



Analisis dan Implementasi Metode Naïve Bayes dan SVM Pada Sentimen Pemilihan Calon Presiden RI

Mediana Zalukhu

Universitas Pembangunan Masyarakat Indonesia (UPMI), Indonesia, medianazalukhu@gmail.com

Info Artikel

Kata Kunci:

Analisis, Implementasi, Naive Bayes, SVM, Pemilihan

Keywords:

Analysis, Implementation, Naive Bayes, SVM, Selection



Lisensi: cc-by-sa

Copyright © 2023 Mediana Zalukhu

Abstrak

Analisis sentimen merupakan ilmu yang berguna untuk menganalisis pendapat seseorang, sentiment seseorang, evaluasi seseorang, sikap seseorang dan emosi seseorang ke dalam bahasa tertulis. Penelitian ini menggunakan search techniques dalam pengambilan data, sehingga pengambilan data dilakukan dengan efektif dan efisien. Search techniques dalam penelitian ini menggunakan Boolean searching dengan operator "AND". Data yang sudah didapat dilabeli positif, netral dan negatif oleh penulis kemudian dikoreksi oleh ahli bahasa. Setelah itu dilakukan preProcessing baik itu mengubah kata tidak baku menjadi baku atau biasa disebut normalisasi menggunakan kamus dan mencari akar kata yaitu stemming dengan bantuan aplikasi Sastrawi Master. Selanjutnya dilakukan juga tokenisasi N-Gram, Unigram, Bigram, dan Trigram terhadap kalimat, kemudian menghilangkan kata-kata yang umum digunakan dan tidak mempunyai Informasi yang berharga pada suatu konteks atau biasa disebut stopword removal, dan mempertahankan emoticon karena emoticon merupakan simbol yang menunjukkan ekspresi seseorang ke dalam tulisan. Akurasi yang terbaik dalam penelitian ini adalah dengan dilakukan normalisasi dan stemming pada data sebesar 89,2655% menggunakan metode SVM, dan kemudian data yang dinormalisasi saja sebesar 88,7006% menggunakan metode SVM. Dalam penelitian ini, tidak ada ujicoba terhadap data yang dilakukan stemming saja, dikarenakan tahap yang harus dilakukan dalam stemming adalah melakukan normalisasi terlebih dahulu terhadap data.

Abstract

Sentiment analysis is a science that is useful for analyzing someone's opinion, someone's sentiment, someone's evaluation, someone's attitude and someone's emotions into written language. This research uses search techniques in data collection, so that data collection is carried out effectively and efficiently. Search techniques in this research use Boolean searching with the "AND" operator. The data that has been obtained is labeled positive, neutral and negative by the author and then corrected by a linguist. After that, preprocessing is carried out, namely changing non-standard words to standard or what is usually called normalization using a dictionary and looking for the root of the word, namely stemming, with the help of the Sastrawi Master application. Next, N-Gram, Unigram, Bigram, and Trigram tokenization is also carried out on sentences, then removing words that are commonly used and do not have valuable information in a context or usually called stopword removal, and retaining emoticons because emoticons are symbols that show expressions. someone into writing. The best accuracy in this research was by normalizing and stemming the data at 89.2655% using the SVM method, and then just normalizing the data at 88.7006% using the SVM method. In this research, there were no trials on the data that were carried out by stemming alone, because the step that must be carried out in stemming is to normalize the data first.

1. PENDAHULUAN

Isi Tanggal 20 Oktober 2014, Indonesia telah melewati pesta rakyat, yaitu pelantikan Presiden dan Wakil Presiden terpilih Joko Widodo - Jusuf Kalla yang resmi menjabat sebagai Presiden dan Wakil Presiden periode 2014-2019 . Terpilihnya Joko Widodo – Jusuf Kalla ini tidak lepas dari pandangan masyarakat yang pro maupun kontra terhadap kinerja maupun individu khususnya Bapak Joko Widodo. Jenjang karir Jokowi sangat cepat, mulai tahun 2005 beliau menjabat sebagai walikota Solo, kemudian

tahun 2012 menjadi Gubernur Ibukota RI yaitu DKI Jakarta, hingga saat ini tahun 2014 beliau sudah menjabat sebagai Presiden Republik Indonesia (RI) .

