



**PREDIKSI STATUS PENJUALAN DENGAN METODE KLASIFIKASI
NAÏVE BAYES (STUDI KASUS : SEVEN INC JOGJA)**

TUGAS AKHIR

ROSALINA
41517320018

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

2021

UNIVERSITAS
MERCU BUANA



**PREDIKSI STATUS PENJUALAN DENGAN METODE KLASIFIKASI
NAÏVE BAYES (STUDI KASUS : SEVEN INC JOGJA)**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:
ROSALINA
41517320018

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021

UNIVERSITAS
MERCU BUANA

LEMBAR PERNYATAAN ORISINALITAS

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41517320018

Nama : Rosalina

Judul Tugas Akhir : Prediksi Status Penjualan Dengan Metode Klasifikasi Naive Bayes (Studi Kasus : Seven Inc Jogja)

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 04 Maret 2022



Rosalina



UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Rosalina
NIM : 41517320018
Judul Tugas Akhir : Prediksi Status Penjualan Dengan Metode
Klasifikasi *Naïve Bayes* (Studi Kasus : Seven Inc
Jogja)

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 04 Maret 2022



Rosalina

UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN LUARAN TUGAS AKHIR

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Rosalina
 NIM : 41517320018
 Judul Tugas Akhir : Prediksi Status Penjualan Dengan Metode
 Klasifikasi Naïve Bayes (Studi Kasus : Seven Inc
 Jogja)

Menyatakan bahwa :

Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	Diajukan ✓
		Jurnal Nasional Terakreditasi	
		Jurnal International Tidak Bereputasi	Diterima
		Jurnal International Bereputasi	
Disubmit/dipublikasikan di :	Nama Jurnal :		
	ISSN :		
2	Kertas Kerja, Merupakan material hasil penelitian sebagai kelengkapan Artikel Jurnal. Terdiri dari (minimal 4)	Literatur Review	[✓]
		Hasil analisa & perancangan aplikasi	[✓]
		Source code	[✓]
		Dataset	[✓]
		Tahapan eksperimen	[✓]
		Hasil eksperimen seluruhnya	[✓]
3	HAKI Disubmit / Terdaftar	HAKI	Diajukan
		Paten	Tercatat
		No & Tanggal Permohonan	
		No & Tanggal Pencatatan	

Demikian pernyataan ini saya buat dengan sebenarnya.

Mengetahui
 Dosen Pembimbing TA

Jakarta, 30 Maret 2022


 Giri Purnama S.Pd, M.Kom


 Rosalina

LEMBAR PERSETUJUAN

Nama Mahasiswa : Rosalina
NIM : 41517320018
Judul Tugas Akhir : Prediksi Status Penjualan Dengan Metode
Klasifikasi Naïve Bayes (Studi Kasus : Seven Inc
Jogja)

Tugas Akhir ini telah diperiksa dan disetujui

Jakarta, 01 Februari 2022

Menyetujui,



(Giri Purnama S.Pd, M.Kom)
Dosen Pembimbing

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517320018
Nama : Rosalina
Judul Tugas Akhir : Prediksi Status Penjualan Dengan Metode
Klasifikasi Naïve Bayes (Studi Kasus : Seven Inc
Jogja)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 19 Maret 2022

Digitally signed by: SUKMA WARDHANA

Signed at: Mar 19, 2022 07:37:41

Sukma Wardhana

(Sukma Wardhana, S.Kom, M.Kom)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517320018
Nama : Rosalina
Judul Tugas Akhir : Prediksi Status Penjualan Dengan Metode
Klasifikasi Naïve Bayes (Studi Kasus : Seven Inc
Jogja)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 14 Maret 2022



(Sabar Rudianto, M.Kom)


UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517320018
Nama : Rosalina
Judul Tugas Akhir : Prediksi Status Penjualan Dengan Metode
Klasifikasi Naïve Bayes (Studi Kasus : Seven Inc
Jogja)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 30 Maret 2022


(Umniy Salamah, S.Kom, MMSI)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : 41517320018

Nama : Rosalina

Judul Tugas Akhir : Prediksi Status Penjualan Dengan Metode Klasifikasi Naïve Bayes (Studi Kasus : Seven Inc Jogja)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 30 Maret 2022

Menyetujui,



(Giri Purnama, S.Pd, M.Kom)
Dosen Pembimbing

Mengetahui,



(Wawan Gunawan, S.Kom., MT)
Koord. Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM,)
Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas anugerah-Nya, sehingga penulis dapat menyelesaikan tugas akhir ini dengan judul “Prediksi Status Penjualan Dengan Metode Klasifikasi Naïve Bayes (Studi Kasus : Seven Inc Jogja)” dan untuk memenuhi salah satu syarat kelulusan program sarjana jurusan Teknik Informatika Universitas Mercu Buana. Pada kesempatan ini, penulis ingin mengucapkan terima kasih untuk semua yang sudah memberikan bimbingan, dukungan, saran, dan nasehat dalam menyelesaikan tugas akhir ini. Oleh karena itu, dengan rasa bersyukur penulis menyampaikan rasa terima kasih kepada :

1. Tuhan Yang Maha Esa atas berkat dan karunia-Nya untuk penulis bisa menyelesaikan tugas akhir ini,
2. Orang tua yang selalu memberikan doa dan dukungan bagi kesuksesan penulis,
3. Bapak Ir. Emil R. Kaburuan, Ph.D., IPM, selaku Kepala Prodi Teknik Informatika Universitas Mercubuana,
4. Bapak Giri Purnama, S.Pd, M.Kom selaku Dosen Pembimbing,
5. Bapak Sukma Wardhana, S.Kom, M.Kom selaku Ketua Penguji,
6. Bapak Sabar Rudianto, M.Kom selaku Dosen Penguji Pertama,
7. Ibu Umniy Salamah, S.Kom, MMSI selaku Dosen Penguji Kedua,
8. Bapak Rekario Danny selaku Founder Seven Inc Jogja,
9. Serta kepada semua pihak yang terlibat dalam penyusunan laporan kerja praktek ini.

