

# ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PELAKSANAAN IMPLEMENTASI KURIKULUM MERDEKA

*Faizal Riza*

*Program Studi Teknik Informatika, Institut Teknologi Budi Utomo Jakarta,  
faizalriza@itbu.ac.id*

## **Abstrak**

Undang Undang Nomor 20 tahun 2003 menyatakan bahwa kurikulum adalah seperangkat rencana pembelajaran yang berkaitan dengan tujuan, isi, bahan ajar dan cara yang digunakan dan dijadikan sebagai pedoman dalam penyelenggara kegiatan pembelajaran untuk mencapai sebuah tujuan pendidikan nasional. Pandemi Covid-19 yang hampir berlangsung selama 2 tahun mengakibatkan ketertinggalan capaian pembelajaran (learning loss) untuk peserta didik. Kurikulum Merdeka yang dirancang pemerintah Republik Indonesia adalah salah satu alat dalam memulihkan learning loss. Tujuan dari penelitian ini adalah untuk mendapatkan algoritma terbaik dalam mengklasifikasikan tanggapan sentimen positif dan negatif dari berbagai elemen pendidikan khususnya warga sekolah. Media sosial twitter dipilih sebagai media untuk mendapatkan opini masyarakat tentang implementasi kurikulum merdeka. Beragam komentar dikumpulkan dan diklasifikasikan menjadi sebuah dataset untuk menilai komentar sentiment mengenai implementasi kurikulum merdeka apakah positif atau negatif. Analisis sentimen diproses dengan menggunakan tools rapidminer dan menggunakan komparasi algoritma naïve bayes, support vector machine serta k-nearest neighbor. Hasil analisis menunjukkan bahwa algoritma support vector machine dengan sampel bootstrapping memiliki akurasi lebih unggul sebesar 90,31% untuk sentiment analisis pada topik implementasi kurikulum merdeka.

Kata Kunci : kurikulum merdeka, k-nearest neighbor, naïve bayes, sampel bootstrapping, sentimen analisis, support vector machine.

## **1. PENDAHULUAN**

Antisipasi dampak pandemi terhadap ketertinggalan pembelajaran (learning loss) dan kesenjangan pembelajaran (learning gap) telah dilakukan oleh Kementerian Pendidikan dan Kebudayaan Republik Indonesia melalui penetapan Kurikulum Merdeka menggantikan Kurikulum 2013. Kajian akademik pemulihan pembelajaran ini disusun untuk menelaah berbagai alternatif kurikulum yang dapat digunakan oleh satuan pendidikan dengan keragaman karakteristiknya untuk meningkatkan kualitas proses pembelajaran, mengoptimalkan hasil belajar siswa, serta mengurangi dampak-dampak negatif pandemi COVID-19 bagi pendidikan di Indonesia (Badan Standar Kurikulum dan Asesemen Pendidikan, 2021).

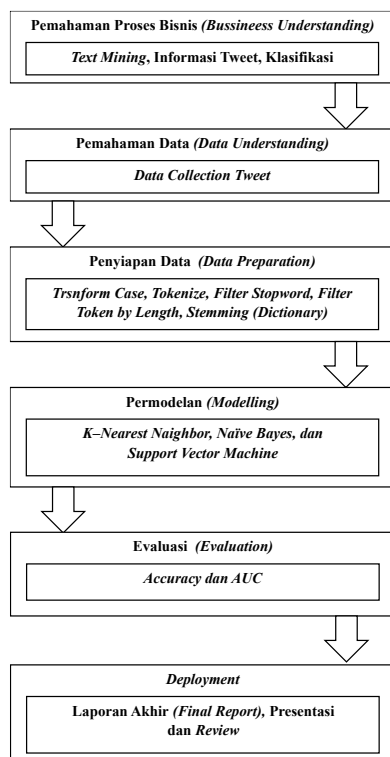
Penelitian terdahulu mengenai analisis sentiment pada sosial media twitter menggunakan naïve bayes classifier terhadap kata kunci "kurikulum 2013" (Pamungkas et al., 2015) menunjukkan hasil penelitian ini menghasilkan sebuah sistem yang dapat mengklasifikasi sentimen secara otomatis dengan hasil pengujian 3000 data latih dan 100 tweet data ujicoba mencapai 91 %. Sedangkan (Zaman, 2020) dalam penelitiannya tentang klasifikasi opini terhadap

