

---

**PENERAPAN DATA MINING PREDIKSI PENJUALAN BARANG  
ELEKTRONIK TERLARIS MENGGUNAKAN ALGORITMA *NAÏVE BAYES*  
(Study Kasus : Planet Cash And Credit Cabang Muara Enim)**

**Novi Pransiska, A.Haidar Mirza, Andri.**

Program Studi Teknik Informatika,  
Fakultas Ilmu Komputer, Universitas Binadarma Palembang  
Jalan Jenderal Ahmad Yani No. 12 Palembang

<sup>1</sup>[novi.pransiska05@yahoo.com](mailto:novi.pransiska05@yahoo.com), <sup>2</sup>[informatika.ubd2016@binadarma.ac.id](mailto:informatika.ubd2016@binadarma.ac.id), <sup>3</sup>[andri@binadarma.ac.id](mailto:andri@binadarma.ac.id)

**ABSTRAK**

Dalam menghadapi persaingan pasar untuk menghasilkan peningkatan pendapatan toko, pihak terkait harus menentukan strategi pemasaran produk yang dijual. Para pelaku bisnis juga harus membutuhkan sebuah analisis untuk memprediksi barang yang paling banyak terjual. Analisis yang dibutuhkan harus bias membantu dan memudahkan para karyawan toko untuk mengetahui barang yang mudah habis terjual dengan memprediksi data penjualan pada tahun-tahun sebelumnya dengan menerapkan data mining menggunakan algoritma *naive bayes* menggunakan data penjualan tahun 2014 sampai dengan tahun 2016 di PT Solusi Prima Artha *planet cash and credit* cabang Muara Enim.

**Kata Kunci:** Prediksi, Data Mining, Algoritma *Naive Bayes*

**ABSTRACT**

*In the face of market competition to produce an increase in store income, related parties must determine the marketing strategy of the products sold. Business people also need to need an analysis to predict the most sold items. The analysis needed must be able to help and facilitate the store employees to find out items that are easily sold out by predicting sales data in previous years by applying data mining using the Naive Bayes algorithm using sales data from 2014 to 2016 at PT Solusi Prima Artha planet cash and credit of Muara Enim branch.*

**Keywords:** Prediction, Data Mining, algorithm *Naive Bayes*

**1. PENDAHULUAN**

Seiring dengan perkembangan zaman persaingan di dunia bisnis dan teknologi semakin berkembang pesat. Sehingga membuat para pelaku bisnis berlomba-lomba untuk menembangkan bisnis mereka dengan memanfaatkan teknologi yang ada dan juga agar bisnis mereka selalu bertahan dan berkembang dalam pesatnya persaingan. Untuk memenuhi keinginan konsumen, para pelaku bisnis harus meningkatkan kualitas produk, penambahan jenis produk yang disukai dan paling sering dibeli konsumen.

Dalam menghadapi persaingan pasar untuk menghasilkan peningkatan pendapatan toko, pihak terkait harus menentukan strategi pemasaran produk yang dijual. Para pelaku bisnis juga harus membutuhkan sebuah analisis untuk memprediksi barang yang paling banyak terjual. Analisis yang dibutuhkan harus bisa membantu dan memudahkan para karyawan toko untuk mengetahui barang yang mudah habis terjual dengan memprediksi data penjualan pada tahun-tahun sebelumnya dengan menerapkan data mining menggunakan algoritma *naive bayes* dan memprediksi harga saham dengan metode SVM (Ferdiansyah, F., Negara, E. S., & Widyanti, Y. 2019, Sutabri, et-al 2018.).

Data mining adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Negara, E. S. 2018, Larose, 2006).

Sedangkan algoritma *naive bayes* adalah salah satu penerapan *theorem bayes* dalam klasifikasi *naive bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output (Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. 2018, Santosa : 2007 : 79).

Dari uraian diatas maka penulis akan melakukan penelitian untuk memprediksi barang terlaris dimasa yang akan datang menggunakan data penjualan tahun 2014 sampai dengan tahun 2016 dengan menerapkan

---

data mining dan algoritma *naive bayes* di PT Solusi Prima Artha *planet cash and credit* cabang Muara Enim.

## 2. TINJAUAN PUSTAKA

### 2.1. Tinjauan Umum

#### 2.1.1. Profil Perusahaan

*Planet cash and credit* diresmikan berdiri pada tanggal 5 Mei 2008. *Planet cash and credit* adalah perusahaan yang bergerak dalam bidang penjualan barang elektronik dan *furniture* baik kredit maupun *cash*.

*Planet cash and credit* telah memiliki banyak cabang di kota-kota besar. *Planet cash and credit* adalah salah satu perusahaan yang maju pesat. Pada perkembangannya *planet cash and credit* telah menjangkau seluruh lapisan masyarakat dengan berbagai kebutuhan yang berkualitas dan berteknologi maksimal, baik dipertanyaan sampai merambah ke daerah-daerah.

*Planet* juga sangat “*concern*” untuk memperkuat jaringan dan distribusi ke semua daerah dan semua aspek, yang mengedepankan kerja sama positif, baik kepada vendor / distributor, konsumen, maupun kepada sesama perusahaan sejenis, sebagai mitra utama perusahaan gratis biaya administrasi, gratis pengiriman, produk bergaransi, kualitas terjamin.

