Jurnal Indonesia : Manajemen Informatika dan Komunikasi
https://journal.stmiki.ac.id

E - ISSN: 2723-7079 | P - ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September

OPEN ACCESS

https://doi.org/10.35870/jimik.v5i3.990

Analisis Sentimen Tanggapan Pengguna Media Sosial X Terhadap Program Beasiswa KIP-Kuliah dengan Menggunakan Algoritma *Support Vector Machine* (SVM)

Ika Amelia [1], Sugiyono [2*], Frencis Matheos Sarimole [3], Tundo [4]

[1] Program Studi Sistem Informasi, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.
[2*,3,4] Program Studi Teknik Informatika, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.

Email: ikaamelia.ac@gmail.com [1], inosoguy007@gmail.com [2*]

Histori Artikel:
Dikirim 23 Juli 2024; *Diterima dalam bentuk revisi* 10 Agustus 2024; *Diterima* 20 Agustus 2024; *Diterbitkan* 20 September 2024.

Abstrak

Beasiswa KIP-Kuliah merupakan program pemerintah Indonesia yang bertujuan memberikan akses pendidikan tinggi bagi mahasiswa dari keluarga kurang mampu. Program ini menjadi topik diskusi hangat di media sosial, termasuk media sosial X. Penelitian ini bertujuan menganalisis sentimen masyarakat pengguna media sosial X terhadap pemberian Beasiswa KIP-Kuliah menggunakan algoritma Support Vector Machine. Obyek penelitian adalah komentar di media sosial X terkait Beasiswa KIP-Kuliah. Metode penelitian meliputi pengumpulan data dengan google collabs, preprocessing data, pelatihan model Support Vector Machine, dan evaluasi model menggunakan RapidMiner. Hasil penelitian menunjukkan bahwa model Support Vector Machine mampu mengklasifikasikan sentimen dengan akurasi 86.27%, namun terdapat bias terhadap sentimen negatif. Mayoritas tanggapan masyarakat bersifat negatif, seringkali terkait penyalahgunaan beasiswa. Saran yang diberikan meliputi pengumpulan data yang lebih seimbang, penggunaan teknik penyeimbangan dataset, penerapan model yang lebih kompleks, serta evaluasi lebih mendalam untuk meningkatkan performa model. Penelitian ini diharapkan menjadi masukan bagi pemerintah dalam meningkatkan mekanisme penyaluran Beasiswa KIP-Kuliah agar lebih tepat sasaran dan mengurangi potensi penyalahgunaan.

Kata Kunci: Beasiswa KIP-Kuliah; Analisis Sentimen; Media Sosial X; Support Vector Machine.

Abstract

The KIP-Kuliah Scholarship is an Indonesian government program which aims to provide access to higher education for students from underprivileged families. This program has become a hot topic of discussion on social media, including social media. The object of research is comments on X's social media regarding the KIP-College Scholarship. Research methods include crawling data using google collabs, data preprocessing, Support Vector Machine model training, and model evaluation using RapidMiner. The research results show that the Support Vector Machine model is able to classify sentiment with an accuracy of 86.27%, but there is a bias towards negative sentiment. The majority of public responses are negative, often regarding misuse of scholarships. The suggestions given include collecting more balanced data, using dataset balancing techniques, implementing more complex models, and more in-depth evaluation to improve model performance. It is hoped that this research will provide input for the government in improving the KIP-College Scholarship distribution mechanism so that it is more targeted and reduces the potential for abuse.

Keyword: KIP-College Scholarship; Sentiment Analysis; Social Media X; Support Vector Machine.

Jurnal Indonesia : Manajemen Informatika dan Komunikasi
https://journal.stmiki.ac.id
E-ISSN: 2723-7079 | P-ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990