kebijakan publik merdeka belajar pada jejaring sosial twitter menggunakan metode naïve bayes dengan seleksi fitur information gain mendapatkan hasil bahwa atribut dalam Naïve Bayes bersifat independen sehingga setiap atribut tidak terikat satu sama lain. Seleksi fitur Information Gain pada Naïve Bayes diharapkan dapat meningkatkan akurasi dengan cara mengurangi fitur atau atribut yang tidak relevan. Berdasarkan skenario uji coba yang telah dilakukan, diperoleh hasil terbaik menggunakan jumlah fitur sebesar 40% dengan nilai akurasi 81.48%, presisi 89.74%, recall 76.09% dan f-measure 82.35%. Penelitian lainnya yang berkaitan dengan analisis sentimen penerapan implementasi kurikulum merdeka adalah penelitian (Kholila, 2021) tentang analisis sentimen terhadap program merdeka belajar – kampus merdeka pada twitter menggunakan support vector machine mendapatkan hasil nilai akurasi F-measure adalah kelas positif sebesar 94.8% dan kelas negatif sebesar 95%. Oleh karena itu maka penelitian ini berusaha mendapatkan komparasi akurasi algoritma naïve bayes, support vector machine dan k-nearest neighbor, untuk sentiment analisis pada topik implementasi kurikulum merdeka.

## 2. METODOLOGI

Penelitian ini menggunakan data dari Twitter berupa tweet dari pengguna aplikasi Twitter dengan topik pembahasan implementasi kurikulum merdeka. Data yang telah di-crawl sebanyak 1.611 dari tanggal 01 Januari 2022 sampai dengan 28 November 2022. Alat penelitian ini memanfaatkan pemrograman twint python untuk data crawling twitter. Selain itu, pada model pada proses pelabelan, peneliti mengkategorikan sentimen positif dan negatif, sedangkan metode klasifikasi yang digunakan adalah naïve bayes, support vector machine dan k-nearest neighbor.

Metodologi penelitian yang digunakan dalam penelitian eksperimen ini dengan menggunakan metode Cross-Industry Standard Process for Data Mining (CRISP-DM) terdiri dari enam tahap yaitu Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment (Riza et al., 2020). Model penelitian ditunjukkan pada gambar 1.



Gambar 1. Kerangka Penelitian  
Sumber : Hasil Olahan Penelitian

- a. Pemahaman Proses Bisnis (*Business Understanding*)  
Setiap proyek penambangan data dimulai dengan menentukan tujuan proyek, termasuk fase pertama, pemahaman bisnis. Bisnis ini bertujuan untuk memaksimalkan waktu kerja dan efisiensi mesin dengan menggunakan analitik prediktif. Target ini kemudian

diubah menjadi data mining dengan mengidentifikasi komponen mesin yang relevan (Riza et al., 2020).

