

Sistem *Opinion Mining* dengan *Metode Pos Tagging* dan *SVM* Untuk Ekstraksi Data Opini Publik pada Layanan JKBM

Luh Ria Atmarani¹, I.A. Dwi Giriantari², Made Sudarma³

Abstract—Sentiment analysis can be used to extract the opinions of lines of text into an information line. One method used is the Hidden Markov Model. HMM is used to provide word class grammatically the sentence. After determining the word class then apply the rule based. By using a rule based the sentence can be specified include opinions or not. Support Vector Machine method is used to classify the opinions in positive or negative opinion. Data used is data on complaints handling and on-line opinion Technical Services Unit Bali Mandara Health Insurance Bali Province. Results of opinion mining process will be tested using the method precision, recall and accuracy. The results show the percentage of the value of precision, recall and accuracy have an average percentage of 89 percent. It shows the method of tagging posts and SVM were able to classify sentences into the opinions

Intisari— Analisis sentiment digunakan untuk mengekstrak opini dari baris baris teks menjadi suatu informasi. Metode yang digunakan adalah *Hidden Markov Model* (HMM). HMM digunakan untuk memberikan kelas kata secara gramatikal pada suatu kalimat. Penentuan aturan kalimat opini dengan menggunakan rule based. Metode *Support Vector Machine* digunakan untuk membagi opini menjadi opini positif atau negatif. Hasil proses *opinion mining* diuji menggunakan metode *precision*, *recall* dan akurasi. Hasil penelitian menunjukkan presentase nilai *precision*, *recall* dan akurasi memiliki rata rata presentase sebesar 89 persen. Ini menunjukkan metode *pos tagging* dan SVM mampu mengklasifikasikan kalimat kedalam opini dan menentukan kalimat ke dalam opini positif dan negatif..

Kata Kunci—*Opinion Mining*, *Hidden Markov Model*, *Rule Based*, *Support Vector Machine*, *Precision*, *Recall*, *Akurasi*.

I. PENDAHULUAN

Jaminan Kesehatan Bali Mandara (JKBM) merupakan program asuransi kesehatan dalam bidang pelayanan kesehatan JKBM ditujukan untuk masyarakat Bali yang belum memiliki Jaminan Kesehatan. Unit Pelaksana Teknis Jaminan kesehatan Masyarakat Bali (UPT JKMB) adalah pengelola program JKBM. Data Perjanjian Kerja Sama (PKS) tahun 2014 menyatakan JKBM memiliki jumlah peserta 2.733.414 peserta

¹Staff UPT JKMB Dinas Kesehatan Provinsi Bali, Mahasiswa Magister Teknik Elektro Bidang Studi Manajemen Sistem Informasi dan Komputer Universitas Udayana, Kampus Sudirman Denpasar Bali (telp.0361-239599; fax:0361-239599; email: ria.atmarani@gmail.com)

²Dosen Magister Teknik Elektro dan Komputer Universitas Udayana, Kampus Sudirman, Denpasar Bali, (Telp. 0361-239599, fax: 0361-239599; email: dayu.giriantari@unud.ac.id)

³Dosen Magister Teknik Elektro dan Komputer Universitas Udayana, Kampus Sudirman, Denpasar Bali, (Telp. 0361-239599, fax: 0361-239599; email: Msudarmai@unud.ac.id)

Luh Ria Atmarani: Sistem *Opinion Mining* dengan ...

yang tersebar di seluruh Bali. Pada tanggal 15 April 2013 JKBM menanggung kecelakaan tunggal, cacat bawaan seperti penderita *hidrocephalus*, bayi tanpa anus, dan bayi tanpa saluran kencing. Dan yang terbaru sesuai Edaran Gubernur, JKBM menanggung persalinan per 1 Januari 2014[1]. Selama berlangsung dari tahun 2010 beragam opini muncul dari masyarakat tentang pelayanan JKBM, baik pada simakrama Gubernur Bali, sosialisasi JKBM ke Kabupaten/Kota, sosialisasi pada petugas kesehatan atau pada media cetak maupun media elektronik.

