (1).

$$t_1^n = \operatorname{argmax}_{t_1^n} p(t_1^n | w_1^n) \quad (1)$$

Dari persamaan (1) dengan diberikan rangkaian tag (t_1^n) dan rangkaian kata (w_1^n), nilai $p(t_1^n | w_1^n)$ masih belum bisa dihitung secara langsung. Dengan menggunakan teorema Bayes

$$p(x | y) = \frac{P(y | x) P(x)}{P(y)} \quad (2)$$

maka persamaan (1) menjadi,

$$t_1^n = \operatorname{argmax}_{t_1^n} \frac{p(w_1^n | t_1^n) p(t_1^n)}{p(w_1^n)} \quad (3)$$

Penyebut dari persamaan (3) yaitu $p(w_1^n)$ bisa dihilangkan. Hal ini dikarenakan dalam proses pencarian sebuah rangkaian tag yang paling sesuai dari semua kemungkinan rangkaian tag, perhitungan $(p(w_1^n | t_1^n) p(t_1^n)) / (p(w_1^n))$ akan dilakukan pada setiap rangkaian tag. Dan nilai $p(w_1^n)$ akan bernilai tetap untuk setiap rangkaian tag. Setelah penyebut dihilangkan, maka persamaan (3) menjadi,

$$t_1^n = \operatorname{argmax}_{t_1^n} p(w_1^n | t_1^n) p(t_1^n) \quad (4)$$

Dari persamaan (4) bisa dikatakan bahwa rangkaian tag yang paling mungkin (t_1^n) untuk suatu string w_1^n merupakan hasil perkalian antara dua buah nilai probabilitas. Probabilitas yang pertama adalah prior probability $p(t_1^n)$ yang merupakan nilai probabilitas transisi suatu state dari state sebelumnya. Probabilitas yang kedua adalah words likelihood $p(w_1^n | t_1^n)$, yang menunjukkan nilai kemiripan atau kemungkinan suatu kata sebagai suatu state.

Untuk memudahkan perhitungan, HMM POS Tagger menggunakan dua buah asumsi. Asumsi yang pertama adalah probabilitas kemunculan suatu kata hanya tergantung pada tag-nya, dan tidak tergantung dengan kata lain di sekitarnya atau tag lain di sekitarnya (persamaan (5)).

$$p(t_1^n) \approx \prod_{i=1}^n p(t_i | t_{i-1}) \quad (5)$$

Asumsi yang kedua adalah probabilitas suatu kemunculan tag hanya bergantung dari tag sebelumnya (persamaan (6)). Asumsi yang disebut sebagai bigram.

$$t_1^n \approx \operatorname{argmax}_{t_1^n} \prod_{i=1}^n p(w_i | t_i) p(t_i | t_{i-1}) \quad (6)$$

Dengan menggunakan dua buah asumsi pada persamaan (5) dan (6) maka persamaan (4) akan menjadi,

$$t_1^n \approx \operatorname{argmax}_{t_1^n} \prod_{i=1}^n p(w_i | t_i) p(t_i | t_{i-1}) \quad (7)$$

Untuk melakukan perhitungan probabilitas transisi tag $p(t_i | t_{(i-1)})$ dan probabilitas kemiripan kata (word likelihood) $p(w_i | t_i)$ diperlukan koleksi data teks yang telah diberikan tag sebelumnya (corpus). Untuk menghitungnya bisa digunakan persamaan (8) dan (9).

$$p(t_i | t_{i-1}) = \frac{C(t_{i-1}, t_i)}{C(t_{i-1})} \quad (8)$$

$$p(w_i | t_i) = \frac{C(t_i, w_i)}{C(t_i)} \quad (9)$$

2.4 Precision dan Recall

Precision adalah rasio jumlah dokumen relevan yang ditemukan dengan total jumlah dokumen yang ditemukan oleh sistem. *Recall* adalah rasio jumlah dokumen relevan yang

ditemukan kembali dengan total jumlah dokumen dalam kumpulan dokumen yang dianggap relevan. Nilai keduanya biasanya ditunjukkan dalam satuan persen (%). Persamaan untuk precision ditunjukkan pada persamaan (10) dan recall pada persamaan (11).

$$Precision = \frac{\#(item\ yang\ relevan\ yang\ ditemukan)}{\#(item\ yang\ ditemukan)} = P(relevant|ditemukan) \quad (10)$$

$$Recall = \frac{\#(item\ yang\ relevan\ yang\ ditemukan)}{\#(semua\ item\ yang\ relevan)} = P(ditemukan|relevan) \quad (11)$$

Kedua persamaan tersebut, dapat lebih disederhanakan dengan memperhatikan tabel 2.

