

Klasifikasi Topik Keluhan Pelanggan Berdasarkan *Tweet* dengan Menggunakan Penggabungan *Feature* Hasil Ekstraksi pada Metode *Support Vector Machine* (SVM)

Enda Esyudha Pratama¹, Bambang Riyanto Trilaksono²

^{1,2}Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung

e-mail: endaesyudha@gmail.com, briyanto@lssk.ee.itb.ac.id

Abstrak—Pemanfaatan *twitter* sebagai layanan *customer service* perusahaan sudah mulai banyak digunakan, tak terkecuali Speedy. Mekanisme yang ada saat ini untuk proses klasifikasi bentuk dan jenis keluhan serta informasi tentang jumlah keluhan lewat *twitter* masih dilakukan secara manual. Belum lagi data *twitter* yang bersifat tidak terstruktur tentunya akan menyulitkan untuk dilakukan analisa dan penggalian informasi dari data tersebut. Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk memproses data teks dari *tweet* pengguna *twitter* yang masuk ke akun @TelkomSpeedy untuk diolah menjadi informasi. Informasi tersebut nantinya digunakan untuk klasifikasi bentuk dan jenis keluhan. Merujuk pada beberapa penelitian terkait, salah satu metode klasifikasi yang paling baik untuk digunakan adalah metode *Support Vector Machine* (SVM). Konsep dari SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane* yang dapat memisahkan *dataset* sesuai dengan kelasnya. Kelas yang digunakan dalam penelitian kali ini berdasarkan topik keluhan pelanggan yaitu *billing*, pemasangan/instalasi, putus (*disconnect*), dan lambat. Faktor penting lainnya dalam hal klasifikasi adalah penentuan *feature* atau atribut kata yang akan digunakan. Metode *feature selection* yang digunakan pada penelitian ini adalah *term frequency* (TF), *document frequency* (DF), *information gain*, dan *chi-square*. Pada penelitian ini juga dilakukan metode penggabungan *feature* yang telah dihasilkan dari beberapa metode *feature selection* sebelumnya. Dari hasil penelitian menunjukkan bahwa SVM mampu melakukan klasifikasi keluhan dengan baik, hal ini dibuktikan dengan akurasi 82,50% untuk klasifikasi bentuk keluhan dan 86,67% untuk klasifikasi jenis keluhan. Sedangkan untuk kombinasi penggunaan *feature* dapat meningkatkan akurasi menjadi 83,33% untuk bentuk keluhan dan 89,17% untuk jenis keluhan.

Kata Kunci—*customer service*, klasifikasi topik keluhan, penggabungan *feature*, *support vector machine*

I. PENDAHULUAN

T*witter* dimanfaatkan untuk berbagai hal, dari sekedar berbagi hal pribadi sampai mendapatkan referensi suatu produk yang ingin dibeli. Pada umumnya, *follower* suatu *brand* atau produk berharap untuk mendapatkan informasi tentang produk tersebut dan layanan lainnya dari media sosial sebelum mereka memutuskan untuk membelinya. Oleh karena itu, terdapat dua *entry point* pemanfaatan *Twitter* yang paling sering digunakan oleh suatu perusahaan, pertama adalah penggunaan media sosial untuk *marketing effort* (promo, *campaign*, dan sebagainya) dan yang kedua adalah penggunaan media sosial untuk layanan pelanggan (*customer services*) [1].

Survei yang dilakukan oleh Aberden Group pada 170 perusahaan menunjukkan sekitar 40% perusahaan menggunakan jalur media sosial sebagai layanan *customer service* dan *support* bagi konsumen mereka. Masih pada survei yang sama dengan pembahasan *tools* yang digunakan untuk *customer service*, hasil survei menunjukkan *facebook page* menjadi *tools* yang paling banyak digunakan sebesar 73%, kemudian *blog* atau *website* sebesar 59%, sedangkan untuk *twitter* sendiri sekitar 51%. Hal ini menunjukkan pemanfaatan *twitter* sebagai sarana *customer service* sudah mulai banyak dilirik oleh perusahaan [2].

Saat ini memang sudah banyak *tools* yang digunakan untuk *social media analysis*. Namun, kebanyakan dari aplikasi tersebut masih berfokus pada *marketing tools* untuk melihat sejauh mana *brand* mereka diketahui atau diperbincangkan di media sosial [3]. Pada umumnya perusahaan membuat suatu divisi khusus yang mengurus media sosial. Disana terdapat beberapa orang yang *online* untuk membalas satu persatu *mention* yang masuk ke akun *twitter* perusahaan [4]. Hal ini tentunya menjadi suatu proses yang tidak efisien terutama dalam hal *time respond*. Sebuah survei di Amerika Serikat yang dilakukan oleh A. T. Kearney menunjukkan bahwa 55% pelanggan menginginkan komentar atau pertanyaannya dibalas saat itu juga atau setidaknya pada hari yang sama [5]. Belum lagi kesulitan yang dihadapi untuk menghasilkan analisa atau informasi tentang kinerja layanan tersebut. Data yang berasal dari *twitter* bersifat tidak terstruktur dan mengandung banyak *noise*. Dibutuhkan suatu *tools* untuk memproses data tersebut menjadi data terstruktur sehingga nantinya dapat menghasilkan informasi yang berguna.

