



## Klasifikasi Data Mining Prediksi Penjualan dengan Metode Appriori (Studi Kasus: Toko *Agu Ate*)

<sup>1</sup>Naomi Dada Kodi, <sup>2</sup>Gergorius Kopong Pati, <sup>3</sup>Agustina P. Setiawi

<sup>1-3</sup>Universitas Stella Maris Sumba, Indonesia

Email [naomidadakodi17@gmail.com](mailto:naomidadakodi17@gmail.com), [gregkopong80@gmail.com](mailto:gregkopong80@gmail.com)

Alamat : Radamata, Kec. Loura, Kabupaten Sumba Barat Daya, Nusa Tenggara Tim.,  
Indonesia

Korespondensi penulis : [naomidadakodi17@gmail.com](mailto:naomidadakodi17@gmail.com)

**Abstract** Databases stored on storage media are rarely used by most of their users and even within a certain period of time the data is deleted because it is considered trash and only fills up the storage media. This assumption is not entirely true, because in fact a large database can provide the information needed for various interests, both for business interests in making decisions and for science and research. The development of information and communication technology in this era often called the millennial, information and communication technology is also increasingly advanced and developing and cannot be avoided. Where the development and progress of information and communication technology is growing very rapidly, such as the need for data processing which is increasing every day and if left alone, the data will be useless. By using the Text Mining technique, the classification method, a sentiment will be known to be positive, neutral or negative. One of the algorithms widely used in sentiment analysis is the Naïve Bayes classification method. This study uses the Naïve Bayes Classifier (NBC) method with tf-idf weighting accompanied by the addition of an emotion icon conversion feature (emoticon) to determine the existing sentiment class from tweets about *Agu Ate Store*. The results of the study show that the Naïve Bayes method without additional features is able to classify sentiment with an accuracy value of 96.44%, while if the tf-idf weighting feature is added along with the emotion icon conversion, the accuracy value can be increased to 98%.

**Keywords:** Sentiment, *Agu Ate*, Appriori Metode

**Abstrak** Database yang tersimpan di media penyimpanan jarang sekali dimanfaatkan oleh sebagian besar penggunanya dan bahkan dalam jangka waktu tertentu data-data tersebut dihapus karena dianggap sampah dan hanya memenuhi media penyimpanan saja. Anggapan tersebut tidak sepenuhnya benar, karena sesungguhnya database dalam ukuran yang besar dapat memberikan informasi yang dibutuhkan untuk berbagai kepentingan, baik untuk kepentingan bisnis dalam mengambil keputusan maupun untuk ilmu pengetahuan dan penelitian. Perkembangan teknologi informasi dan komunikasi pada era yang sering disebut milenial ini, teknologi informasi dan komunikasi juga semakin maju dan berkembang dan tidak dapat dihindari. Dimana perkembangan dan kemajuan teknologi informasi dan komunikasi berkembang dengan sangat cepat, seperti kebutuhan akan pengolahan data yang setiap hari semakin bertambah banyak dan jika dibiarkan maka data tersebut tidak akan berguna. Dengan menggunakan teknik Text Mining metode klasifikasi, akan diketahui suatu sentimen bernilai positif, netral atau negatif. Salah satu algoritme yang banyak digunakan dalam analisis sentimen adalah metode klasifikasi Naïve Bayes. Penelitian ini menggunakan metode Naïve Bayes Classifier (NBC) dengan pembobotan tf-idf disertai penambahan fitur konversi ikon emosi (emoticon) untuk mengetahui kelas sentimen yang ada dari tweet tentang *Toko Agu Ate*. Hasil penelitian menunjukkan bahwa metode Naïve Bayes tanpa penambahan fitur mampu mengklasifikasi sentimen dengan nilai akurasi sebesar 96,44%, sementara jika ditambahkan fitur pembobotan tf-idf disertai konversi ikon emosi mampu meningkatkan nilai akurasi menjadi 98%.

