

ANALISIS ALGORITMA SUPPORT VECTOR MACHINE PADA OPINI MASYARAKAT MENGGUNAKAN TWITTER TERHADAP VAKSIN COVID-19

Support Vector Machine Algorithm Analysis On Public Opinion Using Twitter Against Covid-19 Vaccine

Frizka Fitriana
Teknologi Informasi, Politeknik Sampit
frizkafitriana@gmail.com

ABSTRAK

Penyebaran virus covid-19 yang terjadi dan juga terdapat bahaya yang mengintainya maka butuh penanganan untuk mencegah hal ini semakin berbahaya. Terdapat beberapa cara mencegah penyebaran virus covid-19 yaitu dengan pengembangan vaksin. Vaksin yang dikembangkan ini tidak hanya melindungi penderita covid-19 tetapi bisa juga mengurangi penyebaran penyakit covid-19. Dengan mempertimbangkan pentingnya vaksin covid-19 ini Pemerintah Republik Indonesia telah mengeluarkan vaksinasi untuk masyarakat Indonesia pada akhir tahun 2020. Hal tersebut sudah dikonsultasikan kepada Indonesian Technical Advisory Group on Immunization (ITAGI) yang bertugas memberikan nasehat kepada Menteri Kesehatan.

Dengan adanya vaksin covid-19 tentunya menuai pendapat berbagai pihak. Ada yang menerima vaksin covid-19 ini diterapkan ada juga yang menganggap hal ini tidak baik, selain itu ada juga masyarakat yang bersikap netral. Karena hal tersebut penulis melakukan penelitian mengenai analisis sentimen terhadap opini masyarakat media sosial twitter menggunakan algoritma Support vector machine. Pada penelitian ini akan menguji algoritma Support vector machine apakah sudah efektif dalam menguji 1000 data dengan hal yang diuji mengenai performa akurasi, waktu training dan juga nilai MAE. Dengan hal itu didapatkan hasil pengujian model algoritma Support vector machine memiliki nilai tingkat performa akurasi dengan nilai 85%, performa waktu training dengan nilai 40,60 detik, dan nilai MAE dengan nilai 0,001334 pada 1000 data.

Kata kunci: Analisis Algoritma, Support vector machine, Vaksin covid-19

ABSTRACT

The spread of the COVID-19 virus that occurs and there are also dangers that lurk in it, so handlers are needed to prevent this from becoming more dangerous. There are several ways to prevent the spread of the COVID-19 virus, namely by developing vaccines. The vaccine developed not only protects people with covid-19 but can also reduce the spread of covid-19 disease. By considering the importance of the COVID-19 vaccine, the Government of the Republic of Indonesia has issued vaccinations for the Indonesian people at the end of 2020. This has been consulted with the Indonesian Technical Advisory Group on Immunization (ITAGI) which is tasked with providing advice to the Minister of Health.

With the COVID-19 vaccine, of course, the opinions of various parties have been reaped. There are those who receive the COVID-19 vaccine applied, some think this is not good, besides that there are also people who are neutral. Because of this, the author conducted research on sentiment analysis of social media community opinion using the Support vector machine algorithm. This study will test the Support vector machine algorithm whether it has been effective in testing 1000 data with things tested about accuracy performance, training time and also MAE values. With that, the results of testing the Support vector machine algorithm model have a level value accuracy performance with a value of 85%, training time performance with a value of 40.60 seconds, and MAE value with a value of 0.001334 on 1000 data.

Keywords: Algorithm Analysis, Support vector machine, Covid-19 vaccine

Pendahuluan

Covid-19 merupakan wabah penyakit yang diakibatkan oleh virus. Virus Covid-19 ini telah ditetapkan secara resmi oleh World Health Organization sebagai pandemi global. Wabah penyakit Covid-19 ini sangat berpotensi besar membuat kekebalan tubuh manusia yang menderitanya melemah, dan sudah banyak kasus yang menyebabkan kematian penderita virus ini.

