

# Text Mining pada Sosial Media untuk Mendeteksi Emosi Pengguna Menggunakan Metode Support Vector Machine dan K-Nearest Neighbour

Dwi Ardiada<sup>1</sup>, Made Sudarma<sup>2</sup>, Dwi Giriantari<sup>3</sup>

[Submission: 29-11-2018, Accepted: 31-03-2019]

**Abstract**— Twitter social networking and microblog services that allow users to send and read text-based messages up to 140 characters, known as tweets. A text on a tweet does not only convey information from an information, but also contains information about human behavior including emotions. To detect emotions from text on twitter social media services with unstructured data, one needs to do text analysis by using Text Mining.

In this study propose to conduct text mining research on Social Media to detect users' emotions. From the tests conducted using the Support Vector Machine and K-Nearest Neighbor method, it can produce an average precision value of 0.4564. The recall value is 0.502 and the accuracy value is 0.8104 while the K-Nearest Neighbor method has an average precision value of 0.3421. The recall value is 0.4595 and the accuracy value is 0.797.

The results of testing with the SVM-KNN method have a significant increase in the value of the K-Nearest Neighbor method. SVM-KNN method also has a significant difference in computational time in doing two classifications rather than the K-Nearest Neighbor method which only does one classification.

**Intisari**— Twitter layanan jejaring sosial dan mikroblog yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter, yang dikenal dengan sebutan kicauan (tweet). Sebuah teks pada tweet tidak hanya menyampaikan keterangan dari suatu informasi, tetapi juga berisi informasi tentang perilaku manusia termasuk emosi. Untuk mendeteksi emosi dari teks pada layanan sosial media twitter dengan data yang tidak terstruktur maka perlu dilakukan analisis teks salah satunya dengan menggunakan Text Mining.

Pada penelitian ini mengusulkan melakukan penelitian text mining pada Sosial Media untuk mendeteksi emosi pengguna. Dari Pengujian yang dilakukan dengan metode Support Vector Machine dan K-Nearest Neighbour dapat menghasilkan nilai rata-rata precision sebesar 0.4564. Nilai recall sebesar 0.502 dan pada nilai accuracy sebesar 0.8104 sedangkan dari metode K-Nearest Neighbour nilai rata-rata precision sebesar 0.3421. Nilai recall sebesar 0.4595 dan pada nilai accuracy sebesar 0.797.

<sup>1</sup>Mahasiswa Pascasarjana, Teknik Elektro Fakultas Teknik Universitas Udayana, Jln. Jalan Kamps Bukit Jimbaran 80361 INDONESIA (telp: 0361-703315; fax: 0361-703315; e-mail: [dwiardiada99@gmail.com](mailto:dwiardiada99@gmail.com))

<sup>2, 3</sup> Dosen, Jurusan Teknik Elektro dan Komputer Fakultas Teknik Universitas Udayana, Jln. Jalan Kampus Bukit Jimbaran 80361 INDONESIA (telp: 0361-703315; fax: 0361-4321; e-mail: [msudarma@unud.ac.id](mailto:msudarma@unud.ac.id), [dayu.giriantari@unud.ac.id](mailto:dayu.giriantari@unud.ac.id), [penulis2@unud.ac.id](mailto:penulis2@unud.ac.id))

Hasil dari pengujian dengan metode SVM-KNN memiliki peningkatan nilai yang cukup signifikan dari metode K-Nearest Neighbour. Metode SVM-KNN juga memiliki selisih waktu

Dwi Ardiada: Text Mining Pada Sosial ...

komputasi yang tidak signifikan dalam melakukan dua kali klasifikasi daripada metode K-Nearest Neighbour yang hanya melakukan satu kali klasifikasi

**Kata Kunci**— Emosi, K-Nearest Neighbour, Support Vector Machine, Twitter.

## I. PENDAHULUAN

Data tidak terstruktur banyak terdapat pada layanan sosial media. Layanan sosial media merupakan penyedia sumber daya yang menyediakan data yang cukup besar. Media sosial banyak menyita perhatian masyarakat karena dianggap dapat menjadi tempat untuk berbagi karya, ide, opini tentang isu-isu yang terjadi secara bebas, dan media untuk mengungkapkan berbagai hal mengenai kehidupan pribadinya. Salah satu media sosial yang banyak digunakan masyarakat adalah Twitter.

Twitter layanan jejaring sosial dan mikroblog yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter, yang dikenal dengan sebutan kicauan (tweet)[1].

