
Klasifikasi *Level Customer Priority* PT. Internet Madju Abad Milenindo

Kadek Sri Nopiani¹, I Nyoman Rudy Hendrawan², Erma Sulisty Rini³

Sistem Informasi¹⁾², Bisnis Digital³

Institut Teknologi dan Bisnis STIKOM Bali

Denpasar, Indonesia

e-mail: 210030425@stikom-bali.ac.id¹, rudyhendrawan@stikom-bali.ac.id², erma@stikom-bali.ac.id³

Abstrak

Peningkatan penggunaan teknologi informasi dan komunikasi di Indonesia mendorong persaingan ketat di industri penyedia layanan internet. Penelitian ini bertujuan untuk mengklasifikasikan prioritas pelanggan pada PT. Internet Madju Abad Milenindo (GlobalXtreme) menggunakan metode *K-Nearest Neighbor (KNN)* dan *Support Vector Machine (SVM)*. Dataset yang digunakan merupakan data buatan berdasarkan atribut asli perusahaan, meliputi total pembelian, frekuensi pembelian, dan lama berlangganan, yang mencerminkan nilai, aktivitas, dan loyalitas pelanggan. Metode penelitian meliputi *preprocessing data*, *Feature encoding*, dan *splitting dataset* menjadi *training*, *validation*, dan *testing set*. Model dioptimalkan menggunakan *GridSearchCV*, dengan evaluasi menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil penelitian kedua metode mencapai akurasi keseluruhan 96%. Model *SVM* menunjukkan performa terbaik pada beberapa kelas, dengan nilai *F1-score* tertinggi mencapai 0,99 untuk kelas prioritas tertinggi. Model *KNN* menunjukkan kinerja yang konsisten dengan nilai rata-rata tertimbang *F1-score* sebesar 0,95. Namun, terdapat kelemahan dalam mengklasifikasikan kelas dengan jumlah sampel kecil. Penelitian ini memberikan panduan bagi perusahaan untuk mengidentifikasi pola dalam data pelanggan, sehingga memungkinkan pengambilan keputusan berbasis data dalam menentukan prioritas pelanggan. Penelitian selanjutnya, disarankan untuk menambah jumlah data, terutama pada kelas yang kurang terwakili, serta mempertimbangkan penggunaan teknik *ensemble* guna meningkatkan akurasi dan konsistensi hasil klasifikasi.

Kata kunci: Klasifikasi Prioritas Pelanggan, *K-Nearest Neighbor (KNN)*, *Support Vector Machine (SVM)*

1. Pendahuluan

Penggunaan teknologi informasi di Indonesia, termasuk internet, mengalami pertumbuhan signifikan dalam beberapa tahun terakhir [1]. Berdasarkan data tahun 2022, sekitar 66,48% penduduk Indonesia telah mengakses internet, meningkat dari 62,10% pada tahun sebelumnya. Hal ini mencerminkan adopsi teknologi digital yang semakin pesat [2]. Di Bali, GlobalXtreme sebagai salah satu penyedia layanan internet terkemuka, menawarkan layanan *broadband* yang dirancang untuk memenuhi kebutuhan pelanggan, baik individu maupun bisnis. Perusahaan harus mengutamakan strategi berbasis pelanggan untuk meningkatkan kebahagiaan pelanggan, loyalitas, dan daya saing pasar dalam menghadapi persaingan industri yang semakin ketat.

Permasalahan terbesar yang dihadapi oleh perusahaan seperti GlobalXtreme adalah bagaimana mengklasifikasikan prioritas pelanggan secara efektif untuk meningkatkan retensi dan kepuasan. Dengan pertumbuhan penggunaan teknologi digital, terdapat kebutuhan mendesak untuk memanfaatkan data pelanggan dalam pengambilan keputusan strategis [3]. Data pelanggan yang meliputi total pembelian, frekuensi pembelian, dan lama pelanggan dapat digunakan untuk mencerminkan nilai, aktivitas, serta loyalitas pelanggan, sehingga perusahaan mampu menentukan prioritas mereka secara lebih efisien.