Karena bahaya yang besar ini banyak muncul solusi yang diterapkan oleh pemerintah, salah satunya dengan adanya vaksinasi. Vaksin ini tidak hanya melindungi penderita yang terpapar virus, tapi juga bisa mengurangi penyebaran virus Covid-19 ini.

Dengan keputusan adanya vaksinasi untuk menangani virus Covid-19 yang sudah menyebar ini menuai banyak kontroversial. Banyak sekali para kalangan yang menyampaikan pendapatnya melalui media sosial twitter. Twitter sendiri merupakan media sosial yang banyak digunakan para kalangan untuk menyampaikan pendapatnya. Dan berdasarkan data dari Global Digital Statistic "Digital, Social & Mobile in 2019" di We Are Social (2019), pada tahun 2019 jumlah pengguna media sosial di Indonesia yaitu berjumlah lebih dari 150 juta pengguna.

Pada pengujian algoritma yang menggunakan data dari tweet masyarakat pada twitter mengenai vaksin Covid-19 akan menggunakan algoritma Support Vector Machine. Algoritma Support Vector Machine dikenal sebagai algoritma machine learning yang sangat baik dalam hal klasifikasi pembobotan yang akan diproses.

Penelitian terdahulu (Ratino, 2020) yang menggunakan metode Support Vector Machine dan Naïve Bayes dengan penambahan Particle Swarm Optimization (PSO) menghasilkan akurasi terbaik dalam menganalisa komentar Instagram terhadap informasi berkaitan dengan COVID-19. Dalam penelitian ini pengumpulan data mengambil dari Instagram pada akun World Health Organization (WHO) dengan postingan berkisar tentang covid-19. Dan hasil dari penelitian analisis sentimen terhadap data komentar Instagram mengenai Covid-19 yang dilakukan dengan Support Vector Machine dengan nilai akurasi 80,23% dan Naïve Bayes dengan nilai akurasi 78,02%. Dan hasil Support Vector Machine dengan Particle Swarm Optimization dengan nilai akurasi 81,16% dan Naïve Bayes dengan Particle Swarm Optimization dengan memperoleh nilai akurasi 79,07%.

Sedangkan penelitian yang dilakukan oleh Rian Tineges, Agung Triayudi dan Ira Diana Sholihati (2020) yaitu Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM). Hasil yang didapat akurasi sebesar 87% dengan ketepatan antara hasil prediksi dengan data sebenarnya (precision) sebesar 86%, tingkat keberhasilan sistem dalam memprediksi sebuah data (recall) sebesar 95%, tingkat kesalahan semua data yang diprediksi (error rate) sebesar 13%, sedangkan untuk nilai perbandingan rata-rata

precision dan recall (f1- score) adalah sebesar 90%.

Berdasarkan dari kesimpulan beberapa pendapat sebelumnya dapat ditarik kesimpulan bahwa algoritma Support Vector Machine (SVM) menjadi salah satu algoritma machine learning yang dikenal cukup baik dalam melakukan klasifikasi berdasarkan pembobotan yang diproses dalam algoritma tersebut. Oleh karena itu pada penelitian ini akan dilakukan dengan menganalisis algoritma Support Vector Machine (SVM) pada data postingan twitter berdasarkan kata kunci vaksin covid-19 dengan 1000 data. Dan yang akan diuji berupa performa akurasi, waktu training dan juga nilai MAE.

