

Perbandingan Metode Naïve Bayes Classifier dan Support Vector Machine Untuk Klasifikasi Cyber Harassment Pada Twitter

Comparison of Naïve Bayes Classifier and Support Vector Machine Methods for Classifying Cyber Harassment on Twitter

Rezi Iwardani Saputri¹, Siti Khomsah², Novian Adi Prasetyo^{*3}

^{1,3}Program Studi Teknik Informatika, Institut Teknologi Telkom Purwokerto

²Program Studi Sains Data, Institut Teknologi Telkom Purwokerto

e-mail: ¹17102019@ittelkom-pwt.ac.id, ²siti@ittelkom-pwt.ac.id, ^{3*}novian@ittelkompwt.ac.id

Abstrak

Cyber Harassment dapat disebut juga dengan pelecehan online dapat berupa mengancam atau melecehkan melalui email, pesan instan, media sosial atau memposting informasi secara online. Kasus ini kerap terjadi di media sosial seperti salah satunya adalah Twitter. Untuk itu dibutuhkan sebuah metode klasifikasi yang tepat agar mengatasi kasus Cyber Harassment dari data Twitter. Pada penelitian ini menggunakan bahasa pemrograman Python dan menggunakan dua metode yaitu Naïve Bayes Classifier dan Support Vector Machine untuk membandingkan metode yang memiliki akurasi yang baik dan mengetahui kinerja metode masing-masing. Pada metode Naïve Bayes Classifier menggunakan model Complement Naïve Bayes dan Support Vector Machine menggunakan model Support Vector Classification (SVC). Hasil kinerja masing – masing metode dengan pembagian data training dan data testing yaitu 80% : 20% menunjukkan metode Naïve Bayes Classifier dengan accuracy 86.30%, precision 84.51% dan recall 87.21%. dan Support Vector Machine dengan accuracy 89.56%, precision 83.62% dan recall 94.5%. Dengan demikian metode Support Vector Machine lebih baik dari metode Naïve Bayes Classifier dan dapat diimplementasikan untuk kasus Cyber Harassment di Twitter.

Kata kunci: Cyber Harassment, Twitter, Python, Naïve Bayes Classifier dan Support Vector Machine

Abstract

Cyber harassment, also known as online harassment, can be in the form of threatening or harassing via email, instant message, social media or, posting information online. This case often occurs on social media, such as Twitter. For that, we need an appropriate classification method to overcome the Cyber Harassment case from Twitter. This study uses the Python programming language and uses two methods, namely the Naïve Bayes Classifier and Support Vector Machine to compare methods that have good accuracy and determine the performance of each method. The Naïve Bayes Classifier method uses the Complement Naïve Bayes model and the Support Vector Machine uses the Support Vector Classification (SVC) model. The results of the performance of each method by dividing the training data and testing data, namely 80%: 20%, indicate the Naïve Bayes Classifier method with an accuracy of 86.30%, precision of 84.51% and, a recall of 87.21%. and Support Vector Machine with 89.56% accuracy, 83.62% precision and, 94.5% recall. Therefore the Support Vector Machine method is better than the Naïve Bayes Classifier method and can be implemented for the Cyber Harassment case on Twitter.

Keywords: Cyber Harassment, Twitter, Python, Naïve Bayes Classifier and Support Vector Machine



1. PENDAHULUAN

Perkembangan teknologi seiring waktu semakin cepat, khususnya *smartphone* yang berbasis android. Adanya *smartphone* yang berbasis android ini berdampak pula semakin banyaknya media sosial yang bermunculan. Salah satu media sosial yang masih sering di gunakan masyarakat ialah *Twitter*. Berdasarkan peringkat media sosial berdasarkan jumlah pengguna di Indonesia *Twitter* berada di posisi ke-lima setelah *Youtube*, *Whatsapp*, *Facebook* dan *Instagram* [1]. Banyak aktivitas yang dapat dilakukan melalui *Twitter*, seperti membuat *tweet*, *me-re-tweet* milik akun orang lain, menyukai *tweet* serta mengirim pesan pribadi antar akun atau disebut *direct message* [2].

Setelah mengetahui banyaknya aktivitas yang dapat dilakukan, tidak jarang banyak pihak yang memanfaatkan *Twitter* untuk melakukan kejahatan yang dapat merugikan pihak lain yang biasa disebut dengan *Cybercrime*. *Cybercrime* adalah suatu kejadian atau kasus yang berhubungan dengan teknologi komputer (dunia maya) dimana seorang korban menderita berupa kerugian dan seorang pelaku dengan sengaja memperoleh keuntungan [3]. Jenis-jenis kejahatan kategori *cybercrime* yaitu *cyber terorisme*, *cyber pornography* dan *cyber harassment* [4]. *Cyber Harassment* bisa disebut juga dengan pelecehan online dapat berupa mengancam atau melecehkan melalui email, pesan instan, atau memposting informasi secara online. Dalam kasus ini menargetkan terhadap orang tertentu, baik dengan secara langsung menghubungi mereka atau dengan menyebarkan informasi pribadi mereka, menyebabkan mereka tertekan, takut, atau marah [5].

Pada tahun 2016, menurut lembaga Country Business Head Twitter Indonesia, menyatakan bahwa kicauan oleh pengguna *twitter* Indonesia dari bulan Januari hingga Desember 2016 yakni mencapai 4,1 miliar *tweet* [6]. Sebanyak 4,1 milyar *tweet* tersebut akan dikatakan percuma apabila disalah gunakan ke sesuatu yang negatif, tentu hal ini bertentangan dengan etika berinternet yang diatur oleh pemerintah yang terdapat di Undang-Undang Nomor 11 tahun 2008 tentang Informasi dan Transaksi Elektronik (UU ITE) [7].

