Penerapan NBC dan SVM Terhadap Klasifikasi Opini Sistem Pendidikan Menggunakan Data Twitter

Nadya Hidayatun Najah[1], Naim Rochmawati[2]

[1,2]Jurusan Teknik Informatika/Teknik Informatika, Universitas Negeri Surabaya

[1]nadyanajah16051204003@mhs.unesa.ac.id

[2]naimrochmawati@unesa.ac.id

***Abstrak*— Penyebaran berita saat ini dapat diakses dengan mudah melalui media sosial. Twitter merupakan salah satu media sosial yang sangat populer di kalangan pengguna internet. Di sini pengguna dapat dengan bebas berkomentar dan menulis mengenai berita apapun. Salah satu berita yang banyak dibicarakan adalah mengenai sistem pendidikan yang dibuat oleh pemerintah khususnya kurikulum pendidikan. Dari media sosial tersebut, mengizinkan penggunanya untuk dapat mengetahui reaksi pengguna lain terhadap berita yang sedang dibicarakan. Untuk mengklasifikasikan bentuk sentimen yang diberikan apakah itu sentimen positif, negatif, atau netral dapat menggunakan *text mining*, yaitu proses menggali informasi dalam sekumpulan dokumen besar secara otomatis. Kurikulum pendidikan dipilih untuk dilakukan klasifikasi sentimen karena pembahasan kurikulum pendidikan sangat sensitif terhadap opini-opini masyarakat. Dalam mengklasifikasikannya penulis menggunakan algoritma *Naïve Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM). Dari proses pengklasifikasian dengan *Naïve Bayes Classifier* menghasilkan *accuracy* 64%, *precision* 100%, *recall* 54%, dan *F-Measure* 70%, sedangkan dengan *Support Vector Machine* menghasilkan *accuracy* 96%, *precision* 96%, *recall* 100%, dan *F-Measure* 98%. Hal ini menunjukkan bahwa metode SVM mampu mengklasifikasikan data komentar lebih baik daripada menggunakan metode NBC.**

Kata Kunci— *Naïve Bayes Classifier*, *Support Vector Machine*, Klasifikasi, Twitter, Kurikulum Pendidikan

I. PENDAHULUAN

Teknologi merupakan sarana prasarana untuk menyediakan kebutuhan kelangsungan hidup manusia agar memudahkan dalam pengerjaannya. Seiring dengan perkembangan teknologi informasi dan banyaknya kebutuhan akan nilai tambah dari *database* berskala besar, *data mining* menjadi salah satu bidang yang memiliki perkembangan sangat cepat.

Data mining menjadi alat yang semakin penting untuk mengubah data menjadi informasi, seperti untuk pengawasan, penipuan deteksi, pemasaran, dan penemuan ilmiah. Salah satu metode yang digunakan dalam pengolahan *data mining* adalah klasifikasi. Adapun algoritma pengklasifikasian dalam *data mining*, diantaranya *decision tree*, k-NN, *Naïve Bayes Classifier* (NBC), *Classification and Regression Trees* (CART), *Support Vector Machine* (SVM), dan lain-lain.

Perkembangan internet sebagai *new media* (*the second media age*) menjadi sebuah era baru dimana teknologi interaktif dan komunikasi jaringan khususnya dunia maya bisa mengubah masyarakat [1]. Media sosial merupakan salah satu media *online* untuk berekspresi dan berpendapat mengenai berbagai macam topik. Sosial media menjadi media baru komunikasi dan alat kolaborasi yang memudahkan banyak jenis interaksi yang sebelumnya tidak tersedia untuk orang biasa [2]. Salah satunya adalah twitter.

Twitter merupakan media terbesar untuk masyarakat dalam mengemukakan opini mengenai isu yang sedang hangat dibicarakan. Opini-opini dalam twitter bisa dijadikan sebagai cara untuk menilai atas suatu topik tertentu, seperti kebijakan pemerintahan, *public figure*, jasa, produk, film, dan lain sebagainya. Opini-opini tersebut dapat diklasifikan menjadi komentar positif, komentar negatif, dan komentar netral.

