

# Mendeteksi *Cyberhate* pada Twitter Menggunakan *Text Classification* dan *Crowdsourced Labeling*

Hadi Kurniawan Sidiq<sup>1</sup>, Dana Sulistyono Kusumo<sup>2</sup>, Indra Lukmana Sardi<sup>3</sup>

**Abstract**—During the 2019 presidential election campaign in Indonesia, a lot of support was made by the community with various forms of support, such as poster distribution or even content on social media. For example, in social media such as Twitter, there were many support tags during the presidential election, such as #2019gantipresiden, #2019tetapjokowi, and other hashtags related to the Indonesian presidential election. However, many hate speeches are contained in tweets with the related hashtag. Hate speech on the internet (cyberhate) could cause disputes between support groups of the two presidential candidates which cause conflicts such as riots and other actions that harm the country. This study uses the SVM algorithm to detect cyberhate that produces the best accuracy of 97%. Also, this study applies crowdsourced labeling in dataset labeling which results in 98% valid data.

**Intisari**—Pada masa kampanye pemilihan umum presiden 2019 di Indonesia, banyak dukungan yang dilakukan masyarakat dengan berbagai macam bentuk dukungan, seperti dukungan berupa selebaran dan juga tulisan pada media sosial. Pada media sosial seperti Twitter misalnya, selama masa pemilihan presiden terdapat banyak tagar-tagar dukungan seperti #2019gantipresiden, #2019tetapjokowi, dan tagar lainnya, yang terkait pemilihan presiden Indonesia. Namun, tidak jarang terdapat ujaran kebencian atau *hate speech* yang terkandung pada *tweet* dengan tagar tersebut. *Hate speech* yang disebarluaskan di internet (*cyberhate*) dapat menyebabkan munculnya perselisihan antar individu maupun kelompok. Pada masa kampanye pemilihan presiden ini, *cyberhate* dapat menyebabkan perselisihan antar kelompok pendukung kedua pasangan calon presiden dan dapat berakibat konflik seperti kerusuhan dan aksi lainnya yang merugikan negara. Dalam makalah ini digunakan algoritme SVM untuk mendeteksi *cyberhate* yang menghasilkan keakuratan terbaiknya sebesar 97%. Selain itu, juga diterapkan *crowdsourced labeling* dalam pelabelan *dataset* yang digunakan, yang menghasilkan validitas data sebesar 98% data valid dengan tujuh data yang salah.

**Keyword**— *Crowdsourced Labeling, Cyberhate Tweets, Hate Speech Detection, Text Classification.*

## I. PENDAHULUAN

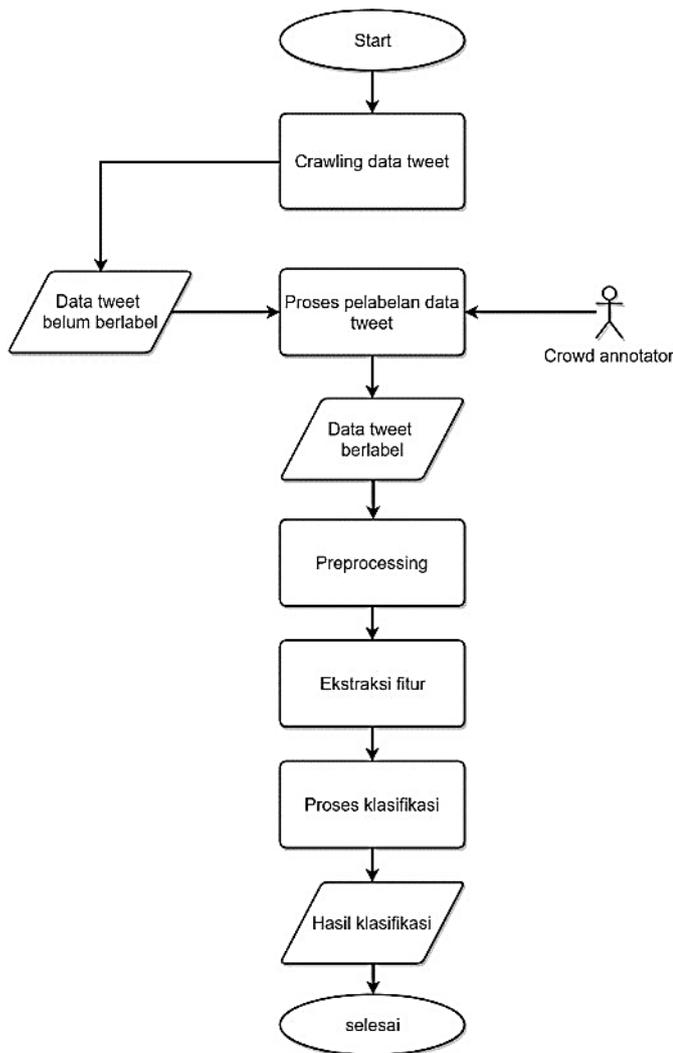
Pada masa kampanye pemilihan presiden 2019 di Indonesia, banyak dukungan yang dilakukan masyarakat dengan berbagai macam bentuk dukungan, seperti dukungan dalam bentuk selebaran maupun tulisan di media sosial. Misalnya pada media sosial seperti Twitter, dalam masa pemilihan presiden terdapat banyak tagar-tagar dukungan, seperti #2019gantipresiden,

