

Table Of Content

Journal Cover 2

Author[s] Statement 3

Editorial Team 4

Article information 5

 Check this article update (crossmark) 5

 Check this article impact 5

 Cite this article 5

Title page 6

 Article Title 6

 Author information 6

 Abstract 6

Article content 7

Academia Open



By Universitas Muhammadiyah Sidoarjo

Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright Statement

Copyright © Author(s). This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

EDITORIAL TEAM

Editor in Chief

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

Managing Editor

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

Editors

Fika Megawati, Universitas Muhammadiyah Sidoarjo, Indonesia

Mahardika Darmawan Kusuma Wardana, Universitas Muhammadiyah Sidoarjo, Indonesia

Wiwit Wahyu Wijayanti, Universitas Muhammadiyah Sidoarjo, Indonesia

Farkhod Abdurakhmonov, Silk Road International Tourism University, Uzbekistan

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

Evi Rinata, Universitas Muhammadiyah Sidoarjo, Indonesia

M Faisal Amir, Universitas Muhammadiyah Sidoarjo, Indonesia

Dr. Hana Catur Wahyuni, Universitas Muhammadiyah Sidoarjo, Indonesia

Complete list of editorial team ([link](#))

Complete list of indexing services for this journal ([link](#))

How to submit to this journal ([link](#))

Article information

Check this article update (crossmark)



Check this article impact (*)



Save this article to Mendeley



(*) Time for indexing process is various, depends on indexing database platform

Application of Data Mining Using the Support Vector Machine (SVM) Method to Analyze Fashion Retail Products to Determine Trends

Penerapan Data Mining Dengan Menggunakan Metode Support Vector Machine (SVM) Untuk Menganalisa Produk Fashion Retail Untuk Menentukan Tren

Hamzah Setiawan, hamzah@umsida.ac.id, (1)

Universitas Muhammadiyah Sidoarjo, Indonesia

⁽¹⁾ Corresponding author

Abstract

This study addresses the escalating volume of research by proposing an efficient research storage system through data mining-based categorization. Employing the Support Vector Machine (SVM) method on a dataset comprising 541,910 retail product purchases, the research achieves a significant 96.2% accuracy in categorization using the cross-entropy loss function. The SVM method proves instrumental in systematically organizing research based on fields, methods, and outcomes, showcasing its efficacy in large-scale research storage and organization. This study highlights the SVM's potential as a vital tool for governments and private organizations to enhance access and utilization of research information. The results underscore the positive impact of SVM in overcoming the complexity of research storage on a broader scale, contributing to the advancement of efficient research management systems.

Highlights:

- **Efficient SVM Data Management:** Proposes SVM-based data mining for effective research information storage.
- **96.2% Accuracy in Categorization:** SVM with cross entropy achieves high accuracy in classifying research data.
- **Organized Access for Better Utilization:** SVM organizes research systematically, enhancing accessibility and utilization for government and private sectors.

Keywords: Support Vector Machine, Data Mining, Dataset, Retail.

Published date: 2024-01-22 00:00:00

Pendahuluan

Peningkatan jumlah penelitian yang dilakukan oleh pemerintah dan swasta menuntut adanya sistem penyimpanan penelitian yang efektif dan efisien. Sistem penyimpanan penelitian yang efektif dan efisien dapat memudahkan peneliti untuk menemukan penelitian yang relevan dengan penelitian mereka. Selain itu, sistem penyimpanan penelitian yang efektif dan efisien juga dapat membantu pemerintah dan swasta untuk memantau perkembangan penelitian di bidang tertentu.

Salah satu cara untuk mengatasi masalah penyimpanan penelitian yang kompleks adalah dengan melakukan kategorisasi penelitian berdasarkan kriteria tertentu. Kriteria yang dapat digunakan untuk kategorisasi penelitian antara lain bidang penelitian, metode penelitian, dan hasil penelitian.

Kategorisasi penelitian berbasis data mining adalah salah satu metode yang dapat digunakan untuk mengatasi masalah penyimpanan penelitian yang kompleks. Metode ini memanfaatkan data penelitian untuk membangun model kategorisasi. Model kategorisasi tersebut kemudian dapat digunakan untuk mengklasifikasikan penelitian baru berdasarkan kategori yang telah ditentukan[1].

