

ANALISIS SENTIMEN PROGRAM BANTUAN SOSIAL TUNAI PADA SOSIAL MEDIA TWITTER MENGGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE*

Muhammad Fadilah Dzukaidah
Program Studi Informatika, Fakultas Ilmu Komputer
Universitas Pembangunan Nasional Veteran Jakarta
Jalan RS. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12450
mfadilahdz@gmail.com

Abstrak. Pada masa ini pertumbuhan media online sebagai alat komunikasi sudah sangat pesat, bermacam-macam informasi dan konten yang tersedia menjadi daya tarik tersendiri bagi penggunanya. Selain berbagai informasi media online juga dapat menjadi sarana yang digunakan untuk menyampaikan isu-isu, kritik, saran serta opini-opini publik. Sebagai salah satu media online tersebut yaitu twitter. Pada twitter terdapat tweet yang dicirikan dalam dua kelas sentimen, yang merupakan sentimen positif dan sentimen negatif. Perhitungan pada opini ini menggunakan Support Vector Machine (SVM), perhitungan ini dilakukan agar dapat memproses urutan perasaan dalam tweet. Data ini diperoleh dengan memanfaatkan antarmuka Pemrograman Application Programming Interface (API) pada Twitter. Di dapat data jumlah tweet 237 sebagai data latih dan 60 sebagai data uji agar mendapatkan hasil pengujian dengan algoritma Support Vector Machine sebesar 88,33% nilai akurasi 91,37% presisi dan 96,36% nilai recall. Besarnya akurasi menunjukkan pada algoritma Support Vector Machine dapat digunakan dalam klasifikasi terhadap Program Bantuan Sosial Tunai.

Kata Kunci: Klasifikasi, *Support Vector Machine*, *Tweet*.

1 Pendahuluan

Pada zaman era digital pertumbuhan media online sebagai alat komunikasi sangat pesat, bermacam-macam informasi dan konten yang tersedia menjadi daya tarik tersendiri bagi penggunanya. Selain berbagai informasi media online juga dapat menjadi sarana yang digunakan untuk menyampaikan isu-isu, kritik, saran serta opini-opini publik. Sebagai salah satu media online tersebut yaitu twitter.

Twitter merupakan aplikasi perangkat lunak yang secara langsung sering dipilih masyarakat sebagai salah satu alat komunikasi yang sedang marak digunakan. Twitter juga dapat mengirim pesan pendek yang dengan Batasan karakter maksimal 140 karakter.

Dari opini itulah dapat terlihat mana yang bersifat positif ataupun negatif. Tetapi melalui opini tersebut, Banyaknya opini tersebut menyulitkan pengguna pada media tersebut mengambil kesimpulan. Hal ini mempengaruhi hasil dari akurasi nilai.

Pada media online manapun belum terdapat adanya fitur tersebut. Agar mendapatkan akurasi nilai dapat melakukan analisis sentimen dan pada opini dilakukan klasifikasi terhadap tweet untuk mendapatkan dua nilai kelas yaitu sentimen positif atau sentimen negatif. Dalam melakukan klasifikasi diperlukan metode Support vector machine (SVM) supaya dapat mengelompokkan kedalam dua nilai kelas yaitu kelas sentimen positif dan kelas sentimen negatif. Pada proses klasifikasi Svm dapat menaikkan jumlah akuratnya nilai, maka dari itu para pengguna memperoleh data nilai positif atau negatif dengan presisi tinggi dari pembagian klasifikasi.

Pada penelitian ini Support vector machine (SVM) mengklasifikasikan beberapa model, yang akan dibandingkan dengan model lain, hasil yang diperoleh akan lebih akurat. Solusi ini dapat memainkan peran terbaik pada data berdimensi tinggi, sehingga dianggap cocok untuk klasifikasi analisis sentimen. Ada beberapa jenis model kernel di SVM, yang menghasilkan nilai presisi yang lebih tinggi dan dapat disesuaikan berdasarkan data pada penelitian.

