

ANALISIS KLASIFIKASI ALGORITMA C.45 DALAM MEMPREDIKSI TINGKAT PENJUALAN PERLENGKAPAN IBADAH

Novrianto¹
Erlin Elisa²

¹Program Studi Sistem Informasi, Universitas Putera Batam

²Program Studi Sistem Informasi, Universitas Putera Batam

email: pb201510002@upbatam.ac.id

ABSTRACT

This research was conducted at a shop that sells goods or equipment for Buddhist worship called the Sinar Mas Indah shop. The Sinar Mas Indah shop operates every day except major holidays such as Chinese New Year. So far, the Sinar Mas Indah store has never used sales transaction data for analysis for the benefit of future business progress, sales transaction data is only stored in archive form, if it is old and feels no longer needed, it will simply be thrown away. , if used to predict future sales then decisions to increase sales in the future can be easily achieved. This research utilized datamining techniques with the C4.5 algorithm by testing the results using Weka software, from the 5 variables tested, category, brand, quantity, quality and price with a total of 160 test data. After calculating the entropy and gain values, the highest gain value was obtained. making consideration of consumer purchasing factors, namely Price, Quality, Quantity and category, into the main factors in decision making and there are results for Correctly Classified Instances 109 Data with a percentage value of 68.125% then for Incorrectly Classified Instances 51 data with a percentage value of 31.875%.

Keywords: *Datamining, Prediction, C4.5 Algorithm, Weka, Sales*

PENDAHULUAN

Perengkapan ibadah merupakan bagian integral dalam praktik keagamaan di berbagai agama dan kepercayaan di seluruh dunia salah satunya kepercayaan agama Buddha. Perlengkapan ibadah mencakup berbagai jenis barang, seperti pakaian ritual, dupa, lilin, minyak sembahyang, kertas sembahyang dan masih banyak lagi. Perlengkapan ibadah dapat mencerminkan berbagai aspek dalam masyarakat, termasuk kepatuhan agama, perubahan sosial, dan tren budaya. Oleh karena itu, memprediksi tingkat penjualan perlengkapan ibadah dapat memiliki implikasi yang signifikan dalam berbagai konteks.dengan paragraf

berikutnya, dan justify. Pendahuluan jurnal berisikan latar belakang permasalahan penelitian, yang diperkuat dengan urgensi penelitian dan rasionalisasi penelitian yang dapat diambil dari hasil penelitian terdahulu. Selain itu pendahuluan berisikan tujuan penelitian dan rencana pemecahan masalah pada penelitian.

Toko Sinar Mas Indah menjual berbagai perlengkapan untuk ibadah mulai dari kertas sembahyang, dupa, lilin, minyak sembahyang dan lain sebagainya. Toko Sinar Mas Indah juga menjual perlengkapan ibadah di media sosial seperti facebook, instagram, dan whatsapp guna memudahkan pembeli

untuk mendapatkan barang atau perlengkapan ibadah yang mereka inginkan, dan efektifitas lainnya yaitu tahap pembayaran, toko Sinar Mas Indah juga mendukung sistem pembayaran bank contohnya transfer, qris, dan ada juga sistem pembayaran secara tunai.

Selama ini toko Sinar Mas Indah belum pernah menggunakan data transaksi penjualan untuk di analisis demi kepentingan kemajuan usaha dimasa yang akan datang, data-data transaksi penjualan hanya di simpan dalam bentuk arsip, jika sudah lama dan merasa tidak di butuhkan lagi maka akan dibuang begitu saja, apabila dimanfaatkan untuk memprediksi penjualan kedepannya maka keputusan untuk peningkatan penjualan dimasa yang akan datang dapat mudah dicapai

Pada masa sekarang, teknologi dan data telah menjadi bagian tak terpisahkan dalam berbagai aspek kehidupan kita. Dengan adanya kemajuan teknologi informasi, kita memiliki akses lebih besar ke data dan alat analisis data yang kuat. Salah satu algoritma yang digunakan dalam analisis klasifikasi data adalah algoritma C4.5, Algoritma ini telah terbukti efektif dalam memprediksi berbagai macam kejadian dan pola dalam data, sesuai dengan penelitian sebelumnya C4.5 dimanfaatkan untuk memprediksi tingkat penjualan pada sebuah toko bangunan dengan hasil yang menunjukkan bahwasanya harga, kualitas dan garansi menjadi pilihan utama konsumen untuk membeli suatu produk (Elisa, 2022).

