



ANALISIS PENERAPAN DATA MINING PADA PENJUALAN KERUPUK RAMBAK MENGGUNAKAN METODE NAÏVE BAYES CLASSIFER UNTUK OPTIMASI STRATEGI PEMASARAN

Sri Ayu Wulandari¹, Heri Kuswara², Nicodias Palasara³

^{1,3} STMIK Nusa Mandiri Jakarta

² Universitas Bina Sarana Informatika

Article Info:

Dikirim: 23 Oktober 2020

Direvisi: 25 November 2020

Diterima: 30 Desember 2020

Tersedia Online: 31 Desember 2020

Penulis Korespondensi:

Sri Ayu Wulandari

STMIK Nusa Mandiri Jakarta,

Bekasi, Indonesia

Email: dariiland18@gmail.com

Abstrak: Perubahan perilaku konsumen terhadap perkembangan zaman, hal ini menyebabkan para pelaku usaha kecil rumahan berinovasi supaya bisnis yang dijalankan tetap bertahan. Salah satu cara inovasi dengan melakukan analisa penjualan untuk menghasilkan suatu target yang diharapkan. Penelitian ini dilakukan pada usaha kerupuk rambak Rumahan di Bekasi. Penelitian ini dilaksanakan pada bulan maret 2020. Banyaknya usaha kerupuk rambak ini menyebabkan persaingan yang ketat. Untuk memenangkan persaingan usaha melakukan strategi pemasaran penetapan harga dan kualitas produk kerupuk rambak. Tujuan penelitian ini untuk mengetahui bagaimana pengaruh strategi pemasaran produk terhadap kerupuk rambak ini. Jenis penelitian dengan pengumpulan data penjualan berjumlah 120 data, data penjualan tersebut pada tahun 2019. Metode analisis data yang digunakan dalam penelitian ini menggunakan metode data mining yaitu Naïve Bayes. Proses pada metode ini yaitu proses analisa pola data penjualan yang telah ada sebelumnya berdasarkan atribut harga dan restock produk. Hasil dari analisa ini akan digunakan untuk kepentingan usaha rumahan kerupuk rambak dalam upaya strategi pemasaran. Pengetahuan ini juga dapat memberikan suatu informasi penting seperti hasil prediksi strategi pemasaran yang dapat digunakan untuk efektivitas dan efisiensi pemasaran dan peningkatan penjualan.

Kata kunci: kerupuk rambak; naive bayes; strategi pemasaran; optimasi; data mining.

Abstract: Changes in consumer behavior towards the times, this causes small-home business actors to innovate so that the business they run is sustainable. One way to innovate is by analyzing sales to produce an expected target. This research was conducted on a rambak rambak cracker business in Bekasi. This research was conducted in March 2020. The large number of rambak cracker businesses causes intense competition. To win the business competition, implement a marketing strategy of pricing and quality of rambak cracker products. The purpose of this study was to determine how the effect of product marketing strategies on this rambak cracker. This type of research is the collection of sales data totaling 120 data, sales data is in 2019. The data analysis method used in this study uses the data mining method, namely Naïve Bayes. The process in this method is the process of analyzing existing sales data patterns based on price attributes and product restocks. The results of this analysis will be used for the benefit of the rambak cracker home business in marketing strategy efforts. This knowledge can also provide important information such as the prediction results of marketing strategies that can be used for marketing effectiveness and efficiency and increasing sales.

Keywords: rambak crackers; naive bayes; marketing strategies; optimization; data mining.

1. PENDAHULUAN

1.1 Latar Belakang

Pada data mining penggalian informasi ini sangat berguna untuk mengingatkan keuntungan ataupun membantu dalam penyusunan strategi pemasaran. Salah satu penggunaan data mining dengan metode *Naïve Bayes Classifier* dalam data penjualan umumnya adalah untuk mengetahui strategi pemasaran pada penjualan lebih efektif dan efisien. Dengan informasi ini suatu usaha dapat mengetahui tingkat ketertarikan pembeli terhadap suatu produk. Sehingga usaha rumahan ini dapat mengetahui serta menentukan target atau sasaran pasar dengan lebih rinci.

Pada pembahasan sebelumnya yang didapat bahwa yang ingin dibuat pada penelitian ini ialah manfaat metode *Naïve Bayes Classifier* yaitu dapat diterapkan dalam memprediksi pada strategi pembelian dari analisis data penjualan di masa lampau. Data penjualan kerupuk rambak pada setiap bulan dalam tahun 2019 menjelaskan bahwa pada bulan Januari sampai Desember rumah makan sate anda bungur selalu melonjak dalam penjualan. Untuk rumah makan seperti sate, soto, bakso mie ayam lainnya setiap bulan nya pada tahun 2019 mengalami naik turun dalam penjualan mungkin karena factor persaingan pada produksi kerupuk lainnya didalam satu tempat.

Data yang telah dipaparkan sebelumnya, masih banyak permasalahannya yang terjadi di produksi kerupuk rumahan ini. Produksi kerupuk rumahan ini umumnya memproduksi kerupuk dari bahan-bahan yang mudah dijangkau. Pelanggan produksi kerupuk ini banyak peminatnya karena membantu para konsumen meningkatkan produk penjualannya juga, tetapi dalam strategi pemasaran masih kurang ini terlihat dari jumlah orderan dari perbulannya masih terbilang rendah.

Pada penelitian ini hanya fokus pada bagaimana cara menerapkan data mining menggunakan metode *Naïve Bayes Classifier* pada data penjualan dan transaksi penjualan untuk mendapatkan informasi yang baru kemudian digunakan untuk mengoptimasi strategi pemasaran produk agar bisa tercapai dengan maksimal. Hal ini dikarenakan butuhnya strategi pemasaran yang baik menghadapi pesaing bisnis lainnya dan juga bisa menambah omset penjualan. Permasalahan yang dialami sejak dulu dalam pemasaran yaitu cenderung menggunakan cara konvensional hanya menggunakan mulut ke mulut saja. Masih kurang memahami bagaimana cara mengoptimasi pemasaran yang baik, serta kurangnya pengetahuan tentang teknik pengolahan data penjualan dalam mengatasi solusi terbaik dari permasalahan tersebut.