Mengenai hal-hal yang sering dijadikan pembicaraan hangat masyarakat, yaitu mengenai pendidikan. Indonesia yang menganut sistem pendidikan nasional dan berupaya untuk memberikan pengetahuan akademis, mengasah keterampilan, dan membina sikap positif para siswa. Tetapi sistem pendidikan nasional di Indonesia masih belum dapat dilaksanakan sebagaimana mestinya, maka dalam hal ini sistem harus menyesuaikan kurikulum yang ada saat ini. Oleh sebab itu, kurikulum di Indonesia sering mengalami perubahan karena mengikuti perkembangan masyarakat maupun ilmu pengetahuan serta teknologi. Sehingga bisa berdampak baik dan buruk bagi mutu pendidikan.

Berdasarkan penjelasan di atas, penulisan akan melakukan klasifikasi sentimen analisis terhadap opini masyarakat mengenai kurikulum pendidikan melalui data twitter dengan menggunakan algoritma NBC dan SVM yang dikolaborasikan dengan ekstraksi fitur TF-IDF.

II. PENELITIAN TERKAIT

Penelitian terkait mengenai sentimen analisis untuk klasifikasi data twitter, diantaranya seperti penelitian pada [3] peneliti mengelompokkan *tweet* yang mengandung informasi

mengenai amnesti pajak. Dalam pengklasifikasian amnesti pajak dengan menggunakan NBC, menggunakan ekstraksi fitur unigram. Sehingga pada penelitian ini menghasilkan data *training* sebesar 80% dari 578 data *tweet* amnesti pajak dengan akurasi sebesar 53,45%.

Pada penelitian [4] peneliti menggunakan 3300 data *tweet* tentang analisis kurikulum 2013. Data tersebut dikelompokkan secara manual dan dibagi untuk sentimen positif, negatif dan netral. 3000 data tweet dan 1000 tweet tiap kategori sentimen digunakan sebagai proses latih. Hasil dari penelitian ini, sistem dapat mengklasifikasi sentimen secara otomatis dengan hasil pengujian 3000 data latih dan 1000 tweet data uji coba mencapai 91%.

Selanjutnya pada penelitian [5] menggunakan algoritma SVM dengan fitur *Lexicon Based*. Dengan menggunakan parameter C bernilai 10 dan *learning rate* senilai 0,03 serta menggunakan *Lexicon Based Features* dengan iterasi sebanyak 50 kali maka menghasilkan *accuracy* sebesar 40%, *precision* 40%, *recall* 100%, dan *f-measure* sebesar 57,14%.

Penerapan *Naïve Bayes* digunakan untuk analisis sentimen [6]. Penelitian tersebut bertujuan untuk mengetahui pendapat masyarakat mengenai Rohingnya yang akan dibagi menjadi 3 kelompok yaitu positif, negatif, dan netral. Dari penelitian ini diperoleh hasil bahwa positif dan negatif seimbang, sedangkan netral lebih sedikit. Jumlah tweet terbanyak tidak diketahui. Hal ini disebabkan sulitnya mengolah data teks, ini mungkin karena ketidaklengkapan kamus atau mungkin karena hal lain.

Lalu penelitian [7] dari 936 pengujian data, ada 703 komentar netral, maka 102 komentar positif, dan 4 komentar negatif. Sehingga dapat disimpulkan bahwa pengguna media sosial di Indonesia adalah netral terhadap kampanye anti-LGBT, tetapi lebih mendukung kampanye Anti-LGBT daripada mereka yang menolak. Berdasarkan hasil keakuratan algoritma *Naïve Bayes* dalam jumlah 83,43%, yang lebih tinggi dari akurasi Algoritma *Decision Tree* dan Algoritma *Random Forest* dapat disimpulkan bahwa algoritma *Naïve Bayes* sangat baik digunakan untuk analisis perilaku sentimen di kasus kampanye Anti-LGBT di Twitter media sosial.