Tabel 2. Variabel untuk Perhitungan Precision dan Recall

	Relevan	Tak Relevan
Ditemukan	<i>True positives</i> (tp)	<i>False positives</i> (fp)
Tak Ditemukan	<i>False negatives</i> (fn)	<i>True negatives</i> (tn)

Dengan memperhatikan Tabel 2, persamaan (10) dan (11) dapat disederhanakan menjadi,

$$precision = tp/(tp + fp) \quad (12)$$

$$recall = tp/(tp + fn) \quad (13)$$

3. METODE

Pada bagian ini diuraikan tentang data yang digunakan dalam penelitian, metode pengolahan data, kerangka solusi yang ditawarkan atas permasalahan *document subjectivity* pada sistem opinion mining, serta metode pengujian yang digunakan.

3.1 Data

Pada penelitian ini digunakan data sekunder, yaitu data yang berasal dari referensi atau penelitian sebelumnya. Data sekunder yang digunakan dalam penelitian ini sebagai berikut:

- a. Data klas kata (Tag).

Data ini digunakan sebagai acuan tag atau kelas kata yang akan digunakan dalam proses POS Tagging. Data kelas kata (tag)

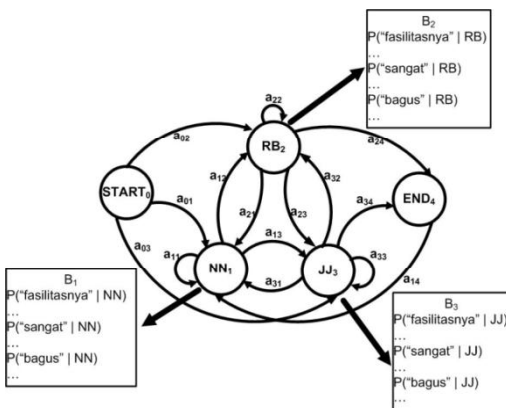
ini diperoleh dari hasil penelitian yang telah dilakukan oleh Alfian Farizki W dkk pada tahun 2010.

b. Data training POS Tagging.

Data training POS Tagging ini berupa corpus yang di dalamnya terdapat kumpulan kalimat, dimana masing-masing kata yang menyusun kalimat tersebut telah diberikan tag atau kelas kata. Data ini didapatkan dari hasil penelitian yang telah dilakukan oleh Universitas Indonesia sebagai salah satu wakil dari Indonesia dalam proyek Pan Localization (PANL10N, <http://www.panl10.net>) [6]. Selain menggunakan tagged corpus (dataset) dari UI, juga dilakukan modifikasi tagged corpus yang digunakan oleh Alfian Farizki dkk. Kedua tagged corpus tersebut dipakai untuk dibandingkan hasilnya.

3.1 Metode Pengolahan Data

Untuk mengetahui apakah suatu kalimat termasuk kalimat opini atau bukan (*document subjectivity*), pertama kali dilakukan POS Tagging untuk menentukan tag atau kelas kata pada setiap kata yang menyusun suatu kalimat. Metode yang digunakan dalam POS Tagging adalah menggunakan Hidden Markov Model (HMM). Kemudian hasil POS Tagging akan dianalisis menggunakan aturan kebahasaan (*rule-based*) untuk mengetahui susunan kata dalam kalimat yang mencerminkan opini. Ilustrasi HMM POS Tagging ditunjukkan pada Gambar 1.

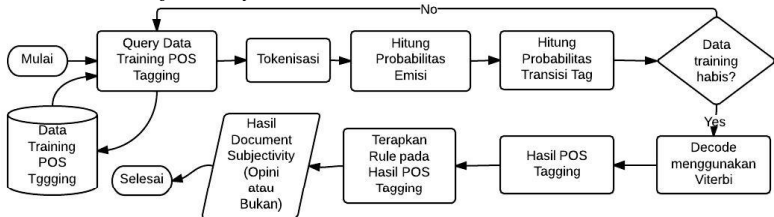


Gambar 1. Proses HMM Pos Tagging

Dari ilustrasi yang diberikan di Gambar 1, kata yang akan dicari *tag*-nya adalah “fasilitasnya”, “sangat” dan “bagus”. Kata tersebut menjadi *observed state*. Sedangkan *hidden state*-nya adalah “NN” (kata benda), “RB” (kata keterangan) dan “JJ” (kata sifat). B_i merupakan probabilitas emisi untuk *hidden state* i dan a_{ij} adalah probabilitas transisi dari *state* k ke *state* j .

