

# IMPLEMENTASI DATA MINING MINAT *CUSTOMER* TOKO DURIAN KIMHUI DENGAN ALGORITMA APRIORI

Vincentsius Oliver Bryan<sup>1</sup>

Koko Handoko<sup>2</sup>

<sup>1</sup>Program Studi Teknik Informatika, Universitas Putera Batam

<sup>2</sup>Mahasiswa Studi Teknik Informatika, Universitas Putera

*email:* [pb190210070@upbatam.ac.id](mailto:pb190210070@upbatam.ac.id)

## ABSTRACT

Tanaman durian adalah tanaman perkebunan yang berasal dari negara yang beriklim tropis. Tanaman durian ini sangat berpotensi untuk dikembangkan yang didukung oleh beberapa faktor seperti minat yang tinggi terhadap buah durian, Tanjung Batu merupakan produksi dari tanaman durian yang di ambil bibitnya dari negara Malaysia langsung dan ditanam di pulau Tanjung Batu, orang Batam sangat mengenal durian yang berasal dari Tanjung Batu dikarenakan rasa dan buahnya hampir sama persis seperti buah yang ditanam di negara Malaysia jenis - jenisnya sama seperti yang ada di negara Malaysia itu sendiri hanya berbeda tempat ditanamnya saja banyak orang batam yang membeli bibit durian dari Malaysia untuk ditanam dikarenakan ada yang hobby ada juga yang ingin memulai bisnis buah durian, dengan objek penelitian durian peneliti melakukan observasi terhadap penjualan durian dengan menggunakan metode algoritma apriori dibantu dengan software Rapidminer dan peneliti akan mendapatkan hasil dari perhitungan penjualan yang paling banyak diminati dan dicari oleh konsumen yaitu didapati jenis durian Musangking yang paling banyak diminati oleh konsumen karena durian Musangking memiliki rasa yang berbeda diantara durian lainnya dan juga musangking memiliki warna daging yang cantik oleh karena itu durian musangking memiliki banyak peminat, dan juga durian musangking mendapatkan nilai yang paling tinggi dalam perhitungan di aplikasi tersebut yaitu mendapatkan 100% nilai support tertinggi.

*Keywords:* Asosiasi, Data Mining, Algoritma Apriori, Durian. Google Assistant.

## PENDAHULUAN

Proses penambangani data untuk menemukan atau mencari koneksi tersembunyi dan memprediksi tren masa depan memiliki sejarah panjang. Kadang-kadang disebut sebagai penemuan pengetahuan dalam *database*, istilah penambangan data tidak diciptakan sampai tahun 1990-ani. Tetapi fondasinya terdiri dari tiga disiplin ilmu yang saling terkait: statistik studi numerik tentang hubungan dalam data, kecerdasan buatan kecerdasan mirip manusia yang ditampilkan oleh perangkat lunak dan atau mesin, dan pembelajaran

mesin yang dapat belajar dari data untuk membuat prediksi algoritma. Lama menjadi baru karena teknik penambangan data terus berkembang untuk mengikuti potensi data besar yang tidak terbatas dan daya komputasi yang terjangkau.

Durian sudah ada dari abad ke 7 masehi. Di Indonesia, durian adalah buah-buahan yang tumbuh di hutan liar. Durian memiliki rasa yang berbeda diantara buah-buahan lainnya, maka dari itu banyak orang yang menyukai durian dan mengambil bibit buah tersebut untuk

ditanam kebun masing-masing. Durian sendiri bisa digolongkan buah-buah yang paling laku diantara buah lainnya dan memiliki lebih banyak peminatnya. Durian semakin lama semakin banyak peminatnya dan banyak juga orang yang mau membuka bisnis buah durian itu sendiri dikarenakan memiliki prospek yang bagus untuk kota besar contohnya kota Batam.

Solusi yang dapat diterapkan didalam persaingan bisnis ini yaitu dari penelitian ini kita bisa melihat kebutuhan pelanggan sehingga pelanggan menjadi lebih loyal menggunakan sistem digital contohnya data mining, dan kita juga bisa menggunakan algoritma apriori untuk membantu kita dalam menjalankan bisnis kita agar lebih mudah untuk kita mendapatkan hasil dari data penjualan kita sendiri.

