

Sentiment Analysis on Hybrid Learning Using Support Vector Machine Algorithm

Analisis Sentimen mengenai Pembelajaran *Hybrid* Menggunakan Algoritma *Support Vector Machine*

Fabianus I Made Jeven¹, Gerson Feoh^{2*}, I Made Dwi Ardiada³

^{1,2,3}Prodi Studi Teknik Informatika, Universitas Dhyana Pura, Bali, Indonesia

(*) Corresponding Author: gerson.feoh@undhirabali.ac.id

Article info

Keywords:

Sentiment Analysis, User Satisfaction Level, Hybrid Learning, Twitter, Support Vector Machine

Abstract

Social media Twitter has become a popular platform for users to share information, interact and gain knowledge including in the field of Education with hybrid learning methods that combine online and offline learning methods. However, the level of Twitter users' satisfaction with hybrid learning methods is still a major concern in improving the quality of learning. Therefore, this research aims to conduct a sentiment analysis of Twitter users' satisfaction level towards hybrid learning using the Support Vector Machine algorithm. In this research, the author uses the Support Vector Machine algorithm to determine the accuracy generated by the Support Vector Machine method and to analyze the sentiment of Twitter users' satisfaction regarding hybrid learning. The results of the classification with the Support Vector Machine algorithm method after testing obtained the results of an accuracy level of 68.33%, negative precision of 71.15%, positive precision of 50%, negative recall of 90.24%, and positive recall of 21.05%. The results showed that students were less satisfied with the application of hybrid learning methods, this could be detrimental to themselves due to a lack of understanding of the material provided. This statement is based on 69.5% negative reviews and 30.5% positive reviews.

Kata kunci:

Analisis Sentimen, Tingkat Kepuasan Pengguna, Pembelajaran Hybrid, Twitter, Support Vector Machine

Abstrak

Media sosial Twitter telah menjadi *platform* populer bagi pengguna untuk berbagi informasi, berinteraksi, dan memperoleh pengetahuan termasuk dalam bidang Pendidikan dengan metode pembelajaran *hybrid* yaitu menggabungkan metode pembelajaran secara *online* dan *offline*. Namun, tingkat kepuasan pengguna Twitter terhadap metode pembelajaran *hybrid* masih menjadi perhatian utama dalam meningkatkan kualitas pembelajaran. Oleh karena itu, penelitian ini bertujuan untuk melakukan analisis sentimen tingkat kepuasan pengguna Twitter terhadap pembelajaran *hybrid* menggunakan algoritma *Support Vector Machine*. Dalam penelitian ini, penulis menggunakan algoritma *Support Vector Machine* untuk mengetahui akurasi yang dihasilkan oleh metode *Support Vector Machine* dan untuk menganalisis sentimen kepuasan pengguna Twitter terkait pembelajaran *hybrid*. Hasil dari klasifikasi dengan metode algoritma *Support Vector Machine* setelah dilakukan pengujian mendapatkan hasil tingkat *accuracy* 68,33%, *precision* negatif 71,15%,

precision positif 50%, *recall* negatif 90,24% dan *recall* positif 21,05%. Hasil penelitian menunjukkan bahwa kurang puasnya mahasiswa dengan diterapkan metode pembelajaran *hybrid*, hal ini dapat merugikan dirinya sendiri dikarenakan kurangnya pemahaman dari materi yang diberikan pernyataan ini berdasarkan 69,5% ulasan negatif dan 30.5% ulasan positif.

PENDAHULUAN

Pandemi Covid-19 telah melanda berbagai wilayah termasuk Indonesia. Selain itu, virus ini juga mendorong banyak negara untuk menerapkan langkah-langkah seperti penggunaan masker, pembatasan jarak, menjaga kebersihan tangan, dan tindakan lainnya (Putra, Triayudi, & Andrianingsih., 2022). Pemerintah Indonesia juga menerapkan status penguncian wilayah yang telah menghentikan kegiatan masyarakat, baik perusahaan, pemerintah maupun swasta, pengusaha, transportasi, pariwisata, dan banyak sektor lainnya, salah satunya adalah pada sektor bidang Pendidikan (Setiawan, *et al.*, 2021).