KAJIAN TEORI

2.1 Penjualan

Penjualan merupakan suatu sistem kerja yang terdiri dari berbagai komponen yang saling terkait dan berinteraksi untuk merumuskan rencana strategi di sektor

usaha dengan tujuan memenuhi kebutuhan konsumen dan pencapaian tujuan tertentu (Surahmat, 2023).

2.2 Algoritma C4.5

Salah satu algoritma yang digunakan untuk mengubah informasi besar menjadi pohon keputusan (decision tree) yang menggambarkan aturan adalah C4.5. Terdapat beberapa perbaikan yang telah diterapkan pada C4.5, seperti kemampuan untuk mengatasi data yang hilang (missing value), data berkelanjutan (data kontinu), dan teknik pruning. Algoritma C4.5 diperkenalkan oleh J. Ross Quinlan sebagai pengembangan dari algoritma ID3, dan algoritma ini digunakan untuk membentuk pohon keputusan. Pohon keputusan dianggap sebagai salah satu pendekatan yang paling terkenal dalam klasifikasi. Pohon keputusan terdiri dari simpul-simpul, dan simpul pertama disebut akar, yang tidak memiliki simpul masukan (Adriansa et al., 2022)

2.3 Data Mining

Menambang data dari *database* yang selama ini, dapat menghasilkan informasi baru seperti penelitian analisa peramalan trend mata uang kripto sehingga dapat membantu orang yang tertarik dalam investasi dibidang ini dapat mengetahui trend bitcoin saat ini, jadi dapat dikatakan, Data Mining merupakan tahap perolehan data atau biasa lebih dikenal proses penambang data yang bertujuan untuk mencari pola-pola yang menarik dengan maksud untuk mendapatkan pengetahuan dari suatu dataset yang memiliki volume data yang sangat besar, seperti yang ditemukan dalam Big Data (Julianto et al., 2022).

2.4 KDD (*Knowledge Discovery in Database*)

Dalam menganalisis data untuk mencari informasi baru tentunya ada beberapa

langkah yang harus di jalani seperti langkah pada *Knowledge Discovery in Database*, langkah ini telah digunakan pada penelitian dalam pemilihan jenis laptop dimana metode ini digunakan untuk memperoleh pengetahuan dari basis data yang ada, sehingga keputusan diambil berdasarkan pengetahuan yang diperoleh melalui tahapan yang membentuk dasar pengetahuan, dengan hasil penelitian terdapat tiga variabel yang sangat mempengaruhi pemilihan laptop, yaitu harga, prosesor dan masa pakai baterai (Baginda, 2023).

2.5 Decision Tree

Pohon keputusan disebutkan sebagai metode yang sangat didukung dan dikenal dengan baik dalam kelompok klasifikasi dan prediksi. Metode keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang mewakili aturan. Aturan dapat dengan mudah dipahami dalam bahasa alami. Dan mereka juga dapat diungkapkan dalam bentuk basis data seperti bahasa Structure Query untuk mencari catatan dalam kategori tertentu (Darmawan, 2018).

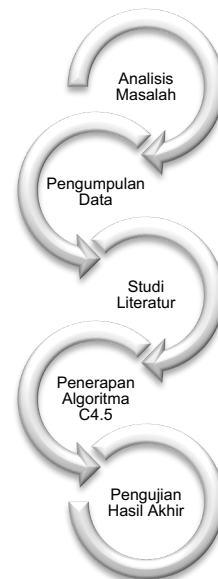
2.6 Waikato Environment for Knowledge Analysis (WEKA)