#### 2.1.2. Visi dan Misi *Planet Cash and Credit*

Untuk mendukung kelangsungan pengelolaan bisnis perusahaan, *Planet Cash and Credit* telah mengembangkan visi dan misi perusahaan yang dapat digunakan sebagai acuan oleh perusahaan dalam menjalankan setiap kegiatan yang dilakukan di dalam perusahaan.

##### 1. Visi *Planet Cash and Credit*

*Kami adalah perusahaan pembawa perubahan yang menciptakan kemanfaatan bagi masyarakat berdasarkan prinsip saling menumbuhkan.*

##### 2. Misi *Planet Cash and Credit*

*Perusahaan cash dan kredit terdepan di In*

#### 1.2.1. Penerapan

Pengertian Penerapan Menurut J.S Badudu dan Sutan Mohammad Zain, penerapan adalah hal, cara atau hasil (Badudu & Zain, 1996:1487).

Adapun menurut Lukman Ali, penerapan adalah mempraktekkan, memasang (Ali, 1995:1044). Berdasarkan pengertian tersebut dapat disimpulkan bahwa penerapan merupakan sebuah tindakan yang dilakukan baik secara individu maupun kelompok dengan maksud untuk mencapai tujuan yang telah dirumuskan. Adapun unsur-unsur penerapan meliputi :

1. Adanya program yang dilaksanakan
2. Adanya kelompok target, yaitu masyarakat yang menjadi sasaran dan diharapkan akan menerima manfaat dari program tersebut.
3. Adanya pelaksanaan, baik organisasi atau perorangan yang bertanggung jawab dalam pengelolaan, pelaksanaan maupun pengawasan dari proses penerapan tersebut (Wahab, 1990:45).

#### 2.2.2 Data Mining

*Data mining* merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar. (Turban et al, 2005).

Menurut *Gartner Group* data mining adalah suatu proses ekstraksi data untuk menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Antoni, D., Negara, E. S., & Suweno, S. 2015, Larose, 2006).

#### 2.2.3. Algoritma *Naive Bayes*

*Naive Bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu. Keuntungan penggunaan *Naive Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (*Training Data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Dalam metode *naive bayes* data *string* yang bersifat konstan dibedakan dengan data numerik yang bersifat kontinyu, perbedaan ini akan terlihat pada saat menentukan.

*Naive Bayes* adalah salah satu penerapan theorem Bayes dalam klasifikasi, *Naive Bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output". ((Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. 2018, Santosa : 2007 : 79).

Menurut Olson dan Delen (2008:102) menjelaskan *Naive bayes* untuk setiap kelas keputusan, menghitung probabilitas dengan syarat bahwa kelas keputusan adalah benar, mengingat vektor informasi obyek. Keuntungan *Naive Bayesian* :

1. Menangani kuantitatif dan data diskrit
2. Kokoh untuk titik noise yang diisolasi, misalkan titik yang dirata-ratakan ketika mengestimasi peluang bersyarat data.
3. Hanya memerlukan sejumlah kecil data pelatihan untuk mengestimasi parameter (rata-rata dan variansi dari variabel) yang dibutuhkan untuk klasifikasi.
4. Menangani nilai yang hilang dengan mengabaikan instansi selama perhitungan estimasi peluang
5. Cepat dan efisiensi ruang
6. Kokoh terhadap atribut yang tidak relevan

### 3. ANALISIS DATA MINING

*Data Mining* merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. Analisa *data mining* mengikuti tahapan dalam *Knowledge Discovery in Database* (KDD) yaitu *data selection*, *data cleaning*, *data integration* dan *transformation*. Setelah Proses di atas selesai maka barulah *data source* (data mentah) biar diolah ke dalam *data mining*.

#### 3.1. Sumber Data

Data yang digunakan dalam penelitian ini bersumber dari data penjualan Planet *Cash and Credit* Muara Enim pada tahun 2014 sampai dengan tahun 2016 dengan kategori Kipas Angin, TV, Air Conditioner, Mesin Cuci, Setrika, Blender, DVD Player, Rice Cooker. Data penjualan ini terdiri dari tabel penjualan dari data barang berhasil terjual. Data yang dicatat adalah *id\_detailtrans*, *id\_transaksi*, *tgl\_trans*, *nm\_item* dan *qty*. Proses pendataan dilakukan di data penjualan. Setelah direkapitulasi kemudian disebarakan. Berikut adalah tabel data penjualan :

id_detailtrans	id_transaksi	tgl_trans	nm_item	qty
527953	5	03/01/2016	Kipas Angin	2
527954	5	03/01/2016	Air Conditioner	1
527955	6	03/01/2016	Kipas Angin	1
527956	7	03/01/2016	Kipas Angin	2
527957	7	03/01/2016	Kipas Angin	1
527958	8	03/01/2016	Mesin Cuci	2
527959	8	03/01/2016	Kipas Angin	2
527960	9	04/01/2016	Kipas Angin	2
527961	9	04/01/2016	Kipas Angin	2
527962	10	05/01/2016	Kipas Angin	2
527963	10	05/01/2016	Setrika	1
527964	11	05/01/2016	Air Conditioner	2
527965	12	05/01/2016	Blender	2
527966	12	05/01/2016	Kipas Angin	1
527967	13	06/01/2016	Kipas Angin	1

Gambar 3.1. gambar tabel data penjualan

Pada gambar diatas adalah tabel data penjualan pada *Planet Cash and Credit* pada tahun 2014 sampai dengan tahun 2016 dimana pada data diatas terdiri dari 5 atribut yaitu atribut *id\_detailtrans*, *id\_transaksi*, *tgl\_trans*, *nm\_item* dan *qty*. Pada tabel data penjualan Planet *Cash and Credit* cabang Muara Enim.