1.2 Identifikasi Masalah

Berdasarkan pemaparan latar belakang diatas yang telah diuraikan maka identifikasi masalah yang menjadi focus masalah pada penelitian ini sebagai berikut:

- a) Kurang optimalnya produksi kerupuk rumahan ini dalam mengolah data penjualan.
- b) Belum adanya analisa data yang dapat membantu dalam mengoptimasi pemasaran pada produksi ini.
- c) Penerapan metode yang dapat menganalisa data.

1.3 Maksud dan Tujuan

- a) Mengetahui seberapa besar dampak penerapan metode *Naïve Bayes Classifier* dalam upaya optimasi strategi pemasaran.
- b) Membuat pengetahuan baru hasil analisa menggunakan metode *Naïve Bayes Classifier*.
- c) Menggunakan hasil analisa untuk optimasi strategi pemasaran.

1.4 Metode Penelitian

- a) Observasi
Observasi, suatu teknik pengumpulan data yang digunakan bila penelitian berkenaan dengan perilaku manusia, proses kerja, gejala-gejala alam dan bila responden yang diamati tidak terlalu besar. Data yang akan terkumpul dari observasi berupa informasi yang didapat melalui pengamatan peneliti ketika melihat secara langsung bagaimana proses bisnis dibidang usaha [1].
- b) Dokumentasi
Suatu cara untuk memperoleh data dan informasi dalam bentuk buku, arsip, dokumen, tulisan angka dan gambar yang berupa laporan serta keterangan yang dapat mendukung penelitian. Dokumentasi yang digunakan dalam penelitian ini meliputi profil usaha kerupuk rambak [2].
- c) Studi Pustaka
Studi pustaka yang berkaitan dengan kajian dan referensi lain yang berkaitan dengan nilai, budaya dan norma yang berkembang pada situasi sisoal yang diteliti, selain itu studi kepustakaan sangat penting dalam melakukan penelitian, hal ini dikarenakan penelitian tidak akan lepas dari literatural-literatural ilmiah. Dara diperoleh dari data yang relevan terhadap permasalahan yang akan diteliti dengan melakukan studi pustaka lainnya seperti buku, jurnal, artikel, penelitian terdahulu [3].

1.5 Ruang Lingkup

Untuk menghindari bias atau ketakaburan dalam memahami konteks penelitian dalam menguraikan hasil temuan dilapangan, analisa dan pembuatan laporan penelitian ini, peneliti memandang perlu memberikan batasan-batasan sesuai dengan focus peneliti sehingga pembahasan yang dipaparkan lebih jelas. Ruang lingkup peneliti pada skripsi ini adalah penerapan optimasi strategi pemasaran usaha kerupuk rambak rumahan dalam meningkatkan pendapatan di Bekasi, adapun alasan penelitian dilokasi ini yaitu orang tua yang mengelola usaha keupuk rambak sehingga memudahkan dalam mencari, mendapatkan data dan informasi secara langsung dari pemilik usaha terutama hal-hal yang terkait dengan focus penelitian.

1.6 Hipotesa

- H0 : Diduga usaha kerupuk rambak yang dijalankan efisien.
H1 : Usaha Kerupuk rambak yang dijalankan mempunyai resiko dalam strategi pemasaran.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dimulai dengan melakukan pengumpulan data. Data yang digunakan dalam penelitian ini adalah data yang sudah ada diperoleh dari pemilik usaha kerupuk rambak industri rumahan ini untuk sebuah penelitian. Adapun data yang digunakan adalah analisis penerapan pada penjualan kerupuk rambak yang terdiri dari 120 record data dan 4 atribut dan dibagi kedalam dua kelas yaitu "ADA" berjumlah 61 record data dan "TIDAK" berjumlah 59 record data. Pada pengolahan data awal, data dibersihkan jika ada data yang tidak lengkap atau data yang tidak konsisten maka data-data tersebut apabila ada dan ditemukan kesalahan dari data tersebut.

a) Metode Usulan

Setelah melakukan pengumpulan dan pengolahan data awal, tahap selanjutnya memilih metode yang akan digunakan pada tahap analisis. Dalam tahap ini dilakukan pembagian untuk data training dan data testing.

b) Eksperimen dan Pengujian Metode

Menjelaskan bagaimana eksperimen dilakukan dengan peneliti sampai awal data. Melakukan perhitungan dengan metode yang dipilih untuk nilai akurasi yang baik.

c) Evaluasi

Tahap ini adalah tahap akhir dari penelitian. Setelah melakukan tahap analisa maka hasilnya nilai akurasi. Kemudian dari hasil akurasi dievaluasi dan diambil kesimpulannya.

2.2 Instrument Penelitian

Pada penelitian data penjualan kerupuk rambak ini, data yang digunakan adalah data primer atau data peneliti secara langsung dari sumber asli (tidak melalui media perantara) merupakan data yang diambil dari hasil penjualan kerupuk rambak dalam 1 tahun pada tahun 2019 untuk digunakan penelitian. Sehingga dalam proses pengumpulan data, diperlukan kuesioner atau alat pengumpulan.

2.3 Metode Pengumpulan Data, Populasi, dan Sampel Penelitian

Data diperoleh dari pemilik usaha pada tahun 2019 digunakan oleh peneliti untuk data penelitian. Data yang digunakan dalam penelitian ini adalah data penjualan kerupuk rambak.

