

Klasifikasi Keluhan Menggunakan Metode *Support Vector Machine* (SVM) (Studi Kasus : Akun Facebook Group iRaise Helpdesk)

Fatmawati¹, Muhammad Affandes²

^{1,2}Teknik Informatika UIN Sultan Syarif Kasim Riau

Jl. H.R. Soebrantas no. 155 KM. 18 Simpang Baru, Pekanbaru 28293

fatmawati11@students.uin-suska.ac.id¹, affandes@uin-suska.ac.id²

Abstrak - Facebook Group iRaise Helpdesk merupakan salah satu layanan media sosial yang digunakan pihak PTIPD UIN Suska Riau sebagai layanan pelanggan (*customer services*) sistem akademik. Mengingat sistem akademik baru mengalami peralihan yang sebelumnya bernama SIMAK menjadi iRaise, sehingga masih ada permasalahan yang ditimbulkan, dan menjadi keluhan bagi penggunaannya. Untuk pengolahan data keluhan, pihak PTIPD masih menggunakan proses manual dengan menggunakan *microsoft word* dan *excel*. Sehingga pada penelitian ini akan dilakukan pengklasifikasian permasalahan sistem iRaise pada kategori *multiclass* yaitu: login, krs, nilai dan personal. Dengan menggunakan metode *Support Vector Machine* (SVM) dengan kernel RBF. Jumlah dataset sebanyak 1040 data keluhan. Pengujian dilakukan menggunakan aplikasi RapidMiner dan diuji dengan menggunakan *10-Fold cross validation* dan diukur dengan *confussion matrix*. Dari hasil uji coba aplikasi menunjukkan akurasi tertinggi sebesar 95.67% pengujian tanpa menggunakan *feature selection* pada titik $C=2$ dan $\gamma = 0.09$.

Kata kunci - *confussion matrix*, *cross validation*, *iraise*, keluhan, klasifikasi, *rapidminer*, *support vector machine*

PENDAHULUAN

Pusat Teknologi Informasi dan Pangkalan Data (PTIPD) Universitas Islam Negeri Sultan Syarif Kasim Riau (UIN Suska Riau) memanfaatkan media sosial sebagai *customer service* bagi mahasiswa, dosen maupun pegawai yang ingin menyampaikan keluhan seputar permasalahan sistem akademik. Mengingat sistem akademik, baru mengalami masa peralihan yang sebelumnya bernama SIMAK menjadi iRaise, sehingga masih ada permasalahan - permasalahan dalam penggunaannya. Berdasarkan fakta yang terjadi di lapangan pada layanan *Customer Care Center* (C3) PTIPD, C3 telah membuka layanan keluhan via langsung dengan membuka nomor antrian kurang lebih 100 antrian setiap hari-nya dengan berbagai jenis keluhan

mengenai permasalahan-permasalahan akademik, akan tetapi berdasarkan *presentase* 75% lebih ke permasalahan sistem iRaise dilihat dari rekap laporan harian C3.

Untuk mengurangi perpanjangan antrian layanan dilapangan, pihak PTIPD membagi tugas layanan lewat media sosial, salah satunya pada Facebook Group dengan alamat iRaise Helpdesk bermanfaat bagi kalangan mahasiswa dan pegawai untuk menyampaikan sebuah permasalahan seputar permasalahan terhadap iRaise.

Untuk penanganan pengolahan data lewat media sosial, pihak PTIPD masih menggunakan proses manual dalam menentukan data bentuk keluhan, dengan menggunakan *microsoft word* dan *excel*.

Sehingga penelitian ini akan melakukan suatu proses *data mining* untuk klasifikasi keluhan mengenai permasalahan iRaise pada akun Facebook Group iRaise Helpdesk dengan menggunakan metode *Support Vector Machine* (SVM).

LANDASAN TEORI

A. iRaise

Integrated Academic Information System (iRaise) merupakan sebuah sistem informasi yang berfungsi sebagai integrator informasi akademik yang ada di berbagai unit akademik (program studi/ fakultas) sekaligus sebagai sarana komunikasi antar civitas akademika kampus. Sistem iRaise yang ada pada UIN Suska Riau memungkinkan mahasiswa untuk melihat dan mengolah data-data perkuliahan mereka. Data-data tersebut antara lain berupa biodata mahasiswa, daftar mata kuliah, jadwal kuliah, informasi transkrip nilai, informasi KHS (Kartu Hasil Studi), *download* tugas, informasi pengisian KRS (Kartu Rencana Studi).

