



Analisis *Sentiment* Twitter Berbasis *Grid Search Algorithm* (GSA) dengan Metode *Support Vector Machine* (SVM)

Twitter Sentiment Analysis Based on Grid Search Algorithm (GSA) with Support Vector Machine (SVM) Method

Dedi Wirasasmita* dan Efi Anisa

Sekolah Tinggi Teknologi Duta Bangsa, Jl. Kalibaru Timur, Kalibaru Medan Satria, Jawa Barat 17530, Indonesia

Informasi artikel:

Diterima:
28/09/2022
Direvisi:
21/11/2022
Disetujui:
23/11/2022

Abstract

Twitter is a social networking service that has undergone tremendous growth and is gaining worldwide popularity at an accelerated rate. Twitter allows for the expression of unbiased thoughts on a variety of issues and can assist businesses in providing public feedback on well-known brands and items. Twitter is having trouble with good and negative answers. Researchers evaluated English-language tweets to determine the proportion of positive and negative replies to popular companies and items. This study will explore Twitter sentiment analysis utilizing the Grid Search Algorithm (GSA) and the support vector machine (SVM) technique. GSA is utilized by the feature selection model to optimize the classification procedure. In this work, training data and testing data are required to do sentiment analysis. Sanders Twitter 0.2 utilizes a dataset consisting of tweets retrieved from Twitter using the search terms @apple, #google, #microsoft, and #twitter. The collected dataset was manually annotated and included 654 negatives, 570 positives, 2503 neutrals, and 1786 irrelevant entries. Data are loaded, tokenized, weighted, preprocessed, filtered, and classified to conduct a sentiment analysis. The application's sentiment analysis achieved a degree of accuracy of up to 79% based on testing. The ratio of neutral and bad tweets on data sandboxes tends to be greater than the percentage of positive tweets, hence optimization rather than accuracy is obtained.

Keywords: *classification, SVM, GSA, Twitter, accuracy.*

SDGs:



Abstrak

Twitter adalah layanan jejaring sosial yang mengalami perkembangan pesat dan cepat meraih popularitas di dunia. Twitter dapat mengekspresikan opini yang objektif tentang topik yang berbeda, serta dapat membantu dalam bidang bisnis untuk menyediakan opini terhadap konsumen mengenai merk dan produk yang populer. Permasalahan pada tanggapan-tanggapan positif dan negatif yang beredar melalui *tweets*. Melalui postingan *tweet* tersebut, maka dilakukan analisis *tweet* berbahasa Inggris untuk mengetahui seberapa banyak tanggapan positif dan negatif tentang merk dan produk yang populer. Penelitian ini akan membahas mengenai Analisis sentiment Twitter berbasis *Grid Search Algorithm* (GSA) menggunakan metode *Support Vector Machine* (SVM). Untuk mengoptimalkan proses klasifikasi model seleksi fitur menggunakan GSA. Data training dan data testing di perlukan dalam penelitian ini untuk melakukan analisis sentiment. Dataset yang digunakan Sanders Twitter 0.2 berupa tweets pada Twitter yang diperoleh melalui pencarian dengan kata kunci @apple, #google, #microsoft, #twitter. Dataset yang diperoleh dilabel dengan cara manual yang terdiri atas data 654 negatif, 570 positif, 2503 netral, 1786 irrelevant. Analisis sentiment dilakukan dari tahapan *loading data, tokenizing, weighting, preprocessing*, Filtering dan klasifikasi. Berdasarkan uji coba yang dilakukan, tingkat keakuratan Analisis sentiment pada aplikasi diperoleh nilai akurasi hingga 79,00%. Persentase tweets data sanders cenderung lebih tinggi tweets netral dan negatif daripada positif, sehingga optimalisasi dari pada akurasi tercapai.

Kata Kunci: klasifikasi, SVM, GSA, Twitter, akurasi.

*Penulis Korespondensi. Tel: +6221 2908 2747 ; Hp: +62 812 1016 9472
email : dedi.wirasamita@sttdb.ac.id



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

1. PENDAHULUAN

Twitter merupakan layanan jejaring sosial yang mengalami perkembangan pesat dan cepat meraih popularitas di dunia. Twitter mampu untuk mengekspresikan suatu opini objektif dengan topik yang berbeda, serta dapat membantu dalam bidang bisnis untuk menyediakan opini terhadap konsumen mengenai merk dan produk yang populer.