#### 3.2. Data Selection (Pemilihan Data)

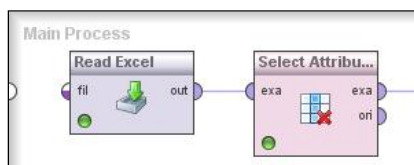
*Data selection* merupakan proses seleksi atau memilih atribut yang akan digunakan dalam proses *data mining* karena tidak semua atribut pada sumber data dapat digunakan seluruhnya. Pada *Planet Cash and Credit* cabang Muara Enim terdapat data penjualan ..

Dikarenakan data tersebut dibutuhkan untuk proses *data mining* maka data tersebut akan dilakukan pengurangan dan data tersebut akan melalui proses seleksi data. Pada proses seleksi data pada data penjualan atribut yang dipilih hanya atribut *id\_detailtrans*, *id\_transaksi*, *tgl\_trans*, *nm\_item* dan *qty*.

Pada gambar diatas merupakan *file selection* Data\_jual pada Planet *Cash and Credit*. Dan pada *file selection* Data\_jual diatas tidak lengkap untuk lebih jelasnya bisa dilihat di lampiran.

### 3.3 Data Integration (Integrasi Data)

Dalam penelitian kali ini di asumsikan bahwa data yang diambil sudah berupa tabel-tabel dalam suatu database. Untuk proses *mining* data penjualan dari *access* dengan *primary key* *id\_detailtrans*. Setelah itu baru dilakukan proses *mining*. Proses integrasi data dilakukan ketika proses ETL (*ekstrak, transform, and load*) ketika membangun data warehouse, dalam proses ETL data *source* dengan *key* *id\_detailtrans*.



Gambar 3.2. Proses ETL

Berikut adalah hasil proses ETL dengan menggunakan *tools rapid miner*. Dengan format yang sesuai untuk tahap selanjutnya.

Row No	nm Item	id_detailtrans	id_transaksi	tgl_trans	qty
1	Kipas Angin	DT760	5	Jan 3, 2014	2
2	Air Condition	DT761	5	Jan 3, 2014	1
3	Kipas Angin	DT760	5	Jan 3, 2014	1
4	?	?	?	?	?
5	Kipas Angin	DT767	7	Jan 3, 2014	1
6	Mesin Cuci	DT768	8	Jan 3, 2014	2
7	Kipas Angin	DT769	8	Jan 3, 2014	2
8	Kipas Angin	DT760	8	Jan 4, 2014	2
9	Kipas Angin	DT761	9	Jan 4, 2014	2
10	Kipas Angin	DT762	10	Jan 5, 2014	2
11	Botol	DT763	10	Jan 5, 2014	1
12	Air Condition	DT764	11	Jan 5, 2014	2
13	Espresso	DT765	12	Jan 5, 2014	2
14	Kipas Angin	DT766	12	Jan 5, 2014	1
15	Kipas Angin	DT767	13	Jan 6, 2014	1
16	TV	DT768	14	Jan 7, 2014	1
17	Kipas Angin	DT769	14	Jan 7, 2014	2
18	TV	DT770	15	Jan 7, 2014	2
19	Kipas Angin	DT771	15	Jan 7, 2014	1
20	Mesin Cuci	DT772	16	Jan 7, 2014	2
21	Espresso	DT773	17	Jan 7, 2014	2

Gambar 3.3. Hasil Proses ETL

### 3.4. Data Cleaning (Pembersihan Data)

Pembersihan data sebenarnya merupakan tahap awal dari proses *KDD*. merupakan proses menghilangkan *noise* dan data yang telah digabungkan dilakukan pembersihan, membuang data yang kosong dan memastikan data tersebut relevan atau terkait satu sama lain. Pembersihan data juga akan mempengaruhi performansi dari teknik *data mining* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya, pembersihan data terhadap *noise* yang ditemukan berupa *missing value*, *inkonsisten data*, dan *redundant data*.

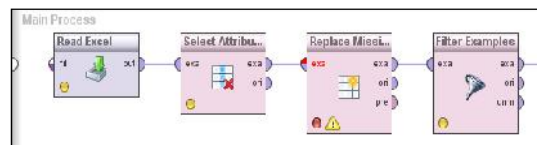
Pada *dataset* contoh Data\_hasil ditemukan data yang *missing value* maka dilakukan *cleaning data* terhadap data dengan *missing value* yang dimaksud. Contoh *dataset* dengan atribut yang *missing value* dapat dilihat pada gambar dibawah ini :

ExampleSet: (1020 examples, 1 special attribute, 4 regular attributes)

Row No.	nm_item	id_data trans	id_transaksi	tgl_trans	qty
1	Kipas Angin	57753	5	Jan 3, 2014	2
2	Air Condition	57754	5	Jan 3, 2014	1
3	Kipas Angin	57755	6	Jan 3, 2014	1
4	?	?	?	?	?
5	Kipas Angin	57757	7	Jan 3, 2014	1
6	Mesin Cuci	57750	0	Jan 3, 2014	2

Gambar 3.4. Data yang *missing value*

Pada gambar 3.4. diatas merupakan contoh data yang *missing value* terlihat bahwa *record* ke-5 semua atribut kosong maka keadaan seperti diatas bisa dikatakan bahwa atribut tersebut *missing value*. Untuk melakukan pembersihan data yang *missing value* dilakukan ketahap proses *cleaning* dari record 5. Dibawah ini adalah tahapan *cleaning* menggunakan tools aplikasi *kattle Rapid miner*.