### KAJIAN TEORI

#### 2.1 Knowledge Discovery in Database (KDD)

Saat ini, konsep Data Mining semakin dikenal sebagai tools penting dalam manajemen informasi karena jumlah informasi yang semakin besar jumlahnya. Data Mining sendiri sering disebut sebagai *Knowledge Discovery in Database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola hubungan dalam set data berukuran besar. Ada beberapa macam pendekatan berbeda yang diklasifikasikan sebagai Teknik pencarian informasi / pengetahuan dalam KDD. Ada pendekatan kuantitatif, seperti pendekatan probabilistik seperti logika induktif, pencarian pola, dan analisis pohon keputusan. (Handoko & Sabda Lesmana, 2018)

#### 2.2 Data Mining

Dalam dunia ilmu komputer, data mining dikenal sebagai teknik penggalian data untuk menemukan pola tersembunyi dalam sekumpulan data untuk menghasilkan pengetahuan baru. Istilah "data mining" mengacu pada proses menambang data atau menambang informasi penting dari database yang sangat besar. (Tarigan, Hardinata, Qurniawan, Safii, & Winanjaya, 2022).

#### 2.3 Algoritma Apriori

Analisis asosiasi, juga dikenal sebagai analisis basket perdagangan, adalah teknik data mining untuk menemukan aturan suatu komposisi, dan algoritma apriori adalah algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk menentukan itemsets sering untuk aturan asosiasi Boolean (Luthfi, 2009).

Salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi tinggi (frequent pattern mining). (Kusumo, Sedyono, & Marwata, 2019)

$$\text{Support}(A) = \frac{\text{Jumlah transaksi mengandung } A}{\text{Total transaksi}} \times 100\%$$

Nilai *Support* dari 2 item diperoleh dengan menggunakan rumus:

$$\text{Support}(A, B) = P(A \cap B)$$

$$\text{Support}(A, B) = \frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{transaksi}} \times 100\%$$

Rumus 2.4 Perhitungan Support  
Sumber: (Sugiyono, 2016)

Pembentukan aturan asosiasi yaitu setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat kepercayaan

minimum. Kepercayaan aturan A U B diperoleh dengan menggunakan rumus

$$Confidence = P(B|A) = \frac{\sum \text{transaksi mengandung A dan B}}{\sum \text{transaksi mengandung A}} \times 100\%$$

berikut.:

Rumus 2.2 Pembentukan Aturan Asosiasi

Sumber: (Sugiyono, 2016)

Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan Support  $\times$  Confidence. Aturan diambil sebanyak vaturan yang memiliki hasil terbesar.

#### 2.4 Rapidminer

Penerapan teknik data mining dengan algoritma *K-Means* pada penelitian ini yaitu untuk menentukan pola pelaksanaan tilang dalam proses razia yang membutuhkan *software* pendukung data mining yang bersifat *open source* yaitu *Rapidminer*.

*Rapidminer* merupakan perangkat lunak yang bersifat terbuka (*open source*). *Rapidminer* adalah sebuah solusi untuk melakukan analisis terhadap datamining, *text* mining dan analisis prediksi. *Rapidminer* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. *Rapidminer* memiliki kurang lebih 500 operator data mining termasuk operator untuk *input*, *output*, data *preprocessing* dan visualisasi (Nabila et al., 2021).

Aplikasi *Rapidminer* adalah alat pengolahan data mining yang memungkinkan pengujian penerapan

algoritma apriori. Aplikasi ini memungkinkan analisis teks, mengekstrak pola dari set data, dan menggabungkannya dengan metode statistika, kecerdasan buatan, dan *database* untuk mendapatkan informasi berkualitas tinggi dari data yang diolah. Hasil pengujian menunjukkan perbedaan pola usia gelandangan dan pengemis yang mungkin menjadi gelandangan dan pengemis dari pengujian dengan aplikasi *Rapidminer* dan perhitungan manual menggunakan Algoritma Apriori. (Agustin & Muharmi, 2018).

#### 2.5 Tanagra

Tanagra adalah program data mining yang dirancang untuk memungkinkan akses ke berbagai algoritma data mining, termasuk analisis data, pembelajaran mesin, *database*, dan pembelajaran statistik. Tujuan Titanagra adalah : (Putri, Budianita, Syafria, & Insani, 2022).

Menurut (Zaki Muhammad et al., 2018) Perusahaan dengan nama yang sama mengembangkan perangkat lunak data ilmu pengetahuan yang disebut *Rapidminer Platform* ini menawarkan lingkungan terpadu untuk analisis prediktif, penambahan teks, pembelajaran mesin, dan pembelajaran mendalam.