Pengambilan data dalam penelitian ini dilakukan menggunakan search techniques yang berasal dari University of Reading, agar pengambilan data dapat dilakukan dengan efektif dan efisien yang selanjutnya dilakukan koreksi label positif, netral dan negatif oleh pihak ketiga yaitu ahli bahasa Indonesia yang sudah berkecimpung lama dalam ilmu bahasa Indonesia dan sudah sepuluh tahun menjadi pengajar bahasa Indonesia.

Analisis sentimen dan opinion mining adalah bidang studi yang menganalisis pendapat seseorang, sentiment seseorang, evaluasi seseorang, sikap seseorang dan emosi seseorang ke dalam bahasa tertulis [1]–[3]. Analisis sentimen telah banyak menerima perhatian semenjak penelitian Pang, Turney, Goldberg dan Zhu . Teknik analisis sentimen dapat mendukung banyak keputusan dalam banyak skenario . Penelitian ini menggunakan tiga class attribute, yaitu positif, netral dan negatif, karena di internet komentar yang muncul dapat berupa komentar positif, netral dan negatif [4], [5].

Normalisasi pada penelitian ini menggunakan kamus KBBA (Kamus Besar Bahasa Alay) yang didapat dari Nurfalah Adiyasa karena kamus yang menjadi penelitiannya di-share untuk kepentingan penelitian selanjutnya. Stemmer yang dipakai menggunakan Sastrawi Master karena merupakan Library PHP untuk stemming bahasa Indonesia, mudah diintegrasikan dengan framework atau package lainnya, mempunyai API yang sederhana dan mudah digunakan [6]–[8].

Setelah preprocessing, kita perlu mempertimbangkan apakah baik untuk melakukan studi analisis sentimen agar kita dapat melihat perbandingan keakuratan data yang dinormalisasi dan historis [6], [9]. Dengan melihat hasil akurasi yang diperoleh, Anda dapat memutuskan apakah preprocessing harus digunakan untuk mendapatkan hasil akurasi terbaik. Metode klasifikasi yang digunakan dalam penelitian ini adalah Naive Bayes dan SVM [10], [11].

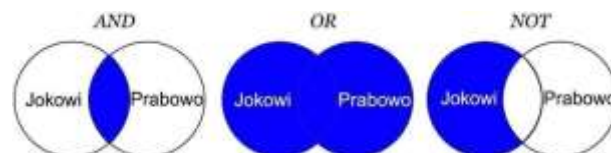
2. METODE PENELITIAN

Pengambilan data dilakukan dari sebelum pemilihan umum 9 Juli 2014 hingga setelah pemilihan Bapak Ir menggunakan teknik pencarian Boolean “AND” yang bersumber dari media sosial Twitter, Facebook dan blog politik. H. Joko Widodo sebagai Presiden 2014- 2019, 20 Oktober 2014. Survei tersebut memuat berita atau komentar netral terkait kepribadian dan kinerja Jokowi. Untuk kata kunci yang dimasukkan tidak hanya “Jokowi and jujur”, “Jokowi and munafik”, dan “Jokowi and membahas” saja, melainkan dilakukan juga untuk data positif, netral dan negatif dengan memasukkan kata kunci.

2.1 Search Techniques

Search techniques pada penelitian ini menggunakan salah satu metode pencarian yang dikemukakan oleh Jackie Skinner dari University of Reading yang merupakan universitas top 1% dunia. Metode dalam search techniques yang dipakai dalam penelitian ini adalah menggunakan operator pencarian [12]. Metode ini menggabungkan kata-kata pencarian mencakup sinonim dan dikenal juga sebagai Boolean searching. Pada metode ini memungkinkan untuk memasukkan banyak kata ataupun konsep dalam pencarian.

Pada Gambar 1 bagian berwarna biru mengindikasikan hasil yang didapat berdasarkan operator “AND”, “OR” dan “NOT”.