Pandemi Covid-19 pada saat ini dirasakan sudah mulai membaik, oleh karena itu pemerintah Indonesia sudah mulai membuka kembali sektor usaha yang ada di Indonesia khususnya bidang pendidikan perguruan tinggi, salah satu strategi yang digunakan oleh instansi pendidikan perguruan tinggi adalah dengan cara metode pembelajaran *hybrid*, hal ini juga sesuai dengan kondisi pendidikan saat ini sudah mulai pada fase normal, hal ini dapat berdampak pada institusi pendidikan khususnya perguruan tinggi yang harus menerapkan pembelajaran *hybrid*. Metode pembelajaran *hybrid* ini adalah metode pembelajaran yang menggabungkan belajar secara *Online* dan perkuliahan secara tatap muka (Watrianthos, *et al.*, 2022). Dengan adanya metode Pembelajaran *hybrid* ini dapat menjadi topik dari pengguna Twitter dalam memberikan argumennya dalam media sosial Twitter. Twitter merupakan *platform* media sosial yang memungkinkan setiap individu dari berbagai penjuru dunia untuk saling berkoneksi dan berdiskusi tentang berbagai topik secara langsung serta memungkinkan mereka untuk memberikan pandangan pribadi mereka mengenai berbagai hal yang terjadi (Natasuwarna, 2020).

Dalam menganalisis penelitian ini, penulis memanfaatkan pendekatan Metode *Support Vector Machine*. Menurut Pintoko & Muslim (2018). Metode *Support Vector Machine* termasuk dalam kategori teknik klasifikasi dalam kategori pengelompokan data teks dan pemrosesan data yang digunakan dalam analisis sentimen, terutama ketika berhadapan dengan data dari *platform* Twitter. *Support Vector Machine* berfungsi untuk menemukan garis pemisah terbaik (*hyperplane*) antara dua kelompok sentimen (positif dan negatif) dengan *margin optimal*, sehingga efektif dalam menangani permasalahan klasifikasi teks.

METODE

Metode Analisis Text Mining

Penelitian ini menggunakan metode analisis *text mining*, yang diawali dengan tahap *crawling* data Twitter menggunakan *jupyter notebook* dengan kata kunci pencarian “kuliah *hybrid*”. Setelah itu, dilakukan analisis seleksi data yang melibatkan pemisahan data yang relevan untuk digunakan dari data yang tidak relevan. Tahap selanjutnya *labelling* data yaitu dengan memberi label pada *dataset* berdasarkan emotikon dari pengguna Twitter, dilanjutkan dengan *pre-processing* dan pembobotan TF-IDF pada tahap

ini ada lima proses yaitu *cleaning case folding*, *tokenizing*, *stopword removal* dan *stemming* tujuan dari proses ini adalah untuk mempersiapkan dan membersihkan data mentah sebelum dijalankan melalui proses klasifikasi. Kemudian selanjutnya tahap proses klasifikasi dengan metode *Support Vector Machine* pada proses ini akan menghasilkan prediksi untuk label sentimen. dan tahap yang terakhir *confusion matrix* untuk melakukan pengujian sehingga menghasilkan *accuracy*, *precision* dan *recall*.

HASIL DAN PEMBAHASAN

Hasil

Pada proses ini mendapatkan hasil yang berupa informasi dan pengetahuan yang baru dengan topik analisis sentimen pada data *mining* dalam proses klasifikasi tingkat kepuasan pengguna Twitter terhadap pembelajaran *hybrid* menggunakan algoritma *Support Vector Machine*.

Pengujian Data Dengan Perbandingan 70:30

Pada bagian pengujian ini, hasil diperoleh dari langkah pemisahan data dalam perbandingan 70:30. Ini mengarah pada penggunaan 70% data untuk proses *training* dari total 140 data *tweet*, sementara sisanya, yaitu 30% data, digunakan untuk *testing* dari kumpulan data *tweet* sebanyak 60 data. Berikut gambar 1. hasil *confusion matrix* dengan *RapidMiner*.

accuracy: 68.33%

	true Negatif	true Positif	class precision
pred. Negatif	37	15	71.15%
pred. Positif	4	4	50.00%
class recall	90.24%	21.05%	

Gambar 1. Hasil Confusion Matrix Dengan Data Testing 30%

Gambar 1 menunjukkan hasil *confusion matrix* yang dihasilkan dengan memanfaatkan data *testing* sebesar 30% dari keseluruhan *dataset* yang terdiri dari 200 *tweet* dan jumlah data *testing* terdiri dari 60 data *tweet*. dimana hasil prediksi ini akan dilakukan tahap perbandingan antara data *tweet* label berdasarkan emotikon dengan hasil prediksi dari *RapidMiner* dan adapun di dalamnya terdapat hasil *true negatif*, *true positif*, *prediction negatif*, dan *prediction positif*.