- b. Pemahaman Data (*Data Understanding*)  
Tujuan proyek data mining dibentuk berdasarkan pengalaman dan asumsi yang mumpuni. Pada fase Data Understanding, informasi tentang skenario perawatan prediktif disembunyikan untuk mendeteksi kesalahan, konsep yang valid untuk mencari pola frekuensi baru dalam aliran data gerakan sensor (Kusrini & Taufiq Emha, 2009).
- c. Penyiapan Data (*Data Preparation*)  
Pada tahap Data Preparation, peneliti mengumpulkan data yang relevan dan menyiapkan data mining yang menggunakan preprocessing, seperti reduksi data, filtering, dan pembuatan fitur yang berkaitan dengan tujuan proyek data mining (Kusrini & Taufiq Emha, 2009)
- d. Permodelan (*Modelling*)  
Pada fase Permodelan data mining, alur kerja dibangun untuk menemukan pengaturan parameter yang diinginkan dan algoritma yang dipilih untuk dieksekusi. Tugas data mining adalah pada data yang telah diproses sebelumnya (Amra & Maghari, 2017).
- e. Evaluasi (*Evaluation*)  
Pada fase Evaluasi, menguji model terhadap kumpulan data nyata dalam skenario produksi dan menilai hasil penambangan data berdasarkan tujuan bisnis. Untuk tujuan ini, hasilkan kumpulan data uji mengikuti langkah-langkah yang dikembangkan dalam fase "Persiapan Data" dan "Permodelan", tidak termasuk langkah pelabelan (Zy, 2017).
- f. Implementasi (*Deployment*)  
Setelah evaluasi berhasil, gunakan model pelatihan dalam produksi pada fase Penerapan. Penyebarannya membutuhkan pengaturan yang stabil untuk akuisisi data, termasuk infrastruktur pemrosesan data (Bustami, 2014).

## 3. ANALISIS DAN PEMBAHASAN

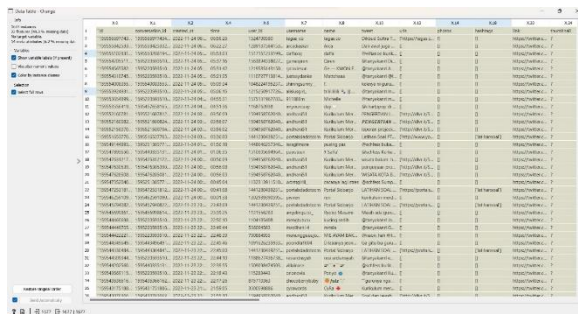
Berikut adalah tahapan tahapan penelitian yang dilakukan:

- a. Pemahaman Proses Bisnis (*Business Understanding*)  
Pada tahap ini dilakukan pemahaman terhadap objek penelitian dengan cara mencari informasi melalui media sosial Twitter tentang Kurikulum Merdeka, mengungkapkan berbagai macam pendapat, baik negatif maupun positif, pada tweet pengguna media sosial. Implementasi pemahaman bisnis membantu menentukan pendekatan analisis sentimen terbaik dan

model yang cocok berdasarkan perbandingan hasil algoritma. Algoritma yang digunakan adalah naïve bayes, support vector machine dan k-nearest neighbor.

**b. Pemahaman Data (Data Understanding)**

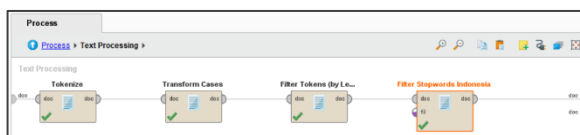
Tahap *Data Understanding* memiliki tujuan untuk mengumpulkan, mengidentifikasi, dan memahami data yang dimiliki. Data juga harus dapat diverifikasi. Data yang digunakan dalam penelitian ini adalah data dari ulasan pengguna Twitter dalam menanggapi kata kunci kurikulum merdeka, yang terdiri dari kategori positif dan negatif. Seperti yang ditunjukkan pada Gambar 2.



Gambar 1. Tweet tentang Kurikulum Merdeka  
Sumber : Hasil Olahan Penelitian

**c. Penyiapan Data (Data Preparation)**

Tahap Persiapan Data diperoleh dari crawling data menggunakan programing twint phyton untuk twitter crawling data dari review pada aplikasi twitter kata kunci implementasi kurikulum merdeka , terdiri dari kategori positif dan negatif. Pembersihan data untuk mengurangi duplikasi dan redundan data, lalu transform case, transform remove url, tokenizing, @anotation removal, stopword dan removal seperti pada gambar 3 .



