

Perbandingan Teknik Data Mining Untuk Prediksi Penjualan Produk UMKM Batik Karangdowo

COMPARISON OF DATA MINING TECHNIQUES FOR PREDICTING SALES OF PRODUCTS SMALL AND MEDIUM ENTERPRISES BATIK KARANGDOWO

Muhamad Rizaludin^{*1}, M.Rudi Fanani², Husni Hidayat³, Nur Hadian⁴

^{1,2,3,4} ITSNU Pekalongan: Jl. Karangdowo No. 9 Kedungwuni Kab. Pekalongan 51173, Indonesia

^{1,2,3,4} Program Studi Teknologi Informasi ITSNU Pekalongan

e-mail: ¹rizal.lonly@gmail.com, ²idurinanaf@gmail.com, ³husni.hidayat11@gmail.com,
⁴nurhadian97@gmail.com

Abstrak

Usaha Mikro, Kecil, dan Menengah (UMKM) memiliki peran penting dalam menyerap tenaga kerja untuk berkembang, termasuk dalam usaha bertahan krisis ekonomi. Selama ini UMKM Batik Karangdowo belum melakukan sebuah inovasi baru baik dalam hal produk dan penjualan yang luas, hal ini disebabkan karena pemilik usaha belum melihat secara detail tentang bagaimana penjualan dan apa yang sedang tren dipasaran menyangkut produk batik. Untuk itu peneliti menggunakan 2 teknik data mining seperti *k-Nearest Neighbor* (k-NN) dan *Naïve Bayes* yang bertujuan untuk mengetahui penjualan produk, perluasan pangsa pasar dan menambah inovasi dalam memperoleh keuntungan baru pada usaha UMKM Batik Karangdowo. Hasil pengujian data dengan menggunakan *software* Rapidminer pada atribut yang mempengaruhi adalah Nama Produk, Jenis Produk, dan Total Penjualan Produk dengan k-NN dan *Naïve Bayes* memiliki Tingkat akurasi 98,59%, presisi 100%, dan sensitivitas 95% dengan jenis produk yang laris adalah “Cargo Pendek Mocca 29-30”. Dimana penggunaan algoritma k-NN dan *Naïve Bayes* sama-sama menghasilkan performa yang sangat baik, sehingga dapat memprediksi penjualan produk UMKM Batik Karangdowo.

Kata kunci — Algoritma k-NN, Algoritma *Naïve Bayes*, Rapidminer, *Data Mining*

Abstract

Micro, Small, and Medium Enterprises play a crucial role in absorbing employment for growth, particularly in weathering economic crises. Batik Karangdowo have not yet ventured into new innovations in terms of both products and broadening sales. This is attributed to the business owner's lack of detailed insights into sales dynamics and market trends related to batik products. Therefore, researchers employed two data mining techniques, namely k-Nearest Neighbor (k-NN) and Naïve Bayes, with the aim of understanding product sales, expanding market share, and fostering innovation to gain new profits for Batik Karangdowo. The testing result using Rapidminer software on influencing attributes such as Product Name, Product Type, and Total Product Sales showed that both k-NN and Naïve Bayes achieved an accuracy rate of 98,59%, precision of 100%, and sensitivity of 95%. The identified best-selling product was “Cargo Shorts Mocca 29-30”. The use of both k-NN and Naïve Bayes algorithms demonstrated excellent performance, enabling the prediction of product sales for Batik Karangdowo.

Keyword — *k-NN Algorithm, Naïve Bayes Algorithm, Rapidminer, Data Mining*

1. PENDAHULUAN

Usaha Mikro, Kecil, dan Menengah (UMKM) memiliki peran penting yang menyerap banyak tenaga kerja untuk berkembang, termasuk dalam usaha bertahan krisis ekonomi. UMKM dianggap sebagai sumber penting dalam menciptakan lapangan kerja dan pembangunan ekonomi di pedesaan [1] [9]. Dalam upaya pembangunan ekonomi kerakyatan, UMKM untuk alat kebijakan pemangunan nasional. Karena UMKM Indonesia menyediakan lapangan kerja, pengentasan kemiskinan, pengangguran, distribusi pendapatan yang tidak merata dan urbanisasi yang berlebihan [2] [3] [8].