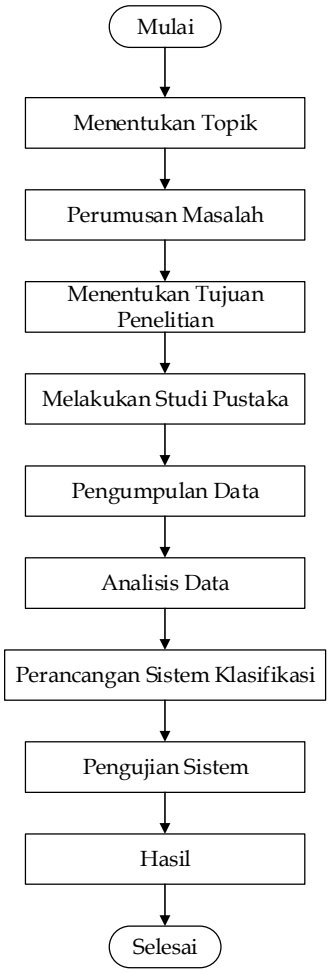
Penerapan lain dari *Naïve Bayes* pada penelitian [8] untuk mendapatkan pengetahuan dari opini online dan mengkategorikannya dalam tiga kelompok besar, yaitu: Distro yang paling banyak membahas Aspek, Polaritas sentimen dan Polaritas aspek distro dari opini online. mampu mengklasifikasikan opini pengguna toko ritel online dengan hasil presisi 97.25%, recall 89.83%, dan akurasi 89.21%.

Support Vector Machine pada [9] mengklasifikasikan posting twitter KAI dengan metode OAA *Multiclass* SVM dengan lima fitur pembobotan yang berbeda yaitu unigram, bigram, trigram, unigram + bigram, dan *wordcloud* untuk menganalisis data tweet. Memberikan hasil akurasi tertinggi adalah model TF-IDF unigram yang digabungkan dengan OAA Multiclass SVM dengan gamma 0.7 yaitu 80.59.

Dari beberapa penelitian yang telah dilakukan, penulis akan melakukan penelitian terrhadap tweet mengenai kurikulum pendidikan dengan membandingkan metode *Naïve Bayes Classifier* dan *Support Vector Machine*.

III. METODOLOGI PENELITIAN

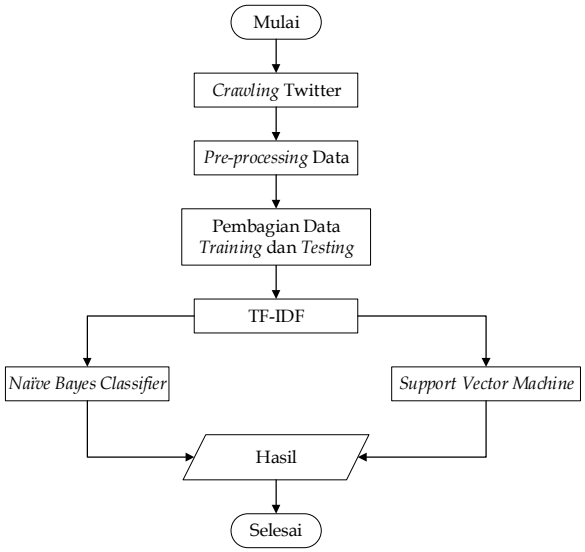
Metodologi penelitian merupakan pedoman dalam pelaksanaan penelitian agar alur serta hasil dari penelitian yang dicapai tidak menyimpang dari tujuan yang telah ditentukan. Gbr. 1 berikut merupakan diagram alur penelitian ini.