Gambar 2. Proses *Data Preparation*  
Sumber : Hasil Olahan Penelitian

**1) Transform Case**

Operator yang digunakan pada tahapan ini adalah untuk mengubah huruf kapital yang masih ada pada text akan diubah menjadi huruf kecil semua. Hal ini dilakukan agar ketika dilakukan proses ke dalam model klasifikasi terdapat keseragaman huruf dan

tidak terjadi kesalahan dalam proses tokenize.

Tabel 1. *Transform Case*

Data Sebelum	Data Sesudah
@tanyakanr1 MALES MAH MALES AJA ANJ LAGAIN KURIKULUM MERDEKA kg bgtu konsepnya dah prasaan	@tanyakanr1 males mah males aja anj lagain kurikulum merdeka kg bgtu konsepnya dah prasaan
@charlppop di sekolah yang aku kerjain malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir banget biar anak2 ga ketinggalan pelajaran gimana bisaaaa itu guru leha2ðŸŸ'	@charlppop di sekolah yang aku kerjain malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir banget biar anak2 ga ketinggalan pelajaran gimana bisaaaa itu guru leha2ðŸŸ'
https://t.co/K9PnzXDWQ	https://t.co/k9pnzxdwq

Sumber Data : Hasil Olahan Data Penelitian

**2) Transformation remove url**

Transformation Remove URL, dalam proses ini link atau URL yang terkandung pada tweet akan dihilangkan. Hal ini bertujuan untuk menjadikan kata atau komentar terseleksi hanya tweetnya saja.

Table 2. Transformation remove url

Data Sebelum	Data Sesudah
@charlppop di sekolah yang aku kerja malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir banget biar anak2 ga ketinggalan pelajaran gimana bisaaaa itu guru leha2ðŸŸ'	@charlppop di sekolah yang aku kerja malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir banget biar anak2 ga ketinggalan pelajaran gimana bisaaaa itu guru leha2ðŸŸ'
https://t.co/K9PnzXDWQ	

Sumber Data : Hasil Olahan Data Penelitian

**3) Tokenization**

Kemudian hasil dari proses *Transformation Remove URL* dilanjutkan oleh proses *Tokenization* (Regexp) yaitu semua kata yang ada didalam tiap dokumen dikumpulkan dan dihilangkan tanda baca, angka, simbol, karakter khusus atau apapun yang bukan huruf.

Table 3. *Tokenization* (Regexp)

Data Sebelum	Data Sesudah
@charlppop di sekolah yang aku kerja malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir banget biar anak2 ga ketinggalan	@charlppop di sekolah yang aku kerja malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir

pelajaran gimana bisaaaa itu guru leha2ðŸŸ’ https://t.co/K9PnzXDWQ	banget biar anak2 ga ketinggalan pelajaran gimana bisaaaa itu guru leha
---	---

Sumber Data : Hasil Olahan Data Penelitian

4) @anotation removal

Teks diurai berdasarkan white space. Dalam proses ini, semua anotasi (@) yang terkandung dalam tweet dihilangkan dan mengubah seluruh huruf kapital menjadi huruh kecil. Tujuannya adalah karena annotation (@) biasanya merujuk pada yang melakukan tweet/komentar.

Table 4. @Anotation removal

Data Sebelum	Data Sesudah
@charlppop di sekolah yang aku kerja malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir banget biar anak2 ga ketinggalan pelajaran gimana bisaaaa itu guru leha2ðŸŸ’ https://t.co/K9PnzXDWQ	charlppop di sekolah yang aku kerja malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir banget biar anak2 ga ketinggalan pelajaran gimana bisaaaa itu guru leha

Sumber Data : Hasil Olahan Data Penelitian

5) Filter stopword removal

Selanjutnya adalah penggunaan operator *Stopword Removal (by Dictionary)* yang berfungsi untuk menghilangkan kata-kata yang tidak hubungan dengan isi text. Maka dengan operator *Stopword Removal (by Dictionary)* peneliti dapat mendaftarkan kata yang harusnya dihapus dari text.