Penelitian kali ini bertujuan untuk menghasilkan *tools* yang dapat melakukan proses klasifikasi bentuk keluhan dan jenis keluhan secara otomatis dari data *twitter*. *Tools* ini nantinya akan mampu melakukan klasifikasi mana yang merupakan bentuk keluhan dan bukan keluhan dari setiap *mention* yang masuk pada suatu *account Twitter*. Dari daftar keluhan atau pertanyaan konsumen tersebut, akan dilakukan proses klasifikasi untuk mengelompokan jenis keluhan atau pertanyaan yang disampaikan oleh konsumen. *Tools* ini juga menampilkan informasi tentang jumlah keluhan dan jumlah konsumen yang menyampaikan keluhan. Penelitian kali ini akan dilakukan pada layanan *customer service* akun *twitter* dari @TelkomSpeedy. Mekanisme yang ada saat ini untuk penanganan keluhan lewat sosial media menggunakan notifikasi *email* untuk melihat setiap *mention* yang masuk. Proses klasifikasi keluhan, informasi jumlah keluhan, dan jumlah pelanggan yang menyampaikan keluhan masih

dilakukan secara manual. Dengan adanya penelitian kali ini diharapkan dapat memberikan informasi tentang keluhan pelanggan secara lebih akurat dan menentukan langkah strategis untuk penanganan keluhan tersebut yang nantinya berdampak pada peningkatan kualitas layanan *customer service*.

Metode klasifikasi yang digunakan pada penelitian kali ini adalah algoritma *Support Vector Machine* (SVM). Berdasarkan penelitian yang dilakukan oleh Aqsath, algoritma SVM memiliki tingkat akurasi paling tinggi sebesar 87% dalam hal klasifikasi teks. Oleh karena itu, pada penelitian kali ini metode klasifikasi teks menggunakan algoritma SVM [6].

II. DASAR TEORI

A. Pemanfaatan *Twitter* Sebagai *Customer Service*

Media sosial saat ini telah mengubah cara orang melakukan bisnis. Pelanggan dan pemilik bisnis memiliki pilihan lebih dari sebelumnya. Bahkan banyak orang melakukan keluhan lewat media sosial, baik itu di *Twitter* maupun *Facebook*. Media sosial menjadi bagian penting dari budaya bisnis, maka menggunakan media sosial untuk layanan pelanggan adalah perpindahan dari sebuah konsep menuju kebutuhan pasar [7].

Saat ini, sebuah keluhan dari pelanggan di media sosial dapat berpengaruh dan memberi efek yang besar pada reputasi perusahaan. Lebih lanjut lagi, percakapan di media sosial juga dapat digunakan sebagai sistem peringatan awal untuk isu yang berkembang sekitar produk dan layanan perusahaan.

Survei Avaya Asia Pacific Customer Experience Index menemukan data tentang konsumen di Asia Pasifik semakin menyukai komunikasi multi-saluran dalam layanan pelanggan. Tercatat 75% konsumen lebih menyukai dukungan layanan pelanggan diberikan melalui beberapa saluran, naik dari 60% dibandingkan 2012. Hal ini menunjukkan bahwa perusahaan-perusahaan saat ini harus segera memiliki strategi pengalaman pelanggan multi-saluran yang lengkap [8].

B. *Support Vector Machine* (SVM)

Support Vector Machine (SVM) adalah sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur (*feature space*) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan learning bias yang berasal dari teori pembelajaran statistik. Prinsip dasar SVM adalah pengklasifikasi linier, dan selanjutnya dikembangkan agar dapat bekerja pada permasalahan nonlinier. dengan memasukkan konsep *kernel trick* pada ruang kerja berdimensi tinggi. Perkembangan ini memberikan minat penelitian di bidang pengenalan pola untuk investigasi potensi kemampuan SVM secara teoritis maupun dari segi aplikasi [9].

Linearly separable data merupakan data yang dapat dipisahkan secara linear. Misalkan $x_i \in \{x_n, \dots, x_1\}$ adalah dataset dan $y_i \in \{+1, -1\}$ adalah label kelas dari data x_i . Fungsi yang digunakan untuk memisahkan kelas adalah dengan

menggunakan fungsi *linear*, dimana fungsi tersebut didefinisikan sebagai berikut:

$$g(x) = \text{sign}(f(x))$$

$$\text{dengan } f(x) = (w^T x + b) \tag{1}$$

dimana, w = normal bidang
 b = posisi bidang relatif terhadap pusat koordinat

Untuk pencarian bidang pemisah terbaik dengan nilai margin terbesar dapat dirumuskan menjadi masalah optimasi *constraint SVM* untuk kasus klasifikasi *linear* dalam *primal space*, yaitu:

$$\min \frac{1}{2} |w|^2$$

$$\text{s. t. } y_i (x_i \cdot w + b) - 1 \geq 0 \tag{2}$$

dimana, x_i = data input
 y_i = output dari x_i

Pada persamaan optimasi *constraint* untuk meminimalkan fungsi objektif $1/2|w|^2$ atau memaksimalkan $w^T w$ yaitu dengan memperhatikan pembatas $y_i(x_i \cdot w + b) \geq 1$. Bila output data $y_i = +1$, maka pembatas menjadi $(x_i \cdot w + b) \geq 1$, sebaliknya $y_i = -1$, maka pembatas menjadi $(x_i \cdot w + b) \leq 1$. Dalam beberapa kasus, terdapat beberapa data yang tidak dapat diklasifikasikan secara benar (*infeasible*), maka dapat dinyatakan melalui persamaan berikut,

