

IMPLEMENTASI DATA MINING PADA PENGATURAN DATA INVOICE DISTRIBUTOR MENGGUNAKAN ALGORITMA FP GROWTH

Piter Prasetyo Sudarto¹,
Koko Handoko²

¹Mahasiswa Program Studi Teknik Informatika, Universitas Putera Batam

²Dosen Program Studi Teknik Informatika Universitas Putera Batam
email: pb190210068@upbatam.ac.id

ABSTRACT

A research was conducted at PT. Wyssa Artha Sejahtera a distribution company, to explore the processing of sales transaction data using the FP Growth algorithm to support better decision-making. The company deals with a large volume of daily sales transaction data involving various types of products, both wet and dry. The study aimed to extract useful information from the sales data, such as sales trends and the most popular products among customers. Data was collected through observations and direct interviews with the company owner from January 2022 to December 2022, using the proactive method for more comprehensive and accurate information. Upon collecting the sales transaction data, the research identified certain products with higher sales than others. RapidMiner software was utilized for processing the sales data, which proved to be suitable for implementing the FP Growth algorithm, especially for wet and dry product types with increasing sales transaction data. In the testing phase, RapidMiner successfully discovered item sets 1, 2, 3, and 4, along with their corresponding support and confidence values. The results of this research carry significant implications for PT. Wyssa Artha Sejahtera in making more informed decisions. By analyzing the sales transaction data, the company can devise more effective strategies to boost sales and meet customer demands. Furthermore, the findings can serve as a valuable reference for other companies in similar industries, helping to enhance their decision-making processes. The FP Growth algorithm analysis revealed that item sets 2, 3, and 4 had a minimum support of 20% and a minimum confidence of 70%.

Keywords : Association Rule, Data Mining ,FP-Growth ,RapidMiner

PENDAHULUAN

Di era komputerisasi ini, kebutuhan manusia akan informasi mendorong pesatnya perkembangan teknologi di bidang TIK. Teknologi yang berkembang didukung oleh sarana dan kebutuhan masyarakat yang beroperasi dengan globalisasi. Sistem informasi dan teknologi informasi memberikan dukungan untuk membuat keputusan yang tepat berdasarkan informasi yang

tersedia. Dalam perkembangan teknologi informasi yang ada dapat dengan mudah mengolah data, dapat menghasilkan informasi yang dibutuhkan secara akurat dan tepat waktu, serta biaya yang lebih efektif.

PT.Wyssa Artha Sejahtera merupakan distributor yang bergerak pada bidang pangan yang mendistribusi berbagai merek Oriental Food Industries Sdn. Bhd. Distribusi barang yang

dilakukan oleh PT.Wyssa Artha Sejahtera merupakan distributor pertama di Batam yang melakukan distribusi merek Oriental di area Batam, sehingga distribusi yang dilakukan akan memiliki jumlah transaksi yang besar. Distribusi yang besar ini akan memberikan berbagai macam data terutama Invoice yang merupakan dokumen yang berisikan isi transaksi dan pembayaran yang terjadi antara penjual kepada pembeli dalam jumlah yang besar.

Invoices yang terdapat pada di PT. Wyssa Artha Sejahtera memiliki berbagai macam data penjualan yang berdampak menumpuknya data-data yang tidak di olah secara efektif oleh staff yang bekerja di PT. Wyssa Artha Sejahtera. Masalah ini menimbulkan banyak kendala terutama pada penjualan yang tidak memiliki peningkatan yang konsisten dan kurangnya efektifitas penjualan dikarenakan pola transaksi tidak pernah dianalisa pola transaksi penjualan tersebut.

Berdasarkan pendahuluan di atas maka diciptakan karya ilmiah **IMPLEMENTASI DATA MINING PADA PENGATURAN DATA INVOICE DISTRIBUTOR MENGGUNAKAN ALGORITMA FP GROWTH**

KAJIAN TEORI

2.1 *Knowledge Discovery in Database (KDD)*

KDD merupakan pengumpulan dan pemanfaatan data historis dengan tujuan mengidentifikasi keteraturan, pola, atau hubungan dalam catatan dalam database besar yang sebelumnya tidak diketahui dan memiliki nilai (Bunda, 2020). Dalam pengumpulan data, KDD menggunakan data historis untuk menemukan

pengetahuan, informasi, keteraturan, pola atau hubungan dalam data yang berukuran besar. KDD merupakan suatu kegiatan yang dilakukan dalam rangka memperoleh informasi baru dari data yang ada (Fajrin & Handoko, 2018)