2 Landasan Teori

2.1 Analisis Sentimen

Penggalan opini adalah topik penelitian penting yang saat ini sedang berlangsung. Penambangan opini adalah cabang penelitian dari penambangan teks. Opinion mining mencakup tiga sub-proses, yaitu subjektivitas dokumen, orientasi opini, dan deteksi target. Cari tahu bagian mana dari kalimat yang menjadi objek opini, opini negatif atau opini positif. [6]

Dapat diartikan Analisis sentimen proses menganalisis topik dari dokumen teks dalam sistem penelitian untuk mencari tahu bagian mana dari kalimat yang menjadi nilai positif atau negatif.

2.2 Text analysis

Text Analysis merupakan teknik penelitian yang berfokus pada identifikasi teks supaya mendapatkan pengetahuan serta informasi yang baru. [1]

Berdasarkan kutipan diatas text analysis yaitu tahapan pencarian informasi yang berbentuk teks dari data yang belum terstruktur.

2.3 Text Mining

Informasi berkualitas tinggi biasanyaskaaaaadasa diperoleh dengan memprediksi pola dan tren melalui pembelajaran model statistik. Proses penambangan teks yang umum termasuk klasifikasi teks, ekstraksi konsep/entitas, pembuatan taksonomi granular, penalaran pada dokumen, dan hubungan entitas pada pemodelan.[4]

Jadi text mining adalah proses mengambil informasi untuk pengolahan data yang berbentuk teks menjadi informasi berkualitas tinggi dengan sebuah informasi yang diinginkan.

2.3.1 Pra proses

Pra proses data adalah teknik data mining awal yang digunakan untuk mengubah data mentah yang dikumpulkan dari berbagai sumber atau biasa dikenal dengan data mentah menjadi informasi yang lebih jelas, agar dapat digunakan untuk diproses selanjutnya.

2.3.2 Case folding

Case folding Merupakan bentuk pra-pemrosesan teks yang sangat sederhana dan paling efektif, biarpun terabaikan. Case Folding itu sendiri berfungsi mengubah huruf kapital menjadi huruf kecil, maka semua huruf besar diubah semua sehingga dapat dilakukan tahap klasifikasi.

2.3.3 Normalisasi kata

Kata-kata tidak baku yang sering muncul dalam tweet, oleh sebab itu untuk kata yang tidak baku atau tidak ada pada kamus akan diterjemahkan dengan kata terdekat menggunakan kamus yang telah dibuat.

2.3.4 Stemming

Stemming adalah proses menghapus kata yang terdapat imbuhan dengan Mendeteksi kata dasar dari sebuah kata yang dinamakan Teknik Stemming. [7]

Jadi, stemming merupakan proses untuk mencari kata dasar yang berisi imbuhan supaya dapat menyederhanakan kata dari hasil proses kata tokenisasi.

2.3.5 Stopword Removal

Proses penyaringan yang kata yang bertujuan untuk mengetahui apakah suatu kata termasuk dalam stopword (stoplist). Contoh stopword: ini, dia, untuk, aku dan sebagainya.

2.3.6 Tokenisasi

Bertujuan untuk menjadikan sebuah kalimat yang bermakna dengan cara memecahkan kalimat menjadi kata untuk memaksimalkan hasil klasifikasi.

2.4 Pembobotan TF-IDF

TF-IDF merupakan metode yang digunakan untuk membobotkan kata pada sebuah dokumen yang berdasarkan banyaknya jumlah kata untuk menentukan bobot dalam kata. [5]

Ada dua proses perhitungan untuk pembobotan TF-IDF, yang pertama adalah proses pelatihan, dimana algoritma klasifikasi menganalisis data pelatihan yang ada. Setelah analisis data latih selesai, data akan disajikan dalam bentuk aturan klasifikasi. Kemudian masuk ke proses kedua, tahap klasifikasi, di mana tes data digunakan untuk memprediksi keakuratan proses klasifikasi. TF-IDF berbobot memiliki persamaan berikut:

Keterangan:

tf : Banyaknya kata pada data dokumen
idf : Banyaknya Dokumen
D : Banyaknya dataset
DF : Banyaknya dataset berisiterm (t)