Gambar 1. Operator Boolean “AND”, “OR” dan “NOT”

2.2 Data Mining

Data mining adalah aplikasi algoritma untuk mengekstrak pola dari data sehingga pengetahuan yang tersembunyi dalam data dapat ditemukan [10]. Data yang tidak diolah tidak memiliki nilai, dan sebaliknya, data yang diolah dengan baik menjadi informasi berharga yang dapat digunakan untuk berbagai tujuan [13], khususnya dalam penelitian ini untuk kepentingan manajemen reputasi Bapak Jokowi.

2.3 Analisis Sentimen

Analisis sentimen dan opinion mining adalah bidang studi yang menganalisis pendapat seseorang, sentiment seseorang, evaluasi seseorang, sikap seseorang dan emosi seseorang ke dalam bahasa tertulis [14]. Analisis sentimen telah banyak menerima perhatian semenjak penelitian Pang,

Turney, Goldberg dan Zhu . Analisis sentimen menghadirkan banyak tantangan, seperti penilaian yang dibuat tentang subjek atau tujuan yang dibuat pada dokumen atau bagian dari dokumen, dan apakah opini yang diungkapkan positif atau negatif [15], [16]. Di lain waktu, Anda harus mencari kekuatan pendapat atau tujuan seseorang di luar kalimat yang diberikan [17], [18]. Sentimen dalam sebuah teks menjadi sangat penting di beberapa aplikasi, seperti pengolahan dan penyimpulan review seseorang, buku dan film, kemudian analisis opini dalam politik, serta mengklasifikasikan kiriman dan komentar blog [14].

2.4 Cleansing

Pada proses ini, data yang diambil belum tentu bersih, terutama data yang diambil dari Twitter [19]. Normalisasi kata komentar yang diberikan seseorang tidak semuanya menggunakan bahasa baku [20], banyak sekali yang menggunakan kata gaul, misalnya : “gue”, “loe”, dan lain lain, serta tidak jarang pula yang menggunakan potongan kata, misalnya : “yg”, “brp”, “bgm”, “smoga” dan lain lain. Kata yang tidak di normalisasi terlebih dahulu akan dikenali weka sebagai kata yang berbeda, misalnya ‘semoga’ dan ‘semoga’ yang seharusnya memiliki makna yang sama akan menjadi beda makna dikarenakan penulisannya yang berbeda. Untuk itu dilakukan normalisasi kata dari yang tidak baku menjadi kata baku [21]. Untuk normalisasi ini menggunakan bantuan kamus KBBA yang didapat dari Nurfalah Adiyasa.

2.5 Aplikasi WEKA

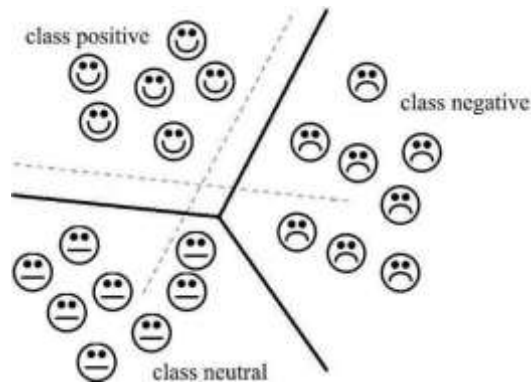
Aplikasi WEKA adalah sebuah aplikasi yang dipakai untuk pembelajaran mesin (machine learning) yang dimaksudkan untuk membantu dalam penerapan teknik pembelajaran mesin keberbagai masalah yang ada dalam dunia nyata [14];, khususnya dalam penelitian ini berhubungan dengan politik dan manajemen reputasi. Pembelajaran yang ada dalam WEKA meliputi visualisasi hasil, manipulasi data, hubungan antar basisdata, cross- validation, dan perbandingan set aturan untuk melengkapi penelitian yang berhubungan dengan pembelajaran mesin dasar [2]. Selain itu, WEKA menyediakan lingkungan untuk keperluan klasifikasi, regresi, clustering, dan fitur seleksi . Dengan WEKA, dapat diketahui akurasi dari mulai preProcessing hingga hasil akhir dengan metode tertentu dengan cepat dibandingkan perhitungan manual.