#2019tetapjokowi, dan tagar lainnya yang terkait pemilihan presiden Indonesia. Namun, terdapat banyak ujaran kebencian atau *hate speech* yang terkandung pada *tweet* dengan tagar tersebut. *Hate speech* atau ujaran kebencian adalah komunikasi apa pun yang meremehkan seseorang atau kelompok atas dasar beberapa karakteristik seperti ras, etnis, jenis kelamin, orientasi seksual, kebangsaan, agama, atau karakteristik lainnya [1]. Di Indonesia, kasus *hate speech* menjadi kasus yang paling sering dilaporkan ke kepolisian [2]. *Hate speech* yang disebarluaskan di internet (*cyberhate*) dapat menyebabkan munculnya perselisihan antar individu maupun kelompok. Pada masa kampanye pemilihan presiden ini, *cyberhate* dapat menyebabkan perselisihan antar kelompok pendukung kedua pasangan calon presiden dan dapat berakibat konflik seperti kerusuhan dan aksi lainnya yang merugikan negara. Pihak berwenang telah melakukan tindakan penanganan *cyberhate* ini dengan diaturnya hukuman bagi pelaku *cyberhate* dalam surat edaran Kepolisian Republik Indonesia tentang ujaran kebencian. Namun, tetap saja masih banyak pelaku *cyberhate* yang menyebarkan ujaran kebencian di internet yang sulit ditemukan, karena membutuhkan sumber daya yang memadai. Oleh karena itu, diperlukan penanganan *cyberhate* untuk membantu pihak berwajib dalam menindaklanjuti pelakunya. Penelitian mengenai penanganan *cyberhate* sudah ada sebelumnya, seperti pada penelitian yang melakukan pendeteksian *hate speech* pada Twitter dengan menggunakan beberapa metode klasifikasi, seperti *Naïve Bayes*, kNN, *Support Vector Machine (SVM)*, *maximum entropy*, dan juga *random forest* yang menghasilkan keakuratan sistem dalam mendeteksi *hate speech* [3].

Pada makalah ini, diadopsi beberapa metode yang digunakan pada penelitian sebelumnya [3], [4]. Pada penelitian sebelumnya, terdapat perbandingan penggunaan metode pengklasifikasian teks yang digunakan, seperti SVM, *random forest*, dan kNN. SVM memiliki kinerja akurasi yang cukup baik dibandingkan *random forest* dan kNN, yakni sebesar 81,3% dengan *dataset* yang imbang [3]. Oleh sebab itu, makalah ini mencoba mengklasifikasikan teks yang mengandung *cyberhate* pada Twitter menggunakan pendekatan algoritme SVM untuk mendapatkan akurasi yang baik.