Salah satu usaha yang sekarang lagi marak adalah UMKM batik, dimana UMKM batik merupakan usaha yang sangat disukai pelanggan karena batik sebagai ikon budaya penting di Indonesia [4]. Batik adalah kerajinan yang memiliki nilai seni tinggi dan bagian dari budaya Indonesia sejak lama. Walaupun UMKM batik memiliki tantangan tidaklah mudah, namun

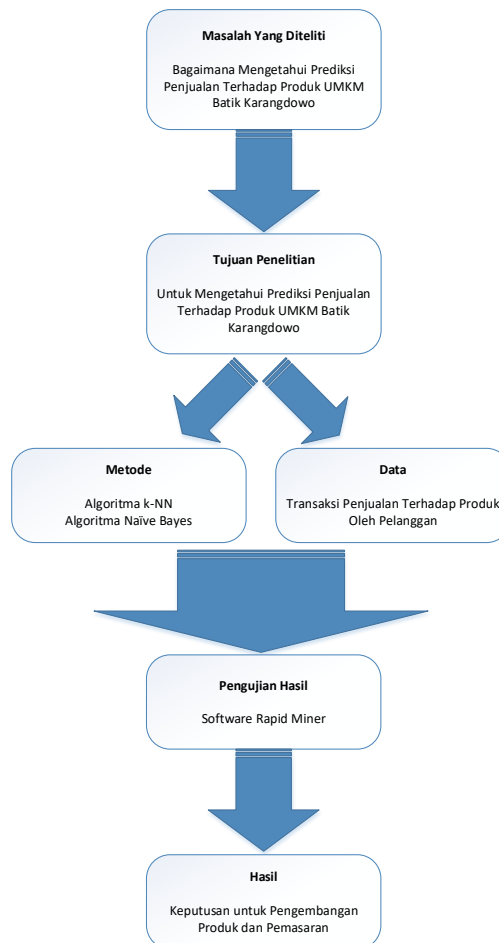
UMKM ini merupakan peluang usaha yang paling menjanjikan untuk ditekuni, dilihat dari segi modal UMKM batik ini tidak membutuhkan modal yang besar dalam memulainya [5] [6].

Demikian juga pada kota Pekalongan dikenal sebagai kota batik karena sejarah dan relasi kota dengan batik, kontribusi terhadap perkembangan batik, dan adanya sentra kerajinan batik sebagai mata pencaharian warga. Perkembangan UMKM batik Pekalongan terus ditingkatkan dalam sektor pariwisata. Sebagaimana yang ketahui setiap daerah memiliki ikon batik yang khas untuk menarik wisatawan. Oleh karena itu para pengusaha UMKM batik bersaing untuk membuat produk lebih baik dan laku dipasaran, dimana salah satunya adalah UMKM batik Karangdowo. UMKM batik Karangdowo terletak di daerah Pekalongan dengan produk usaha seperti batik. Selama ini UMKM batik Karangdowo belum melakukan inovasi baru dalam hal produk dan penjualan yang lebih luas, disebabkan karena pemilik usaha belum mengetahui penjualan terhadap produk yang laris dan tren produk batik dipasaran. Saat ini UMKM batik Karangdowo sudah memiliki pelanggan tetap, tetapi perluasan penjualan produk belum dilakukan secara menyeluruh misal se kota Pekalongan atau bisa dijadikan sebagai salah satu ciri khas produk batik di kota Pekalongan.

Untuk itu melalui penelitian ini, peneliti akan menggunakan 2 teknik *data mining* seperti *k-Nearest Neighbor* (k-NN) dan *Naïve Bayes* yang bertujuan untuk mengetahui penjualan produk, perluasan pangsa pasar, dan menambah inovasi dalam memperoleh keuntungan baru pada usaha UMKM batik Karangdowo.

2. METODE PENELITIAN

Tahapan penelitian ini dapat dilihat pada gambar sebagai berikut ini :



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Penjualan terhadap produk yang dihasilkan oleh UMKM batik Karangdowo dengan mencari informasi guna memperoleh data dengan cara berikut :

1. Observasi

Melakukan tinjauan langsung pada objek yang akan di teliti untuk mengetahui tentang kendala yang dihadapi oleh UMKM batik Karangdowo terutama untuk peningkatan penjualan dan perluasan pemasaran produk.