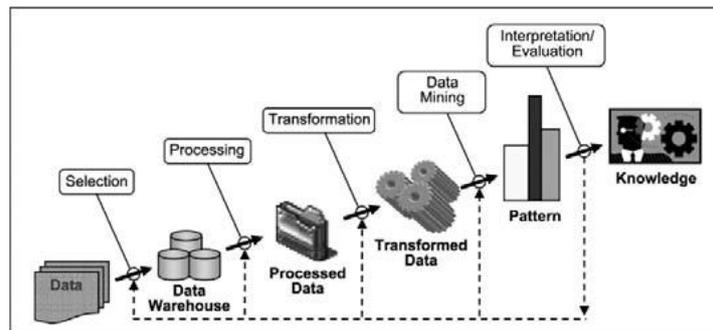
Populasi yang digunakan terdiri dari 120 record data pada data penjualan kerupuk rambak dan dibagi kedalam dua kelas yaitu kelas ADA berjumlah 61 record dan kelas TIDAK berjumlah 59 record. Data berjumlah 120 record. Data sampel yang diambil secara acak pada data penjualan kerupuk rambak berjumlah 120 data. Pada teknik pengambilan sampel secara probabilitas teknik sampling atau random/acak sampling merupakan teknik sampling yang dilakukan dengan memberikan peluang pada semua data penjualan untuk menjadi sampel.

Data ini digunakan peneliti pada tanggal 2020-02-01 yang bersumber dari pemilik usaha. Dataset penjualan terdiri dari beberapa atribut antara lain Bulan, Tempat Penjualan, Harga, Terjual, Stock.

2.4 Metode Analisa Data

Pada metode ini bahwa analisa penerapan data penjualan kerupuk menggunakan data mining *Naïve Bayes Classifier (NBC)* digunakan untuk klasifikasi probabilitas sederhana yang mengaplikasikan rumus *Teorema Bayes* dengan menggunakan asumsi ketidakpastian yang tinggi. Keuntungan penggunaan *Naïve Bayes Classifier (NBC)* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan proses pengklasifikasian. Karena digunakan sebagai variable independen, maka varian dari suatu variable dalam sebuah kelas dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari matriks kovarians.

Metode penelitian yang digunakan peneliti adalah metode kuantitatif yang dilakukan untuk menekankan analisa pada data-data numeric(angka), untuk mendapatkan gambaran berdasarkan data penjualan yang dihasilkan. Dan mengumpulkan data analisa menjadi informasi baru untuk menganalisa masalah yang diteliti. Adapun tahapan yang dilakukan dengan Knowledge Discovery Data (KDD) sebagai berikut :



Gambar 1. Tahapan KDD [1]

a) **Data Selection**

Menganalisa himpunan data, pemilihan himpunan data pada variable atau sampel adata, dimana discovery akan dilakukan. Seleksi data dari sekumpulan data dilakukan dengan tahap sebelum informasi KDD dimulai. Data pada hasil seleksi digunakan untuk proses data mining, disimpan dalam berkas.

b) **Pre-processing/Cleaning**

Proses pembersihan data merupakan operasi dasar. Sebelum proses data mining dilaksanakan perlu proses pembersihan pada data yang menjadi proses KDD. Proses pembersihan pada duplikasi data, untuk memeriksa kesalahan data seperti kesalahan cetak.

c) **Transformation**

Pencarian gambar bergantung pada goal yang ingin dicapai. Proses transformasi pada data yang akan dipilih sehingga data tersebut sesuai untuk data mining. Proses ini bergantung pada jenis atau pola informasi yang akan dicari dalam data yang digunakan.

d) **Data Mining**

Pada data mining ini untuk pemilihan goal dari proses tahapan KDD misalnya klasifikasi, regresi, clustering. Pemilihan data mining untuk suatu pencarian proses data mining yaitu proses mencari pola atau informasi yang menarik dalam data terpilih dengan menggunakan teknik metode tertentu.

e) **Evaluation**

Pola ini data mining membuat suatu proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti. Tahap ini adalah bagian dari proses tahapan KDD yang mencakup pemeriksaan apakah pola atau informasi yang didapat bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

3. **HASIL DAN PEMBAHASAN**

3.1 **Pengumpulan Dataset**

Pengumpulan dataset yang diperlukan dalam penelitian ini, Penulis memperoleh dataset yang akan digunakan dalam penelitian ini diperoleh langsung dari pemilik usaha kerupuk rambak rumahan.

Dataset penjualan terdiri dari beberapa atribut antara lain Bulan, Tempat Penjualan, Harga, Terjual, Stock. Persiapan data memenuhi semua kegiatan untuk membuat dataset penjualan yang akan digunakan ke dalam teknik pemodelan, dari data mentah merupakan dataset penjualan yang selanjutnya akan diproses dengan data mining.

Berikut adalah data keseluruhan yang akan digunakan dalam penelitian ini yang didapat dari penjualan kerupuk pada tahun 2019.

Tabel 1. Dataset [1]

No	Bulan	Tempat Penjualan	Harga	Terjual	Restock
1	Januari	Sate Anda Bungur	150000	15	Ada
2	Januari	Sate Anda Pup	75000	4	Tidak
3	Januari	Bakso Malang Arema	20000	10	Ada
4	Januari	Bakso Rusuk	20000	10	Tidak
5	Januari	Soto Kudus Kalimalang	150000	4	Tidak
6	Januari	Soto Kudus Pulogebang	150000	4	Tidak