Sebuah teks pada tweet tidak hanya menyampaikan keterangan dari suatu informasi, tetapi juga berisi informasi tentang perilaku manusia termasuk emosi. Emosi merupakan keadaan kompleks dari pikiran yang dipengaruhi oleh peristiwa eksternal, perubahan fisiologis, atau hubungan dengan orang lain. Dengan tidak adanya kontak tatap muka untuk mendeteksi ekspresi wajah dan intonasi dalam suara, opsi alternatifnya adalah menguraikan emosi dari teks di layanan sosial media. Studi penelitian pendeteksian emosi telah menyelidiki deteksi emosi dalam prosodi, perubahan keadaan fisiologis, ekspresi wajah dan teks. Namun, ada kekurangan penelitian dalam mendeteksi emosi dari teks dibandingkan dengan area lain dari deteksi emosi [2].

Untuk mendeteksi emosi dari teks pada layanan sosial media twitter dengan data yang tidak terstruktur maka perlu dilakukan analisis teks salah satunya dengan menggunakan Text Mining. Text mining mencoba untuk mengekstrak informasi yang berguna dari sumber data melalui identifikasi dan eksplorasi dari suatu pola menarik. Sumber data berupa sekumpulan dokumen dan pola menarik yang tidak ditemukan dalam bentuk database record, tetapi dalam data teks yang tidak terstruktur.

Beberapa penelitian mengenai deteksi emosi telah dilakukan contohnya pada penelitian yang dilakukan Chaitail G. Patil dan Sandip S. Patil menyebutkan penggunaan metode Support Vector Machine dan dataset ISEAR memiliki akurasi tertinggi yaitu 71.64% sedangkan Metode Naive Bayes

p-ISSN: 1693 – 2951; e-ISSN: 2503-2372



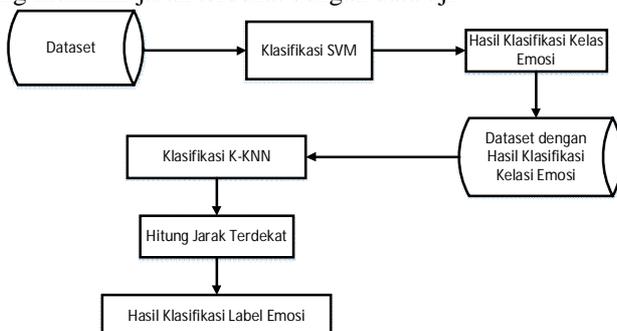
9 772503 237085

Classifier akurasi 60.8% dan yang terendah pada metode Vector Space Model 34.8% dalam untuk Ekstraksi Emosi dari Headline News [3]. Namun Pada penelitian Arifin and Ketut Eddy Purnama melakukan Klasifikasi Emosi Dalam Teks Bahasa Indonesia menggunakan metode K-Nearest Neighbour. Pada penelitian yang dilakukan penulis melakukan klasifikasi emosi pada artikel yang ada di internet kemudian dilakukan pengujian antara metode Naïve Bayes dengan K-Nearest Neighbour. Hasil dari penelitian tersebut didapat metode K-Nearest Neighbour menghasilkan nilai akurasi 71.26% yang lebih tinggi daripada metode Naïve Bayes dengan nilai akurasi 58.01% [4].

Berdasarkan latar belakang dan beberapa penelitian sebelumnya maka penulis melalui penelitian ini mengusulkan melakukan penelitian implementasi text mining pada Sosial Media untuk mendeteksi emosi pengguna. Metode klasifikasi yang digunakan yaitu metode Support Vector Machine untuk klasifikasi kelas emosi dan metode K-Nearest Neighbour untuk klasifikasi kategori emosi. Metode tersebut digunakan karena metode Support Vector Machine memiliki nilai akurasi tertinggi pada penelitian sebelumnya serta Support Vector Machine secara teoritik dikembangkan untuk problem klasifikasi dengan dua class yang sangat tepat untuk klasifikasi kelas emosi[5]. Sedangkan Metode K-Nearest Neighbour digunakan karena pada penelitian sebelumnya K-Nearest Neighbour memiliki akurasi yang lebih tinggi daripada metode Naive Bayes dan Metode K-Nearest Neighbour melakukan pelatihannya sangat cepat dan Efektif jika data pelatihan besar yang sangat cocok dengan penggunaan ISEAR dataset [6]. Deteksi emosi berbasis teks seperti yang disebutkan sebelumnya dapat digunakan dalam bisnis, pendidikan, psikologi, dan bidang lain mana pun yang paling penting untuk memahami dan menafsirkan emosi.