Penemuan pengetahuan dalam database (KDD) terdiri dari ekstraksi informasi implisit potensial dan satu set data yang tidak diketahui (Simanjuntak & Elisa, 2019). Dengan memasukkan hasil proses data mining, yaitu proses penemuan pengetahuan, informasi diubah menjadi hasil yang akurat dan lebih mudah dipahami. Maka berdasarkan pengertian di atas dapat disimpulkan bahwa pengertian dari KDD adalah proses penting untuk menemukan dan mengenali pola dalam data dan bentuk yang dihasilkan baru, berguna, dan mudah dipahami. Hal ini dapat diperoleh dari informasi yang dihasilkan oleh data mining.

2.2 *Data Mining*

Data mining adalah dapat bermanfaat dan dimengerti dalam suatu database yang sangat besar, data mining merupakan proses iterative dan interactive untuk menemukan pola atau model baru yang sempurna mendefinisikan bahwa pengambilan keputusan di waktu yang akan datang dapat dibantu dengan data mining yang berisi pencarian trend atau pola yang diinginkan dalam database besar, yang mungkin saja menggunakan perangkat pendukung keputusan yang lain dari pola-pola yang dikenali perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan untuk kemudian dapat dipelajari serta dipahami dengan lebih teliti.

2.3 FP Growth

memiliki keunggulan dalam memproses data yang besar dan kompleks. Hal ini dikarenakan algoritma ini hanya memerlukan satu pemindaian data transaksi untuk membangun FP-Tree, yang dapat menghemat waktu pemrosesan data. Selain itu, FP-Growth juga memiliki kemampuan untuk menghasilkan itemset yang lebih besar dan kompleks dibandingkan dengan algoritma-aplikasi lainnya (Kumar & Mohbey, 2022)

Dalam konteks penelitian, FP-Growth dapat digunakan untuk mengatasi masalah yang melibatkan analisis pola distribusi barang dalam transaksi. Dengan menerapkan FP-Growth pada data historis transaksi yang besar, peneliti dapat mengidentifikasi pola-pola frekuensi tinggi yang dapat memberikan wawasan tentang tren penjualan, asosiasi antara item barang, dan preferensi pelanggan. Dengan demikian, FP-Growth memberikan kontribusi dalam pengambilan keputusan yang lebih baik dalam hal strategi pemasaran, manajemen persediaan, dan pengembangan produk. (Bunda, 2020).

$Confident P(A \cap B) = \frac{\Sigma(A \cap B)}{\Sigma A} \times 100\%$ <p>Rumus Confident</p>
$Support (A) = \frac{Jumlah\ Transaksi\ Yang\ Mengandung\ A}{Total\ Transaksi} \times 100\%$ <p>Rumus Support</p>
$Support (A,B) = \frac{Jumlah\ Transaksi\ Yang\ Mengandung\ A\ dan\ B}{Total\ Transaksi} \times 100\%$ <p>Rumus Support 2 Itemset</p>

Rumus 2.3 Rumus Support dan Confident
Sumber : Data Peneliti, 2023

2.4 Aturan Asosiasi

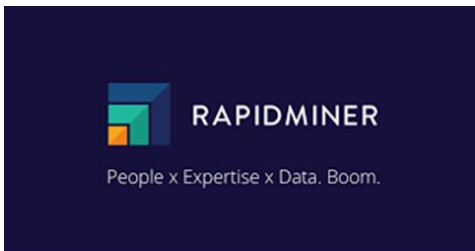
Tujuannya adalah untuk menunjukkan nilai asosiatif antara jenis produk yang dibeli pelanggan dan memungkinkan mereka untuk melihat pola dalam bentuk produk yang sering dibeli pelanggan. Dengan mengetahui jenis produk yang biasa dibeli, hal ini dapat dijadikan dasar untuk memutuskan produk mana yang pantas untuk diiklankan kepada pelanggan tersebut.