Berbagai studi telah menunjukkan pentingnya penerapan metode *machine learning* dalam klasifikasi dan pengambilan keputusan berbasis data. Hendrawan et al. (2021) menggunakan metode *Naïve Bayes* agar mengelompokkan lama studi dan predikat kelulusan mahasiswa berdasarkan data atribut seperti jenis kelamin, program studi, dan tahun masuk. Hasil menunjukkan akurasi 0,74 untuk klasifikasi lama studi dan 0,61 untuk predikat kelulusan pada Program Studi Sistem Komputer, serta akurasi 0,73 dan 0,67 masing-masing untuk kelompok data Sistem Informasi, membuktikan relevansi metode ini dalam analisis data pendidikan [4]. Triandini et al. (2022) menerapkan model *machine learning* berbasis regresi, termasuk *TensorFlow*, *XGBoost*, dan *LightGBM*, untuk memprediksi curah hujan harian di Stasiun Meteorologi Ngurah Rai. Penelitian ini menunjukkan bahwa meskipun *neural network* memberikan prediksi yang kuat, model *gradient boosting* memberikan peningkatan akurasi yang signifikan setelah dilakukan optimasi

preprocessing data, memberikan peluang besar untuk aplikasi prakiraan cuaca berbasis data [5]. Lantara et al. (2022) mengembangkan dataset gerakan dasar Tari Bali (GETARI) yang mencakup enam gerakan dasar. Dataset ini diklasifikasikan menggunakan model *pre-trained* VGG-LSTM dan I3D, di mana model I3D mencapai akurasi hingga 0,97, *precision* 0,98, *recall* 0,98, dan *F1* 0,98. Penelitian ini menunjukkan potensi besar *machine learning* dalam pelestarian budaya melalui klasifikasi gerakan tari tradisional [6].

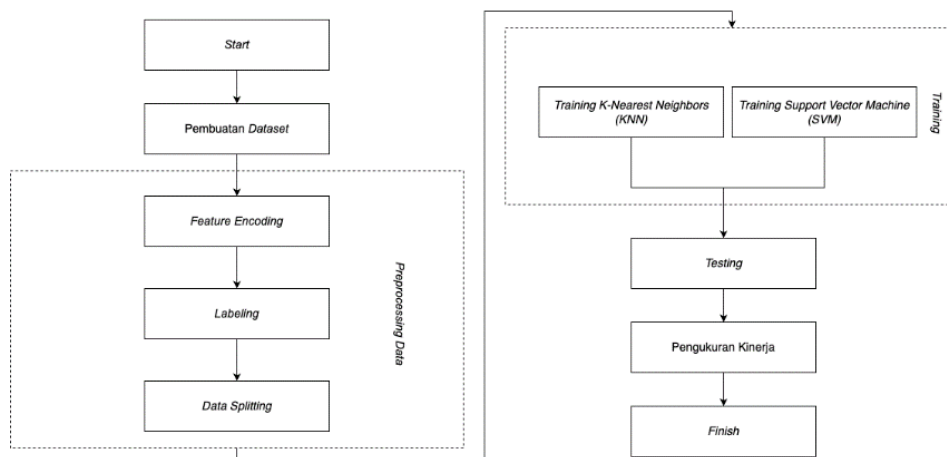
Penelitian-penelitian ini menunjukkan potensi besar *machine learning* dalam mendukung klasifikasi dan prediksi di berbagai domain, termasuk analisis data pendidikan, prediksi cuaca, dan pelestarian budaya.

Penelitian ini berfokus pada klasifikasi prioritas pelanggan di GlobalXtreme menggunakan KNN dan *Support Vector Machine* (SVM). Kedua metode terbukti efektif dalam studi sebelumnya, dengan tingkat akurasi masing-masing 81,32% untuk KNN dan 81,58% untuk SVM [7]. Data yang digunakan adalah data sintesis berbasis atribut asli pelanggan GlobalXtreme, yang mencakup informasi penting seperti total pembelian, frekuensi pembelian, dan lama pelanggan. Dengan pendekatan ini, penelitian bertujuan menghasilkan klasifikasi yang akurat untuk membantu perusahaan menetapkan strategi prioritas pelanggan secara lebih efektif.

Hasil temuan diharapkan memberikan panduan praktis bagi GlobalXtreme dalam memanfaatkan data pelanggan untuk meningkatkan kepuasan dan loyalitas. Selain itu, diharapkan menjadi referensi bagi penelitian selanjutnya dalam mengembangkan pendekatan *machine learning* untuk klasifikasi pelanggan, serta mendorong pengembangan model yang lebih canggih melalui optimasi parameter dan penggunaan dataset yang lebih representatif.