**Kata kunci:** Sentimen, *Toko Sahu Ate*, Metode Appriori

### 1. PENDAHULUAN

Database yang tersimpan di media penyimpanan jarang sekali dimanfaatkan oleh sebagian besar penggunanya dan bahkan dalam jangka waktu tertentu data-data tersebut dihapus karena dianggap sampah dan hanya memenuhi media penyimpanan saja. Anggapan tersebut tidak sepenuhnya benar, karena sesungguhnya database dalam ukuran yang besar

dapat memberikan informasi yang dibutuhkan untuk berbagai kepentingan, baik untuk kepentingan bisnis dalam mengambil keputusan maupun untuk ilmu pengetahuan dan penelitian. Perkembangan teknologi informasi dan komunikasi pada era yang sering disebut milenial ini, teknologi informasi dan komunikasi juga semakin maju dan berkembang dan tidak dapat dihindari. Dimana perkembangan dan kemajuan teknologi informasi dan komunikasi berkembang dengan sangat cepat, seperti kebutuhan akan pengolahan data yang setiap hari semakin bertambah banyak dan jika dibiarkan maka data tersebut tidak akan berguna.

Para pengambil keputusan berusaha untuk memanfaatkan kebutuhan terhadap keinginan konsumen (Wahyudi, 2022). Sistem penjualan ini, bukan hanya sekedar transaksi jual dan beli, didalamnya namun juga bisa terintegrasi perhitungan akuntansi, manajemen barang dan stok, manajemen penggajian karyawan, melakukan perhitungan hutang piutang, dan berbagai macam fungsi dan fitur lainnya (Handayani, 2020). Dengan adanya prediksi penjualan dalam kebutuhan perusahaan ataupun toko yang akan memasarkan dan bertransaksi dengan produk atau barang yang dijual, dirasa akan menjadi lebih efektif dan efisien, karena terbantu dengan terotomatisasinya sistem dan mengurangi kesalahan informasi seperti pencatatan penjualan dan akuntansi perusahaan atau toko (Pascalina, 2022).

Data Mining adalah proses mencari pola menarik dalam kumpulan data dengan menerapkan teknik tertentu. Proses untuk menemukan hubungan dari banyak data dalam sebuah basis data yang besar (Wanto, 2020). Proses dalam Data Mining ini terdapat berbagai jenis seperti klasifikasi, klustering, prediksi serta asosiasi.

Toko Agu Ate adalah salah satu toko yang ada di Kabupaten Sumba Barat. Toko Agu Ate menjual lemari, meja, kursi, tempat tidur dan berbagai macam furniture rumah lainnya. Dimana dalam melakukan pengolahan data transaksi penjualan dilakukan secara konvensional, Selain memenuhi keinginan konsumen, pelaku bisnis harus meningkatkan kualitas produk ataupun penambahan jenis produk yang terlaris dan paling sering dibeli oleh konsumen (Wahyudi, 2022). Dalam menghadapi persaingan pasar, untuk dapat meningkatkan penjualan produk yang dijual pada perusahaan ataupun toko, pihak terkait harus memikirkan strategi untuk pemasaran produk dan apa saja yang harus dilakukan untuk meningkatkan produktivitas penjualan (Cahyono, 2023). Salah satunya ialah memprediksi dan menganalisis data transaksi penjualan yang digunakan (Bhandari, 2023).

Dengan memprediksi dan menganalisis data transaksi penjualan, kita dapat memperoleh dan memprediksi suatu pola produk yang paling sering dibeli oleh konsumen (Lutfiani, 2022). Data yang dihasilkan akan digunakan sebagai pendukung pengambilan keputusan dalam membuat solusi bisnis, dengan menerapkan dukungan teknologi informasi seperti penggunaan

data mining dengan metode association rule menggunakan algoritma Apriori (Priambodo, 2022).

Pada penelitian ini menggunakan Algoritma Apriori yang merupakan teknik data mining untuk menemukan suatu aturan asosiasi antara suatu kombinasi item untuk menemukan pola pada data transaksi penjualan (Watini, 2023).