2.6 Naïve Bayes

Naive Bayes adalah algoritma analisis statistik yang menggunakan probabilitas Bayesian untuk melakukan pemrosesan data pada data numerik [22]. Klasifikasi Bayesian adalah klasifikasi statistik yang dapat memprediksi kelas anggota probabilitas. Dalam klasifikasi Bayesian sederhana, yang lebih dikenal sebagai pengklasifikasi Nave-Bayes, pengaruh nilai atribut kelas dapat diasumsikan tidak terpengaruh atau tidak terpengaruh oleh nilai atribut lainnya [23]. Asumsi ini disebut independensi kelas dan diciptakan untuk memfasilitasi komputasi. Pemahaman ini dianggap “naif”. Sederhananya, kita berasumsi bahwa kemunculan istilah kata dalam sebuah kalimat tidak dipengaruhi oleh kata lain, jadi dalam analisis sentimen ada bobot untuk setiap kemunculan kata, yang dihitung sebagai jumlah bobot untuk keseluruhan kalimat. Apakah kalimat itu positif, netral, atau negatif.

2.7 SVM

Support vector machine (SVM) adalah jenis model vektor berdasarkan pengklasifikasi yang memerlukan konversi teks ke vektor sebelum dapat digunakan untuk klasifikasi [24]. Key idea dari SVM adalah untuk menemukan permukaan keputusan (Hyperlane) yang maksimal dari setiap titik data [25]. Untuk melakukan training mesin yang didukung oleh vektor atau biasa disebut Support Vector Machine (SVM) memerlukan solusi Quadratic Programming (QP) yang sangat besar [12]. Quadratic Programming adalah masalah matematika untuk menemukan vektor “x” yang meminimalkan fungsi kuadrat , dengan melakukan pembagian kelas menggunakan hyperlink maka masing-masing kelas positif, netral dan negatif dapat dibagi berdasarkan area masing-masing sehingga ketika terdapat data baru dapat ditentukan kelasnya berdasarkan area positif, netral maupun negatif. SVM untuk multi class dapat diperlihatkan pada Gambar 2.



Gambar 2. SVM untuk kelas positif, netral dan negatif

3. HASIL DAN ANALISIS

Data yang pencariannya tanpa menggunakan Boolean searching mendapatkan hasil pencarian di Google sebanyak 25.700.000 yang dapat dilihat pada Gambar 3. Dengan menggunakan Boolean searching untuk pencarian data positif dengan kata kunci “Jokowi and jujur” pada Gambar 4 mendapatkan hasil sebanyak 1.030.000 atau terjadi pengurangan data sebesar 24.670.000 (95,99%). Untuk pencarian data negatif dengan kata kunci “Jokowi and munafik” pada Gambar 5 mendapatkan hasil sebanyak 463.000 atau terjadi pengurangan sebesar 25.237.000 (98,198%) dan untuk pencarian data netral dengan kata kunci “Jokowi and membahas” pada Gambar 6 mendapatkan hasil sebanyak 922.000 atau terjadi pengurangan sebesar 24.778.000 (96,41%).



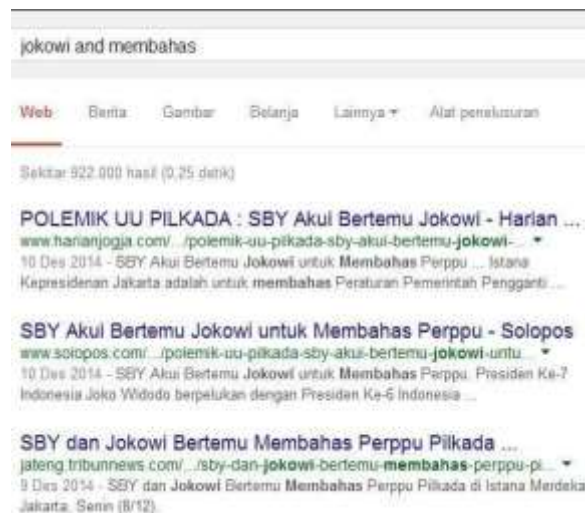
Gambar 3. Hasil pencarian Jokowi di Google tanpa Boolean searching



Gambar 4. Hasil pencarian Jokowi di Google dengan Boolean searching “Jokowi and jujur”.