Gbr. 1 Diagram alur penelitian

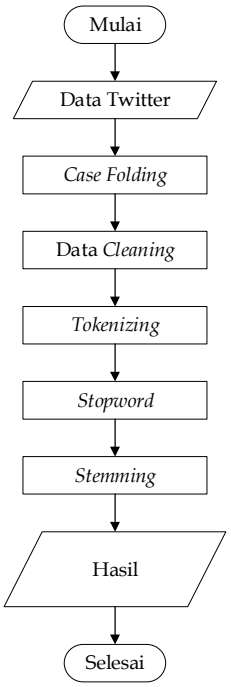
Penelitian ini mengenai analisa pengklasifikasian data *tweets* opini pengguna twitter terhadap kurikulum pendidikan menggunakan algoritma *Naïve Bayes Classifier* dan *Support Vector Machine*. Data twitter diperoleh dari proses *crawling*,

selanjutnya data akan memasuki proses *pre-processing* sebelum data terbagi menjadi 90% data *training* dan 10% data *testing*. Setelah data terbagi, data akan dihitung pembobotan fitur TF-IDF. Adapun gambaran umum sistem seperti pada Gbr. 2 berikut :



Gbr. 2 Gambaran umum sistem

Dari gambaran umum sistem, proses *pre-processing* data dapat diuraikan pada Gbr. 3 berikut :

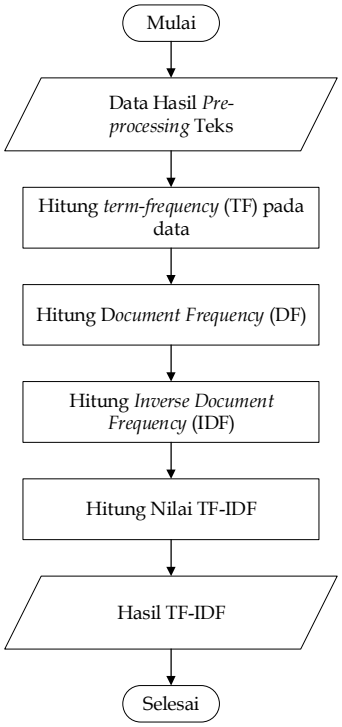


Gbr. 3 Proses *pre-processing* data

Pengambilan data twitter dengan teknik *crawling* menggunakan bahasa pemrograman python. Kata kunci yang digunakan adalah kurikulum. Data yang diperoleh akan disimpan dalam format txt. Kemudian barulah memasuki proses *pre-processing*, dimana *pre-processing* merupakan tahapan pengolahan data mentah, tidak terstruktur dan memiliki format asli diubah menjadi terstruktur, lebih mudah diolah dan efektif untuk menghasilkan data yang siap digunakan. Tahapan yang dimaksud antara lain :

a. *Case Folding*
   Dalam tahap *case folding*, mengonversi data *tweet* yang terkumpul dalam dokumen menjadi huruf kecil.
b. *Data Cleaning*
   Data *cleaning* atau pembersihan data merupakan proses menghilangkan data yang tidak digunakan. Pembersihan yang dilakukan disini adalah menghilangkan simbol & newline, RT atau via, @, http, #, dan lainnya.
c. *Tokenizing*
   Tahap berikutnya yaitu melakukan *tokenizing* untuk memisahkan seluruh karakter suatu teks berdasarkan kata yang menyusunnya.
d. *Stopword*
   Stopword adalah menghapus kata umum yang sering muncul dan memiliki arti yang tidak deskriptif (tidak memiliki makna) pada dokumen, maka perlu dihapus.
e. *Stemming*
   Stemming adalah proses mengubah kata berimbuhan menjadi kata dasar sesuai Kamus Besar Bahasa Indonesia (KBBI).

