# SENTIMENT ANALISYS OF KOMISI PEMILIHAN UMUM INDONESIA (KPU) INSTITUTION ON 2019 ELECTION IN TWITTER SOCIAL MEDIA USING SUPPORT VECTOR MACHINE (SVM) CLASSIFICATION METHOD AND LEXICON BASED FEATURES

Dr. Miftah Andriansyah Ssi., MMSi, Roni Budianto

Faculty of Computer Science and Information Technology, Gunadarma University

Jakarta, Indonesia

Andriansyah.miftah@gmail.com

# ABSTRACT

Indonesia has completed elections in 2019. Elections are a process in all citizens in a democratic country that uses its voting rights to elect a president and some leadership in a new government held simultaneously. The institution responsible for holding elections in Indonesia is “Komisi Pemilihan Umum” also known as KPU. No matter how good the KPU's performance is, do not get away from mistakes and public criticism, especially in this digital era where everyone is free to talk on social media. Based on the above considerations, the researcher wants to find out about people's responses about this using sentiment analysis. Sentiment analysis will be carried out on Twitter social media using 1000 data taken in the post-election period (28 April - 22 May) 2019. Sentiment analysis is done using vector machine support and lexicon-based features classification methods using 80% composition of training data and 20% testing data. The sentiments produced were 589 (58.9%) negative sentiments and 411 positive (41.1%) with an accuracy level of 80.55% and 74% of precision.

Keywords : *support vector machine, svm, twitter*, kpu, sentiment analysis

1. **INTRODUCTION**

Indonesia has completed elections in 2019. Elections are a process in all citizens in a democratic country that uses its voting rights to elect a president and some leadership in a new government held simultaneously. The institution responsible for holding elections in Indonesia is “Komisi Pemilihan Umum” also known as KPU. No matter how good the KPU's performance is, do not get away from mistakes and public criticism, especially in this digital era where everyone is free to talk on social media. Twitter is one of the famous social media. Twitter users in Indonesia are quite numerous. According to the site we are social, Twitter users as of January 2019 totaled 6.43 million people. Twitter is used to express thoughts which are then known as tweets. Based on the above considerations, the researcher wants to find out about people's responses about this using sentiment analysis on twitter social media. Sentiment analysis is often used to monitor social media because it allows us to get a broad picture of public opinion on a particular topic. The purpose of sentiment analysis is to determine the behavior or opinion of the author by paying attention to the topic being read. And therefore, sentiment analysis will be conducted on the KPU. Using #kpu or #kpuid Hastag. This research takes and uses data through Twitter API with a total of 1000 data and using bahasa indonesia. Data was taken on vulnerable period from April 28 2019 to May 22 2019 (post-election) and processed using the support vector machine classification method and lexicon based features

1. **REVIEW OF RELATED LITERATURE**
2. **Machine Learning**

Machine learning is a technique for inferring data with a mathematical approach that reflects data patterns [12]. The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

**Analisis Sentimen**

Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini [11]. Analisis sentimen adalah analisis yang dilakukan pada pendapat seseorang tentang suatu organisasi, topik atau perusahaan tertentu. Analisis sentimen atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentiment yang terkandung dalam suatu kalimat opini.

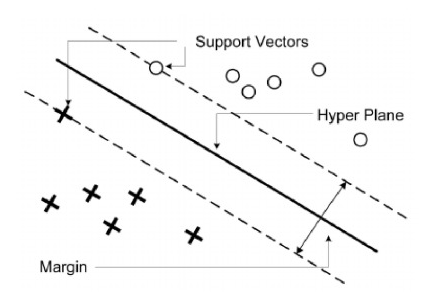
Tujuan pertama ketika seseorang berurusan dengan analisis sentimen biasanya terdiri dalam membedakan antara kalimat subjektif dan objektif. Jika kalimat yang diberikan diklasifikasikan sebagai obyektif, tidak ada tugas mendasar lainnya yang diperlukan, sedangkan jika kalimat tersebut diklasifikasikan sebagai subyektif, polaritasnya perlu diperkirakan [2]. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan negatif atau positif.

**Twitter**

Twitter adalah jejaring sosial yang dimiliki dan dikelola oleh Twitter Inc. Twitter menawarkan jaringan sosial mikroblog sehingga memungkinkan penggunanya untuk membaca dan mengirimkan pesan *tweets*. Mikroblog adalah suatu bentuk blog yang memungkinkan penggunanya untuk menulis teks pembaharuan singkat dan mempublikasikannya.