Gambar 5. Hasil pencarian Jokowi di Google dengan Boolean searching “Jokowi and munafik”



Gambar 6. Hasil pencarian Jokowi di Google dengan Boolean searching “Jokowi and membahas”

Reduksi data yang cukup besar memungkinkan pengambilan data Jokowi berjalan efektif dan efisien. Jumlah total data yang digunakan dalam penelitian ini adalah 59 positif, 60 netral dan 58 negatif. Pada penelitian ini dilakukan preprocessing yang berbeda dilakukan pada semua data dalam penelitian ini. Teknik preprocessing yang digunakan dalam penelitian ini dapat dilihat pada kolom Preprocessing Techniques, teknik preProcessing sebagai berikut:

Teknik preProcessing A

TF-IDF = Yes ; Lowercase = Yes ; minTermFreq = 1 ; Normalize all data ; Stopwords = Yes ; Tokenizer = N-Gram ; dan Emoticon = Yes.

Teknik preProcessing B

TF-IDF = Yes ; Lowercase = Yes ; minTermFreq = 1 ; Normalize all data ; Stopwords = Yes ; Tokenizer = N-Gram ; dan Emoticon = No.

Teknik preProcessing C

TF-IDF = Yes ; Lowercase = Yes ; minTermFreq = 1 ; Normalize all data ; Stopwords = No ; Tokenizer = N-Gram ; dan Emoticon = Yes.

Teknik preProcessing D

TF-IDF = Yes ; Lowercase = Yes ; minTermFreq = 1 ; Normalize all data ; Stopwords = No ; Tokenizer = Unigram ; dan Emoticon = Yes.

Teknik preProcessing E

TF-IDF = Yes ; Lowercase = Yes ; minTermFreq = 1 ; Normalize all data ; Stopwords = No ; Tokenizer =

Bigram ; dan Emoticon = Yes.

Teknik preProcessing F

TF-IDF = Yes ; Lowercase = Yes ; minTermFreq = 1 ; Normalize all data ; Stopwords = No ; Tokenizer = Trigram ; dan Emoticon = Yes.

Data Sudah Dinormalisasi (Percobaan 1)

Percobaan 1 adalah data yang sudah dilakukan normalisasi tetapi belum dilakukan stemming dapat dilihat pada Tabel 1.

Tabel 1. Data Sudah Dinormalisasi

TPre	Met	TF-IDF	LC	MTF	N	SW	T	Emo	TTB(s)	Hasil (%)
A	NB	Yes	Yes	1	1	Yes	N-Gram	Yes	1,08	83,0508
A	SVM	Yes	Yes	1	1	Yes	N-Gram	Yes	0,66	83,0508
B	NB	Yes	Yes	1	1	Yes	N-Gram	No	1,13	80,226
B	SVM	Yes	Yes	1	1	Yes	N-Gram	No	0,55	79,096
C	NB	Yes	Yes	1	1	No	N-Gram	Yes	1,14	86,4407
C	SVM	Yes	Yes	1	1	No	N-Gram	Yes	0,5	88,7006
D	NB	Yes	Yes	1	1	No	Unigram	Yes	0,23	88,1356
D	SVM	Yes	Yes	1	1	No	Unigram	Yes	0,22	88,1356
E	NB	Yes	Yes	1	1	No	Bigram	Yes	0,5	67,7966
E	SVM	Yes	Yes	1	1	No	Bigram	Yes	0,74	66,6667
F	NB	Yes	Yes	1	1	No	Trigram	Yes	0,59	42,3729
F	SVM	Yes	Yes	1	1	No	Trigram	Yes	0,44	40,678

Data Sudah Dinormalisasi dan Stemming (Percobaan 2)

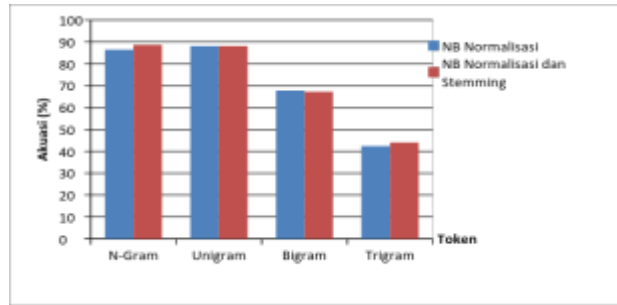
Percobaan 2 ini adalah data yang sudah dinormalisasi dan stemming, hasil akurasi pada jenis data ini dapat dilihat pada Tabel 2.