Metode

Metode pertama yang dilakukan penelitian ini adalah diagnosa. Diagnosa sendiri memiliki beberapa tahap sebagai berikut:

- a. Studi Literatur
Studi literatur merupakan langkah awal yang dilakukan dengan mempelajari arsip/buku, jurnal, makalah dan laporan penelitian sebelumnya sebagai pedoman/landasan dalam menyusun penelitian ini. Studi Literatur memberikan penjelasan mengenai teori-teori yang dapat membantu dalam menyelesaikan penelitian ini.
- b. Pengumpulan data dan Informasi
Pada tahapan ini dilakukan pengumpulan data dan informasi. Dalam hal ini terdapat beberapa tahap pengumpulan data dan informasi yaitu melakukan observasi dan dokumentasi
- c. Requirements Definition
Pada tahap ini akan dilakukan identifikasi kebutuhan data yang akan digunakan dalam penelitian ini, dalam hal ini data yang akan digunakan adalah postingan pada sosial media twitter berkaitan dengan hastag vaksin covid-19.

Selain dari tahap diagnosa, ada juga tahapan lainnya yaitu tahapan Action Taking. Tahapan atau metode Action Taking ini terdapat beberapa poin. Yaitu sebagai berikut:

- a. **Pemodelan Algoritma**
 Pada tahap ini pemodelan metode untuk setiap model algoritma yaitu Support Vector Machine yang akan dilakukan. Pemodelan algoritma ini dengan melakukan pembuatan model atau arsitektur Support Vector Machine dengan konfigurasi.
- b. **Pemodelan Proses**
 Sedangkan tahap ini akan dilakukan pemodelan proses dengan menggunakan UML atau Unified Modelling Language untuk membantu dalam pendeskripsian dan desain sistem perangkat lunak yang akan digunakan dalam penelitian ini.
- c. **Pemodelan Data**
 Kemudian akan dilakukan 3 tahap analisis yaitu pengambilan data dari twitter dengan parameter vaksin covid-19, pre-processing data serta pelabelan yang akan dilakukan oleh pakar untuk menentukan label.
- d. **Pemodelan Sistem**
 Tahapan ini akan merancang desain antar muka sistem guna sebagai acuan sistem yang akan dibuat. Diantaranya desain antarmuka pengelola dataset, antarmuka pengolahan data training, antarmuka hasil perbandingan dan antarmuka pediksi.

Sedangkan tahapan yang terakhir adalah evaluasi. Tahapan evaluasi ini melakukan evaluasi hasil metode yang SVM yang digunakan dan selanjutnya dihitung hasil presentase performa pada metode. Tahap terakhir diperoleh kesimpulan presentase performa pada metode SVM yang menggunakan data twitter melalui kata kunci vaksin covid-19.

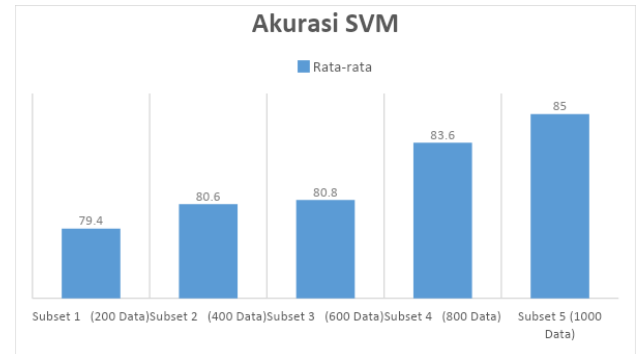
Hasil Dan Pembahasan

Pada penelitian ini akan menguji atau menganalisis algoritma Support Vector Machine. Hal yang akan di uji berupa performa akurasi, waktu training dan juga nilai MAE pada algoritma Support Vector Machine menggunakan data twitter berupa keyword vaksin covid-19.

1. Pengujian Support Vector Machine

- a. **Performa Akurasi Support Vector Machine**
 Berdasarkan hasil lima percobaan maka diperoleh hasil akurasi pada penerapan algoritma svm adalah

subset ke-1 adalah 79,40%, subset ke-2 adalah 80,60%, subset ke-3 adalah 80,80%, subset ke-4 adalah 83,60% dan subset ke-5 adalah 85,00%, maka berdasarkan hasil percobaan akurasi paling tinggi didapat pada percobaan subset ke-5 dengan nilai akurasi sebesar 85,00%.