Analisis sentiment atau opinion mining bisa dipakai dalam memperoleh suatu gambaran secara umum persepsi oleh masyarakat terhadap kualitas layanan, apakah cenderung positif, negatif atau netral (Rustam dkk., 2019). Sentimen memiliki dua nilai yaitu positif atau negatif tetapi bisa dengan mengkatagorikan menjadi baik, sangat baik, buruk, dan sangat buruk (Naiknaware, Kawathekar dan Deshmukh, 2017). Beberapa penelitian terdahulu yang telah dilakukan terhadap analisa sentiment Twitter ataupun yang lain dengan metode SVM antara lain sebagai berikut:

- Naive Bayes (NB), SVM. Pada coordinator minister of maritime affairs (Wardhani dkk., 2018).
- Support Vector Machine (SVM), NB, Indonesian presidential (Hayatin, Marthasari dan Nuarini, 2020).
- NB, SVM dan DT pada gojek's service (PD dan Haryoko, 2019).
- NB pada transportasi online (Que, Iriani dan Purnomo, 2020).

NB mempunyai kelebihan kokoh terhadap atribut yang tidak relevan serta Cepat dan Efisiensi ruang (PD dan Haryoko, 2019). Tapi memiliki kelemahan yaitu harus berasumsi antar fitur yang satu dengan fitur yang lain tidak saling berkaitan (*independent*), sementara dalam kenyataannya, keterkaitan itu ada dan Keterkaitan tersebut tidak dapat dimodelkan oleh NB (Luqyana, Cholissodin dan Perdana, 2018; PD dan Haryoko, 2019). DT memiliki kelebihan mudah dalam proses pengelompokan data teks, relatif transparan, mudah dimengerti dan memiliki sejumlah kualitas yang tinggi. Tapi memiliki kelemahan mengeksploitasi fitur yang relatif independen satu sama lain sehingga membatasi kemampuan DT (Alasadi dan Bhaya, 2017). KNN

memiliki kelebihan mudah direpresentasikan, memiliki ketangguhan terhadap training data yang memiliki banyak *noise*, dan dalam proses pengelompokan cukup efektif. Tetapi dalam akurasi memiliki kelemahan, yang disebabkan penetapan nilai k yang sama disemua kategori tanpa adanya perhitungan jumlah dokumen yang dimiliki pada setiap kategori (Que, Iriani dan Purnomo, 2020). SVM mempunyai kelebihan kinerja dalam hasil eksperimen sangat baik serta tingkat ketergantungan rendah pada dataset yang berdimensi. Tapi memiliki kelemahan pada pengkategorian atribut yang kurang optimal jika ada missing value dan memerlukan proses ulang (Hayatin, Marthasari dan Nuarini, 2020).

Penerapan metode SVM dapat menghasilkan tingkat keakuratan yang paling baik dibandingkan metode lainnya, yaitu dengan nilai tingkat keakuratan lebih tinggi (Wardhani dkk., 2018; Iqbal dkk., 2019). Namun SVM juga memiliki kelemahan yaitu sulitnya dalam menentukan fitur yang paling sesuai dan optimal dalam bobot atribut yang dipakai sehingga mengakibatkan rendahnya tingkat nilai akurasi (Al Rivan, Rachmat dan Ayustin, 2020). Metode Grid Search (GS) adalah metode sebagai pemilihan fitur yang terbukti efektif yang bisa membantu pemilihan fitur pada metode klasifikasi seperti SVM (PD dan Haryoko, 2019; Que, Iriani dan Purnomo, 2020). Berdasarkan hal tersebut dalam penelitian akan menerapkan GS untuk memilih fitur SVM yang paling sesuai dan optimal, agar lebih akurat hasil analisis nilai sentimennya.