Gambar 3.5. Proses *Cleaning*

Berikut adalah hasil proses *cleaning* dengan menggunakan tools *rapid miner* Dengan format yang sesuai untuk tahap selanjutnya.

Row No.	id_data trans	id_transaksi	tgl_trans	nm_item	qty
1	57753	5	Jan 3, 2014	Kipas Angin	2
2	57754	5	Jan 3, 2014	Air Condition	1
3	57755	6	Jan 3, 2014	Kipas Angin	1
4	57755	7	Jan 3, 2014	Kipas Angin	2
5	57757	7	Jan 3, 2014	Kipas Angin	1
6	57758	8	Jan 3, 2014	Mesin Cuci	2
7	57759	8	Jan 3, 2014	Kipas Angin	2
8	57760	9	Jan 4, 2014	Kipas Angin	2
9	57761	9	Jan 4, 2014	Kipas Angin	2
10	57762	10	Jan 5, 2014	Kipas Angin	2
11	57763	10	Jan 5, 2014	Detrika	1
12	57764	11	Jan 5, 2014	Air Condition	2
13	57765	12	Jan 5, 2014	Blender	2
14	57766	12	Jan 5, 2014	Kipas Angin	1
15	57767	13	Jan 6, 2014	Kipas Angin	1
16	57768	14	Jan 7, 2014	TV	1
17	57769	14	Jan 7, 2014	Kipas Angin	2
18	57770	15	Jan 7, 2014	TV	2
19	57771	15	Jan 7, 2014	Kipas Angin	1
20	57772	15	Jan 7, 2014	Mesin Cuci	2
21	57773	17	Jan 7, 2014	Blender	2

Gambar 3.6. Hasil proses *cleaning*

### 3.5. Data Transformation

Tahap *transformation* data merupakan tahap merubah data kedalam bentuk yang sesuai untuk di mining. Tahap ini juga berguna untuk membentuk format data yang diterima di perangkat lunak data mining yang memprosesnya. Biasanya perangkat lunak data mining menggunakan format data .csv atau *excel*. Data yang diperoleh terkadang belum tentu sesuai dengan format dan bentuk yang biasa diterima oleh perangkat lunak yang menjalankan algoritma *naive bayes* seperti *rapid miner*, kita harus mengubahnya ke dalam bentuk tertentu. Gambar dibawah ini adalah data dalam bentuk awal. Namun tidak bisa diterima oleh perangkat lunak *rapid miner* untuk ditindak lanjuti. Proses transformasi data dapat dilihat pada gambar 3.7. dibawah ini :

[illegible]

Gambar 3.7. data awal sebelum di transformasi

Selanjutnya *export* dari *microsoft access* ke *microsoft excel*. Hasilnya dapat dilihat pada gambar

	A	B	C	D	E
1	id_defect	id_transaksi	tgl_transaksi	nm_item	qty
2	C2672	3275	31-Dec-16	TV	1
3	C2672	3275	31-Dec-16	Kipas Angin	2
4	C2671	3274	31-Dec-16	TV	2
5	C2671	3274	30-Dec-16	Kipas Angin	2
6	C2669	3272	30-Dec-16	Mesin Cuci	1
7	C2668	3271	30-Dec-16	Kipas Angin	2
8	C2667	3271	30-Dec-16	Kipas Angin	1
9	C2668	3272	30-Dec-16	Kipas Angin	2
10	C2665	3270	30-Dec-16	Setrika	2
11	C2681	3269	29-Dec-16	Blender	2
12	C2681	3269	29-Dec-16	TV	1
13	C2682	3268	28-Dec-16	Kipas Angin	2
14	C2681	3268	28-Dec-16	Blender	2
15	C2663	3267	27-Dec-16	Blender	1
16	C2653	3267	27-Dec-16	TV	1
17	C2653	3265	25-Dec-16	DVD Player	1
18	C2627	3265	25-Dec-16	Setrika	1
19	C2655	3265	25-Dec-16	Mesin Cuci	1
20	C2653	3264	25-Dec-16	Mesin Cuci	2
21	C2654	3264	25-Dec-16	Rice Cooker	2
22	C2655	3265	25-Dec-16	Kipas Angin	2
23	C2652	3265	25-Dec-16	Kipas Angin	2
24	C2651	3262	25-Dec-16	DVD Player	2
25	C2652	3261	23-Dec-16	Kipas Angin	1

Gambar 3.8. data penjualan yang telah diubah ke dalam format .csv atau *microsoft excel*

Setelah data awal berhasil diubah kedalam bentuk .csv , maka tampilan data penjualan item yang dihasilkan dapat dilihat pada gambar 3.9. dibawah ini .