Waikato Environment for Knowledge Analysis, yang disingkat sebagai WEKA, adalah kumpulan alat praktis untuk mesin learning yang dikembangkan di Universitas Waikato, Selandia Baru, yang memiliki beragam aplikasi dan keahlian. WEKA memiliki sejumlah keunggulan, terutama dalam hal klasifikasi, yang sangat efektif dalam menyelesaikan permasalahan penambangan data di dunia nyata. WEKA juga mudah digunakan dalam berbagai tahap, serta menyediakan alat-alat untuk pra-pemrosesan data,

klasifikasi, regresi, pengelompokan, aturan asosiasi, dan visualisasi. Aplikasi ini mendukung beberapa metode dalam penambangan data, terutama pohon keputusan, yang melibatkan node keputusan yang menghubungkan berbagai bagian mulai dari simpul akar hingga simpul daun (akhir). Ini merupakan bagian menarik dalam metode klasifikasi, di mana indikasi pada simpul keputusan diuji, dan setiap tindakan mengarah ke bagian yang sesuai. Setiap bagian kemudian di arahkan ke simpul lain atau simpul akhir untuk menghasilkan keputusan akhir (Baginda, 2023).

METODE PENELITIAN

Berikut merupakan desain penelitian yang peneliti rangkum dalam bentuk diagram antara lain:



Gambar 1. Diagram Desain Penelitian (Sumber: Data Penelitian, 2024)

HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian dengan Algoritma C4.5

Prediksi penjualan pada toko Sinar Mas Indah dengan menggunakan informasi yang telah dikumpulkan selama ini. Untuk proses analisis data peneliti akan memanfaatkan data transaksi

4.1.1 Pra Proses Data

Penetapan format untuk proses penelitian lebih lanjut melibatkan seleksi informasi-informasi yang telah dikumpulkan sebelumnya untuk proses

penjualan selama ini, dalam hal ini data yang dipakai sebanyak 160 data sample transaksi dengan variabel yang telah ditentukan antara lain kategori, merek, kuantitas, kualitas, dan harga, sedangkan variabel keputusan yang digunakan adalah beli dan tidak beli.

penelitian, sehingga memperoleh sejumlah atribut yang telah ditentukan. Data untuk proses penelitian lebih lanjut disajikan dalam tabel berikut.

Tabel 1. Data Penjualan Barang

Kategori	Merek	Kuantitas	Kualitas	Harga	Keputusan
KS 5858	Bagus	Sedang	Bagus	Tinggi	Beli
KS 1299	Sedang	Banyak	Bagus	Menengah	Beli
KS 999	Sedang	Sedikit	Tidak Bagus	Rendah	Beli
KS 999	Biasa	Sedikit	Tidak Bagus	Rendah	Tidak Beli
KS 1299	Bagus	Sedikit	Bagus	Rendah	Beli
KS 999	Biasa	Sedang	Bagus	Rendah	Beli
KS 1299	Sedang	Sedang	Cukup	Menengah	Tidak Beli
DP SH	Biasa	Sedang	Tidak Bagus	Menengah	Beli
.....
.....
.....
KS 999	Biasa	Banyak	Tidak Bagus	Rendah	Tidak Beli
KS 1299	Sedang	Sedang	Cukup	Menengah	Beli

(Sumber: Data Penelitian, 2024)

4.1.2 Hasil Perhitungan Manual C4.5

Data-data yang telah diseleksi akan diubah ke dalam bentuk format yang kemudian akan diaplikasikan. dibawah ini merupakan pengelompokan yang

didasarkan pada atribut-atribut tertentu antara lain:

Tabel 2. Klasifikasi Atribut Kategor

Atribut	Nilai
Kategori	KS 5858
	KS 1299
	KS 999
	LL HC
	DP XTG
	LL YY
	DP YZX
	DP SH
	LL KY

Tabel 3. Klasifikasi Atribut Merek

Atribut	Nilai
Merek	Bagus
	Sedang
	Biasa

Tabel 4. Klasifikasi Atribut Kuantitas

Atribut	Nilai
Kuantitas	Banyak
	Sedang
	Sedikit

Tabel 7. Perhitungan Node 1

Variabel	Kasus	Beli	Tidak Beli	Entropy	Gain
Total	160	105	55	0,9283	
Kategori					
KS 5858	23	8	15	0,9321	0,0924
KS 1299	32	24	8	0,8112	
KS 999	41	23	18	0,9892	