Dari uraian diatas maka saya tertarik untuk melakukan penelitian dengan judul “Klasifikasi Data Mining Dalam Memprediksi Penjualan Pada Toko Agu Ate dengan Metode Appriori”.

## **2. METODE**

Untuk membantu dalam pelaksanaan penelitian ini, maka perlu adanya kerangka kerja yang jelas. Kerangka kerja penelitian merupakan langkah-langkah yang menggambarkan tahapan-tahapan kegiatan yang akan dilakukan selama penelitian agar penelitian ini tercapai sesuai sasaran yang diinginkan. Adapun tahapan-tahapan kerangka kerja penelitian yang digunakan sebagai berikut :

### **1 Studi Literatur**

Pada tahap ini dilakukan pencarian landasan-landasan teori yang diperoleh dari berbagai buku dan juga sumber dari internet untuk melengkapi konsep dan teori, sehingga memiliki landasan dan keilmuan yang baik dan sesuai.

### **2 Pengumpulan data**

Pada tahap ini dilakukan pengumpulan data untuk mendapatkan data-data yang diperlukan. Oleh sebab itu penulis menggunakan metode pengumpulan data dengan cara melakukan observasi dan wawancara dengan pihak-pihak yang terkait dengan penelitian yang dilakukan.

### **3 Analisis data**

Setelah pengumpulan data selesai, dilanjutkan dengan analisa terhadap data yang dikumpulkan dari berbagai sumber ilmiah yang telah dilakukan sebelumnya. Hal ini bertujuan untuk melakukan pengelompokan terhadap data-data apa saja yang dibutuhkan dalam pengembangan sistem nantiya sehingga akan memudahkan penulis dalam melakukan analisa.

### **Metode Pengumpulan Data**

Untuk memperoleh data sebagai bahan pembuatan skripsi, maka dilakukan beberapa cara teknik pengumpulan data yaitu:

#### **1. Observasi**

Yaitu pengumpulan data dengan cara melakukan pengamatan secara langsung di Puskesmas Waimangura guna memperoleh data-data yang dapat membantu dalam proses

perancangan sistem yang meliputi prosedut pembelian obat, entri data, persediaan obat dan laporan.

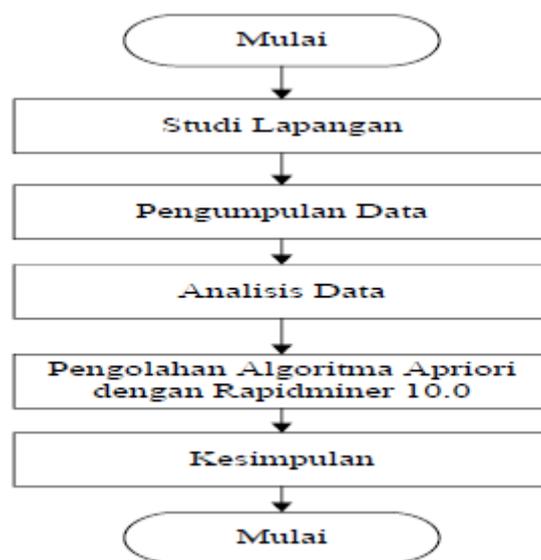
## 2. Wawancara

Metode wawancara ini dilakukan dengan Wawancara secara langsung kepada Pimpinan untuk memperoleh data-data yang dibutuhkan dalam penelitian.

## 3. Studi Pustaka

Studi pustaka dilakukan dengan membaca buku, jurnal, referensi yang menguraikan teori-teori mempunyai studi kasus yang berkaitan dengan perancangan sistem informasi persediaan obat pada Puskesmas dan hal-hal terkait lainnya.

Berikut pada gambar 31 merupakan alur dari tahapan penelitian.