Opini yang muncul dari masyarakat merupakan hal yang penting pada sebuah instansi pelayanan publik. Selama ini, opini yang muncul tentang pelayanan JKBM tidak pernah ditampung dan dianalisis sentimennya, apakah opini pelayanan JKBM berorientasi positif atau negatif.. Selama ini ribuan data yang ada pada bagian penanganan keluhan tidak pernah dikelola dengan baik. Analisis sentimen atau *opinion mining* adalah pengolahan bahasa alami untuk melacak sikap, perasaan atau penilaian dari masyarakat tentang topik tertentu, produk atau jasa [2]. Data opini dapat mengungkap hal hal yang belum terungkap pada pelayanan JKBM. Penelitian ini untuk melakukan penerapan metode *opinion mining* untuk menganalisis data opini masyarakat baik peserta JKBM maupun non peserta sehingga opini itu dapat dimanfaatkan untuk membantu bagian penanganan keluhan dalam menganalisis data opini peserta JKBM.

Penggunaan *opinion mining* digunakan untuk mengekstrak pendapat dari teks dan menjadi suatu informasi. Ini adalah bahan studi yang banyak dikaji, dan tidak diragukan lagi nilai komersialnya [3]. Karena banyaknya informasi (perkiraan umum diatas 80%) disimpan sebagai teks, dan *opinion mining* diyakini mempunyai potensi nilai komersial tinggi [4]. Metode *Support Vector Machine* atau SVM bisa digunakan untuk menggolongkan opini ke dalam opini positif dan negatif. Sedangkan dari *natural language processing* metode yang dapat digunakan untuk menyelesaikan masalah pada *Opinion Mining yaitu Part of Speech (POS) tagging*. *Pos Tagging* digunakan untuk memberi kelas kata secara gramatikal pada tiap kata dalam kalimat. *Pos Tagging* menggunakan *Hidden Markov Model* (HMM). HMM merupakan model statistik pada sistem yang melakukan perhitungan probabilitas pada satu kejadian yang tidak bisa diamati berdasarkan kejadian yang bisa diamati [5]. Pada hasil *POS tagging* akan diberikan aturan untuk mengetahui suatu dokumen termasuk opini atau bukan, dan membandingkan bagian kalimat yang merupakan obyek yang menjadi target opini. Kalimat yang dibaca opini kemudian diklasifikasikan menjadi opini negatif dan positif menggunakan SVM.

Penelitian – penelitian terkait metode untuk *opinion mining* dan *pos tagging* telah banyak dilakukan antara lain Kathryn Widiyanti melakukan penelitian menggunakan pelabelan kelas



kata teks berbahasa Indonesia. Dalam penelitian ini *Pos Tagging* menggunakan metode *Hidden Markov Model* dan *Rule Based* untuk menentukan suatu kalimat opini atau bukan. Pravesh Kumar dan Mohid Shahid, melakukan perbandingan metode menggunakan *N-Gram Feature*, *Naive Bayes*, *MCD* dan *SVM* pada *opinion mining* untuk *movie reviews* dan *product reviews*. Dan *SVM* memiliki tingkat akurasi lebih baik diantara metode lainnya. Law, et al dalam penelitiannya yang berjudul *Sentiment Classification Of Online Reviews To Travel Destinations By Supervised Machine Learning Approaches*, membandingkan metode *naive bayes*, *N-gram* dan *SVM classifier*. Pada penelitian ini didapatkan akurasi *SVM* lebih baik dari metode klasifikasi lainnya.

Melihat permasalahan yang ada, pada penelitian ini akan dibangun suatu aplikasi yang untuk ekstraksi data opini publik pada *JKBM*. Metode yang akan digunakan dalam melakukan analisis opini adalah *SVM* untuk klasifikasi opini dan penentuan *Pos Tagging* dengan *HMM* dan *Rule Based* sehingga diharapkan memperoleh klasifikasi data opini yang lebih akurat untuk mengetahui sentimen publik pada layanan *JKBM*.