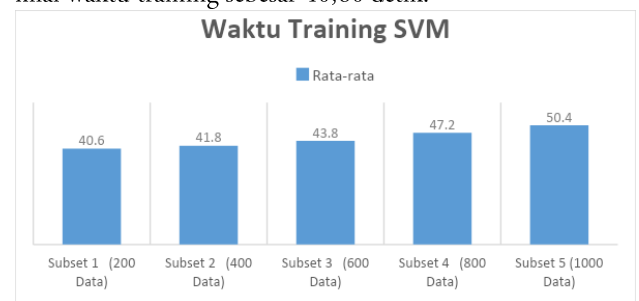


Gambar 1. Grafik Akurasi *Support Vector Machine*

Berdasarkan gambar 1 menunjukkan bahwa performa akurasi akan mengalami peningkatan pada setiap subset atau penambahan data. Oleh karena itu grafik akurasi pada penerapan algoritma Support Vector Machine secara berturut-turut dari Fold 1 s.d. 5 seperti pada Gambar 1.

b. Waktu Training Support Vector Machine

Berdasarkan hasil lima percobaan diatas maka diperoleh hasil waktu training pada penerapan algoritma svm adalah subset ke-1 adalah 40,60 detik, subset ke-2 adalah 41,80 detik, subset ke-3 adalah 43,80 detik, subset ke-4 adalah 47,20 detik dan subset ke-5 adalah 50,40 detik, maka berdasarkan hasil percobaan waktu training paling cepat didapat pada percobaan subset ke-1 dengan nilai waktu training sebesar 40,60 detik.

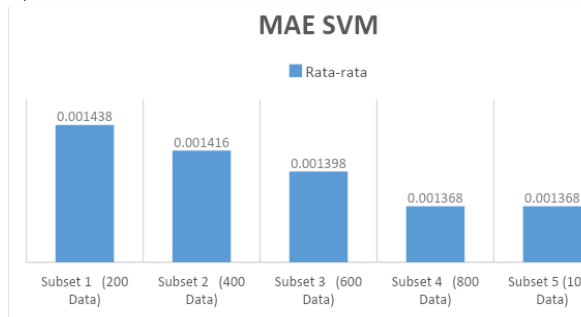


Gambar 2 Grafik Waktu training *Support Vector Machine*

Berdasarkan gambar 2 menunjukkan bahwa performa waktu training akan mengalami penurunan pada setiap subset atau penambahan data. Oleh karena itu grafik waktu training pada penerapan algoritma Support Vector Machine secara berturut-turut dari Fold 1 s.d. 5 seperti pada Gambar 2.

c. Nilai MAE Support Vector Machine

Berdasarkan hasil lima percobaan diatas maka diperoleh hasil MAE pada penerapan algoritma SVM adalah subset ke-1 adalah 0,001438, subset ke-2 adalah 0,001416, subset ke-3 adalah 0,001398, subset ke-4 adalah 0,001368 dan subset ke-5 adalah 0,001334, maka berdasarkan hasil percobaan MAE paling baik didapat pada percobaan subset ke-5 dengan nilai MAE sebesar 0,001334.



Gambar 3 Grafik MAE SVM

Berdasarkan gambar 3 menunjukkan bahwa performa MAE akan mengalami peningkatan pada setiap subset atau penambahan data. Oleh karena itu grafik MAE pada penerapan algoritma SVM sistem klasifikasi sentimen vaksin covid-19 secara berturut-turut dari Fold 1 s.d. 5 seperti pada Gambar 3.

2. Specifying learning (Representasi Hasil)

Pada bagian ini hasil percobaan dari implementasi Support vector machine akan representasi hasil. Hasil yang akan dibandingkan merupakan rata-rata terbaik dari keseluruhan percobaan pengujian dengan parameter Akurasi, Waktu training dan MAE yang telah dilakukan pada sub bab sebelumnya. Hasil nilai rata-rata performa pada percobaan terbaik dapat dilihat pada Tabel 1.