***Support Vector Machine***

*Support vector machine (SVM)* are supervised learning models that are used to build classifiers and regressors. An SVM finds the best separating boundary between the two sets of points by solving a system of mathematical equations. adalah suatu teknik yang relatif baru (1995) untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi, yang sangat populer belakangan ini. *SVM* berada dalam satu kelas dengan *ANN* dalam hal fungsi dan kondisi permasalahan yang bisa diselesaikan. Keduanya masuk dalam kelas *supervised learning*, dimana dalam implementasinya perlu adanya tahap training dan disusul tahap testing [21]. *Support Vector Machine (SVM)* Adalah salah satu algoritma pada *machine learning* yang paling efektif dari sisi praktis maupun teoritis. *SVM* berusaha menemukan *hyperplane* dengan memaksimalkan jarak antar kelas. Dengan cara ini, *SVM* dapat menjamin kemampuan generalisasi yang tinggi untuk data-data yang akan datang [7]. *SVM* berusaha menemukan *hyperplane* terbaik dengan memaksimalkan *margin* antar kelas.

****

Gambar 2.1 Ilustrasi Support Vector Machine

Proses pembelajaran *SVM* adalah untuk menentukan *support vector*, hanya cukup mengetahui fungsi kernel yang dipakai dan tidak perlu mengetahui wujud dari fungsi non-linier. *SVM* dapat memisahkan data secara *linear* dan non linear.

1. *Hard Margin SVM*

*Hard Margin* atau *Linearly sparable data* adalah data yang digunakan untuk perhitungan svm yang dapat dipisahkan secara *linear*. Pada kasus ini data yang digunakan adalah data dengan dua kelas yang sudah terpisah secara linear. Data latih dinyatakan oleh *(xi , xj )* dan *xi*adalah atribut set untuk data latih kelas ke – i yakni *xi* = { *x1, x2, x3* ... *xi* }. Untuk y {1,-1} menyatakan label kelas. Pendefinisian persamaan suatu *hyperplane* pemisah dituliskan dengan :

*w \* xi +b* = 0 (2.1)

Dimana,

w = bobot

b= bias

x= *variable* input

Dengan ketentuan jika *w \* xi* + b > 0 untuk yi = +1 adalah *hyperplane*-pendukung (*supporting* *hyperplane*) dari kelas +1, dan *w \* xi* + b < 0 untuk yi = -1 adalah *hyperplane* pendukung dari kelas -1, margin antara dua kelas dapat dihitung dengan mencari jarak pembatas antara kedua kelas dan titik terdekatnya dengan mencari titik minimal dengan persamaan (2.2):

*min* *t*(*w*) = ||*w*||2  (2.2)

sedangkan subjek *constrain* / kendala persamaannya adalah sebagai berikut:

*y(xi* . *w + b)*-1 ≥ 0, i (2.3)

Permasalahan ini dapat dipecahkan dengan berbagai teknik komputasi. Lebih mudah dengan mengubah persamaan diatas ke dalam fungsi lagrangian dan menyederhanakannya menjadi persamaan berikut :

L (*w,b,a*) = ||*w*||2 - i *yi* (*wTxi+b*)+ i  (2.4)

Dimana *a*i adalah lagrange multiplier yang bernilai nol atau positif (*a*i ≥ 0). Nilai optimal dari (2.4) dapat dihitung dengan meminimalkan L terhadap *a*i sekaligus terhadap *w* dan *b* dengan persamaan berikut:

i *-*  , *a*i *a*jyi xi (2.5)

Dengan memperhatikan persamaan berikut : i yi = 0 , *a*i ≥ 0 (*i,j =* 1,..,n) (2.6)

Diperoleh nilai *a*i yang nantinya akan digunakan untuk menemukan nilai bobot (*w*). Terdapat nilai *a*i untuk setiap data latih. Nilai *a*i yang > 0 adalah merupakan support vector dan dapat digunakan untuk mencari hyperplane(*b*).

**Bahasa Pemrograman Python**

*Python* adalah *general*-*purpose*, *high*-*level* *programming* *language*. Filosofi desain *python* menitikberatkan pada *code* *readability*, dan syntax yang memungkinkan programmer untuk mengekspresikan konsepnya dengan *lines* *of* *code* yang lebih sedikit dari bahasa lainnya [4].