Tabel 2. Data Sudah Normal dan Stemming

TPre	Met	TF-IDF	LC	MTF	N	SW	T	Emo	TTB(s)	Hasil (%)
A	NB	Yes	Yes	1	1	Yes	N-Gram	Yes	1,15	83,0508
A	SVM	Yes	Yes	1	1	Yes	N-Gram	Yes	0,61	81,9209
B	NB	Yes	Yes	1	1	Yes	N-Gram	No	0,92	81,9209
B	SVM	Yes	Yes	1	1	Yes	N-Gram	No	0,56	81,3559
C	NB	Yes	Yes	1	1	No	N-Gram	Yes	1,03	88,7006
C	SVM	Yes	Yes	1	1	No	N-Gram	Yes	0,59	88,7006
D	NB	Yes	Yes	1	1	No	Unigram	Yes	0,34	88,1356
D	SVM	Yes	Yes	1	1	No	Unigram	Yes	0,36	89,2655
E	NB	Yes	Yes	1	1	No	Bigram	Yes	0,66	67,2316
E	SVM	Yes	Yes	1	1	No	Bigram	Yes	0,27	66,1017
F	NB	Yes	Yes	1	1	No	Trigram	Yes	0,49	44,0678
F	SVM	Yes	Yes	1	1	No	Trigram	Yes	0,38	44,0678

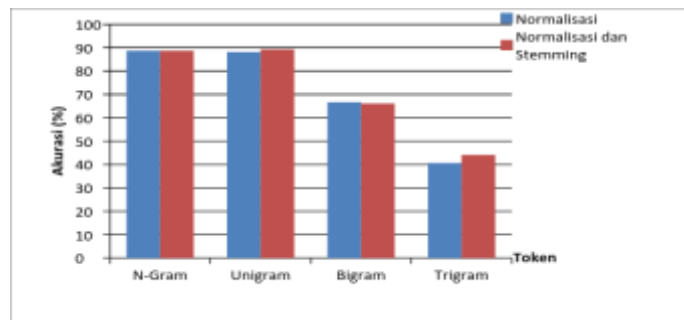
Pada Tabel 2 terdapat peningkatan akurasi dibandingkan Tabel 1 pada beberapa teknik preProcessing. Selain itu, terjadi perubahan akurasi tertinggi dibanding Tabel 1, pada Tabel 2 akurasi tertinggi dengan menggunakan token Unigram dan metode SVM dengan hasil akurasi sebesar 89,2655%.

Grafik perbandingan token pada metode Naive Bayes berdasarkan percobaan 1 dan 2 dapat dilihat pada Gambar 7.



Gambar 7. Grafik perbandingan metode Naive Bayes percobaan 1 dan 2

Dan grafik perbandingan token pada metode SVM berdasarkan percobaan 1 dan 2 dapat dilihat pada Gambar 8.



Gambar 8. Grafik perbandingan token pada metode SVM percobaan 1 dan 2

Pada Grafik yang terlihat pada Gambar 7 dan Gambar 8 memperlihatkan bahwa metode N- Gram dan Unigram memperlihatkan akurasi yang lebih baik dibanding Bigram dan Trigram, kemudian Bigram lebih baik dibandingkan Trigram. Unigram dan N-Gram lebih tinggi dikarenakan masing-masing membagi kalimat menjadi satu kata, sehingga kemungkinan mendapatkan kata yang sama sangat besar sedangkan pada Bigram membagi kalimat per dua kata, sehingga dapat diartikan dua kata tersebut menjadi satu arti dan lebih kecil kemungkinan bertemu dua kata yang sama tersebut di kalimat yang lain, terlebih lagi trigram yang membagi kalimat per tiga kata, lebih kecil lagi kemungkinan bertemu kembali kepada tiga kata yang sama yang sudah dibagi di tiap kalimat.