$TF-IDF = tf * idf$

$IDF = \log \left[\frac{D}{DF} \right]$

2.5 Support Vector Machine

Analisis sentimen menggunakan cara mengklasifikasikan sentimen ke dalam kelas positif / negatif dengan algoritma Support Vector Machine.[2] Support Vector Machine akan mengklasifikasikan apakah opini publik termasuk dalam sentimen positif atau negatif dengan mencari hyperlane terbaik dalam dua kategori klasifikasi. Menambahkan fungsi kernel ke Support Vector Machine dukungan berguna untuk memproses data yang dipisahkan secara non-linear.[3]

Dapat di simpulkan support vector machine adalah teknik klasifikasi dan regresi yang digunakan untuk memprediksi nilai kasus tertentu yang berfungsi untuk memisahkan variabel-variabel dengan nilai optimal atau biasa disebut dengan hyperlane. Bentuk hyperplane yaitu line dua dimensi sedangkan flat plane lebih dari dua dimensi. Hyperlane adalah generalisasi garis lurus dalam ruang dua dimensi, flat plane dalam ruang tiga dimensi, dan ruang arsitektur dalam ruang empat dimensi. Dalam persamaan ruang kartesius dijelaskan bahwa $ax+by+c=0$ adalah persamaan linear dalam ruang dua dimensi. Dari persamaan ruang dapat membuat persamaan ruang lebih dari dua dimensi, dengan mengubah notasi_variabel dan konstanta menjadi x menjadi x_1 , y menjadi x_2 , a menjadi w_1 , b menjadi w_2 .

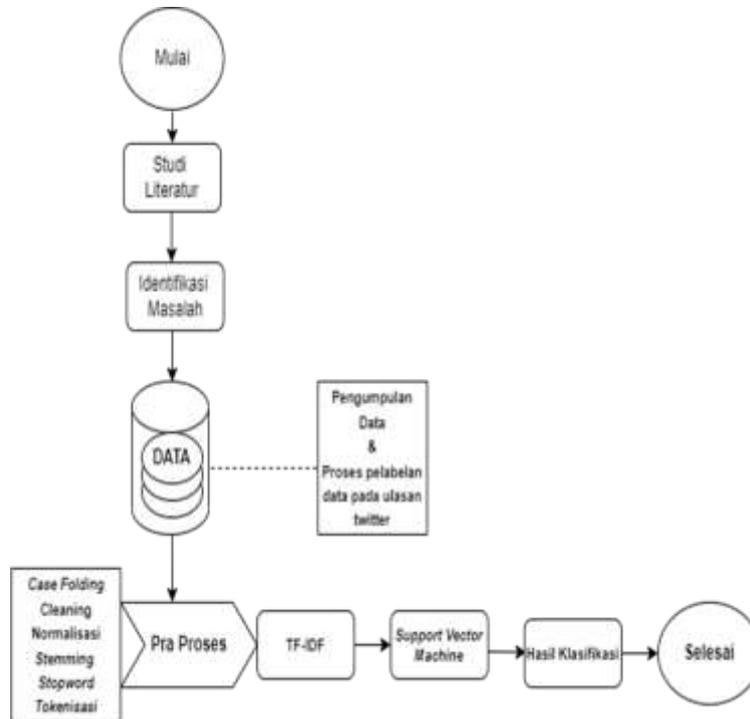
2.6 Python

Python adalah bahasa pemrograman untuk perangkat lunak yang tidak ada batasan untuk menyalin atau mendistribusikan dan distandarisasi untuk disertakan dalam paket distribusi. Berisi source codenya, debugger, dan penganalisis, yang berisi antarmuka untuk antarmuka layanan. [8]

Berdasarkan pernyataan diatas python dapat digunakan di berbagai operasi sistem, dengan bahasa pemrograman yang berisi source codenya, debugger dan sebagainya.

3 Metode Penelitian

3.1 Kerangka Pikir



Gambar 1. Kerangka Pikir

Pada gambar diatas merupakan kerangka pikir yang berisi proses tahapan peneliti guna mencapai hasil klasifikasi yang telah menjadi tujuan penelitian berikut yang akan di jabarkan subbaw ini.