Association rule adalah konsep penting dalam analisis data yang digunakan untuk mengidentifikasi hubungan antara item-item dalam dataset. Dalam penelitian yang dilakukan oleh Agrawal, Imielinski, dan Swami (1993), association rule dijelaskan sebagai aturan yang menghubungkan antara suatu kelompok item atau atribut dengan kelompok item atau atribut lainnya dalam dataset. Aturan asosiasi ini sering digunakan dalam data mining untuk menemukan keterkaitan antara item atau atribut yang sering muncul bersama dalam suatu konteks. Proses penemuan association rule dilakukan melalui dua langkah utama, yaitu pembentukan kandidat rule dan pengujian kandidat rule. Dalam langkah pertama, kandidat rule dibangun dengan memilih item atau atribut yang sering muncul bersama dalam dataset. Kemudian, pada langkah kedua, kandidat rule tersebut diuji dengan menghitung tingkat dukungan dan tingkat kepercayaan untuk memilih rule yang memiliki nilai di atas batas ambang tertentu. Tingkat dukungan mengukur frekuensi kemunculan rule dalam dataset, sedangkan tingkat kepercayaan mengukur sejauh mana rule tersebut dapat dipercaya.

Penerapan association rule dalam analisis data memiliki banyak manfaat dalam berbagai bidang. Misalnya, dalam analisis keranjang belanja, association rule dapat membantu toko atau perusahaan untuk mengidentifikasi item yang sering dibeli bersama-sama oleh konsumen. Dengan mengetahui hubungan antara item, toko dapat meningkatkan strategi penjualan lintas produk dan meningkatkan pengalaman belanja pelanggan. Selain itu, association rule juga digunakan dalam rekomendasi sistem untuk menghasilkan rekomendasi produk yang relevan berdasarkan perilaku konsumen dan preferensi.

2.5 RapidMiner

Sebuah perangkat lunak yang dapat digunakan untuk mengolah data dengan berbagai macam fungsi dan operator unik. Perangkat ini digunakan mulai dari tahap pengolahan dan peletakan data hingga machine learning yang mampu menghasilkan perubahan pada data yang akan dimanipulasi (Fajrin & Handoko, 2018).



Gambar 1 RapidMiner

Sumber: <https://rapidminer.com>

2.6 Kerangka Pemikiran

Bentuk rangka pemikiran yang akan dikembangkan pada penerapan *Fp-Growth* pada PT. Wyssa Artha Sejahtera ini adalah sebagai berikut:



Gambar 2 Kerangka Pemikiran

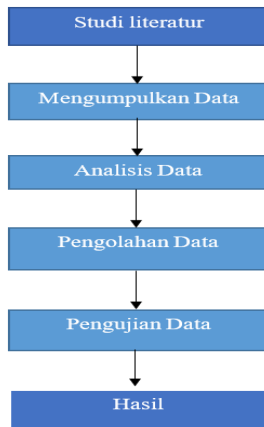
Sumber: Data Peneliti, 2023

Pada peta di atas, pada proses pertama penelitian ini data tentang penjualan barang sebagai proses input. bentuk input data merupakan produk/barang seperti merk dan jenis produk, serta jumlah total transaksi yang dilakukan ini serta disaring terlebih dahulu. Selain itu, data ini diproses oleh algoritma setelah data diproses dan mencapai nilai dukungan minimum, apriori kepercayaan minimal. Kemudian uji data yang akan kita gunakan menggunakan perangkat lunak RapidMiner untuk mendapatkan hasil yang akurat dan menemukan hasil penelitian tersebut sehingga hasil dari proses adalah produk yang paling sering dibeli konsumen berguna untuk toko yang ingin mengetahui pola penjualan.

METODE PENELITIAN

Metode penelitian yang dilakukan untuk penelitian ini adalah penelitian terapan. Metode penelitian ini menerapkan algoritma *FP-Growth* ke pengaturan *invoices* data penjualan PT.Wyssa Artha Sejahtera untuk menganalisis pola penjualan yang muncul dari sampel data transaksi Januari 2022 – Desember 2022. Dari sampel data tersebut akan di temukan pola penjualan setelah pembangkitan *FP-Tree* dan hasil akhirnya akan menemukan *Frequent Itemset Pattern* yang kemudian akan di implementasikan menggunakan aturan

asosiasi dengan menentukan nilai minimal *support* dan *confidence* pada masing *itemset*.