Undang-Undang Nomor 11 tahun 2008 tentang Informasi dan Transaksi Elektronik disahkan pada tanggal 21 April 2008 sebagaimana telah diubah menjadi Undang-Undang Nomor 19 Tahun 2016 (UU ITE) dan menjadi *cyber law* pertama di Indonesia [8]. Menurut Komnas Perempuan pada tahun 2017 mencatat sebanyak 91 laporan pelecehan online. Selain itu pelecehan seksual dalam transportasi online juga sering terjadi, yakni dengan mencuri data milik konsumennya. Di tahun 2017 kasus siber yang dilaporkan atau tindakan aktif kepolisian mencapai 5.061 kasus yaitu dengan penyelesaian 3.325 kasus untuk kasus kebencian [9]. Hal ini menandakan penanganan kasus *harassment* di Indonesia masih tergolong lambat, sehingga korban merasa ragu untuk melaporkan kepihak berwenang seperti polisi atau orang tua mereka.

Pada penelitian ini menggunakan proses *text mining*, yaitu suatu teknik yang digunakan untuk menangani masalah klasifikasi, pengelompokan, ekstraksi informasi, dan pencarian informasi [10] serta memiliki sifat hampir sama dengan data mining namun hanya fokus pada teks daripada bentuk data yang lebih terstruktur [11]. Melakukan secara khusus melibatkan penggunaan teknologi *Natural Language Processing* (NLP), yang menerapkan prinsip-prinsip linguistik komputasional untuk menguraikan dan menginterpretasikan set data. *Natural Language Processing* berkaitan dengan komputer dapat digunakan untuk memahami dan memanipulasi teks bahasa alami (*natural language*) untuk mendapatkan informasi tertentu [12].

Pada penelitian ini menggunakan TF-IDF (*Term Frequency-Invers Document Frequency*) sebagai seleksi fitur. Metode TF-IDF ini terkenal efisien, mudah, dan memiliki hasil yang akurat [13]. Secara sederhana, metode TF-IDF digunakan untuk mengetahui berapa sering kata yang muncul dalam suatu dokumen. Pada penelitian yang dilakukan oleh Rizky Tri Wahyuni, Dhidik Prastiyanto dan Eko Suprpto yang berjudul Penerapan Algoritma *Cosine Similarity* dan



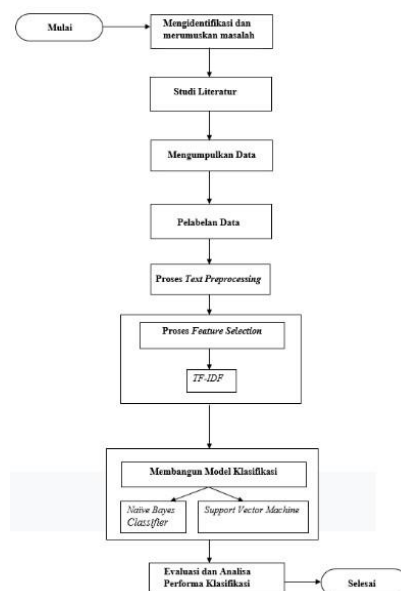
pembobotan TF-IDF pada sistem Klasifikasi Dokumen Skripsi yang menunjukkan hasil penelitian dengan presentase tingkat ketepatan klasifikasi sistem dengan sebesar 98% [14].

Ada beberapa penelitian yang sudah dilakukan seperti oleh Tri Mueri Sandes yang berjudul Klasifikasi *Tweet* Pelecehan Online Pada *Twitter* Dengan Menggunakan Metode *Naïve Bayes* pada tahun 2018 [15]. Pada penelitian tersebut menggunakan metode *naïve bayes* yang memiliki akurasi 85% dan kategori nya yang terdapat lima kelas yaitu *Racist*, *Embarrass*, *Physically threatned*, *Sexual Harassed* dan *Neutral* berbeda dengan penelitian ini hanya terdapat dua kelas. Sedangkan penelitian berikutnya yaitu dilakukan oleh Amer Saeed Ali Al-Katheri & Maheyzah Md Siraj yang berjudul “*Classification of Sexual Harassment on Facebook Using Term Weigthing Schemes*” pada tahun 2018 [16]. Penelitian ini tentang pelecehan seksual yang terjadi pada media sosial *Facebook*. Metode yang digunakan *Support Vector Machine* dengan tiga pembobotan yaitu *Entropy*, *TFIDF*, dan *Modified TFIDF*. Penelitian ini menggunakan metode SVM untuk mengklasifikasi dua dataset berdasarkan akurasi dan presisi serta hasil percobaan menunjukkan bahwa *Modified TFIDF* lebih tinggi dibandingkan *Entropy* dan *TFIDF*.

Setelah mempelajari penelitian terdahulu, penelitian ini akan membandingkan kinerja dua metode yaitu *Naïve Bayes Classifier* dan *Support Vector Machine*. Pemilihan metode *Naïve Bayes Classifier* karena merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki kemampuan dan akurasi tinggi [15]. Metode *Support Vector Machine* (SVM) sebagai pembanding dengan *Naïve Bayes Classifier* karena merupakan suatu teknik untuk melakukan prediksi, baik prediksi dalam kasus regresi maupun klasifikasi [17]. Metode *Naïve Bayes Classifier* dan *Support Vector Machine* termasuk klasifikasi *supervised learning* [18]. Kedua metode ini dipilih karena memiliki persamaan dalam melakukan proses klasifikasi serta dapat membandingkan metode yang lebih tepat dalam kasus ini.