**Flowchart**

Flowchart adalah suatu bagan dengan simbol-simbiol tertentu yang menggambarkan urutan proses secara mendetail dan hubungan suatu proses (Instruksi) dengan proses lainya dalam suatu *program* [17]. Tujuan membuat flowchart

1. Menggambarkan suatu tahapan penyelesaian masalah
2. Secara sederhana, terurai, dan jelas
3. Menggunakan simbol-simbol standar

Simbol-simbol dari flowchart memiliki fungsi yang berbeda antara satu simbol dengan simbol lainya.

***Confusion Matrix***

*Confusion Matrix* adalah tabel yang digunakan untuk memahami kinerja model klasifikasi [3]. *Confusion* *matrix* digambarkan dengan tabel yang menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan. *Confusion* *Matrix* seringkali digunakan untuk mengukur Akurasi, Presisi, dan *Recall* pada sebuah model klasifikasi. Akurasi menggambarkan seberapa akurat sistem melakukan klasifikasi, presisi adalah kecocokan antara bagian data yang diambil dengan informasi yang dibutuhkan, recall menggambarkan berapa % data berkategori positif yang terklasifikasikan dengan benar oleh sistem sedangkan, F1 score adalah keseimbangan antara presisi dan recall.

Akurasi = \* 100%

Presisi = \* 100%

Recall = \* 100%

F1 Score = 2 \*

Dimana,

TP = *True* Positif

TN = *True* Negatif

FP = *False* Positif

FN = *False* Negatif

**Komisi Pemilihan Umum Indonesia**

Komisi pemilihan umum atau KPU adalah lembaga yang menyelenggarakan pemilu di Indonesia. Menurut UU No.22 Tahun 2007 Pasal 1 ayat 6, Komisi Pemilihan Umum, selanjutnya disebut KPU, adalah lembaga Penyelenggara Pemilu yang bersifat nasional, tetap, dan mandiri.

**Pembobotan TF-IDF**

Dalam pencarian informasi, TF-IDF (*Term Frequency - Inverse Document Frequency*) adalah statistik numerik yang dimaksudkan untuk mencerminkan betapa pentingnya sebuah kata dalam koleksi atau corpus [4]. TF-IDF adalah metode yang digunakan untuk menghitung bobotan setiap kata yang telah diekstrak.

Model pembobotan TF-IDF merupakan metode yang mengintegrasikan model *term* *frequency* (tf) dan *inverse* *document* *frequency* (idf).*Term* *frequency* (tf) merupakan proses untuk menghitung jumlah kemunculan *term* dalam satu dokumen dan inverse *document* *frequency* (idf) digunakan untuk menghitung *term* yang muncul di berbagai dokumen (komentar) yang dianggap sebagai term umum, yang dinilai tidak penting [17].

Interface menu TF-IDF berisi informasi tentang kata-kata yang telah dipecah dari tweet dan telah memiliki nilai TF-IDF sendiri. Rumus yang digunakan adalah :

TFIDF = TF x IDF = TF x loge

Dimana,

TF = *Term* *Frequency*, banyaknya sebuah kata/istilah muncul

DF = *Document* *Frecuency*, jumlah dokumen di mana kata/istilah tersebut muncul. Minimal 1 dokumen.

| D | = Total seluruh dokumen

***Lexicon Based Features***

*Lexcion Based Features* merupakan pendekatan yang menggunakan suatu kamus sentimen berisi kata positif dan kata negatif yang dibandingkan dan dicocokkan dengan kata pada kalimat untuk diketahui tingkat polaritasnya.

*Lexcion Based Features* merupakan suatu kesepakatan dalam pendekatan yang meliputi frase, bentuk ekspresi, atau konten yang berupa teks yang umumnya terdapat pada obrolan, dialog, *post*, *review*,dan lainnya [17]. Dalam proses klasifikasi menggunakan *lexicon* *based*, dapat digunakan rumus sebagai berikut :