Beberapa karakteristik *Rapidminer* adalah sebagai berikut:

1. Ditulis dengan bahasa pemrograman Java sehingga dapat dijalankan di berbagai sistem operasi.
2. Proses penemuan pengetahuan dimodelkan sebagai operator *trees*.
3. Representasi XML internal untuk memastikan format standar pertukaran data.
4. Bahasa *scripting* memungkinkan untuk eksperimen skala besar dan otomatisasi eksperimen.
5. Konsep *multi-layer* untuk menjamin

tampilan data yang efisien dan menjamin penanganan data.

- Memiliki GUI, *command line mode* dan Java API yang dapat dipanggil dari program lain.

Beberapa Fitur dari *Rapidminer*, antara lain:

- Banyaknya algoritma *datamining*, seperti *decision tree* dan *self-organizationmap*.
- Bentuk grafis yang canggih, seperti tumpang tindih diagram histogram, *tree chart* dan 3D *Scatter plots*.
- Banyaknya variasi *plugin*, seperti *text plugin* untuk melakukan analisis teks.
- Menyediakan prosedur *datamining* dan *machine learning* termasuk ETL (*extraction, transformation, loading*), data *preprocessing*, visualisasi, *modelling* dan evaluasi.
- Proses *datamining* tersusun atas operator-operator yang *nestable*, dideskripsikan dengan XML dan dibuat dengan GUI.
- Mengintegrasikan proyek *datamining* Weka dan statistika.



**Gambar 2.2** Rapidminer

Sumber: <https://rapidminer.com/>

## 2.5 Kerangka Pemikiran

Kerangka pemikiran dalam penelitian ini adalah dasar pemikiran dalam penerapan algoritma apriori untuk menghitung minat dari konsumen Toko KimHui yang berlokasi di Batam

dengan metode algoritma apriori kita bisa mendapatkan perhitungan berapa banyak orang yang minat dengan item yang ada di Toko tersebut adapun kerangka pemikiran seperti berikut:



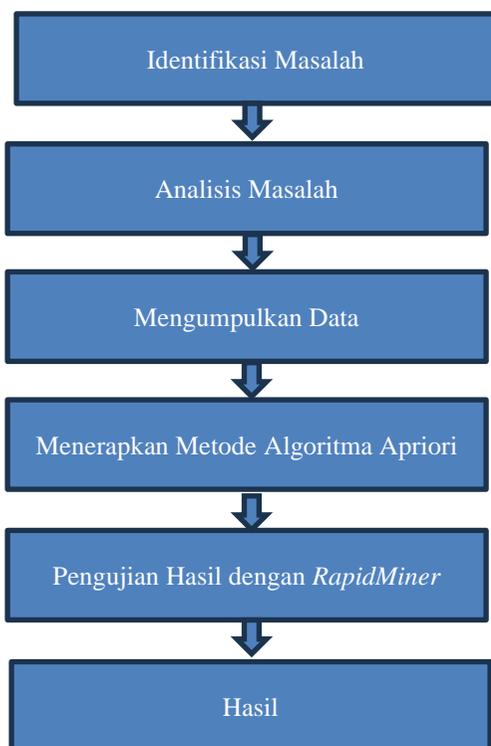
**Tabel 2. 1** Kerangka Pemikiran  
Sumber : (Data Peneliti, 2023)

Pada penjelasan diatas proses pertama yaitu mengumpulkan data penjualan barang sebagai proses input. Bentuk data merupakan produk buah durian dan jenis durian lainnya, serta jumlah total transaksi yang dilakukan ini disaring terlebih dahulu. Proses selanjutnya yaitu algoritma apriori setelah data dikumpulkan kita uji datanya menggunakan rumus tersebut dan juga kita uji dengan aplikasi Rapidminer untuk mendapatkan hasil yang akurat dari data tersebut dan yang terakhir yaitu hasil dari penjualan barang yang paling diminati oleh konsumen toko dan berguna untuk pendataan barang.

## METODE PENELITIAN

Metode penelitian yang dipakai dalam penelitian ini adalah metode eksperimen atau penelitian terapan. Yaitu menerapkan algoritma apriori ke dalam sistem serta menganalisis tren pembelian pada Toko Durian KimHui, dimana data transaksi yang diambil pada bulan Januari sampai dengan Desember 2022 dijadikan sebagai sampel data. Dari data-data transaksi yang digunakan tadi akan dilakukan pembentukan itemset dengan minimum support yang telah ditentukan setelah semua pembentukan itemset untuk menentukan pola frekuensi

transaksi yang sering terjadi selanjutnya akan dilakukan pembentukan aturan asosiasi yang memenuhi syarat minimum *confidence* yang telah dilakukan.