Selain metode klasifikasi, makalah ini juga menerapkan sistem *crowdsourcing* dalam pelabelan *dataset* yang akan diklasifikasikan. *Crowdsourcing* dapat dimanfaatkan dalam menilai sebuah konflik publik karena permasalahan yang muncul merupakan masalah publik [5]. Penelitian tersebut mencoba melibatkan penilaian publik dalam menyelesaikan masalah *cyberhate* dengan menerapkan sistem *crowdsourcing* untuk menilai sebuah *tweet* mengandung *cyberhate* atau tidak. Data hasil *crowdsourcing* ini kemudian diuji validitasnya oleh ahli.

<sup>1,2,3</sup> Fakultas Informatika, Universitas Telkom, Jl. Telekomunikasi Terusan Buah Batu Bandung, 40257 INDONESIA (Tlp: +62-22-756-9131; fax: +62-22-7565-930; email: hadikaes@student.telkomuniversity.ac.id, danakusumo@telkomuniversity.ac.id, indraluk@telkomuniversity.ac.id)



Gbr. 1 Alur perancangan sistem.

## II. METODE

### A. Alur Perancangan Sistem

Alur perancangan sistem merupakan gambaran sistem yang diterapkan pada makalah ini, dimulai dari *crawling* data menggunakan API Twitter, lalu melakukan pelabelan data menggunakan sistem *crowdsourcing* hingga menjadi sebuah *dataset* yang digunakan sebagai data latih untuk melatih algoritme SVM, untuk kemudian diuji dengan melakukan pendeteksian *cyberhate* pada data uji. Hasil dari kinerja SVM dalam mendeteksi *cyberhate* pada data uji dievaluasi menggunakan beberapa metode berbeda, yang merupakan fokus pada makalah ini.

Berdasarkan rancangan sistem pada Gbr. 1, langkah pertama yang dilakukan adalah melakukan *crawling* data *tweet* menggunakan API Twitter. Data *tweet* yang diambil dari hasil *crawling* pada Twitter lalu diberi label melalui *crowdsourcing* dengan menggunakan *crowdsourced labeling system*. Proses ini menghasilkan *tweet* yang sudah diberi label.

Langkah selanjutnya adalah melakukan *pre-processing* pada data yang sudah berlabel. Pada tahap ini dilakukan proses

penghapusan URL, tanda baca, dan *stopwords* serta dilakukan *stemming* agar diperoleh kata yang hendak dibuat menjadi fitur. Selanjutnya, pada tahap ekstraksi fitur, dilakukan pengelompokan kata menggunakan metode *bag of words* dan pembobotan fitur menggunakan TF-IDF [6], [7]. Hal ini dilakukan untuk membuat fitur yang memiliki nilai, agar dapat diklasifikasikan oleh SVM. Setelah itu, kumpulan fitur ini diklasifikasikan menggunakan algoritme SVM yang akan menghasilkan nilai *recall*, presisi, dan akurasi dari pendeteksian sentimen *hate speech* pada data uji yang telah ditentukan.