$$\min \frac{1}{2} |w|^2 + C \left(\sum_{i=1}^n \xi_i \right)$$

$$\text{s. t. } y_i (w \cdot x_i + b) \geq 1 - \xi_i \tag{3}$$

Nilai C (*Complexity*) adalah nilai yang dipilih sebelum dilakukan optimasi dengan proses *Quadratic Programming*. Nilai C memiliki rentang antara nol sampai positif tak hingga ($0 < C < \infty$). Tujuan adanya nilai C (*Complexity*) adalah untuk meminimalkan *error* dan memperkecil nilai *slack variabel*. Jika nilai C mendekati nol, maka lebar margin pada bidang pembatas menjadi maksimum dan jumlah data yang dilatih yang berada dalam margin atau yang ada posisi yang salah tidak akan dipedulikan. Hal ini berarti akan mengurangi tingkat akurasi pada proses *training*, sehingga mengakibatkan data uji tidak dapat diklasifikasikan dengan baik.

Dalam kasus *machine learning*, *kernel trick* merupakan metode yang menggunakan algoritma *linier classifier* untuk menyelesaikan permasalahan nonlinier dengan cara memetakan dimensi input ke ruang dimensi yang lebih tinggi, sehingga membuat *linier classifier* di ruang dimensi yang baru setara dengan *non linear classifier* di ruang dimensi asli. Dengan *kernel*, fungsi pemetaan tidak pernah dihitung secara eksplisit, karena ruang dimensi tinggi yang digunakan memungkinkan pada dimensi yang tak terbatas. Menurut Hsu[10], berikut ini adalah beberapa fungsi kernel yang umum digunakan antara lain:

- *Linear* : $K(x_i, x_j) = x_i^T x_j$
- *Polynomial* : $K(x_i, x_j) = (\gamma \cdot x_i^T x_j + r)^d, \gamma > 0$
- *RBF* : $K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2), \gamma > 0$
- *Sigmoid* : $K(x_i, x_j) = \tanh(\gamma \cdot x_i^T x_j + r)$

C. Penelitian Terkait

Penelitian tentang pemanfaatan *twitter* untuk berbagai macam kebutuhan telah banyak dilakukan sebelumnya. Zhichao dalam penelitiannya menunjukkan dampak dari pemanfaatan sosial media pada layanan pelanggan terhadap kepuasan pelanggan. Dalam penelitiannya, setiap akun sosial media dari tiap pelanggan di kelompokkan (*profiling*) berdasarkan parameter tertentu seperti jenis kelamin, umur, dan sebagainya [11]. Zhiheng dalam penelitiannya yang berjudul *Discovering User Interest on Twitter with a Modified Author-Topic Model* mencoba menemukan ketertarikan seseorang (*user interest*) berdasarkan *tweet* yang dihasilkan [12]. Hasil dari penelitiannya menunjukkan *tweet* yang dihasilkan dari *user* memiliki pengaruh yang besar dan sangat berkaitan dengan *interest* dari *user* tersebut. Penelitian lainnya yang dilakukan oleh Finin, dimana dia memprediksi tentang suatu kejadian (*event*) yang sedang terjadi berdasarkan data yang diambil dari *twitter*. Hal ini menunjukkan pula tentang pemanfaatan data dari *twitter* dapat dijadikan kumpulan *dataset* yang akurat [13].

Aqsath telah melakukan penelitian untuk *sentiment classification* dengan menggunakan data *twitter* untuk *tweet* yang berbahasa Indonesia. Metode yang digunakan untuk *feature selection* yaitu kamus kata (*dictionary*) sedangkan untuk metode klasifikasi yang digunakan *Support Vector Machine* (SVM). Klasifikasi kelas yang digunakan yaitu netral, positif, dan negatif. Penelitian tersebut menunjukkan tingkat akurasi algoritma SVM sebesar 86,66% [6].

Algoritma SVM juga memiliki kinerja yang baik dalam hal kategorisasi teks bahasa Indonesia. Fatimah Wulandini dan Anto Satriyo Nugroho melakukan pengujian terhadap dokumen berbahasa Indonesia dengan menggunakan data latih sebanyak 240 dokumen dan data uji sebanyak 120 dokumen. Metode *feature selection* yang digunakan yaitu pengindeksan kata (*indexing word*). Dari hasil penelitian didapat akurasi SVM sebesar 92,5% lebih baik dibandingkan dengan algoritma lainnya yang masing-masing memiliki akurasi NBC(90%), kNN(27,5%), dan C45(77,5%) [14]. Penelitian yang dilakukan oleh Watters menggunakan 600 dokumen untuk dikategorisasi [15]. Metode *feature selection* yang digunakan yaitu *document frequency* (DF) dengan mengambil nilai parameter (*threshold*) secara acak. Di dalam penelitiannya dia mencoba membandingkan antara dua buah algoritma yaitu SVM dengan ANN (*Artificial Neural Network*). Hasil kinerja menunjukkan tingkat akurasi SVM sebesar 82%, jauh lebih baik dibandingkan dengan Algoritma ANN yang hanya 58,53%.