1. Pendahuluan

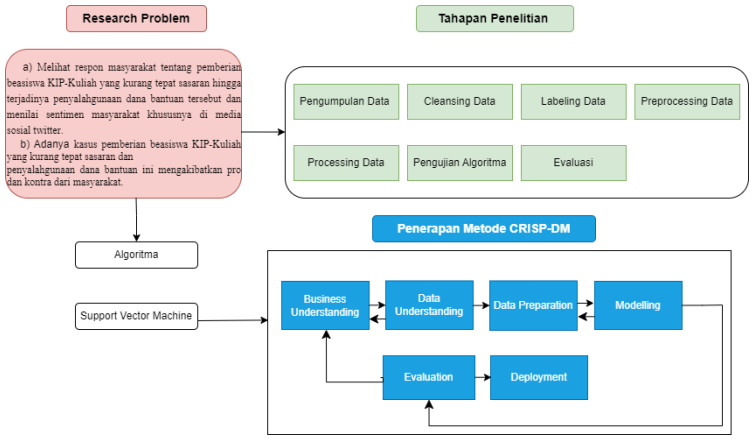
Program KIP-Kuliah merupakan salah satu upaya pemerintah Indonesia untuk meningkatkan akses pendidikan tinggi bagi mahasiswa yang berasal dari keluarga dengan keterbatasan ekonomi. Program ini bertujuan untuk mengurangi kesenjangan akses pendidikan dan mendorong inklusi sosial di kalangan mahasiswa dari latar belakang ekonomi rendah. Kriteria penerima beasiswa mencakup prestasi akademik, penghasilan orang tua, jumlah tanggungan keluarga, luas tanah, serta status sebagai penerima bantuan sosial seperti Kartu Indonesia Pintar (KIP) atau Kartu Keluarga Sejahtera (KKS). Meskipun dirancang dengan tujuan baik, proses seleksi penerima sering kali tidak optimal, sehingga banyak mahasiswa yang sebenarnya lebih membutuhkan justru tidak mendapatkan bantuan yang mereka perlukan (Yuliana *et al.*, 2022). Dalam era digital, media sosial menjadi platform utama bagi masyarakat untuk menyuarakan pendapat mengenai berbagai isu, termasuk kebijakan pemerintah. Media sosial X, misalnya, memiliki 206 juta pengguna aktif di seluruh dunia, dengan 18,4 juta pengguna di Indonesia, menjadikannya salah satu platform terbesar di mana diskusi mengenai program KIP-Kuliah berlangsung (Annur, 2024). Media ini digunakan oleh masyarakat untuk berkomunikasi melalui teks, foto, maupun video, dan sering kali menjadi wadah untuk mengekspresikan ketidakpuasan atau kritik. Beberapa unggahan di media sosial X menunjukkan adanya penyalahgunaan dana beasiswa KIP-Kuliah oleh para penerimanya, seperti gaya hidup mewah yang dipamerkan oleh mahasiswa penerima beasiswa. Unggahan tersebut memicu protes publik yang mempertanyakan efektivitas dan mekanisme penyaluran beasiswa yang dianggap kurang tepat sasaran. Contohnya, beberapa kasus melibatkan mahasiswa yang memamerkan barang-barang mewah, yang mengundang kecaman karena dianggap tidak menghargai kesempatan yang diberikan oleh pemerintah (Larasati *et al.*, 2023).

Penyalahgunaan beasiswa KIP-Kuliah bukanlah masalah yang bersifat insidental. Berbagai platform media sosial sering menampilkan unggahan serupa yang menunjukkan pola yang sama, yaitu penggunaan dana beasiswa untuk kebutuhan konsumtif yang tidak sesuai dengan tujuan program. Hal ini menimbulkan kekhawatiran publik terhadap efektivitas program KIP-Kuliah dalam mencapai tujuannya, yakni membantu mahasiswa yang benar-benar membutuhkan akses pendidikan tinggi. Banyak pihak yang mempertanyakan sistem seleksi dan distribusi beasiswa yang dinilai kurang tepat, sehingga menyebabkan ketimpangan dan ketidakadilan dalam akses terhadap bantuan pendidikan. Akibatnya, mahasiswa yang seharusnya menerima bantuan malah tidak mendapatkannya, sementara penerima yang tidak memenuhi kriteria justru menikmati manfaat beasiswa tanpa kebutuhan mendesak (Witanti *et al.*, 2023). Untuk memahami sentimen masyarakat terhadap program KIP-Kuliah, diperlukan metode analisis yang tepat dan dapat diandalkan. Support Vector Machine (*SVM*) merupakan salah satu teknik klasifikasi yang digunakan dalam penelitian ini untuk mengkategorikan sentimen publik terhadap pemberian beasiswa KIP-Kuliah. *SVM* banyak digunakan dalam klasifikasi data teks karena kemampuannya dalam menangani data berdimensi tinggi dan memisahkan kelas data dengan margin optimal (Pamungkas, 2020). Penelitian ini bertujuan untuk mengklasifikasi sentimen positif dan negatif dari komentar publik mengenai program beasiswa KIP-Kuliah di media sosial X, menggunakan dataset yang terdiri dari 1.148 data komentar yang diambil pada periode 29 April – 1 Mei 2023. Hasil dari penelitian ini diharapkan dapat memberikan gambaran mengenai sentimen publik terhadap program ini serta memberikan masukan kepada pemerintah untuk memperbaiki mekanisme penyaluran beasiswa agar lebih tepat sasaran dan mengurangi potensi penyalahgunaan di masa depan.

Jurnal Indonesia : Manajemen Informatika dan Komunikasi
https://journal.stmiki.ac.id
E-ISSN: 2723-7079 | P-ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990

2. Metode Penelitian

Pada penelitian memiliki beberapa tahapan alur metode penelitian sebagai berikut:



Gambar 1. Metode Penelitian

Metode penelitian ini mencakup beberapa tahapan yang dirancang untuk mencapai tujuan utama, yaitu menganalisis sentimen publik terhadap pemberian beasiswa KIP-Kuliah menggunakan algoritma *Support Vector Machine* (SVM). Tahap pertama adalah pengumpulan data yang dilakukan melalui teknik *crawling* menggunakan Google Colab dengan memanfaatkan *Twitter Auth Token*. Dalam proses ini, tweet yang diambil merupakan tweet berbahasa Indonesia dengan kata kunci "KIPK" dan batasan hingga 1.000 tweet. Hasil dari proses ini adalah terkumpulnya 1.148 tweet yang kemudian dijadikan sebagai sampel dalam penelitian ini. Pemilihan Google Colab sebagai platform pengumpulan data didasarkan pada kemampuannya dalam memproses data dalam skala besar secara efisien melalui komputasi berbasis *cloud*. Tahap berikutnya adalah *preprocessing* data, yang bertujuan untuk mengubah data mentah menjadi bentuk yang siap diolah lebih lanjut. Proses ini melibatkan beberapa langkah, yaitu *tokenizing*, *case folding*, *filter stopwords*, dan *stemming*. Langkah pertama, *tokenizing*, adalah memecah teks menjadi unit-unit terkecil yang bermakna, seperti kata atau karakter. Sebagai contoh, kalimat "Salah sasaran ini" dipecah menjadi ["Salah", "sasaran", "ini"]. Langkah ini penting untuk memungkinkan analisis mendetail pada level kata. Langkah kedua, *case folding*, mengubah semua huruf dalam teks menjadi huruf kecil (*lowercase*) untuk menyamakan kata yang seharusnya sama meskipun memiliki variasi huruf besar dan kecil, misalnya "Bagus" dan "bagus". Ini dilakukan untuk mengurangi keragaman bentuk kata yang tidak perlu.

Selanjutnya, *filter stopwords* adalah langkah di mana kata-kata umum yang tidak signifikan seperti "saya", "aku", "mereka" dihapus dari teks. Langkah ini penting untuk membersihkan data dari kata-kata yang dapat mengganggu akurasi model klasifikasi. Tahap terakhir dalam *preprocessing* adalah *stemming*, yaitu mengubah kata-kata menjadi bentuk dasarnya, seperti "bermain", "bermainnya", dan "memainkan" yang semuanya disederhanakan menjadi "main". Meskipun *stemming* efektif untuk menangkap makna dasar dari kata, perlu diwaspadai bahwa proses ini kadang bisa menyamakan kata yang memiliki arti berbeda. Penelitian ini juga menggunakan kerangka kerja *Cross-Industry Standard Process for Data Mining* (*CRISP-DM*), yang merupakan standar industri dalam praktik data mining. *CRISP-DM* memberikan panduan melalui enam tahapan utama: pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan implementasi. Model ini dipilih karena fleksibilitasnya dalam berbagai aplikasi data mining dan telah banyak digunakan untuk membantu proses pengambilan keputusan berdasarkan data. Algoritma *Support Vector Machine* (SVM) digunakan sebagai teknik klasifikasi utama dalam penelitian ini. SVM adalah algoritma pembelajaran terawasi yang bekerja dengan cara memisahkan data menjadi dua kelas dalam ruang berdimensi tinggi menggunakan *hyperplane* optimal. Tujuan utama SVM adalah menemukan *hyperplane* yang dapat memaksimalkan jarak

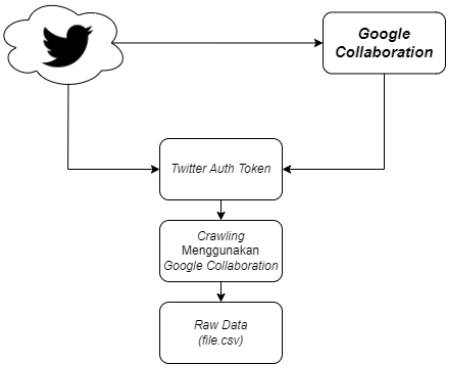
Jurnal Indonesia : Manajemen Informatika dan Komunikasi
https://journal.stmiki.ac.id
E-ISSN: 2723-7079 | P-ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990

atau margin antara dua kelas, di mana margin ini adalah jarak antara *hyperplane* dengan data yang paling dekat dari kedua kelas. Data yang berdekatan dengan *hyperplane* dan berpengaruh terhadap posisinya disebut *support vectors*, yang berperan penting dalam membentuk batas pemisah yang optimal (Rifkie, 2021).

3. Hasil dan Pembahasan

3.1 Hasil

Pada penelitian analisis sentimen untuk mengetahui bagaimana opini masyarakat mengenai pemberian beasiswa KIP-Kuliah, melalui implementasi algoritma *Support Vector Machine* (SVM) dengan menggunakan software RapidMiner. Dalam melakukan pengambilan data atau crawling data menggunakan Google Collabs diperlukan Twitter API Key. Kemudian setelah proses crawling data selesai, maka hasilnya disimpan dalam bentuk file csv.