3.1.1 Studi literatur

Penelitian literatur dalam penelitian ini berguna untuk mencari sumber literatur tentang analisis sentimen, text mining, metode klasifikasi, dan algoritma support vector machine yang telah dibahas pada penelitian sebelumnya. Menggunakan situs e-book, web, dan jurnal terkait untuk mencari dan mendiskusikan dokumen terkait.

3.1.2 Identifikasi Masalah

Peneliti mencari atau menemukan masalah yang ada, lalu mengidentifikasi sumber permasalahan tersebut, kemudian akan di proses dalam penelitian ini. Dengan menggunakan algoritma Support Vector Machine untuk mengklasifikasikan tweet ulasan terhadap bantuan sosial tunai.

3.1.3 Akuisisi Data

Akuisisi data berfungsi untuk pengumpulan data yang menggunakan indeks informasi dari tweet pengguna twitter pada halaman URL dengan menyertakan fitur API (Application Programming Interface) sebagai data latih dan uji untuk penambahan dataset yang lebih besar.

3.1.3.1 Pengumpulan Data

Pada penelitian ini pengumpulan data dilakukan pada tahap pertama yang dilakukan dengan mencari 1000 data yang di sortir menjadi 297 data. Scraping data ini dilakukan menggunakan API Twitter yang memakai bahasa python. Proses Scraping ini dilakukan secara otomatis untuk mendapatkan data review Bantuan Sosial Tunai yang melalui komentar pada Twitter. Proses ini mendapatkan hasil akhir sebuah kumpulan banyaknya data yang berformat CSV.

3.1.3.2 Pelabelan Data

Pada tahap ini yaitu mengidentifikasi review yang telah dilakukan pada scaping dengan mengumpulkan data tersebut secara manual pada setiap komentar yang hanya dapat memiliki satu nilai. Yang mana nilai tersebut berbentuk positif atau negatif. Pelabelan data ini untuk menjaga substansi informasi review, pelabelan ini dilakukan sebelum tahap pra proses. Semua ulasan pada tahap ini hanya dibagi menjadi dua kategori, ulasan yang bernilai positif dan ulasan yang bernilai negatif.

Tabel Error! No text of specified style in document..1.2 Pelabelan Data

Tweet	Anatator 1	Anatator 2	Anatator 3	Final Anatator
Salah satu contoh isi tweet dari sebuah ulasan	Positif	Positif	Positif	Positif
Salah satu contoh isi tweet dari sebuah ulasan	Positif	Positif	Negatif	Positif
Salah satu contoh isi tweet dari sebuah ulasan	Positif	Negatif	Negatif	Negatif
Salah satu contoh isi tweet dari sebuah ulasan	Negatif	Negatif	Negatif	Negatif

Pelabelan data pada tabel diatas dilakukan oleh tiga anatator yang berbeda agar menghindari subjektivitas dan objektivitas dari pemberian klasifikasi sentimen pada ulasan twitter. Dari ketiga nilai sentimen yang diberikan oleh masing masing responden akan di dapatkan hasil final atau suatu kesimpulan dari isi nilai ulasan tweet seperti pada tabel 3.1.3.1 Final nilai dari sebuah ulasan ialah positif jika terdapat tiga atau dua user yang menyampaikan nilai positif terhadap ulasan data twitter yang sama, dan juga sebaliknya final nilai dari sebuah ulasan merupakan negatif jika terdapat tiga atau dua user yang menyampaikan nilai negatif pada ulasan yang sama.

Pada pelabelan dilakukan dengan tiga orang atau lebih, di hindarkan bahwa ada kesalahan yang terjadi, maka metode yang dikenal sebagai leiss kappa vallue digunakan untuk melakukan dan memeriksa reliabilitas antar anatator.