Setelah melalui tahap *pre-processing* langkah selanjutnya adalah pembobotan fitur TF-IDF seperti Gbr. 4 berikut :



Gbr. 4 Skema perhitungan TF-IDF

Pembobotan fitur adalah proses pemberian nilai pada setiap fitur dan sesuai dengan relevansi serta pengaruhnya terhadap hasil klasifikasi [3]. Nilai tersebut digunakan sebagai seleksi fitur berdasarkan minimal bobot yang telah dihitung pada setiap fitur.

Teknik ekstraksi fitur pada salah satu proses pemberian nilai pada setiap kata yang terdapat pada *tweets* latih (data *training*) adalah pembobotan TF-IDF [5]. Pembobotan TF-IDF digunakan sebagai pemberian skor berdasarkan frekuensi kemunculan kata didalam dokumen. Berikut adalah rumus TF-IDF :

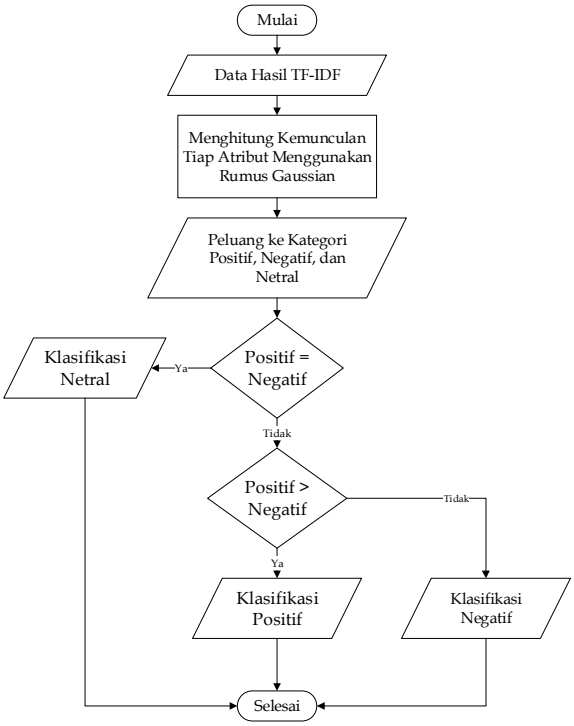
$$TF - IDF = TF \times IDF \qquad (1)$$

dimana,

$$TF = \frac{jumlah\ kemunculan\ term\ pada\ satu\ dokumen}{jumlah\ seluruh\ term\ dalam\ satu\ dokumen} \qquad (2)$$

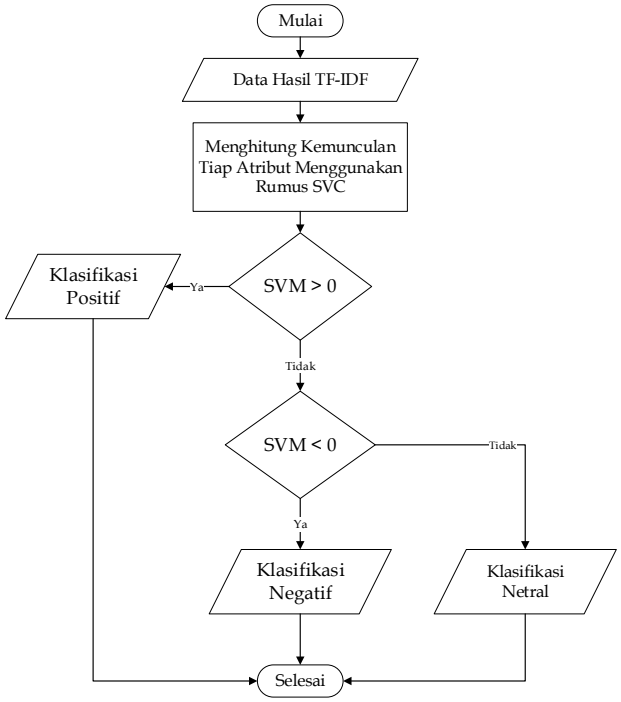
$$IDF = \log \frac{jumlah\ seluruh\ dokumen}{jumlah\ dokumen\ suatu\ term\ muncul} \qquad (3)$$

Setelah didapat hasil pembobotan fitur TF-IDF, dilakukan proses pengklasifikasian data serta menghasilkan nilai akurasi untuk mengetahui kinerja sistem pada pengklasifikasian data *tweet* tentang kurikulum pendidikan. Pengklasifikasian dilakukan menggunakan dua algoritma. Yang pertama adalah algoritma *Naïve Bayes Classifier* seperti Gbr. 5 berikut.