2. Wawancara

Peneliti melakukan sesi tanya jawab dengan pemilik usaha UMKM batik Karangdowo bertujuan untuk mengetahui kendala secara detail dihadapi oleh mereka selama ini.

3. Studi literatur

Mempelajari konsep *data mining* dengan mengumpulkan bahan pustaka, *literature*, dan karya ilmiah yang berkaitan dengan penelitian ini.

2.2. Desain Penelitian

Penelitian menggunakan model CRISP-DM (*Cross Industry Standart Process for Data Mining*) dengan langkah-langkah adalah [7] [10] :

1. Pemahaman Bisnis (*Bussiness Understanding*)

Fase ini memiliki tujuan untuk mengetahui penjualan terhadap produk UMKM batik Karangdowo agar memberikan manfaat bagi pelaku usaha khususnya pada bidang teknologi dan bisnis.

2. Pemahaman Data (*Data Understanding*)

Fase ini mengambil data permintaan terhadap produk oleh pelanggan pada UMKM batik Pieter Jackson. Data tersebut akan memulai proses latih untuk pembentukan k-NN dan *Naive Bayes*, dan proses uji hasil dengan k-NN dan *Naive Bayes*. Data berhubungan dengan transaksi penjualan terhadap produk selama 1 bulan mulai dari Nama Produk, Jenis Produk, dan Total Penjualan Produk. Adapun penjelasan setiap atribut diuraikan pada tabel di bawah ini :

Tabel 1. Atribut Data Transaksi Penjualan Produk

No	Nama Atribut	Keterangan
1	Nama Produk	Nama dari berbagai jenis produk
2	Jenis Produk	Jenis berdasarkan ukuran produk
3	Total Penjualan Produk	Total penjualan produk dalam 1 bulan

Seluruh data berjumlah 600 *record*. Dari sumber data tersebut maka dapat menganalisis dan memprediksi data permintaan produk dengan algoritma k-NN dan *Naive Bayes*.

3. Pengolahan Data (*Data Preparation*)

Fase ini *dataset* akan dipilih dengan kondisi atribut paling dekat dengan keputusan atau label prediksi dan memastikan kondisi konsistensi nilai atribut.

4. Pemodelan (*Modelling*)

Fase ini penerapan *data mining* untuk prediksi penjualan terhadap produk menggunakan algoritma k-NN dan *Naive Bayes* dengan perhitungan manual dan di uji dengan Rapidminer menggunakan operator k-nn dan *naive bayes* kemudian evaluasi nilai akurasi data. *Dataset* berjumlah 80 *record* dengan atribut adalah Nama Produk, Jenis Produk, Total Penjualan Produk, dan Laris.

5. Evaluasi (*Evaluation*)

Dataset di evaluasi hingga memperoleh hasil keputusan, berfungsi untuk menentukan hasil nilai untuk pemodelan dan membentuk hasil keputusan.

6. Penyebaran (*Deployment*)

Setelah evaluasi akan terjadi penyebaran yang membentuk keputusan. Hasil akurasi memiliki fungsi sebagai materi tes dan pembuktian sebuah kasus.

3. HASIL DAN PEMBAHASAN

Penelitian menggunakan algoritma k-NN dan *Naïve Bayes* dalam menentukan hasil akurasi. Langkah-langkah dengan menguji dataset penelitian dengan *software* Rapidminer sebagai berikut :

3.1. Import Data

Langkah pertama adalah *import* data latih yang akan digunakan ke dalam *software* Rapidminer.

