

Klasifikasi Penjualan Kemeja Pria Menggunakan Algoritma Naïve Bayes (Studi Kasus: Toko Online Shopee)

Ilham Hardika^{*1}, Ratna Nur Tiara Shanty², Dwi Cahyono³
Program Studi Teknik Informatika Universitas Dr. Soetomo Surabaya
ilham.hardika@gmail.com¹, ratnanurtiara@unitomo.ac.id², dwik@unitomo.ac.id³

ABSTRAK

Penelitian ini membahas pembuatan sistem klasifikasi penjualan produk kemeja pria menggunakan algoritma Naïve Bayes pada platform Toko Online Shopee. Pendahuluan membahas latar belakang pentingnya strategi perusahaan dalam menghadapi kompetisi bisnis yang semakin beragam dan kompetitif, terutama dalam konteks perkembangan belanja online. Kemeja, sebagai produk yang signifikan dalam dunia fashion, menjadi fokus penelitian untuk menganalisis pola penjualan berdasarkan data transaksi. Metode penelitian menggunakan algoritma Naïve Bayes, sebuah classifier statistik yang memprediksi probabilitas kelas. Algoritma ini didasarkan pada teorema Bayes dan telah terbukti akurat dan cepat dalam mengklasifikasikan data besar. Tahap pengujian akurasi dilakukan melalui matriks confusion pada Jupyter Notebook. Hasil penelitian menunjukkan tingkat akurasi sebesar 77,8%. Kesimpulan dari penelitian ini menyatakan bahwa sistem klasifikasi menggunakan algoritma Naïve Bayes mampu mengklasifikasikan penjualan kemeja pria dengan tingkat akurasi yang baik. Respons positif juga diperoleh dari responden yang berpendapat bahwa sistem ini membantu dalam melakukan klasifikasi penjualan kemeja. Penelitian ini memberikan kontribusi penting terutama karena belum ada sistem serupa yang dapat melakukan klasifikasi produk kemeja pria. Tingkat akurasi yang dihasilkan oleh algoritma Naïve Bayes memberikan keyakinan dalam penggunaan sistem ini untuk membantu para penjual kemeja dalam menentukan model kemeja dengan tingkat penjualan yang baik.

Kata kunci: data mining, klasifikasi, kemeja pria, Naïve Bayes

This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Histori Naskah

Naskah di-Kirim : 01-12-2023

Naskah di-Revisi: 17-03-2024

Naskah di-Terima : 14-04-2024

I. PENDAHULUAN

Perkembangan dunia usaha saat ini semakin beragam dan kompetitif, laju perkembangan sangat sulit diprediksi, sehingga memiliki strategi yang tepat adalah kunci bagi suatu perusahaan agar dapat terus bersaing dan bertahan. Berkembangnya internet banyak hal yang timbul, salah satunya adalah berbelanja produk atau jasa secara online, berbelanja secara online telah menjadi alternatif cara pembelian produk atau pun jasa, penjualan secara online berkembang baik dari segi pelayanan, efektifitas, keamanan dan juga popularitas.

Kemeja adalah salah satu pakaian yang dibuat untuk menutupi lengan, dada, bahu, dan kerah. Biasanya dibuat sesuai selera pemakainya, terkadang bisa dijahit menjadi kemeja lengan panjang atau lengan pendek. Salah satu kegunaan utamanya adalah sebagai seragam kerja atau formal. Kemeja juga bisa digunakan untuk fashion favorit pria dan wanita. Oleh karena itu, penting untuk mengetahui lebih banyak tentang pakaian ini. Apalagi bagi mereka yang tidak bisa lepas dari kemeja setiap hari untuk menunjang aktivitasnya. Kemeja memiliki berbagai model untuk digunakan di berbagai acara yang berbeda. Mulai dari bahan, panjang lengan, hingga motif, sehingga dapat mewakili penampilan yang nyaman sekaligus rapi.

Penggunaan data transaksi penjualan sebenarnya dapat digunakan untuk membantu menganalisis pola pembelian yang dilakukan konsumen, bukan hanya digunakan sebagai pendataan. Dalam kumpulan data ini, sebenarnya terdapat informasi yang dapat digunakan untuk melihat pola pembelian produk yang terjadi pada setiap transaksi [1].