Dapat dilihat dari penjelasan di atas, penelitian ini akan dituangkan ke dalam Proposal Tugas Akhir yang berjudul “Perbandingan Metode *Naïve Bayes Classifier* dan *Support Vector Machine* Untuk Klasifikasi *Cyber Harassment* Pada *Twitter*”. Alasan pengambilan judul ini berharap dari penelitian yang dilakukan terhadap data yang ada di *Twitter* untuk mengklasifikasi jenis kasus *cyber harassment* secara cepat dan akurat dengan perbandingan dua metode tersebut serta kelebihan nya belum ada yang melakukan penelitian ini.

2. METODE PENELITIAN



Gambar 1. Diagram Alir Penelitian

Langkah pertama yang dilakukan pertama kali yaitu mengidentifikasi dan merumuskan masalah tentang apa yang akan diteliti. Menentukan bidang, topik, masalah penelitian serta mengusulkan metode yang akan digunakan pada penelitian. Pada tahap ini mempelajari masalah yang masih terjadi di kehidupan sehari-hari, lalu tujuan, ruang lingkup serta metodologi penelitian. Pada penelitian ini ide atau masalah sudah ada penelitian terdahulunya, namun terdapat perbedaan pada algoritma atau metode yang di usulkan.

1.1 Studi literatur

Langkah selanjutnya dengan melakukan studi literatur atau riset kepustakaan. Pada tahap ini, peneliti melakukan pengumpulan data – data yang berkaitan dengan topik permasalahan yaitu tentang klasifikasi, kasus *Cyber harassment*, tentang tweet di *Twitter* serta algoritma yang akan diusulkan yaitu *naives bayes* dan *support vector machine*. Data – data untuk penelitian ini diperoleh dari jurnal, buku elektronik, situs internet dan media elektronik. Studi literatur memiliki tujuan untuk memperkuat permasalahan yang dibahas pada penelitian ini serta menjadi dasar untuk melakukan pengembangan selanjutnya.

1.2 Mengumpulkan Data

Langkah berikutnya adalah mengumpulkan data, data yang dikumpulkan berupa *tweet* pada *Twitter* tentang *cyber harassment*. Mengumpulkan data ini bertujuan agar mendapatkan informasi tentang kasus *cyber harassment* ini untuk dapat di analisis dengan cara klasifikasi menggunakan metode *Naive Bayes* dan *Support Vector Machine*. Dalam mengumpulkan data, peneliti melakukan dengan cara *crawling* menggunakan RapidMiner.

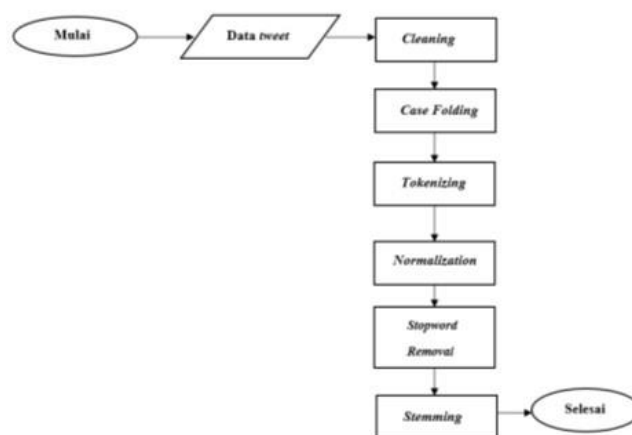
1.3 Pelabelan Data

Setelah mengumpulkan beberapa data, data tersebut masih bersifat mentah maka dibutuhkan label pada data tersebut dan bertujuan agar bisa melakukan klasifikasi. Label “1”

menandakan bahwa teks mengandung unsur pelecehan atau *harassment*. Sedangkan label “0” menandakan bahwa teks tidak mengandung unsur pelecehan atau *harassment*. Proses pelabelan data dilakukan secara manual yang dilakukan oleh peneliti dan beberapa orang sebagai Responden. Data *tweet* menggunakan bahasa Indonesia sehingga penelitian ini membutuhkan Kamus Besar Bahasa Indonesia (KBBI).

1.4 Text Preprocessing

Setelah melakukan pelabelan pada data maka selanjutnya langkah untuk *text preprocessing* yaitu *cleaning* dengan cara penghapusan retweet, URL, *emoticon*, dan selanjutnya *case folding*, *tokenizing*, *normalization*, *stopword removal* dan *stemming*.



Gambar 2. Proses *Text Preprocessing*

1.5 Proses *Feature Selection*

Selanjutnya data akan melakukan proses *feature selection* atau seleksi fitur menggunakan pembobotan *TF-IDF*. Adanya pembobotan fitur ini sebelum klasifikasi dilakukan agar membantu meningkatkan akurasi klasifikasi.

1.6 Membangun Model Klasifikasi

Selanjutnya membangun model klasifikasi menggunakan metode *Naïve Bayes Classifier* dan *Support Vector Machine* yang merupakan bagian data mining, proses ini merupakan inti dari proses ekstraksi pengetahuan dari data.

1.7 Evaluasi dan Analisa Performa Klasifikasi

Pada langkah evaluasi ini parameter yang digunakan adalah *confusion matrix*. Tujuan dari adanya evaluasi ini agar dapat melihat nilai *accuracy*, *precision* dan *recall* dengan melihat performa model klasifikasi *Naïve Bayes Classifier* dan SVM untuk kasus *cyber harassment* pada Twitter. Setelah mendapatkan hasil dari evaluasi maka dapat di tarik kesimpulan dari penelitian ini.

3. HASIL DAN PEMBAHASAN

1.1 Pengumpulan Data

Pada penelitian ini menggunakan data hasil dari *crawling tweet* pada Twitter dengan total 2300 data. Data tweet yang di unduh adalah data yang menggunakan berbahasa indonesia dan

mengacu dari beberapa tweet yang mengandung kata kunci jenis *harassment* yang populer dibicarakan di Twitter. Rentang waktu data yang di dapatkan yaitu pada bulan Oktober 2018 hingga bulan Januari 2021. Data *crawling* ini terdapat banyak yang menggunakan bahasa asing yaitu bukan Bahasa Indonesia maka peneliti melakukan pembersihan dan apabila masih data yang terdapat bahasa asing karena terdapat Bahasa Indonesia maka kata tersebut diabaikan.