Tabel 5. Klasifikasi Atribut Kualitas

Atribut	Nilai
Harga	Tinggi
	Menengah
	Rendah

Tabel 6. Klasifikasi Atribut Harga

Kemudian langkah berikutnya Menentukan atribut sebagai *node root* berdasarkan perhitungan nilai *gain* dari atribut yang tersedia, untuk mencari atribut yang memiliki nilai *gain* tertinggi. untuk analisis kasus ini peneliti melakukan iterasi hingga node 3 berikut hasil pergitungan node 1, 2 dan 3 pada tabel dibawah ini:

Atribut	Nilai
Kualitas	Bagus
	Cukup
	Tidak Bagus

LL HC	10	8	2	0,7219	
DP XTG	10	7	3	0,8812	
LL YY	10	6	4	0,9709	
DP YZX	15	12	3	0,7219	
DP SH	10	9	1	0,4689	
LL KY	9	8	1	0,5032	
Merek					
Bagus	46	27	19	0,9780	
Sedang	55	43	12	0,7568	0,0275
Biasa	59	35	24	0,9747	
Kuantitas					
Banyak	11	8	3	0,8453	
Sedang	53	32	21	0,9687	0,0048
Sedikit	96	65	31	0,9075	
Kualitas					
Bagus	49	35	14	0,8631	
Cukup	51	37	14	0,8478	0,0214
Tidak Bagus	60	33	27	0,9927	
Harga					
Tinggi	23	8	15	0,9321	0,0548
Menengah	76	57	19	0,8112	
Rendah	61	40	21	0,9288	

Tabel 8. Perhitungan Node 2

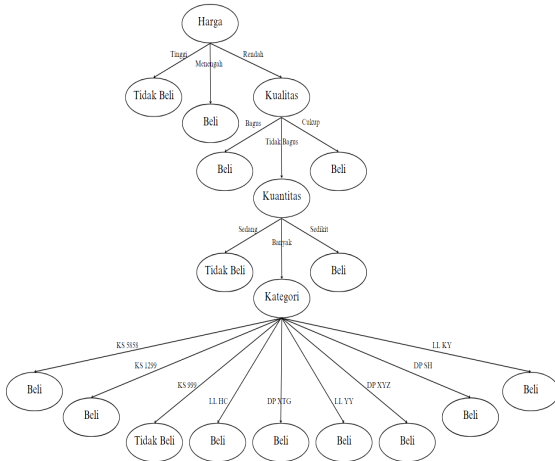
Variabel	Kasus	Beli	Tidak Beli	Entropy	Gain
Total	61	40	21	0,9288	
Kategori					
KS 1299	1	1	0	0	0,0713
KS 999	41	23	18	0,9892	
LL HC	10	8	2	0,7219	
LL KY	9	8	1	0,5032	
Merek					
Bagus	1	1	0	0	0,0988
Sedang	14	13	1	0,3712	
Biasa	46	26	20	0,9876	
Kuantitas					
Banyak	6	4	2	0,9182	0,0246
Sedang	19	10	9	0,9980	
Sedikit	36	26	10	0,8524	
Kualitas					
Bagus	5	5	0	0	0,1304
Cukup	12	11	1	0,4138	
Tidak Bagus	44	24	20	0,9940	

Tabel 9. Perhitungan Node 3

Variabel	Kasus	Beli	Tidak Beli	Entropy	Gain
Total	44	24	20	0,9940	
Kategori					
KS 999	35	17	18	0,9994	0,0427
LL HC	9	7	2	0,7642	
Merek					
Sedang	2	2	0	0	0,0410
Biasa	42	22	20	0,9983	
Kuantitas					
Banyak	5	3	2	0,9709	0,0739
Sedang	11	3	8	0,8453	
Sedikit	28	18	10	0,9402	



Dengan demikian maka terbentuklah pohon keputusan sebagai berikut :



Gambar 2. Pohon Keputusan

Dari Pohon keputusan tersebut maka di dapat aturan berikut:

1. IF Harga = Tinggi: Tidak Beli
2. IF Harga = Menengah: Beli
3. Harga = Rendah and Kualitas = Bagus: Beli
4. IF Kualitas = Tidak Bagus and Kuantitas = Sedang: than keputusan =Tidak Beli
5. IF Kuantitas = Banyak and Kategori = KS 5858: KS 1299: than keputusan = Beli
6. IF Kategori = KS 999: Tidak Beli
7. IF Kategori = LL HC:DP XTG: LL YY: Beli DP YZX : DP SH than keputusan Beli
8. IF Kategori = LL KY: Beli, Kuantitas = Sedikit : Beli and Kualitas = Cukup: Beli

Berdasarkan hasil aturan rules yang terbentuk maka dilihat bahwasanya faktor yang menjadi penilaian utama dalam prediksi penjualan yaitu harga, kualitas, kuantitas dan kategori.