### B. Pelabelan Data

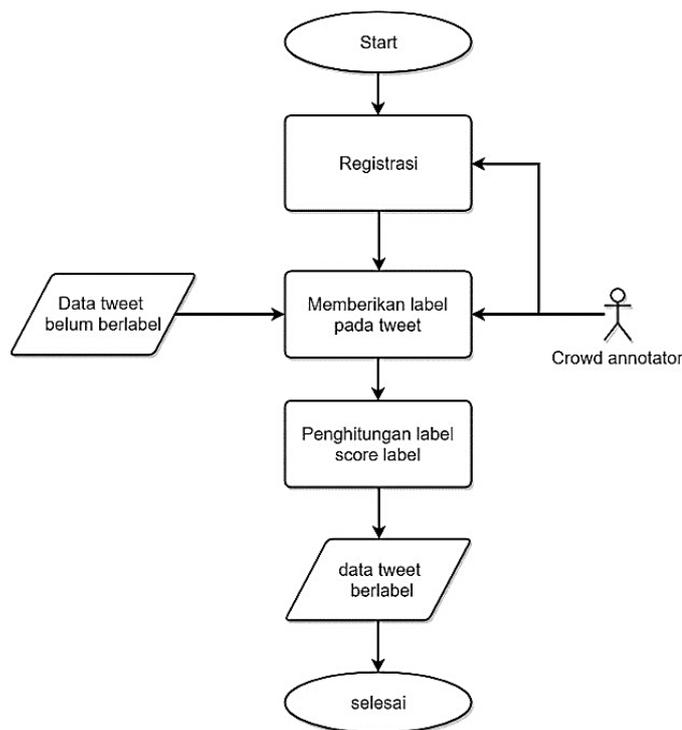
Pada sebuah penelitian, pelabelan *dataset* dilakukan menggunakan Mechanical Turk, yakni sebuah sistem *crowdsourced* dari Amazon untuk pelabelan teks dalam bahasa Inggris [8]. Sistem pemberian label melalui *crowdsourced labeling* yang dibangun merujuk pada penelitian sebelumnya [9]. Sistem ini dibuat dengan menggunakan model skala Likert [10] dengan meniadakan *vote* netral untuk mendapatkan skor pasti dari label yang ditentukan, dalam hal ini *hate speech* dan *non-hate speech*. Sistem ini dibuat dalam bentuk web dengan menampilkan data *tweet* yang sudah dikumpulkan, lalu pemberi label (*annotator*) menentukan tingkat *cyberhate* yang terkandung dalam *tweet* yang ditampilkan, dengan pilihan tingkat keyakinan 1 sampai 4. Pada penelitian sebelumnya, komposisi *annotator* berjumlah 75 orang, yang terdiri atas 48 mahasiswa, lima guru, dan 20 pegawai swasta [7]. *Annotator* pada sistem ini terdiri atas 96 orang, dengan komposisi 46 mahasiswa, 28 pegawai negeri dan swasta, dan 22 wiraswasta, dengan rentang usia 20 sampai 25 tahun.

Proses pelabelan data dilakukan melalui sistem *crowdsourcing*. Pelabelan data ini bertujuan memberikan label untuk menentukan sebuah *tweet* merupakan *hate speech* atau bukan. Perancangan sistem pelabelan yang dibuat ditunjukkan pada Gbr. 2.

Pada Gbr. 2 terlihat bahwa alur perancangan sistem pelabelan dimulai dari proses registrasi yang dilakukan oleh *annotator*, kemudian *annotator* memberikan label pada *tweet* yang belum diberi label. Setelah itu, *tweet* yang sudah diberi label diperhitungkan skor labelnya untuk mengategorikan label. Selanjutnya, *tweet* yang sudah memiliki label disimpan pada daftar *tweet* berlabel. *Annotator* ini memiliki rentang usia 20 sampai 25 tahun, yaitu yang telah dianggap dewasa menurut *de facto* dan dapat membedakan *hate speech* atau bukan. *Annotator* dapat memberikan pilihan satu dari empat pilihan pada suatu *tweet*. Tujuan diberikannya empat pilihan pada pelabelan adalah untuk mendapatkan tingkat keyakinan *annotator* pada *tweet* tersebut, karena tingkat keyakinan *annotator* berbeda-beda [7]. Tabel I memperlihatkan keterangan mengenai pilihan pada sistem pelabelan menggunakan skala Likert [7].

Pada Tabel I ditunjukkan bobot masing-masing label pada skala Likert. Langkah-langkah penghitungan skor label menggunakan metode skala Likert pada data *tweet* adalah sebagai berikut.

1. Pilihan 1 dan 2 digunakan untuk menghitung skor total dari skor *non-cyberhate*. Pilihan 3 dan 4 digunakan untuk menghitung skor total dari skor *cyberhate*.



Gbr. 2 Alur pelabelan data.