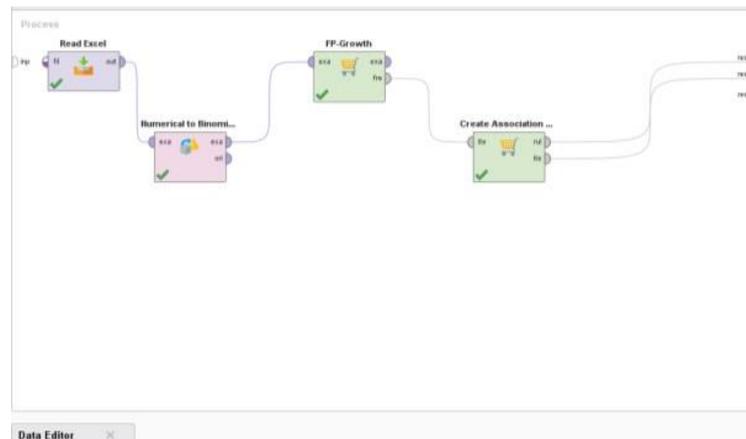
Gambar 1. Tahapan Penelitian

Implementasi Algoritma Fp-growth, dengan target mendapatkan hasil dengan RapidMiner Studio Versi 10.1. Adapun data-data yang akan diolah adalah sebagaimana tabel 1 berikut.

Table 1. Golongan Bahan Penjualan Toko

No	Bahan Penjualan
1	Bahan Dapur
2	Bahan Makan
3	Bahan Minuman
4	Bahan Bangunan
5	Bahan Kendaraan

Setelah data-data tersebut dikumpulkan, kemudian diolah dengan menggunakan RapidMiner. Langkahlangkah yang dilakukan, adalah menyusun konfigurasi operator algoritma Apriori, dengan terlebih dahulu mengatur data set, termasuk memilih type data. Konfigurasi model, tampak seperti gambar 2 berikut.



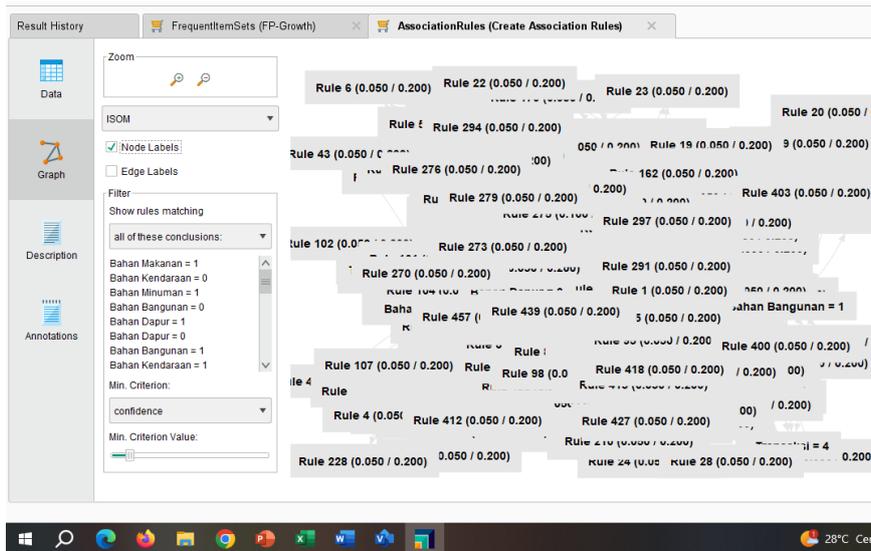
Gambar 2. Pengolahan Data Set Dengan Rapidminer Studio 10.1

Pada gambar 2 Utama dibentuk melalui langkah-langkah proses dengan menghubungkan masing-masing operator, yaitu: Read Excel menghubungkan ke Numeric ke Binominal, Numeric ke Binominal menghubungkan ke Fp-Growth, Fp-Growth menghubungkan ke Proses Hasil, Fp-Growth menghubungkan ke buat Aturan Asosiasi dan Buat Aturan Asosiasi yang menghubungkan ke Proses Hasil.pertama pada operator read excel, klik parameter import configuration wizard untuk mengimpor data transaksi yang berbentuk tabular. setelah itu pada operator FPGrowth, pada parameternya penulis harus mengisi minimal support sebesar 0.02. kemudian pada operator Create Association Rules, pada parameternya penulis akan mengisikan confidence minimal 0.5. terakhir klik run dan akan muncul hasil perhitungan. Kemudian hasil Rule tersebut dibentuk menjadi 2 rule dari hasil Rapidminer 10.1 sebagai gambar berikut.