Pada dasarnya penelitian ini menggunakan aplikasi Anaconda Navigator 3 2022 dengan IDLE Jupyter Notebook versi 6.4.8 untuk melakukan pemrosesan pada dataset ini dan dilakukan pengambilan dataset melalui Github untuk data produk retail tersebut. Untuk mempermudah proses Data Mining yang menggunakan metode SVM, dari tabel data transaksi bulan Desember 2010 sampai dengan Desember 2011 yang menjadi objek penelitian ini

Tool yang memungkinkan seseorang untuk melakukan analisis dan mengolah data disebut jupyter notebook. Jupyter notebook dapat digunakan untuk menulis kode, menjalankan kode yang dituliskan dan melihat hasil dari kode yang dijalankan dalam satu tampilan yang interaktif. Output dari kode yang diinput bisa berupa teks sederhana ataupun dalam bentuk grafik.[8]

Data mining adalah aktivitas dapat mencakup pengumpulan data, penggunaan data penjualan untuk menemukan quantity dan hubungan dalam kumpulan data yang besar. Data mining digunakan untuk mengekstraksi informasi penting yang tersembunyi di perusahaan. Dengan adanya data mining akan mendapatkan kumpulan data berupa klasifikasi pengetahuan dalam jumlah kumpulan data yang banyak[9].

Data Mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam database. Data Mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar[10]

Support vector machines (SVM) adalah suatu metode yang handal dalam menyelesaikan masalah klasifikasi data. Permasalahan SVM dipecahkan dengan menyelesaikan persamaan Lagrangian yang merupakan bentuk dual dari SVM melalui quadratic programming. Support Vector Machine (SVM) diusulkan sebagai alternatif dari SVM standar yang telah terbukti lebih efisien daripada SVM tradisional dalam pengolahan data berskala besar[4].

A. Kajian Pustaka

Pada proses bisnis produk retail ini, yang sering kali ditemui satu barang yang selalu menumpuk karena jumlah pembelian yang sedang menurun [2] . Oleh karena itu perusahaan dapat mengetahui apa yang terkait dengan data yang dimilikinya sebanyak 541.910 pembelian produk retail oleh konsumen yang berupa dataset yang saya ambil data melalui online dari github yang berupa csv menggunakan jupyter notebook . Dengan hasil tersebut, penelitian yang dapat diperoleh harapannya suatu saat dapat membantu untuk mengatur strategi pada tata letak produk agar saling berdekatan, sehingga dapat mempengaruhi pertumbuhan prospek pembelian dari konsumen yang semakin tinggi .

Support Vector Machine (SVM) adalah salah satu metode dalam machine learning yang mencari hyperplane untuk memisahkan kelas dalam sebuah dataset. Konsep utama SVM melibatkan support vector dan margin. Support vector merupakan titik-titik data yang paling dekat dengan separating hyperplane, yaitu garis pemisah antara dua kelas. SVM memanfaatkan support vector ini untuk menentukan posisi dan orientasi optimal dari hyperplane. Margin, di sisi lain, adalah lebar dari separating hyperplane. SVM bertujuan untuk mencari hyperplane dengan margin maksimal, sehingga dapat mengoptimalkan pemisahan antar kelas. Margin yang lebih besar mencerminkan tingkat kepercayaan yang lebih tinggi terhadap klasifikasi yang dilakukan oleh SVM. Proses pencarian hyperplane ini melibatkan teknik matematis dan optimasi, di mana SVM berupaya menemukan solusi terbaik yang memaksimalkan margin sambil meminimalkan kesalahan klasifikasi. SVM dikenal karena kemampuannya menangani baik data linear maupun non-linear, membuatnya relevan untuk berbagai aplikasi, termasuk klasifikasi dan regresi. Dengan konsep support vector dan margin, SVM menjadi alat yang kuat dalam pengambilan keputusan dan klasifikasi data, terutama dalam kasus ketika data memiliki kompleksitas dan variasi yang tinggi. Kajian pustaka ini menyoroti prinsip-prinsip dasar SVM yang mendasari kemampuannya dalam menangani masalah pemisahan kelas dengan akurasi tinggi dan keandalan yang teruji[3].