<i>No</i>	<i>Bulan</i>	<i>Tempat Penjualan</i>	<i>Harga</i>	<i>Terjual</i>	<i>Restock</i>
7	Januari	Sate Amanda Cakung	75000	10	Ada
8	Januari	Bakso Asli Solo Wahana	20000	10	Ada
9	Januari	Soto Ayam Lamongan	20000	4	Tidak
10	Januari	Mie Ayam Bakso Rudal	20000	4	Tidak
11	Februari	Sate Anda Bungur	150000	14	Ada
12	Februari	Sate Anda Pup	75000	5	Tidak
13	Februari	Bakso Malang Arema	20000	8	Tidak
14	Februari	Bakso Rusuk	20000	8	Tidak
15	Februari	Soto Kudus Kalimalang	150000	10	Ada
16	Februari	Soto Kudus Pulogebang	150000	10	Ada
17	Februari	Sate Amanda Cakung	75000	8	Tidak
18	Februari	Bakso Asli Solo Wahana	20000	8	Tidak
19	Februari	Soto Ayam Lamongan	20000	3	Tidak
20	Februari	Mie Ayam Bakso Rudal	20000	5	Tidak
21	Maret	Sate Anda Bungur	150000	15	Ada
22	Maret	Sate Anda Pup	75000	10	Ada
23	Maret	Bakso Malang Arema	20000	10	Ada
24	Maret	Bakso Rusuk	20000	9	Tidak
25	Maret	Soto Kudus Kalimalang	150000	8	Tidak
26	Maret	Soto Kudus Pulogebang	150000	8	Tidak
27	Maret	Sate Amanda Cakung	75000	11	Ada
28	Maret	Bakso Asli Solo Wahana	20000	11	Ada
29	Maret	Soto Ayam Lamongan	20000	7	Tidak
30	Maret	Mie Ayam Bakso Rudal	20000	4	Tidak
31	April	Sate Anda Bungur	150000	9	Tidak
32	April	Sate Anda Pup	75000	10	Ada
33	April	Bakso Malang Arema	20000	9	Tidak
34	April	Bakso Rusuk	20000	10	Ada
35	April	Soto Kudus Kalimalang	150000	10	Ada
36	April	Soto Kudus Pulogebang	150000	10	Ada
37	April	Sate Amanda Cakung	75000	9	Tidak
38	April	Bakso Asli Solo Wahana	20000	8	Tidak
39	April	Soto Ayam Lamongan	20000	10	Ada
40	April	Mie Ayam Bakso Rudal	20000	11	Ada
41	Mei	Sate Anda Bungur	150000	15	Ada
42	Mei	Sate Anda Pup	75000	4	Tidak
43	Mei	Bakso Malang Arema	20000	10	Ada
44	Mei	Bakso Rusuk	20000	10	Ada
45	Mei	Soto Kudus Kalimalang	150000	4	Tidak
46	Mei	Soto Kudus Pulogebang	150000	4	Tidak
47	Mei	Sate Amanda Cakung	75000	10	Ada
48	Mei	Bakso Asli Solo Wahana	20000	10	Ada
49	Mei	Soto Ayam Lamongan	20000	4	Tidak
50	Mei	Mie Ayam Bakso Rudal	20000	4	Tidak
51	Juni	Sate Anda Bungur	150000	9	Tidak
52	Juni	Sate Anda Pup	75000	10	Ada
53	Juni	Bakso Malang Arema	20000	9	Tidak

<i>No</i>	<i>Bulan</i>	<i>Tempat Penjualan</i>	<i>Harga</i>	<i>Terjual</i>	<i>Restock</i>
54	Juni	Bakso Rusuk	20000	10	Ada
55	Juni	Soto Kudus Kalimalang	150000	10	Ada
56	Juni	Soto Kudus Pulogebang	150000	10	Ada
57	Juni	Sate Amanda Cakung	75000	9	Tidak
58	Juni	Bakso Asli Solo Wahana	20000	8	Tidak
59	Juni	Soto Ayam Lamongan	20000	10	Ada
60	Juni	Mie Ayam Bakso Rudal	20000	11	Ada
61	Juli	Sate Anda Bungur	150000	14	Ada
62	Juli	Sate Anda Pup	75000	5	Tidak
63	Juli	Bakso Malang Arema	20000	8	Tidak
64	Juli	Bakso Rusuk	20000	8	Tidak
65	Juli	Soto Kudus Kalimalang	150000	10	Ada
66	Juli	Soto Kudus Pulogebang	150000	10	Ada
67	Juli	Sate Amanda Cakung	75000	8	Tidak
68	Juli	Bakso Asli Solo Wahana	20000	8	Tidak
69	Juli	Soto Ayam Lamongan	20000	3	Tidak
70	Juli	Mie Ayam Bakso Rudal	20000	5	Tidak
71	Agustus	Sate Anda Bungur	150000	9	Tidak
72	Agustus	Sate Anda Pup	75000	10	Ada
73	Agustus	Bakso Malang Arema	20000	10	Ada
74	Agustus	Bakso Rusuk	20000	10	Ada
75	Agustus	Soto Kudus Kalimalang	150000	9	Tidak
76	Agustus	Soto Kudus Pulogebang	150000	9	Tidak
77	Agustus	Sate Amanda Cakung	75000	10	Ada
78	Agustus	Bakso Asli Solo Wahana	20000	11	Ada
79	Agustus	Soto Ayam Lamongan	20000	8	Tidak
80	Agustus	Mie Ayam Bakso Rudal	20000	10	Ada
81	September	Sate Anda Bungur	150000	15	Ada
82	September	Sate Anda Pup	75000	10	Ada
83	September	Bakso Malang Arema	20000	9	Tidak
84	September	Bakso Rusuk	20000	9	Tidak
85	September	Soto Kudus Kalimalang	150000	8	Tidak
86	September	Soto Kudus Pulogebang	150000	9	Tidak
87	September	Sate Amanda Cakung	75000	11	Ada
88	September	Bakso Asli Solo Wahana	20000	11	Ada
89	September	Soto Ayam Lamongan	20000	7	Tidak
90	September	Mie Ayam Bakso Rudal	20000	4	Tidak
91	Oktober	Sate Anda Bungur	150000	15	Ada
92	Oktober	Sate Anda Pup	75000	10	Ada
93	Oktober	Bakso Malang Arema	20000	10	Ada
94	Oktober	Bakso Rusuk	20000	9	Tidak
95	Oktober	Soto Kudus Kalimalang	150000	8	Tidak
96	Oktober	Soto Kudus Pulogebang	150000	8	Tidak
97	Oktober	Sate Amanda Cakung	75000	11	Ada
98	Oktober	Bakso Asli Solo Wahana	20000	11	Ada
99	Oktober	Soto Ayam Lamongan	20000	7	Tidak
100	Oktober	Mie Ayam Bakso Rudal	20000	4	Tidak