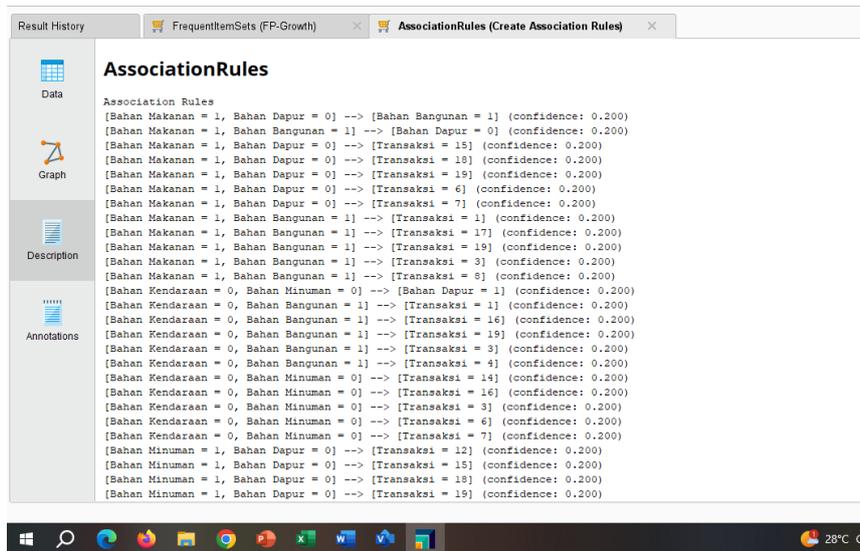
The screenshot shows the 'AssociationRules (Create Association Rules)' operator results. The table displays the association rules, their support, and confidence values. The table is as follows:

is	Conclusion	Support	Confidence
lakanan = 1	Bahan Kendaraan = 1	0.200	0.286
lakanan = 1	Bahan Minuman = 0	0.200	0.286
endaraan = 1	Bahan Minuman = 0	0.100	0.286
linuman = 0	Bahan Kendaraan = 1	0.100	0.286
lakanan = 1	Bahan Bangunan = 0, Bahan Dapur = 0	0.200	0.286
endaraan = 1	Bahan Makanan = 1, Bahan Bangunan = 0	0.100	0.286
linuman = 0	Bahan Makanan = 1, Bahan Bangunan = 0	0.100	0.286
lakanan = 1	Bahan Dapur = 1, Bahan Bangunan = 1	0.200	0.286
endaraan = 1	Bahan Makanan = 1, Bahan Dapur = 1	0.100	0.286
linuman = 0	Bahan Makanan = 1, Bahan Dapur = 1	0.100	0.286
endaraan = 1	Bahan Makanan = 1, Bahan Dapur = 0	0.100	0.286
linuman = 0	Bahan Makanan = 1, Bahan Dapur = 0	0.100	0.286
endaraan = 1	Bahan Makanan = 1, Bahan Bangunan = 1	0.100	0.286
linuman = 0	Bahan Makanan = 1, Bahan Bangunan = 1	0.100	0.286
endaraan = 0, Bahan Dapur = 1	Bahan Bangunan = 1	0.100	0.286
linuman = 0	Bahan Kendaraan = 0, Bahan Bangunan = 1	0.100	0.286

Gambar 3. Hasil Runing



Gambar 4. Grafik Rule



Gambar 5. Description Hasil

Hasil proses implementasi algoritma apriori pada perhitungan Rapidminer 10.1 untuk data transaksi penjualan barang sebanyak 20 data dengan memberikan limit minimal 0.150 Support dan 1 Confidence yang telah ditetapkan. Kemudian rekapan untuk perolehan data transaksi ditunjukkan pada tabel 2.