2. Metode Penelitian

Penelitian dilakukan pada PT. Internet Madju Abad Milenindo selama tiga bulan, dari Desember 2024 hingga Februari 2025, dengan tujuan mengklasifikasikan level prioritas pelanggan. Data yang digunakan merupakan data buatan yang mencakup atribut seperti nama, tanggal lahir, alamat, user ID, paket, harga, tanggal pemasangan, total pembelian, lama berlangganan, frekuensi, dan prioritas. Proses penelitian meliputi *preprocessing* data dengan *encoding*, pemisahan fitur dan label, serta pembagian dataset menjadi data latih, validasi, dan uji. Dua algoritma klasifikasi, yaitu *K-Nearest Neighbors* (KNN) dan *Support Vector Machine* (SVM), digunakan untuk melatih model. Proses pelatihan dilakukan dengan menentukan parameter yang optimal melalui eksperimen, kemudian model diuji dengan data uji untuk mengukur kinerjanya menggunakan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*. Adapun metode penelitian ini melibatkan langkah-langkah pada Gambar 1.



Gambar 1. Metode Penelitian

2.1 Pembuatan Dataset

Dikarenakan adanya kebijakan privasi pelanggan pada perusahaan tersebut, penelitian ini akan menggunakan data buatan yang dibuat berdasarkan atribut asli yang telah ditentukan. Atribut yang dimasukkan meliputi Nama, Tanggal Lahir, Alamat, User ID, Paket, Harga Paket, Alat, Harga Alat, Tanggal Instalasi, Lama Berlangganan, Frekuensi Pembelian, dan Prioritas. Dataset dibuat menggunakan bahasa pemrograman *Python*, dimulai dari *list dictionary* yang diubah menjadi *DataFrame* dengan *Pandas*, lalu disimpan dalam format *CSV*. Dataset ini digunakan untuk klasifikasi level prioritas pelanggan

berdasarkan atribut yang ada, dengan prioritas ditentukan oleh rentang nominal pembelian [8]. Hasil pembuatan dataset terdiri dari 3733 baris dan 13 kolom pada Tabel 1.

Tabel 1. Sampel Dataset

No	Nama	Umur (TTL)	Alamat	Email	User ID	Paket	Harga Paket	Harga Alat	Total Pembelian	Tanggal Pemasangan	Lama Berlangganan (bulan)	Frekuensi Pembelian	Prioritas
0	Linda Martinez	29/01/1973	Kerobokan	linda.martinez@hotmail.com	lmartinez49UI	SIGNATURE 300	3330000	888000	4218000	12/09/2020	52	6	4
1	Laura Williams	21/02/1971	Denpasar Timur	laura.williams@hotmail.com	lwilliams1WH	SIGNATURE 500	5550000	888000	6438000	26/07/2022	30	10	3
2	Michael Brown	23/10/1996	Gianyar	michael.brown@yahoo.com	mbrown84EF	SIGNATURE 300	3330000	1320900	4650900	05/05/2022	32	4	1
3	Dewa Sari	14/03/1999	Tabanan	dewa.sari@gmail.com	dsari16ZA	LITE 100	333000	888000	1221000	12/11/2021	38	9	5
4	Emily Brown	17/12/2000	Gianyar	emily.brown@outlook.com	ebrown67FJ	SIGNATURE 300	3330000	1320900	4650900	18/12/2022	25	1	2
...
3732	Laura Smith	26/03/1957	Tabanan	laura.smith@icloud.com	lsmith55GE	DEDICATED INTERNET ACCESS	11100000	1320900	12420900	28/02/2020	59	1	3

2.2 Preprocessing Data

Proses *preprocessing* data dimulai dengan menghapus fitur yang tidak relevan, seperti data pribadi (Nama, *Email*, *User ID*) dan informasi temporal (Tanggal Pemasangan) menggunakan fungsi *drop()*. Kemudian, dilakukan pemeriksaan kualitas data dengan mengecek nilai null dan data duplikat. Selanjutnya, outlier ditangani dengan metode *Interquartile Range (IQR)*, mengubah nilai ekstrem menjadi batas terdekat untuk mengurangi dampaknya tanpa kehilangan informasi penting. Hasil *preprocessing* menunjukkan dataset dengan 3733 baris dan 9 kolom, siap untuk analisis lanjut, pada Tabel 2.