3.2 Alat Yang Digunakan

Alat Pendukung pada Penelitian ini menggunakan (Hardware dan software) yang di pakai dalam pembuatan laporan tugas akhir yaitu;

a) Perangkat Keras

Spesifikasi Perangkat Keras (laptop) yang digunakan sebagai berikut:

- 1) Sistem Model : Laptop Asus X442URR
- 2) Processor : Intel Core i5-8250U CPU @ 1,6 GHz (8CPUs)
- 3) Memory : RAM 8 Gb
- 4) Hardisk : 1TB

b) Perangkat Lunak

Perangkat Lunak yang dipakai adalah:

- 1) Operating System: Windows 11 Home Single Language 64-bit (10.0, Build 22000)
- 2) Aplikasi yang dibutuhkan dalam penelitian:
 - Microsoft Word 2019
 - Python dan Google Colab

4 Hasil dan Pembahasan

4.1 Data

Dalam penelitian ini data berasal dari situs web mikroblog Twitter, dalam penelitian ini merupakan hasil crawling dengan memakai fitur Application Programming Interface (API) yang telah menjadi bagian dari situs web mikroblog Twitter. Proses crawling data dalam penelitian ini menggunakan bahasa pemrograman Python yang dibantu oleh perangkat lunak google colab serta library – library pendukung yang ada saat ini. Hasil crawling data disimpan dalam bentuk Comma-Separated Values (.csv).

Penelitian ini menggunakan data yang berupa tweet pada twitter yang memuat #bantuansosialtunai dan #bst. Hasil proses crawling data ini selanjutnya akan diberi label tweet positif dan negatif.

b'RT @tatakujiyati: ""Bantuan sosial tunai tak boleh disambungkan dg persyaratan apapun," karena bantuan sosial dimaksudkan untuk menyambung hidup""

Sifa: b'BPK juga menemukan adanya indikasi bansos yang tidak sesuai ketentuan dalam penyalurannya hingga kasus yang bermasalah kepada penerimanya. Untuk bantuan Sosial Tunai (BST) saja yang tidak sasaran

Gambar Error! No text of specified style in document. Data

4.1.1 Pelabelan Data

Pelabelan data ini merupakan teks atau ulasan yang memerlukan bantuan dari spesialis multibahasa yang secara otomatis yang dapat mengembangkan struktur klasifikasi ulasan ini. Tetapi dalam penelitian ini, prosedur kategori ini dilakukan secara manual oleh tiga orang anatator yang berlatar infomatika dan public relation. Kategori ini membagi kumpulan data ulasan menjadi dua kategori yaitu positif dan negatif. Hasil final diambil berdasarkan jumlah kelas yang mayoritas di antara ketiga anatator tersebut.

Tabel Error! No text of specified style in document..2 Pelabelan Data

Tweet	Anatator 1	Anatator 2	Anatator 3	Final Anatator
Bantuan sosial tunai tak boleh disambungkan dg persyaratan apapun karena bantuan sosial dimaksudkan untuk menyambung hidup.	1	1	1	1
BPK juga menemukan adanya indikasi bansos yang tidak sesuai ketentuan dalam penyalurannya hingga kasus yang bermasalah kepada penerimanya. Untuk bantuan Sosial Tunai (BST) saja yang tidak sasaran	0	0	0	0

4.2 Pra proses

Setelah data diperoleh melalui proses crawling menggunakan Application Programming Interface yang telah menjadi fitur dari Twitter. Data akan melalui tahap pra proses, tahap ini penting karena mempengaruhi nilai akurasi. Pra proses sendiri adalah proses menghilangkan noise dari data untuk mendapatkan fitur utama dari sebuah teks dokumen, proses ini melalui lima tahap, yaitu Case Folding, Noise Removal, Stopword Removal, Stemming dan Tokenisasi.

4.3 Pembobotan kata TF-IDF

Data dalam tahap pra proses mendapatkan hasil yang diisolasi pada sebuah kalimat. Tetapi, menjelang berakhirnya pra proses, setiap kata tidak mempunyai bobot serta nilai, sehingga diperlukan proses pembobotan kata dengan menggunakan metode TF-IDF. Proses ini membuat kata yang terkandung dalam tweet mendapat nilai dan kemudian melakukan klasifikasi yang dimana dapat menambah dari hasil klasifikasi tersebut.