3.3 Kerangka Solusi Masalah

Proses dalam sistem akan diawali dengan *query* data *training* untuk *POS Tagging*. Pada tiap data *training* akan dilakukan tokenisasi (memisahkan kata per kata), kemudian dilanjutkan dengan perhitungan nilai probabilitas emisi dan transisi. Probabilitas emisi menunjukkan nilai probabilitas kemunculan suatu *tag* untuk suatu kata kata tertentu, sedangkan probabilitas transisi menunjukkan nilai probabilitas kemunculan transisi suatu *tag* yang diikuti dengan suatu *tag* selanjutnya. Setelah data *training* habis, akan dilakukan proses *decode* menggunakan *Viterbi*. Pada hasil *POS Tagging* akan dianalisis *rule* dan akan ditentukan apakah kalimat termasuk ke dalam klasifikasi opini atau tidak. Hasil dari *POS Tagging* adalah berupa kalimat atau teks dimana tiap kata dalam kalimat tersebut masing-masing sudah memiliki *tag*/kelas kata. Gambar 2 menggambarkan urutan proses yang akan ditempuh untuk menyelesaikan masalah *document subjectivity*.

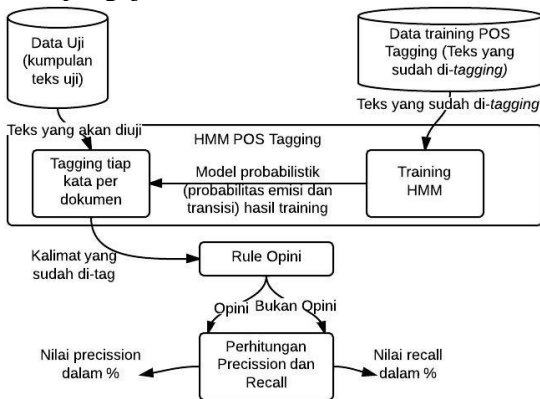


Gambar 2. Kerangka Proses

3.4 Metode Pengujian

Pengujian dilakukan dengan mencari nilai *precision* dan *recall*. Pengujian ini ditujukan untuk mengukur akurasi sistem pada proses *document subjectivity*. Perhitungan nilai *precision* dan *recall* dilakukan setelah semua dokumen yang akan diuji, dideteksi sebagai opini atau bukan. Hasil deteksi tersebut

kemudian dianalisis validitasnya. Analisis dilakukan oleh peneliti dengan mengecek satu persatu dokumen. Gambar 3 menunjukkan tahapan proses pengujian.



Gambar 3. Tahapan Pengujian

4. PEMBAHASAN DAN ANALISIS

Pada bagian ini akan diuraikan perancangan *rule* dan hasil pengujian yang telah dilakukan.

4.1 Perancangan Rule Kalimat Opini

Rule yang digunakan untuk mengetahui apakah suatu kalimat termasuk ke dalam opini atau bukan ditunjukkan pada tabel 3.

Tabel 3. *Rule* Opini

No	Rule	Contoh
1	RB JJ	sangat bagus, dengan bagus, benar-benar bagus, seperti jelek
2	RB VB	semoga berjalan, semoga membawa hikmah, seandainya datang
3	NN JJ	bukunya bagus, pakaiannya rapi, perkataannya halus, jalannya jelek
4	NN VB	Pelajarannya membosankan, perkataannya menjengkelkan
5	JJ VB	mudah difahami, gampang dimaafkan, cepat beradaptasi

6	CK JJ	bagus atau baik, tetapi malas
7	JJ BB	sama bagus
8	VB VB	membuat merinding, membikin pusing
9	JJ RB	indah sekali, bagus sekali
10	VB JJ	membikin bingung
11	NEG JJ	tidak seindah, tidak semudah
12	NEG VB	tidak mengerti, tidak memahami, bukan mengajar
13	PRP VBI	saya menyukai, kita suka
14	PRP VBT	kita suka
15	VBT NN	memiliki kedekatan, memiliki kepekaan
16	MD VBT	Perlu mengambil referensi
17	MD VBI	Perlu dikembangkan

4.2 Hasil Pengujian Document Subjectivity

Pengujian yang dilakukan pada subproses *document subjectivity* ditujukan untuk mengetahui akurasi sistem dalam mendeteksi suatu dokumen apakah termasuk ke dalam kategori opini atau tidak. Pengujian dilakukan pada 575 data yang tersimpan dalam basisdata. Data yang digunakan tersebut berupa dokumen teks. Data tersebut didapatkan melalui kuisioner *online* yang disebar melalui media *website*.

Karena proses *document subjectivity* melibatkan proses *POS Tagging* (menggunakan *HMM*) di dalamnya, proses pengujian juga ditujukan untuk mengetahui pengaruh penggunaan *dataset* (*tagged dataset*) yang berbeda untuk proses *training HMM*.