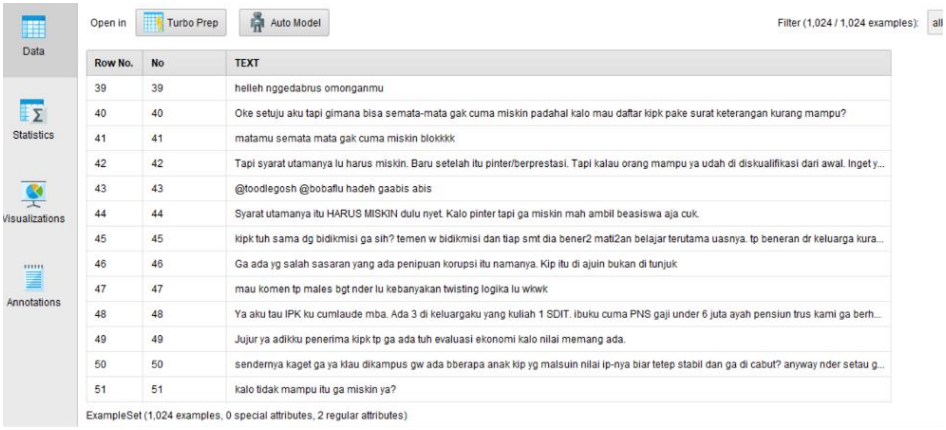


Gambar 2. Proses *Crawling* Data

Pengujian ini dapat dilakukan dengan menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang memiliki 6 (enam) tahapan yaitu :

1) Pemahaman Bisnis (*Business Understanding*)
   KIP Kuliah (Kartu Indonesia Pintar Kuliah) adalah program bantuan pendidikan dari pemerintah Indonesia yang ditujukan untuk siswa-siswa berprestasi dari keluarga kurang mampu agar dapat melanjutkan pendidikan ke jenjang perguruan tinggi. Program ini bertujuan untuk memberikan akses pendidikan yang merata dan mendukung pemerataan pendidikan tinggi di seluruh Indonesia. Berbagai tanggapan, opini, dan persepsi masyarakat, terutama pengguna media sosial, terhadap program ini diharapkan dapat memberikan informasi yang berguna bagi pemerintah dalam mengevaluasi dan meningkatkan program KIP Kuliah, serta mengidentifikasi area yang perlu diperbaiki berdasarkan sentimen masyarakat.
2) Pemahaman Data (Data *Understanding*)
   Pada tahapan ini peneliti berusaha memahami data yang akan digunakan dalam analisis. Pada penelitian ini, data yang digunakan adalah data tweet dari media sosial X yang mengandung kata kunci "KIP-Kuliah" dengan dua atribut yaitu tweet dan sentimen.
3) Data *Preparation*
   Pada tahap data preparation dilakukan beberapa tahapan antara lain sebagai berikut:
   a) Pengumpulan Data
      Pengumpulan data awal yaitu melakukan pengumpulan data yang dibutuhkan untuk mendukung dalam melakukan pemahaman data. Adapun sumber data utama yang digunakan dalam penelitian ini adalah data postingan Beasiswa KIP-Kuliah pada tanggal 29 April – 01 Mei 2024.

Jurnal Indonesia : Manajemen Informatika dan Komunikasi
https://journal.stmiki.ac.id
E - ISSN: 2723-7079 | P - ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990



Gambar 3. Hasil *Crawling* Data

b) *Cleansing* Data

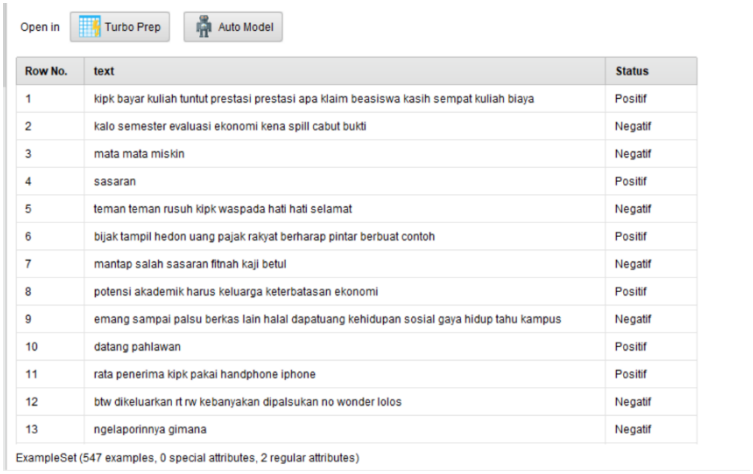
Pada tahapan ini peneliti akan melakukan proses cleansing dengan menggunakan web *http://www.gataframework.com/* sebelum melakukan labeling data agar dataset terhindar dari duplikasi data atau data yang tidak diperlukan. Pada proses ini dilakukan pembersihan dari berbagai noise seperti menghilangkan link URL, username, retweet, digit angka dan karakter.

Gambar 4. Hasil *Cleansing* Data

c) Labeling Data

Kemudian tahap berikutnya adalah melakukan penentuan sentimen berdasarkan masing-masing komentar. Selama proses penentuan sentimen berlangsung, disini peneliti melakukan pemberian sentimen secara manual. Berikut ini data yang sudah dibeli label.