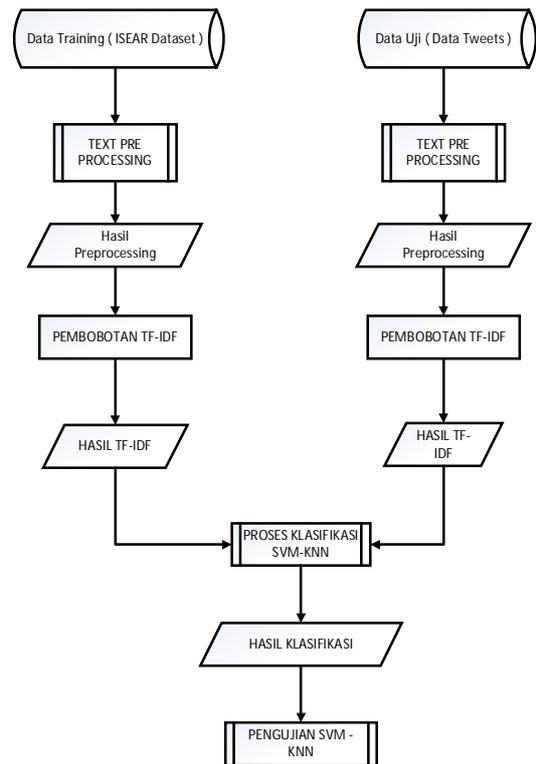
## II. METODE PENELITIAN

Penelitian ini menggunakan metode gabungan SVM dan KNN pada data twitter. Alur metode usulan yang digunakan terlihat pada Gambar 1. Metode SVM polynomial kernel digunakan untuk melakukan klasifikasi pada ISEAR dataset untuk menentukan kelas emosi. Setelah didapatkan label kelas emosi dilakukan seleksi dataset kembali berdasarkan kelas emosi yang telah ditentukan pada tabel kelas emosi[7]. Metode K-NN akan melakukan klasifikasi kembali berdasarkan dataset yang sudah ditentukan kelas emosinya dari metode SVM. K-NN bekerja dengan melakukan perhitungan euclidean distance untuk menentukan label emosi yang memiliki jarak terdekat dengan data uji.



Gambar 1. Metode SVM-KNN

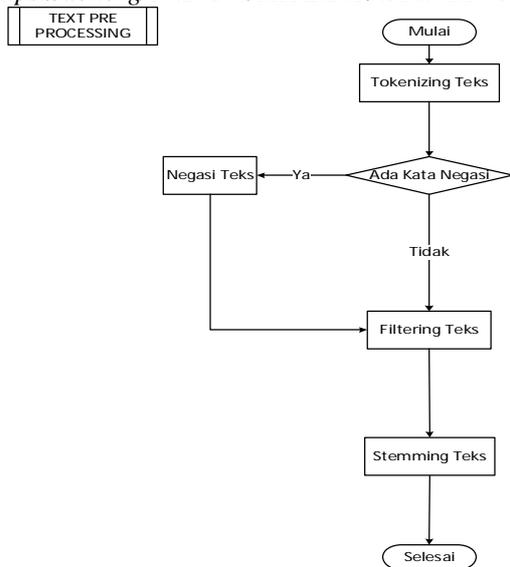
Secara umum proses text mining pada Sosial Media untuk mendeteksi emosi pengguna menggunakan metode SVM-KNN dan Dataset ISEAR ditunjukkan pada Gambar 2:



Gambar 2. Proses Metode dan Teknik Analisis Data

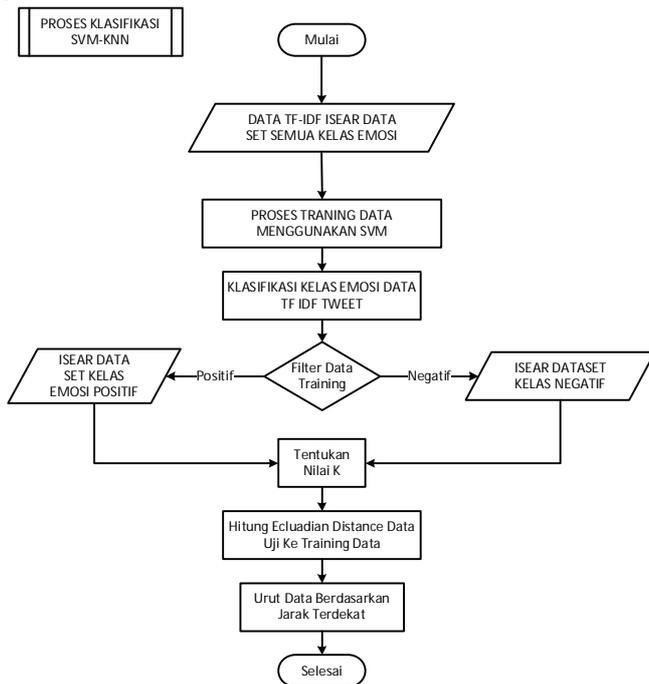
Pada Gambar 2. menjelaskan gambaran sistem yang akan berjalan pada penelitian ini. Proses awal dari sistem ini adalah melakukan pengumpulan data pada twitter dengan menggunakan Stream Twitter API. Setelah melakukan pengumpulan data menggunakan Stream Twitter API kemudian masuk ke Sub Proses Text Preprocessing. Selanjutnya Hasil dari Proses Text Preprocessing adalah melakukan pembobotan menggunakan metode TF – IDF[8]. Pembobotan dilakukan agar bisa dilakukan klasifikasi. Hasil dari pembobotan menggunakan metode TF-IDF kemudian masuk ke sub Proses Klasifikasi menggunakan Metode SVM – KNN untuk mendeteksi emosi dari teks. Proses Selanjutnya dilakukan pengujian terhadap hasil dari klasifikasi sistem. Pengujian dilakukan untuk mendapatkan nilai akurasi, recall dan precision dari pendeteksian yang dilakukan.