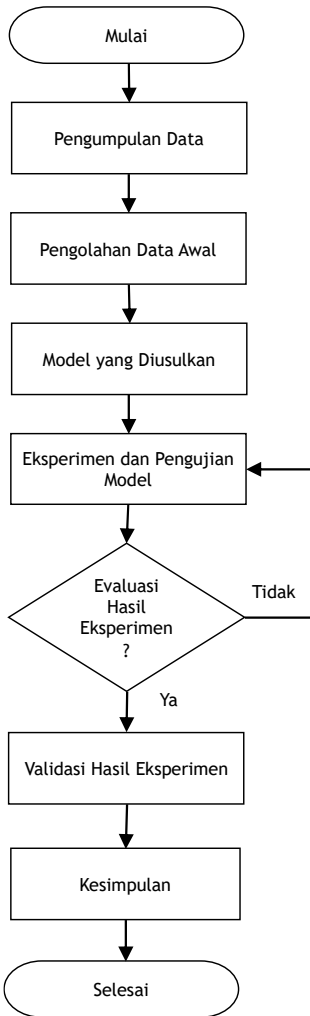
2. METODOLOGI

Dalam menyelesaikan penelitian ini dilakukan beberapa tahap untuk mencapai tujuan seperti yang ditampilkan pada Gambar 1.

2.1. Tahap Pengumpulan Data

Dalam penelitian data yang di gunakan adalah data sekunder, yang mana data ini diperoleh dari halaman web berikut <http://www.sananalytics.com/lab> (Saif dkk., 2013). Dataset Sanders-twitter-0.2 adalah dataset berupa tweets pada Twitter yang diperoleh melalui pencarian dengan kata kunci @apple, #google, #microsoft, #twitter. Dataset dilabel

dengan cara manual yang terdiri atas data 654 negatif, 570 positif, 2503 netral, 1786 irrelevant. Dataset sanders dapat dilihat pada [Tabel 1](#).

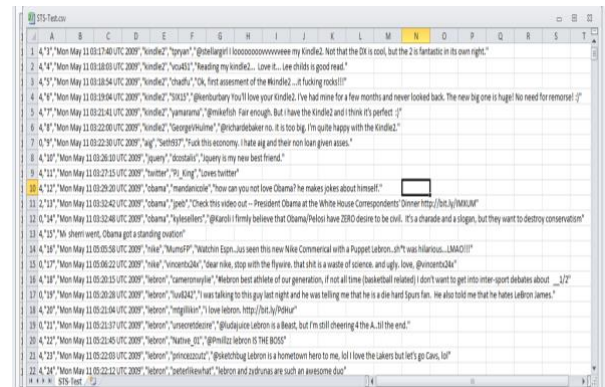


Gambar 1. Tahap penelitian

Tabel 1. Atribut dan data sanders (Saif dkk., 2013)

No	Nama Produk	Sentiment	ID Tweets
1	apple	positive	126415614616154112
2	apple	positive	126404574230740992
3	apple	positive	126402758403305474
4	apple	positive	126397179614068736
5	apple	positive	126395626979196928
6	apple	positive	126394626979196914
7	apple	positive	126395626979196977
8	apple	positive	126395626979196979
9	apple	positive	126395726979296962
10	apple	positive	126395426779196961
11	apple	positive	126495636979496982
12	apple	positive	126395626979196929
13	apple	positive	126395626979196930
14	apple	positive	126395626979196932

Dataset STS-Test adalah dataset berupa tweets pada Twitter yang diperoleh melalui pencarian menggunakan kata kunci identitas produk, nama perusahaan maupun orang. Dataset yang diperoleh dilabel dengan cara manual yang terdiri dari 177 negatif, 182 positif dan 139 netral. Dataset STS-Test ditampilkan pada [Gambar 2](#).



Gambar 2. Dataset STS-test

Dataset tersebut digunakan sebagai data pembandingan dengan dataset utama yaitu dataset Sanders pada tahapan eksperimen dan pengujian

2.2. Pengolahan Data Awal

Dari proses pengumpulan data di peroleh data sebanyak 5.513 data, dari data yang di peroleh tidak semua data dan tidak semua atribut bisa digunakan karena perlu melalui beberapa tahap pengolahan awal data (*preprocessing data*) seperti yang diperlihatkan pada [Gambar 3](#).

Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut:

- 1) **Data Loading:** Data Sanders disimpan dalam format csv, sehingga untuk membaca data tersebut menggunakan pustaka csv pada bahasa python:

```

import csv

data = []

data_file =
csv.reader(open("../Data/Twitter/Sentiment_Analysis/STS-Test.csv"))

data_file.next()

for column in data_file:
    data.append(column[1])
  
```

- 2) **Tokenize:** digunakan untuk memisahkan kata. Atau proses untuk mendapatkan token (term/word) dari dokumen dengan menjadikan spasi dan tanda baca sebagai pemisah antar token
- 3) **Weighting:** yaitu proses pembobotan masing-masing token yang dapat berupa *term count*, *term frequency* (TF), atau *term frequency inversed document frequency* (TFIDF)
- 4) **Filter token:** mengambil kata kata penting hasil tokenizer atau proses penyaringan fitur yang potensial. Misal token yang bukan stopwords, token yang bukan merupakan noise, dan lainnya.
- 5) **Stem:** mencari kata dasar dari setiap kata hasil filter.
- 6) **Filter Stopwords:** menghilangkan kata kata yang sering muncul namun tidak memiliki pengaruh dalam ekstraksi sentiment.
- 7) Proses pembentukan bag-of-words (Feature extraction) dapat menggunakan fungsi CountVectorizer atau TfidfVectorizer pada bahasa python sebagai berikut:

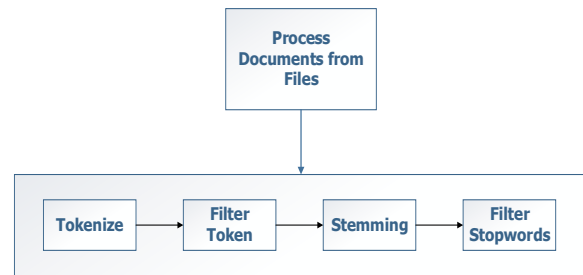
```
import csv
from sklearn.feature_extraction.text
import CountVectorizer, TfidfVectorizer

# Feature Extraction
# -----
#vectorizer = CountVectorizer()
#vectorizer =
TfidfVectorizer(use_idf=False)
#vectorizer = TfidfVectorizer()
vectorizer =
TfidfVectorizer(stop_words='english')
#vectorizer = TfidfVectorizer(min_df=2)
X = vectorizer.fit_transform(data)
```

- 8) Pada tahap penskalaan dalam preprocessing terhadap metode machine learning tertentu, nilai fitur data diharapkan berada pada suatu skala tertentu, seperti bernilai [0,1], [-1,1], atau memiliki mean nol dan standar deviasi satu. Seperti pada script python di bawah ini:

```
from sklearn import preprocessing
# Preprocessing
```

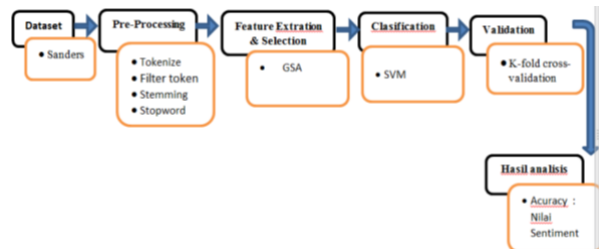
```
#scaler =
preprocessing.StandardScaler(with_mean=False)
#X = scaler.fit_transform(X)
#scaler = preprocessing.MinMaxScaler()
#X =
sparse.csr_matrix(scaler.fit_transform(X.toarray()), dtype=np.float64)
```



Gambar 3. Model proses pengolahan data awal

2.3. Model yang Diusulkan

Berikutnya yaitu tahap pengusulan model, yaitu dengan melakukan pemrosesan data *traning*. Gambar 4 merupakan usulan tahapan model yang di usulkan pada algoritma yang akan dilakukan pengujian.



Gambar 4. Model yang diusulkan

2.4. Eksperimen dan Tahap Pengujian Model

Dalam menyelesaikan analisis sentiment Twitter tahap modeling digunakan dua metode yang pertama algoritma *support vector machine* digunakan untuk clasifikasi data dan yang kedua algoritma *Grid search* yang digunakan dalam memilih fitur.