Table 5. Stopword removal

Data Before	Data After
@charlppop di sekolah yang aku kerja malah padet banget sampe mau ada kegiatan selain berhubungan sama mapel dan project kurikulum merdeka aja kita mikir banget biar anak2 ga ketinggalan pelajaran gimana bisaaaa itu guru leha2ðŸŸ’ https://t.co/K9PnzXDWQ	sekolah kerja malah padet banget sampe kegiatan selain berhubungan mapel project kurikulum merdeka mikir banget anak2 ketinggalan pelajaran gimana bisaaaa guru leha

Sumber Data : Hasil Olahan Data Penelitian

d. Modelling

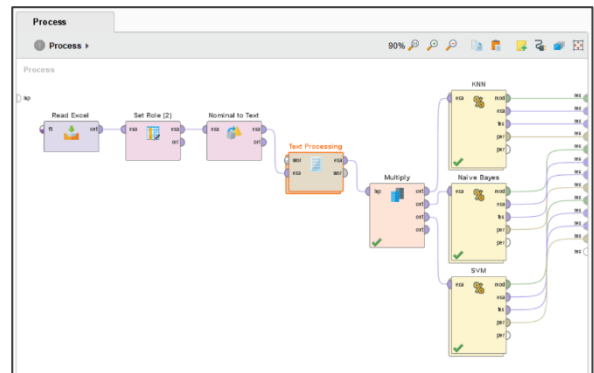
Merupakan tahap pemilihan teknik mining dengan menentukan algoritma yang akan digunakan. Tool yang digunakan adalah RapidMiner versi 9.10. Hasil pengujian model yang dilakukan adalah mengklasifikasikan dengan algoritma naïve bayes, knn, dan svm untuk mendapatkan nilai akurasi terbaik.

Pengaturan dan penggunaan operator serta parameter dalam frameworks Rapid Miner sangat berpengaruh terhadap akurasi dan model yang terbentuk, sebagai pembanding nilai akurasi maka model dibagi dua yaitu model yaitu model dengan menggunakan sample bootstrapping dan model tanpa menggunakan sampel bootstrapping.

Bootstrapping adalah prosedur statistik dengan cara mengubah data dari sampel yang kita peroleh dan melakukan replikasi dari data sampel tersebut (resampling) secara acak untuk diperoleh data simulasi baru. Prosedur ini dapat digunakan untuk menghitung standar error, interval kepercayaan, dan melakukan pengujian hipotesis untuk berbagai jenis analisis statistik (Sain & Vapnik, 2016).

1. Pengujian model algoritma Naïve Bayes, KNN, dan SVM tanpa *sample bootstrapping*

Model pengujian dari algoritma Naïve Bayes, SVM dan KNN menggunakan rapidminer tanpa menggunakan *sample bootstrapping*, diawali dari memasukan data kemudian mengatur *set role* yang nantinya menentukan label disana dan nominal text lalu ke proses dokumen. Pengujian model tanpa sampel bootstrapping dapat dilihat pada gambar 4.