Gambar 3 Desain Penelitian
Sumber: Data Peneliti, 2023

Berdasarkan gambar 3 diatas maka tahapan yang akan dilakukan pada penelitian ini sebagai berikut:

1. Langkah pertama, mengidentifikasi masalah yang sedang diselidiki dan kemudian mengidentifikasi judul yang sedang diselidiki. Fase ini juga mengidentifikasi masalah, yaitu masalah pada PT. Wyssa Artha Sejahtera yang sedang dilakukan pada peneliti. Selanjutnya adalah analisis masalah yang meneliti hal-hal dan hal-hal secara rinci dan merakit dan meneliti bagian-bagian yang dijelaskan dari bagian-bagian.
2. Langkah kedua, data dikumpulkan untuk mencapai tujuan penelitian guna memperoleh informasi yang diperlukan. Dengan mengumpulkan data pada tahun Januari 2022-Desember 2022, peneliti memperoleh data observasi langsung dari pemilik toko, wawancara langsung

ke PT Wyssa Artha Sejahtera, dan awalnya berasal dari bahan penelitian literatur tentang data mining dan referensi ke jurnal dan buku nasional dan internasional.

3. Langkah ketiga, Data yang dikumpulkan dan dianalisis disiapkan dan diproses dalam proses data mining. Yaitu, algoritma FP-Growth untuk menentukan nilai *support* dan *Confidence* minimum untuk setiap produk yang dijual.

4. Langkah keempat, Implementasi dalam penelitian ini adalah penerapan data penjualan barang yang telah di filter menggunakan teknik algoritma FP-Growth yang akan di uji dengan perangkat pembantu RapidMiner.

5. Langkah kelima, hasil tersedia setelah semua tahapan atau proses penelitian. Kemudian akan menggunakan algoritma FP-Growth untuk menguji hasil menggunakan perangkat lunak RapidMiner dan memeriksa hasilnya untuk melihat apakah cocok dengan hasil.

6. Langkah terakhir yang merupakan tahap akhir penelitian, pada tahap ini akan menemukan hasil penelitian selidiki set item terlaris dan tentukan presentase produk penjualan di perusahaan.

HASIL DAN PEMBAHASAN

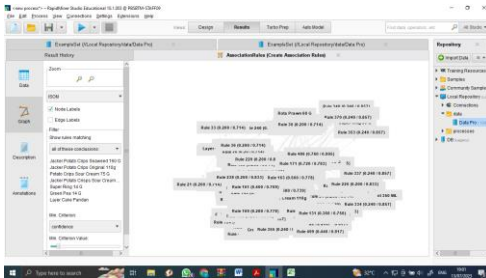
Hasil penelitian yang diolah berdasarkan data transaksi yang akan di sampel dari PT Wyssa Artha Sejahtera sebagai berikut:

Tabel 1. Data Transaksi

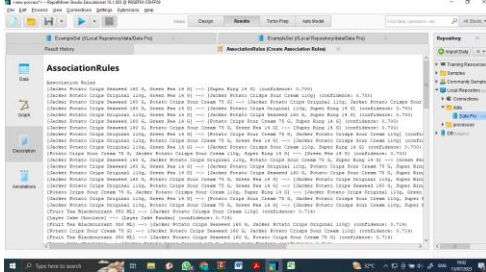
No Invoices	<i>ITEM YANG DIBELI</i>
CR 01002	Jacker Potato Crips Sour Cream 110g, Jacker Potato Crips Original 110g, Jacker Potato Crips Seaweed 160 G, Rota Prawn 14 G, Jacker Potato Crips Sour Cream 75 G
CR 02005	Layer Cake Banana 16 G, Rota Prawn 14 G, Super Ring 14 G
CR 03008	Fudo Swiss Roll Chocolate, Fudo Swiss Roll Pandan, Super Ring 14 G , Jacker Potato Crips Sour Cream 75 G
CR 04011	Layer Cake Chcolate, Layer Cake Pandan, Green Pea 14 G, Rota Prawn 14 G, Super Ring 14 G, Fruit Tea Apel 350 ML
CR 05014	Layer Cake Chcolate, Rota Prawn 14 G, Super Ring 14 G, Jacker Potato Crips Original 110g
CR 06017	Super Ring 14 G , Super Ring 60 G, Jacker Potato Crips Sour Cream 75 G, Layer Cake Chcolate
CR 07020	Rota Prawn 14 G, Super Ring 14 G, Rota Prawn 60 G, Fruit Tea Apel 350 ML, Fruit Tea Blackcurrant 350 ML, Fruit Tea Freeze 350 ML,
CR 08023	Jacker Potato Crips Original 110G, Jacker Potato Crips Sour Cream 110G, Cheese Ball 60 G, Green Pea 14 G, Rota Prawn 60 G, Fruit Tea Blackcurrant 350 ML
CR 09026	Green Pea 14 G, Rota Prawn 14 G, Super Ring 14 G, Jacker Potato Crips Seaweed 160 G, Fruit Tea Blackcurrant 350 ML
CR 10029	Rota Prawn 60 G, Super Ring 60 G, Fruit Tea Blackcurrant 350 ML, Cheese Ball 60 G
CR 11033	Rota Prawn 60 G, Cheese Ball 60 G, Green Pea 14 G, Rota Prawn 14 G, Super Ring 60 G, Jacker Potato Crips Sour Cream 110g
CR 12032	Layer Cake Banana, Rota Prawn 14 G, Rota Prawn 60 G, Super Ring 14 G, Green Pea 14 G, Jacker Potato Crips Sour Cream 110g
CR 05005	Super Ring 60 G, Layer Cake Chcolate, Layer Cake Pandan, Fruit Tea Blackcurrant 350 ML
CR 06008	Super Ring 60 G, Layer Cake Pandan, Layer