Untuk subproses Text preprocessing ditunjuk pada Gambar 3. pada proses ini Data tweet tersebut akan dilakukan tokenizing dengan cara dilakukan pemisahan – pemisahan tiap kata pada setiap data tweet. Proses selanjutnya dilakukan filtering. Pada proses filtering ini dilakukan proses seleksi terhadap kata-kata yang dihasilkan dari proses tokenizing, selanjutnya Algoritma stop list akan membuang kata-kata yang tidak penting seperti kata ganti, kata keterangan, kata sambung, kata depan dan kata sandang. dan algoritma word list akan menyimpan kata-kata yang penting. Selanjutnya Sistem melakukan proses stemming, Proses Stemming dilakukan untuk mencari kata dasar dari setiap kata yang telah lolos proses filtering.



Gambar 3. Sub Proses Text Preprocessing

Selanjutnya Sistem akan melakukan Analisis dan Pembobotan menggunakan Metode TF IDF. Pada Proses Metode TF IDF ini dilakukan penentuan bobot data berdasarkan kemunculan term (istilah). Semakin sering sebuah istilah muncul, semakin tinggi bobot data untuk istilah tersebut, dan sebaliknya. Untuk Sub Proses Klasifikasi SVM –KNN ditunjuk pada Gambar 4.



Gambar 4. Sub Proses Klasifikasi SVM-KNN

Pada proses ini Hasil Dari pembobotan TF –IDF pada ISEAR dataset dilakukan Training data menggunakan metode SVM. Selanjutnya Data TF-IDF dari data Tweet dilakukan klasifikasi menggunakan Metode SVM untuk menentukan kelas Emosinya. Hasil dari Klasifikasi SVM akan melakukan filter pada Data Traning pada parameter kelas emosi. Hasil Filter Data Traning digunakan kembali untuk melakukan klasifikasi menggunakan metode KNN. Dengan menggunakan Dwi Ardiada: Text Mining Pada Sosial ...

Metode KNN diperoleh data yang memiliki jarak terdekat yang digunakan untuk mendeteksi emosi pengguna. Hasil dari pengujian ini akan dilakukan evaluasi dengan menghitung recall, precision dan accuracy pada data yang diuji pada penelitian ini.

### III. HASIL DAN PEMBAHASAN

Aplikasi deteksi emosi pengguna dibuat menggunakan Bahasa pemrograman PHP pada Sublime Text 3. Data tweets yang akan diproses disimpan pada table tb\_datauji yang nantinya akan dibaca oleh aplikasi deteksi emosi. Jumlah data yang akan diuji sejumlah 294 data tweet. Untuk melakukan pendeteksian emosi memerlukan data latih sebagai proses mesin pembelajaran. Penelitian ini menggunakan ISEAR dataset sebagai data latihnya. Pada Tabel I merupakan salah satu data latih pada ISEAR dataset.

TABEL I  
 CONTOH DATA LATIH ISEAR DATASET

No	Label Indonesia	Teks Indonesia
1	senang	Pada hari-hari ketika saya merasa dekat dengan pasangan saya dan teman-teman lainnya Ketika saya merasa damai dengan diri sendiri dan juga mengalami kontak dekat dengan orang-orang yang sangat saya hargai
2	Takut	Setiap kali saya membayangkan bahwa seseorang yang saya cintai atau saya dapat menghubungi penyakit yang serius bahkan kematian
3	marah	Ketika saya jelas diperlakukan dengan tidak adil dan tidak memiliki kemungkinan untuk menjelaskan hal ini

#### A. Data Uji Twitter

Pada penelitian ini menggunakan data uji sejumlah 294 data tweets yang sudah dilabelkan berdasarkan emosi yang ditulis pada hastag tweets dibagian akhir [9]. Data uji pada penelitian ini diambil pada Bulan Agustus 2018 sampai Bulan Oktober 2018 yang hanya berbahasa Indonesia saja. Salah satu sampel data uji terlampir pada table II.