4.2 Pembahasan dan Pengujian

Tahap selanjutnya akan dilakukan pengujian dengan memanfaatkan aplikasi *open source weka* untuk mengetahui hasil prediksi *Classifier Output* sebagai berikut:

```

Classifier output
=====
= Summary =
Correctly Classified Instances      109      68.125 %
Incorrectly Classified Instances    51       31.875 %
Kappa statistic                    0.2154
Mean absolute error                 0.4084
Root mean squared error             0.4679
Relative absolute error             90.3444 %
Root relative squared error         98.4631 %
Total Number of Instances          160

= Detailed Accuracy By Class =
Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area | Class
-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----
Beli  | 0.681   | 0.487   | 0.660     | 0.681   | 0.655     | 0.230 | 0.609   | 0.443   | Beli
Tidak Beli | 0.319   | 0.513   | 0.340     | 0.319   | 0.325     | 0.230 | 0.391   | 0.557   | Tidak Beli

= Confusion Matrix =
a b <-- classified as
1 14 | a = Beli
7 18 | b = Tidak Beli
    
```

Gambar 3. Classifier Output

Berdasarkan gambar diatas tertera bahwasanya dari data training sebanyak 160 data hasil *Correctly Classified Instances* 109 Data dengan nilai persentase 68.125 % kemudian untuk *Incorrectly Classified Instances* 51 data dengan nilai persentase 31.875 %. Dengan nilai *Kappa statistic* 0.2154 *Mean absolute error* 0.4084, *Root mean squared error* 0.4679, *Relative absolute error* 90.3444%, *Root relative squared error* 98.4631 %

SIMPULAN

Hasil analisis data penjualan menggunakan algoritma C4.5 menunjukkan prediksi tingkat penjualan perlengkapan ibadah di toko Sinar Mas Indah dengan faktor yang menjadi pertimbangan yaitu Harga, Kualitas, Kuantitas dan Kategori dimana atribut tersebut menunjukkan nilai gain tertinggi pada masing-masing *node*.

Pohon keputusan yang dibentuk dalam analisis algoritma C4.5 pada toko Sinar Mas Indah untuk memprediksi tingkat penjualan perlengkapan ibadah menunjukkan kesamaan dengan hasil manual dengan atribut Harga, Kualitas, Kuantitas dan Kategori merupakan hasil dari pohon keputusan dengan hasil *Correctly Classified Instances* 109 Data dengan nilai persentase 68.125 % kemudian untuk *Incorrectly Classified Instances* 51 data dengan nilai persentase 31.875%.