Spositif =

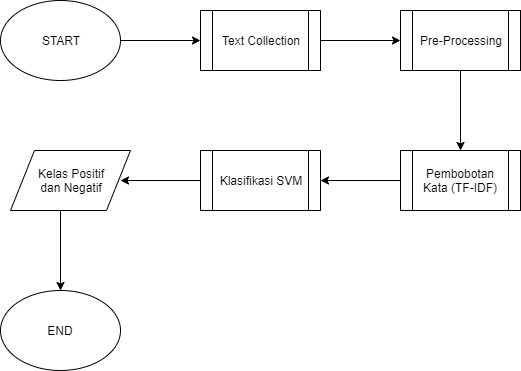
Snegatif =

Dimana, (*Spositif*) adalah sentimen positif hasil dari jumlah bobot positif yang memiliki nilai positif lebih besar, dan *(Snegatif)* adalah sentimen negatif yang didapatkan dari jumlah bobot negatif yang lebih besar. Jika total jumlah nilai positif lebih besar dari jumlah nilai negatif maka kalimat akan berorientasi positif. Namun jika total jumlah nilai positif kurang dari jumlah nilai negatif maka kalimat akan berorientasi negatif [14]. Dari persamaan nilai sentimen dalam suatu kalimat maka, diperoleh untuk menentukan orientasi sentimen.

1. **METHOD**

**Alur Proses Analisis Sentimen**

Proses analisis sentimen pada penelitian ini memiliki berberapa tahapan yang harus dilakukan untuk mendapatkan hasil yang diinginkan. Berikut adalah tahapan atau metode yang digunakan dalam melakukan analisis sentimen pada penelitian ini.

****

**Gambar 3.1** Alur Proses Analisis Sentimen

Gambar 3.1 menunjukan langkah – langkah yang harus di lakukan, berikut penjelasannya:

1. *Text* *Collection*, disebut juga sebagai ekstraksi data, adalah proses pengambilan data dengan menggunakan twitter API.
2. Pre-Proces, adalah proses setelah data diekstraksi, yaitu *cleansing* untuk penyaringan dan pembersihan data tweet dari kata dan karakter yang tidak perlu, *case folding* untuk mengubah seluruh tweet menjadi huruf kecil dan *stemming* untuk mengubah kalimat menjadi kata dasar.
3. Pembobotan kata, pembobotan kata menggunakan TF-IDF (*Term Frequency – Inverse Term Frequency*) yang melakukan perhitungan bobot setiap kata yang telah di ekstrak.
4. *Support* *Vector* *Machine*, pada proses ini dilakukan training data SVM dan *Testing* *data*.
5. Menampilkan *Output*, *output* hasil akhir akan ditampilkan.

**Ekstraksi Tweet**

Ekstraksi tweet adalah pengambilan data dari *Twitter* yang akan dijadikan sebagai objek penelitian dengan kata pencarian “#kpuid” dan “#kpu” yang dilakukan dengan menggunakan *Twitter API* dan bahasa pemrograman Python. Ekstraksi data pada halaman web menggunakan algoritma *partial tree alignment* yang memetakan sebuah website seperti sebuah pohon dan mengambil datanya pada setiap elemen html.

Hasil dari pencarian dan pengambilan data dapat dilihat pada Tabel 3.1.

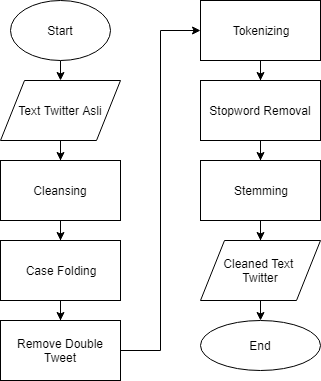
**Tabel 3.1** Hasil Ekstraksi Data Twitter

|  |  |
| --- | --- |
| **No** | **Text Tweet** |
| 1. | Bersihkan KPU dari praktek demokrasi kotor, https://t.co/07RXap5JOH |
| 2. | RT @KPU\_ID: Anak Kelas ll MTS Bobol Situs KPU Pemilu 2019 |
| 3. | @reno\_2 @Dody\_Lucas\_\_ @AkunTofa @KPU\_ID @bawaslu\_RI Mantap... |
| 4. | @Nadiku18 Yaa itu mah 02 nya gak legowo sama hasil situng |

**Pre-*Processing* *Data***

Pada pre-*processing* *data* adalah tahapan yang harus dilakukan sebelum melakukan klasifikasi. Pre-*processing* dilakukan untuk melakukan perubahan bentuk data menjadi lebih terstruktur dan sesuai dengan kebutuhan untuk pemrosesan lebih lanjut. Berberapa tujuan dari pre-*processing* *data* adalah :