[illegible]

Gambar 3.9. data penjualan dalam format .CSV

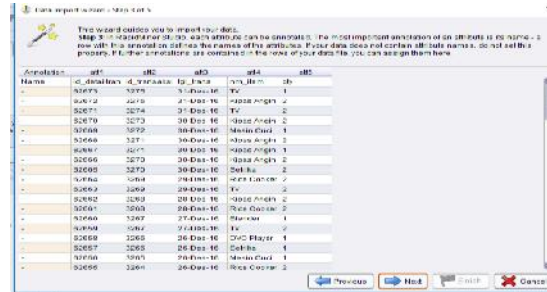
#### 4.1 HASIL DAN PEMBAHASAN

#### 4.1.1. Data Mining

Setelah melakukan transformasi, tahap berikutnya adalah melakukan proses mining data. Proses yang dilakukan bertujuan untuk memprediksi atau mencari nilai tertinggi dari suatu data. Untuk dapat mengumpulkan data, kita memerlukan cara atau proses tertentu. Hasil akhir yang akan diambil adalah

prediksi barang terlaris setiap tahunnya. Dengan mengetahui tingkatan minat konsumen. Tahapan ini adalah inti dari tahapan KDD (*knowledge discovery in database*).

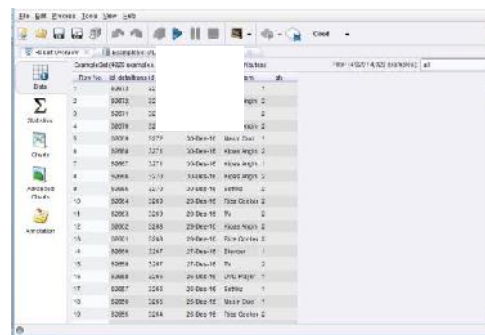
Pada proses data mining, setelah data diubah ke dalam bentuk format .csv, data tersebut langsung dapat digunakan untuk diproses. *Rapidminer* dapat membaca atribut-atributnya seperti pada gambar 4.1 berikut :



id_detailtrans	id_transaksi	tp	jml	jml
300079	300079	30-Dec-16	TV	1
300072	300072	30-Dec-16	Handy Phone	2
300075	300075	30-Dec-16	TV	2
300070	300070	30-Dec-16	Handy Phone	2
300068	300068	30-Dec-16	Motor Cycle	1
300066	300066	30-Dec-16	Handy Phone	2
300067	300067	30-Dec-16	Handy Phone	1
300065	300065	30-Dec-16	Handy Phone	2
300063	300063	30-Dec-16	Handy Phone	2
300062	300062	30-Dec-16	Handy Phone	2
300061	300061	30-Dec-16	Handy Phone	2
300060	300060	30-Dec-16	Handy Phone	2
300059	300059	30-Dec-16	Handy Phone	2
300058	300058	30-Dec-16	Handy Phone	2
300057	300057	30-Dec-16	Handy Phone	2

Gambar 4.1. Data Penjualan setelah dibaca oleh rapid miner

Total data sebanyak 4929 dan id\_detailtrans menjadi *primary key* nya. Bisa dilihat pada gambar 4.2. dibawah ini :

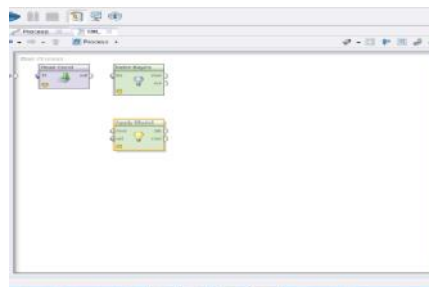


id_detailtrans	id_transaksi	tp	jml	jml
300079	300079	30-Dec-16	TV	1
300072	300072	30-Dec-16	Handy Phone	2
300075	300075	30-Dec-16	TV	2
300070	300070	30-Dec-16	Handy Phone	2
300068	300068	30-Dec-16	Motor Cycle	1
300066	300066	30-Dec-16	Handy Phone	2
300067	300067	30-Dec-16	Handy Phone	1
300065	300065	30-Dec-16	Handy Phone	2
300063	300063	30-Dec-16	Handy Phone	2
300062	300062	30-Dec-16	Handy Phone	2
300061	300061	30-Dec-16	Handy Phone	2
300060	300060	30-Dec-16	Handy Phone	2
300059	300059	30-Dec-16	Handy Phone	2
300058	300058	30-Dec-16	Handy Phone	2
300057	300057	30-Dec-16	Handy Phone	2
300056	300056	30-Dec-16	Handy Phone	2

Gambar 4.2. Data Penjualan Yang Siap Diproses

Untuk menjalankan algoritma *naive bayes* kita harus melewati tahapan-tahapan sebagai berikut :

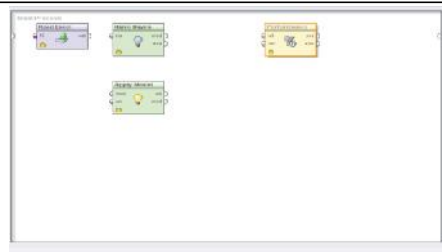
Tahapan *apply model* pembuatan model *apply model* dilakukan dengan menambahkan komponen *apply model* dengan mengetikkan pada kota operator.