Tabel 2. Dataset Setelah *Preprocessing*

No	Alamat	Paket	Harga Paket	Harga Alat	Total Pembelian	Lama Berlangganan (bulan)	Frekuensi Pembelian	Prioritas	Umur
0	Kerobokan	SIGNATURE 300	3330000	888000	4218000	52	6	4	51
1	Denpasar Timur	SIGNATURE 500	5550000	888000	6438000	30	10	3	53
2	Gianyar	SIGNATURE 300	3330000	1320900	4650900	32	4	1	28
3	Tabanan	LITE 100	333000	888000	1221000	38	9	5	25
4	Gianyar	SIGNATURE 300	3330000	1320900	4650900	25	1	2	24
...
3731	Denpasar Selatan	LITE 100	333000	888000	1221000	38	8	3	29
3732	Tabanan	DEDICATED INTERNET ACCESS	11100000	1320900	12420900	59	1	3	67

a. Feature Encoding

Feature encoding adalah proses mengubah data kategoris menjadi bentuk numerik yang dapat diproses oleh algoritma pembelajaran mesin [9]. Pada umumnya, algoritma seperti *Support Vector Machine (SVM)* dan *K-Nearest Neighbors (KNN)* tidak dapat bekerja langsung dengan data kategori, sehingga fitur-fitur tersebut perlu diubah menjadi representasi numerik. Hasil dari proses *feature encoding* dapat dilihat pada Tabel 3.

Tabel 3. *Feature Encoding*

Nama	Umur (TTL)	Alamat	Email	User ID	Paket	Harga Paket	Harga Alat	Total Pembelian	Tanggal Pemasangan	Lama Berlangganan (bulan)	Frekuensi Pembelian	Prioritas
135	29/01/1973	7	674	2453	6	3330000	888000	4218000	-9.22E+09	52	6	4
129	21/02/1971	3	644	2634	7	5550000	888000	6438000	-9.22E+09	30	10	3
150	23/10/1996	5	752	2693	6	3330000	1320900	4650900	-9.22E+09	32	4	1
36	14/03/1999	9	180	603	1	333000	888000	1221000	-9.22E+09	38	9	5

40	17/12/2000	5	203	760	6	3330000	1320900	4650900	-9.22E+09	25	1	2
----	------------	---	-----	-----	---	---------	---------	---------	-----------	----	---	---

b. Labeling

Pada proses *labeling*, data dipisahkan menjadi fitur (X) yang mencakup semua kolom kecuali 'Prioritas', yang berfungsi sebagai label (Y). Fitur (X) meliputi informasi seperti nama, umur, alamat, ID pengguna, paket, harga, dan lainnya, sementara label (Y) adalah kolom 'Prioritas' yang menunjukkan tingkat prioritas. Pemisahan ini penting untuk membangun model prediktif. Hasil *labeling* dapat dilihat pada Tabel 4.

Tabel 4. *Labeling*

	Fitur (X)											Fitur (Y)	
	Nama	Umur (TTL)	Alamat	Email	User ID	Paket	Harga Paket	Harga Alat	Total Pembelian	Tanggal Pemasangan	Lama Berlangganan (bulan)	Frekuensi Pembelian	Prioritas
0	135	29-01-1973	7	674	2453	6	3330000	888000	4218000	1.60E+15	52	6	4
1	129	21-02-1971	3	644	2634	7	5550000	888000	6438000	1.66E+15	30	10	3
2	76	26-01-1973	4	381	1505	1	3330000	888000	1221000	1.61E+15	48	8	4
3	40	17-12-2000	5	203	760	6	3330000	1320900	4650900	1.67E+15	25	1	2
4	12	14-11-1959	10	64	304	3	1110000	888000	1998000	1.66E+15	28	5	4

c. Data Splitting

Dalam penelitian ini menggunakan 3733 data, data dibagi menjadi tiga subset yaitu *train set* (60%), *validation set* (20%), dan *test set* (20%) menggunakan fungsi *train_test_split* dari *Scikit-Learn*. *Train set* digunakan untuk melatih model, *validation set* untuk mengevaluasi model selama pelatihan, dan *test set* untuk menguji kinerja akhir model. Pembagian dilakukan secara bertahap, yaitu pertama membagi data menjadi *train set* dan *temporary data* (60:40), kemudian *temporary data* dibagi lagi menjadi *validation set* dan *test set* (50:50). Hasil pembagian menghasilkan *train set* sebanyak 2.239 data, serta *validation set* dan *test set* masing-masing sebanyak 747 data. Hasil akhir pembagian menunjukkan komposisi 60:20:20, pada Tabel 5.