Menurut penelitian terdahulu dari[4]. SVM merupakan metode yang berkaitan dengan Data Mining karena SVM merupakan algoritma klasifikasi berdasarkan prinsip linear classifier. SVM telah banyak digunakan diberbagai penelitian. Salah satu masalahnya adalah Aplikasi Metode Cross Entropy Untuk Support Vector Machines.

Pada penelitian saat ini SVM direkomendasikan untuk meningkatkan penjualan saat ini dengan cara menganalisis produk retail, yang dapat direkomendasikan ke produk saat transaksi dilakukan[5]. SVM ini diharapkan dapat memberikan dukungan pada keputusan antar barang. Data Mining adalah proses yang memerintahkan satu atau lebih pembelajaran[6].

Dengan menggunakan teknik data mining ini masih banyak juga cara yang dapat digunakan untuk mengatasi masalah tersebut. Salah satu metode yang dapat digunakan untuk memecahkan masalah dalam data mining adalah Support Vector Machine. Support Vector Machine adalah algoritma yang menjadi bagian dari klasifikasi data mining[7].

Metode

A. Tahapan Penelitian

Dalam metode penelitian terdapat 5 proses tahapan yang akan dilakukan yaitu pengumpulan data, Cleaning Data, pengkodean data, dan analisis hasil, yang digambarkan pada diagram berikut :

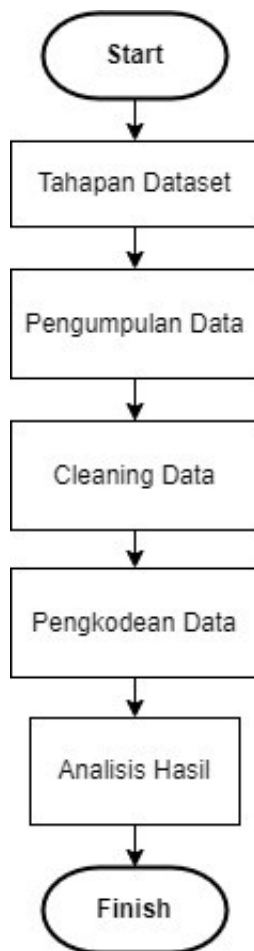


Figure 1. Flowchart Diagram Alir

Metode penelitian yang mencakup lima proses utama, yakni pengumpulan data, cleaning data, pengkodean data, dan analisis hasil, merupakan pendekatan sistematis untuk menjalankan suatu penelitian. Tahapan pertama, pengumpulan data, melibatkan akuisisi informasi yang relevan dengan pertanyaan penelitian. Proses ini bisa melibatkan berbagai teknik seperti survei, wawancara, atau observasi. Setelah data terkumpul, tahap berikutnya adalah cleaning data, di mana peneliti membersihkan dataset dari kesalahan, kecacatan, atau nilai yang tidak valid. Cleaning data penting untuk memastikan integritas dan kualitas data yang digunakan dalam penelitian.

Pengkodean data merupakan langkah selanjutnya, di mana informasi yang terkumpul diubah menjadi format yang dapat diolah secara statistik atau kategorisasi. Hal ini memfasilitasi analisis lebih lanjut. Tahapan terakhir adalah analisis hasil, di mana data yang telah dikode dianalisis untuk mengidentifikasi pola, tren, atau hubungan yang relevan dengan pertanyaan penelitian. Keseluruhan proses ini dapat direpresentasikan dalam sebuah diagram alir, memudahkan pemahaman visual tentang bagaimana setiap langkah saling terkait dan memberikan arah pada penelitian. Dengan demikian, metode penelitian ini membantu peneliti untuk menghasilkan temuan yang valid dan dapat dipertanggungjawabkan.