Jurnal Indonesia : Manajemen Informatika dan Komunikasi
https://journal.stmiki.ac.id
E - ISSN: 2723-7079 | P - ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990


Open in Turbo Prep Auto Model

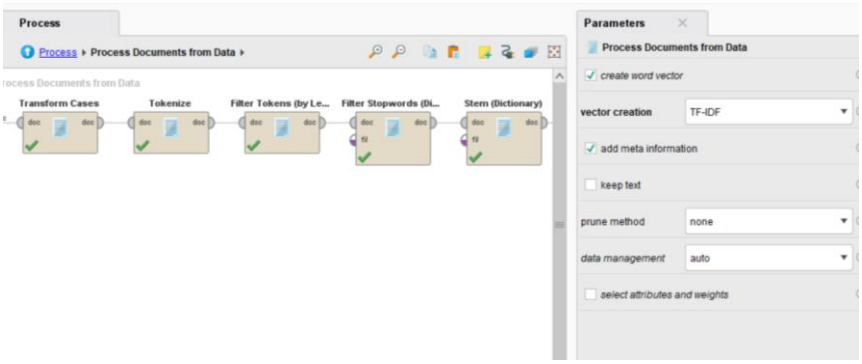
Row No.	text	Status
1	kipk bayar kuliah tuntut prestasi prestasi apa klaim beasiswa kasih sempat kuliah biaya	Positif
2	kalo semester evaluasi ekonomi kena spill cabut bukti	Negatif
3	mata mata miskin	Negatif
4	sasaran	Positif
5	teman teman rusuh kipk waspada hati hati selamat	Negatif
6	bijak tampil hedon uang pajak rakyat berharap pintar berbuat contoh	Positif
7	mantap salah sasaran fitnah kaji betul	Negatif
8	potensi akademik harus keluarga keterbatasan ekonomi	Positif
9	emang sampai palsu berkas lain halal dapatuang kehidupan sosial gaya hidup tahu kampus	Negatif
10	datang pahlawan	Positif
11	rata penerima kipk pakai handphone iphone	Positif
12	btw dikeluarkan rt rw kebanyakan dipalsukan no wonder lolos	Negatif
13	ngelaporinnya gimana	Negatif

ExampleSet (547 examples, 0 special attributes, 2 regular attributes)

Gambar 5. Data Telah Dilabeli Sentimen

4) *Modeling*

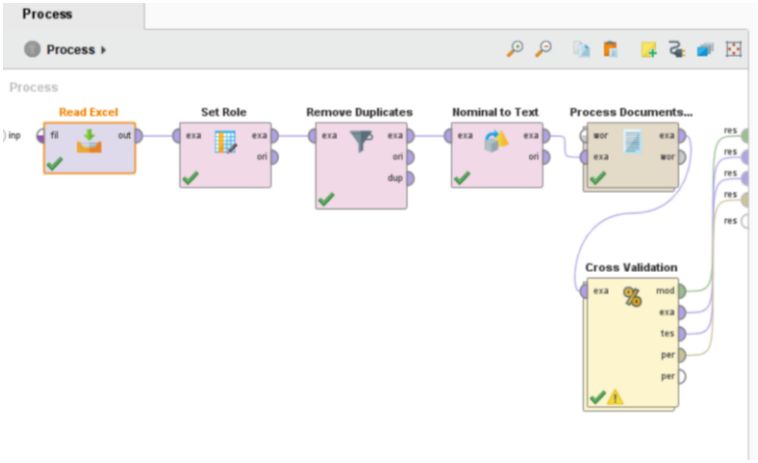
Pada tahapan ini peneliti akan melakukan preprocessing pada dataset hal ini ditunjukkan untuk menyiapkan data yang bersih dan bebas dari noise. Pada proses ini juga untuk menghitung pembobotan kata untuk keperluan pada proses modeling nanti. Pada tahap *Preprocessing* data yang ada di dalam operator *Process Documents from* Data, berisikan beberapa tahapan yaitu *Transform Case* untuk merubah seluruh kata yang ada di dalam data menjadi *lower case* atau huruf kecil dengan tujuan data yang akan diolah memiliki kesamaan untuk meminimalisir kesalahan, Tokenize untuk menghilangkan karakter-karakter atau simbol-simbol tertentu dan menghapus data yang bukan merupakan huruf, Filter Token (*by Length*) untuk menghapus data dengan ketentuan minimal terdapat empat karakter dan maksimal 25 karakter, dan Filter *Stopwords* (*Dictionary*) untuk menghapus kata-kata yang tidak dibutuhkan tanpa menghilangkan makna dari kata tersebut apabila berdiri sendiri. Dan Stem (*Dictionary*) untuk mengurangi kata-kata ke bentuk dasarnya (stem). Berikut *text preprocessing* dapat dilihat pada gambar 6. Setelah melewati tahapan tersebut, maka data tweet siap untuk dipakai pada tahap selanjutnya.



Gambar 6. Proses *Text Preprocessing*

Untuk tahapan *modeling* peneliti akan melakukan pengukuran performa klasifikasi dengan pemodelan menggunakan split data. Pemodelan ini data akan dibagi dengan rasio 0.8 dan 0.2. Pada tahapan ini akan dilakukan pengukuran performa dengan software RapidMiner.