Meesad dalam penelitiannya mencoba menunjukkan kinerja SVM dalam kasus klasifikasi dokumen untuk *digital library* [16]. Metode *feature selection* yang digunakan yaitu *chi-*

square, *information gain*, dll. Dari penelitiannya didapat kinerja SVM (92,20%) lebih baik dibandingkan dengan NBC(91,70%) dan ID3(86,20%).

III. ANALISIS DAN PERANCANGAN

A. Dataset

Dataset yang digunakan pada penelitian kali ini berasal dari *mention tweet* yang masuk pada akun *Twitter @SpeedyTelkomsel*. Proses pengumpulan *dataset* ini dilakukan dengan menggunakan suatu program *crawler* berbasis *web* yang memanfaatkan layanan *twitter API*. Jumlah *dataset* yang digunakan berjumlah 600 *tweet* dengan rincian sebagai berikut:

Tabel 1. Rincian Dataset

Keluhan-Bukan Keluhan				Data Training (80%)	Data Testing (20%)
Tweet Keluhan		Tweet Bukan Keluhan			
300		300		480	120
600					
Jenis Keluhan				Data Training (80%)	Data Testing (20%)
Billing	Pemasangan	Disconnect	Lambat		
150	150	150	150	480	120
600					

B. Text preprocessing

Text preprocessing merupakan sekumpulan tahapan yang harus dilakukan untuk mempersiapkan kumpulan *dataset* menjadi data masukan (*input*) pada proses selanjutnya yaitu klasifikasi menggunakan SVM. Adapun beberapa tahapan yang dilakukan pada *text preprocessing* ini yaitu *tokenizing*, *stopword removal*, dan *stemming*.

Proses *tokenizing* merupakan proses memisahkan setiap kata dalam suatu kalimat sehingga menghasilkan kumpulan kata-kata yang berdiri sendiri. Pemisahan kata dilakukan dengan cara menemukan spasi (*space*) antar kata. Pada proses ini pula dilakukan penghapusan tanda baca. Langkah selanjutnya adalah melakukan proses *filtering*. Pada proses ini, setiap kata yang telah berdiri sendiri akan diidentifikasi untuk menentukan kata itu akan digunakan atau dihapus. Kata-kata yang dihapus adalah kata-kata yang termasuk dalam *stoplist*. *Stoplist* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words*. Contoh *stopwords* adalah yang, dan, di, dari, atau, pada, saat, dan lain sebagainya. Pada penelitian kali ini berfokus pada bentuk keluhan dari *tweet text*, maka kata-kata yang mengandung makna dari *tweet entity* seperti *mention*, *retweet*, *hashtag*, dan *link url* juga akan dihapus. Setiap kata juga akan dibersihkan dari simbol atau kode karakter *numeric (noisy text)*, seperti : ('~&#[0-9]+;'). Adapun beberapa tahapan yang dilakukan pada proses *stemming* secara detail yaitu sebagai berikut:

1) Cek kata apakah ada di kamus, jika ada maka kata ditemukan. Namun jika tidak, maka lanjutkan ke tahap berikutnya.

2) Hapus *inflection suffix*, yaitu : "-lah", "-kah", "-ku", "-mu", atau "-nya"

3) Cek *prefix* dan *suffix* yang tidak diperbolehkan, yaitu : ("be-" dan "-i"), ("di-" dan "-an"), ("ke-" dan "-i, -kan"), ("me-" dan "-an"), ("se-" dan "-i, -kan")

4) Hapus *derivation suffix*, yaitu : "-i", "-an", "-kan"

5) Hapus *derivation prefix*, yaitu : "di-", "ke-", "se-", "te-", "be-", "me-", atau "pe"

C. Ekstraksi Features

Pada penelitian kali ini, untuk mendapatkan nilai *threshold* parameter yang terbaik maka akan dilakukan pengamatan terhadap distribusi frekuensi kemunculan kata dan jumlah *feature*. Nilai *threshold* yang terbaik adalah titik dimana frekuensi kemunculan kata dan jumlah *feature* mulai konstan.

Dari pengamatan terhadap nilai *threshold*, didapat sejumlah fitur kata yang akan digunakan dari setiap metode ekstraksi. Untuk metode *term frequency* (TF) dan *document frequency* (DF), jumlah *feature* yang dihasilkan mulai konstan pada saat nilai *threshold* berada pada kisaran 15-20. Hal ini dapat diartikan bahwa *feature* yang akan digunakan dalam penelitian kali ini adalah kata atau *term* yang memiliki frekuensi kemunculan kata di atas 15 kali. Berdasarkan parameter tersebut, maka didapat jumlah *feature* untuk tiap proses klasifikasinya. Untuk proses klasifikasi keluhan-bukan keluhan dengan metode TF menggunakan 51 *feature* dan metode DF menggunakan 44 *feature*. Sedangkan untuk proses klasifikasi jenis keluhan dengan metode TF menggunakan 47 *feature* dan metode DF menggunakan 44 *feature*.