Gambar 4.3. Tahapan Apply Model

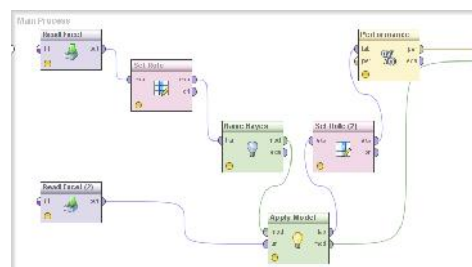
Setelah tahapan *apply model* selanjutnya tahapan *performance* pembuatan model *performance* dilakukan dengan menambahkan komponen *performance* dengan mengetikkan pada kota operator bisa dilihat pada gambar 4.4. sebagai berikut :





Gambar 4.4. Tahapan *performance*

Kemudian tahapan *process* pembuatan model data akan dilakukan dengan menggunakan aplikasi *rapid miner*. Proses ini digunakan untuk membuat *class* mengenai data penjualan.



Gambar 4.5. Tahapan *Process*

Hasil *distribution modelnaive bayes* dilakukan untuk memperoleh informasi melalui data testing berdasarkan kategori produk dengan nilai penjualan terbanyak yaitu produk Kipas Angindengan nilai 0,336.

```
Distribution model for label nm_item

Class Kipas Angin (0.336)
3 distributions

Class Air Conditioner (0.082)
2 distributions

Class Mesin Cuci (0.083)
3 distributions

Class Setrika (0.160)
2 distributions

Class Blender (0.081)
3 distributions

Class TV (0.089)
3 distributions

Class Rice Cooker (0.085)
2 distributions

Class DVD Player (0.084)
3 distributions
```

Gambar 4.6. Hasil *Distribution Model Naive Bayes*

Model distribusi *naive bayes* untuk nm\_item adalah sebagai berikut :  
Kipas angin 0.336 , TV 0.089 , mesin cuci 0.083, setrika 0.160, rice cooker 0.085 , blender 0.081 , DVD player 0.084 , air conditioner 0.082.

#### 4.1.2. Perhitungan Manual Algoritma *Naive bayes*

Melakukan proses prediksi menggunakan metode *naive bayes* berdasarkan kriteria yang ditentukan. Melakukan proses prediksi menggunakan metode *naive bayes* berdasarkan kriteria yang ditentukan dengan menggunakan rumus.

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

Di mana :



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

Tabel 4.1. Diketahui Data Training Penjualan Barang

	id_detailtrans	id_transaksi	nm_item (X1)	Tgl_trans (X2)	qty (X3)
	62673	3275	Kipas Angin	01-12-2016	2
	62672	3275	Mesin Cuci	01-11-2016	1
	62671	3274	Mesin Cuci	01-12-2016	1
	62670	3273	Kipas Angin	01-11-2016	2
	62669	3272	Mesin Cuci	01-10-2016	1
	62668	3271	Mesin Cuci	01-11-2016	1
	62667	3271	Kipas Angin	01-10-2016	2
	62666	3270	Mesin Cuci	01-12-2016	1
	62665	3270	Mesin Cuci	01-11-2016	1
	62673	3275	Mesin Cuci	01-12-2016	1

Menggunakan metode Naive Bayes, tentukan X(X1=Kategori,X2=Tanggal Transaksi,X3=2) ?

Jawab:

Menggunakan metode Naive Bayes, tentukan X(X1=Kategori,X2=Tanggal Transaksi,X3=2) ?

Jawab:

$P(Y=\text{KIPAS ANGIN}) = 3/10$ ,  $P(Y=\text{MESIN CUCI}) = 7/10$

$P(X1=\text{MESIN CUCI}|Y=\text{KIPAS ANGIN}) = 3/3 = 1$

$P(X1=\text{MESIN CUCI}|Y=\text{MESIN CUCI}) = 3/7$

$P(X2=\text{Tanggal Transaksi}|Y=\text{KIPAS ANGIN}) = 1/3$

$P(X2=\text{Tanggal Transaksi}|Y=\text{MESIN CUCI}) = 3/7$

$P(X3=2|Y=\text{KIPAS ANGIN}) = 1/3$

$P(X3=2|Y=\text{MESIN CUCI}) = 1/7$

$P(X1=\text{MESIN CUCI}, X2=\text{Tanggal Transaksi}, X3=2|Y=\text{KIPAS ANGIN})$

$= \{P(X1=\text{MESIN CUCI}|Y=\text{KIPAS ANGIN}) \cdot$

$P(X2=\text{Tanggal Transaksi}|Y=\text{KIPAS ANGIN}) \cdot$

$P(X3=2|Y=\text{KIPAS ANGIN}) \cdot P(Y=\text{KIPAS ANGIN})\} = \{(1) \cdot (1/3) \cdot (1/3) \cdot (3/10)\}$

$= 1,96$

$P(X1 = \text{MESIN CUCI}, X2 = \text{Tanggal Transaksi}, X3 = 2 | Y = \text{MESIN CUCI}) : = \{ P ( P ( X1 = \text{MESIN CUCI} | Y = \text{MESIN CUCI}) \cdot$

$P(X2 = \text{Tanggal Transaksi} | Y = \text{MESIN CUCI}) \cdot$

$P(X3 = 2 | Y = \text{MESIN CUCI}) \cdot$

$P(Y = \text{MESIN CUCI})\} = \{(3/7) \cdot (3/7) \cdot (1/7) \cdot (7/10)\}$

$= 1,7$

Sehingga Keputusannya adalah KIPAS ANGIN.