Tabel 4.3.1 Pelebelan tweet

Data Latih Pra proses		
Dokumen	Sudah di Pra proses	Label
D1	['bantu', 'sosial', 'tunai', 'cair']	Positif
D2	['bantu', 'sosial', 'tunai', 'sambung', 'syarat', 'bantu', 'sosial', 'maksud', 'sambung', 'hidup']	Positif
D3	['bantu', 'sosial', 'tunai', 'sifat', 'populis', 'gambar', 'raja', 'perkara', 'utama']	Negatif
D4	['anjing', 'keluarga', 'sesak', 'sekarat', 'bicara', 'bilang', 'jahat', 'kerja', 'harap', 'kasih', 'kaya', 'bantu', 'perintah', 'marah', 'bantu', 'langsung', 'tunai', 'kaca', 'nenek']	Negatif

Tabel Error! No text of specified style in document..3 Nilai TF-IDF

Kata	TF				D F	IDF	TF-IDF			
	D 1	D 2	D 3	D 4			D 1	D 2	D 3	D 4
bantu	1	2	1	2	4	0	0	0	0	0
bicara	0	0	0	1	1	0.602	0	0	0	0.602
bilang	0	0	0	1	1	0.602	0	0	0	0.602
Cair	1	0	0	0	1	0.602	0.602	0	0	0
gambar	0	0	1	0	1	0.602	0	0	0.602	0
harap	0	0	0	1	1	0.602	0	0	0	0.602
hidup	0	1	0	0	1	0.602	0	0.602	0	0
jahat	0	0	0	0	1	0.602	0	0	0	0.602
kasih	0	0	0	1	1	0.602	0	0	0	0.602
kerja	0	0	0	1	1	0.602	0	0	0	0.602
langsung	0	0	0	1	1	0.602	0	0	0	0.602
maksud	0	1	0	0	1	0.602	0	0.602	0	0
marah	0	0	0	1	1	0.602	0	0	0	0.602
nenek	0	0	0	1	1	0.6020	0	0	0	0.602
populis	0	0	1	0	1	0.602	0	0	0	0.602
perkara	0	0	1	0	1	0.602	0	0	0.602	0
raja	0	0	1	0	1	0.602	0	0	0.602	0
sambung	0	2	0	0	2	0.301	0	0.301	0	0
sekarat	0	0	0	1	1	0.602	0	0	0	0.602
syarat	0	1	0	0	1	0.602	0	0	0	0.602
sesak	0	0	0	1	1	0.602	0	0	0	0.602
sosial	1	2	1	0	4	0	0	0	00	0
tunai	1	1	1	1	4	0	0	0	0	0
utama	0	0	1	0	1	0.602	0	0	0.602	0

4.4 Proses Klasifikasi

Setelah kata diberi nilai atau bobot memakai TF-IDF berikutnya nilai pada kata mulai dikelompokkan berdasarkan kelasnya dan diakumulasikan. Pada penelitian sentimen dalam algoritma SVM dengan kernel Radial Basis Function (RBF) akan dihitung otomatis pada algoritma SVM tersebut menggunakan library SCIKIT – Learn. Dengan membagi data latih dan data uji dengan test size = 0,20 dan rs 42. Dengan hasil data latih 80% dan data uji 20% dari jumlah data hasil crawling tweet. Penelitian ini melakukan percobaan pada 4 kernel yaitu kernel Linear, Sigmoid, RBF dan juga poly.

<i>Kernel</i>	<i>Akurasi</i>
<i>Linear</i>	86,67%
<i>Sigmoid</i>	85,00%
<i>Radial Basis Function</i>	88,33%
<i>Polynomial</i>	86,66%

4.5 Hasil Klasifikasi

Hasil akhir pada analisis sentimen ini merupakan hasil klasifikasi SVM dengan 4 kernel yaitu, linear, sigmoid Radial Basis Function dan polynomial. Pada analisis ini peneliti menggunakan 80% data latih dan 20% data uji dimana 237 sebagai data latih dan 60 sebagai data uji yang berlabel positif dan negatif. Setelah dilakukan klasifikasi SVM dengan 4 kernel terlihat bahwa kernel yang memiliki akurasi tertinggi yaitu kernel RBF (Radial Basis Function).