- Penghitungan skor *cyberhate*. Karena pilihan kategori *tweet* menggunakan skala Likert, untuk menghitung nilai fungsi label “*Cyberhate*” yang didefinisikan sebagai  $C$ , setiap pilihan 3 atau 4 dikalikan dengan total partisipan yang melakukan pilihan yang memilih pilihan tersebut, menggunakan (1).

$$C = \frac{\sum_{k=3}^4 \text{Weight}_k \times \text{total\_annotator}_k}{\sum_{k=1}^4 \text{Weight}_k \times \text{total\_annotator}_k} \quad (1)$$

- Penghitungan skor *non-cyberhate*. Setelah mendapatkan nilai  $C$  pada (1), ditentukan nilai dari *non-cyberhate* ( $nC$ ), menggunakan (2).

$$nC = C - 1 \quad (2)$$

- Menentukan probabilitas terbesar dari hasil kedua skor yang didefinisikan sebagai  $M$ , menggunakan (3).

$$M = \text{Max}(C, nC) \quad (3)$$

Dari (3) didapatkan nilai probabilitas terbesar yang akan menjadi label final.

### C. Support Vector Machine (SVM)

Algoritme SVM adalah salah satu algoritme berbasis statistik untuk mengklasifikasi data, yang merupakan salah satu bentuk dari konsep *supervised learning*. Konsep utama dari algoritme ini adalah mengklasifikasikan data ke dalam kelas positif dan negatif menggunakan fungsi linear atau nonlinear [11].

Sebelum mengklasifikasikan data, SVM mengubah data numerik menjadi *data point* ( $X, Y$ ). Kemudian, SVM membentuk sebuah fungsi linear dengan rumusan (1).

TABEL I  
KETERANGAN DESKRIPSI BOBOT OPSI PADA SISTEM PELABELAN

Pilihan	Bobot	Deskripsi
1	$ -2 $	Sangat bukan <i>hate speech</i>
2	$ -1 $	Bukan <i>hate speech</i>
3	1	<i>Hate speech</i>
4	2	Sangat <i>hate speech</i>

$$(x) = \langle W \cdot X \rangle + b \quad (4)$$

dengan fungsi  $(x)$  merupakan fungsi *real-value*,  $W$  adalah *weight vector*, dan  $\langle W \cdot X \rangle$  merupakan perkalian *dot* dari  $W$  dan  $X$ . Setelah data diubah menjadi numerik, SVM memberikan label pada masing-masing data, sehingga terbentuk sebuah nilai yang menjadi garis pemisah antar dua kelas yang diklasifikasikan. Pada kasus dalam makalah ini, kelas yang dipisahkan adalah kelas *hate speech* dan *non-hate speech*. Garis pemisah ini disebut juga sebagai *hyperplane*. Rumus perhitungan *hyperplane* ditunjukkan dalam (5).

$$W \cdot X + b = 0 \quad (5)$$

dengan  $W$  sebagai nilai parameter *hyperplane*,  $X$  adalah data *input* SVM,  $b$  merupakan parameter *hyperplane* yang dicari (nilai *bias*), dan  $f$  merupakan fungsi *hyperplane*.

### III. HASIL DAN PEMBAHASAN

Pada makalah ini, hasil pengujian yang dihasilkan ada dua. Hasil yang pertama merupakan hasil pengujian validasi pelabelan oleh *crowd*, dan yang kedua merupakan hasil pengujian dari klasifikasi *hate speech*.

#### A. Skenario Pengujian

Skenario pengujian dalam makalah ini dibagi menjadi dua. Pertama, pengujian terhadap *dataset* yang telah dilabeli oleh *crowd* divalidasi oleh ahli untuk mengetahui validitas data yang diberi label oleh sistem *crowdsourced labeling*. Pengujian kedua adalah pengujian pada model klasifikasi ini, menggunakan tiga skenario, yakni sebagai berikut.

- Pengujian berdasarkan perbandingan komposisi data latih dan data uji.
- Pengujian berdasarkan metode SVM, penerapan beberapa macam *kernel* seperti linear, RBF, dan polinomial. Komposisi data yang digunakan diambil dari komposisi paling optimal dari skenario pertama. Dicari *kernel* dengan tingkat akurasi, presisi, dan *recall* tertinggi.
- Pengujian berdasarkan *k-fold cross validation*, dengan dibagi menjadi 5 *fold*. Untuk metode SVM, data yang digunakan adalah komposisi paling optimal dari skenario pertama dan digunakan *kernel* yang paling optimal dalam skenario kedua. Dari pengujian, dicari *fold* dengan nilai akurasi, presisi, dan *recall* tertinggi.