1. Untuk mempermudah sehingga mempermudah pemilihan teknik dan metode *data* *mining* yang tepat
2. Untuk meningkatkan kualitas *data* sehingga hasil dari data mining menjadi lebih baik
3. Untuk meningkatkan efisiensi dan kemudahan proses penambangan *data*

****

**Gambar 3.2** Alur Pre-Processing

***Cleansing* *Data***

*Cleansing* *data* adalah salah satu langkah dalam proses pre-*processing* untuk membersihkan *data* yang ‘kotor’ supaya dapat digunakan untuk proses lebih lanjut. Langkah – langkah untuk melakukan *cleansing* adalah:

**clean.png**

**Gambar 3.3** Alur Cleansing

**Klasifikasi dengan *Lexicon* *Based* *Features***

Dalam proses labeling ini, dilakukan berdasarkan kata positif dan negatif yang terdapat pada tweet yang telah di bersihkan melalui tahap pre-*processing*. Klasifikasi ini dilakukan dengan kata-kata yang terdapat dalam database kata positif dan database kata negatif. Klasifikasi dilakukan dengan cara mencari jumlah kata positif dan negatif di dalam tweet, apabila dalam satu kalimat jumlah kata positif lebih besar dari kata bersentimen negatif, akan digolongkan sebagai sentimen positif dan diberi skor 1. Namun, jika sebaliknya kalimat akan digolongkan sebagai sentimen negatif dan diberi skor -1.

**Pembobotan TF-IDF**

*Term Frequency – Inverse Document Frequency* (TF-IDF) adalah metode yang digunakan untuk menghitung bobotan setiap kata yang telah diekstrak. TF-IDF berfungsi untuk menghitung seluruh jumlah kata dan total kemunculannya pada sebuah dokumen serta menghasilkan sebuah kamus data yang kemudian akan digunakan untuk proses training pada SVM.

**Tahap Pelatihan dengan *Support Vector Machine***

Metode yang dilakukan pada penelitian ini adalah dengan menggunakan *Support Vector Machine* (SVM). Alasan digunakannya metode klasifikasi *Support Vector Machine* pada penelitian ini karena metode SVM berusaha menemukan satu hyperplane dengan memaksimalkan jarak antar kelas (margin). Dengan memaksimalkan margin, SVM akan memiliki kemampuaan generalisasi yang tinggi terhadap data – data yang akan datang. Secara umum, SVM digunakan untuk klasifikasi dua kelas walaupun dapat digunakan untuk menyelesaikan masalah multi kelas.

**Pengujian Data**

Pengujian data dilakukan setelah nilai *w* dan *b* ditemukan, pengujian dilakukan guna mengetahui apakah data yang diuji masuk ke kelas positif atau masuk ke kelas negatif berdasarkan model yang telah dibuat

1. **RESULT**

**Pengujian**

Pada tahap ini dilakukan pengujian dengan menggunakan data latih dan data uji dengan perbandingan 80:20 [12]. Hasil dari klasifikasi menunjukan nilai akurasi sebesar 80.5%, presisi sebesar 74%, recall sebesar 77%, dan f1-score sebesar 75.4%. Perhitungan untuk menghitung akurasi, presisi dan recall, dilakukan dengan menggunakan *confusion* *matrix*. *confusion* *matrix* pada penelitian ini dapat dilihat pada tabel 4.1.

**Tabel 4.1** Implementasi *Confusion* *Matrix*

|  |  |  |
| --- | --- | --- |
| Prediction | True Values | |
| True | False |
| True | TP  62 | FP  18 |
| False | FN  21 | TN  99 |

Dengan nilai *confusion* *matrix* diatas maka, dapat dilakukan perhitungan Akurasi, Presisi dan Recall dengan perhitungan sebagai berikut:

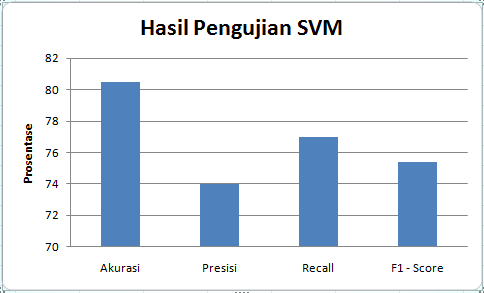
Akurasi = \* 100 = 80.5%

Presisi = \* 100 = 74%

Recall = \* 100 = 77%

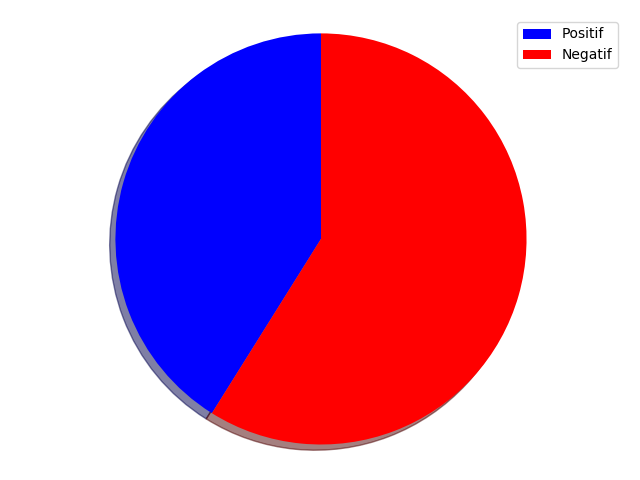
F1 Score = 2\* = 75.4%

Dengan perhitungan diatas, telah diketahui hasil dari akurasi adalah 80.5%, presisi 74%, recall 77% dan F1 Score 75.4%. histogram dari hasil pengujian dapat dilihat pada gambar 4.1



**Gambar 4.1** Hasil SVM

Penelitian ini mengklasikan 1000 data tweet ke dalam 2 kelas, 589(58.9%) kedalam kelas negatif, dan 411(41.1%) ke dalam kelas negatif dengan menggunakan lexicon based features. Diagram pie untuk prosentase sentimen pada tweet dapat dilihat pada gambar 4.2.



**Gambar 4.2** Prosentase Kalimat Sentimen

**Implementasi Tampilan Website**

Pada proses ini dilakukan implementasi tampilan website sesuai dengan rancangan yang dibuat pada bab sebelumnya. Hasil dari penelitian akan ditampilkan pada sebuah website dengan alamat url [http://kpuanalysissentiment.epizy.com](http://kpuanalysissentiment.epizy.com/).



**Gambar 4.3** Halaman Home

****

**Gambar 4.4** Halaman Tweet Positif

****

**Gambar 4.5** Halaman Tweet Negatif

1. **PENUTUP**

**Kesimpulan**

Penelitian ini menggunakan dokumen berbahasa Indonesia yang didapatkan dari jejaring sosial Twitter*.* Setelah data didapatkan dilakukan pre-processing untuk membersihkan dan membuat data menjadi bentuk yang lebih mudah diolah. Klasifikasi kelas sentimen pada penelitian ini dibagi menjadi 2 kategori yaitu positif dan negatif. Selanjutnya dilakukan pembobotan TF-IDF agar dapat diklasifikasi dengan metode *support vector machine* (SVM).

Berdasarkan Penelitian yang telah dilakukan, diperoleh kesimpulan bahwa penelitian analisis sentimen terhadap komisi pemilihan umum Indonesia dengan metode klasifikasi *support* *vector* *machine* (SVM) memperoleh hasil sentimen negatif sebesar 589 (58.9%) dan sentimen positif sebesar 411 (41.1%), yang berarti bahwa kpu mendapatkan lebih banyak tanggapan negatif di jejaring sosial twitter.

Penelitian ini juga menghasilkan tingkat akurasi sebesar 80.55%, presisi sebesar 74%, recall sebesar 77% dan f1-score sebesar 75.4% dengan komposisi data latih 80% dan data uji 20%. Hasil dari penelitian berhasil ditampilkan pada sebuah website dengan alamat url [http://kpuanalysissentiment.epizy.com](http://kpuanalysissentiment.epizy.com/).

**Saran**

Adapun berberapa saran yang diberikan pada penelitian ini untuk pengembangan lebih lanjut supaya mendapatkan hasil yang lebih baik adalah:

1. Data di bersihkan dari typo dan slangword supaya kamus dapat memberi nilai dengan benar
2. Memperbanyak kamus untuk pelabelan / scoring supaya hasil label lebih akurat
3. Menggunakan data dari media sosial lain selain Twitter untuk dapat mengetahui perbandingan keakuratan sentimen yang dihasilkan oleh metode klasifikasi yang digunakan.