II. METODE

A. Hidden Markov Model

Hidden Markov Model dikembangkan dari *Markov Chain* yaitu keadaan yang akan datang dari suatu *sequence* tidak ditentukan dari keadaan saat ini, tapi juga perpindahan dari suatu *state* ke *state sequence* yang lain. *State sequence* inilah adalah bagian *hidden* dari suatu *hidden markov models*[6]. Pada *POS Tagging*, data yang diamati adalah kumpulan kata, dari kumpulan kata akan ditentukan tiap kata yang masuk dalam kelas kata yang tepat. Sebagai gambaran proses *HMM*, satu kalimat tersebut akan diberikan *alur tag* yang paling sesuai[7]. Rumus ditunjukkan pada persamaan (1). Persamaan dibawah adalah persamaan dari *Hidden Markov Model* pada kasus *Part of Speech Tagging*

$$Tag_n = \text{Max}(P(\text{word}_i | \text{tag}_i) * P(\text{tag}_i | \text{tag}_{i-1})) \quad (1)$$

Pada *HMM POS Tagger* asumsi probabilitas suatu kemunculan *tag* menggunakan *bigram* yang mana probabilitas kemunculan hanya bergantung dari *tag* sebelumnya, dan akan ditunjukkan pada persamaan (2)

$$t_{1-n} = \arg \max t_1 \dots t_n P(t_1) * \prod_{i=2}^n P(t_i | t_{i-1}) * \prod_{i=1}^n P(w_i | t_i) \quad (2)$$

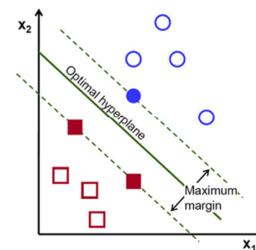
B. Rule Based

Rule Based ini merupakan metode yang menggunakan aturan bahasa (*grammar*) untuk mendapatkan kelas kata pada suatu kata dalam suatu kalimat. Metode *Rule Base* ini memiliki 2 arsitektur. Metode yang pertama adalah metode *Rule base* yang menggunakan kamus untuk menandai kata dengan kelas kata (*leksikon*) [8]. Pada penelitian ini proses *rule based* diawali dengan memproses hasil dari *Hidden Markov Model* yang berupa kata berikut kelas katanya akan dipecah menjadi kalimat-kalimat dengan parameter titik, koma, tanda tanya dan tanda seru. Setelah itu kata akan diambil kelas katanya. Kemudian dari kelas kata pertama sampai terakhir

akandicocokkan dengan *rule* yang sudah ada dikamus aturan. Jika semua susunan aturan pada kalimat ada pada kamus aturan, maka sistem menampilkan kata dengan kelas katanya sebagai output. Jika ada perbedaan kelas kata dengan kelas kata pada kamus, maka sistem memberi tanda pada kata pada hasil *HMM*

C. Support Vector Machine (SVM)

SVM adalah proses pembelajaran terbimbing dan menganalisa suatu data sertaproses pengenalan pola, digunakan metode ini biasanya digunakan pada klasifikasi dan analisis regresi. *SVM* yang standar memperoleh data dan memprediksi setiap masuk yang diberikan, kemungkinan masukan adalah anggota pada salah satu kelas dari dua kelas yang ada, yang menjadikan *SVM* sebagai penggolong *non probabilistik linier biner* [9]. *SVM* sebagai pengklasifikasi akan diberi data latih, masing-masing diberi tanda sebagai milik salah satu dari dua kategori, algoritma pelatihan *SVM* membangun sebuah model yang memprediksi apakah data yang baru jatuh ke dalam suatu kategori atau yang lain. *Hyperlane* pemisah terbaik antara kedua kelas diperoleh dengan mengukur margin dari *hyperlane* serta mencari margin terbesar. Margin adalah jarak antara *hyperlane* tersebut dengan data terdekat dari masing masing kelas [10]. Usaha mencari lokasi *hyperlane* adalah inti dari proses pembelajaran pada *SVM*. Ada banyak *hyperplane* yang mengklasifikasikan data. *Hyperplane* yang paling baik yaitu satu garis yang mewakili pemisahan atau margin terbesar, antara dua kelas. Kita memilih *hyperplane* yang jarak dari dan ke titik data terdekat di setiap sisi dimaksimalkan. Jika sebuah garis maksimal ditemukan, maka dikenal sebagai *hyperplane* maksimum margin dan *linier classifier* yang didefinisikannya dikenal sebagai pengklasifikasi margin maksimal [11]. Gambar 1 menunjukkan *hyperlane* margin



Gambar 1: Hyperlane Margin Maximal

$$\vec{w} \cdot \vec{x} + b = 0 \quad (3)$$

Pola (x_i) yang merupakan kelas -1 (sample negatif) akan dirumuskan menjadi *pattern* dan memenuhi pertidaksamaan