1	Nama Produk	Jenis Produk	Total Penjualan Produk	Laris
2	Terbaru Celana Pendek Pria...	.Cargo Panjang Abu 33-34	75.000	ya
3	Terbaru Celana Pendek Pria...	.Cargo Panjang Abu 35-36	87.000	ya
4	Terbaru Celana Pendek Pria...	.Cargo Panjang Cream 29-30	82.000	ya
5	Terbaru Celana Pendek Pria...	.Cargo Panjang Cream 31-32	80.000	ya
6	Terbaru Celana Pendek Pria...	.Cargo Panjang Cream 33-34	45.000	tidak
7	Terbaru Celana Pendek Pria...	.Cargo Panjang Hitam 29-30	82.000	ya
8	Terbaru Celana Pendek Pria...	.Cargo Panjang Hitam 31-32	132.000	ya
9	Terbaru Celana Pendek Pria...	.Cargo Panjang Hitam 33-34	92.000	ya
10	Terbaru Celana Pendek Pria...	.Cargo Panjang Mocca 27-28	96.000	ya
11	Terbaru Celana Pendek Pria...	.Cargo Panjang Mocca 33-34	88.000	ya
12	Terbaru Celana Pendek Pria...	.Cargo Pendek Abu 27-28	63.000	ya
13	Terbaru Celana Pendek Pria...	.Cargo Pendek Abu 29-30	66.000	ya
14	Terbaru Celana Pendek Pria...	.Cargo Pendek Abu 31-32	141.000	ya

Gambar 1. Import Data

Dataset terdiri dari 80 data dengan tiga atribut dan satu label. Dilanjutkan menentukan tipe data pada tiap-tiap atribut. Berikut tampilan tipe data tersebut :

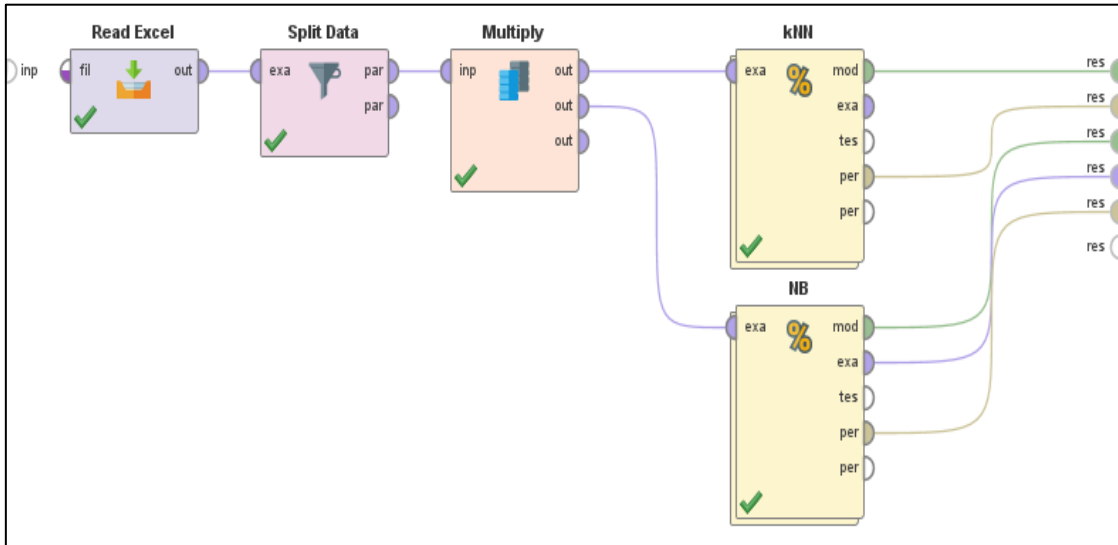
column index	attribute meta data information			
0	Nama Produk	<input checked="" type="checkbox"/> column ...	polynomi...	attribute ▼
1	Jenis Produk	<input checked="" type="checkbox"/> column ...	polynomi...	attribute ▼
2	Total Penjuala	<input checked="" type="checkbox"/> column ...	integer ▼	attribute ▼
3	Laris	<input checked="" type="checkbox"/> column ...	polynomi...	label ▼