Tabel 2. Data Transaksi

Transaksi	Bahan Dapur	Bahan Makanan	Bahan Minuman	Bahan Bangunan	Bahan Kendaraan
1	1	1	1	1	0
2	1	0	1	0	1
3	1	1	0	1	0
4	0	0	1	1	0
5	1	1	1	0	0
6	0	1	0	0	0
7	0	1	0	0	0
8	1	1	1	1	1
9	1	1	1	0	0
10	1	0	0	0	1
11	1	1	1	0	0
12	0	0	1	1	1
13	1	1	1	0	0
14	0	0	0	0	0
15	0	1	1	0	1
16	0	0	0	1	0
17	1	1	0	1	1
18	0	1	1	0	1
19	0	1	1	1	0
20	1	1	1	0	0

Berdasarkan hasil yang telah diolah oleh Rapidminer, maka kita bisa melihat besaran pengaruh Jenis barang terhadap barang lainnya, serta bisa melihat nilai suport, baik minimal dan maksimalnya juga. Selanjutnya, untuk melihat akurasi keberhasilan perhitungan ini, dengan menggunakan description pada gambar Dan hasil penelitian ini memperlihatkan bahwa jika anda membeli barang dapur maka anda akan membeli bahan makanan.

### 3. SIMPULAN DAN SARAN

Berdasarkan penelitian yang dilakukan penulis, penulis dapat menarik kesimpulan bahwa penjualan barang pada toko yang paling banyak terjual dapat ditemukan menggunakan algoritma appriori, dengan melihat bahan yang memenuhi minimal support 2,4% dan minimal confidence 50%. Produk yang paling banyak pembeli adalah bahan dapur dan bahan makanan terbatas dengan kesimpulan bahwa bahan dapur dan bahan makanan dengan nilai support 0.350 atau 35% dan nilai kepercayaan 0,86 atau 86%.

### DAFTAR PUSTAKA

- Mukti, R. A. (2021). Sistem informasi jurnal elektronik berbasis web pada Universitas Diponegoro. *Jurnal Teknoinfo*, 15(1), 38. <https://doi.org/10.33365/jti.v15i1.473>
- Nasution, L. H., & Aliwijaya, A. (2023). Strategi preservasi digital tradisi lisan Makkobar dalam upacara perkawinan adat Mandailing (Studi kasus di Kota Padangsidempuan). *Abstrak Pendahuluan Literasi Tidak Hanya Sekedar Membaca Dan Menulis*.
- Purnamasari, A. I., & Ali, I. (2024). Analisis sentimen komentar berita Detik.com menggunakan algoritma support vector machine (SVM). *Jurnal Teknoinfo*, 8(3), 3175–3181.

- Pustaka, T. (2024). Implementasi long short-term memory dalam analisis sentimen pengguna aplikasi Twitter yang mengandung ujaran kebencian. *Jurnal Teknoinfo*, 8(3), 3170–3174.
- Trisnawati, W., & Wibowo, A. (2024). Sentiment analysis of ICT service users using Naive Bayes classifier and SVM methods with TF-IDF text weighting. *Jurnal Teknoinfo*, 5(3), 709–719.
- Wahyudin, Y., & Rahayu, D. N. (2020). Analisis metode pengembangan sistem informasi berbasis website: A literature review. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 15(3), 26–40. <https://doi.org/10.35969/interkom.v15i3.74>
- Winoto, D., Aditia, V. D., Sorisa, C., Priskila, R., & Pranatawijaya, V. H. (2024). Analisis sentimen pada ulasan pengguna terhadap aplikasi pembelajaran bahasa Duolingo: Menggunakan algoritma Naïve Bayes dan K-Nearest Neighbor. *Jurnal Teknoinfo*, 8(3), 3230–3236.