4. KESIMPULAN

Berdasarkan penelitian yang dilakukan, dapat ditarik kesimpulan sebagai berikut, Boolean searching dengan operator “AND” sangat efektif dan efisien dalam pencarian data, selain itu cukup berpengaruh terhadap akurasi analisis sentimen, karena menghasilkan tidak semua data mengenai Jokowi, melainkan yang mengandung kata yang dimasukkan. Berdasarkan penelitian yang dilakukan, akurasi yang dihasilkan ketika data dilakukan stemming, terdapat peningkatan rata-rata sebesar 0,85% untuk metode Naive Bayes dan 0,85% untuk metode SVM. Akurasi yang dihasilkan metode SVM tidak selalu unggul dibandingkan metode Naive Bayes, begitu pula sebaliknya. Untuk metode yang paling tinggi di masing- masing percobaan adalah metode SVM pada percobaan 1 mendapatkan akurasi 88,7006% untuk teknik preProcessing C, percobaan 2 mendapat akurasi 89,2655% untuk teknik preProcessing D. Dengan melakukan normalisasi dan stemming, hasil yang didapat lebih tinggi dibandingkan dengan data yang hanya dilakukan normalisasi saja, dikarenakan dengan adanya normalisasi, kata-kata yang tidak sesuai dengan KBBA dapat disesuaikan dengan standar yang ada, ditambah lagi dengan adanya stemming, katakata yang dianggap berbeda karena terdapat imbuhan, dapat diubah terlebih dahulu berdasarkan kata dasarnya, sehingga variasi kata semakin sedikit, semakin sering muncul, dan dapat memberikan bobot positif, negatif ataupun netral terhadap kata tersebut semakin baik.

REFERENSI

- [1] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, “ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI,” *Jurnal Teknoinfo*, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.
- [2] M. R. A. Nasution and M. Hayaty, “Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter,” *Jurnal Informatika*, vol. 6, no. 2, pp. 226–235, Sep. 2019, doi: 10.31311/ji.v6i2.5129.