<i>No</i>	<i>Bulan</i>	<i>Tempat Penjualan</i>	<i>Harga</i>	<i>Terjual</i>	<i>Restock</i>
101	November	Sate Anda Bungur	150000	9	Tidak
102	November	Sate Anda Pup	75000	10	Ada
103	November	Bakso Malang Arema	20000	9	Tidak
104	November	Bakso Rusuk	20000	10	Ada
105	November	Soto Kudus Kalimalang	150000	10	Ada
106	November	Soto Kudus Pulogebang	150000	10	Ada
107	November	Sate Amanda Cakung	75000	9	Tidak
108	November	Bakso Asli Solo Wahana	20000	8	Tidak
109	November	Soto Ayam Lamongan	20000	10	Ada
110	November	Mie Ayam Bakso Rudal	20000	11	Ada
111	Desember	Sate Anda Bungur	150000	15	Ada
112	Desember	Sate Anda Pup	75000	10	Ada
113	Desember	Bakso Malang Arema	20000	10	Ada
114	Desember	Bakso Rusuk	20000	10	Ada
115	Desember	Soto Kudus Kalimalang	150000	11	Ada
116	Desember	Soto Kudus Pulogebang	150000	4	Tidak
117	Desember	Sate Amanda Cakung	75000	10	Ada
118	Desember	Bakso Asli Solo Wahana	20000	10	Ada
119	Desember	Soto Ayam Lamongan	20000	4	Tidak
120	Desember	Mie Ayam Bakso Rudal	20000	4	Tidak

3.2 Pengolahan Data Mentah(Pre-Processing Data)

Dataset diatas berjumlah 120 record yang terdiri dari 61 record ADA dan 59 record TIDAK dikolom *classification*. Data diatas tersebut perlu dicek terlebih dahulu apakah ada data yang kosong atau tidak sehingga data diatas tidak memiliki noise atau kesalahan saat dianalisa.

Berdasarkan pengecekan yang dilakukan oleh penulis secara manual dan tidak terjadi noise pada data diatas. Maka data tersebut bisa diolah ke proses selanjutnya. Untuk memeriksa data noise dapat menggunakan aplikasi pendukung jika data akan diolah berjumlah besar. Bisa menggunakan *rapidminer* untuk waktu yang efisien.

Pada tahap ini merupakan tahap untuk memastikan data penjualan yang dipilih telah layak untuk dilakukan proses pengolahan. Data yang digunakan berjenis numerik seperti harga, dan terjual harus dilakukan dengan proses inisialisasi data dahulu ke dalam bentuk nominal untuk melakukan inisialisasi restock dapat dilakukan.

- a) Barang yang terjual lebih dari 10 dalam sebulan diberi inisial pada atribut RESTOCK “ADA”.
- b) Barang yang terjual kurang atau sama dengan dari 10 dalam sebulan diberi inisial pada atribut RESTOCK “TIDAK”.

Tabel 2. Kelas restock [2]

<i>Terjual</i>	<i>Restock</i>
15	Ada
4	Tidak
10	Ada
14	Ada
5	Tidak
8	Tidak
9	Tidak
11	Ada
7	Tidak
3	Tidak

3.3 Menentukan Metode Yang Digunakan

Setelah semua data yang akan digunakan dalam penelitian sudah siap, maka tahap selanjutnya adalah menentukan metode yang akan digunakan untuk analisa dengan perhitungan *Naïve Bayes*. *Naïve Bayes* sendiri merupakan salah satu metode data mining yang sering digunakan untuk klasifikasi. Metode *Naïve Bayes* ini adalah salah satu metode yang paling sederhana dibandingkan dengan metode *data mining* lainnya. Tingkat akurasi perhitungan yang cukup tinggi. Pemodelan merupakan suatu teknik yang melibatkan data mining yaitu dengan melakukan pemilihan teknik pada data mining dan menentukan algoritma yang akan digunakan. Tool yang digunakan adalah Rapid Miner Versi 9.6.001

Dalam tahap ini juga penulis membagi data yang ada sebanyak 120 record kedalam dua jenis data yaitu data training dan data testing. Data training sendiri berisi 120 record yang akan digunakan untuk menghitung probabilitas atribut tiap kelas dan untuk data testing berisi 120 record digunakan untuk melakukan pengujian terhadap algoritma yang telah dihitung sebelumnya dan data tersebut untuk testing dipilih secara acak.