B. Text Preprocessing

Text preprocessing merupakan tahapan sangat penting dalam melakukan proses klasifikasi data teks. Tujuan dilakukannya *text preprocessing* yaitu untuk menghilangkan *noise*, menyeragamkan bentuk kata dan mengurangi *volume* kata.^[1]

Data yang masuk sebelum diolah harus terlebih dahulu melalui tahapan *preprocessing* agar sistem mengerti sebelum melanjutkan proses *text minning*

pada data tersebut. Berikut tahapan dari *text preprocessing* [2]:

- a) *Cleaning*
- b) *Tokenizing (Tokenization)*
- c) *Case Folding*
- d) *Spelling Normalization*
- e) *Filtering*
- f) *Stemming*
- g) *Feature Selection*
- h) *Pembobotan (Term Weighting)*

[1] Pelabelan Data

Didalam permasalahan klasifikasi hal yang paling utama adalah melakukan pelabelan dengan benar agar data yang terkumpul atau terklasifikasi sesuai dengan kelas yang seharusnya. Pada penelitian ini akan dilakukan pelabelan terhadap postingan keluhan. Untuk melakukan pelabelan ini menurut [3], definisi keluhan (*complain*) merupakan sebuah kata yang sering berkonotasi negatif, pada umumnya dipersepsikan sebagai kesalahan, masalah, *stres*, frustrasi, kemarahan, konflik, hukuman, tuntutan, ganti rugi dan sejenisnya. *Complain* merupakan bagian dari proses yang berawal dari konsumen yang merasakan ketidakpuasan setelah menerima pelayanan atau melakukan transaksi. Pada penelitian ini keluhan dibagi menjadi 4 kategori yaitu:

- a) Login
- b) KRS
- c) Nilai
- d) Personal

C. Support Vector Machine (SVM)

SVM adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. SVM memiliki prinsip dasar *linier classifier* yaitu kasus klasifikasi yang secara *linier* dapat dipisahkan, namun SVM telah dikembangkan agar dapat bekerja pada problem *non-linier* dengan memasukkan konsep *kernel* pada ruang kerja berdimensi tinggi. Fungsi *kernel* yang digunakan untuk memetakan dimensi awal (dimensi yang lebih rendah) himpunan data ke dimensi baru (dimensi yang relatif lebih tinggi).^[4]

Menurut [5] macam fungsi kernel diantaranya:

1. kernel linier :

$$K(u, v) = uv^T \quad (1)$$

2. kernel polinomial

$$K(u, v) = (1 + uv^T)^d, \quad d \geq 2 \quad (2)$$

3. kernel RBF (Radial Basis Function) :

$$K(u, v) = \exp(-\gamma \|u - v\|^2), \quad \gamma > 0 \quad (3)$$

4. Kernel Gaussian : :

$$K(u, v) = \exp\left(-\frac{\|u - v\|^2}{2\sigma^2}\right) \quad (4)$$

Menurut (Santosa, 2007) *hyperplane* klasifikasi *linier* SVM dinotasikan:

$$f(x) = w^T \cdot x + b \quad (5)$$

Sehingga diperoleh persamaan:

$$[(w^T \cdot x_i) + b] \geq 1 \text{ untuk } y_i = +1$$

$$[(w^T \cdot x_i) + b] \leq -1 \text{ untuk } y_i = -1$$

Dengan x_i = himpunan data *training*, $i=1,2,\dots,n$ dan y_i = label kelas dari x_i untuk mendapatkan *hyperplane* terbaik adalah dengan mencari *hyperplane* yang terletak ditengah-tengah antara dua bidang pembatas kelas.

Mencari *hyperplane* terbaik dapat digunakan metode *Quadratic Programming (QP) Problem* yaitu meminimalkan $\frac{1}{2}w^T w$, dengan menggunakan fungsi *Lagrange Multiplier* yang telah ditransformasi sebagai berikut:^[4]

$$L(w, b, \alpha) = \frac{1}{2}w^T w - \sum_{i=1}^n \alpha_i y_i (w^T \cdot x_i) - b \sum_{i=1}^n \alpha_i y_i + \sum_{i=1}^n \alpha_i \quad (5)$$

berdasarkan persamaan 5, maka persamaan 6 menjadi sebagai berikut:

$$L_d = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (6)$$

dan diperoleh dual problem:

$$\max_d L_d = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j$$

Dengan batasan $\alpha_i \geq 0, i = 1, 2, \dots, n$ dan $\sum_{i=1}^n \alpha_i y_i = 0$ data *training* dengan $\alpha_i \geq 0$ terletak pada *hyperplane* disebut *support vector*, data *training* yang tidak terletak pada *hyperplane* tersebut mempunyai $\alpha_i = 0$. Setelah solusi permasalahan *quadratic programming* ditemukan (nilai α_i), maka kelas dari data yang akan di prediksi atau data testing dapat ditentukan berdasarkan nilai fungsi tersebut.^[4]