- [3] S. Styawati, N. Hendrastuty, and A. R. Isnain, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. 3, pp. 150–155, Oct. 2021, doi: 10.30591/jpit.v6i3.2870.
- [4] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *SMATIKA JURNAL*, vol. 10, no. 02, pp. 71–76, Dec. 2020, doi: 10.32664/smatika.v10i02.455.
- [5] D. Gunawan, D. Riana, D. Ardiansyah, and F. Akbar, "Komparasi Algoritma Support Vector Machine Dan Naïve Bayes Dengan Algoritma Genetika Pada Analisis Sentimen Calon Gubernur Jabar 2018-2023," *Jurnal Teknik Komputer*, vol. 6, no. 1, pp. 121–129, Jan. 2020, doi: 10.31294/jtk.v6i1.6866.
- [6] A. A. Aldino, D. Darwis, A. T. Prastowo, and C. Sujana, "Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency," *J Phys Conf Ser*, vol. 1751, no. 1, p. 12038, Jan. 2021, doi: 10.1088/1742-6596/1751/1/012038.
- [7] M. P. K. Putra and W. -, "A Novel Method for Handling Partial Occlusion on Person Re-identification using Partial Siamese Network," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, 2021, doi: 10.14569/ijacsa.2021.0120735.
- [8] R. N. N. Az-zahra, M. P. Tyo, F. Rahman, M. Z. Sidiq, and R. Sanjaya, "Analisis Kualitas Layanan Terhadap Kepuasan Pengguna Sistem Informasi Kawasan Agrowisata Cicantayan Menggunakan Webqual 4.0," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 8, no. 1, p. 13, Apr. 2022, doi: 10.26418/jp.v8i1.51056.
- [9] E. Indrayuni, "Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes," *Jurnal Khatulistiwa Informatika*, vol. 7, no. 1, pp. 29–36, Jun. 2019, doi: 10.31294/jki.v7i1.1.
- [10] A. R. Isnain, N. S. Marga, and D. Alita, "Sentiment Analysis Of Government Policy On Corona Case Using Naive Bayes Algorithm," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 1, p. 55, Jan. 2021, doi: 10.22146/ijccs.60718.
- [11] H. Tuhuteru and A. Iriani, "Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 3, no. 3, pp. 394–401, Oct. 2018, doi: 10.30591/jpit.v3i3.977.
- [12] A. Nurkholis, D. Alita, and A. Munandar, "Comparison of Kernel Support Vector Machine Multi-Class in PPKM Sentiment Analysis on Twitter," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 2, pp. 227–233, Apr. 2022, doi: 10.29207/resti.v6i2.3906.
- [13] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Riset Komputer)*, vol. 8, no. 6, p. 219, Dec. 2021, doi: 10.30865/jurikom.v8i6.3655.
- [14] D. Alita, "Multiclass SVM Algorithm for Sarcasm Text in Twitter," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 1, pp. 118–128, Mar. 2021, doi: 10.35957/jatisi.v8i1.646.
- [15] N. Hendrastuty, "Rancang Bangun Aplikasi Monitoring Santri Berbasis Android (Studi Kasus: Pesantren Nurul Ikhwan Maros)," *Jurnal Data Mining Dan Sistem Informasi*, vol. 2, no. 2, pp. 21–34, 2021.
- [16] I. Kurniawan and A. Susanto, "Implementasi Metode K-Means dan Naïve Bayes Classifier untuk Analisis Sentimen Pemilihan Presiden (Pilpres) 2019," *Eksplora Informatika*, vol. 9, no. 1, pp. 1–10, Sep. 2019, doi: 10.30864/eksplora.v9i1.237.
- [17] S. Styawati, A. R. Isnain, N. Hendrastuty, and L. Andraini, "Comparison of Support Vector Machine and Naïve Bayes on Twitter Data Sentiment Analysis," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. 1, pp. 56–60, 2021.
- [18] D. Darwis, N. Siskawati, and Z. Abidin, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN REVIEW DATA TWITTER BMKG NASIONAL," *Jurnal Tekno Kompak*, vol. 15, no. 1, p. 131, Feb. 2021, doi: 10.33365/jtk.v15i1.744.
- [19] H. Syah and A. Witanti, "ANALISIS SENTIMEN MASYARAKAT TERHADAP VAKSINASI COVID-19 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)," *Jurnal Sistem Informasi dan Informatika (Simika)*, vol. 5, no. 1, pp. 59–67, Feb. 2022, doi: 10.47080/simika.v5i1.1411.
- [20] S. Rahayu and J. J. Purnama, "KLASIFIKASI KONSUMSI ENERGI INDUSTRI BAJA MENGGUNAKAN TEKNIK DATA MINING," *Jurnal Teknoinfo*, vol. 16, no. 2, p. 395, Jul. 2022, doi: 10.33365/jti.v16i2.1984.
- [21] S. D. Asri, D. Ramayanti, A. D. Putra, and Y. T. Utami, "DETEKSI RODA KENDARAAN DENGAN CIRCLE HOUGH TRANSFORM (CHT) DAN SUPPORT VECTOR MACHINE (SVM)," *Jurnal Teknoinfo*, vol. 16, no. 2, p. 427, Jul. 2022, doi: 10.33365/jti.v16i2.1952.
- [22] P. S. Dewi, C. K. Sastradipraja, and D. Gustian, "Sistem Pendukung Keputusan Kenaikan Jabatan Menggunakan Metode Algoritma Naïve Bayes Classifier," *Jurnal Teknologi dan Informasi*, vol. 11, no. 1, pp. 66–80, Mar. 2021, doi: 10.34010/jati.v11i1.3593.
- [23] S. eka Yuliana Putri, "PENERAPAN MODEL NAIVE BAYES UNTUK MEMPREDIKSI POTENSI PENDAFTARAN SISWA DI SMK TAMAN SISWA TELUK BETUNG BERBASIS WEB," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 1, no. 1, pp. 93–99, Jun. 2020, doi: 10.33365/jatika.v1i1.228.

- [24] D. Alita, Y. Fernando, and H. Sulistiani, "IMPLEMENTASI ALGORITMA MULTICLASS SVM PADA OPINI PUBLIK BERBAHASA INDONESIA DI TWITTER," *Jurnal Tekno Kompak*, vol. 14, no. 2, p. 86, Aug. 2020, doi: 10.33365/jtk.v14i2.792.
- [25] S. Styawati and K. Mustofa, "A Support Vector Machine-Firefly Algorithm for Movie Opinion Data Classification," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 13, no. 3, p. 219, Jul. 2019, doi: 10.22146/ijccs.41302.