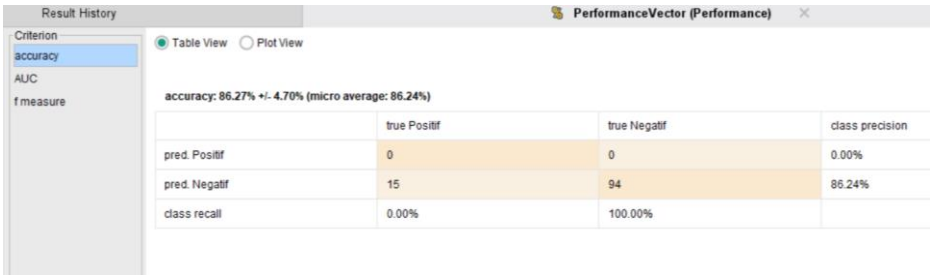
Jurnal Indonesia : Manajemen Informatika dan Komunikasi
https://journal.stmiki.ac.id
E - ISSN: 2723-7079 | P - ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990



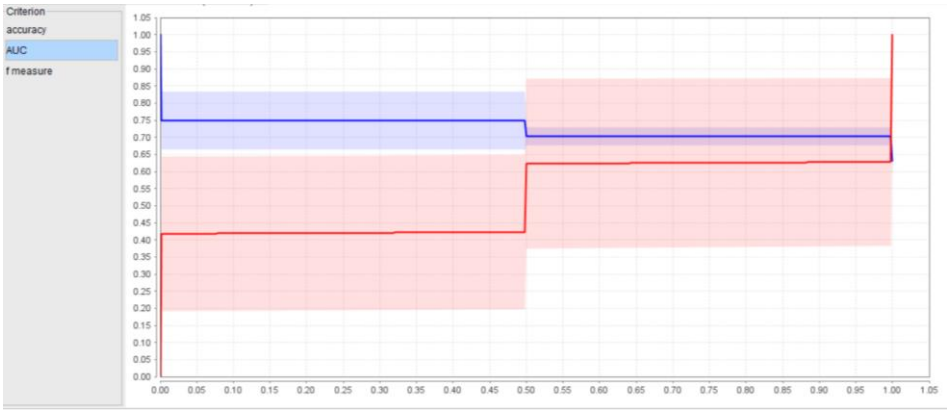
Gambar 7. Proses *Modeling*

5) Evaluasi

Pada penelitian ini akan dilakukan pengujian menggunakan *confussion matrix* untuk melihat hasil pengujian data yang diperoleh dari tahapan *modeling* dengan menggunakan algoritma *Support Vector Machine*. Berikut ini adalah *confussion matrix* hasil dari tahapan *modeling* yang dapat dilihat hasil dari perhitungan *RapidMiner* yang menunjukkan hasil dari model *Support Vector Machine*.

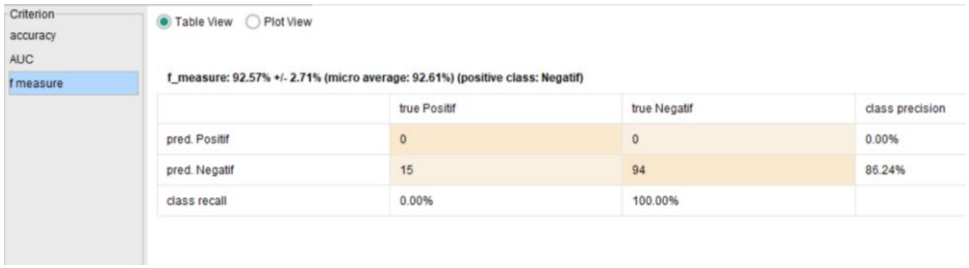


Gambar 8. Hasil Proses Akurasi



Gambar 9. Grafik *Area Under Curve* (AUC)

Jurnal Indonesia : Manajemen Informatika dan Komunikasi
https://journal.stmiki.ac.id
E-ISSN: 2723-7079 | P-ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990



Gambar 10. Hasil F *Measure*

Berdasarkan gambar diatas dibuat kesimpulan bahwa model yang digunakan untuk mengklasifikasikan sentimen terhadap pemberian beasiswa KIP-Kuliah menunjukkan hasil akurasi rata-rata sebesar 86.27% dengan variasi 4.70%. Ini menunjukkan bahwa secara umum, model ini dapat membuat prediksi yang benar sebanyak 86.27% dari keseluruhan data. Dengan hasil dari confusion matrix menunjukkan bahwa model tidak berhasil memprediksi kelas positif dengan baik. Tidak ada instance yang diprediksi sebagai positif (*True Positives* = 0), sementara semua instance negatif diprediksi dengan benar (True Negatives = 94). Sebanyak 15 instance positif diprediksi sebagai negatif (*False Negatives* = 15). Nilai AUC sebesar 0.523 menunjukkan bahwa model ini hampir setara dengan tebakan acak dalam membedakan antara sentimen positif dan negatif. Sementara Nilai F-measure sebesar 92.57% menunjukkan bahwa model ini sangat baik dalam memprediksi kelas negatif, tetapi tidak efektif untuk kelas positif. Model ini kurang baik dalam mengklasifikasikan sentimen positif karena ketidakseimbangan kelas dalam data pelatihan, mungkin disebabkan oleh dominasi komentar negatif terkait penyalagunaan beasiswa KIP-Kuliah. Untuk memperbaiki performa model, penting untuk menyeimbangkan data dan memastikan data pelatihan mencakup representasi yang memadai dari kedua sentimen positif dan negatif.