Pada metode *information gain* (IG), jumlah frekuensi fitur yang dihasilkan mulai konstan pada saat nilai IG berada pada nilai $\geq 0,02$ untuk keluhan-bukan keluhan dan $\geq 0,03$ untuk jenis keluhan. Berdasarkan parameter tersebut didapat jumlah fitur yang dihasilkan untuk klasifikasi bentuk keluhan sebanyak 25 atribut dan untuk klasifikasi jenis keluhan sebanyak 46 atribut.

Untuk metode *chi-square*, jumlah *feature* yang dihasilkan mulai konstan pada saat nilai *threshold* berada pada kisaran 6-15 untuk bentuk keluhan dan 16-20 untuk jenis keluhan. Berdasarkan parameter tersebut, maka didapat jumlah *feature* untuk tiap proses klasifikasinya. Untuk proses klasifikasi keluhan-bukan keluhan menggunakan 29 *feature*. Sedangkan untuk jenis klasifikasi menggunakan 37 *feature*. Adapun rincian jumlah *feature* yang digunakan dapat dilihat pada tabel berikut.

Tabel 2.
Jumlah Feature Tiap Metode

Klasifikasi	Metode Ekstraksi			
	TF	DF	Inf.Gain	Chisquare
Keluhan-Bkn Klhn	51	44	25	29
Jenis Keluhan	47	44	46	37

Setelah didapat beberapa kumpulan kata atau *term* sebagai kumpulan *feature* dari beberapa metode ekstraksi, penelitian ini juga menggunakan metode penggabungan *feature* dengan menggunakan operasi gabungan (*union*) dari set *feature* yang telah dihasilkan oleh masing-masing metode. Sebagai contoh untuk metode ekstraksi x menggunakan *feature* kata putus dan *wifi*. Sedangkan metode ekstraksi y menggunakan *feature* kata putus dan lambat. Maka gabungan *feature* yang digunakan yaitu kata putus, *wifi*, dan lambat.

D. Data Text to Vector

Model ruang vektor digunakan untuk memberikan setiap *feature* dalam dokumen sebuah ID (dimensi) dan sebuah bobot berdasarkan seberapa penting keberadaannya dalam dokumen (*tweet*). Adapun contoh format data *input* adalah :

0,0,0,0,3,0,8,0,0,0,0,0,2,2,0,0,0,1,1,1,0,0,0,3,billing
1,5,6,0,0,0,2,1,0,7,0,0,2,,0,1,1,1,0,0,0,2,0.1putus

Angka menunjukkan bobot fitur pada setiap *tweet*. Setiap bobot dipisahkan oleh koma (.). Nilai angka yang muncul sebanyak jumlah fitur yang digunakan. Sedangkan nilai lainnya sesuai dengan perhitungan bobot yang digunakan. Pada bagian akhir dari baris data vektor merupakan nama kelas.

E. Fungsi Kernel dan Estimasi Parameter

Berdasarkan *dataset* yang digunakan, maka dalam pada penelitian kali ini akan diterapkan SVM *nonlinear*. Fungsi *kernel* yang digunakan adalah fungsi *kernel* RBF karena memiliki performansi yang paling baik dibandingkan dengan *kernel* linier pada parameter tertentu maupun *kernel* polinomial. Pada penelitian ini estimasi parameter terbaik akan dilakukan dengan menggunakan *grid search*. *Grid search* bertujuan membuat *grid* parameter dari setiap pasangan (C, γ). Parameter nilai (C, γ) ditentukan terlebih dahulu dengan rentang nilai 0,1 sampai 0,9. Kemudian memasang setiap nilai parameter (C, γ) tersebut. Untuk melihat rataan akurasi dari data latih pada setiap pasangan nilai (C, γ) digunakan metode *10-fold cross validation*. Pasangan nilai (C, γ) yang menghasilkan rataan akurasi terbaik akan digunakan untuk proses *training* terhadap keseluruhan data uji.

10-fold cross validation dilakukan pada data latih yang akan dibagi menjadi 10 *subset* (segmen) sama banyak. Akan dilakukan 10 iterasi proses *training* dan *testing*, dengan 9/10 segmen sebagai data latih dan 1/10 segmen sebagai data uji secara bergantian. Sehingga untuk setiap *subset* (segmen) berkesempatan menjadi data uji.

F. Training & Testing

Masukan atau *input* dalam membangun suatu model dalam SVM berupa data latih yang telah diubah ke dalam bentuk vektor. Selain itu, terdapat pula input parameter lainnya seperti *C* dan γ serta metoda *kernel* yang dipilih sebagaimana telah dijelaskan sebelumnya. Data latih dalam bentuk vektor disimpan dalam suatu *file* bernama *train.arff*. Untuk membangkitkan suatu model, sistem akan memanggil suatu perintah dalam program dan menyimpan *output* model tersebut ke dalam suatu file *nama_file.model*.

Untuk menguji keakuratan dari model yang dibangun, akan dilakukan evaluasi terhadap model tersebut menggunakan data uji. Sama halnya dengan data latih, data uji yang akan digunakan juga harus terlebih dahulu diubah ke dalam bentuk vektor dan disimpan ke dalam suatu berkas *testing.arff*.

Kemudian sistem akan menjalankan suatu perintah program untuk melakukan pengujian. Sistem akan memberikan informasi keakuratan dari model dengan menghitung presentasi data yang diklasifikasikan secara benar terhadap jumlah data uji. Jika pada keluaran menunjukkan nilai atau label yang sama dengan nilai yang ada pada data uji, maka dapat dikatakan sistem melakukan klasifikasi secara benar. Begitu pula jika kondisi yang terjadi tidak sesuai, maka sistem akan menilai hal itu sebagai ketidakakuratan proses klasifikasi.