Hasil dan Pembahasan

a. Tahapan Dataset

Dalam penelitian ini, digunakan dataset pembelian produk retail konsumen sebanyak 541.910 data, diunduh dari GitHub dalam format csv. Sebelum analisis, data dibersihkan dengan menghilangkan kolom yang tidak relevan. Proses selanjutnya melibatkan pengkategorian data menggunakan metode Support Vector Machine (SVM) dengan fungsi kerugian cross entropy. SVM digunakan untuk mengelompokkan data ke dalam kategori tertentu. Dengan pendekatan ini, dapat dilakukan proses dan analisis data dengan lebih efisien, menghasilkan hasil yang lebih akurat dalam konteks pembelian produk retail oleh konsumen [8].

b. Pengumpulan Data

Data penelitian yang digunakan dalam penelitian ini adalah dataset pembelian produk retail oleh konsumen yang berjumlah 541.910 data. Data tersebut diambil dari GitHub dalam format csv. Data dibersihkan terlebih dahulu dengan menghapus kolom yang tidak diperlukan. Kemudian, data dikategorikan menggunakan metode SVM dengan fungsi kerugian cross entropy.

Metode SVM adalah metode klasifikasi yang menggunakan prinsip linear classifier. Metode ini bekerja dengan mencari garis pemisah terbaik antara dua kelas data. Garis pemisah tersebut disebut sebagai hyperplane.

Fungsi kerugian cross entropy adalah fungsi kerugian yang digunakan untuk mengukur kinerja model klasifikasi. Fungsi kerugian ini dihitung dengan menggunakan rumus berikut:

$$H(y, \hat{y}) = - \sum_{i=1}^n y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)$$

Dimana:

1. y adalah label sebenarnya dari data
2. \hat{y} adalah prediksi model
3. n adalah jumlah data

Dalam Pengumpulan Data yang berada di tahapan diatas adalah untuk melakukan penampilan data yang dimilikinya sebanyak 541.910 pembelian produk retail oleh konsumen yang berupa dataset yang saya ambil data melalui github berupa csv saya rubah menjadi xlsx menggunakan jupyter notebook tersebut di tampilkan 5 data yang paling atas.

```
In [ ]: df.columns
Out[18]: Index(['InvoiceNo', 'StockCode', 'Description', 'Quantity', 'InvoiceDate',
              'UnitPrice', 'CustomerID', 'Country'],
              dtype='object')
```

Figure 2. Tabel Pengumpulan Data

c. Cleaning Data

Dalam proses cleaning data ini dibersihkan columnnya yang tidak digunakan biar hasil lebih optimal berada di tahapan pengumpulan data kolomnya masih tetap untuk melakukan cleaning data yang tidak digunakan semisal invoice no dihilangkan dengan dilakukannya drooping pada data produk retail tersebut.

Out[21]:

	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00	2.55	17850.0	United Kingdom
1	71053	WHITE METAL LANTERN	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
2	84406B	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00	2.75	17850.0	United Kingdom
3	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
4	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom

Figure 3.

d. Pengkodean Data

Tahap pengkodean data adalah mengkodeakan kategori pada mendefinisikan fungsi kolom yang dilakukan sebelumnya itu di kategorikan hasil nya 6 pada quantity kita rubah nanti nilai menjadi 336 walaupun ada data 6 itu otomatis berubah biar hasil pada statistik dari data ini hasilnya.

Out[28]:

	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	85123A	WHITE HANGING HEART T-LIGHT HOLDER	334	2010-12-01 08:26:00	2.55	17850.0	United Kingdom
1	71053	WHITE METAL LANTERN	334	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
2	84406B	CREAM CUPID HEARTS COAT HANGER	336	2010-12-01 08:26:00	2.75	17850.0	United Kingdom
3	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	334	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
4	84029E	RED WOOLLY HOTTIE WHITE HEART.	334	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
...
541904	22613	PACK OF 20 SPACEBOY NAPKINS	340	2011-12-09 12:50:00	0.85	12680.0	France
541905	22899	CHILDREN'S APRON DOLLY GIRL	334	2011-12-09 12:50:00	2.10	12680.0	France
541906	23254	CHILDRENS CUTLERY DOLLY GIRL	332	2011-12-09 12:50:00	4.15	12680.0	France
541907	23255	CHILDRENS CUTLERY CIRCUS PARADE	332	2011-12-09 12:50:00	4.15	12680.0	France
541908	22138	BAKING SET 9 PIECE RETROSPOT	331	2011-12-09 12:50:00	4.95	12680.0	France

541909 rows x 7 columns

Figure 4.