6) *Deployment*
Deployment merupakan tahap akhir dari dalam pembuatan laporan hasil dari kegiatan data *mining*. Laporan akhir yang berisi tentang pengetahuan yang diperoleh atau pengenalan pola dalam proses data *mining*.

3.2 Pembahasan

Pembahasan dalam penelitian ini mengkaji hasil analisis sentimen publik terhadap program beasiswa KIP-Kuliah menggunakan algoritma *Support Vector Machine* (SVM). Berdasarkan hasil klasifikasi, algoritma SVM mampu mencapai akurasi sebesar 86,27%, namun terdapat bias signifikan terhadap sentimen negatif. Hal ini sejalan dengan temuan A. Witanti *et al.* (2022) dalam penelitian mereka mengenai sentimen masyarakat terhadap vaksinasi COVID-19 di media sosial Twitter, di mana dominasi sentimen negatif juga terlihat. Bias negatif ini dapat disebabkan oleh ketidakseimbangan dalam distribusi data sentimen yang digunakan untuk pelatihan model, yang lebih banyak mengandung komentar negatif dibandingkan positif. Dalam hal beasiswa KIP-Kuliah, kritik terhadap penggunaan dana beasiswa untuk hal-hal konsumtif yang tidak relevan dengan tujuan pendidikan menjadi salah satu sumber sentimen negatif yang dominan di media sosial (Larasati *et al.*). Menurut Yuliana *et al.* (2022), proses seleksi penerima beasiswa KIP-Kuliah perlu ditingkatkan untuk memastikan bahwa bantuan diberikan kepada mahasiswa yang benar-benar membutuhkan. Penelitian mereka yang menggunakan metode *K-Means Clustering* menunjukkan bahwa seleksi yang lebih akurat dapat mengurangi ketidakpuasan publik dan meningkatkan persepsi positif terhadap program ini. Kritik terhadap mekanisme seleksi yang dianggap tidak tepat sasaran tercermin dalam berbagai unggahan di media sosial yang mengungkapkan penerima beasiswa memamerkan gaya hidup mewah, memicu kemarahan dan kecaman dari masyarakat luas.

E-ISSN: 2723-7079 | P-ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990

Pramesti dan Pratiwi (2023) dalam studi mereka menggunakan SVM dan *Decision Tree* untuk menganalisis sentimen terhadap program Merdeka Belajar Kampus Merdeka (MBKM), menemukan bahwa distribusi data yang tidak seimbang dapat mengakibatkan bias dalam model klasifikasi. Hal ini memperkuat pentingnya penerapan teknik penyeimbangan data, seperti *oversampling* pada kelas minoritas atau penerapan algoritma yang lebih kompleks, untuk memperbaiki performa model dan menghasilkan klasifikasi yang lebih akurat. Dalam hal ini, meskipun algoritma SVM memiliki kemampuan yang baik dalam mengklasifikasikan data berdimensi tinggi, bias yang muncul terhadap sentimen negatif menunjukkan bahwa algoritma ini tidak selalu optimal dalam situasi data yang tidak seimbang. Pamungkas (2021) juga menyoroti bahwa penggunaan algoritma yang lebih kompleks atau penggabungan model (*ensemble methods*) dapat memberikan hasil yang lebih stabil dan akurat dalam klasifikasi sentimen.

Hasil penelitian ini memberikan informasi penting bagi pembuat kebijakan untuk mengevaluasi efektivitas program beasiswa KIP-Kuliah dan mempertimbangkan perbaikan pada mekanisme seleksi dan distribusi. Dengan adanya dominasi sentimen negatif yang terdeteksi, pemerintah perlu melakukan langkah-langkah strategis seperti peningkatan transparansi seleksi, pengawasan penggunaan dana, serta edukasi kepada penerima beasiswa tentang pentingnya memanfaatkan bantuan sesuai dengan tujuan program. Bagaskoro *et al.* (2023) juga menekankan bahwa analisis sentimen dapat menjadi alat yang efektif untuk memahami persepsi publik dan mengidentifikasi area perbaikan dalam kebijakan publik.

Dengan demikian, pembahasan ini menegaskan perlunya evaluasi berkelanjutan terhadap program beasiswa KIP-Kuliah dan penerapan strategi yang lebih adaptif untuk meningkatkan penerimaan dan kepercayaan publik. Penelitian ini juga menunjukkan pentingnya analisis data yang lebih seimbang untuk mengurangi bias dan memastikan bahwa hasil klasifikasi benar-benar merefleksikan opini publik secara objektif.