Terdapat beberapa penelitian yang dilakukan oleh H. D. Wijaya dan S. Dwiasnati mengangkat studi kasus tentang klasifikasi produk, salah satunya terdapat pada penelitian yaitu Sistem Klasifikasi Penjualan pada sebuah apotek yang ada di Jakarta dengan menggunakan algoritma Naïve Bayes. Penelitian ini menggunakan 150 dataset dan menghasilkan tingkat accuracy dengan nilai 88.00% [2].

Pada penelitian lainnya yang dilakukan oleh F. S. M. Kom. dan J. S. Kom tentang analisis penjualan pakaian pada perusahaan perseorangan prima jaya lestari menggunakan metode autoregressive integrated moving average (arima) untuk menganalisis penjualan yang perlu ditingkatkan pada periode berikutnya dengan menggunakan data penjualan sebelumnya [3].

Sampai saat ini belum ada sistem yang mampu melakukan klasifikasi produk kemeja pria sehingga mempengaruhi para penjual kemeja atau para pelaku usaha yang ingin memulai usaha dalam bidang penjualan kemeja. Dalam hal ini sistem

klasifikasi kemeja pria sangat dibutuhkan guna menentukan model kemeja yang memiliki tingkat penjualan yang tinggi ataupun rendah.

Dalam menentukan klasifikasi produk pada sistem ini menggunakan algoritma naïve bayes. Algoritma Naive Bayes adalah klasifikasi statistik yang dapat digunakan untuk memprediksi probabilitas milik suatu kelas [2]. Algoritma naïve bayes termasuk dalam salah satu algoritma data mining. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstrak dan mengidentifikasi informasi yang berguna dan pengetahuan terkait dari sejumlah data yang besar [4].

Berdasarkan uraian persoalan yang dihadapi, maka penelitian ini bertujuan untuk membuat sistem klasifikasi penjualan produk kemeja pria menggunakan algoritma naïve bayes yang berguna bagi para penjual kemeja dalam menentukan model kemeja yang memiliki tingkat penjualan tinggi ataupun rendah.

Yang menjadi batasan dalam pembuatan aplikasi ini adalah :

1. Dataset yang penulis analisa ini difokuskan pada data penjualan kemeja pria pada toko online Shopee M231 Official Store.
2. Dataset yang penulis analisa ini difokuskan untuk Data mining klasifikasi penjualan Kemeja dengan variabel meliputi bahan, warna kemeja, tipe kerah, tipe lengan, motif, posisi motif, warna motif, dan harga.serta data target berupa klasifikasi penjualan.
3. Penelitian ini menggunakan teknik scrapping dalam pengambilan data pada website Shopee.
4. Penelitian ini menggunakan Algoritma Naïve Bayes.
5. Aplikasi ini hanya dibuat dengan basis WEB dengan menggunakan framework Flask dan Jupyter Notebook untuk merancang pembuatan web dengan bahasa Python, HTML dan CSS.

Sehingga mendapatkan tujuan dalam pembuatan aplikasi ini yaitu membuat sistem klasifikasi penjualan produk kemeja pria yang memiliki kemampuan untuk melakukan klasifikasi penjualan kemeja pria dan melihat hasil klasifikasi yang dibuat dalam bentuk website

Diharapkan sistem klasifikasi penjualan produk kemeja pria ini dapat memberikan banyak manfaat antara lain adalah pengguna dapat melakukan dan menampilkan hasil klasifikasi dengan lebih cepat dan tepat serta sistem ini dibangun dalam bentuk website sehingga mempermudah pengguna dalam mengaksesnya