Gbr. 5 *Flowchart* Algoritma *Naïve Bayes Classifier*

Selanjutnya klasifikasi dilakukan menggunakan SVM dimulai dengan mengubah *text* menjadi data vektor. Vektor dalam penelitian ini mempunyai dua komponen yaitu dimensi (*word id*) dan bobot (nilai TF-IDF). Model ruang vektor bertujuan untuk memberikan ID (dimensi) dan sebuah bobot pada setiap kata dalam dokumen berdasarkan seberapa penting keberadaannya dalam dokumen (posisi dokumen dalam dimensi itu). Metode SVM mencoba untuk menemukan garis terbaik dan membagi tiga kelas, kemudian mengklasifikasikan dokumen uji berdasarkan pada sisi mana garis tersebut muncul. Flowchart *Support Vector Machine* ditunjukkan pada Gbr. 6 berikut ini.



Gbr. 6 *Flowchart* Algoritma *Support Vector Machine*

IV. HASIL DAN PEMBAHASAN

Pada penelitian ini klasifikasi sentimen dimasukkan kedalam tiga klasifikasi, yaitu netral, positif, dan negatif. Pengklasifikasian tersebut menggunakan algoritma *Naïve Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM). Proses awal yang dilakukan dalam pengklasifikasian adalah dengan melakukan *crawling* data twitter. Data yang diambil adalah *tweet* mengenai kurikulum pendidikan. Dari *dataset* yang didapatkan, data akan dibagi menjadi dua, yaitu 90% data *training* dan 10% data *testing*.

1) Tampilan Program

Untuk menjalankan sistem klasifikasi ini menggunakan *Command Line Interface* (CLI) seperti Gbr. 7, Gbr. 8, dan Gbr. 9 berikut :

```
Terminal: Local
D:\KULIAH\SKRIPSI\NBC-SVM Nadya>python text_tfidf_nbc_svm.py
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\Chocochasew\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\Chocochasew\AppData\Roaming\nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\Chocochasew\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
Naive Bayes
Test set score: 0.6444444444444445
SVM
Test set score: 0.9555555555555556
```

Gbr. 7 Tampilan hasil *accuracy* NBC dan SVM

```
Terminal: Local
D:\KULIAH\SKRIPSI\NBC-SVM Nadya>python RecallPrecisionNBC.py
Precision: 1.0
Recall: 0.5387096774193548
F-measure: 0.7002096436058701
```

Gbr. 8 Tampilan hasil *recall, precision, dan f-measure* NBC

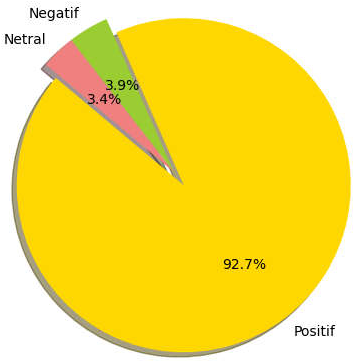
```
Terminal: Local
D:\KULIAH\SKRIPSI\NBC-SVM Nadya>python RecallPrecisionSVM.py
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\Chocochasew\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\Chocochasew\AppData\Roaming\nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\Chocochasew\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
Classifiction Report :
C:\Program Files\Python37\lib\site-packages\sklearn\metrics\_classification.py
.0 in labels with no predicted samples. Use `zero_division` parameter to cont
  _warn_prf(average, modifier, msg_start, len(result))
              precision    recall  f1-score

                   0.96      1.00      0.98
```