5. Penelitian terkait

Beberapa penelitian klasifikasi *text* yang pernah dilakukan (Tegar, 2013) melakukan pengklasifikasian topik dan analisis sentimen dalam media sosial dengan metode SVM. Penelitian (Maya, 2014) melakukan klasifikasi terhadap dokumen *E-Complaint* kampus dengan menggunakan *Directed Acyclic Graph Multy-Class SVM* menghasilkan akurasi sebesar 82.61%, dengan menggunakan *kernel Gaussian RBF* dan *stemming* dalam 70% data latih. (Enda, 2014) penelitian ini menunjukkan bahwa SVM mampu melakukan klasifikasi keluhan dengan baik, hal ini dibuktikan dengan tingkat akurasi klasifikasi sebesar 83,33% untuk bentuk keluhan dan 89,17% untuk jenis keluhan.

ANALISA

A. Analisa Data

Pada penelitian ini data yang digunakan bersumber dari postingan pelanggan pada akun Facebook Group iRaise Helpdesk. Data yang diambil adalah data postingan yang hanya mengandung kalimat keluhan mengenai permasalahan sistem iRaise.

Selanjutnya data keluhan dilakukan pelabelan secara manual dengan 4 kelas kategori yaitu: Login, KRS, Nilai, dan Personal. dengan dataset berjumlah 1040 data yang berisi 260 data pada masing-masing kelas keluhan, yaitu 260 pada kelas login, 260 kelas krs, 260 nilai dan 260 kelas personal.

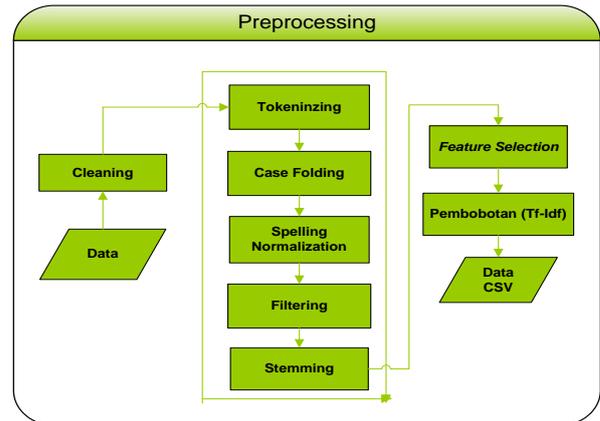
Berikut hasil data keluhan yang telah diberi label seperti tabel bawah ini:

Tabel 3 Keluhan Sistem iRaise

No	Keluhan	Kategori
1	Nama: Riswan Hasibuan, NIM: 11521105581.E-mail: Riswan.hasber05@gmail.com Kenapa iRaise saya tidak bisa dibuka? Selalu saja muncul tulisan error.	Login
2	Pas saya mengisi krs untuk smster 4 kok matakuliah nya gx ada timbul satu pun	KRS
3	Saya setiap ngisi audit, selalu kembali ke audit yang sudah di isi, tidak ke audit selanjutnya.	Personal
4	nama: aila risda wati, fak: fekonsos, jurusan: akuntansi Nilai saya sudah ada beberapa yg keluar tpi bobot nya masih nol.	Nilai
...		...
1040	knapa sewaktu saya login ke iraise, ada pemberitahuan pengajuan cuti anda sedang di proses, sedangkan saya tidak pernah mengajukan cuti.	Personal

B. Analisa Text Preprocessing

Text Preprocessing merupakan langkah penting dalam melakukan analisa klasifikasi keluhan yang bertujuan untuk membersihkan data dari unsur-unsur yang ada yang tidak dibutuhkan. Berikut ini merupakan tahapan *preprocessing*:



Gambar 2 Tahapan *Preprocessing*

Berikut ini penjelasan dari tahapan *text preprocessing* sebagai berikut:

1. *Cleaning*

Adapun kata atau karakter yang dihilangkan adalah identitas pengguna seperti: alamat url (*email.com*), *username*, *nim*, *fakultas*, *jurusan*. Proses *cleaning* dengan menghilangkan identitas pengguna dilakukan untuk membersihkan data dari kata yang tidak penting dan tidak mempengaruhi proses klasifikasi. Berikut proses *cleaning* data keluhan:

Tabel 2 proses *Cleaning*

No	Keluhan	Kategori
1	Kenapa iRaise saya tidak bisa dibuka? Selalu saja muncul tulisan error.	Login
2	Pas saya mengisi krs untuk smster 4 kok matakuliah nya gx ada timbul satu pun	KRS
3	Saya setiap ngisi audit, selalu kembali ke audit yang sudah di isi, tidak ke audit selanjutnya.	Personal
4	Nilai saya sudah ada beberapa yg keluar tpi bobot nya masih nol.	Nilai

2. *Tokenizing*

Tokenizing yaitu proses melakukan pemecahan terhadap kalimat data keluhan menjadi sebuah kata tunggal. Pada saat bersamaan, tokenisasi juga membuang beberapa angka, karakter tertentu yang dianggap sebagai tanda baca. contoh berikut proses *tokenizing* :

Tabel 3 proses *Tokenizing*

Keluhan 1	Keluhan 2	Keluhan 3	Keluhan 4
Kenapa iRaise saya tidak bisa dibuka Selalu muncul tulisan error	pas saya mengisi krs untuk smster kok matakuliah nya gx ada timbul satupun	Saya setiap ngisi audit yang sudah di isi tidak ke audit selanjutnya	Nilai saya sudah ada beberapa yg keluar tpi bobot nya masih nol

3. *Case folding*

Case folding merupakan proses penyeragaman bentuk huruf pada dokumen dengan mengubah huruf besar menjadi huruf kecil atau sebaliknya.

4. *Spelling Normalization*

Merupakan proses perbaikan kata yang tidak sesuai dengan penulisan kata yang sebenarnya misalnya “gk” diubah menjadi “tidak”.

Tabel 4 Proses *Spelling Normalization*

Keluhan 1	Keluhan 2	Keluhan 3	Keluhan 4
kenapa iraise saya tidak bisa dibuka selalu muncul tulisan rusak:erro r	saat:pas saya mengisi krs untuk semester: smster kok matakulia h nya tidak:gx ada timbul satupun	saya setiap isi:ngisi audit yang sudah di isi tidak ke audit lanjut:selan jutnya	nilai saya sudah ada beberapa yang:yg keluar tapi:tpi bobot nya masih nol

5. *Filtering*

Filtering adalah tahap mengambil kata-kata penting dari hasil token.

Tabel 5 Proses *Filtering*

<i>Filtering</i>			
iraise	tulisan	matakuliah	Nilai
tidak	rusak	audit	Keluar
muncul	mengisi	selanjutnya	Bobot
dibuka	krs	semester	Nol

6. *Stemming*

Pada tahap ini kata berimbuhan akan diubah dalam bentuk kata dasar dengan menggunakan algoritma Nazief dan Adriani berdasarkan Kamus Besar Bahasa Indonesia (KBBI).

Tabel 6 Proses *Stemming*

<i>Stemming</i>			
iraise	tulis	matakuliah	nilai
Tidak	rusak	audit	luar
muncul	isi	lanjut	bobot
Buka	krs	semester	nol

7. Ekstraksi *Feature*

Pada penelitian ini pemilihan *feature selection* dilakukan dengan 3 cara ekstraksi *feature* yaitu pemilihan *feature* ekstraksi *threshold*, *ranking selection threshold up*, *Ranking selection threshold down*.

8. Pembobotan

Merupakan proses merubah kata menjadi bentuk *vektor*, dimana tiap kata dihitung sebagai satu fitur, dengan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF).