TABEL II  
 CONTOH DATA UJI TWITTER

Foto	Nama	Tweet	Label
	kin	Ramai gak kahwin time muda ni. Ramai juga yang bercerai time muda #takut	takut
	MF Sam bg.	@ruhutsitompul Loe punya mulut bikin berguna dan bermanfaat kek,.jgn bau comberan mulu yg keluarr...#jijik	jijik
	Hery Hutabarat	Trotoar Jalan Margonda Sudah Berbulan-bulan Berlubang <a href="https://t.co/ueRm9NAnaV">https://t.co/ueRm9NAnaV</a> @pemkotdepek #malu	Malu



### B. Tokenizing

Untuk melakukan text Preprocessing pada penelitian ini dilakukan tahap tokenizing pada suatu data. Pada tokenizing ini menghilangkan karakter simbol dan angka [10]. Berikut hasil tokenizing pada data latih terdapat pada tabel III.

TABEL III  
CONTOH HASIL TOKENIZING

No.	Label	Teks	Tokenizing
1	senang	Pada hari-hari ketika saya merasa dekat dengan pasangan saya dan teman-teman lainnya Ketika saya merasa damai dengan diri sendiri dan juga mengalami kontak dekat dengan orang-orang yang sangat saya hargai	Pada hari hari ketika saya merasa dekat dengan pasangan saya dan teman teman lainnya Ketika saya merasa damai dengan diri sendiri dan juga mengalami kontak dekat dengan orang orang yang sangat saya hargai
2	Takut	Setiap kali saya membayangkan bahwa seseorang yang saya cintai atau saya dapat menghubungi penyakit yang serius bahkan kematian	Setiap kali saya membayangkan bahwa seseorang yang saya cintai atau saya dapat menghubungi penyakit yang serius bahkan kematian
3	marah	Ketika saya jelas diperlakukan dengan tidak adil dan tidak memiliki kemungkinan untuk menjelaskan hal ini	Ketika saya jelas diperlakukan dengan tidak adil dan tidak memiliki kemungkinan untuk menjelaskan hal ini

### C. Negasi

Pada Proses Negasi pada Text Preprocessing diperlukan Data yang sudah melewati proses tokenizing. Pada Proses Negasi melakukan transformasi teks yang diawali kata tidak kemudian dilakukan antonim berdasarkan data negasi yang sudah diinputkan. Berikut hasil proses Negasi data Latih dalam tabel IV.

TABEL IV  
CONTOH HASIL NEGASI

No	Label	Teks Tokenizing	Hasil Negasi
1	senang	Pada hari hari ketika saya merasa dekat dengan pasangan saya dan teman teman lainnya Ketika saya merasa damai dengan diri sendiri dan juga mengalami kontak dekat dengan orang orang yang sangat saya hargai	pada hari hari ketika saya merasa dekat dengan pasangan saya dan teman teman lainnya ketika saya merasa damai dengan diri sendiri dan juga mengalami kontak dekat dengan orang orang yang sangat saya hargai
2	Takut	Setiap kali saya membayangkan bahwa seseorang yang saya cintai atau saya dapat menghubungi penyakit yang serius	setiap kali saya membayangkan bahwa seseorang yang saya cintai atau saya dapat menghubungi penyakit yang serius bahkan kematian

		bahkan kematian	
3	marah	Ketika saya jelas diperlakukan dengan tidak adil dan tidak memiliki kemungkinan untuk menjelaskan hal ini	ketika saya jelas diperlakukan dengan timpang dan tidak memiliki kemungkinan untuk menjelaskan hal ini

### D. Filtering

Pada Proses Filtering pada Text Preprocessing diperlukan Data yang sudah melewati proses negasi. Pada Proses Negasi melakukan penghapusan kata – kata yang ada pada daftar data stoplist. Berikut hasil proses Filtering data Latih dalam tabel V.