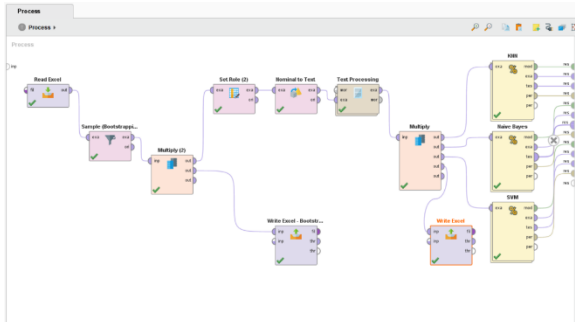


Gambar 3. Desain. Model Algoritma Naïve Bayes, SVM dan KNN tanpa *sample bootstrapping*  
Sumber : Hasil Olahan Penelitian

2. Pengujian model algoritma Naïve Bayes, KNN, dan SVM dengan sampel bootstrapping

Model pengujian dari algoritma Naïve Bayes, SVM dan KNN menggunakan rapidminer

dengan menggunakan sample bootstrapping dapat dilihat pada figure 5.



Gambar 5. Desain. Model Algoritma Naïve Bayes, SVM dan KNN dengan *sample bootstrapping*  
 Sumber : Hasil Olahan Penelitian

**e. Evaluation**

Tahapan evaluasi bertujuan untuk menentukan nilai kegunaan dari model yang telah berhasil dibuat pada langkah sebelumnya. Untuk evaluasi digunakan *10-fold cross validation*. Dari hasil pengujian model dari tiga algoritma yang dipakai adalah untuk menghasilkan sebuah nilai Accuracy (*Confusion Matrix*) dan *Area Under Curve* (AUC). Maka mendapatkan hasil grafik ROC dengan nilai *Area Under Curve* (AUC).

Berdasarkan hasil analisis dari algoritma *Naïve Bayes*, *SVM* dan *KNN*, maka dapat dirangkum hasilnya seperti tabel 6.

Table 6. Perbandingan *Algorithm Performance*

Metode	Accuracy	AUC
SVM	82,00%	0,796
SVM + Bootstrapping	90,31%	0,863
Naif Bayes	82,17%	0,624
Naïve Bayes + Bootstrapping	88,57%	0,670
KNN	71,67%	0,740
KNN + Bootstrap	74,75%	0,842

Sumber : Hasil Olahan Penelitian

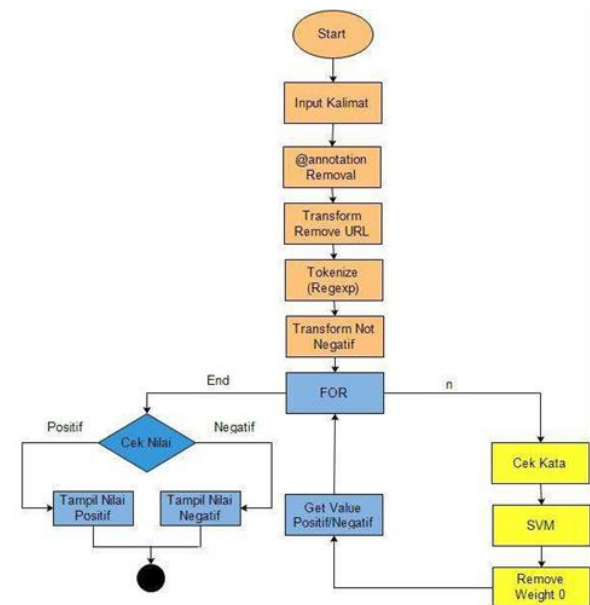
Dari hasil perbandingan performance ketiga algoritma diatas dengan menggunakan sample bootstrapping dan tanpa menggunakan bootstrapping terdapat perbedaan nilai akurasi dan AUC yang sangat signifikan. Pada model tanpa menggunakan sample bootstrapping algoritma Naïve Bayes memiliki nilai akurasi yang paling tinggi 82,17% mengungguli algoritma lainnya

tetapi memiliki nilai AUC paling kecil sebesar 0,624% sedangkan hasil pengujian dengan menggunakan sample bootstrapping algoritma SVM nilai akurasinya paling tinggi dibandingkan dengan algoritma Naïve Bayes dan Algoritma KNN dengan nilai akurasi 90,31% dan nilai AUC paling tinggi juga sebesar 0,863% dengan kategori excellent classification. Berdasarkan hasil pengujian dapat disimpulkan bahwa model dengan menggunakan parameter sample bootstrapping dapat menaikkan nilai akurasi dan akurasi tertinggi dihasilkan dengan menggunakan metode sample bootstrapping pada saat preprocessing data sebelum diolah menggunakan algoritma SVM. Maka dari itu penelitian yang digunakan pada dataset implementasi kurikulum merdeka ini menggunakan algoritma SVM dengan sample bootstrapping.