IV. PENGUJIAN

A. Pengujian Parameter (*C,γ*) Terbaik pada Fungsi Kernel

Proses klasifikasi SVM pada penelitian kali ini menggunakan fungsi *kernel Gaussian RBF* dimana pada kernel tersebut memerlukan parameter *C* dan γ pada prosesnya. Untuk mendapatkan nilai parameter terbaik, akan dilakukan beberapa tahapan terhadap dataset. Langkah pertama dilakukan dengan membuat *grid* parameter pada setiap pasangan nilai parameter. Parameter nilai *C* dan γ ditentukan terlebih dahulu secara manual dengan rentang nilai masing-masing 0,1 sampai dengan 0,9. Pasangan nilai *C* dan γ terbaik adalah yang memberikan nilai rata-ran akurasi paling tinggi pada proses klasifikasi. Adapun hasil pengujian *grid search* dapat dilihat pada tabel di bawah ini.

Tabel 3. *Grid search* bentuk keluhan

γ C	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0.1	75.63	76.67	77.71	77.50	79.38	77.50	74.58
0.2	77.29	77.5	77.29	80.42	81.88	82.08	81.67
0.3	77.29	78.13	78.96	81.04	82.08	82.08	81.88
0.4	76.46	78.54	80.00	81.67	82.08	82.50	81.88
0.5	77.08	78.98	80.21	82.29	82.08	82.29	82.08
0.6	77.29	80.00	81.25	82.29	82.29	82.29	81.88
0.7	77.71	80.83	82.29	82.08	82.29	82.29	82.29
0.8	77.91	80.83	81.46	82.29	82.50	83.75	83.13
0.9	79.38	81.25	81.67	82.29	82.92	83.54	83.33

Tabel 4. *Grid Search* Jenis Keluhan

γ C	0,1	0,2	0,3	0,4	0,5	0,6	0,7
0.1	69.79	72.71	71.67	71.35	67.50	64.79	60.63
0.2	74.17	74.38	74.17	73.13	71.46	68.33	65.83
0.3	73.75	74.17	73.75	73.96	72.08	70.83	69.17
0.4	74.38	73.54	74.79	75.21	74.17	72.50	71.04
0.5	73.54	75.00	75.21	75.21	75.00	73.75	72.29
0.6	73.75	75.00	75.42	75.83	76.46	75.21	73.54
0.7	74.38	74.79	75.83	76.67	77.08	76.04	74.58
0.8	74.38	75.21	76.04	77.08	76.67	76.04	75.21
0.9	74.38	75.21	76.04	76.25	76.88	75.83	75.00

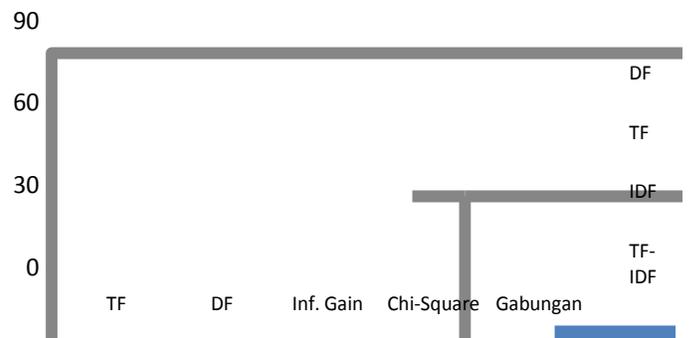
Dari tabel tersebut dapat dilihat terdapat beberapa pasangan nilai parameter yang memberikan akurasi paling baik sebesar 83,75% dan 77,08%. Adapun pasangan nilai terbaik yaitu ($C=0.8, \gamma=0.8$) untuk klasifikasi bentuk keluhan dan ($C=0.7, \gamma=0.5$) untuk klasifikasi jenis keluhan. Pasangan nilai parameter tersebut akan digunakan pada tahap selanjutnya untuk menguji tingkat akurasi klasifikasi SVM terhadap data uji (*testing*).

B. Hasil Akurasi Beberapa Metode Ekstraksi Fitur

Untuk menentukan metode mana yang paling baik dalam hal klasifikasi, maka akan dilakukan pengujian pada setiap metode untuk melihat tingkat akurasi terbaik yang dapat dihasilkan. Adapun hasil perbandingan akurasi dari tiap metode sebagai berikut.