Pengumpulan data di lakukan dengan *Crawling* menggunakan *tools* berupa Rapidminer dan rangkaian operator yang terhubung dengan API Twitter. Untuk mengumpulkan kata-kata yang mengandung *Cyberharassment* dengan membaca referensi dari berbagai sumber terkait *Cyberharassment*. Berikut kata kunci yang digunakan saat mengumpulkan data pada tabel sebagai berikut:

Tabel 1. Kata kunci

Jenis Harassment	Kata Kunci
Racial Harassment	Cina, kafir, agama, sesat
Sexual Harassment	Pelacur, mesum, pelakor, lonte, pantat
Appearance-related Harassment	Gendut, jelek, buta, botak, kecil, besar
Intellectual Harassment	Tolol, bodoh, bego
Political Harassment	Kadrun, cebong, buzzer

1.2 Pelabelan Data

Tahap selanjutnya setelah melakukan pengumpulan data adalah pelabelan data. Setelah mengumpulkan beberapa data, data tersebut masih bersifat mentah maka dibutuhkan label pada data tersebut dan bertujuan agar bisa melakukan klasifikasi.

Proses pelabelan dapat dilakukan filter sebagai berikut:

1. Pelabelan dilakukan mengandung *Cyberharassment* atau tidak pada data *tweet* berdasarkan kata kunci yang sudah ditentukan.
2. Selanjutnya memberi label pada seluruh data *tweet* sesuai yang telah ditentukan yaitu label kelas 1 yang berarti mengandung *cyberharassment* dan 0 yang berarti tidak mengandung *cyberharassment*.

	Text	Label
2	Apa yg bisa diharapkan dari politisi gaya pelacur, pramagtis, tidak jelas warna politiknya, tidak punya prinsip dlm berpolitik, sama saja intelektual	1
3	Syahganda Buka-bukaan Propaganda Hasil Survei Denny JA Di Depan Bawaslu https://politik.rnol.co/read/2019/04/24/387417/	1
4	Di @PDemokrat banyak org2 yg punya akal sehat demi kemajuan bangsa, tapi ada beberapa orang beguandal disertai pelacur politik yg	1
5	Cypridophobia adalah fobia atau takut pada pelacur atau penularan penyakit kelamin	0
6	Siapalah yang bagi idea suh kerja malam tu? Pelacur Intelektual #KPUCurangRakyatMelawan	1
7	#KPU Ngejar QuickCount https://twitter.com/SuwandaBen/status/1123805929210810368 #e	1
8	Lembaga Pelacur LSI DennyJA dan Germo @Metro_TV rupanya mengejar target "Diatas 50% Propinsi" dlm opini publik, simultan dg operasi perampokan suara dimana-mana.	1
8	Hina betul hidup mereka https://twitter.com/Metro_TV/status/1123983002730737664 #e	1
9	Bangunlah!!! Gaada orang yang sukses di kasar kecusul pelacur #selfreminder	1
10	abaikan semua pelacur agama! hidup NKRI! hidup Presiden Joko Widodo	1

Gambar 3. Data yang sudah diberi Label

Data tersebut kurang lebih sebanyak 2300 *tweet* yang kemudian di lakukan proses pelabelan data secara manual oleh peneliti, 23 orang sebagai Responden dan 1 orang sebagai responden inti. Peneliti membuat pembagian setiap 1 orang Responden memberi label kurang



lebih sebanyak 100 data. Setelah Responden memberi label dan di terima oleh peneliti namun di lihat kembali untuk *cross check* pada data tersebut. Apabila ada perbedaan antara Responden dan peneliti pada suatu data maka peneliti meminta responden inti untuk menilai data tersebut agar dapat menjadi nilai label dengan hasil keputusan terbanyak.

3.3 Text Preprocessing

3.3.1 Import Data Tweet

Data Tweet yang digunakan adalah hasil pengumpulan data dengan *crawling* dan sudah diberikan label. Data berasal dari berbagai kata kunci kemudian di bentuk kedalam satu *file* dengan format.

3.3.2 Cleaning

Cleaning yaitu dengan cara kata atau karakter yang akan dihilangkan seperti simbol, link URL, taggar (#), nama pengguna atau mention (@namaakun), *emoticon* dan RT (*Retweet*).

Tabel 2. *Cleaning*

Sebelum	Sesudah
@RAMujiyono: Semakin hari semakin banyak org yg buta nurani dan cenderung tak waras !	Semakin hari semakin banyak org yg buta nurani dan cenderung tak waras

3.3.3 Case Folding

Case folding merupakan tahap yang mengubah semua huruf dalam data menjadi lower case (huruf kecil). Hasil *case folding* pada Tabel 3.

Tabel 3. *Case folding*

Sebelum	Sesudah
Semakin hari semakin banyak org yg buta nurani dan cenderung tak waras	semakin hari semakin banyak org yg buta nurani dan cenderung tak waras

3.3.4 Tokenizing

Tokenizing merupakan tahap pemisah atau pemotongan setiap kata atau *term*. Library pada tokenizing menggunakan *nlTK*. Setelah melakukan implementasi ke pemrograman, maka terdapat contoh data yang sebelumnya dan setelah melakukan tokenizing pada Tabel 4.