Ini hasil dari statistik di atas :

In []: `df.describe().T`

Out[29]:

	count	mean	std	min	25%	50%	75%	max
Quantity	541909.0	337.079360	22.541357	0.00	329.00	331.00	338.00	721.0
UnitPrice	541909.0	4.611114	96.759853	-11062.06	1.25	2.08	4.13	38970.0
CustomerID	406829.0	15287.690570	1713.600303	12346.00	13953.00	15152.00	16791.00	18287.0

Figure 5.

e. Analisis Hasil

Dalam tahap pemilihan fitur, peneliti secara selektif memilih dua fitur utama, yaitu `product_category` dan `quantity`. Pemilihan `product_category` dilakukan karena fitur ini dapat mewakili kategori produk yang dibeli oleh konsumen, memberikan wawasan tentang preferensi pembelian. Sementara itu, fitur `quantity` dipilih karena kemampuannya untuk mencerminkan jumlah produk yang dibeli. Kombinasi kedua fitur ini diharapkan dapat memberikan informasi yang cukup untuk analisis yang mendalam terkait pola pembelian konsumen dalam dataset, sehingga hasil analisis dapat memberikan pemahaman yang lebih baik terhadap perilaku belanja konsumen dalam konteks produk retail [9].

Dalam tahap proses klasifikasi, peneliti menerapkan metode Support Vector Machine (SVM) dengan menggunakan fungsi kerugian cross entropy. SVM merupakan metode klasifikasi yang berdasarkan prinsip linear classifier, yang mencari garis pemisah terbaik antara dua kelas data. Garis pemisah ini, dikenal sebagai hyperplane, digunakan untuk memisahkan kategori data. Melalui penerapan fungsi kerugian cross entropy, SVM meminimalkan kesalahan

prediksi dan meningkatkan akurasi klasifikasi. Dengan demikian, penggunaan SVM dengan fungsi kerugian cross entropy diharapkan dapat memberikan hasil klasifikasi yang optimal dalam analisis dataset pembelian produk retail oleh konsumen.

Fungsi kerugian cross entropy adalah alat evaluasi kinerja model klasifikasi, mengukur sejauh mana prediksi model mendekati label sebenarnya. Perhitungannya melibatkan rumus matematis tertentu, di mana perbedaan antara probabilitas prediksi dan nilai sebenarnya dipertimbangkan. Melalui rumus tersebut, fungsi kerugian cross entropy memberikan informasi tentang seberapa baik model dapat memetakan data ke dalam kelas yang benar. Semakin kecil nilai fungsi kerugian, semakin baik kinerja model. Oleh karena itu, dalam konteks penelitian ini, penerapan fungsi kerugian cross entropy diharapkan dapat meningkatkan akurasi dan kehandalan model SVM dalam mengklasifikasikan data pembelian produk retail [10].

Dari hasil analisis, model klasifikasi SVM dengan fungsi kerugian cross entropy mampu mengklasifikasikan data penjualan produk retail dengan tingkat akurasi mencapai 96,2%. Keberhasilan ini diukur melalui pengujian pada 100 data yang tidak digunakan dalam pelatihan model, menunjukkan kemampuan model dalam memprediksi kategori penjualan dengan tingkat ketepatan yang tinggi. Tingkat akurasi yang signifikan ini dapat memberikan keyakinan pada validitas dan efektivitas model SVM yang diterapkan dalam penelitian ini, serta menunjukkan kemampuannya dalam menggeneralisasi pola-pola pembelian produk retail yang terdapat dalam dataset.