Tabel 1 Hasil Perbandingan Tingkat Performa

Algoritma	Akurasi (%)	Waktu (Detik)	MAE
SVM	85	40.60	0,001334

Berdasarkan hasil pada Tabel 1 dapat ditarik suatu kesimpulan bahwa nilai performa untuk akurasi model algoritma Support vector machine memiliki tingkat akurasi yang baik dengan nilai akurasi sebesar 85% pada konfigurasi dataset sebanyak 1000 data.

Kemudian untuk nilai tingkat performa waktu training memiliki nilai waktu sebesar 40.60 detik dengan model algoritma Support vector machine pada konfigurasi dataset sebanyak 1000 data.

Kemudian untuk nilai tingkat performa MAE Support vector machine memiliki nilai MAE sebesar 0,001334 pada konfigurasi dataset sebanyak 1000 data.

Simpulan Dan Saran

Berdasarkan hasil pengujian pada semua skenario maka didapatkan suatu temuan bahwa tingkat performa algoritma SVM pada parameter pengujian nilai tingkat performa akurasi dengan nilai 85%, serta performa waktu training dengan nilai 40,60 detik, dan parameter pengujian MAE algoritma SVM memiliki nilai MAE 0,001334.

Pengaruh jumlah records terhadap model algoritma yang digunakan memiliki pengaruh baik atau tren kenaikan terhadap nilai akurasi dan MAE dimana nilai akurasi, dan MAE akan meningkat jika jumlah records semakin bertambah, akan tetapi berdampak buruk atau tren penurunan pada kecepatan waktu training di mana waktu training akan melambat jika jumlah records semakin bertambah.

Adapun saran yang dapat digunakan oleh peneliti-peneliti selanjutnya untuk meningkatkan hasil kesimpulan pada penelitian ini, adalah sebagai berikut :

1. Pada penelitian ini tahap preprocessing dilakukan secara manual dan pada penelitian selanjutnya tahap preprocessing dapat dilakukan dengan mengimplementasikan teknik feature selection menggunakan teknik atau algoritma tertentu.
2. Pada penelitian selanjutnya dapat dilakukan pengujian dengan mengimplementasikan model-model algoritma machine learning lain serta pendekatan statistik ataupun squensial lain.

Pustaka Acuan

- Kementerian Kesehatan RI. 2020. Penetapan Jenis Vaksin Untuk Pelaksanaan Vaksinasi Coronavirus Disease 2019 (Covid-19). Kementerian Kesehatan RI Nomor HK.01.07/Menkes/9860/2020.
- Nareza, M. (2021). Jangan Termakan Isu, Ini Fakta Penting Vaksin COVID-19. <https://www.alodokter.com/>, diakses 25 Januari 2021
- World Health Organization (WHO). 2021. Coronavirus disease (COVID-19) pandemic <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>, diakses 25 Januari 2021
- Our World In Data. 2021. Coronavirus (COVID-19) Vaccinations. <https://ourworldindata.org/covid-vaccinations>, diakses 25 Januari 2021
- Global Digital Statistic. 2021. Digital, Social & Mobile in 2019. <https://wearesocial.com/blog/2021/01/digital-use-around-the-world-in-January-2020>, diakses 25 Januari 2021
- Ratino, Noor, H., Sita A., Windu, G. 2020. "Sentimen Analisis Informasi Covid-19 menggunakan Support Vector Machine dan Naïve Bayes". Jurnal JUPITER Vol. 12 No. 2 Bulan 10 Tahun 2020, Hal 1 - 11
- Rian, T., Agung, T., Ira, D. 2020. "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)". Jurnal Media Informatika Budidarma. Volume 4, Nomor 3, Juli 2020, Page 650-658.