Tabel 4. *Tokenizing*

Sebelum	Sesudah
semakin hari semakin banyak org yg buta nurani dan cenderung tak waras	“semakin”, “hari”, “semakin”, “banyak”, “org”, “yg”, “buta”, “nurani”, “dan”, “cenderung”, “tak”, “waras”

3.3.5 Normalization

Normalization bertujuan untuk menyeragamkan term yang memiliki makna yang sama namun berbeda penulisan biasanya faktor bahasa gaul atau *slang* yang disingkat. Pertama yang dilakukan pada *normalization* adalah membuat kamus *slang* secara manual. Pada bagian ini peneliti menggunakan Kamus Besar Bahasa Indonesia sebagai acuan data yang tepat.

```
slangs={'ndhal':'ndahal','sm':'sama','pju':'hejitu','nyantik':'ranik','syuka':'suka','nggak':'tidak','ane':'aku','kimi':'keiru',
'yg':'yang','aq':'aku','ajg':'anjing','mndng':'mending','gak':'tidak',
'kagak':'tidak','dgn':'dengan','smpahhh':'smpah','piso':'pisau','lngsung':'langsung','anjg':'anjing','ajg':'anjing',
'gw':'gue','lg':'lagi','tp':'tapi','ngk':'tidak','gua':'gue','nggak':'tidak','kn':'kamu','y':'ya','eng':'emang',
'bner':'benar','org':'orang','bgt':'banget','dn':'dari','udh':'sudah','bilang':'bilang','utk':'untuk','pake':'pakai',
'sbg':'sehagai','kn':'karena','kop':'kenapa','smpn':'sampai','gblk':'goblok','tau':'tau','alesan':'alasan','w':'gun',
'cewe':'cewek','cowo':'cowok','ga':'tidak','td':'tadi','kaga':'tidak','trs':'terus','skrng':'sekarang','km':'kamu','lo':'lu',
'dripda':'daripada','kalo':'kalau','jgnn':'jangan','jg':'juga','nrk':'mereka','nls':'malas','tmi':'cuma','grgr':'karena',
'kadrunista':'kadrun','china':'cina','nebesar':'besar','gk':'tidak','china':'cina'}
```

Gambar 4. Kamus *slang*

Setelah kamus *slang* telah dibuat, lalu melakukan implementasi ke pemrograman, maka terdapat contoh data yang sebelumnya dan setelah melakukan tokenizing pada Tabel 5.

Tabel 5. *Normalization*

Sebelum	Sesudah
“semakin”, “hari”, “semakin”, “banyak”, “org”, “yg”, “buta”, “nurani”, “dan”, “cenderung”, “tak”, “waras”	“semakin”, “hari”, “semakin”, “banyak”, “orang”, “yang”, “buta”, “nurani”, “dan”, “cenderung”, “tak”, “waras”

3.3.6 Stopword Removal

Stopword Removal merupakan tahap menghapus kata yang tidak penting (tidak memiliki arti). Pada tahap ini menggunakan library NLTK untuk *filtering* terhadap *DataFrame*. Terdapat contoh data yang sebelumnya dan setelah melakukan *stopword removal* pada Tabel 6.

Tabel 6. *Stopword Removal*

Sebelum	Sesudah
“semakin”, “hari”, “semakin”, “banyak”, “orang”, “yang”, “buta”, “nurani”, “dan”, “cenderung”, “tak”, “waras”	“semakin”, “hari” , “semakin”, “banyak”, “orang”, “yang” , “buta”, “nurani”, “dan” , “cenderung”, “tak”, “waras”

3.3.7 Stemming

Stemming merupakan tahap yang mengubah kata yang berimbuhan menjadi kata dasar. Stemming pada penelitian ini menggunakan *library Sastrawi* dan *library swifter* untuk mempercepat proses stemming pada *DataFrame* karena fungsi *stemmer.stem()* *library sastrawi* yang lambat.

Setelah melakukan implementasi ke pemrograman, maka terdapat contoh *output* dari *stemming* pada Tabel 7.



Tabel 7. Stemming

Sebelum	Sesudah
“semakin”, “semakin”, “banyak”, “orang”, “yang”, “buta”, “nurani”, “dan”, “cenderung”, “tak”, “waras”	“makin”, “makin”, “banyak”, “orang”, “buta”, “nurani”, “cenderung”, “tak”, “waras”

Setelah menyelesaikan semua rangkaian atau proses dari *text preprocessing* maka akan menghasilkan *output preprocessing* seperti Gambar 5.

```

2300 2300
0 apa politis gaya lacur pragmatik tak jelas warna politik tak punya prinsip politik sana tak milik integritas cukup yakin po
litis model begini tak suka rakyat 1
1 kalau jadi gugat memang benar memang lacur buka propaganda survei denyi ja depan bawaslu politik baca 1
2 banyak punya akal maju bangsa beberapa orang begundal serta lacur politik grasak grusuk satu tak twitter com status 1
3 fobia takut lacur tular sakit kelamin 0
4 ramai kerja waktu malam ni biasa lacur idea kerja malam tu 1
5 lacur twitter com suwandaben status 1
6 lembaga lacur lsi denyija germo rupa kejar target atas propinsi opini publik simultan dg operasi ampo suara mana benar twit
ter com status 1
7 gaada orang sukses kasur lacur selfreinder 1
8 abai semua lacur agama nkri presiden joko widodo 1
9 mungkin saat semua segala jenis kotor bangsa massive awal laknat lacur agama biadab ancam selamat nkri 1
10 tak sukses capai atas kasur lacur bajing 1
11 gaya politis gaya lacur pragmatik yakin politis model begini tak suka rakyat 1
12 gugat memang benar memang lacur 1
13 banyak punya akal maju bangsa beberapa orang begundal serta lacur politik grasak grusuk satu tak twitter com status 1
14 ramai kerja waktu malam ni biasa lacur idea kerja malam tu 1
    
```

Gambar 5. Hasil *Preprocessing*

3.4 Feature Selection

Setelah proses *text preprocessing* selesai, langkah selanjutnya adalah melakukan *feature selection* (seleksi fitur) yaitu pembobotan kata atau term yang pada penelitian ini menggunakan *TF-IDF*. Implementasi *TF-IDF* dengan pemrograman python menggunakan pengujian *library sklearn.feature_extraction.text*.