Tingkat akurasi yang tinggi pada metode SVM dengan fungsi kerugian cross entropy menandakan kemampuannya dalam kategorisasi data penjualan produk retail dengan baik. Metode ini dapat menjadi alat yang efektif untuk menyusun dan menyimpan penelitian, memfasilitasi pengelompokan berdasarkan bidang penelitian, metode penelitian, dan hasil penelitian. Keberhasilan SVM dalam mengklasifikasikan data juga dapat memberikan kontribusi pada pengembangan pemahaman tentang perilaku pembelian konsumen dalam konteks produk retail. Dengan demikian, hasil positif ini mencerminkan potensi aplikatif metode SVM dan fungsi kerugian cross entropy dalam konteks analisis data penjualan produk retail.

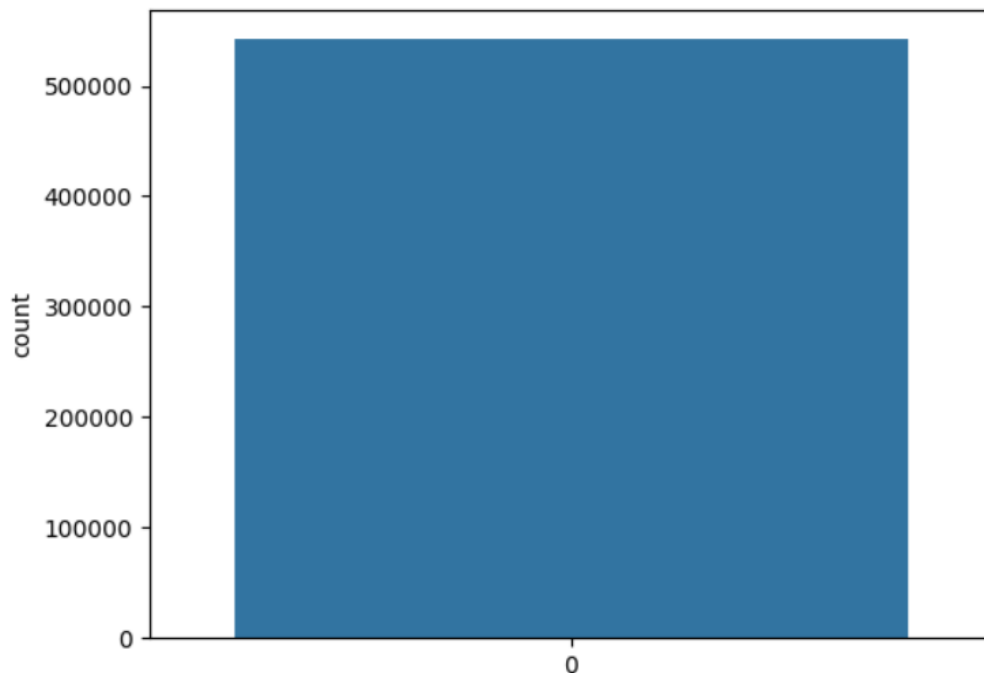


Figure 6.

Simpulan

Penelitian ini berhasil melalui seluruh tahap penerapan data mining dengan menggunakan metode Support Vector Machine (SVM) dalam menganalisis data penjualan produk retail. Proses pengolahan data dimulai dengan penghapusan kolom yang tidak diperlukan, bertujuan untuk menyederhanakan dataset dan meningkatkan efisiensi analisis. Fokus utama penelitian ini adalah pada nilai quantity, dan hasil analisis menunjukkan pencapaian nilai terbaik dalam konteks penjualan produk retail. Dalam pengoptimalan model, parameter khusus dengan nilai encode 6 mengalami transformasi menjadi 336. Hal ini bertujuan untuk memperkuat kemampuan SVM dalam mengklasifikasikan data dengan presisi yang lebih tinggi. Transformasi ini dapat dianggap sebagai strategi untuk meningkatkan daya beda dalam mengidentifikasi pola dan karakteristik yang lebih halus dalam dataset penjualan.