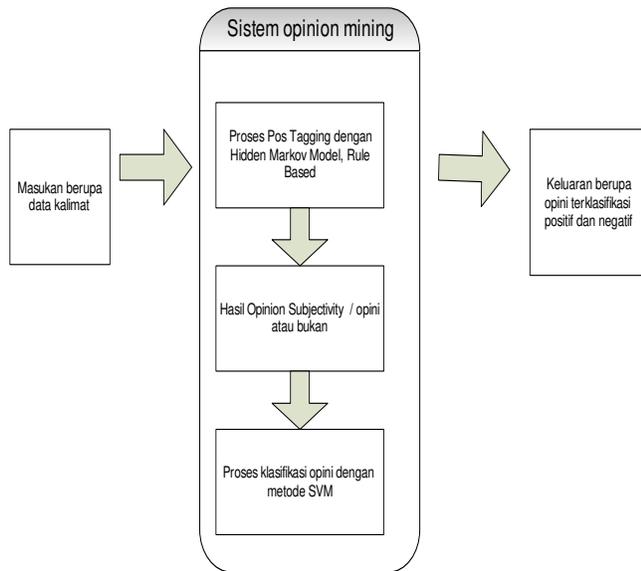
$$\vec{w} \cdot \vec{w} + b \leq -1 \quad (4)$$

Untuk pola (x_i) yang termasuk kelas +1 (sample positif)

$$\vec{w} \cdot \vec{w} + b \geq +1 \quad (5)$$

III. GAMBARAN UMUM SISTEM

Gambaran umum sistem dari penelitian ini dapat digambarkan dengan *Block Diagram*. *Block Diagram* menggambarkan alur sistem secara umum yang berupa masukan, proses dan keluaran dari sistem. Secara keseluruhan gambaran umum dari sistem pada penelitian ini dapat dilihat pada gambar berikut.



Gambar 2: Gambaran Umum Sistem

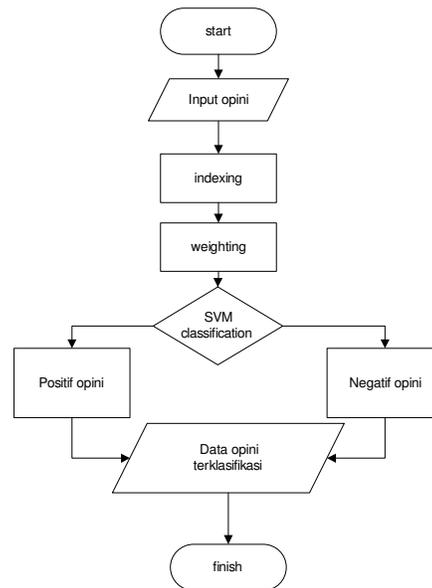
IV. ALUR PENELITIAN

Diagram alur penelitian dapat dilihat pada gambar 3. Langkah – langkah proses pda alur penelitian adalah sebagai berikut :

1. Proses dimulai dengan pengumpulan data. Data diperoleh dari data penanganan keluhan dan kuesioner online tentang pendapat masyarakat seputar JKBM.
2. Proses dilanjutkan dengan pengambilan data dan akan dilakukan pemecahan kata atau tokenisasi. Kemudian tiap kata akan diperoleh probabilitasnya menggunakan HMM. Proses HMM melibatkan proses *hidden state* dan *observed state*. Setelah diperoleh hasil dari HMM, maka hasil tersebut akan diproses dengan *Rule Based* agar dapat diklasifikasi ke dalam data opini atau bukan dengan kamus aturan sehingga memperoleh data berupa *document subjectivity*.³⁵ Tagset Bahasa Indonesia digunakan pada proses *Pos Tagging* dengan mengacu pada KBBI Kateglo. Contoh Tagset yaitu NN (*Common Noun*) contohnya Mobil, PRP (*Personal Pronouns*) contohnya Saya, Kamu, VBI (*Intransitive Verb*) contohnya Pergi, VBT (*Transitive Verb*) contohnya Membeli, JJ (*Adjective*) contohnya kecewa, dan RB (*Adverb*) contohnya Sementara, Nanti