II. METODE

A. Deskripsi Sistem

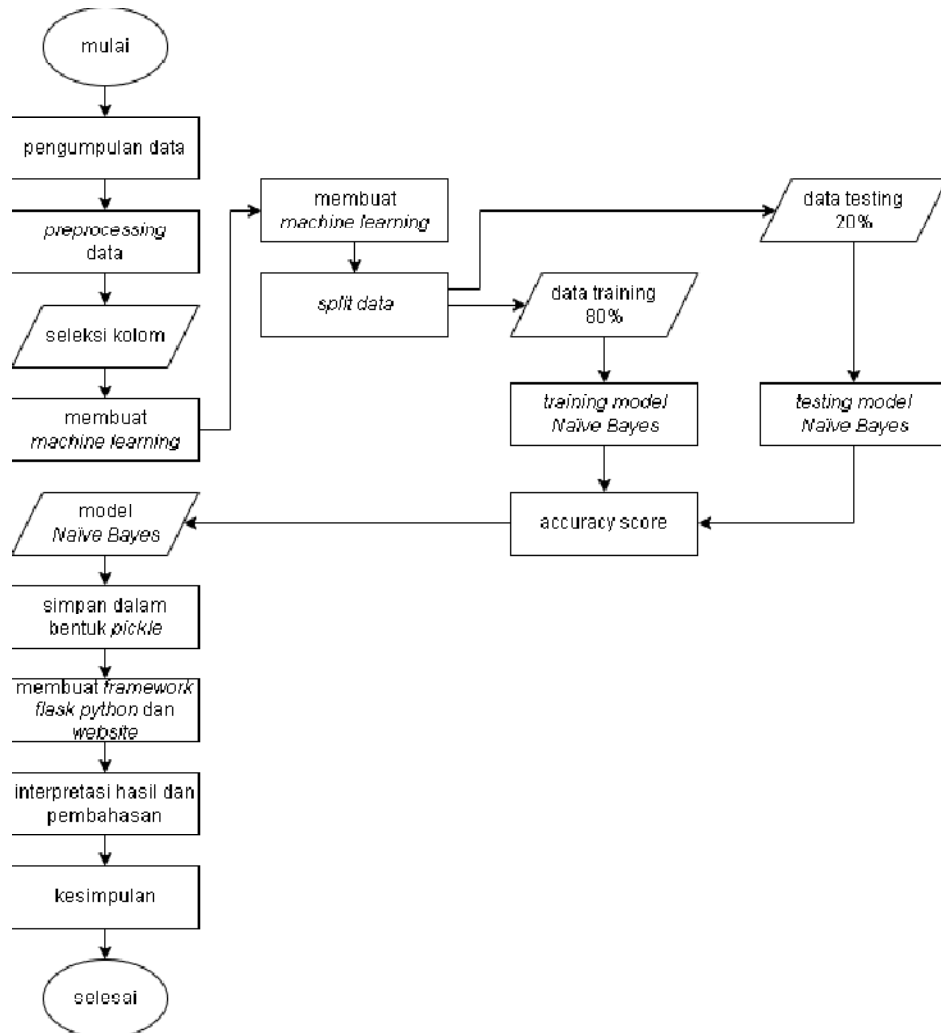
Pada penelitian ini, kami mengimplementasikan Klasifikasi Penjualan Produk Kemeja berbasis Web. Sistem ini dirancang untuk menentukan klasifikasi penjualan produk kemeja pada platform toko online Shopee. Fokus utama sistem ini adalah membantu menentukan model kemeja yang memiliki tingkat penjualan tinggi atau rendah di pasar. Dalam konteks kesulitan klasifikasi produk kemeja, terutama karena banyaknya model serupa, sistem ini memanfaatkan data penjualan dan atribut-atribut seperti jenis bahan, warna, tipe kerah, tipe lengan, motif, posisi motif, warna motif, dan harga. Dengan menggunakan data target berupa total penjualan, diharapkan sistem dapat membantu pengelompokkan produk kemeja berdasarkan atribut-atribut yang berpengaruh terhadap tingkat penjualan. Alur System ini dapat dilihat pada Gambar

bahan	warna_kemeja	tipe_kerah	tipe_lengan	motif	posisi_motif	warna_motif	harga
Katun Oxford	Hijau	Straight Collar	Lengan Panjang	Polos	Tidak Ada	Tidak Ada	149000
Katun Oxford	Army	Straight Collar	Lengan Panjang	Polos	Tidak Ada	Tidak Ada	149000
Katun Oxford	Coklat Muda	Straight Collar	Lengan Pendek	Polos	Tidak Ada	Tidak Ada	129000
Katun Oxford	Coklat	Straight Collar	Lengan Panjang	Polos	Tidak Ada	Tidak Ada	149000
Katun Oxford	Khaki	Straight Collar	Lengan Panjang	Polos	Tidak Ada	Tidak Ada	149000
...
Katun	Coklat	Straight Collar	Lengan Panjang	Batik	Area Bawah	Krem Coklat	169000
Katun	Navy	Straight Collar	Lengan Panjang	Batik	Area Atas	Krem Coklat	169000
Katun	Coklat	Straight Collar	Lengan Panjang	Batik	Full	Krem Hitam	169000
Katun	Putih	Straight Collar	Lengan Panjang	Batik	Full	Coklat Hitam	169000
Katun	Biru	Straight Collar	Lengan Panjang	Batik	Full	Putih Hitam	169000