Gambar 2. Pemilihan Tipe Data

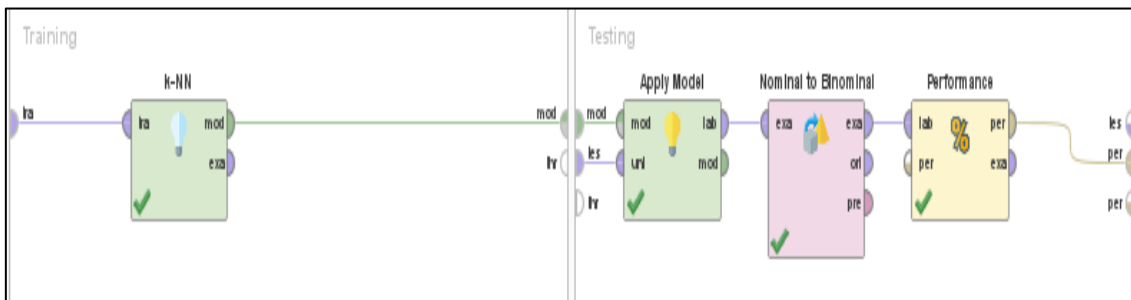
Polynomial berfungsi jumlah *value* terdiri dari lebih dua *value* pada tiap atribut. Atribut “Laris” sebagai label yaitu parameter keputusan, sedangkan tiga atribut yang lain adalah sebagai *attribute*.

Operator yang digunakan adalah *split* data dan *cross validation*. Operator *split* data berguna untuk memisah atau membagi menjadi data latih dan data uji guna untuk mengetahui performa dari model yang dibentuk oleh k-NN dan *Naïve Bayes*. Proporsi data latih di ambil 90% dan data uji di ambil 10% dari perbandingan seluruh *dataset*. Sedangkan operator *cross validation* berguna sebagai pengukur keakuratan sebuah model yang sedang dijalankan. Kemudian, di dalam operator *cross validation* terdapat *process*, pada *process* tersebut membutuhkan operator yang dapat menghasilkan tujuan pengujian. Karena pada operator *process* yang digunakan adalah k-nn dan *naïve bayes*. Kriteria yang digunakan k-NN dengan parameter k adalah 5 (lima) dan perhitungan jarak terdekat menggunakan *Mixed Euclidean*

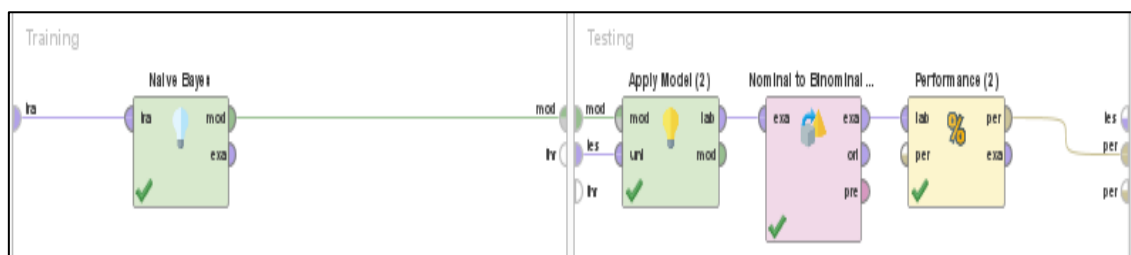
Distance, sedangkan kriteria *Naïve Bayes* berdasarkan performa *dataset*. Dalam perhitungan akurasi menggunakan operator *performance*.



Gambar 3. Tampilan Proses k-NN dan *Naïve Bayes*



Gambar 4. Tampilan Proses *Validation* k-NN



Gambar 5. Tampilan Proses *Validation* *Naïve Bayes*

3.2. Hasil Akurasi Data dengan k-NN dan *Naïve Bayes*

1. Akurasi, untuk mengukur seberapa baik model membuat prediksi yang benar dari total prediksi yang dilakukan. Akurasi yang didapatkan adalah 98,59%. Berikut perhitungan akurasi yaitu :

$$Akurasi = \frac{51 + 19}{51 + 1 + 19} \times 100\% = \frac{70}{71} \times 100\% = 98,59\%$$

accuracy: 98.57% +/- 4.52% (micro average: 98.59%)			
	true ya	true tidak	class precision
pred. ya	51	1	98.08%
pred. tidak	0	19	100.00%
class recall	100.00%	95.00%	

Gambar 6. Hasil Akurasi k-NN dan *Naïve Bayes*

2. Presisi, untuk mengukur seberapa baik model membuat prediksi yang benar untuk kelas positif dari total prediksi positif yang dilakukan. Presisi yang didapatkan adalah 100%. Berikut perhitungan presisi yaitu :

$$Presisi = \frac{19}{19} \times 100\% = 100\%$$

precision: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: tidak)			
	true ya	true tidak	class precision
pred. ya	51	1	98.08%
pred. tidak	0	19	100.00%
class recall	100.00%	95.00%	