**Gambar 3.1** Desain Penelitian  
**Sumber :** (Data Peneliti ,2023)

Berdasarkan gambar 3.1 diatas maka tahapan yang akan dilakukan bagian ini menjelaskan metodologi yang digunakan dalam penelitian ini :

## HASIL DAN PEMBAHASAN

### 1. Identifikasi Masalah

Dalam penelitian ini langkah utamanya adalah mengidentifikasi permasalahan yang muncul pada objek penelitian dimana

penulis melakukan studi lapangan terhadap permasalahan yang muncul, pada tahap ini penulis melakukan wawancara dan observasi untuk mendapatkan informasi mengenai penjualan Toko Durian KimHui.

### 2. Analisa Masalah

Dari masalah yang teridentifikasi pada tahap analisi, dilanjutkan dengan mengidentifikasi elemen situasi masalah dan mencari tahu komponen mana yang kritis disini penulis hanya menganalisis permasalahan durian mana yang diminati oleh konsumen Toko Durian KumHui.

### 3. Mengumpulkan Data

Tahap selanjutnya adalah pengumpulan data, dimana penulis mengumpulkan informasi tentang minat konsumen Toko KimHui data yang terkumpul mulai dari tahun 2022.

### 4. Menerapkan Algoritma Apriori

Setelah tahap pengumpulan data, dianalisis menggunakan Teknik data mining yaitu algoritma apriori untuk mengetahui tingkat segmental dari minat konsumen Toko Kimhui.

### 5. Pengujian Hasil Dengan Rapidminer

Pada Langkah selanjutnya hasil analisis diuji Kembali dengan *software Rapidminer* dan hasil pengolahan manual dibandingkan dengan sistem.

### 6. Hasil

Pada Langkah terakhir, penulis merangkum hasil klasifikasi dari minat durian yang ada di Toko KimHui dan dari informasi yang akan terkumpul akan dilakukan stok yang lebih banyak terhadap jenis durian yang diminat oleh konsumen Toko KimHui.

MK	DURIHITA M	UDANGMERA H	X O	10 1	Tek Kah	D1 3	9 9	SUPE R	KAMPUN G
1	1	1	1	0	0	0	0	1	1
1	1	1	1	0	1	0	1	1	1
1	1	1	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0
1	0	0	0	0	1	0	1	1	1

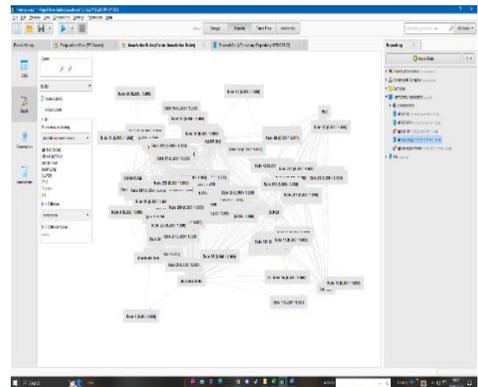
**Tabel 1.** Jumlah Produk  
**Sumber :** Data Penelitian, 2023

Setelah semua *tools* sudah lengkap dimasukkan klik ikon *Play* tunggu beberapa detik untuk mendapatkan hasil perhitungan. Setelah beberapa saat akan muncul sebuah tab yang baru yaitu *Association Rules* yaitu sebuah tabel yang berisi seluruh *itemset* yang memenuhi parameter *FP-Growth* dan *Association Rules*. Dan totalnya terdapat 470 data yang ditemukan. sebagai berikut:

**Gambar 3.2** Hasil Association Rules  
**Sumber:** (Data Penelitian, 2023)

Dibawah ini adalah hasil dalam bentuk grafik *view* perhitungan menggunakan aplikasi *RapidMiner* dan bisa kita lihat dari gambar tersebut dapat kita simpulkan bahwa durian Musangking memiliki

ketertarikan yang sangat tinggi diantara durian lainnya.



**Gambar 3.3** Hasil dalam bentuk Grafik View  
**Sumber :** (Data Penelitian, 2023)

**SIMPULAN**

Dalam penerapan langkah-langkah hasil dari algoritma apriori ini konsumen banyak yang memilih jenis durian Musangking dikarenakan jenis durian ini selalu ada setiap harinya dan memiliki rasa yang berbeda dari durian lainnya

Variabel data jenis durian di dalam peneltian ini terdapat 10 jenis durian jadi perhitungan algoritma apriori dengan *support minimum* 0.4 (40%) terdapat 4 pola kombinasi satu itemset yang

mencapai *support minimum*, dengan *support* tertinggi Musangking sebesar 100% untuk kombinasi 2 *itemset* ada 10 pola kombinasi dengan *support* tertinggi (Musangking, Udang Merah, Duri Hitam, dan Super) yaitu 80% untuk kombinasi *itemset* 3 ada 3 pola kombinasi tertinggi dengan *support* tertinggi (Musangking, Udang Merah, Duri Hitam, dan Super,) yaitu 60% dan 80%