DAFTAR PUSTAKA

- Adriansa, M., Yulianti, L., & Elfianty, L. (2022). Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5. *Jurnal Teknik Informatika UNIKA Santo Thomas*, 07(21), 115–121. <https://doi.org/10.54367/jtiust.v7i1.1983>
- Amrin, A., & Pahlevi, O. (2021). Data Mining Optimization Based on Particle Swarm Optimization For Diagnosis of Inflammatory Liver Disease. *Journal of Informatics and Telecommunication Engineering*, 5(1), 152–159. <https://doi.org/10.31289/jite.v5i1.5312>
- Arif, A., & Yulianto, A. (2022). Klasifikasi Penjualan Makanan Hewan Peliharaan di Muezza Petshop01 Bogor Menggunakan Metode Algoritma C4.5. *Remik*, 6(4), 840–854. <https://doi.org/10.33395/remik.v6i4.11859>
- Baginda, B. (2023). Implementasi Data Mining Dalam Pemilihan Laptop Berbasis Algoritma C4.5 Pada Software WEKA. *Jurnal Minfo Polgan*, 12(1), 1065–1073. <https://doi.org/10.33395/jmp.v12i1.12582>
- Darmawan, E. (2018). C4.5 Algorithm Application for Prediction of Self Candidate New Students in Higher Education. *Jurnal Online Informatika*, 3(1), 22. <https://doi.org/10.15575/join.v3i1.171>
- Diansyah, T. M., & Exprada, Y. (2022). Penerapan Algoritma C4.5 Untuk Memprediksi Penjualan Lele Pada Kolam Pancing Galatama. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 3(4), 567. <https://doi.org/10.30865/json.v3i4.4264>
- Elisa, E. (2022). Penerapan Algoritma C4.5 Untuk Memprediksi Penjualan Barang Pada Pt Batam Bangun Prathama. *Jurnal Comasie*, 7(1), 127–133.
- Iftitah, A., & Setyadi, R. (2023). Penerapan Algoritma C.45 Untuk Analisis Pengadaan Peralatan dan Mesin Kantor. *Journal of Information System Research (JOSH)*, 4(2), 434–442. <https://doi.org/10.47065/josh.v4i2.2673>
- Julianto, I. T., Kurniadi, D., Nashrulloh, M. R., & Mulyani, A. (2022). Comparison of Data Mining Algorithm For Forecasting Bitcoin Crypto Currency Trends. *Jurnal Teknik Informatika (Jutif)*, 3(2), 245–248. <https://doi.org/10.20884/1.jutif.2022.3.2.194>
- Mulyanda, S., Defit, S., & Sumijan. (2023). Analisis Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Harga Pasar Mobil Bekas. *Jurnal KomtekInfo*, 10, 116–121. <https://doi.org/10.35134/komtekinfo.v10i3.427>
- Rusdiansyah, R., Supendar, H., Suharyanti, N., & Tuslaela, T. (2022). Data Mining implementation on SMUN Scholarship recipient candidates using the C4.5 algorithm. *Sinkron*, 7(4), 2208–2213. <https://doi.org/10.33395/sinkron.v7i4.11767>
- Salsabilla, T., & Sulastri, S. (2022). Implementasi Algoritma C4.5 Untuk Klasifikasi Produk Laris Sepeda Motor Honda Pada Cv Cendana Motor Cepiring. *Rabit: Jurnal Teknologi Dan Sistem Informasi Univrab*, 7(2), 164–171. <https://doi.org/10.36341/rabit.v7i2.2489>
- Sulastri, R. (2022). Identifikasi Tingkat Penjualan Produk Herbal HWI Menggunakan Algoritma C4.5. *Jurnal Informatika Ekonomi Bisnis*, 4(3), 84–89. <https://doi.org/10.37034/infec.v4i3.141>
- Surahmat, A. (2023). Rancang Bangun Aplikasi Sistem Penjualan Pada Percetakan Cubic Art. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 81–86. <https://doi.org/10.36040/jati.v7i1.6064>
- Tanyu, B. F., Abbaspour, A., Alimohammadlou, Y., & Tecuci, G. (2021). Landslide susceptibility analyses using Random Forest, C4.5, and C5.0 with balanced and unbalanced datasets. *Catena*, 203(April), 105355. <https://doi.org/10.1016/j.catena.2021.105355>
- Wahono, H., & Riana, D. (2020). Prediksi Calon Pendoron Darah Potensial Dengan Algoritma Naïve Bayes, K-Nearest Neighbors dan Decision Tree C4.5. *JURIKOM (Jurnal Riset Komputer)*, 7(1), 7.

- <https://doi.org/10.30865/jurikom.v7i1.1953>
Wijaya, A., & Mutatkin Bakti, A. (2023). Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Memprediksi Penjualan Alat – Alat Telekomunikasi Dan IT (Information Technology) Pada PT Julian Adiputra Utama. *Jupiter*, 15, 597–606.
- Yusuf Maulana, Riki Winanjaya, & Fitri Rizki. (2022). Penerapan Data Mining dengan Algoritma C.45 Dalam Memprediksi Penjualan Tempe. *Bulletin of Computer Science Research*, 2(2), 53–58.
<https://doi.org/10.47065/bulletincsr.v2i2.163>



Penulis Pertama
Novrianto, merupakan mahasiswa Prodi Sistem Infomasi Universitas Putera Batam.



Penulis Kedua Erlin Elisa merupakan Dosen Prodi Sistem Informasi Universitas Putera Batam. Penulis banyak berkecimpung di bidang Sistem Informasi