Gambar 8. Hasil Presisi k-NN dan *Naive Bayes*

3. Sensitivitas (*recall*), untuk menggambarkan seberapa baik suatu model dalam mengidentifikasi kelas positif dengan benar. Sensitivitas yang didapatkan adalah 95%. Berikut perhitungan sensitivitas yaitu :

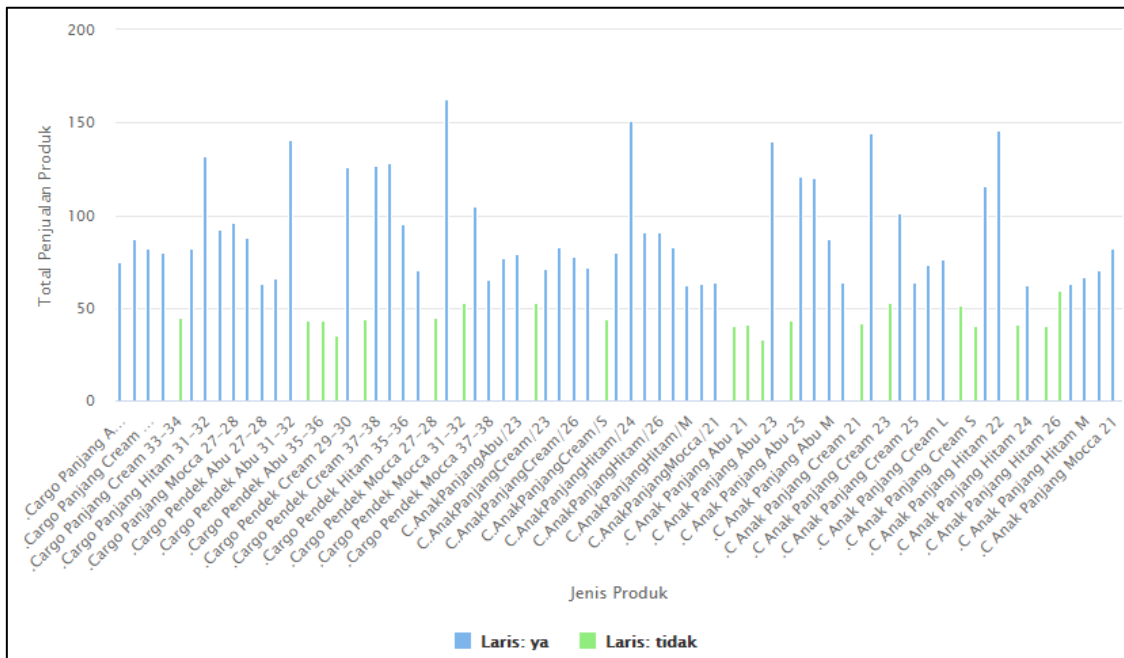
$$Sensitivitas = \frac{19}{19 + 1} \times 100\% = \frac{19}{20} \times 100\% = 95\%$$

recall: 95.00% +/- 15.81% (micro average: 95.00%) (positive class: tidak)			
	true ya	true tidak	class precision
pred. ya	51	1	98.08%
pred. tidak	0	19	100.00%
class recall	100.00%	95.00%	

Gambar 9. Hasil Sensitivitas k-NN dan *Naive Bayes*

3.3. Hasil Pengujian Data

Dari pengujian data dengan *software* Rapidminer pada UMKM batik Karangdowo didapatkan bahwa jenis produk yang laris adalah “Cargo Pendek Mocca 29-30”. Kesimpulan dari hasil pengujian data UMKM batik Karangdowo dengan algoritma k-NN dan *Naive Bayes* sama-sama menghasilkan performa yang sangat baik, sehingga dapat memprediksi data penjualan terhadap produk UMKM batik Karangdowo.