Gambar 3: Gambaran Alur Penelitian

3. Pada klasifikasi dengan svm dimulai dengan mengubah kalimat yang terklasifikasi opini menjadi data vektor. Data vector memiliki 2 komponen yaitu dimensi dan bobot, yang mana bobot sering dituangkan dalam nilai tf-idf, tfidf adalah nilai bobot kata berdasarkan perhitungan tfidf. Kemudian proses dilanjutkan dengan pemberian skor dimana jika diatas 0 termasuk dalam kelas positif dan dibawah 0 termasuk kelas negatif. Data tersebut nantinya merupakan opini terklasifikasi positif atau negatif. Prose ditunjukkan pada gambar dibawah ini.



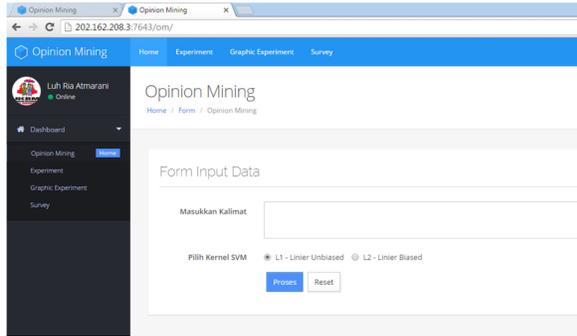
Gambar 4: Diagram Proses SVM



- Proses pengujian sistem dengan membandingkan hasil *opinion classification* pada data uji dengan metode perhitungan *precision, recall* dan akurasi.

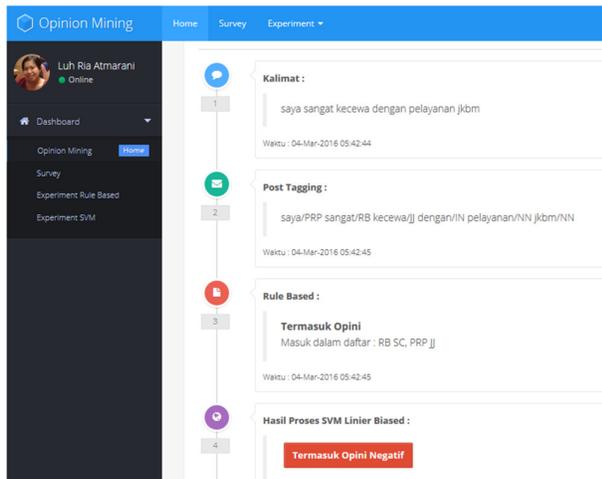
V. TAMPILAN SISTEM

Tampilan sistem berfungsi untuk memudahkan user memasukkan kalimat dan melihat proses yang dilakukan serta hasil dari tahapan penelitian. Berikut adalah antar muka dari halaman input kalimat



Gambar 5: tampilan input kalimat

Kalimat kemudian diproses dengan *pos tagging*. Setelah diperoleh hasil, kemudian data tersebut diproses menggunakan *rule based* untuk menentukan kalimat tersebut opini atau bukan. Setelah itu data opini akan dirubah kedalam bentuk *vector* yang kemudian akan diklasifikasi dengan menggunakan *support vector machine*. Tampilan hasil proses proses *pos tagging* dengan *Hidden Markov Model* dan *Rule Based* serta klasifikasi menggunakan *Support Vector Machine* dapat dilihat pada gambar berikut.



Gambar 6: tampilan proses *pos tagging*

Pada gambar 6 talimat diinputkan pada kolom masukkan kalimat dan kernel SVM yang dipilih adalah SVM *Linier Unbiased* dengan kalimat input “Saya sangat kecewa dengan pelayanan JKBM”. Hasil *Pos Tagging* yaitu Saya/PRP

sangat/RB kecewa/jj dengan/NN pelayanan/NN jkmb/NN, pada *rule based* termasuk opini dan hasil yang ditampilkan adalah opini negatif. Opini negatif dibaca dari data vektor yang muncul di *output* dengan nilai vektor -0.37119843. Tanda minus pada nilai tersebut menunjukkan opini terklasifikasi sebagai opini negatif