Gambar 2. Data yang diambil

B. Pengambilan Data

Data yang diambil difokuskan pada data penjualan kemeja pada toko online Shopee M231 meliputi atribut bahan, warna kemeja, tipe kerah, tipe lengan, motif, posisi motif, warna motif, dan harga sebagai atribut input serta status tingkat penjualan sebagai target yang memiliki value tinggi dan rendah. Pada Gambar 2 adalah contoh sampel data yang telah diambil menggunakan metode scrapping, yang memiliki jumlah 403 data dan 9 kolom seperti pada Gambar 2



Gambar1. Flowchart

C. Perancangan System Naïve Bayes

Dari Gambar 3 dapat diketahui bahwa terdapat dua alur proses untuk model klasifikasi yaitu data training dan data testing. Pada dasarnya data training dan data testing memiliki alur hampir sama. Data yang telah di training akan melalui proses algoritma klasifikasi Naïve Bayes dan model klasifikasi akan disimpan yang kemudian akan di uji coba menggunakan data testing dengan data yang baru berdasarkan pelatihan model dari data yang lama. Terdapat 7 tahapan yaitu :

1. Open Data

Open data adalah proses dimana data akan diload

2. Preprocessing

Terdapat banyak teknik untuk melakukan persiapan data sebelum dilakukan implementasi klasifikasi, salah satu teknik yang digunakan dalam penelitian ini adalah handling missing value atau penanganan pada nilai yang kosong

3. Algoritma Naïve Bayes

Algoritma Naïve Bayes dapat melakukan perhitungan yang efisien menangani atribut tipe diskrit dan tipe diskrit-numerik, dengan cara melakukan setiap probabilitas

4. Test & Score

Yaitu tahap dimana kita dapat melihat nilai Akurasi, Presisi, dan Recall dari data yang telah dilakukan proses Klasifikasi

5. Save Model

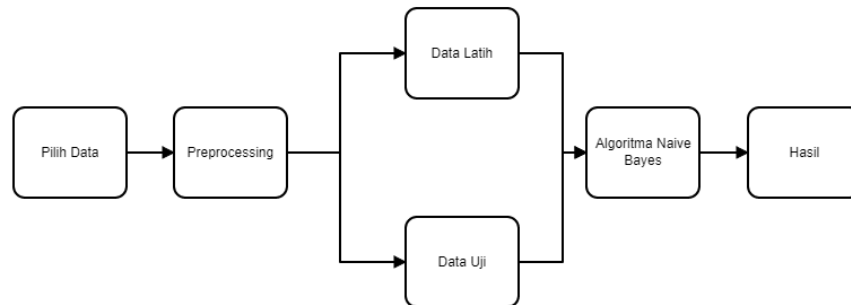
Adalah teknik untuk meyimpan model dari alur data yang telah di training

6. Aply Model

Menerapkan model yang telah disimpan

7. Result

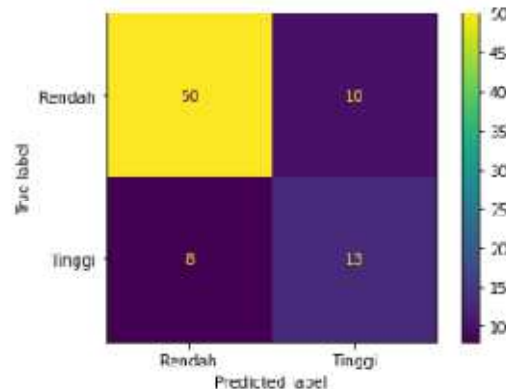
Tahap untuk melihat hasil prediksi dari data yang telah di implementasikan menggunakan modelr



Gambar 2. Perancangan System Naïve Bayes

D. Uji Akurasi

Metode pengukuran keakuratan sistem dilakukan melalui dua pendekatan, yakni menggunakan aplikasi Jupyter Notebook. menggunakan metrik akurasi (accuracy_score) untuk menghitung seberapa sering hasil yang dihasilkan oleh sistem cocok dengan data yang sebenarnya. Metrik ini dihitung dengan membagi jumlah keputusan yang benar oleh sistem dengan total keputusan yang dibuat. Selain itu, pengujian akurasi juga menggunakan metode Confusion Matrix untuk mengevaluasi performa model klasifikasi, menampilkan jumlah prediksi yang benar dan salah.