Setelah perhitungan manual dari algoritma apriori dilanjutkan menggunakan aplikasi *Rapidminer* sehingga dari perhitungan tersebut aplikasi *Rapidminer* bisa membantu penelitian ini mendapatkan perhitungan yang paling akurat dan jenis yang paling diminati yaitu durian Musangking oleh karena itu aplikasi tersebut sangat membantu proses perhitungan yang dilakukan peneliti. yang diperoleh dari penelitian). Sedangkan, pengayaan kedua membahas kontribusi penelitian dalam *body of knowledge* topik yang dibahas di jurnal penelitian

### DAFTAR PUSTAKA

- Agustin, W., & Muharmi, Y. (2018). Apriori Algorithm through RapidMiner for Age Patterns of Homeless and Beggars. *Indonesian Journal of Artificial Intelligence and Data Mining*, 1(2), 86. <https://doi.org/10.24014/ijaidm.v1i2.5670>
- Fuady, Z., & Sasongko, P. S. (2018). Aplikasi Data Mining Menggunakan Algoritma Apriori Untuk Analisis Pola Penjualan (Studi Kasus: Apotek Keluargku Semarang). *Jurnal Masyarakat Informatika*, 9(1), 34–39. <https://doi.org/10.14710/jmasif.9.1.31479>
- Handoko, K., & Lesmana, L. S. (2018). Data Mining Pada Jumlah Penumpang Menggunakan Metode Clustering. *Snistek*, (1), 97–102.
- Handoko, K., & Sabda Lesmana, L. (2018). Computer Based Information System Journal PENGELOMPOKKAN DATA MINING PADA JUMLAH PENUMPANG DI BANDARA HANG NADIM INFORMASI ARTIKEL KATA KUNCI. *Cbis Journal*, 02. Retrieved from <http://ejournal.upbatam.ac.id/index.php/cbis>
- Hasugian, P. S., & Panjaitan, S. (2018). The Data Mining of Cell Phone Most Interested Using Apriorial Algorithm. *Infokum*, 7(1), 7–15. Retrieved from <http://infor.seaninstitute.org/index.php/infokum/index>
- Kusumo, H., Sedyono, E., & Marwata, M. (2019). Analisis Algoritma Apriori untuk Mendukung Strategi Promosi Perguruan Tinggi. *Walisongo Journal of Information Technology*, 1(1), 49. <https://doi.org/10.21580/wjit.2019.1.1.4000>
- M. Afdal, M. A., & Rosadi, M. (2019). Penerapan Association Rule Mining Untuk Analisis Penempatan Tata Letak Buku Di Perpustakaan Menggunakan Algoritma Apriori. *Jurnal Ilmiah Rekayasa Dan Manajemen Sistem Informasi*, 5(1), 99. <https://doi.org/10.24014/rmsi.v5i1.7379>
- Putri, H. N., Budianita, E., Syafria, F., & Insani, F. (2022). Penerapan Algoritma Apriori Dalam Menentukan Pola Perilaku Dan Gaya Hidup Terhadap Penderita

- Hipertensi. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 5(3), 450–458.  
<https://doi.org/10.32672/jnkti.v5i3.4402>
- Simanjuntak, P., Suharyanto, C. E., Sitohang, S., & Handoko, K. (2022). Data Mining Untuk Klasifikasi Status Pandemi Covid 19. *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, 5(2), 327.  
<https://doi.org/10.37600/tekinkom.v5i2.620>
- Sugiyono. (2016). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta.
- Syahdan, S. Al, & Sindar, A. (2018). Data Mining Penjualan Produk Dengan Metode Apriori Pada Indomaret Galang Kota. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 1(2).  
<https://doi.org/10.32672/jnkti.v1i2.771>
- Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., Safii, M., & Winanjaya, R. (2022). Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang. *Jurnal Janitra Informatika Dan Sistem Informasi*, 2(1), 9–19.  
<https://doi.org/10.25008/janitra.v2i1.142>



Biodata Penulis pertama, Vincentius Oliver Bryan, merupakan mahasiswa Prodi Teknik Informatika Universitas Putera Batam



Biodata Penulis kedua, Koko Handoko, S.Kom., M.Kom. merupakan Dosen Prodi Teknik Informatika Universitas Putera Batam.