Tabel 7 Proses TF-IDF

kata	TF				DF	IDF = $\log(n/df)$ $n=6$
	Data 1	Data 2	Data 3	Data 4		
iraise	1	0	0	0	1	0.6020
Rusak	1	0	0	0	1	0.6020
Matakuliah	0	1	0	0	1	0.6020
Nilai	0	0	0	1	1	0.6020
bobot	0	0	0	1	1	0.6020
Krs	0	1	0	0	1	0.6020
nol	0	0	0	1	1	0.6020
audit	0	0	3	0	1	0.6020

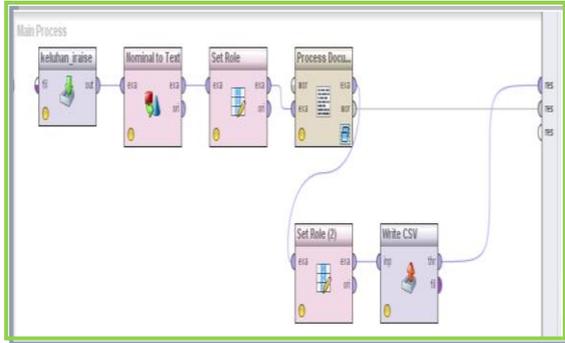
Tabel 8 Hasil TF-IDF

kata	TF-IDF=(TFxIDF)			
	Data 1	Data 2	Data 3	Data 4
iraise	0.6020	0	0	0
Rusak	0.6020	0	0	0
Matakuliah	0	0.6020	0	0
Nilai	0	0	0	0.6020
bobot	0	0	0	0.6020
Krs	0	0.6020	0	0
Nol	0	0	0	0.6020
audit	0	0	1.806	0

PENGUJIAN

Pengujian pada penelitian ini menggunakan *tools* RapidMiner untuk menghasilkan suatu *output* klasifikasi dan keakurasian data.

Hal pertama dilakukan adalah memilih file dataset berjumlah 1040 data keluhan berlabel yang akan dieksekusi dan dilakukan tahapan *preprocessing*. Proses ini meliputi tahapan membaca data dari *database*, *nominal to text*, *set role*, *process document*, *set role* dan terakhir menyimpan data dalam bentuk format CSV. Berikut tampilan proses *text preprocessing* data pada RapidMiner.



Gambar 2 Tahapan Text Preprocessing

Setelah seluruh proses *text preprocessing* dilakukan, dengan hasil kata di ubah dalam bentuk *vektor* dengan menghitung TF-IDF dan di simpan dalam bentuk CSV. Berikut tampilan hasil data dalam bentuk nilai *vektor*:

adab	administrasi	agama	agroteknologi	ajar	aju	ajuk	akademik
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0.272
0	0	0	0	0	0	0.490	0
0	0	0	0	0	0	0	0

Gambar 3 Data CSV

1. Klasifikasi Data

Pada tahapan klasifikasi ada dua proses umum yang harus dilakukan yaitu: *Read CSV*, dan *Validation model*.

a. Read CSV

data yang memiliki bobot nilai hasil *text preprocessing* dalam bentuk CSV dipanggil untuk dilakukan proses *validation*. Dengan cara membaca data *file* CSV dengan memilih menu *import configuration wizard*.

b. Validation Model

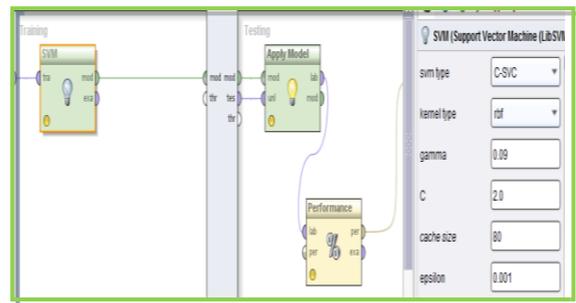
Pada proses *validation* semua fungsi, *training* dan *testing* akan dilakukan. Fungsi yang digunakan pada penelitian ini yaitu fungsi LibSVM dengan

kernel RBF dengan fungsi nilai (c, γ) , dan menggunakan *10-fold cross validation*.

nilai pada parameter C dan γ berdasarkan rentang nilai:

- $100 \leq c \leq 1000$ dan $100 \leq \gamma \leq 1000$
- $10 \leq c \leq 100$ dan $10 \leq \gamma \leq 100$
- $1 \leq c \leq 10$ dan $1 \leq \gamma \leq 10$
- $0 \leq c \leq 1$ dan $0.01 \leq \gamma \leq 1$
- $0 \leq c \leq 3$ dan $0.01 \leq \gamma \leq 10$

Berikut tampilan dari proses *validation* berisi pengujian dengan proses training fungsi LibSVM dan *Testing* dengan parameter C dan γ .