TABEL V  
CONTOH HASIL FILTERING

No	Label	Teks Negasi	Hasil Filtering
1	senang	pada hari hari ketika saya merasa dekat dengan pasangan saya dan teman teman lainnya ketika saya merasa damai dengan diri sendiri dan juga mengalami kontak dekat dengan orang orang yang sangat saya hargai	pasangan teman teman damai mengalami kontak orang orang hargai
2	Takut	setiap kali saya membayangkan bahwa seseorang yang saya cintai atau saya dapat menghubungi penyakit yang serius bahkan kematian	kali membayangkan cintai menghubungi penyakit serius kematian
3	marah	ketika saya jelas diperlakukan dengan timpang dan tidak memiliki kemungkinan untuk menjelaskan hal ini	diperlakukan timpang tidak memiliki

### E. Stemming

Pada Proses Stemming pada Text Preprocessing diperlukan Data yang sudah melewati proses filtering. Pada Proses Stemming melakukan pencarian kata dasar yang ada pada daftar data kata dasar. Proses pencarian kata dasar menerapkan Algoritma Nazief Stemmer[11]. Berikut hasil proses Stemming data Latih dalam tabel VI.

TABEL VI  
CONTOH HASIL STEMMING

No	Label	Teks Filtering	Hasil Stemming
1	senang	pasangan teman teman damai mengalami kontak orang orang hargai	pasang teman teman damai alami kontak orang orang harga
2	Takut	kali membayangkan cintai menghubungi penyakit serius kematian	kali bayang cinta hubung sakit serius mati
3	marah	diperlakukan timpang tidak memiliki	laku timpang tidak milik

### F. Pembobotan TF-IDF

Sebelum Pada klasifikasi emosi data yang sudah dilakukan text preprocessing memerlukan pembobotan untuk dapat

diklasifikasi. Dalam penelitian ini data latih akan dilakukan pembobotan dengan menggunakan metode TF-IDF. Pembobotan TF IDF bisa dilakukan apabila ada sebuah data uji yang digunakan sebagai query dalam metode TF-IDF. Berikut pada Gambar 5. akan diperlihatkan proses aplikasi dalam melakukan pembobotan menggunakan metode TF IDF

kategori senang, 22 data testing pada kategori takut, 35 data testing pada kategori marah, 187 data testing pada kategori sedih, 17 data testing pada kategori jijik, 24 data testing pada kategori malu.

Sedangkan pada data latih ISEAR dataset terdapat 1091 data latih pada kategori senang, 1093 data latih pada kategori takut, 1094 data latih pada kategori marah, 1094 data latih pada kategori sedih, 1094 data latih pada kategori jijik dan 1094 data latih pada kategori malu. Dari data tersebut diklasifikasi untuk mendeteksi emosi dengan metode yang diusulkan untuk melakukan perbandingan. Dari tahap evaluasi pengajuan yang dilakukan maka didapatkan hasil pengujian klasifikasi emosi. Berikut pada tabel VII merupakan hasil pengujian menggunakan metode K-Nearest Neighbour.

TABEL VII

HASIL EVALUASI KESELURUHAN KATEGORI EMOSI METODE K-NN

Kategori Emosi	Precision	Recall	Accuracy
jijik	0.2727	0.5294	0.8912
Malu	0.2333	0.2917	0.8639
marah	0.2549	0.3714	0.7959
Sedih	0.9028	0.3476	0.5612
senang	0.0741	0.4444	0.8129
Takut	0.3148	0.7727	0.8571
Rata-Rata	0.3421	0.4595	0.797

Gambar 5. Pembobotan TF IDF

### G. Klasifikasi Emosi

Dari Hasil Pembobotan Menggunakan Metode TF IDF dilakukan proses Klasifikasi untuk mendeteksi emosi. Proses Klasifikasi pada penelitian ini melakukan perbandingan Klasifikasi sekali metode K-Nearest Neighbour dengan dua kali Klasifikasi yaitu dengan metode Support Vector Machine (SVM) untuk menentukan Kelas Emosi setelah itu diklasifikasi kembali dengan Metode K-Nearest Neighbour untuk mendeteksi kategori emosinya. Pada metode K-Nearest Neighbour menggunakan nilai K= 6. Pada Metode Support Vector Machine menggunakan kernel polynomial non Linier Kernel dengan parameter tetap yang di default p=2 dan b=1. Berikut hasil klasifikasi emosi pada kedua metode yang dibandingkan terlampir pada Gambar 6.