Gambar 3. Confusion Matrix Python

Fungsi ini akan mengembalikan objek yang digunakan untuk menampilkan confusion matrix dari suatu model klasifikasi. Objek ini dapat digunakan untuk menampilkan confusion matrix dalam bentuk visual. Setelah itu dilakukan pemanggilan method `.plt.show()` pada objek `cm_display` untuk menampilkan visualisasi dari confusion matrix seperti pada Gambar 4.

Setelah menjalankan kode di atas, maka diperoleh hasil akurasi dari model yang telah dilatih dengan data latih dan diuji dengan data uji, yaitu sebesar 0,778 atau 77,8%. Nilai akurasi tersebut menunjukkan seberapa baik model dapat memprediksi kelas dari data uji dengan benar.

III. HASIL DAN PEMBAHASAN

3.1. Implementasi Program

Implementasi sistem klasifikasi kemeja pria berbasis website berdasarkan rancangan sistem dan desain yang dibangun menggunakan HTML, CSS, Python dan Framework Flask adalah sebagai berikut:

3.1.1. Tampilan Halaman Home

Sebagai halaman pertama yang dilihat user, halaman home memberikan pengalaman dan tampilan yang baik bagi user, halaman ini berisi informasi penting mengenai website, seperti deskripsi, navigasi, dan tautan ke halaman lain. serta memperhatikan desain dan konten halaman home, agar dapat memberikan pengalaman yang baik bagi user seperti yang dapat dilihat pada Gambar 6..

```
rb_clf = GaussianNB()
rb_clf.fit(X_train, y_train)
y_pred = rb_clf.predict(X_test)
round(accuracy_score(y_test, y_pred), 3)

0.778
```

Gambar 4. Sintaks Akurasi Naive Bayes menggunakan Python

3.1.2. Tampilan Halaman Input Data

Pada halaman penilaian yang ditunjukkan pada Gambar 7 user akan diminta untuk menginput data atribut kemeja meliputi data bahan, warna kemeja, tipe kerah, tipe lengan, motif, posisi motif, warna motif, dan harga yang diperlukan dan dapat menekan tombol prediksi untuk masuk proses analisa.

3.1.3 Tampilan Halaman Output Data

Pada Gambar 3.3 ditunjukkan hasil dari pengujian data yang telah dilakukan input oleh user, sistem akan melakukan pengolahan data secara otomatis menggunakan algoritma naïve bayes sebagai algoritma klasifikasi dengan cara menghitung probabilitas setiap kelas dan menghasilkan hasil prediksi dengan kategori tingkat penjualan rendah dan tingkat penjualan tinggi.

3.1.4 Tampilan Halaman Input File

Pada halaman input file yang ditunjukkan pada Gambar 9 user akan diminta untuk menginput dile csv berisikan atribut kemeja meliputi data bahan, warna kemeja, tipe kerah, tipe lengan, motif, posisi motif, warna motif, dan harga yang diperlukan dan dapat menekan tombol prediksi untuk masuk proses analisa.

3.1.5 Tampilan Halaman Output File

Pada Gambar 10 ditunjukkan hasil dari pengujian data yang telah dilakukan input oleh user, sistem akan melakukan pengolahan data secara otomatis menggunakan algoritma naïve bayes sebagai algoritma klasifikasi dengan cara menghitung probabilitas setiap kelas dan menghasilkan hasil akurasi dan prediksi dengan kategori tingkat penjualan rendah dan tingkat penjualan tinggi secara massive dalam bentuk tabel.

IV. KESIMPULAN

Kesimpulan dari pengujian sistem Klasifikasi Penjualan Kemeja Pria dengan menggunakan Algoritma Naïve Bayes (Studi Kasus: Toko Online Shopee) dapat diuraikan sebagai berikut. Sistem mampu mengklasifikasi prediksi penjualan kemeja pria dengan tingkat akurasi sebesar 77,8% pada Jupyter Notebook, berdasarkan 403 data termasuk dalam sistem. Selain itu, hasil uji penerimaan menunjukkan bahwa 84,4% responden berpendapat bahwa sistem ini memberikan kontribusi positif dalam melakukan klasifikasi penjualan kemeja. Dengan demikian, penggunaan algoritma Naïve Bayes dalam klasifikasi penjualan kemeja pria di Shopee terbukti efektif dan dapat membantu para penjual dalam menentukan model kemeja yang berpotensi memiliki tingkat penjualan yang baik.