- 1) **Support Vector Machine(SVM)** merupakan sekumpulan metode yang berkaitan erat dengan metode pembelajaran, yang berhubungan dengan masalah klasifikasi dan regresi (Luqyana, Cholissodin dan Perdana, 2018).

- 2) **Grid Serach Algorithm(GSA)** merupakan salah satu tahapan yang harus dilakuakan berkaitab dengan pemilihan fitur model untuk direkomendasikan dalam metode classification (Elgeldawi dkk., 2021).

2.5. Tahap Evaluasi dan Tahap Validasi

Pada tahap ini Model yang akan di usalkan untuk menyelesaikan penelitian ini adalah dengan menggunakan *support vector machine* dan *support vector machine* berbasis *Grid Search Algorithm* sebagai pemilihan fitur.

Berdasarkan model yang ada tahap selanjutnya adalah melakukan pengujian menggunakan *confusion matrix* untuk mengetahui tingkat akurasi. Pengujian juga diukur dengan menggunakan *ROC Curve*.

Pada tahap awal penggunaan algoritma SVM adalah menentukan besarnya nilai *weight*. Tahap berikutnya darinilai terbesar yang diperoleh dari akurasi dan AUC, nilai *weight* akan di gunakan sebagai dasar nilai dalam mencari nilai tertinggi akurasi dan AUCnya. Sedangkan penerapan algoritma *Support Vector Machine (SVM) berbasis Grid Search Alorghm* merujuk pada algoritma terhadap nilai *weightnya*. Tahapan selanjutnya adalah mencari nilai *weight*, setelah ditemukannya nilai akurasi yang paling ideal dari parameter tersebut.

Tahapan berikutnya setelah didapatkan nilai akurasi yang paling ideal melalui parameter yang ada adalah menentukan nilai *weight* untuk membentuk struktur algoritma yang ideal sehingga diperoleh pemecahan masalah tersebut.

3. HASIL DAN PEMBAHASAN

Berdasarkan Dalam penelitian ini besarnya Nilai akurasi sentiment twitter ditentukan melalui tahapan percobaan dengan memasukkan C, epsilon. Pengujian ini dilakukan dengan metode SVM terhadap nilai akurasi **Weighting, Filtering, Preprocessing (skala)**, serta **selection feature** terhadap dataset yang sudah disediakan diantaranya STS-Test, STS_Gold, Sanders, dan Capres2014. Tools yang digunakan dalam pengujian ini menggunakan Bahasa Pemrograman Python Canopy. Berikut ditampilkan hasil percobaan yang sudah dilakukan dalam penentuan

nilai akurasi sentiment. Untuk loading data menggunakan Dataset STS-Test disimpan dalam format csv, sehingga untuk membaca data tersebut menggunakan pustaka csv:

```
# Data Loading
print "Data Loading ..."
data_file = csv.reader(open("D:\kuliah\KULIAH
S2\data minning\DATA\STS-Test.csv"))
data_file.next()
for column in data_file:
data.append(column[5])
target.append(column[0])
```

3.1. Hasil Pengujian Weighting

Weighting, Yaitu Proses Pembobotan Masing-Masing Token Yang Dapat Berupa Term Count, Term Frequency (Tf), Atau Term Frequency Inversed Document Frequency (Tfidf). Sehingga dari hasil pengujian pada Metode SVM pada proses Weighting dapat dilihat pada Tabel 2.

Tabel 2. Hasil pengujian weighting

Dataset Sanders			
Term	Count	TF	TFIDF
Akurasi	65.99%	66.60%	68.20%

3.2. Hasil Pengujian Filtering

Filtering, yaitu proses penyaringan fitur yang potensial. Misal Token yang bukan *stopwords*, token yang bukan merupakan *noise* dan lainnya serta menggunakan *weighting* TFIDF. Sehingga dari hasil pengujian pada Metode SVM pada proses Filtering dapat dilihat pada Tabel 3.