Dari hasil pengujian keseluruhan dalam klasifikasi emosi dengan metode K-Nearest Neighbour nilai rata-rata precision yaitu 0.3421. Sedangkan nilai recall yaitu 0.4595 dan pada nilai accuracy yaitu pada 0.797. Sedangkan pada hasil pengujian menggunakan metode SVM-KNN didapat hasil pada table VIII seperti berikut :

TABEL VIII

HASIL EVALUASI KESELURUHAN KATEGORI EMOSI METODE SVM K-NN

Kategori	Precision	Recall	Accuracy
jijik	0.1591	0.4118	0.8401
Malu	0.3333	0.3333	0.8912
marah	0.2923	0.5429	0.7891
Sedih	0.881	0.3957	0.5816
senang	0.8333	0.5556	0.983
Takut	0.2394	0.7727	0.7993
Rata-Rata	0.4564	0.502	0.8141

Foto	Nama	Username	Tweet	Label Sebenarnya	Kelas SVM	Label KNN	Label SVM-KNN
	Mira Siregar	uberfunk	Nonton film Sultan Agung ketemu sodara banyak banget padahal ngga janjian 🤔 #senang	senang	negatif	jijik	jijik
	danialraees	danialraees8	Pencarian buddy setel #senang	senang	positif	senang	senang
	Dewi	wiwipas245623	Ups!!! ternyata sampe 9000 yahh, yang ngelike makasih lah buat tmm yg sering ngelike status2 aku. #senang	senang	positif	marah	senang

Gambar 6. Klasifikasi Emosi

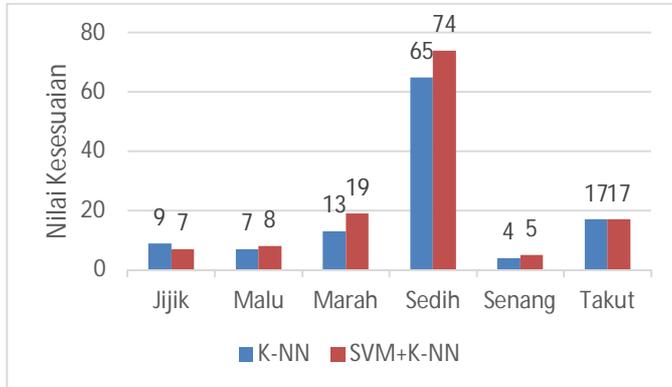
### H. Evaluasi Pengujian

Pada Hasil Klasifikasi Emosi yang telah dilakukan penelitian ini melakukan Evaluasi pengujian dengan menggunakan metode confusion matrix untuk menentukan Recall, Precision dan Accuracy pada sebuah klasifikasi yang dilakukan. Klasifikasi yang dilakukan pada penelitian ini menggunakan data uji sejumlah 294 data dan dengan data latih sebanyak 6561 dari ISEAR dataset. Data testing diperoleh dari media sosial twitter dengan menggunakan Twitter Api pada Bulan Agustus 2018 Sampai Tanggal 2 Oktober 2018. Data Uji dilabelkan berdasarkan hastag berdasarkan kategori emosi. Terdapat 9 data testing pada

Dari hasil pengujian keseluruhan dalam klasifikasi emosi dengan metode SVM-KNN nilai rata-rata precision yaitu 0.4564. Sedangkan nilai recall yaitu 0.502 dan pada nilai accuracy yaitu pada 0.8141. Perbandingan Evaluasi Hasil Pengujian Dari Metode SVM-KNN dengan K-NN menunjukkan adanya perbedaan nilai precision, recall dan accuracy yang disebabkan dari pengurangan data training pada klasifikasi SVM-KNN berdasarkan hasil klasifikasi Kelas Emosi dengan metode SVM. Apabila kelas Emosinya



Positif maka Klasifikasi selanjutnya hanya menggunakan kategori emosi yang memiliki kelas positif Demikian juga sebaliknya. Dari hal tersebut juga mempengaruhi Klasifikasi K-NN pada nilai K. Selain itu hasil pengujian tersebut juga menunjukkan perbandingan nilai kesesuaian dalam Klasifikasi Emosi pada Gambar 8. Berikut hasil Evaluasi Kesesuaian Klasifikasi Seluruh Emosi.



Gambar 7 Grafik Hasil Kesesuaian Keseluruhan Kategori Emosi

Selain itu dilakukan pengujian Komputasi Waktu yang diperlukan dalam klasifikasi emosi. Dari pengujian komputasi waktu didapatkan hasil pada table IX sebagai berikut :

TABEL IX  
HASIL EVALUASI WAKTU KOMPUTASI KATEGORI EMOSI

Kategori Emosi	Waktu Komputasi Klasifikasi ( dalam detik )	
	K-NN	SVM-K-NN
Jijik	±5.0006	±5.4353
Malu	±9.1735	±11.5552
Marah	±14.7334	±15.8777
Sedih	±72.8879	±77.572
Senang	±3.3227	±3.7488
Takut	±7.8605	±8.8471

Pada Pengujian Komputasi waktu Metode SVM-KNN memiliki selisih waktu yang tidak signifikan dari metode KNN itu disebabkan karena metode SVM-KNN melakukan dua kali klasifikasi yaitu klasifikasi pertama menggunakan Metode SVM dan Klasifikasi kedua menggunakan Metode K-NN yang menggunakan data latih berdasarkan hasil dari klasifikasi SVM.