VI. EKSPERIMEN

Pengujian data diperoleh dari data penanganan keluhan dan data opini online. Data uji terdiri dari 30, 50, 100, 500, 1000, 1500, 2000, 2500, 3000 dan 3300 data uji. Berikut sampel dari data 30 data uji

TABELI
TABEL 30 DATA UJI

No	Kalimat	Sentimen	
		Klasifikasi Secara Manual	Klasifikasi Oleh Sistem
1	didesa saya belum dapat kartu jkmb tolong dicek di desa duda, tepatnya banjar padang tunggal. Saya kalau berobat mesti urus surat keterangan tidak memiliki jaminan kesehatan dulu, karena tidak adanya kartu jkmb	Opini negatif	Opini negative
2	saya punya orangtua punya kartu jkmb terus dia sakit jantung kenapa kok terus dikasi resep untuk beli obat dan harganya lumayan tinggi saya merasa heran apa untungnya kartu jkmb	Opini negatif	Opini negatif
3	tolong obat obatan yang ditanggung jkmb jangan sampai kosong dipuskesmas karena itu yang menyebabkan adanya pembayaran atas obat yang tidak ditanggung	Opini Negatif	Opini Negatif
4	Terima kasih JKBM atas bantuannya selama lima hari anak kami rawat nginap di RSUD Buleleng .JKBM memang sangat bermanfaat bagi kami.	Opini positif	Opini negatif
5	sistem verifikasi dipuskesmas dan rumah sakit masih belum sempurna	Opini negatif	Opini negatif
6	untuk pasien gawat darurat seharusnya di tolong dahulu bukan lagi harus mengurus administrasinya, bagaimana jika waktu menunggu administrasi pasien tambah gawat	Opini negatif	Opini negatif
7	pelaksanaan program jkmb di RSUD Bangli kurang mendapatkan pelayanan, begitu pula untuk urusan obat seperti antibiotik ada kesan pasien	Opini negatif	Opini negatif

No	Kalimat	Sentimen	
		Klasifikasi Secara Manual	Klasifikasi Oleh Sistem
	digiring membeli antibiotik diluar tanggungan jkbm, mohon diperhatikan		
8	mengurus jkbm dikantor desa sering susah sering sekali kantor kepala desa dalam kondisi terkunci seperti di desa batuagung banjar batuagung	Opini negatif	Opini negatif
9	banyak masyarakat desa abiansemal badung yang tidak memperoleh kartu tetapi ada juga warga seperti i wayan yudana yang memperoleh 3 kartu sekaligus	Opini negatif	Opini negatif
10	pelaksanaan program jkbm di RSUD Bangli kurang mendapatkan pelayanan, begitu pula untuk urusan obat seperti antibiotik ada kesan pasien digiring membeli antibiotik diluar tanggungan jkbm, mohon diperhatikan	Opini negatif	Opini negatif

Data pengujian 30 data opini yang digunakan pada Tabel 1 terdiri dari pengujian manual terdiri dari 20 opini negatif dan 10 opini positif. Sedangkan hasil klasifikasi *Linier Unbiased* yang dilakukan oleh sistem, diperoleh 20 data opini termasuk sentimen negatif dan 8 opini termasuk opini positif, dan 2 opini positif dibaca negatif. Berikut ini adalah Tabel *confusion matrix* yang digunakan sebagai acuan untuk menghitung nilai *precision*, *recall* dan akurasi. Tabel nilai matriks ditunjukkan pada tabel 2

TABEL II
TABEL NILAI CONFUSION MATRIX UNBIASED

		Kelas Prediksi	
		positif	negatif
Kelas Sebenarnya	Positif	8	2
	negatif	0	20

Pada SVM, menggunakan SVM linier dengan *liner biased* dan *unbiased*. Untuk perhitungan akurasi, *Precision* dan *Recall Linier Unbiased* adalah sebagai berikut

TABEL III
TABEL NILAI CONFUSION MATRIX BIASED

		Kelas Prediksi	
		positif	negatif
Kelas Sebenarnya	Positif	10	0
	negatif	0	20

Berdasarkan pengujian, didapatkan hasil akurasi *Linier Unbiased* klasifikasi opini dari sistem *opinion mining* dengan HMM, *rule based* dan SVM sebesar 93% dengan *precision* sebesar 100% dan *recall* sebesar 80%. Sedangkan hasil *Linier* Luh Ria Atmarani: Sistem *Opinion Mining* dengan ...

biased klasifikasi opini dari sistem *opinion mining* dengan HMM, *rule based* dan SVM dengan *precision*, *recall* dan akurasi sebesar 100%.