#### B. Dataset

*Dataset* yang digunakan pada pengujian ini dibagi menjadi dua jenis, yakni latih dan uji. Pembagian *dataset* dilakukan dengan komposisi data yang berbeda-beda, sesuai dengan skenario pengujian yang dilakukan. Hal ini ditujukan untuk

TABEL II  
PEMBAGIAN *DATASET*

Jumlah <i>Dataset</i>	Pembagian Data	
	Latih	Uji
50:50	381	381
60:40	457	305
70:30	533	229
80:20	610	152
90:10	686	76

Gbr. 3 Tampilan halaman registrasi *annotator*.

menguji keakuratan model yang digunakan dalam pendeteksian *cyberhate* pada data uji. Tabel II menyajikan detail dari pembagian *dataset* yang dilakukan.

Pada Tabel II diketahui bahwa *dataset* dibagi menjadi beberapa bagian. Hal ini ditujukan untuk menguji hasil terbaik dari kinerja sistem klasifikasi berdasarkan pembagian data latih dan data uji.

### C. Eksperimen Sistem Pelabelan

Pelabelan *dataset* menggunakan sistem *crowdsourced labeling system* dalam bentuk *website* sederhana. Sebelum *annotator* (pelabel *dataset*) memberikan label, *annotator* harus mengisi *form* registrasi terlebih dahulu untuk mengidentifikasi jumlah *annotator* yang memberikan label data *tweet* pada *website*.

Pada Gbr. 3 ditunjukkan formulir registrasi bagi *annotator* yang berisikan nama, umur, *email*, dan juga status dari *annotator*. Setelah itu, *annotator* diberi instruksi dan arahan dalam memberikan label pada data *tweet* yang akan ditampilkan. Setiap *annotator* melabeli tiga puluh data *tweet* seperti pada Gbr. 4.

Pada Gbr. 4 digambarkan bentuk tampilan *tweet* yang akan diberi label oleh *annotator*. Pada gambar terlihat juga informasi tagar asal *tweet* dan juga pilihan label yang dapat dipilih oleh *annotator*. Dari hasil pelabelan *crowdsourced* ini, diperoleh data yang terlabeli sebanyak 762 dari 1.045 data dalam kurun

### Tweets

Gbr. 4 Tampilan proses pelabelan *tweet*.

waktu tiga minggu (11 juni 2019 sampai 2 Juli 2019). Pelabelan dilakukan oleh total 96 *annotator* dengan komposisi 46 mahasiswa, 28 pegawai negeri atau swasta, dan 22 wiraswasta. *Dataset* yang dihasilkan memiliki 311 data berlabel *hate speech* dan 451 data berlabel *non-hate speech*.

### D. Hasil Pengujian Validasi Keakuratan *Dataset*

Setelah dilakukan pelabelan oleh *annotator* pada *crowdsourced labeling system*, *dataset* kemudian diuji untuk dievaluasi oleh *expert* guna menentukan keakuratan data. Dari hasil evaluasi yang dilakukan, tingkat keakuratan data dari *crowdsourced labeling* mencapai 99%, dengan jumlah kesalahan pelabelan sebanyak tujuh data dari 762 data yang telah diberi label.

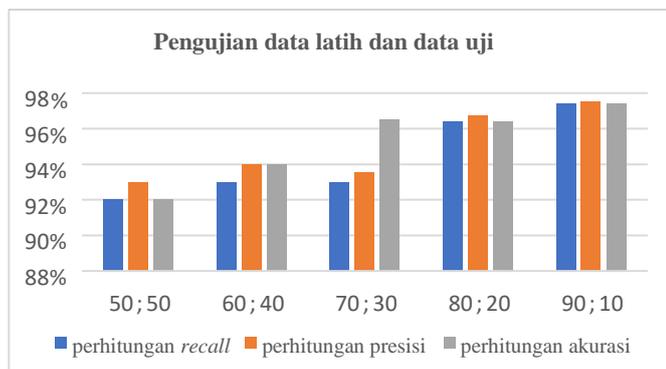
### E. Analisis Pengujian Klasifikasi Menggunakan Pembagian *Dataset*

Dalam pengujian skenario pertama dilakukan pengujian dengan perbandingan komposisi data latih dan data uji. Dari perbandingan komposisi data, digunakan hasil yang paling optimal untuk pengujian selanjutnya.