#### IV. KESIMPULAN

Berdasarkan uji coba yang telah dilakukan dalam penelitian ini dapat disimpulkan beberapa hal sebagai berikut:

1. Penelitian ini menunjukkan hasil klasifikasi emosi dengan metode Support Vector Machine dan K-Nearest Neighbour mengalami peningkatan nilai presisi , recall , accuracy dan kesesuaian yang signifikan dibandingkan dengan metode K-Nearest Neighbour pada klasifikasi emosi. Hal ini terlihat dari hasil pengujian yang dilakukan dalam mendeteksi emosi dengan metode Support Vector Machine dan K-Nearest Neighbour mendapatkan hasil dengan nilai rata-rata precision sebesar

0.4564.nilai recall sebesar 0.502 dan pada nilai accuracy sebesar 0.8104 dalam melakukan klasifikasi emosi sedangkan dari metode K-Nearest Neighbour mendapatkan hasil dengan nilai rata-rata precision sebesar 0.3421.nilai recall sebesar 0.4595 dan pada nilai accuracy sebesar 0.797.

2. Metode dua kali klasifikasi Support Vector Machine dan K-Nearest Neighbour memiliki selisih waktu yang tidak signifikan dengan metode K-Nearest Neighbour yang melakukan satu kali klasifikasi. Hal ini terlihat dari hasil komputasi waktu klasifikasi yang dilakukan. Itu disebabkan karena adanya pengurangan data training pada klasifikasi SVM-KNN yang menggunakan acuan hasil klasifikasi Kelas Emosi dengan metode Support Vector Machine. Apabila kelas Emosinya Positif maka Klasifikasi selanjutnya hanya menggunakan data training dengan kategori emosi yang memiliki kelas positif saja demikian juga sebaliknya.

#### REFERENSI

- [1] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter , a Social Network or a News Media ?," *19th Int. Conf. World Wide Web*, pp. 591–600, 2010.
- [2] V. P. Haji Binali, Chen Wu, "Computational Approaches for Emotion Detection in Text," *IEEE Int. Conf. Digit. Ecosyst. Technol. (IEEE DEST 2010)*, 2010.
- [3] C. G. Patil and S. S. Patil, "Use of Porter Stemming Algorithm and SVM for Emotion Extraction from News Headlines," *Int. J. Electron. Commun. Soft Comput. Sci. Eng.*, vol. 2, no. 7, pp. 2277–9477, 2013.
- [4] Arifin, "Classification of Emotions in Indonesian Texts Using K-NN Method," *Int. J. Inf. Electron. Eng.*, 2012.
- [5] C. Cortes and V. Vapnik, "Support vector machine," *Mach. Learn.*, pp. 1303–1308, 1995.
- [6] R. I. Ndaumanu and M. R. Arief, Kusriani, "Analisis Prediksi Tingkat Pengunduran Diri Mahasiswa dengan Metode K-Nearest Neighbor," *Jatisi*, vol. 1, no. 1, pp. 1–15, 2014.
- [7] I. N. W. Tirtayani, Luh Ayu, Noce Maylani Asril, "Perkembangan Sosial Emosional pada Anak Usia Dini," *Univ. Terbuka Repos.*, pp. 1–43, 2013.
- [8] N. G. Yudiarta, M. Sudarma, and W. G. Ariastina, "Pengelompokan Berita Pada Unstructured Textual Data," vol. 17, no. 3, pp. 339–344, 2018.
- [9] A. P. Sujana, "Memanfaatkan Big Data Untuk Mendeteksi Emosi," *Tek. Komput. Unikom*, vol. 2, no. 2, pp. 1–4, 2013.
- [10] N. L. Ratniasih, M. Sudarma, N. Gunantara, and A. A. U. Sistem, "Penerapan Text Mining dalam Spam Filtering untuk Aplikasi Chat," *Tekno. Elektro*, vol. 16, no. 3, pp. 13–18, 2017.
- [11] M. Adriani, J. Asian, B. Nazief, and H. E. Williams, "Stemming Indonesian : A Confix Stripping Approach," *ACM Trans. Asian Lang. Inf. Process.*, vol. 6, no. 4, pp. 1–33, 2007.