A. Hasil Uji Sistem dengan SVM Linier Unbiased

Dari proses pengujian data 30 data uji diperoleh nilai *precision* sebesar 100% dan *recall* sebesar 80%, serta akurasi sejumlah 93%. Pada data 50 data uji diperoleh nilai akurasi sebesar 90% dengan *precision* sebesar 96% dan *recall* sebesar 85%. Pada 100 data uji diperoleh nilai *precision* sebesar 100% dan *recall* sebesar 80%, sedangkan dari 500 data uji diperoleh nilai *precision* sebesar 96,15% dan *recall* sebesar 98,80%. Untuk data sejumlah 1000 uji diperoleh *recall* sebesar 95,95% dan *precision* sebesar 99,40%. Berikut Tabel 4 dengan persentase nilai *precision* dan *recall* dari masing masing jumlah data uji.

TABELIV
PERSENTASE PRECISION, RECALL, AKURASI LINIER UNBIASED

Jumlah data uji	precision	Recall	Akurasi
30 data uji	100%	80%	93%
50 data uji	96%	85%	90%
100 data uji	96,15%	100%	98%
500 data uji	96,11%	98,80%	97,40%
1000 data uji	95,95%	99,40%	97,60%
1500 data uji	97%	99,20%	98,07%
2000 data uji	96,41%	99,40%	97,85%
2500 data uji	97,25%	99,04%	98,12%
3000 data uji	95,61%	98,80%	97,13%
3300 data uji	96,01%	99,00%	97,21%

B. Hasil Uji Sistem dengan SVM Linier biased

Pada pengujian SVM *Linier biased*, setelah sistem melakukan pengujian dengan jumlah data uji yang berbeda beda maka nilai *precision* dan *recall* dan akurasinya ditunjukkan pada tabel 5

TABELV
RSENTASE PRECISION, RECALL, AKURASI LINIER BIASED

Jumlah data uji	precision	Recall	Akurasi
30 data uji	100%	100%	100%
50 data uji	96%	93%	94%
100 data uji	98,04%	100%	99%
500 data uji	96,11%	98,80%	97,40%
1000 data uji	96,13%	99,40%	97,70%
1500 data uji	97,00%	99,20%	98,07%
2000 data uji	96,32%	99,40%	97,80%
2500 data uji	97,02%	99,04%	98%
3000 data uji	97,24%	98,47%	97,83%
3300 data uji	97,21%	96,01%	99%

Dari proses pengujian data dengan *Linier Biased* data 30 data uji (20 opini negatif dan 10 opini positif) diperoleh nilai *precision* sebesar 100% dan *recall* sebesar 100%, serta akurasi sejumlah 100%. Pada data 50 data uji (28 opini positif dan 22 opini negatif) diperoleh nilai *precision* sebesar 90,48%, *recall* sebesar 95% dan akurasi sebesar 94%. Pada 100 data (50 opini positif dan 50 opini negatif) diperoleh nilai *precision* sebesar 98,04% dan *recall* sebesar 100%, sedangkan dari 500 data opini diperoleh nilai *precision* sebesar 96,11% dan *recall*



sebesar 98,80 %. Untuk data sejumlah 1000 opini diperoleh *recall* sebesar 96,13 % dan *precision* sebesar 99,40% . Dan dari 3300 data uji diperoleh nilai *precision* 97,21%, nilai *recall* 96,01% dan akurasi sebesar 99%.