	Cake Chcolate,Cheese Ball 14 G, Green Pea 14 G, Super Ring 14 G, Cheese Ball 60 G,
CR 07011	Fruit Tea Apel 350 ML, Fruit Tea Blackcurrant 350 ML, Fruit Tea Freeze 350 ML
CR 08014	Super Ring 14 G, Teh Botol Original 350 ML
CR 09014	Cheese Ball 60 G, Teh Botol Original 350 ML
CR 10017	Super Ring 14 G, Green Pea 14 G, Cheese Ball 60 G,
CR 11021	Jacker Potato Crips Sour Cream 110 G, Super Ring 14 G, Super Ring 60 G
CR 12002	Cheese Ball 60 G, Green Pea 14 G, Rota Prawn 14 G Super Ring 14 G
CR 00105	Rota Prawn 14 G, Green Pea 14 G, Super Ring 60 G, Jacker Potato Crisps Sour Cream 110g, Fruit Tea Freeze 350 ML, Fruit Tea Blackcurrant 350 ML
CR 00301	Layer Cake Chcolate, Layer Cake Pandan, Green Pea 14 G,Super Ring 14 G, Teh Botol Original 350ml, Rota Prawn 60 G
CR 00901	Rota Prawn 14 G
CR 11003	Rota Prawn 14 G, Cheese Ball 60 G, Rota Prawn 60 G, Super Ring 60 G
CR 07207	Layer Cake Banana, Layer Cake Chcolate, Layer Cake Pandan,Super Ring 14 G, Green Pea 14 G, Rota Prawn 14 G

Langkah pertama yakni mengubah seluruh data transaksi menjadi tabulasi untuk di letakkan pada *RapidMiner* sebagai berikut:



Gambar 8 Hasil Data Graph
Sumber: Data peneliti,2023



Gambar 9 Hasil Data Deskriptif
Sumber: Data peneliti,2023

SIMPULAN

Pada penelitian yang dilakukan berdasarkan perhitungan manual yang telah diuji dengan *RapidMiner* dihasilkan akan memberikan 2 frequent itemset yang paling sering di beli / dijumpai pada data transaksi yang disampel yakni dengan minimal nilai support 20% dan Confidence 70%.Maka pola kombinasi terbaik yang didapatkan berupa Pembelian Layer Cake Pandan pasti akan membeli Layer Cake Chocolate dengan Confidence 100% dan nilai support tertinggi berupa setiap pembelian Rota Prawn 14 pasti akan membeli Super Ring 14 dengan support 32%.

DAFTAR PUSTAKA

AHMAD ADRI. (2021). Implementasi Data mining Menggunakan Algoritma

Apriori. Paper Knowledge . Toward a Media History of Documents , 6(2), 1–77.