Tabel 3. Dataset penjualan [2]

No	Bulan	Tempat Penjualan	Harga	Terjual	Re-stok
1	Januari	Sate anda bungur	150.000	15	Ada
2	Januari	Sate Anda Pup	75.000	4	Tidak
3	Januari	Bakso Malang Arema	20.000	10	Ada
4	Januari	Bakso Rusuk	20.000	10	Tidak
5	Januari	Soto Kudus Kalimalang	150.000	4	Tidak
6	Januari	Soto Kudus Pulogebang	150.000	4	Tidak
7	Januari	Sate Amanda Cakung	75.000	10	Ada
8	Januari	Bakso Asli Solo Wahana	20.000	10	Ada
9	Januari	Soto Ayam Lamongan	20.000	4	Tidak
10	Januari	Mie Ayam Bakso Rudal	20.000	4	Tidak
11	Februari	Sate Anda Bungur	150.000	14	Ada
12	Februari	Sate Anda Pup	75.000	5	Tidak
13	Februari	Bakso Malang Arema	20.000	8	Tidak
14	Februari	Bakso Rusuk	20.000	8	Tidak
15	Februari	Soto Kudus Kalimalang	150.000	10	Ada
16	Februari	Soto Kudus Pulogebang	150.000	10	Ada
17	Februari	Sate Amanda Cakung	75.000	8	Tidak
18	Februari	Bakso Asli Solo Wahana	20.000	8	Tidak
19	Februari	Soto Ayam Lamongan	20.000	3	Tidak
20	Februari	Mie Ayam Bakso Rudal	20.000	5	Tidak
21	Maret	Sate Anda Bungur	150.000	15	Ada
22...120	Maret	Sate Anda Pup	75.000	10	Ada

3.4 Eksperimen dan Pengujian Metode

Pada tahap ini yaitu perhitungan pencarian kelayakan dengan metode Data Mining *Naïve Bayes* untuk mencari nilai probabilitas dari masing-masing kelas. Untuk penerapan penjualan kerupuk rambak sebagai optimasi strategi pemasaran akan ditentukan 2 kelas yaitu kelas “ADA” dan “TIDAK”. Data mining merupakan penguraian dari sekumpulan data menjadi informasi yang memiliki potensi secara implisit yang sebelumnya tidak diketahui [4]. Dalam Teorema Bayes probabilitas atau peluang bersyarat dinyatakan sebagai berikut:

$$P(H|X) = \frac{P(X|H) P(H)}{P(X)} \quad (1)$$

Keterangan :

- X : Data dengan class yang belum diketahui
- H : Hipotesis data X merupakan suatu class spesifik
- P(H|X) : Probabilitas hipotesis H berdasar kondisi X
- P(H) : Probabilitas hipotesis H (prior probability)
- P(X|H) : probabilitas berdasarkan kondisi pada hipotesis
- P(X) : Probabilitas dari X

Cara penghitungannya adalah dengan mencari berapa jumlah data yang ADA dan TIDAK dari total keseluruhan data training, lalu membaginya dengan total keseluruhan data.

Berikut adalah perhitungan probabilitas berdasarkan data yang akan digunakan dalam penelitian ini.

- a) Menghitung probabilitas kelas
 Probabilitas ADA = $C1(\text{Class Restock} = \text{"ADA"}) = \text{jumlah "ADA" pada kolom Restock} = 61/120 = 0,508$
 Probabilitas TIDAK = $C2(\text{Class Restock} = \text{"TIDAK"}) = \text{jumlah "TIDAK" pada kolom Restock} = 59/120 = 0,491$
- b) Menghitung probabilitas Tempat Penjualan (Mie Ayam Bakso Rudal)
 $P(\text{Tempat Penjualan "Mie Ayam Bakso Rudal"} | \text{class class Restock} = \text{"ADA"}) = 4/12 = 0,33333333333333$
 $P(\text{Tempat Penjualan "Mie Ayam Bakso Rudal"} | \text{class class Restock} = \text{"TIDAK"}) = 8/12 = 0,66666666666666$
 Berikut adalah daftar lengkap hasil perhitungan probabilitas Tempat Penjualan (Mie Ayam Bakso Rudal) terhadap kelas restock.

Tabel 4. Probabilitas tempat penjualan (mie ayam bakso rudal)

	<i>Ada</i>	<i>Tidak</i>
Mie Ayam Bakso Rudal	0,33333333333333	0,66666666666666

Setelah probabilitas dari semua kriteria maka tahap selanjutnya adalah melakukan perhitungan dengan menggunakan probabilitas kelas restock ADA dengan mengkalikan seluruh probabilitas dari atribut yang ada berdasarkan kelas ADA.

$$0,33333333333333 \times 0,66666666666666 \times 0,5 \times 0,66666666666666 \times 0,58333333333333 \times 0,25 \times 0,58333333333333 \times 0,5 \times 0,41666666666666 \times 0,58333333333333 \times 0,508 = \mathbf{0,04671380244}$$

Selanjutnya penulis juga akan melakukan perhitungan dengan menggunakan probabilitas kelas restock TIDAK.

$$0,66666666666666 \times 0,33333333333333 \times 0,5 \times 0,33333333333333 \times 0,41666666666666 \times 0,75 \times 0,41666666666666 \times 0,5 \times 0,58333333333333 \times 0,41666666666666 \times 0,491 = \mathbf{0,0345782160869}$$

Tabel 5. Perkalian Probabilitas Akhir Setiap Kelas [5]

<i>Tempat Penjualan</i>	<i>Re-Stok</i>	
	<i>Ada</i>	<i>Tidak</i>
Mie Ayam Bakso Rudal	0,33333333333333	0,66666666666666
Sate Anda Bungur	0,66666666666666	0,33333333333333
Bakso Malang Arema	0,5	0,5
Sate Anda Pup	0,66666666666666	0,33333333333333
Bakso Rusuk	0,58333333333333	0,41666666666666
Soto Ayam Lamongan	0,25	0,75
Sate Amanda Cakung	0,58333333333333	0,41666666666666
Soto Kudus Kalimalang	0,5	0,5
Soto Kudus Pulogebang	0,41666666666666	0,58333333333333
Bakso Asli Solo Wahana	0,58333333333333	0,41666666666666
Jumlah	0,04671380244	0,0345782160869