Pada pengujian ini menggunakan perbandingan 70:30 dimana 70% data *training* dari 140 data *tweet* dan 30 % data *testing* dari 60 dari data *tweet* dan total data *tweet* sebanyak 200. Dari pengujian ini mendapatkan hasil **accuracy 68,33%**, **precision negatif 71,15%**, **precision positif 50%**, **recall negatif 90,24%** dan **recall positif 21,05%** untuk pembelajaran *hybrid* dengan metode *Support Vector Machine* pada *RapidMiner*.

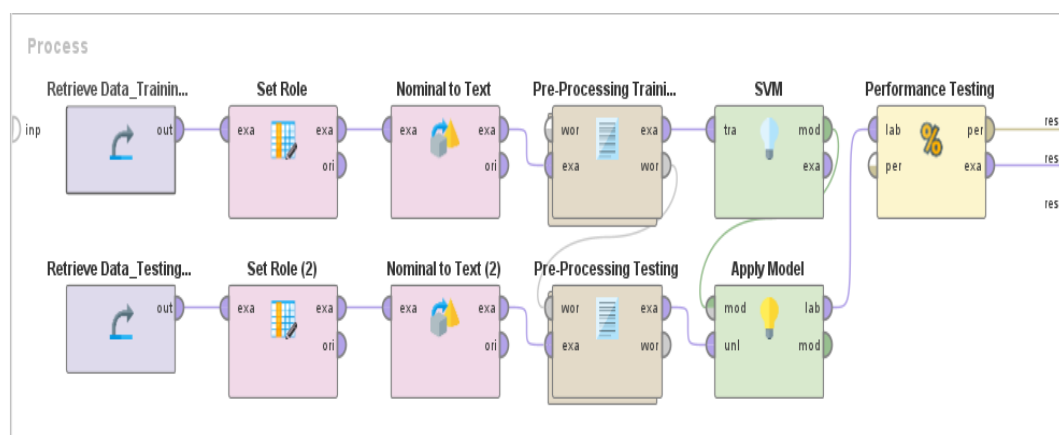
Pembahasan

Dalam konteks penelitian ini, proses *crawling* data itu mendapatkan data sebanyak 3.600 data *tweet*. Dalam pengumpulan data *tweet* menggunakan Bahasa Indonesia dengan kata kunci pencarian “kuliah *hybrid*”. Langkah berikutnya adalah melaksanakan dua tahap seleksi data. Dalam tahap pertama seleksi data, data dipilah antara yang relevan dan tidak relevan untuk penelitian. Data yang dianggap relevan ditentukan berdasarkan emotikon yang digunakan oleh pengguna Twitter dan mendapatkan hasil 520 data *tweet*, selanjutnya dilakukan seleksi data yang kedua untuk menghilangkan kalimat yang mengandung iklan, komentar dan kata-kata yang memiliki makna yang tidak penting dan menghasilkan 200 data *tweet*. Tahap selanjutnya adalah *labelling* data *tweet* pada tahap ini dilakukan proses pemberian label secara berdasarkan emotikon dari pengguna Twitter.

Tahap selanjutnya dilakukan *Pre-processing* data dan pembobotan TF-IDF Dalam *RapidMiner* proses ini diawali dengan tahap *import* data *tweet* kedalam *RapidMiner* dimana file CSV yang sudah diberi label sebelumnya akan di *import* kedalam *RapidMiner* untuk proses selanjutnya yaitu *Cleaning Data Tweet* Proses berguna untuk membersihkan data *tweet* dari URL, @, *hashtag*, simbol-simbol dan *mention*.

Tahap selanjutnya adalah *Case Folding* dimana data *tweet* yang sudah dibersihkan akan dikonversi menjadi huruf kecil (*lowercase*) secara keseluruhan, dilanjutkan proses *Tokenizing* yaitu membagi kalimat menjadi kata perkata, setelah proses *tokenizing* dilanjutkan dengan proses *Stopword Removal* adalah proses mengeliminasi kata-kata yang tidak memiliki nilai penting dan termasuk dalam daftar kamus *stopword* Bahasa Indonesia dalam penelitian ini jumlah kata dalam kamus *stopword* berjumlah 668 kata dan tahap yang terakhir adalah *Stemming* dimana proses ini dilakukan untuk mengubah kata imbuhan menjadi kata dasar dan kata dasar ini terdapat dalam kamus *stemming* Bahasa Indonesia jumlah kata dasar yang digunakan pada penelitian ini berjumlah 238 kata.

Langkah berikutnya dalam penelitian ini melibatkan pembagian antara data *training* dan data *testing* dengan perbandingan 70:30. Ini berarti 70% dari total 140 data *tweet* digunakan sebagai data *training*, sedangkan 30% dari total 60 data *tweet* digunakan sebagai data *testing*, dengan jumlah total data *tweet* mencapai 200 data *tweet*.