Tagged dataset yang digunakan dalam pengujian ini ada 2 yaitu:

1. *Dataset* yang dikeluarkan oleh Universitas Indonesia yang pernah digunakan oleh Femphy Pisceldo dkk. pada penelitiannya tahun 2009.

2. *Dataset* yang dirumuskan sendiri oleh peneliti, yang merupakan hasil modifikasi *dataset* yang telah digunakan oleh Alfian Farizki pada penelitiannya tahun 2010.

Tabel 4 menunjukkan hasil pengujian dengan menggunakan *dataset* hasil modifikasi yang dilakukan sendiri oleh peneliti.

Tabel 4. Hasil Pengujian *Document Subjectivity* Menggunakan *Dataset* Hasil Modifikasi oleh Peneliti

		Dikenali sebagai opini oleh pakar (ekspektasi pakar)	
		True	False
Dideteksi sebagai opini oleh sistem (hasil observasi sistem)	True	501 (a)	2 (b)
	False	70 (c)	2 (d)

Catatan: total dokumen = 575, total dokumen yang dikenali sebagai opini oleh sistem = 503, total dokumen yang tidak dikenali sebagai opini oleh sistem = 72

Dari Tabel 4 nilai $precision = 501 / (501+2) = 0.99$ dan $Recall = 501 / (501+70) = 0.88$

Hasil pengujian yang menggunakan *dataset* keluaran Universitas Indonesia ditunjukkan pada Tabel 5.

Tabel 5. Hasil Pengujian *Document Subjectivity* Menggunakan *Dataset* dari Universitas Indonesia

		Dikenali sebagai opini oleh pakar (ekspektasi pakar)	
		True	False
Dideteksi sebagai opini oleh sistem (hasil observasi sistem)	True	411 (a)	2 (b)
	False	160 (c)	2 (d)

Catatan: total dokumen = 575, total dokumen yang dikenali sebagai opini oleh sistem = 413, total dokumen yang tidak dikenali sebagai opini oleh sistem = 162

Dari Tabel 5 nilai $precision = 411 / (411+2) = 0.99$ dan $recall = 411 / (411+160) = 0.72$

Penggunaan 2 *dataset* yang berbeda pada proses *training HMM POS Tagging* menghasilkan nilai *precision* dan *recall* yang berbeda. Ini menunjukkan bahwa penggunaan *dataset* yang relevan dan bagus berpengaruh pada akurasi sistem. Semakin lengkap koleksi kata dan *tag* yang ada dalam *dataset*, dan semakin valid *tag* yang diberikan pada masing-masing kata dalam *dataset*

tersebut, maka semakin bagus dataset tersebut. Akurasi hasil *POS Tagging* yang dihasilkan oleh *HMM* ditentukan oleh dataset yang digunakan pada saat *training*. Desain *rule* berpengaruh pada akurasi deteksi *document subjectivity*. *Rule* didapatkan dari hasil observasi atas struktur kalimat yang tergolong opini.

5. PENUTUP

Dari proses pengujian dan analisis yang telah dilakukan, kesimpulan yang dapat diambil antara lain:

- 1) Penggunaan metode rule-based pada proses *document subjectivity* pada sistem *opinion mining*, bisa berfungsi dengan baik. Hal ini terlihat dari nilai *precision* dan *recall* hasil pengujian hasil Implementasi *Hidden Markov Model* pada proses *POS Tagging* yang digunakan dalam proses *document subjectivity* dapat berfungsi dengan baik. Demikian juga dengan rancangan *rule* yang digunakan, dapat memberikan kontribusi yang baik dalam hasil deteksi *document subjectivity*. Nilai akurasi ditunjukkan dengan parameter uji *precision* dan *recall* yang semua berkisar di atas 0.99 dan 0.88 (dari skala 1).
- 2) Penggunaan *dataset* yang berbeda pada proses *training POS Tagging* berakibat pada hasil akurasi yang berbeda. Semakin bagus *dataset*, maka akurasi sistem *document subjectivity* juga akan semakin bagus.

6. DAFTAR PUSTAKA

- Femphy Pisceldo, Manurung, R., Adriani, Mirna. 2009. Probabilistic Part-of-Speech Tagging for bahasa Indonesia. Third International MALINDO Workshop, colocated event ACL-IJCNLP 2009, Singapore.
- Jurafsky, Daniel dan Martin, H. James. 2007. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice-Hall
- Kessler, J dan Nicolov, N. 2009. Targeting sentiment expressions through supervised ranking of linguistic configurations. Proceedings of the Third International AAI Conference on Weblogs and Social Media, San Jose, California, USA,