Keterangan :

P(Y) : Probabilitas Total Keseluruhan Nama Item P

(X1) : Probabilitas PerNama ItemP

(X2) : Probabilitas Tgl Transaksi

P(X3) : Probabilitas Qty

#### 4.1.3. Persebaran (*Deployment*)

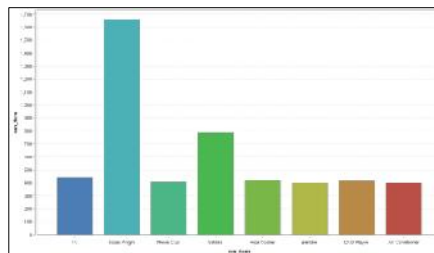
##### 4.1.3.1. Hasil Analisis

*Deployment* merupakan sebuah kegiatan akhir dalam sebuah laporan datamining. Laporan akhir yang berisi pengenalan pola pada data, berdasarkan penelitian yang dilakukan dalam analisis memprediksi penjualan barang elektronik pada planet *cash and credit* data dari tahun 2014 sampai dengan tahun 2016. Berikut hasil yang diperoleh dengan menggunakan *algoritma naive bayes*.

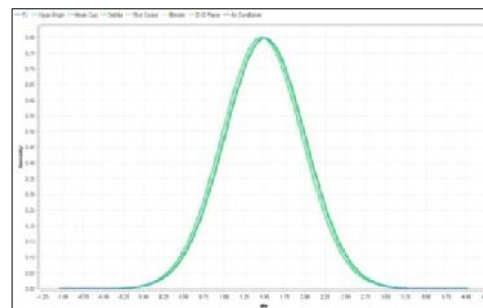
Tabel 4.2. *Nominal Value*

Index	Nominal value	Absolute count	Fraction
1	Kipas Angin	1356	0.336
2	Setrika	737	0.160
3	TV	441	0.089
4	Rice Cooker	420	0.085
5	DVD Player	415	0.084
6	Mesin Cuci	410	0.083
7	Air Conditioner	402	0.082
8	Blender	398	0.081

Tabel diatas merupakan tabel hasil dengan dari proses mining dengan menggunakan *algoritma naive bayes*, daritabel diatas dapat dilihat perbandingan dari produk banyak terjual yaitu dari produk kipas angin dengan nilai 0.336.



Gambar 4.7. Grafik Penjualan Produk



Gambar 4.8. Persentase Grafik Penjualan Produk

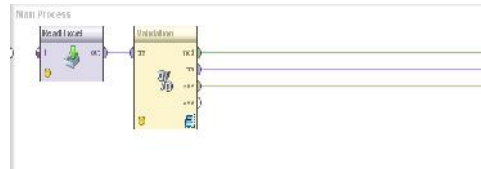
Dapat dilihat dari gambar diatas perbandingan penjualan produk yang banyak terjual menggunakan data yang didapat dari planet *cash and credit*, dari proses perhitungan menggunakan algoritma *naive bayes* dengan nilai *fraction* kipas angin 0.336, setrika 0.160, TV 0.089, *ricecooker* 0.085, *DVD player* 0.084, mesin cuci 0.083, air *conditioner* 0.082, blender 0.081. Berdasarkan hasil perhitungan data mining dan proses pengujian tingkat akurasi dengan menggunakan *rapidminer*, dapat ditarik kesimpulan produk yang

terjual dengan pencapaian 80% dimana nilai ini membuktikan bahwa model yang dibangun dapat digunakan untuk melakukan prediksi.

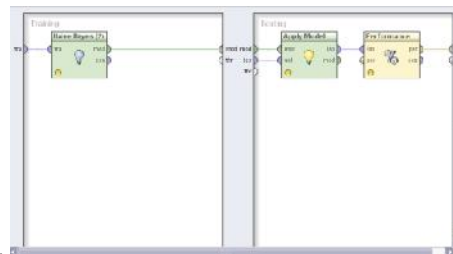
#### 4.1.4. Simulasi Algoritma *Naive Bayes* Kedalam Aplikasi *Rapid Miner*

##### 4.1.4.1. Tahapan *Validation*

*Validation* adalah suatu tindakan yang membuktikan bahwa suatu proses/metode dapat memberikan hasil yang konsisten sesuai dengan spesifikasi yang telah ditetapkan dan terdokumentasi dengan baik.



Gambar 4.9. Tahapan *Validation*



Gambar 5.0. Proses *training* dan *testing*

Pada tahapan *validation* terdapat proses *training* dan *testing*. Pada proses *training*/latihan ada tahapan *naive bayes* setelah melakukan *training* proses selanjutnya adalah proses *testing* digunakan untuk melakukan evaluasi terhadap performa algoritma tersebut.