Sebagai hasilnya, metode SVM berhasil mengidentifikasi nilai quantity sebesar 541909.0 dari dataset, memberikan kontribusi signifikan pada pemahaman tentang perilaku penjualan produk retail. Hasil temuan penelitian ini mengungkapkan keunggulan metode SVM dalam konteks data mining untuk analisis penjualan produk retail. Dengan akurasi yang ditingkatkan, SVM mampu secara efektif mengklasifikasikan dan mengidentifikasi nilai quantity dengan tingkat kepercayaan yang tinggi. Implikasinya adalah bahwa penggunaan SVM dapat memberikan wawasan yang lebih mendalam dan akurat terkait dengan faktor-faktor yang memengaruhi kuantitas penjualan, membantu pemangku kepentingan untuk membuat keputusan yang lebih baik dan strategi yang lebih terinformasi. Dengan demikian, penelitian ini tidak hanya memvalidasi keefektifan SVM dalam analisis penjualan produk retail tetapi juga memberikan kontribusi pada pemahaman praktis dalam menerapkan teknik data mining untuk mengoptimalkan hasil bisnis. Temuan ini memiliki implikasi penting dalam mendukung pengambilan keputusan yang lebih baik, meningkatkan efisiensi operasional, dan merangsang inovasi dalam strategi pemasaran dan manajemen persediaan dalam konteks industri ritel.

References

1. G. Darussalam and A. G. Arief, "Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian," *Resti*, vol. 1, no. 1, pp. 19-25, 2017.
2. R. Takdirillah, "Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan," *Edumatic Journal Pendidikan Informatika*, vol. 4, no. 1, pp. 37-46, 2020, doi: 10.29408/edumatic.v4i1.2081.
3. A. Handayanto, K. Latifa, N. D. Saputro, and R. R. Waliyansyah, "Analisis dan Penerapan Algoritma Support Vector Machine (SVM) dalam Data Mining untuk Menunjang Strategi Promosi (Analysis and Application of Algorithm Support Vector Machine (SVM) in Data Mining to Support Promotional Strategies)," *JUITA Journal of Informatics*, vol. 7, no. 2, pp. 71-79, 2019.
4. R. R. Fiska, "Penerapan Teknik Data Mining dengan Metode Support Vector Machine (SVM) untuk Memprediksi Siswa yang Berpeluang Drop Out (Studi Kasus di SMKN 1 Sutera)," *SATIN - Sains dan Teknologi Informasi*, vol. 3, no. 1, pp. 15-23, 2017, doi: 10.33372/stn.v3i1.200.
5. I. A. Ashari, A. Wirasto, and D. N. Triwibowo, "Implementasi Market Basket Analysis dengan Algoritma Apriori untuk Analisis Pendapatan Usaha Retail Implementation of Market Basket Analysis with Apriori Algorithm for Retail Business Income Analysis," vol. 21, no. 3, 2022, doi: 10.30812/matrik.v21i3.1439.
6. N. Anwar, F. Adikara, R. Setiyati, R. Satria, and A. Satriawan, "Data Mining Menggunakan Metode Algoritma Apriori Pada Vending Machine Product Display," *JBASE - Journal of Business Audit and Information Systems*, vol. 4, no. 2, pp. 23-31, 2021, doi: 10.30813/jbase.v4i2.3004.
7. J. Jtik et al., "Analisis Sentimen pada Komen Twitter Pawang Hujan Mandalika dengan Support Vector Machine (SVM) dan Naïve Bayes," vol. 7, no. 2, pp. 0-6, 2023.
8. J. G. Susanto and S. Budi, "Penerapan Data Science Pada Dataset Pokemon," vol. 4, no. November, pp. 243-254, 2022.
9. M. Pratama Putra, M. Ariandi, M. Bina Darma, and D. Bina Darma, "Penerapan Data Mining Untuk Memprediksi Tingkat Ketepatan Jumlah Penjualan Produk Air Mineral Pada Pt. Mars Lestari," pp. 20-33, 2022, [Online]. Available: <http://eprints.binadarma.ac.id/16685/>.
10. A. Agung and A. Putri, "Penerapan Data Mining Untuk Mengestimasi Laju Data Mining Usage to Estimate Civil Growth in Denpasar," vol. 6, no. 1, pp. 37-44, 2023.