Tabel 5. *Data Splitting*

Data Split	Presentase	Jumlah data
<i>Train</i>	60%	2239
<i>Validation</i>	20%	747
<i>Test</i>	20%	747
Total	100%	3733

3. Hasil dan Pembahasan

3.1 Training

Model *K-Nearest Neighbors* (KNN) dioptimalkan menggunakan *GridSearchCV* dengan parameter *n_neighbors*, *weights*, dan *metric* [10]. Hasil terbaik diperoleh dengan *n_neighbors*=2, *weights*='distance', dan *metric*='manhattan'. Konfigurasi ini menunjukkan model bekerja optimal dengan dua tetangga terdekat, pembobotan berdasarkan jarak, dan pengukuran jarak Manhattan. Model final disimpan dalam variabel *knn_model* untuk prediksi selanjutnya. Proses pencarian parameter terbaik untuk model *Support Vector Machine* (SVM) menggunakan *GridSearchCV* mengoptimalkan parameter C, gamma, dan kernel. Hasil terbaik diperoleh dengan C=1, gamma=1, dan kernel RBF. Konfigurasi ini menunjukkan model dengan margin kecil, toleransi kesalahan rendah, dan pengaruh sampel pelatihan lokal. Adapun konfigurasi parameter pada Tabel 6.

Tabel 6. Parameter

Model	Parameter	Nilai
SVM	C	0.1, 1, 10
	gamma	1, 0.1, 0.01
	kernel	linear, rbf
KNN	<i>n_neighbors</i>	2, 4, 6, 8

weights uniform,
 distance
 metric euclidean,
 manhattan

Jumlah repetisi *Grid Search* dihitung dari kombinasi parameter dikalikan jumlah lipatan validasi silang (cv). Adapun hasil *GridSearchCV* tertuang dalam tabel 7 dan 8.

Tabel 7. Hasil *GridSearchCV* SVM

C	Gamma	Kernel	Mean Accuracy	Std Dev
1	0.01	linear	0.958	0.045644
1	1	linear	0.958	0.045644
1	1	rbf	0.958	0.06455
1	0.1	linear	0.958	0.045644
1	0.1	rbf	0.950	0.061237
10	0.01	linear	0.950	0.061237
10	0.1	rbf	0.950	0.061237
10	0.1	linear	0.950	0.061237
10	1	linear	0.950	0.061237
10	0.01	rbf	0.950	0.061237
0.1	1	rbf	0.941	0.062361
10	1	rbf	0.941	0.056519
0.1	0.01	linear	0.941	0.062361
0.1	0.1	linear	0.941	0.062361
0.1	1	linear	0.941	0.062361
1	0.01	rbf	0.908	0.048591
0.1	0.1	rbf	0.900	0.056519
0.1	0.01	rbf	0.466	0.266667

Tabel 8. Hasil *GridSearchCV* KNN

n_neighbors	weights	metric	Mean Accuracy	Std Dev
2	distance	manhattan	0.958	0.045644
2	uniform	euclidean	0.950	0.040825
2	distance	euclidean	0.950	0.040825
4	uniform	euclidean	0.950	0.040825
4	distance	euclidean	0.950	0.040825
6	distance	euclidean	0.950	0.040825
2	distance	manhattan	0.950	0.040825
4	uniform	manhattan	0.950	0.040825
6	uniform	euclidean	0.941	0.042492
8	distance	euclidean	0.941	0.056519
6	uniform	manhattan	0.941	0.056519
6	distance	manhattan	0.941	0.056519
8	uniform	manhattan	0.941	0.056519
8	distance	manhattan	0.941	0.056519
8	uniform	euclidean	0.933	0.056519
2	uniform	manhattan	0.925	0.040825