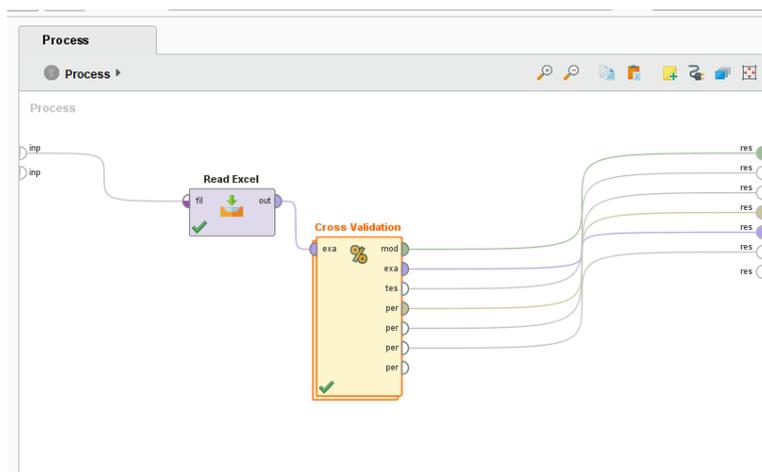
Setelah penulis mengetahui hasil perhitungan dari kedua kelas restock yang ada yaitu ADA dan TIDAK, maka tahap selanjutnya melakukan perbandingan dari kedua hasil tersebut untuk menentukan mana yang paling besar. Nilai paling besar tersebut akan menjadi yang terpilih dari kedua kelas yang ada. Berdasarkan hasil perhitungan diatas maka nilai tersebut adalah **0,04671380244** yang merupakan kelas restock dari ADA. Jika dibandingkan dengan data diatas yang record tersebut memiliki kelas ADA. Maka dapat disimpulkan bahwa proses klasifikasi record pertama ini adalah benar. Jika salah maka hasil dari perhitungan dengan data yang dimiliki kelas berbeda.

3.5 Evaluasi

Pada tahap akhir ini, akan dilakukan beberapa perhitungan untuk menentukan nilai akurasi menggunakan metode naïve bayes. Berdasarkan hasil sebelumnya diketahui bahwa dari 120 data training. Ada 61 data yang diklasifikasi dengan nama kelas restock ADA dan 59 data yang diklasifikasi dengan nama kelas restock TIDAK.

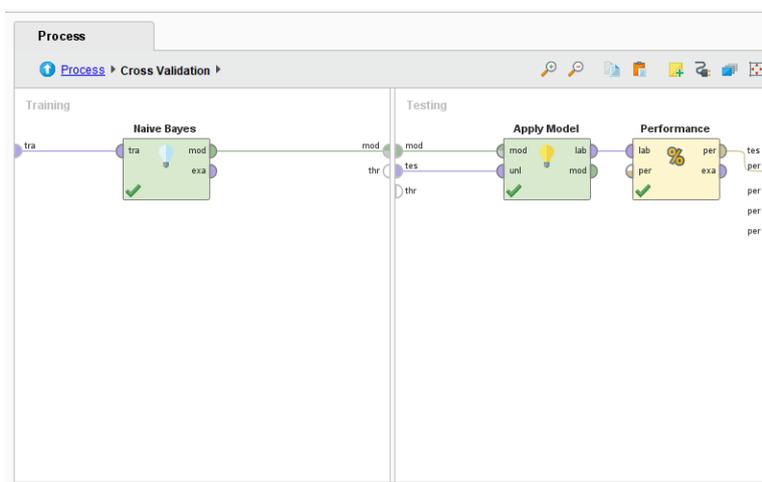
Berdasarkan perhitungan hasil akurasi pada data training adalah 99,17 yang berarti klasifikasi pada penjualan kerupuk menggunakan naïve bayes bisa dipakai karena tingkat akurasinya cukup baik.

Berikut pengolahan data dengan menggunakan Naïve Bayes pada Rapidminer:



Gambar 2. Pengolahan Data dengan Naïve Bayes Pada Rapidminer [6]

Jumlah dataset mencapai 120 record maka penulis menggunakan tools Rapidminer untuk membantu proses perhitungan. Validation membantu menghasilkan sebuah tingkat akurasi berdasarkan dataset pada penjualan yang telah dilakukan proses klasifikasi.



Gambar 3. Proses Klasifikasi [7]

Penelitian ini untuk mengetahui nilai akurasi dari *Naïve Bayes* yang digunakan untuk klasifikasi Restock. Di dalam kolom training terdapat klasifikasi yang ada dalam *Naïve Bayes*. Sedangkan pada kolom testing terdapat sebuah *Apply Model* untuk menjalankan model perhitungan *Naïve Bayes* dan sebuah *Performance* untuk mengukur performa dari model *Naïve Bayes*.

PerformanceVector (Performance) ExampleSet (Read Excel)

Table View Plot View

accuracy: 99.17% +/- 2.64% (micro average: 99.17%)

	true ADA	true TIDAK	class precision
pred. ADA	60	1	98.36%
pred. TIDAK	0	59	100.00%
class recall	100.00%	98.33%	

Gambar 4. Hasil Proses Akurasi Data [8]

Hasil akurasi pada model *Naïve Bayes* menunjukkan tingkat akurasinya 99.17%. sedangkan class precision dan class recall untuk prediksi ADA menunjukkan tingkat precision sebesar 98,36% dan untuk prediksi TIDAK sebesar 100,00%, artinya model klasifikasi kelulusan menggunakan *Naïve Bayes* terbukti baik hal ini dilihat dari tingkat akurasinya yang mencapai 99.17% akan tetapi hal ini perlu di tinjau ulang dari sudut pandang kompleksitas dan jumlah datasetnya.