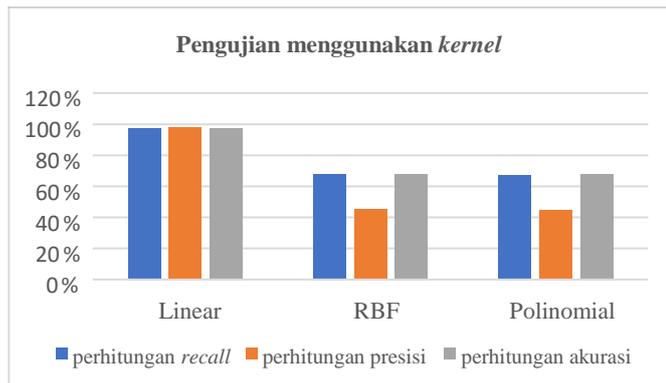
Pada Gbr. 5 terlihat grafik hasil pengujian berdasarkan perbandingan komposisi data latih dan data uji. Dalam pengujian ini, algoritme SVM dengan perbandingan komposisi 90:10 memiliki hasil yang paling optimal, yaitu mencapai 97% nilai akurasi. Dari pengujian dapat dilihat bahwa semakin banyak jumlah data latih, semakin tinggi hasil perhitungan. Hal ini dikarenakan semakin banyaknya *dataset* yang diuji menyebabkan fitur yang dihasilkan juga semakin banyak, sehingga nilai skor dari data yang diklasifikasi juga semakin dekat dengan batas *hyperlane* SVM.

### F. Analisis Pengujian Klasifikasi Menggunakan Kernel SVM

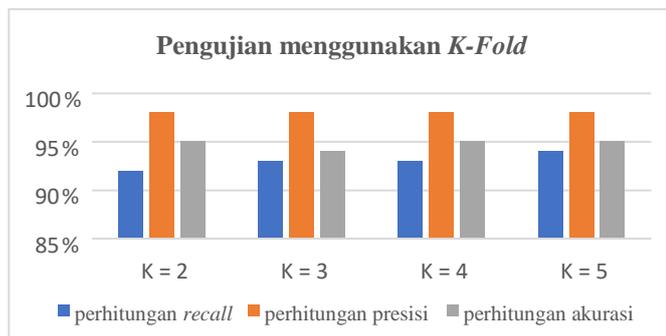
Dalam analisis kedua, digunakan komposisi data yang paling optimal dari pengujian pertama. Masing-masing dibandingkan akurasi, presisi, dan *recall*-nya berdasarkan *kernel* dari metode SVM. Untuk komposisi datanya, digunakan skema 90:10.



Gbr. 5 Grafik hasil penguujian SVM berdasarkan perbandingan komposisi data.



Gbr. 6 Grafik hasil penguujian SVM dengan kernel.



Gbr. 7 Grafik hasil penguujian SVM terhadap jumlah fold.

Pada Gbr. 6 terlihat hasil penguujian SVM dengan berbagai penguujian *kernel* yang berbeda-beda. Dari *kernel* yang diujikan, *kernel* linear memiliki akurasi lebih baik daripada kedua penguujian lainnya. Dalam studi analisis teks, *kernel* linear lebih sering digunakan daripada *kernel* yang lain. Hal ini dikarenakan *kernel* linear dibuat untuk mengklasifikasikan model kelas dengan jumlah kelas yang lebih sedikit, dibandingkan dengan RBF maupun polinomial yang digunakan untuk mengklasifikasikan *multi-class* dengan kompleksitas yang tinggi.

#### G. Analisis Penguujian Kklasifikasi Menggunakan K-Fold

Dalam analisis ketiga, dilakukan perubahan nilai *fold*. Nilai *fold* yang digunakan adalah 2-5 *fold*. Untuk analisis pada

metode SVM, digunakan *kernel* linear dan komposisi data dengan skema 90:10 karena paling optimal nilai akurasinya.