4.6 Evaluasi

Setelah mendapatkan hasil klasifikasi, selanjutnya peneliti melakukan evaluasi terhadap proses klasifikasi. Evaluasi dilakukan dengan menggunakan confusion matrix untuk menguji akurasi algoritma Support Vector Machine dengan harapan peneliti dapat melihat kinerja klasifikasi terhadap data uji. Dapat disimpulkan bahwa nilai true positive 53, true negative 0, false positive 2 dan false negative 5. akurasi pada program ini sebesar 88,33 % dengan nilai presisi 91,37% dan juga recall sebesar 96,36%. Besarnya akurasi algoritma SVM pada kernel RBF sebesar 88,33% maka dapat disimpulkan bahwa metode ini dapat memenuhi Analisis sentimen pada Program Bantuan Sosial Tunai dengan efektif.

5 Penutup

5.1 Kesimpulan

Pada bab ini, pembahasan yang ada pada bab sebelumnya memuat kesimpulan pada penelitian serta saran yang telah dilakukan penulis berdasarkan analisa.

1. Penerapan algoritma Support Vector Machine (SVM) dalam klasifikasi sentimen tweet terhadap Bantuan Sosial Tunai dikerjakan dengan baik.

2. Algoritma SVM dalam klasifikasi sentimen tweet pada Bantuan Sosial Tunai mendapatkan nilai akurasi sebesar 88,33% dengan nilai presisi sebesar 91,37% dan nilai recall sebesar 96,36%.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, penulis memiliki beberapa saran guna untuk mengembangkan penelitian selanjutnya, yaitu:

1. Meningkatkan akurasi klasifikasi dengan menambah jumlah data latih serta mencari formula terbaik dengan menggunakan kernel lain
2. Menambahkan kelas netral pada tahapan pebelan data
3. Menaikkan jumlah data tweet agar memaksimalkan proses klasifikasi supaya meningkatkan akurasi.

Referensi

- [1] Bratawisnu, M.K., Rinaldi, R. and Firdaus, M.F. (2018) "Analisa Persepsi Customer Feedback e-commerce Tokopedia dan Bukalapak Menggunakan Text Network Analysis," *Journal of Information Engineering and Educational Technology*, 2(1). doi:10.26740/jieet.v2n1.p6-12.
- [2] Fachrurrozi, M. et al. (2015) "Analisis Sentimen Pengguna Jejaring Sosial Menggunakan Metode Support Vector Machine," *ePrints Online Repository Sriwijaya University*, 1(1).
- [3] Lukmana, D.T., Subanti, S. and Susanti, Y. (2019) "Analisis Sentimen Terhadap Calon Presiden 2019 Dengan Support Vector Machine Di Twitter," *Seminar Nasional Penelitian Pendidikan Matematika (SNP2M) 2019 UMT [Preprint]*, (2002).
- [4] Pascasarjana, P. and Udayana, U. (2011) "Text mining dengan metode naïve bayes classifier dan support vector machines untuk sentiment analysis," *Universitas Stuttgart [Preprint]*.
- [5] Rahimi, F. and Asyikin, A.N. (2015) "Aplikasi Penilaian Ujian Essay Otomatis Menggunakan Metode Cosine Similarity," *Poros Teknik*, 7(2).
- [6] Rozi, I., Pramono, S. and Dahlan, E. (2012) "Implementasi Opinion Mining (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik Pada Perguruan Tinggi," *Jurnal EECCIS*, 6(1).
- [7] Sholihin, A., Solihin, F. and Rachman, F.H. (2013) "Penerapan Modifikasi Metode Enhanced Confix Stripping Stemmer Pada Teks Berbahasa Madura," *Jurnal Sarjana Teknik Informatika*, 2(1).
- [8] Wijaya, I.W.S. et al. (2021) "Program Menghitung Banyak Bata pada Ruangan Menggunakan Bahasa Python," *TTERS Information Technology Journal*, 2(1).