Gbr. 9 Tampilan hasil *recall, precision, dan f-measure* SVM

2) Hasil *Crawling*

Dari proses *crawling*, didapat 1000 data *tweet* mengenai kurikulum pendidikan. Kemudian dari *dataset* tersebut akan diberi label untuk mengidentifikasi kelas. Proses *labeling* menghasilkan 92,7% *tweet* positif, 3,9% *tweet* negatif, dan 3,4% *tweet* netral. Perbandingan jumlah klasifikasi sentimen tweet kurikulum pendidikan dapat dilihat pada Gbr. 10 dan contoh hasil sentimen dapat dilihat pada Gbr. 11.



Gbr. 10 Perbandingan hasil sentimen

banyak terkait pendidikan dan kurikulum.	Positif
Selama ini kurikulum belajar di sekolah belum nge-cover itu menurutku.	Netral
Pendidikan formal kan di indonesia ada kurikulum sbg acuan?	Positif
Apa setiap ganti menteri hrs ganti kurikulum ?	Negatif
ya lbih baik kuota sih, biar lebih adil dan sesuai kurikulum masing2 sekolah jg.	Positif
Sedang digodok kurikulum baru untuk bidang ilmu	Positif
Kemendikbud Terbitkan Kurikulum Darurat	Positif

Gbr. 11 Contoh data hasil klasifikasi

Kemudian data yang diperoleh diambil 10% untuk di-*testing* kedalam NBC dan SVM. Terdapat perbedaan antara hasil klasifikasi dari NBC dan SVM seperti terlihat

pada Gbr. 12 dan Gbr. 13, karena klasifikasi NBC menggunakan metode probabilitas dan statistik, sedangkan klasifikasi SVM proses mengubah data ke dimensi lebih tinggi menggunakan pemetaan nonlinear.

```
Naive Bayes
['Negatif' 'Positif' 'Negatif' 'Positif' 'Positif' 'Negatif' 'Negatif'
 'Positif' 'Negatif' 'Positif' 'Negatif' 'Positif' 'Negatif' 'Positif'
 'Positif' 'Positif' 'Positif' 'Positif' 'Negatif' 'Positif' 'Negatif'
 'Positif' 'Negatif' 'Negatif' 'Negatif' 'Positif' 'Negatif' 'Negatif'
 'Negatif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Negatif']
```

Gbr. 12 Contoh hasil klasifikasi *Naïve Bayes Classifier*

```
SVM
['Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif'
 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif'
 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif'
 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif'
 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif' 'Positif']
```

Gbr. 13 Contoh hasil klasifiikasi *Support Vector Machine*

3) *Wordcloud*

Untuk dapat menampilkan kata yang sering muncul pada tweet digunakan *wordcloud*. Kata yang ditampilkan dalam *wordcloud* memiliki beragam ukuran, semakin tinggi frekuensinya maka akan ditampilkan semakin besar. Kelebihan dari *wordcloud* adalah citra visual yang dimunculkan memungkinkan untuk secara cepat menangkap intisari penting dari data yang dianalisis. Gbr. 13 berikut merupakan hasil *wordcloud* dari kategori sentimen positif, Gbr. 14 merupakan hasil *wordcloud* sentimen negatif, dan Gbr. 15 merupakan hasil *wordcloud* sentimen netral.