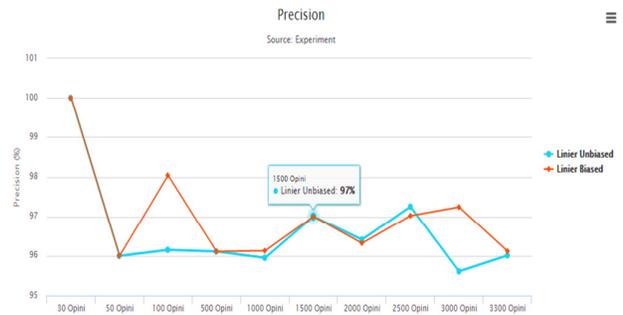
VII. GRAFIK EKSPERIMEN

Grafik eksperimen pada sistem *opinion mining* dengan svm *linier biased* dan *linier unbiased*. Grafik pada sistem ini terdiri dari hasil pengujian akurasi, *precision* dan *recall*. Masing masing hasil pengujian ditunjukkan pada di bawah ini



Gambar 7: Grafik Akurasi Pada Data Opini

Gambar 8 merupakan grafik hasil perhitungan akurasi dari *linier biased* dan *linier unbiased*. Dari beberapa data pengujian terlihat hasil akurasi dengan linier biased memiliki tingkat akurasi yang lebih tinggi dari *linier unbiased*.



Gambar 8: Grafik Precision pada Data Opini

Gambar 9 menunjukkan grafik dengan *precision*. Berdasarkan pengujian *precision*, Secara keseluruhan pada svm *linier biased* tingkat *precision*nya lebih tinggi dibandingkan tingkat *precision* linier unbiased.



Gambar 9: Grafik Recall pada Data Opini

Pada gambar diatas pada grafik *recall* data uji opini terlihat SVM *linier biased* memiliki persentase rata rata *recall* yang lebih tinggi dibandingkan dengan SVM *linier unbiased*

VIII. KESIMPULAN

Penerapan SVM dengan menggunakan klasifikasi *Linier* yang terdiri dari *linier bias* dan *linier unbiased*. Pada percobaan dengan *linier biased* memiliki keakuratan hingga 100 persen. Presentase nilai *Precision*, *recal* dan akurasi pada penelitian *opinion mining* dengan metode *pos tagging* dan *Support Vector Machine* memiliki rata rata 89%. Dengan demikian dapat dikatakan metode *pos tagging* dan *Support Vector Machine* mampu mengklasifikasikan kata kedalam opini dan menentukan kalimat ke dalam opini positif atau negatif

UCAPAN TERIMA KASIH

Terimakasih disampaikan kepada Prof. Ir. Ida Ayu Giriartari, Meng.Sc.,Ph.D dan Dr. Ir. Made Sudarma, M.A.Sc. selaku pembimbing pada penelitian ini. Terimakasih pada Ir IGAP Mahadewi, M.Kes selaku Kepala UPT JKMB Dinas Kesehatan Provinsi Bali atas ijin dan dukungannya untuk melakukan penelitian ini.

REFERENSI

- [1] Tim UPT JKMB, "Pedoman Penyelenggaraan Jaminan Kesehatan Bali Mandara", Denpasar : Dinas Kesehatan Provinsi Bali, 2015.
- [2] Bang Liu, "Sentiment Analysis and Subjectivity in Handbook of Natural Language Processing", 2010.
- [3] Bo Pang and Lilian Lee, "Opinion Mining and Sentiment Analysis Foundations and Trends Information retrieval, vol – volume 2, no. issue 1- 2, pp. 1-135, 2008.
- [4] Jurafsky, DS, "An Introduction to Natural Language Processing, Computational Linguistics and Speech Reconigntion", Pretince – Hall, Inc. New Jersey, 2000.
- [5] Kathryn Widhiyanti, "Pos Tagging Bahasa Indonesia Hidden Markov Model dan Rule Based", 2012
- [6] Pravesh Kumar Singh, Moh Shadid Husain, "Methodological Study of Opinion Mining and Sentiment Analysis Techiques", 2014
- [7] AF Wicaksono, "HMM Based Part of Speech Tagger for Bahasa Indonesia", 2012
- [8] Fam Rashel, Andry Luthfi, Arawinda Dinakaramani, "Building an Indonesian Rule Based Part of Speech Tagger", Universitas Indonesia, Depok, Indonesia.
- [9] Wibisono Y, "Penggunaan Hidden Markov Model untuk Kompresi Kalimat", Program Magister Informatika, Institut Teknologi Bandung, 2008.
- [10] M Rushdi Saleh, MT Martin – valdivia, A. Monrejo-Raez, L A Urena – Lopez, "Expreiments with SVM to Classify Opinion in Different Domain", 2011
- [11] George Forman, "BNS Feature Scalling : An Improved Reprsentation over TF IDF for SVM", 2008