Memperoleh *term* atau kata unik sebanyak 4840 yang dapat dilihat pada Gambar 3.4. Berupa kumpulan *term* atau kata unik.

```

jumlah kolom fitur (jumlah term unik) = 4840
['abai', 'abang', 'abet', 'abi', 'abidzar', 'abin', 'abis', 'abnormal', 'abong', 'abrik', 'absen', 'abu', 'abud', 'acak', 'ac
ar', 'acara', 'acc', 'accu', 'acne', 'ada', 'adab', 'adain', 'adegan', 'adem', 'adik', 'adil', 'admin', 'adrenalin', 'adrena
line', 'adu', 'adulkitporn', 'advertising', 'after', 'agam', 'agama', 'agar', 'agasa', 'agenda', 'agens', 'agent', 'agitas
i', 'ago', 'ai', 'aipek', 'air', 'airmataku', 'aiyu', 'aja', 'ajak', 'ajang', 'ajar', 'akademika', 'akal', 'akang', 'akbar',
'akibat', 'ako', 'akonentar', 'akrab', 'akses', 'aktif', 'aktng', 'aktivitas', 'aktor', 'aku', 'akuarium', 'akun', 'al', 'al
a', 'aladin', 'alak', 'alam', 'alami', 'alas', 'alat', 'alay', 'albert', 'album', 'alebong', 'alesane', 'alex', 'ali', 'alians
i', 'alias', 'alim', 'alir', 'alive', 'all', 'alladin', 'alm', 'almetnya', 'alreit', 'also', 'alternakati', 'alternatif', 'a
lterpnoy', 'alterworld', 'alus', 'asal', 'aman', 'amandel', 'amat', 'amata', 'awbil', 'awbul', 'asburadul', 'ambyar', 'amc
a', 'american', 'amerika', 'amin', 'amit', 'ampas', 'amplop', 'ampo', 'apung', 'amsterdam', 'an', 'ana', 'anak', 'analisis',
'analogi', 'ananta', 'anaqi', 'anarkis', 'anatomi', 'ancan', 'ancen', 'ancol', 'ancur', 'and', 'anda', 'andai', 'andd', 'andi
k', 'andre', 'aneka', 'anette', 'ang', 'anget', 'anggap', 'anggota', 'angguk', 'anggur', 'angin', 'angka', 'angkat', 'angko
t', 'aninasi', 'anime', 'anjay', 'anjing', 'anjir', 'anjrit', 'annuarmusa', 'ano', 'ansur', 'antek', 'antem', 'anti', 'antisi
pasi', 'antusias', 'anus', 'apa', 'apabila', 'aparap', 'apart', 'apas', 'api', 'aplikasi', 'apo', 'app', 'apresiasi', 'apun
g', 'aquarium', 'arab', 'are', 'area', 'arek', 'argentina', 'argumen', 'argument', 'argumentasi', 'ariety', 'arjuna', 'army',
'aroma', 'around', 'arrietty', 'arrow', 'art', 'artdudut', 'arti', 'artis', 'arudra', 'arum', 'arus', 'as', 'asa', 'asal',
'asan', 'asi', 'asik', 'asimetris', 'asin', 'asing', 'ask', 'askmewsuppsit', 'asli', 'aspal', 'asrama', 'assalamualaikum',
'astaga', 'at', 'atas', 'atau', 'ati', 'atlet', 'atmi', 'ato', 'atsuko', 'attack', 'attitude', 'atu', 'atun', 'atur', 'atut',
'au', 'audience', 'audisi', 'audit', 'aurat', 'australia', 'auto', 'awa', 'available', 'avocado', 'aw', 'awal', 'awan', 'awa
    
```

Gambar 6. Term unik

3.5 Membangun Model Klasifikasi

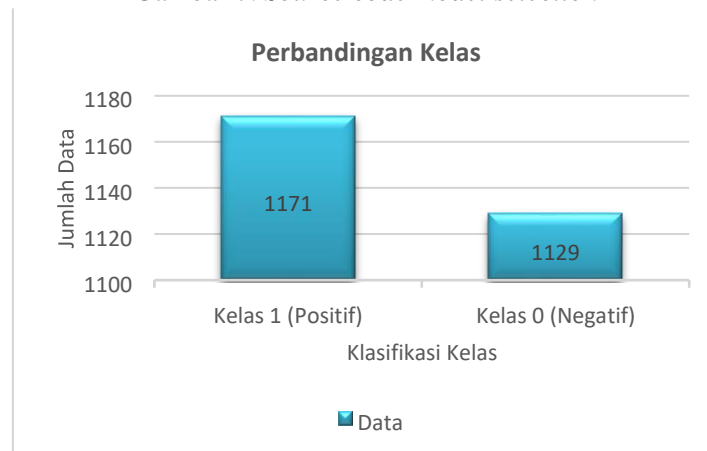
Setelah mendapatkan *term* pada semua dokumen, selanjutnya melakukan klasifikasi. Pengujian klasifikasi menggunakan dua metode yaitu *Naïve Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM). Pada pengujian ini menggunakan pembagian data training dan data testing 80%:20% dengan total data sebanyak 2300 yaitu 1171 dengan label 1 dan 1129 dengan label 0.

```

from sklearn import model_selection, preprocessing #import library
from matplotlib import pyplot as plt #import library
#split train 80% & test 20%
x_train, x_test, y_train, y_test = model_selection.train_test_split(X, y, stratify = y, test_size=0.2)
    
```