Gambar 4 Sub Proses Validation

Berikut tampilan hasil klasifikasi data dengan pengujian parameter c dan γ :

PerformanceVector			
PerformanceVector:			
accuracy: 95.67% +/- 1.16% (mikro: 95.67%)			
ConfusionMatrix:			
True:	Login	KRS	Nilai
Login:	249	6	0
KRS:	5	248	5
Nilai:	0	6	248
Personal:	6	0	7
classification_error: 4.33% +/- 1.16% (mikro: 4.33%)			
ConfusionMatrix:			
True:	Login	KRS	Nilai
Login:	249	6	0
KRS:	5	248	5
Nilai:	0	6	248
Personal:	6	0	7
absolute_error: 0.043 +/- 0.012 (mikro: 0.043 +/- 0.203)			
relative_error: 4.33% +/- 1.16% (mikro: 4.33% +/- 20.35%)			

Gambar 5 Hasil Klasifikasi Data

Pada penelitian ini untuk melihat tingkat akurasi terbaik klasifikasi dalam menentukan kelas suatu data berdasarkan hasil algoritma, maka dilakukan beberapa bentuk model pengujian dengan ekstraksi *feature* dengan jumlah *feature* yang berbeda. pengujian juga menggunakan dua hasil bobot *feature* yaitu berdasarkan hasil kamus *stemming* manual dan algoritma nazief dan Adriani. Berikut tabel pengujian Ekstraksi *feature*:

Tabel 1 Model Pengujian

Model Pengujian	
Ekstrasi Feature	Jumlah Feature
Tanpa Feature selection	1-590

Model Pengujian	
Ekstrasi Feature	Jumlah Feature
Feature Selection	
Ekstrasi Threshold	13-24
Ranking selection threshold down	1-20
Ranking selection threshold up	25-590

Berikut proses Pengujian dilakukan dengan cara *grid search*, dengan *10-fold cross validation*, Pengujian dilakukan berdasarkan *vector* dari hasil proses *stemming* manual dan algoritma nazief dan adriani.

Y \ C	0.01	0.03	0.06	0.09	0.1	0.3	0.6	0.9	1	3
0.0	46.63 %	88.56 %	92.79 %	92.31 %	91.92 %	92.79 %	92.98 %	91.92 %	91.35 %	55.38 %
0.1	15.19 %	62.21 %	72.40 %	78.27 %	81.06 %	89.42 %	89.52 %	88.27 %	86.73 %	34.62 %
0.2	15.19 %	62.21 %	73.85 %	78.27 %	80.77 %	89.42 %	89.52 %	88.27 %	86.73 %	34.62 %
0.3	15.29 %	62.22 %	73.85 %	83.46 %	83.00 %	89.42 %	89.52 %	88.27 %	86.73 %	34.62 %
0.6	15.29 %	88.37 %	92.79 %	92.69 %	91.92 %	90.58 %	89.52 %	88.27 %	86.73 %	34.62 %
0.9	38.27 %	88.27 %	92.69 %	92.02 %	91.92 %	91.15 %	92.50 %	90.77 %	90.00 %	34.62 %
1	46.63 %	88.56 %	92.79 %	92.31 %	91.92 %	92.12 %	92.98 %	91.73 %	91.35 %	55.38 %
2	87.60 %	90.58 %	93.37 %	95.67 %	95.00 %	92.79 %	92.98 %	91.92 %	91.54 %	58.65 %
3	90.00 %	90.58 %	94.13 %	94.81 %	94.62 %	92.79 %	92.98 %	91.92 %	91.54 %	58.65 %
6	92.21 %	93.37 %	94.42 %	94.81 %	94.62 %	92.79 %	92.98 %	91.92 %	91.54 %	58.65 %

Gambar 6 Bentuk Pengujian grid Search

HASIL DAN PEMBAHASAN

Berdasarkan pengujian yang telah dilakukan dengan menggunakan beberapa jenis pengujian ekstraksi *feature* dengan proses *grid search* dengan parameter C dan γ , maka didapat hasil rekapitulasi keseluruhan dapat dilihat pada tabel berikut:

Tabel 2 Hasil Pengujian

Hasil Pengujian		
Ekstrasi Feature	Kamus Algoritma Nazief dan Adriani	Kamus Manual
Tanpa Feature Selection	95.67%	95.48%
Feature Selection		
-Ekstrasi threshold	91.15%	91.06%
-Rankingselection threshold down	91.54%	91.44%
-Rankingselection threshold up	95.48%	94.90%

1. Perbandingan hasil prediksi

Berikut perbandingan hasil prediksi klasifikasi antara algoritma SVM dengan prediksi klasifikasi secara manual. pengujian berdasarkan dataset berjumlah 1040 data. Ketepatan klasifikasi sebanyak 995 data, dan klasifikasi *error* sebanyak 45 data.