Tabel 5. Perbandingan tingkat akurasi dari beberapa metode ekstraksi bentuk keluhan

Format Vektor	Metode Ekstraksi				
	TF	DF	Inf.Gain	Chisquare	TF+IG+Chi
DF	82.50	80.00	79.17	78.33	83.33
TF	79.17	72.50	79.17	77.50	79.17
IDF	68.33	65.00	64.17	66.67	70.83
TF-IDF	76.67	72.50	74.17	74.17	78.33



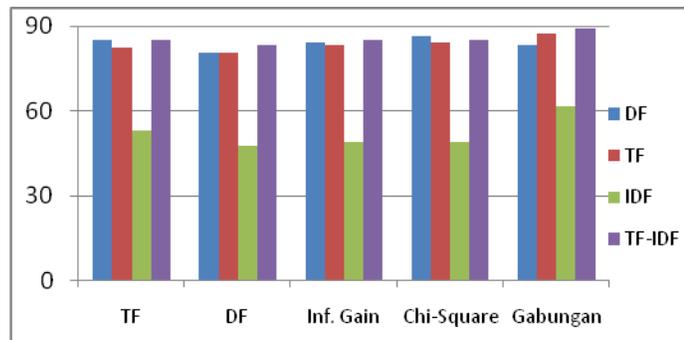
Gambar 1. Grafik Perbandingan Akurasi Bentuk Keluhan

Berdasarkan grafik perbandingan akurasi dari tiap metode ekstraksi untuk proses klasifikasi bentuk keluhan, terlihat bahwa nilai akurasi tertinggi dihasilkan oleh metode *term fequency* (TF) sebesar 82,50%. Metode penggunaan gabungan *feature* dari metode TF, *information gain*, dan *chi-sqaure* dapat meningkatkan akurasi menjadi 83,33%. Sedangkan perbandingan akurasi tiap metode untuk klasifikasi jenis keluhan dapat dilihat sebagai berikut.

Tabel 6.

Perbandingan tingkat akurasi dari beberapa metode ekstraksi jenis keluhan

Format Vektor	Metode Ekstraksi				
	TF	DF	Inf.Gain	Chisquare	TF+IG+Chi
DF	85.00	80.83	84.17	86.67	83.33
TF	82.50	80.83	83.33	84.17	87.50
IDF	53.33	47.50	49.17	49.17	61.67
TF-IDF	85.00	83.33	85.00	85.00	89.17



Gambar 2. Grafik perbandingan akurasi jenis keluhan

Berdasarkan grafik perbandingan akurasi dari tiap metode ekstraksi untuk proses klasifikasi jenis keluhan, terlihat bahwa nilai akurasi terbaik dihasilkan oleh metode ekstraksi *chi-square* sebesar 86,67%. Sedangkan metode penggunaan gabungan *feature* dari metode TF, *information gain*, dan *chi-sqaure* dapat meningkatkan akurasi menjadi 89,17% atau mengalami kenaikan sebesar 2,5%.

C. Analisa Hasil Pengujian dan Kinerja Klasifikasi

Jumlah data uji yang digunakan sebanyak 20% dari total *dataset* atau sebanyak 120 data untuk setiap proses klasifikasinya. Guna mengevaluasi kinerja sistem dalam hal klasifikasi, maka akan digunakan tiga buah parameter yaitu *precision*, *recall*, dan *F-measure*. Adapun kinerja sistem berdasarkan parameter yang telah ditentukan.sebagai berikut:

Tabel 7.

Kinerja Sistem Untuk Proses Klasifikasi Bentuk Keluhan

Kateogri	Precision (%)	Recall (%)	F-measure (%)
Keluhan	87.04	78.33	82.46
Bkn Keluhan	80.30	88.33	84.12
Rata-Rata	83.67	83.33	83.29

Tabel 8.

Kinerja Sistem Untuk Proses Klasifikasi Jenis Keluhan

Kateogri	Precision (%)	Recall (%)	F-measure (%)
Billing	81.25	86.67	83.87
Pemasangan	81.25	86.67	83.87
Putus	100	90.00	94.74
Lambat	96.55	93.33	94.91
Rata-Rata	89.76	89.17	89.34

Dari tabel tersebut, dapat kita lihat kinerja sistem berdasarkan tiga parameter (*precision*, *recall*, dan *F-measure*) menghasilkan nilai di atas 50%. Hal ini menunjukkan kinerja sistem sudah berjalan dengan baik dalam hal klasifikasi. Sedangkan untuk kesalahan klasifikasi, hal ini disebabkan di dalam suatu data memiliki banyak *feature* yang merepresentasikan lebih dari satu kelas.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Penelitian ini telah berhasil membuat suatu *tools* atau aplikasi yang mampu mengklasifikasikan bentuk dan jenis keluhan berdasarkan *tweet* menggunakan metode SVM dengan kernel *Gaussian RBF*. Aplikasi mampu membangun *dataset* dari kumpulan *tweet* yang masuk ke akun @SpeedyTelkomsel. *Dataset* yang digunakan sebanyak 600 *tweet*, dimana 480 *tweet* digunakan sebagai data latih untuk membangun model sedangkan sisanya 120 *tweet* digunakan sebagai data uji untuk mengukur akurasi dari model yang telah dibangun.

Tools atau aplikasi mampu melakukan serangkaian proses *preprocessing* sebagai tahapan persiapan masukan data yang meliputi pelabelan *dataset*, *tokenizing*, dan *stemming*. *Tools* atau aplikasi mampu mendapatkan daftar kata yang digunakan sebagai *feature* dengan menggunakan metode ekstraksi yaitu *term frequency* (TF), *document frequency* (DF), *information gain*, *chi-square*, dan penggabungan dari keempat metode tersebut.

Pasangan nilai parameter *C* dan γ terbaik yang dihasilkan pada penelitian kali ini adalah ($C=0.8$, $\gamma=0.8$) untuk klasifikasi bentuk keluhan dan ($C=0.7$, $\gamma=0.5$) untuk klasifikasi jenis keluhan.