Gambar 7. Source code model selection



Gambar 8. Diagram perbandingan kelas 1 dan 0

3.5.1 Metode *Naïve Bayes Classifier*

Implementasi klasifikasi setiap masing-masing ini menggunakan *library sklearn* pada metode *Naïve Bayes Classifier* dengan model *Complement Naïve Bayes* (ComplementNB). Berikut hasil Klasifikasi dengan metode *Naïve Bayes Classifier* dengan model *Complement Naïve Bayes*.

```

=====
Accuracy : 0.8630434782608696
=====classification_report=====
              precision    recall  f1-score   support

   0             0.87       0.85       0.86         226
   1             0.85       0.88       0.87         234

 accuracy                   0.86         460
 macro avg              0.86       0.86       0.86         460
 weighted avg          0.86       0.86       0.86         460
=====
    
```

Gambar 9. Model Klasifikasi *Naïve Bayes Classifier*

3.5.2 Metode *Support Vector Machine*

Implementasi klasifikasi selanjutnya menggunakan *library sklearn* seperti klasifikasi sebelumnya yaitu *Support Vector Machine* dengan model *Support Vector Machine* (SVC). Pada implementasi ini, menggunakan parameter dengan *default* yaitu ketentuannya sebagai berikut:

Tabel 8. Parameter *Support Vector Machine*

Parameter
<code>class sklearn.svm.SVC(*, C=1.0, kernel='rbf', degree=3, gamma='scale', coef0=0.0, shrinking=True, probability=False, tol=0.001, cache_size=200, class_weight=None, verbose=False, max_iter=1, decision_function_shape='ovr', break_ties=False, random_state=None)</code>

Berikut hasil Klasifikasi dengan metode *Support Vector Machine* dengan model *Support Vector Classification* (SVC).


```

=====
Accuracy : 0.8956521739130435
=====classification_report=====
      precision    recall  f1-score   support

     0       0.94      0.84      0.89       226
     1       0.86      0.95      0.90       234

 accuracy          0.90          0.90          0.90          460
 macro avg         0.90          0.89          0.90          460
 weighted avg     0.90          0.90          0.90          460
=====
    
```

Gambar 10. Model Klasifikasi *Support Vector Machine*

3.6 Evaluasi dan Analisa Performa Klasifikasi

Pada tahap terakhir yaitu melakukan pengujian pada model klasifikasi dengan metode yang digunakan *Confusion Matrix*. Tujuan tahap ini adalah memvalidasi model setelah melakukan *training*. Pada *Confusion Matrix* ini menghasilkan berupa matrix yang terdiri dari nilai *True Positive*, *True Negative*, *False Positive* dan *False Negatif*. Kemudian hasil nilai tersebut dapat menghasilkan nilai *accuracy*, *recall* dan *precision*. Implementasi *confusion matrix* ini ke dalam pemrograman *python* menggunakan *library sklearn*. Berikut merupakan source code library yang digunakan:

Tabel 9. Implementasi *Confusion Matrix*

```

from sklearn.metrics import confusion_matrix #import library print(f"Confusion
Matrix :",confusion_mxatrix(y_test, y_preds))
    
```

Setelah melakukan implementasi source code pada tabel xx maka dapat di peroleh berupa matrix yang terdiri dari *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)* dan *False Negatif (FN)* yang akan ditampilkan pada table xx sebagai berikut:

Tabel 10. *Confusion Matrix Naïve Bayes Classifier*

	Cyberharassment (+)	Non-Cyberharassment (-)
Cyberharassment (+)	(TP) 191	(FP) 35
Non-Cyberharassment (-)	(FN) 28	(TN) 206

Pada perhitungan di atas menyatakan hasil klasifikasi dengan Metode *Naives Bayes Classifier* yaitu *Accuracy* 86.30%, *Precision* 84.51% dan *Recall* 87.21%.

Tabel 11. *Cunfusion Matrix Support Vector Machine*

	Cyberharassment (+)	Non-Cyberharassment (-)
Cyberharassment (+)	(TP) 189	(FP) 37
Non-Cyberharassment (-)	(FN) 11	(TN) 223

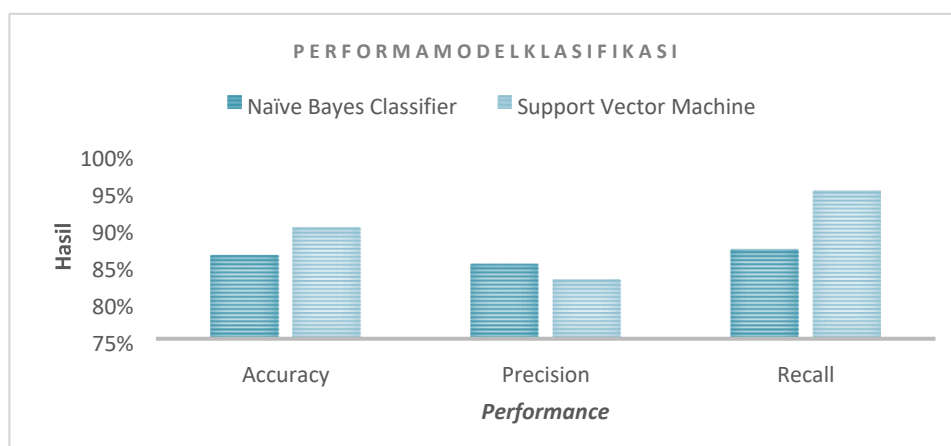


Pada perhitungan di atas menyatakan hasil klasifikasi dengan Metode *Support Vector Machine* yaitu *Accuracy* 89.56%, *Precision* 83.62% dan *Recall* 94.5%.