Row No.	No	Kategori	confidence(login)	confidence(KRS)	confidence(nilai)	confidence(lain-lain)	prediction(kategori)
1	18	lain-lain	0	0	0	1	lain-lain
2	51	nilai	0	1	0	0	KRS
3	53	nilai	0	0	1	0	nilai
4	64	login	0	0	0	1	lain-lain
5	76	lain-lain	0	0	0	1	lain-lain
6	81	login	1	0	0	0	login
7	87	KRS	0	1	0	0	KRS
8	91	nilai	0	0	1	0	nilai
9	96	lain-lain	0	0	0	1	lain-lain
10	106	KRS	0	1	0	0	KRS
11	111	nilai	0	0	1	0	nilai
12	116	lain-lain	0	0	0	1	lain-lain
13	126	KRS	0	1	0	0	KRS
14	154	nilai	0	0	1	0	nilai
15	164	login	1	0	0	0	login

Gambar 1 Perbandingan Hasil Prediksi

KESIMPULAN DAN SARAN

1. Kesimpulan

- Algoritma SVM berhasil diterapkan untuk melakukan klasifikasi data dalam kategori *multiclass* terhadap keluhan sistem iRaise berdasarkan postingan pada akun Facebook iRaise Helpdesk.
- Penelitian ini berhasil mendapatkan nilai akurasi tertinggi sebesar 95.67% terdapat pada titik $c=2$, $\gamma=0.09$ dengan pengujian data tanpa *Feature Selection* dengan Algoritma Nazief dan Adriani, dan untuk data yang sudah dilakukan *feature selection* mencapai nilai akurasi tertinggi sebesar 95.48% pada titik $c=3$, $\gamma=0.09$.
- Berdasarkan hasil pengujian dengan menggunakan dataset 1040 data, didapat hasil perbandingan klasifikasi data antara prediksi manual dengan prediksi model, maka hasil ketepatan prediksi yang dibangun model sebesar 995 data, dan tingkat kesalahan prediksi sebanyak 45 data.

2. Saran

- Penelitian ini dapat dikembangkan lagi dengan menambah data permasalahan keluhan akademik dan kelas kategori keluhan lainnya.
- Diharapkan untuk penelitian selanjutnya dapat dikembangkan dengan membuat sistem aplikasi yang dapat terintegrasi dengan Facebook secara langsung sehingga dapat membantu instansi dalam proses penanganan keluhan lebih efektif dan efisien.

REFERENSI

- [1] Berry, M.W. & KOGan, J. (2010). *Text Mining Application and Theory*. Wiley. United Kingdom
- [2] Feldman, R & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press New York
- [3] Wahjono. (2010). *Pengaruh Kualitas Layanan dan Penanganan Keluhan Terhadap Loyalitas Nasabah*. Universitas Negeri Surabaya
- [4] Santosa, Budi. (2007). *Tutorial Support Vector Machine*. ITS, Surabaya
- [5] Prasetyo, Eko. (2012). *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*, Yogyakarta: ANDI
- [6] Enda, Esyudha P. (2014). *Klasifikasi Keluhan Pelanggan Berdasarkan Tweet Dengan Menggunakan Metode Support Vector Machine (SVM)*. Institut Teknologi Bandung.
- [7] Kusrini, Dkk. (2009). *Algoritma Data Mining*. Yogyakarta.
- [8] Muningsih, Elly. (2011). *Facebook Commerce E-Commerce Pada Media Sosial Facebook yang Modern dan Populer*. Amik BSI Yogyakarta.
- [9] Ahmad, Fathan H. (2004). *Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik pada Twitter*. Seminar Nasional Informatika (Semnas IF), (No.ISSN:1979-2328). Yogyakarta.
- [10] Andi, Nurul H. (2015). *Analisis Sentimen Terhadap Wacana Politik Pada Media Masa Online Menggunakan Algoritma Support Vector Machine dan Naive Bayes*. Jurnal Elektronik Sistem Informasi dan Komputer, (ISSN).