SimpleDistribution

```
Distribution model for label attribute G

Class ADA (0.500)
6 distributions

Class TIDAK (0.500)
6 distributions
```

Gambar 5. Proses Simple Distribution [9]

Hasil klasifikasi dari data penjualan dengan metode *Naïve Bayes* dibagi 2 kelas klasifikasi yaitu class ADA dan TIDAK. Model distribusi untuk label atribut kelas Restock adalah sebagai berikut:

- a) Class ADA : 6 distributions
- b) Class TIDAK : 6 distributions

Percobaan pada penelitian ini menggunakan Rapidminer 9.6.001. Algoritma yang digunakan adalah *Naïve Bayes*. lalu divalidasi untuk ditesting dengan menggunakan *Apply Model* untuk menjalankan sebuah perhitungan algoritma atau model *Naïve Bayes* serta *Performance* untuk mengukur sebuah performa dari model *Naïve Bayes* tersebut.

3.6 Hasil Analisa

Tahapan akhir pada pembuatan laporan hasil kegiatan data mining. Laporan akhir berisi mengenai pengetahuan yang didapat pada data dalam proses data mining. Berdasarkan penelitian yang dilakukan, telah dihasilkan suatu pola, informasi, dan pengetahuan baru dalam proses data mining untuk klasifikasi restock barang berdasarkan data penjualan kerupuk rambak tahun 2019. Dari penelitian dihasilkan suatu informasi dan pengetahuan baru yang sesuai dengan tujuan data mining yaitu pola perhitungan data mining yang berisi data training dan data testing serta mencari nilai probabilitas dari setiap atribut berdasarkan data training dan data testing untuk menghasilkan suatu informasi baru, apakah pada data penjualan kerupuk rambak tahun 2019 tempat penjualan mana saja yang harus di restock/diperbanyak untuk kedepannya . Kemudian untuk menguji tingkat keakurasiannya maka digunakan *Rapidminer* sebagai alat bantu dalam proses pengujian tingkat akurasi dari klasifikasi tersebut. Dari proses perhitungan data mining menggunakan algoritma *Naïve Bayes* dan tingkat keakurasiannya, dihasilkan suatu informasi baru yaitu perhitungan data mining berdasarkan data penjualan produk tahun 2019, untuk menunjukkan kelas restock “ADA” dengan total semua perkalian seluruh probabilitas senilai **0,04671380244**, sedangkan kelas restock “TIDAK” dengan total semua perkalian seluruh probabilitas senilai **0,0345782160869**. Untuk menentukan tingkat akurasi berdasarkan proses klasifikasi menggunakan perhitungan algoritma *Naïve Bayes*, dengan melalui semua tahapan dipastikan tidak ada bagian penting yang terlewatkan, dihasilkan tingkat akurasi sebesar 99.17%. Berdasarkan hasil perhitungan data mining dan proses pengujian tingkat akurasi dengan menggunakan *Rapidminer*, dapat ditarik kesimpulan bahwa tahun 2019 kelas restock “ADA” lebih besar dari kelas “TIDAK”. Sedangkan analisa yang dilakukan terhadap tingkat akurasi menggunakan algoritma *Naïve Bayes* menunjukkan bahwa nilai yang dihasilkan oleh *Naïve Bayes* memiliki tingkat kekuatan yang cukup tinggi. Penerapan dengan metode data mining pada data set penjualan kerupuk rambak mencapai tingkat optimasi sebesar lebih dari 80%.

4. KESIMPULAN

Metode data mining sangat penting untuk menganalisa penerapan data penjualan pada kerupuk rambak dengan metode *naïve bayes*. Pada metode *naïve bayes* ini sangat tepat untuk menganalisa data mining penjualan yang salah satunya untuk mengklasifikasikan atau memprediksi kejadian sebelumnya dengan menghitung probabilitas dari masing-masing kejadian. Rapid miner salah satu software yang membantu untuk membuktikan bahwa hasil perhitungan dengan metode *naïve bayes* nilai akurasi nya baik.

Hasil analisa dari metode ini yang telah penulis lakukan perhitungan, penulis menyimpulkan bahwa metode *Naïve Bayes* bisa digunakan untuk mengklasifikasikan dalam mengoptimasikan strategi pemasaran. Pengujian menggunakan 120 data dengan perhitungan rapid miner dihasilkan untuk nilai accuracy 99,17%. Hasil dari data penjualan dengan metode *Naïve Bayes* membagi 2 kelas yaitu class ADA dengan nilai 0,508 dan class TIDAK dengan nilai 0,491. Analisa data penjualan untuk melihat optimasi strategi pemasaran dengan nilai tertinggi 04671380244 yang merupakan kelas restock dari ADA. Hasil akurasi tersebut membuktikan bahwa hasil

perhitungan ini bisa dijadikan landasan untuk mengklasifikasikan penjualan kerupuk rambak karena hasil dari akurasi untuk data training sudah diatas 80% atau hasil akurasinya tinggi meskipun untuk hasil akurasi data testing tidak terlalu tinggi tapi hasil akurasinya masih bisa dikatakan baik.

UCAPAN TERIMA KASIH

Penulis mengucapkan banyak terima kasih kepada pemilik usaha kerupuk rumahan yang telah memberi dukungan secara financial terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] Sugiyono, "Repository Unpas," *Stud. kepustakaan*, vol. 53, no. 9, pp. 1689–1699, 2015, doi: 10.1017/CBO9781107415324.004.
- [2] universitas negeri Sugiyono, "metode penelitian bab III," *Biomass Chem Eng*, vol. 49, no. 23–6, pp. 40–68, 2015, [Online]. Available: https://eprints.uny.ac.id/53740/4/TAS_BAB_III_13416241020.pdf.
- [3] Paul M. Muchinsky, "pengaruh peserta didik baru," *Psychol. Appl. to Work An Introd. to Ind. Organ. Psychol. Tenth Ed. Paul*, vol. 53, no. 9, pp. 1689–1699, 2012, doi: 10.1017/CBO9781107415324.004.
- [4] "Skripsi_revisi taufik."