3.2 Pengukuran Kinerja

Hasil testing dan evaluasi model klasifikasi *Support Vector Machine* (SVM) menunjukkan akurasi keseluruhan 96% pada 755 data, dengan kelas 5 memiliki performa terbaik yang ditunjukkan oleh *F1-score* 0,99, *precision* 0,98, dan *recall* 1,00 pada 313 sampel. Kelas 3, meskipun memiliki *precision* 0,90, mengalami *recall* yang rendah (0,50), sehingga menghasilkan *F1-score* terendah 0,64, yang dimana hanya setengah data kelas 3 terklasifikasi dengan benar. Hal ini disebabkan oleh jumlah data yang sedikit (support = 18) dan ketidakseimbangan data antar kelas, sehingga model kurang fokus pada kelas ini. Model ini juga menghasilkan rerata tertimbang 0,96, menandakan konsistensi dan kehandalan yang tinggi meskipun terdapat ketidakseimbangan jumlah sampel antar kelas. Visualisasi *confusion matrix* menunjukkan prediksi yang tepat di diagonal utama. Kelas 4 dan 0 memiliki akurasi tinggi, sedangkan kelas 2 menunjukkan prediksi rendah. Hasil testing dan pengukuran kinerja pada Tabel 9 dan Tabel 10.

Tabel 9. Hasil Model SVM

	Precision	Recall	F1-Score	Support
1	0.96	0.98	0.97	251
2	0.92	0.93	0.92	94
3	0.90	0.50	0.64	18
4	0.97	0.91	0.94	79
5	0.98	1.00	0.99	313
Accuracy			0.96	755
Macro Avg	0.95	0.86	0.89	755
Weighted Avg	0.96	0.96	0.96	755

Tabel 10. *Confusion Matrix* SVM

Actual / Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Total
Class 0	247	1	0	0	3	251
Class 1	6	87	0	0	1	94
Class 2	4	9	2	1	1	17
Class 3	2	3	1	72	1	79
Class 4	0	0	0	0	313	313

Sedangkan Model *K-Nearest Neighbor* (KNN) menunjukkan akurasi tinggi sebesar 96% pada 755 sampel data. Kelas 5 memiliki performa terbaik dengan *precision* 0,97, *recall* sempurna (1,00), dan *F1-score* 0,98 dari 313 sampel. Sebaliknya, kelas 3 memiliki *F1-score* terendah 0,60 dari 18 sampel. *F1-score* kelas 3 memiliki nilai terendah (0,60) karena *recall* yang rendah (0,50), yang menunjukkan hanya setengah data kelas 3 terklasifikasi dengan benar. Hal ini disebabkan oleh jumlah data yang sedikit ($\text{support} = 18$) dan ketidakseimbangan data antar kelas, sehingga model kurang fokus pada kelas ini. Matriks konfusi menunjukkan model mampu mengklasifikasikan dengan baik kelas 0 dan 4, sementara kelas 2 memiliki performa terendah dengan hanya 9 prediksi tepat. Beberapa kesalahan terjadi, seperti data kelas 0 yang salah diklasifikasikan sebagai kelas 4 dan data kelas 1 yang salah ke kelas 0. Secara keseluruhan, model cukup andal, meskipun perlu peningkatan untuk klasifikasi kelas minoritas. Hasil pada Tabel 11 dan Tabel 12.

Tabel 11. Hasil Model KNN

	Precision	Recall	F1-Score	Support
1	0.96	0.96	0.96	251
2	0.91	0.94	0.92	94
3	0.75	0.50	0.60	18
4	0.97	0.89	0.93	79
5	0.97	1.00	0.98	313
Accuracy			0.96	755
Macro Avg	0.91	0.86	0.88	755
Weighted Avg	0.95	0.96	0.95	755

Tabel 12. Confusion Matrix KNN

Actual / Predicted	Class 0	Class 1	Class 2	Class 3	Class 4
Class 0	242	2	0	0	7
Class 1	5	88	1	0	0
Class 2	3	4	9	2	0
Class 3	1	3	2	70	3
Class 4	0	0	0	0	313