Tabel 3. Hasil pengujian filtering

Dataset Sanders		
Term	TFIDF	TFIDF
Filtering	Stopword	Stopword
Model seleksi	Grid Search	Two Step GS
Akurasi	69.98%	68.81%

3.3. Hasil Pengujian Preprocessing (Skala)

Preprocessing adalah pra pengolahan data, misal normalisasi, standarisasi, reduksi dimensi. Pada metode machine learning tertentu, nilai fitur data diharapkan berada pada suatu skala tertentu, seperti bernilai [0,1], [-1,1], atau memiliki mean nol dan standar deviasi satu. Pada pengujian Preprocessing ini proses yang digunakan pada Weighting menggunakan TFIDF, proses Filtering menggunakan Stopword dan Skala yang digunakan adalah standar, Stddev=1 dan Min=0 Max=1. Hasil Akurasi Preprocessing (skala) dapat dilihat pada Gambar 5.

```

Python
In [14]: %run "D:\kuliah\KULIAH S2\data mining\Program Python\Program Python\sk-AnalisisSentimen-Svm.py"
Data Loading ...
Learning ...
Nilai C : 0.51365692002
Nilai gamma : 0.840896415254
Estimasi akurasi : 0.688128772636

In [14]: %run "D:\kuliah\KULIAH S2\data mining\Program Python\Program Python\sk-AnalisisSentimen-Svm.py"
Data Loading ...
Learning ...
Nilai C : 64.0
Nilai gamma : 3.62917210389e-05
Estimasi akurasi : 0.649899396378

In [15]: %run "D:\kuliah\KULIAH S2\data mining\Program Python\Program Python\sk-AnalisisSentimen-Svm.py"
Data Loading ...
Learning ...
Nilai C : 4096.0
Nilai gamma : 0.8371627223438
Estimasi akurasi : 0.665995975855
    
```

Gambar 5. Hasil akurasi preprocessing (skala)

Sehingga dari hasil pengujian pada Metode SVM pada proses Preprocessing dapat dilihat pada Tabel 4.

Tabel 4. Hasil pengujian preprocessing

Dataset Sanders			
Term	TFIDF	TFIDF	TFIDF
filtering	Stopword	Stopword	Stopword
Skala	None	Stddev=1	Min=0, Max=1
Akurasi	68.81%	64.98%	66.59%

3.4. Hasil Pengujian Model Selection

Model Selection adalah penentuan parameter model yang tidak dipelajari secara langsung pada tahapan learning, seperti orde polinomial M dan parameter regularisasi λ .

Pada pengujian model seleksi yang digunakan adalah untuk Weighting menggunakan TFIDF, Filtering menggunakan Stopword dan model seleksi yang digunakan adalah Grid Search dan

Two Step Grid Search. Hasil akurasi model seleksi dapat di lihat pada Gambar 6.

```

Python
In [2]: %run "D:\kuliah\KULIAH S2\data mining\Program Pyth
Data Loading ...
Learning ...
Nilai C : 1.68179283051
Nilai gamma : 0.420448207627
Estimasi akurasi : 0.698189134809

In [3]: %run "D:\kuliah\KULIAH S2\data mining\Program Pyth
Data Loading ...
Learning ...
Nilai C : 181.019335984
Nilai gamma : 0.594603557501
Estimasi akurasi : 0.688128772636

In [4]:
    
```

Gambar 6. Hasil akurasi model seleksi

Sehingga dari hasil pengujian pada Metode SVM pada proses model seleksi fitur dapat dilihat pada Tabel 5.

Tabel 5. Hasil pengujian model selection

Dataset Sanders			
Term	TFIDF	TFIDF	TFIDF
Filtering	None	Stopword	Min df=2
Akurasi	68.20%	69.41%	66.80%

3.5. Hasil Pengujian

Dari hasil pengujian berdasarkan dataset Sanders antara lain dengan penggunaan SVM dengan menggabungkan metode MaxEntropy dan Naïve Bayes dapat di lihat pada Gambar 7.

```

Accuracy 0.787587
Confusion Matrix
      |   n   e   n   o   |
      |   g   e   s   |
      |   a   u   i   |
      |   t   t   t   |
      |   i   r   i   |
      |   v   a   v   |
      |   e   l   e   |
-----+-----+
negative | <.> 359   . |
neutral  | .<2373> . |
positive | . 281 <.> |
-----+-----+
(row = reference; col = test)
    
```