**f. Deployment**

Pada tahap selanjutnya adalah implementasi untuk mengukur sentiment pada tweet yang telah diperoleh berdasarkan hasil penelitiannya. Tujuannya untuk menguji aplikasi apakah prediksi dan actual nya sama terhadap data.

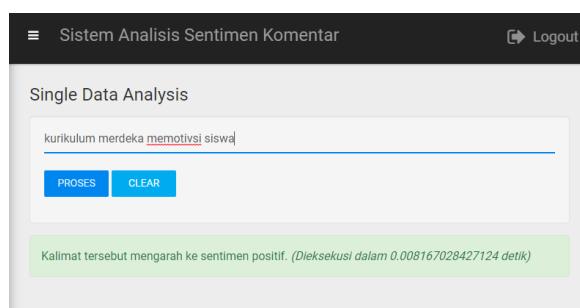
Sebelum melakukan deployment langkah pertama yang dilakukan adalah membuat flowchartnya terlebih dahulu untuk mengetahui alur dari pembuatan program untuk menilai sentiment terhadap tweet.



Gambar 4. Flowchart Kerja Sentiment Analysis  
 Sumber : Hasil Olahan Penelitian

Selanjutnya membuat interface halaman input design. Pada desain ini sentiment diinputkan datanya untuk menilai apakah sentiment yang

dimasukan atau tweet yang dituliskan dari pengguna twitter untuk menilai implementasi kurikulum merdeka akan Positif atau Negatif. Pilih tweet untuk menilai dan melihat bobot dari masing-masing kata sehingga hasilnya bisa dilihat pada hasil dibawahnya seperti ditunjukkan pada gambar 7.



Gambar 5. Tampilan Beranda Aplikasi Analisis Sentiment Kurikulum Merdeka  
Sumber : Hasil Olahan Penelitian

#### 4. KESIMPULAN

Dalam penelitian ini setelah dilakukan preprocessing dan dilakukan pengujian model dengan membandingkan tiga metode data mining yaitu *Naive Bayes*, *Support Vector Machine* dan *K-Nearest Neighbor*, hasil evaluasi dan validasi, diketahui bahwa nilai akurasi untuk menentukan bahwa komentar di twitter untuk sentiment analisis implementasi kurikulum merdeka dengan kategori komentar positif dan negatif, dapat dibuktikan dengan nilai akurasi dan nilai AUC dari masing- masing algoritma dengan menambahkan parameter *sample bootstrapping*. Dari hasil perbandingan performance ketiga algoritma diatas dengan menggunakan *sample bootstrapping* dan tanpa menggunakan *bootstrapping* terdapat perbedaan nilai akurasi dan AUC yang sangat signifikan. Pada model tanpa menggunakan *sample bootstrapping* algoritma *Naive Bayes* memiliki nilai akurasi yang paling tinggi 82,17% mengungguli algoritma lainnya tetapi memiliki nilai AUC paling kecil sebesar 0,624% sedangkan hasil pengujian dengan menggunakan *sample bootstrapping* algoritma SVM nilai akurasinya paling tinggi dibandingkan dengan algoritma *Naive Bayes* dan Algoritma KNN dengan nilai akurasi 90,31% dan nilai AUC paling tinggi juga sebesar 0,863% dengan kategori *excellent classification*. Berdasarkan hasil pengujian dapat disimpulkan bahwa model dengan menggunakan parameter *sample bootstrapping* dapat menaikkan nilai akurasi dan akurasi tertinggi dihasilkan dengan menggunakan metode *sample bootstrapping* pada saat preprosesing data sebelum diolah menggunakan algoritma SVM. Maka dari itu penelitian yang digunakan pada dataset