4. Kesimpulan

Berdasarkan hasil pengolahan dataset dan hasil evaluasi model, algoritma KNN dan SVM menunjukkan performa yang baik dengan akurasi masing-masing 0,96. SVM sedikit lebih unggul pada beberapa kelas, terutama kelas 5, sementara KNN memiliki nilai *precision*, *recall*, dan *F1-score* yang konsisten tinggi. Adapun harapan untuk penelitian selanjutnya dapat mengumpulkan lebih banyak data, terutama untuk kelas yang kurang terwakili. Selain itu, penerapan *cross-validation* dan pemantauan berkala terhadap model akan membantu memastikan akurasi dan relevansi dalam pengambilan keputusan terkait prioritas pelanggan. Dengan optimasi lebih lanjut, kedua model memiliki potensi untuk mendukung pengambilan keputusan yang lebih akurat terkait prioritas pelanggan.

Daftar Pustaka

- [1] A. Alimuddin, "PENGUNAAN INTERNET DAN PELUANG BERWIRSAUSAHA DI INDONESIA," *Jurnal Kewirausahaan dan Bisnis*, vol. 26, no. 2, Art. no. 2, Dec. 2021, doi: 10.20961/jkb.v26i2.50913.
- [2] B. P. S. Indonesia, "Statistik Telekomunikasi Indonesia 2022." Accessed: Jan. 15, 2025. [Online]. Available: <https://www.bps.go.id/id/publication/2023/08/31/131385d0253c6aae7c7a59fa/statistik-telekomunikasi-indonesia-2022.html>
- [3] A. Triwijayati, Y. P. Luciany, Y. Novita, N. Sintesa, and A. Zahrudin, "Strategi Inovasi Bisnis untuk Meningkatkan Daya Saing dan Pertumbuhan Organisasi di Era Digital," *Jurnal Bisnis dan Manajemen West Science*, vol. 2, no. 03, Art. no. 03, Aug. 2023, doi: 10.58812/jbmws.v2i03.564.
- [4] I. N. R. Hendrawan, I. M. A. B. Saputra, G. A. P. C. Dewi, I. G. S. A. Pranata, and N. L. N. Wedasari, "Klasifikasi Lama Studi dan Predikat Kelulusan Mahasiswa menggunakan Metode Naïve Bayes," *Jurnal Eksplora Informatika*, vol. 11, no. 1, Art. no. 1, 2021, doi: 10.30864/eksplora.v11i1.606.
- [5] "(PDF) Regression Based Machine Learning Model for Rainfall Forecasting on Daily Weather Data," in *ResearchGate*, doi: 10.1109/ICIC60109.2023.10381974.
- [6] I. P. P. B. Lantara, I. P. D. Payana, G. A. S. Pratama, W. E. A. Munayana, K. S. Nopiani, and I. N. R. Hendrawan, "GETARI: Dataset untuk Klasifikasi Gerakan Dasar Tari Bali Perempuan," *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI*, vol. 11, no. 3, Art. no. 3, Dec. 2022, doi: 10.23887/janapati.v11i3.52598.
- [7] "(PDF) Perbandingan Kinerja Algoritma K-Nearest Neighbor, Naïve Bayes Classifier dan Support Vector Machine dalam Klasifikasi Tingkah Laku Bully pada Aplikasi Whatsapp," *ResearchGate*, Oct. 2024, doi: 10.30998/faktorexacta.v12i2.4181.
- [8] V. Arinal and I. Rusmarhadi, "Implementasi Data Mining Untuk Menentukan Strategi Penjualan Produk Umkm Raja Geprek Pada Pola Pembelian Konsumen Menggunakan Algoritma Apriori," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 7, no. 5, pp. 1482–1494, Sep. 2024, doi: 10.31539/intecom.s.v7i5.11673.

-
- [9] F. M. Herza, B. Rahmat, and M. M. A. Haromainy, "PENGARUH RFE TERHADAP LOGISTIC REGRESSION DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN HOTEL SHANGRI-LA SURABAYA," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 6, Art. no. 6, Nov. 2024, doi: 10.36040/jati.v8i6.11272.
- [10] G. A. Pradipta, M. Liandana, P. D. W. Ayu, D. P. Hostiadi, and P. S. E. Putra, "Voting Scheme Nearest Neighbors by Difference Distance Metrics Measurement," *JUITA: Jurnal Informatika*, vol. 11, no. 2, p. 165, Nov. 2023, doi: 10.30595/juita.v11i2.19298.