Pada Gbr. 7 didapatkan hasil penguujian SVM menggunakan *K-Fold* dengan jumlah *k* yang berbeda-beda. Dari hasil penguujian, terlihat bahwa jumlah *fold* yang diberikan berpengaruh terhadap perhitungan *recall* dan akurasi, sedangkan untuk perhitungan presisi jumlah *fold* tidak terlalu berpengaruh.

#### IV. KESIMPULAN

Berdasarkan hasil penguujian dan analisis yang telah dilakukan sebelumnya, dapat diambil kesimpulan bahwa data yang dilabeli oleh *crowd* memiliki tingkat akurasi yang tinggi. Dari hasil pengecekan evaluator, hasil kebenaran sebuah data mencapai 99%. Dari 762 data terlabel, hanya terdapat tujuh data yang salah.

Dalam percobaan *kernel* SVM, yang paling optimal adalah yang menggunakan *kernel* linear, dengan rata-rata nilai akurasi, presisi, dan *recall* mencapai 96%. Juga dapat diketahui bahwa jumlah *folding* pada penguujian ini tidak memengaruhi kinerja masing-masing metode yang digunakan.

#### REFERENSI

- [1] I. Alfina, R. Mulia, M.I. Fanany, dan Y. Ekanata, "Hate Speech Detection in the Indonesian Language: A Dataset and Preliminary Study," *2017 Int. Conf. on Advanced Computer Science and Information Systems (ICACSIS)*, 2017, hal. 233-238.
- [2] H. Margono, X. Yi, dan G.K. Raikundalia, "Mining Indonesian Cyberbullying Patterns in Social Networks," *Proc. of Thirty-Seventh Australasian Computer Science Conference*, 2014, hal. 115-124.
- [3] S.H. Pratiwi, "Detection of Hate Speech against Religion on Tweet in the Indonesian Language Using Naïve Bayes Algorithm and Support Vector Machine," B.Sc. Tesis, Universitas Indonesia, Jakarta, Indonesia, 2016.
- [4] I. Alfina, D. Sigmawaty, F. Nurhidayati, dan A.N. Hidayanto, "Utilizing Hashtags for Sentiment Analysis of Tweets in the Political Domain," *Proc. of the 9th Int. Conf. on Machine Learning and Computing*, 2017, hal. 43-47.
- [5] A. Kahl, C. McConnell, dan W. Tsuma, "Crowdsourcing as a Tool in Conflict Prevention," *Conflict Trends*, Vol. 2012, No. 1, hal. 27-34, Jan 2012.
- [6] (2018) "Pembobotan Kata atau Term Weighting TF-IDF," [Online], <https://informatikalogi.com/term-weighting-tfidf>, tanggal akses: 3-Mei-2019.
- [7] J. Ramos, "Using TF-IDF to Determine Word Relevance in Document Queries," *1st Int. Conf. on Machine Learning*, 2003, hal. 1-4.
- [8] A. Kontostathis, K. Reynolds, A. Garron dan L. Edwards, "Detecting Cyberbullying: Query Terms and Techniques," *Proc. of the 5th Annual ACM Web Science Conference (WebSci '13)*, 2013, hal. 195-204.
- [9] H. Nurrahmi dan D. Nurjanah, "Indonesian Twitter Cyberbullying Detection using Text Classification and User Credibility," *Int. Conf. on Information and Communications Technology (ICOIACT)*, 2018, hal 543-548.
- [10] I.E. Allen dan C.A. Seaman, (2007) "Likert Scale and Data Analyses," [Online], <http://asq.org/quality-progress/2007/07/statistics/likertscale-and-data-analyses.html>. tanggal akses: 20-Mei-2019.
- [11] K. Dinakar, B. Jones, C. Havasi, H. Lieberman, dan R. Picard, "Common Sense Reasoning for Detection, Prevention, and Mitigation of Cyberbullying," *ACM Transactions on Interactive Intelligent Systems*, Vol. 2, No. 3, hal. 18:1-30, 2012.