Gbr. 13 *Wordcloud* sentimen positif

Gbr. 14 *Wordcloud* sentimen negatif

Gbr. 15 *Wordcloud* sentimen netral

4) Hasil Pengujian

Dengan mengolaborasikan algoritma NBC dan SVM dengan fitur TF-IDF menghasilkan nilai *accuracy*, *precision*, *recall*, dan *F-Measure* pada algoritma *Naïve Bayes Classifier* dan *Support Vector Machine* sebagai berikut :

TABEL 1
HASIL PENGUJIAN METODE *NAIVE BAYES CLASSIFIER*

Jumlah Data Uji	*Accuracy*	*Precision*	*Recall*	*F-Measure*
100	64%	100%	54%	70%

TABEL 2
HASIL PENGUJIAN METODE *SUPPORT VECTOR MACHINE*

Jumlah Data Uji	*Accuracy*	*Precision*	*Recall*	*F-Measure*
100	96%	96%	100%	98%

V. KESIMPULAN

Berdasarkan hasil pengujian yang dilakukan, kesimpulan yang diperoleh dari penggunaan algoritma *Naïve Bayes Classifier* dan *Support Vector Machine* untuk klasifikasi opini terhadap tweet pendidikan pendidikan adalah :

1. Data yang diperoleh dari *crawling* twitter mengenai kurikulum pendidikan sebanyak 1000 data. Dengan hasil pengategorian sentimen positif cenderung lebih banyak yaitu sebanyak 92,7%, diposisi terbanyak kedua yaitu sentimen negatif sebanyak 3,9%, sedangkan sentimen netral sebanyak 3,4%.
2. Hasil pengujian data dari kolaborasi TF-IDF dengan *Naïve Bayes Classifier* menghasilkan *accuracy* 64%, *precision* 100%, *recall* 54% dan *F-Measure* 70%, sedangkan dengan *Support Vector Machine* menghasilkan *accuracy* 96%, *precision* 96%, *recall* 100%, dan *F-Measure* 98%. Hal ini menunjukkan bahwa algoritma SVM lebih baik dalam

pengklasifikasian opini *tweet* sistem pendidikan daripada menggunakan algoritma NBC.

VI. SARAN

Data tweet yang diperoleh memiliki kelemahan pada tata bahasa yang digunakan, kebanyakan menggunakan bahasa yang tidak baku. Pada penelitian selanjutnya disarankan untuk menambah tahapan *pre-processing* yang lebih terinci agar mendapatkan hasil sentimen yang lebih baik. Selain itu disarankan untuk penelitian selanjutnya dapat menggunakan metode lain yang lebih baik lagi.

UCAPAN TERIMA KASIH

Puji syukur kepada Allah SWT yang telah mempermudah dalam menyelesaikan penelitian ini. Terima kasih kepada orang tua yang selalu memberikan dukungan dalam menuntut ilmu. Serta terima kasih juga kepada seluruh pihak yang telah mendukung jalannya penelitian ini.

REFERENSI

[1] Littlejohn, dkk. *Teori Komunikasi*. Edisi 9. Jakarta: Salemba Humanika. 2009.

[2] Brogan, C. *Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM)*. Wiley, Manhattan. 2010.

[3] Garnis Berliana, dkk. *Klasifikasi Posting Tweet mengenai Kebijakan Pemerintah Menggunakan Naïve Bayesian Clasification*. 2018.

[4] Dyarsa Singgih Pamungkas & Noor Ageng Setiyanto. *Analisis Sentiment Pada Sosial Media Twitter Menggunakan Naive Bayes Classifier Terhadap Kata Kunci "Kurikulum 2013"*. 2015.

[5] Arsya Monica Pravina, Imam Cholissodin, dan Putra Pandhu Adikara. *Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM)*. 2019.

[6] Naim Rochmawati & S C Wibawa. *Opinion Analysis on Rohingya using Twitter Data*. 2018.

[7] Veny Amilia Fitri, dkk. *Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm* 2019.

[8] Potong Fiarni, dkk. *Sentiment Analysis System for Indonesia Online Retail Shop Review Using Hierarchy Naive Bayes Technique*. 2016.

[9] Dhina Nur Fitriana & Yuliant Sibaroni. *Sentiment Analysis on KAI Twitter Post Using Multiclass Support Vector Machine (SVM)*. 2019.