Tabel 12. Perbandingan *Accuracy*, *Precision* dan *Recall*

	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
<i>Naïve Bayes Classifier</i>	86.30%	84.51%	87.21%
<i>Support Vector Machine</i>	89.56%	83.62%	94.5%

Pada Tabel 12 menunjukkan bahwa kinerja *Support Vector Machine* lebih baik dengan *accuracy* 89.56%, *precision* 83.62% dan *recall* 94.5% dibandingkan kinerja *Naïve Bayes Classifier* dengan *accuracy* 86.30%, *precision* 84.51% dan *recall* 87.21%.



Gambar 11. Performa Model Klasifikasi

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dapat diambil beberapa kesimpulan dari seluruh tahap penelitian. Kesimpulan tersebut adalah Penggunaan dua metode yaitu *Naïve Bayes Classifier* dan *Support Vector Machine* pada penelitian ini untuk melakukan klasifikasi data tweet kedalam dua kelas yaitu mengandung *Cyberharassment* dan tidak mengandung *Cyberharassment*. Hasil kinerja masing – masing metode dengan pembagian data training dan data testing yaitu 80% : 20% menunjukkan metode *Naïve Bayes Classifier* dengan *accuracy* 86.30%, *precision* 84.51% dan *recall* 87.21%. dan *Support Vector Machine* dengan *accuracy* 89.56%, *precision* 83.62% dan *recall* 94.5%. Percobaan dengan tanpa menggunakan *stemming* hasil nya tidak jauh berbeda dengan menggunakan *stemming*, namun hasil akhir lebih baik menggunakan *stemming*. Perbandingan metode *Naïve Bayes Classifier* dan *Support Vector Machine* pada kasus *Cyberharassment* di Twitter dapat dilihat pada performa atau kinerja klasifikasi menyatakan bahwa *Support Vector Machine* lebih baik dibandingkan *Naïve Bayes Classifier* sehingga dapat digunakan untuk memberikan solusi terhadap kasus *Cyberharassment* pada Twitter.

5. DAFTAR PUSTAKA

- [1] D. H. Jayani, "10 Media Sosial yang Paling Sering Digunakan di Indonesia," 26 February 2020. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2020/02/26/10media-sosial-yang-paling-sering-digunakan-di-indonesia>.
- [2] Y. Kusuma, *Creative Project: Pintar Twitter*, Jakarta: Grasindo, 2009.
- [3] A. Hamzah, *Aspek-aspek pidana di bidang komputer*, Jakarta: Sinar Grafika, 1987.
- [4] S. Aisyah, "Kejahatan Dunia Maya (Cybercrime)," 29 November 2019. [Online]. Available: <https://pa-sidoarjo.go.id/informasi-pengadilan/223-kejahatan-dunia-mayacybercrime>.
- [5] P. R. Clearinghouse, "Online Harassment & Cyberstalking," 25 Oktober 2018. [Online]. Available: <https://privacyrights.org/consumer-guides/online-harassment-cyberstalking>.
- [6] R. Simangunsong, "Indonesia Hasilkan 4,1 Miliar Cuitan Twitter di Tahun 2016," *Tempo.co*, 6 Desember 2016. [Online]. Available: <https://tekno.tempo.co/read/825782/indonesia-hasilkan-41-miliar-cuitan-twitter-di-tahun2016>. [Accessed 5 Mei 2020].
- [7] R. M. Putri, "Etika berinternet menurut UU ITE," Retno Mertoyoso, 1 September 2015. [Online]. Available: http://mertoyoso97.blogspot.com/p/blog-page_17.html. [Accessed 20 July 2020].
- [8] B. Charlotte, "Cyberstalking Sebagai Perbuatan Melawan Hukum dan Pengaturannya Dalam Hukum Pidana Indonesia," *Universitas Katolik Parahyangan*, 2018.
- [9] M. Arista, "Kerentanan Perempuan di Dunia Maya?," 1 Maret 2019. [Online]. Available: <http://hakasasi.id/article/detail/125?name=Kekerasan+Seksual+Online%3A+Bukti+Kerentanan+Perempuan+di+Dunia+Maya%3F>.
- [10] M. W. Berry and J. Kogan, *Text Mining Application and Theory*, United Kingdom: WILEY, 2010.
- [11] U. D. Arni, "Apa Itu Text Mining?," *Garuda Cyber*, 25 Oktober 2018. [Online]. Available: <https://garudacyber.co.id/artikel/1254-apa-itu-text-mining>. [Accessed 20 July 2020].
- [12] N. Indrawati, "NATURAL LANGUAGE PROCESSING (NLP) BAHASA INDONESIA SEBAGAI PREPROCESSING PADA TEXT," *Docplayer Info*, 2010. [Online]. Available: <https://docplayer.info/32707775-Natural-language-processing-nlp-bahasaindonesia-sebagai-preprocessing-pada-text-mining.html>.
- [13] D. Sierra, "Algoritma TF — IDF," *Medium*, 13 February 2019. [Online]. Available: <https://medium.com/@dltsierra/algoritma-tf-idf-633e17d10a80>. [Accessed 2020 July 25].
- [14] R. T. Wahyuni, D. Prastiyanto and E. Suprpto, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi," *Jurnal Teknik Elektro*, vol. 9, pp. 8-23, 2017.
- [15] T. M. Sandes, "Klasifikasi Tweet Pelecehan Online Pada Twitter dengan Menggunakan Metode Naive Bayes," *Universitas Islam Negeri Sultan Syarif Kasim Riau*, 2018.
- [16] A. S. A. Al-Katheri and M. M. Siraj, "Classification of Sexual Harassment on Facebook," *International Journal of Innovative Computing*, pp. 15-19, 2018.
- [17] A. Mustopa, Hermanto and A. Y. Kuntoro, "Algoritma Klasifikasi Naive Bayes dan Support Vector Machine dalam Layanan Komplain Mahasiswa," *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, pp. 211-220, 2020.
- [18] J. W. G. Putra, *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning*, Tokyo: Wiragotama, 2019.