Gambar 10. Grafik Hasil Pengujian Data Pada Jenis Produk

4. KESIMPULAN

Berdasarkan hasil analisa data menggunakan *data mining* yang telah dijelaskan sebelumnya, maka peneliti dapat menarik kesimpulan bahwa :

1. Dengan menggunakan algoritma k-NN dan *Naïve Bayes* dapat memprediksi untuk pengembangan produk dan pemasaran UMKM batik Karangdowo.
2. Tingkat akurasi dari olah data penjualan produk UMKM batik Karangdowo dengan menggunakan *software* Rapidminer pada atribut yang mempengaruhi adalah Nama Produk, Jenis Produk, dan Total Penjualan Produk dengan k-NN dan *Naïve Bayes* memiliki tingkat akurasi 98,59%, presisi 100%, dan sensitivitas 95% dengan jenis produk yang laris adalah "Cargo Pendek Mocca 29-30".

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Institut Teknologi dan Sains Nahdlatul Ulama (ITSNU) Pekalongan yang telah memberi dukungan finansial terhadap penelitian ini. Terima kasih juga disampaikan kepada pihak UMKM batik Karangdowo yang secara aktif berpartisipasi pada penelitian ini.

DAFTAR PUSTAKA

- [1] Pramana, I.M.A.A, Sudiarsa, I.W, dan Nugraha, P.G.S.C. 2023. Penerapan Algoritma Naïve Bayes Untuk Prediksi Penjualan Produk Terlaris Pada CV Akusara Jaya Abadi. *Jurnal Teknik Informatika dan Sistem Informasi*. Vol.10 No.4. Hal 518-534.
- [2] Ronaldi, A.A., dan Hunafi, N. 2020. Implementasi Data Mining Untuk Prediksi Penjualan Pestisida Pada CV Mitra Artha Sejati Menggunakan Algoritma Naïve Bayes. *eProsiding Teknik Informatika (Protektif)*. Vol.1 No.1. Hal 250-257.
- [3] Nosiell, Sriyanto, dan Maylani, F. 2021. Perbandingan Teknik Data Mining Untuk Prediksi Penjualan Pada UMKM Gerabah. *Seminar Nasional Hasil Penelitian dan Pengabdian Masyarakat*. Hal 72-86.

- [4] Alfani, A., Rozi, F., dan Sukmana, F. 2021. Prediksi Penjualan Produk Unilever Menggunakan Metode k-Nearest Neighbor. *Jurnal Ilmiah Penelitian dan Pembelajaran Informatika (JIPI)*. Vol.6 No.1.
 - [5] Elisa, E., Mardiansyah, Y., dan Fauzi, R. 2022. Pola Pembelian Konsumen Terhadap Produk UMKM Martista Ikhsan Dengan Algoritma Naïve Bayes. *Jurnal Manajemen Informatika & Sistem Informasi (MISI)*. Vol.5 No.2.
 - [6] Yusuf, W., Witri, R., dan Juliane, C. 2022. Model Prediksi Penjualan Jenis Produk Tekstil Menggunakan Algoritma k-Nearest Neighbor (k-NN). *Indonesian Journal on Computer and Information Technology (IJCIT)*. Vol.7 No.1. Hal 1-6.
 - [7] Aprilla, C.C.E., dan Christanto, F.W. 2023. Penerapan Algoritma Naïve Bayes Untuk Prediksi Penjualan Produk Iconnet. *Jurnal Ilmiah Teknologi Informasi Dan Komunikasi (JTIK)*. Vol.14 No.2. Hal 255-265.
 - [8] Raharjo, M.R. 2017. Analisis Algoritma Klasifikasi Dan Asosiasi Terhadap Atribut Data Pelaku Usaha Mikro Kecil Dan Menengah (UMKM). *Technologia*. Vol.8 No.3.
 - [9] Azizah, H.W., Nurdiawan, O., Dwilestari, G., Kaslani, dan Tohidi, E. 2022. Klasifikasi Pemberian Bantuan UMKM Cirebon Dengan Menggunakan Algoritma k-Nearest Neighbor. *Journal of Computer System and Informatics (JoSYC)*. Vol.3 No.3.
 - [10] Rizaludin, M., dan Fikriah, F. 2023. Prediksi Perilaku Pelanggan Pada Produk UMKM Batik Dengan Menggunakan Algoritma Decision Tree. *Teknomatika*. Vol.13 No.2. Hal 8-16.
-