Accuracy (Jumlah Benar / Jumlah Data)									
	klas Kipas Angin	klas Air Conditioner	klas Mesin Cuci	klas Setrika	klas Blender	klas TV	klas Rice Cooker	klas DVD Player	klas prediksi
pred. Kipas Angin	1003	248	236	149	231	255	261	246	64,02%
pred. Air Conditioner	77	112	113	236	111	116	111	112	6,74%
pred. Mesin Cuci	30	118	235	55	119	236	111	211	6,10%
pred. Setrika	187	117	55	1113	55	116	111	57	17,38%
pred. Blender	36	116	112	211	111	116	116	112	1,56%
pred. TV	110	211	211	55	211	211	111	211	11,51%
pred. Rice Cooker	30	116	111	211	111	111	111	111	1,55%
pred. DVD Player	111	111	211	211	111	111	211	111	1,55%
Overall Accuracy	30,11%	2,46%	3,74%	17,38%	3,52%	8,39%	3,57%	4,58%	

Gambar 5.1. hasil simulasi *naive bayes*

Pada hasil simulasi *naive bayes* informasi yang didapatkan adalah *class recall* kipas angin 60,57% , air conditioner 2,99% , mesin cuci 6,10% , setrika 13,34% , blender 3,52% , TV 8,39% , rice cooker 3,57% dan DVD player 4,58% .

## 5. Kesimpulan

Berdasarkan hasil penelitian penulis yang dilakukan serta pembahasan yang dilakukan oleh penulis maka dapat diambil kesimpulan bahwa :

1. Analisis data penjualan barang pada perusahaan Planet *Cash And Credit* ini digunakan untuk menampilkan informasi barang yang laris terjual dan barang kurang laris menggunakan algoritma *naive bayes*.
2. Algoritma *naive bayes* sangat cocok diterapkan dalam memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga memudahkan perusahaan untuk memprediksi peminatan masyarakat terhadap barang elektronik yang dijual.
3. Dengan mengetahui barang yang laris atau kurang laris terjual akan meminimalisir kerugian pada perusahaan dan perusahaan akan lebih selektif dalam menyetok barang elektronik tersebut.

#### DAFTAR PUSTAKA

- Antoni, D., Negara, E. S., & Suweno, S. (2015). Ekstraksi Data Geo-Spatial Twitter (Studi Kasus: Badan Penyelenggara Jaminan Sosial Kesehatan).
- Ferdiansyah, F., Negara, E. S., & Widyanti, Y. (2019). BITCOIN-USD Trading Using SVM to Detect The Current day's Trend in The Market. *Journal of Information Systems and Informatics*, 1(1), 70-77.
- Intan Cahya Gumilang, Drs Sudjalwo, M.Kom., Aris Rakhmadi, S.T.,M.Eng, 2014. Prediksi Persediaan Obat Dengan Metode Naive Bayes, tersedia [http://eprints.ums.ac.id/31340/22/02.NASKAH\\_PUBLIKASI.pdf/](http://eprints.ums.ac.id/31340/22/02.NASKAH_PUBLIKASI.pdf/), diunduh 22 November 2016, Jurnal Rekomendasi.
- Dicky Nofriansyah, Kamil Erwansyah, Mukhlis Ramadhan, 2016, Penerapan Data Mining dengan Algoritma *Naive Bayes Clasifier* untuk Mengetahui Minat Beli Pelanggan terhadap Kartu *Internet* XL, tersedia: [https://lppm.trigunadharma.ac.id/public/fileJurnal/hpgu1%20Dicky\\_Mei2016.pdf](https://lppm.trigunadharma.ac.id/public/fileJurnal/hpgu1%20Dicky_Mei2016.pdf), diunduh 21 Oktober 2016.
- Negara, E. S. (2018). Chapter\_1 Data Mining, Data and Attribute.
- Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. (2018, October). Improving Naïve Bayes in Sentiment Analysis For Hotel Industry in Indonesia. In 2018 Third International Conference on Informatics and Computing (ICIC) (pp. 1-6). IEEE.
- Yuda Septian Nugroho, 2016, Data Mining Menggunakan Algoritma Naive Bayes Untuk Klasifikasi Kelulusan Mahasiswa Dian Nuswantoro, tersedia: [http://eprints.dinus.ac.id/13239/1/jurnal\\_13789.pdf](http://eprints.dinus.ac.id/13239/1/jurnal_13789.pdf), diunduh 02 November 2016.
- Sutrisno, Afriyudi, Widiyanto 2013, Penerapan Data Mining Pada Penjualan Menggunakan Metode Clustering di <http://eprints.binadarma.ac.id/78/1> Diunduh 17 Oktober 2016.
- Andi, Data Mining dan Web Mining, 2009, <http://andyku.wordpress.com/2008/04/17/data-mining-dan-web-mining/>
- Herdianto (2013 : 8), Pengertian Prediksi . Jurnal Tentang Pengertian Prediksi. <http://digilib.umg.ac.id/files/disk1/16/jipptumg--setyowanto-1512-2-babii.pdf>
- Santoso, B. 2007. Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. Yogyakarta: Graha Ilmu.
- Kusrini & Luthfi, E. T. 2009. Algoritma Data Mining. Yogyakarta : Penerbit Andi.

---

Palembang, 08 Maret 2019

Penulis,

(Novi Pransiska)

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,

(A.Haidar Mirza )

(Andri)