implementasi kurikulum merdeka ini menggunakan algoritma SVM dengan *sample bootstrapping*. Pada penelitian ini dapat diketahui bahwa tingkat akurasi yang didapatkan algoritma *Support Vector Machine* lebih unggul dibanding dengan *K-Nearest Neighbor* dan *Naive Bayes*. Untuk itu, penerapan *Support Vector Machine* dapat memberikan solusi terhadap permasalahan pada sentiment analisis implementasi kurikulum merdeka.

Sentimen analisis dengan menggunakan *sample bootstrapping* dan tanpa menggunakan *sample bootstrapping* memiliki perbandingan nilai akurasi dan AUC yang signifikan. Maka dari itu metode algoritma dengan menggunakan *sample bootstrapping* cocok digunakan pada penelitian ini. Accuracy dan AUC terbaik terhadap sentiment analisis diperoleh dengan algoritma SVM dengan penggunaan *sample bootstrapping* dengan nilai akurasi sebesar 90,31% dan AUC 0,863% dan kategori *excellent classification*.

#### DAFTAR PUSTAKA

- Amra, I. A. A., & Maghari, A. Y. A. (2017). Students performance prediction using KNN and Naive Bayesian. *ICIT 2017 - 8th International Conference on Information Technology, Proceedings*. <https://doi.org/10.1109/ICITECH.2017.8079967>
- Badan Standar Kurikulum dan Asesemen Pendidikan. (2021). Kajian Akademik: Kurikulum untuk Pemulihan Pembelajaran. In *Repository Kemdikbud*. Pusat Kurikulum dan Pembelajaran.
- Bustami. (2014). Penerapan Algoritma Naive Bayes. *Jurnal Informatika*.
- Kholila, N. (2021). ANALISIS SENTIMEN TERHADAP PROGRAM MERDEKA BELAJAR –KAMPUS MERDEKA PADA TWITTER MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM). *ANTIVIRUS: Jurnal Ilmiah Teknik Informatika*, 15(2), 1–12. <https://doi.org/10.35457/ANTIVIRUS.V15I2.1866>
- Kusrini, & Taufiq Emha, L. (2009). Algoritma Data Mining Yogyakarta. In *Algoritma Data Mining* (Issue February). Andi. <https://books.google.co.id/books?id=-Ojclag73O8C&printsec=frontcover#v=onepage&q&f=false>
- Pamungkas, D. S., Setiyanto, N. A., & Dolphina, E. (2015). Analisis Sentiment Pada Sosial Media Twitter Menggunakan Naive Bayes Classifier Terhadap Kata Kunci “Kurikulum

- 2013''." *Universitas Dian Nuswantoro*, 14(4), 299–314.
- Riza, F., Rifai, S., Dirgantara, A., Sfenrianto, Rasenda, & Herdyansyah, S. (2020). Information Retrieval Technique for Indonesian PDF Document with Modified Stemming Porter Method Using PHP. *Journal of Physics: Conference Series*, 1477(3). <https://doi.org/10.1088/1742-6596/1477/3/032016>
- Sain, S. R., & Vapnik, V. N. (2016). The Nature of Statistical Learning Theory. *Technometrics*. <https://doi.org/10.2307/1271324>
- Zaman, M. R. T. (2020). *Klasifikasi opini terhadap kebijakan publik Merdeka Belajar pada jejaring sosial Twitter menggunakan metode naïve bayes dengan seleksi fitur information gain*.
- Zy, A. T. (2017). Comparison Algorithm Classification Naive Bayes, Decission Tree and Neural Network for Analysis Sentiment. *Jurnal Pelita Teknologi*.