**DAFTAR PUSTAKA**

[1] Anton Subagia. 2016*. Membuat Web dengan Php 7 dan Database PDO*

*MysqLi*. Elex Media Komputindo, Jakarta.

[2] F. A. Pozzi et al*.* 2017. *Challenges of Sentiment Analysis In Social Networks:*

*An Overview, Sentiment Analysis in Social Networks.* Elsevier, Camridge.

[3] Joshii Prateek. 2016. *Python: Real World Machine Learning*. Packt

Publishing Ltd, Birmingham.

[4] Purbo, Onno W. 2019. *Text Mining – Analisis Medsos, Kekuatan Brand, &*

*Intelijen di Internet*. ANDI, Yogyakarta

[5] Rosa A.S. 2014. *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi*

*Objek*. Informatika, Bandung.

[6] Schneider, K.-M. 2005. *Techniques for Improving the Performance*. *In*

*Proceedings of CICLing, 5-7.*

[7] Suyanto. 2019. *Data Mining untuk Klasifikasi dan Klasterisasi Data*.

Informatika, Bandung.

[8] Teguh Wahyono. 2018. *Fundamental of Python for Machine Learning*. Gava

Media, Yogyakarta.

[9] Yuniar Supardi dan Ading Hermawan. 2018. *Semua Bisa Menjadi*

*Programmer CodeIgniter Basic*. Elex Media Komputindo, Jakarta.

[10] Abu Ahmad. 2017. *Mengenal Artificial Intelligence, Machine Learning,*

*Neural Network, dan Deep Learning*.

[11] Imam Fahru., dkk. 2012. *Implementasi Opinion Mining (Analisis Sentimen)*

*untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi, Jurnal EECCIS Vol. 6, No. 1.*

[12] Jan Wira Gotama Putra. 2019. *Pengenalan Konsep Pembelajaran Mesin dan*

*Deep Learning.*

[13] Moch Ali Ramdhani. 2015. *Permodelan Proses Bisnis Sistem Akademik*

*Menggunakan Pendekatan Business Process Modelling Notation (BPMN) (Studi Kasus Institusi Perguruan Tinggi XYZ)*.

[14] Pamungkas, E. W & Putri, D. P. 2016*. An Experimental Study of Lexicon-*

*Based Sentiment Analysis on Bahasa Indonesia. Proceeding of The* 6th*International Annual Engineering Seminar (INAES).*

[15] Ramadhan W.P., dkk. 2017. *Analisis Sentimen Menggunakan Support Vector*

*Machine dan Maximum Entropy*. Telkom University.Bandung, Indonesia

[16] Sumartini Saraswati. 2011. *Text Mining dengan Metode Naive Bayes*

*Classifier dan Support Vector Machines untuk Sentiment Analysis*. Universitas Udayana.

[17] Wibawanto, Wanda. 2017. *Desain dan Pemrograman Multimedia*

*Pembelajaran Interaktif*. Cerdas Ulet Kreatif, Jember.

[18] <https://www.advernesia.com/blog/python/pengertian-bahasa-pemrograman-

python-dan-kegunaanya/>. 2017. “*Pengertian Bahasa Pemrograman Python*”. Advernesia. 12 Juli 2019.

[19] <https://algorit.ma/blog/data-science/pengertian-text-mining-dan-nlp/>. 2018.

“Pengertian NLP dan Text Mining”. Tommy Wijaya. 14 Agustus 2019.

[20] <http://www.teknologi-bigdata.com/2016/07/web-crawling-di-era-big-

data.html>.2016.” *Pentingnya Web Crawling sebagai Cara Pengumpulan Data di Era Big Data*”.Vijjam Wijaya. 2 September 2019

[21] <http://grahailmu.co.id/previewpdf/978-979-756-224-3-272.pdf>. 2019.

“*Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis*“. Budi Santosa, Graha Ilmu, Yogyakarta. 20 Juli 2019.

[22] <https://inixindojogja.co.id/mengenal-machine-learning>. 2019. “*Mengenal*

*Machine Learning*”. Inixindojogja. 20 Juli 2019

[23] <http://www.teguhwahyono.net>. 2018. “*AI Road Map: Fase-fase Booming*

*Kecerdasan Buatan* “. Wahyono, Teguh. 20 Oktober 2019.

[24] <ilmukomputer.org>. 2016. “*Pengantar BPMN Business Process Modeling*

*Notation*”. IlmuKomputer.Com. Yudho Yudhanto.