Gambar 2. Proses Klasifikasi SVM Pada *RapidMiner*

Proses klasifikasi SVM pada *RapidMiner* dimana Setelah melalui proses *pre-processing* data, langkah berikutnya adalah menjalankan proses klasifikasi dengan metode SVM yang diawali dengan menghubungkan hasil *pre-processing* di data *training* dihubungkan dengan metode SVM, kemudian dari metode SVM dihubungkan dengan operator *apply model* dan *pre-processing* di data *testing* dihubungkan dengan operator *apply model* untuk mendapatkan hasil dari klasifikasi menggunakan metode SVM dilanjutkan proses klasifikasi SVM kemudian dihubungkan dengan operator *performance testing* untuk mengeluarkan hasil dari pengujian dengan perbandingan 70:30.

Hasil dari klasifikasi metode SVM dengan perbandingan 70:30 dengan data *testing* 30% yang berjumlah 60 data *tweet* maka hasil dari *labelling* berdasarkan emotikon berjumlah 41 data *tweet* dengan label negatif dan 19 data *tweet* dengan label positif sedangkan hasil dari prediksi metode SVM berjumlah 52 data *tweet* dengan prediksi label negatif dan 8 data *tweet* dengan prediksi label positif. Tahapan selanjutnya yaitu pengujian antara *Labelling* berdasarkan emotikon dengan hasil prediksi sentimen dari metode SVM dari hasil tersebut bisa menghasilkan *accuracy*, *precision*, dan *recall*.

SIMPULAN

Dari hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa:

1. Berdasarkan hasil analisis tingkat kepuasan pengguna Twitter terhadap pembelajaran *hybrid* menggunakan algoritma *Support Vector Machine* pada analisis sentimen diketahui bahwa mayoritas mahasiswa kurang puas dengan diterapkan metode pembelajaran *hybrid*. Hal ini dapat merugikan dirinya sendiri dikarenakan kurangnya pemahaman dari materi yang diberikan hal ini didasarkan pada 69,5% ulasan negatif dari hasil klasifikasi. Namun disini lain pembelajaran *hybrid* itu menyenangkan karena tidak memerlukan banyak uang untuk berangkat kekampus hal ini didasarkan pada 30,5% ulasan positif dari hasil klasifikasi.
2. Berdasarkan hasil perhitungan *accuracy*, *precision* dan *recall* dari algoritma *Support Vector Machine* yang dilakukan oleh penulis pada penelitian ini, data menunjukkan bahwa hasil ***accuracy* 68,33%** merupakan gambaran model SVM dalam mengklasifikasikan dengan tepat pada analisis sentimen. ***Precision* negatif 71,15%** dan ***precision* positif 50%** merupakan gambaran dari tingkat ketepatan hasil prediksi yang diberikan oleh model SVM pada analisis sentimen. ***Recall* negatif 90,24%** dan ***recall* positif 21,05%** adalah gambaran keberhasilan model SVM dalam menemukan kembali informasi yang relevan dari pembelajaran *hybrid* pada analisis sentimen.

DAFTAR PUSTAKA

- Pintoko, B. M, & Muslim, K. (2018). Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naïve Bayes Classifier. *E-Proceeding of Engineering*, 5(3), 8121–8230.
- Natasuwarna, A. P. (2020). Seleksi Fitur Support Vector Machine pada Analisis Sentimen Keberlanjutan Pembelajaran Daring. *Techno.Com*, 19(4), 437–448. <https://doi.org/10.33633/tc.v19i4.4044>
- Setiawan, H., Utami, E., & Sudarmawan, S. (2021). Analisis Sentimen Twitter Kuliah Online Pasca Covid-19 Menggunakan Algoritma Support Vector Machine dan Naive

- Bayes. *Jurnal Komtika (Komputasi Dan Informatika)*, 5(1), 43–51. <https://doi.org/10.31603/komtika.v5i1.5189>
- Watrianthos, R., Giatman, M., Simatupang, W., Syafriyeti, R., & Daulay, N. K. (2022). Analisis Sentimen Pembelajaran Campuran Menggunakan Twitter Data. *Jurnal Media Informatika Budidarma*, 6(1), 166. <https://doi.org/10.30865/mib.v6i1.3383>
- Putra, T. W., Triayudi, A., & Andrianingsih, A. (2022). Analisis Sentimen Pembelajaran Daring menggunakan Metode Naïve Bayes, KNN, dan Decision Tree. *Jurnal Teknologi Informasi dan Komunikasi*, 6(1). <https://doi.org/10.35870/jtik.v6i1.368>