Gambar 7. Hasil pengujian metode klasifikasi SVM berbasis GA

4. SIMPULAN

Dengan data *training* sebanyak 5513 tersusun atas 654 negatif, 570 positif, 2503 netral, 1786 irrelevant. Disimpan kedalam format csv dan terdiri dari 5 field: 0 : Tweet id 1 : Tweet text 2 : Tweet creation date 3 : Topic used for sentiment 4 : Sentiment label ('positive', 'neutral', 'negative', 'irrelevant'). Dalam mengklasifikasi data testing yang tesusun atas kfoldcross=10 maka nilai akurasi hingga 79,00%.

Akurasi mencapai hingga 69,22% setelah metode preprocessing di tambahkan. Sedangkan melalui metode *stop word removal* akurasi dapat mencapai hingga 73,74%. Kemudian melalui metode *Grid Search* dan *Two Step Grid Search* di dapatkan nilai akurasi hingga 73,89%. dari keseluruhan optimasi diperoleh nilai akurasi hingga 80,33%.

Untuk penelitian lanjut dapat dikembangkan *system online* yang dapat mengampu *dataset* dan memasukan sentiment secara *real time*.

DAFTAR PUSTAKA

- Al Rivan, M.E., Rachmat, N. dan Ayustin, M.R. (2020) 'Klasifikasi Jenis Kacang-Kacangan Berdasarkan Tekstur Menggunakan Jaringan Syaraf Tiruan', *Jurnal Komputer Terapan*, 6(1), hal. 89-98.
- Alasadi, S.A. dan Bhaya, W.S. (2017) 'Review of data preprocessing techniques in data mining', *Journal of Engineering and Applied Sciences*, 12(16), hal. 4102-4107.
- Elgeldawi, E. dkk. (2021) 'Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis', *Informatics*, 8(79), hal. 1-21.
- Hayatin, N., Marthasari, G.I. dan Nuarini, L. (2020) 'Optimization of Sentiment Analysis for Indonesian Presidential Election using Naïve Bayes and Particle Swarm Optimization', *Jurnal Online Informatika*, 5(1), hal. 81-88.
- Iqbal, F. dkk. (2019) 'A hybrid framework for sentiment analysis using genetic algorithm based feature reduction', *IEEE Access*, 7, hal. 14637-14652.
- Luqyana, W.A., Cholissodin, I. dan Perdana, R.S. (2018) 'Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine', *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11), hal. 4704-4713.
- Naiknaware, B.R., Kawathekar, S. dan Deshmukh, S.N. (2017) 'Sentiment analysis of indian government schemes using twitter datasets', *IOSR Journal of Computer Engineering*, 1, hal. 70-78.
- PD, W.M. dan Haryoko (2019) 'Optimization of parameter support vector machine (SVM) using genetic algorithm to review Go-Jek's services', in *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*. 2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia: IEEE, hal. 301-304.
- Que, V.K.S., Iriani, A. dan Purnomo, H.D. (2020) 'Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization', *Jurnal Nasional Teknik Elektro dan Teknologi Informasi* | Vol, 9(2), hal. 162-170.
- Rustam, F. dkk. (2019) 'Tweets classification on the base of sentiments for US airline companies', *Entropy*, 21(1078), hal. 1-22.
- Saif, H. dkk. (2013) 'Evaluation datasets for Twitter sentiment analysis: a survey and a new dataset, the STS-Gold', in *Social and Expressive Media: Approaches and Perspectives from AI (ESSEM 2013)*. 1st Interantional Workshop on Emotion and Sentiment, Italy: The Open University's, hal. 1-13.
- Wardhani, N.K. dkk. (2018) 'Sentiment analysis article news coordinator minister of maritime affairs using algorithm naive bayes and support vector machine with particle swarm optimization', *J. Theor. Appl. Inf. Technol.*, 96(24), hal. 8365-8378.