Berdasarkan pengujian terhadap data uji dengan membandingkan tiap metode ekstraksi *feature*, didapat metode *term fequency* (TF) menghasilkan akurasi paling baik sebesar 82,50% untuk klasifikasi bentuk keluhan. Sedangkan untuk klasifikasi jenis keluhan, metode *chi-square* menghasilkan akurasi paling baik sebesar 86,67%. Penggabungan *feature* yang dihasilkan dari metode TF, *information gain*, dan *chi-sqaure* dapat meningkatkan akurasi menjadi 83,33% untuk klasifikasi bentuk keluhan dan 89,17% untuk klasifikasi jenis keluhan.

Evaluasi kinerja sistem dalam hal klasifikasi dapat dilihat berdasarkan tiga parameter yaitu *precision*, *recall*, dan *f-measure*. Untuk klasifikasi bentuk keluhan, rata-rata nilai yang dihasilkan untuk setiap parameter yaitu 83,67%, 83,33%, dan 83,29%. Sedangkan untuk klasifikasi jenis keluhan, rata-rata nilai yang dihasilkan 89,76%, 89,17%, dan 89,34%. Dari nilai tersebut dapat dilihat kinerja sistem dalam hal klasifikasi sudah cukup baik.

B. Saran

Adapun beberapa saran yang terkait untuk penelitian selanjutnya adalah mengkombinasikan penggunaan *feature* yang didapat secara otomatis dari beberapa metode ekstraksi dengan *feature* yang ditentukan secara manual. Selain itu dengan meningkatkan koleksi data latih sehingga memungkinkan dapat meningkatkan akurasi dari kinerja sistem dalam hal klasifikasi. Pada penelitian ini kalimat yang akan diklasifikasikan dipandang sebagai *bag of words* atau sekumpulan kata-kata. Faktor yang berpengaruh adalah frekuensi kemunculan kata Kedepannya diharapkan dapat diteliti pengklasifikasian kalimat yang juga memperhitungkan faktor susunan kata-kata yang dapat dipisahkan dalam *subject*, *predicate*, dan *object* serta penanganan frase.

DAFTAR PUSTAKA

- [1] Zhang, G., Zhou F., and Lan, Y. Customer Value of Sosial Network Service Website: Key Components and Impacts on Customer Loyalty. Central South University, Changsha, P.R. China. 2010.
- [2] Aberdeen Group. Sosial Media and Customer Survey: From Listening to Engagement (Analyst Insight). 2012.
- [3] Almadhoun, N. M., Dominic, P. D. D., and Woon, L. F. Sosial Media as a Promotional Tool in Higher Education in Malaysia. National Postgraduate Conference (NPC). 2011.
- [4] Mousavi, S. , Demirkan, H. The Key to Sosial Media Implementation: Bridging Customer Relationship Management to Sosial Media. 46th Hawaii International Conference on System Sciences. 2013.
- [5] Kearney, A. T. Sosial Media: Are You Part Of Conversation. 2012.
- [6] Naradhipa, A., R., dan Purwarianti, P., Sentiment Classification for Indonesian Message in Sosial Media, International Conference on Electrical Engineering and Informatics, 2011.
- [7] Arbitya, P.P. Ekspresi Afeksi Dalam Twitter Studi Pada Remaja Followers Akun @Soalcinta. Fakultas Ilmu Sosial dan Politik. Universitas Indonesia. 2012.
- [8] Avaya. *Video dan Media Sosial Jadi Saluran Komunikasi Pelayanan Pelanggan*. Retrieved 2014, from : <http://mix.co.id/research/video-dan-media-sosial-jadi-saluran-komunikasi-pelayanan-pelanggan/>
- [9] Muflikha, L. Ridok, A., Hardono. J. Klasifikasi Kondisi Penderita Penyakit Hepatitis Dengan Menggunakan Metode Support Vector.
- [10] Hsu, C.W., Chang, C.C., Lin, C.J. A Practical Guide to Support Vector Classification. Departement of Computer Science National Taiwan University, Taipei 106, 2010.
- [11] Zhichao, L. *The Impact of Social Networks within Service Teams on Customer Satisfaction*. International Conference on Information Management, Innovation Management and Industrial Engineering. 2012.
- [12] Zhiheng, X., Rong, Lu., and Liang X. Discovering User Interest on Twitter with a Modified Author-Topic Model. International Conferences on Web Intelligence and Intelligent Agent Technology. 2011.
- [13] Finin, T.,Iyengar A., and Joshi, A. Content-based prediction of temporal boundaries for events in Twitter. IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Sosial Computing. 2011.
- [14] Wulandini, F., dkk. A Study on Text Classification for Webmining Based Spatio Temporal Analysis of the Spread of Tropical Diseases. Proc. of International Conference on Advance Computer Science & Information System (ICACSIS), 2010.
- [15] Watters, C., Basu, A., and Shepherd, M. Support Vector Machines for Text Categorization. Proceedings of the 36th Hawaii International Conference on System Sciences. Faculty of Computer Science. Dalhousie University. 2002.
- [16] Meesad, P., Boonrawd, P., Nuipian, V. A Chi-Square-Test for Word Importance Differentiation in Text Classification. International Conference on Information and Electronics Engineering. Singapore. 2011.