Ait-Mlouk, A., Agouti, T., & Gharnati, F. (2017). Mining and prioritization of association rules for big data: multi-criteria decision analysis approach. *Journal of Big Data*, 4(1), 42. <https://doi.org/10.1186/s40537-017-0105-4>

Antari, N. M. D., Agustini, K., & Divayana, D. G. H. (2016). Studi Komparatif Model Pembelajaran Talking Stick Dan Snowball Throwing Terhadap Hasil Belajar Teknologi Informasi Dan Komunikasi (Tik) Siswa Kelas Xi Sma Negeri 1 Seririt Tahun Ajaran 2015/2016. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 13(2), 127–136. <https://doi.org/10.23887/jptk.v13i2.8521>


Bunda, Y. P. (2020). Algoritma FP-Growth Untuk Menganalisa Pola Pembelian Oleh-Oleh (Studi Kasus Di Pusat Oleh-Oleh Ummi Aufa Hakim). *Riau Journal of Computer Science*, 06(01), 34–44.

Data, I., Penjualan, M., Dengan, K., & Apriori, A. (2020). Jurnal Comasie IMPLEMENTASI DATA MINING PENJUALAN KOSMETIK DENGAN. 3.

Dogan, A., & Birant, D. (2021). Machine learning and data mining in manufacturing. *Expert Systems with Applications*, 166, 114060. <https://doi.org/10.1016/J.ESWA.2020.114060>

Fajrin, A. A., & Handoko, K. (2018). Penerapan Data mining Untuk Mengolah Association rule Tata Letak Buku Dengan Metode. *Jurnal Ilmiah Informatika (JIF)*, 2, 60–65.

- Hamid Mughal, M. J. (2018). Data mining: Web data mining techniques, tools and algorithms: An overview. *International Journal of Advanced Computer Science and Applications*, 9(6), 208–215.
<https://doi.org/10.14569/IJACSA.2018.090630>
- Kumar, S., & Mohbey, K. K. (2022). A review on big data based parallel and distributed approaches of pattern mining. *Journal of King Saud University - Computer and Information Sciences*, 34(5), 1639–1662.
<https://doi.org/10.1016/j.jksuci.2019.09.006>
- Kurnia, Y., Isharianto, Y., Giap, Y. C., Hermawan, A., & Riki. (2019). Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm. *Journal of Physics: Conference Series*, 1175(1).
<https://doi.org/10.1088/1742-6596/1175/1/012047>
- Prasetya, A., Andriana, S., & Komalasari, R. T. (2021). Rancang Bangun Inventarisasi Barang menggunakan Algoritma Apriori Sebagai Data mining. *Jurnal JTik (Jurnal Teknologi Informasi Dan Komunikasi)*, 5(4), 354.
<https://doi.org/10.35870/jtik.v5i4.223>
- Rahmawati, F., & Merlina, N. (2018). Metode Data mining Terhadap Data Penjualan Sparepart Mesin Fotocopy Menggunakan Algoritma Apriori. *PIKSEL : Penelitian Ilmu Komputer Sistem Embedded and Logic*, 6(1), 9–20.
<https://doi.org/10.33558/piksel.v6i1.1390>
- Sibarani, A. J. P. (2020). Implementasi Data mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 7(2), 262–276.
<https://doi.org/10.35957/jatisi.v7i2.195>
- Simanjuntak, P., & Elisa, E. (2019). Data mining Untuk Menentukan Pemilihan Celular Card Di Kota Batam. *Journal Information System ...*, 4(2), 1–5.
<https://ejournal.medan.uph.edu/index.php/isd/article/view/283%0Ahttps://ejournal.medan.uph.edu/index.php/isd/article/download/283/143>
- Simanjuntak, P., Suharyanto, C. E., Sitohang, S., & Handoko, K. (2022). Data mining Untuk Klasifikasi Status Pandemi Covid 19. *Jurnal Teknik Informatika Dan Komputer (Tekinkom)*, 5(2), 327.
<https://doi.org/10.37600/tekinkom.v5i2.620>

	<p>Biodata Penulis pertama, Piter Prasetyo Sudarto merupakan mahasiswa Prodi Teknik Informatika Universitas Putera Batam</p>
	<p>Biodata Penulis kedua, Koko Handoko, S.Kom., M.Kom. merupakan Dosen Prodi Teknik Informatika Universitas Putera Batam.</p>