4. Kesimpulan

Berdasarkan hasil analisis dan pengujian yang telah dilakukan pada penelitian ini, menghasilkan kesimpulan bahwa algoritma *Support Vector Machine* (SVM) mampu mengklasifikasikan sentimen masyarakat pengguna media sosial X terhadap pemberian Beasiswa KIP-Kuliah dengan akurasi yang cukup baik. Terdapat bias terhadap sentimen negatif yang disebabkan oleh dominasi data negatif dalam dataset yang digunakan. Bahwa sebagian besar tanggapan masyarakat cenderung negatif, terutama terkait dengan penyalahgunaan beasiswa oleh penerimanya.

5. Daftar Pustaka

Akbar, Y., & Ihsan, A. N. (2023). Analisis Sentimen Twitter Terhadap Opini Masyarakat Pada Sea Games Kamboja 2023 Menggunakan Algoritma Support Vector Machine. *INTECOMS: Journal of Information Technology and Computer Science*, *6*(2), 814-821. DOI: https://doi.org/10.31539/intecoms.v6i2.7670.

Bagaskoro, S. A., Hasanah, A., Bahri, S., Utami, E., & Yaqin, A. (2023). ANALISIS SENTIMEN LPDP (LEMBAGA PENGELOLA DANA PENDIDIKAN) PADA MEDIA SOSIAL TWITTER. *Pseudocode*, *10*(2), 65-73. DOI: https://doi.org/10.33369/pseudocode.10.2.65-73.

Handayani, A., & Zufria, I. (2023). Analisis sentimen terhadap bakal capres ri 2024 di twitter menggunakan algoritma svm. *Journal of Information System Research (JOSH)*, *5*(1), 53-63. DOI: https://doi.org/10.47065/josh.v5i1.4379.

E-ISSN: 2723-7079 | P-ISSN: 2776-8074
Vol. 5 No. 3 (2024) | September
OPEN ACCESS
https://doi.org/10.35870/jimik.v5i3.990

Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*, *10*(1), 18-26. DOI: https://doi.org/10.34148/teknika.v10i1.311.

Larasati, A. D., Dinda, D., Aidah, N. A., Gustiputri, R., & Isyak, S. N. R. (2022). Analisis Kebijakan Program Beasiswa Kartu Indonesia Pintar-Kuliah (Kip-K) Di Universitas Diponegoro. *Jurnal Ilmu Administrasi Dan Studi Kebijakan (JIASK)*, *5*(1), 1-22. DOI: https://doi.org/10.48093/jiask.v5i1.91.

Luqyana, W. A., Cholissodin, I., & Perdana, R. S. (2018). Analisis sentimen cyberbullying pada komentar instagram dengan metode klasifikasi support vector machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *2*(11), 4704-4713.

Manalu, D. R., Tobing, M. C. L., & Yohanna, M. (2022). Analisis sentimen twitter terhadap wacana penundaan pemilu dengan metode support vector machine. *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, *6*(2), 149-156. DOI: https://doi.org/10.46880/jmika.Vol6No2.pp149-156.

Pamungkas, B., Purbaya, M. E., & AK, D. J. (2021). Analisis Sentimen Twitter Menggunakan Metode Support Vector Machine (SVM) pada Kasus Benih Lobster 2020. *Journal of Informatics Information System Software Engineering and Applications (INISTA)*, *3*(2), 10-20. DOI: https://doi.org/10.20895/inista.v3i2.243.

Pramesti, L. A., & Pratiwi, N. (2023). Analisis Sentimen Twitter Terhadap Program MBKM Menggunakan Decision Tree dan Support Vector Machine. *Journal of Information System Research (JOSH)*, *4*(4), 1145-1154. DOI: https://doi.org/10.47065/josh.v4i4.3807.

Slamet, R., Gata, W., Novtariany, A., Hilyati, K., Jariyah, F. A., & Mandiri, U. N. (2022). Twitter Sentiment Analysis of South Korea Artists As Brand Ambassadors of Local Beauty Products. *J. Inf. Technol. Comput. Sci*, *5*(1).

Witanti, A., Wates-Jogjakarta, B. Y. J. R., Sedayu, K., Bantul, K., & Yogyakartalamat, D. I. (2022). ANALISIS SENTIMEN MASYARAKAT TERHADAP VAKSINASI COVID-19 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM). *Jurnal Sistem Informasi dan Informatika (Simika) P-ISSN*, *5*, 2622-6901.

Yuliana, D. T., Fathoni, M. I. A., & Kurniawati, N. (2022). Penentuan Penerima Kartu Indonesia Pintar KIP Kuliah dengan Menggunakan Metode K-Means Clustering. *Journal Focus Action of Research Mathematic (Factor M)*, *5*(1), 127-141. DOI: https://doi.org/10.30762/f_m.v5i1.570.