DAFTAR PUSTAKA

- [1] R. I. M. a. M. N. S. H. A. ichlas, "Aturan Asosiasi untuk Analisis Data Penjualan Produk Menggunakan Algoritma Apriori," *Jurnal Teknik Informatika Unika St. Thomas (JTIUST)*, pp. vol. VI, no. 1, pp. 53-60, 2021.
- [2] H. D. W. a. S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes," *JURNAL INFORMATIKA*, pp. vol. VII, no. 1, pp. 1-7, 2020.
- [3] F. S. M. a. J. S.kom., "ANALISIS PENJUALAN PAKAIAN PADA PERUSAHAAN PERSEORANGAN PRIMA JAYA LESTARI MENGGUNAKAN METODE AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)," *Jurnal Penelitian Teknik Informatika Universitas Prima Indonesia (UNPRI)*, pp. vol. III, no. 2, pp. 496-502, 2020.
- [4] P. B. a. Y. C. H. S. A. K. Lalo, "Implementasi Algoritma C4.5 Untuk Klasifikasi Penjualan Barang di Swalayan Dutalia," *Jurnal Teknik Informatika Unika St. Thomas (JTIUST)*, pp. vol. VI, no. 01, pp. 1-12, 2021. , 2021. .
- [5] M. Sholihin, "Implementasi K-Nearest Neighbor Untuk Klasifikasi Grade Telur Ayam Ras Berdasarkan Ukuran," *INFORM : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, p. Vol. 4 No. 2, 2019.
- [6] N. N. D. K. R. H. Talitha Syahla Janiar Arifin, "The Level of Student Satisfaction with the Online Learning Process During a Pandemic Using the K-means Algorithm," *Inform : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, pp. Vol.6, No.2 ,pp. 123-126, 2021.
- [7] D. R. S. S. a. B. W. P. B. N. Setio, "Klasifikasi dengan Pohon Keputusan Berbasis Algoritme C4.5," *Prosiding Seminar Nasional Matematika*, pp. vol. III, pp. 64-71, 2020.
- [8] N. R. P. P. S. a. M. R. A. Rinaldi, "Analisa Pengambilan Keputusan Pemilihan Bahan Dalam Pembuatan Kemeja Menggunakan Metode TOPSIS.," *Jurnal Sains, Teknologi dan Industri*, pp. vol. XVII, no. 2, pp. 163-172, 2021.
- [9] M. Djufri, "PENERAPAN TEKNIK WEB SCRAPING UNTUK PENGGALIAN POTENSI PAJAK (Studi Kasus pada Online Market Place Tokopedia, Shopee dan Bukalapak)," *urnal BPPK*, pp. vol. 13, no. 2, pp. 65-75, 2020.
- [10] R. A. R. a. R. V. N. L. P. C. Savitri, "Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning.," *Jurnal Teknik Informatika dan Sistem Informasi*, pp. vol. VII, no. 1, pp. 47-58, 2021.
- [11] A. U. T. A. J. L. a. N. E. P. B. G. Kaitelapatay, "ANALISIS SENTIMEN PENGARUH PERISTIWA WADAS TERHADAP GANJAR PRANOWO MENGGUNAKAN SUPPORT VECTOR MACHINE PADA TWITTER," *Jurnal Ilmiah Indonesia*, pp. vol. VIII, no. 12, pp. 32-49, 2022.
- [12] W. R. a. E. Harahap, "Pengolahan Data Hasil Penjualan Online Menggunakan Aplikasi Microsoft Excel.," *Jurnal Matematika*, pp. vol. XX, no. 2, pp. 69-76, 2021.
- [13] H. & W. T. Indriyawati, "Pemodelan Data Mining Pola Kelayakan Kemampuan Lulusan Dengan Kebutuhan Stakeholder Menggunakan Algoritma Apriori," *JITSI : Jurnal Ilmiah Teknologi Sistem Informasi*, vol. II, no. 3, pp. 78-84, 2021.