Penulis menyadari bahwa laporan Tugas Akhir ini masih jauh dari kata sempurna, untuk itu penulis mohon maaf. Harapan terakhir, semoga Tugas Akhir ini dapat berguna bagi penulis khususnya dan bagi pembaca pada umumnya.

Jakarta, 06 Juli 2021
Penulis,



Rosalina

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR... ..	iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN	v
LEMBAR PERSETUJUAN PENGUJI	vi
LEMBAR PENGESAHAN	ix
KATA PENGANTAR.....	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
NASKAH JURNAL	xv
I. PENDAHULUAN	xv
II. REFERENSI TERKAIT	xvi
III. METODOLOGI	xvii
3.1 <i>Data Collection</i>	xviii
3.2 <i>Data Selection</i>	xviii
3.3 <i>Data Pre-Processing</i>	xix
3.4 <i>Data Mining</i>	xx
IV. HASIL DAN PEMBAHASAN.....	xx
V. KESIMPULAN	xxii
KERTAS KERJA.....	xxiii
BAB 1. LITERATUR REVIEW	1
BAB 2. ANALISIS DAN PERANCANGAN.....	4
BAB 3. DATASET.....	5
BAB 4. TAHAPAN EKSPERIMEN.....	12
BAB 5. HASIL SEMUA EKSPERIMEN.....	17
DAFTAR PUSTAKA	21
LAMPIRAN DOKUMEN HAKI.....	23



UNIVERSITAS
MERCU BUANA

DAFTAR GAMBAR

Gambar 1 Tahap Penelitian	xvii
Gambar 2 Data Penjualan 2020 Sebelum Diolah.....	xix
Gambar 3 Flowchart Algoritma <i>Naïve Bayes</i>.....	4
Gambar 4 Tampilan RapidMiner	12
Gambar 5 Gambar Perspektif dan View Pada RapidMiner	13
Gambar 6 Tampilan Process Pada RapidMiner	14
Gambar 7 Dataset Berekstensi .xlsx.....	14
Gambar 8 Tampilan Operator dan Parameter Pada Menu Process	14
Gambar 9 Tampilan Ratio Pada RapidMiner	15
Gambar 10 Tampilan Format Columns	16
Gambar 11 Tampilan Hasil Accuracy Pada Performance Vector	17
Gambar 12 Tampilan Hasil Recall Pada Performance Vector.....	17
Gambar 13 Tampilan Hasil Precision Pada Performance Vector.....	18
Gambar 14 Produk Terlaris pada Jenis Pakaian Formal di Harga 399.000.	18
Gambar 15 Produk Terlaris pada Jenis Pakaian Informal di Harga 300.000	19

DAFTAR TABEL

Tabel 1 Variabel Data Sebelum Diolah.....	xviii
Tabel 2 Variabel Data Sesudah Diolah	xix
Tabel 3 Confusion Matrix	xx
Tabel 4 Dataset Yang Sudah Diolah.....	5
Tabel 5 Perhitungan dengan Confusion Matrix.....	19



NASKAH JURNAL

I. PENDAHULUAN

Perkembangan industri kreatif di Indonesia pada saat ini sangatlah pesat, sehingga persaingan antara para pelaku usaha yang terkait tak bisa dihindarkan. Untuk dapat bersaing di tengah pesatnya perkembangan industri kreatif tersebut, mendorong para pelaku usaha untuk dapat menyusun strategi dalam memenangkan persaingan, yang salah satunya adalah mengolah data yang dimilikinya untuk menjadikannya informasi yang bermanfaat bagi perusahaan.

Seven Inc Jogja adalah salah satu perusahaan digital yang bergerak di bidang industri kreatif, dengan produk awalnya adalah fashion yang pusat operasinya berlokasi di Jogja. Seiring perkembangannya perusahaan ini membutuhkan informasi mengenai produk yang paling laris yang telah dipasarkannya, untuk dapat digunakan sebagai salah satu variabel dalam penyusunan strategi perusahaan. Berdasarkan latar belakang masalah ini, analisis data mining akan coba diterapkan pada data penjualan di Seven Inc Jogja untuk dapat menjawabnya.

Tersedianya data dalam jumlah besar dan kebutuhan akan informasi untuk mendukung pengambilan keputusan sebagai bentuk solusi bisnis maka teknik data mining merupakan solusi yang tepat. Data mining adalah rangkaian proses yang mengeksplorasi nilai tambah dari suatu koleksi data berupa pengetahuan yang belum diketahui. Pengetahuan tentang prediksi status penjualan pakaian dapat diperoleh dari sejumlah besar data penjualan pakaian. Dengan menggunakan data mining, kumpulan data dapat diolah dan diprediksi untuk mengetahui status penjualannya dengan menggunakan metode naïve bayes.

Metode *naïve bayes* merupakan salah satu metode yang termasuk dalam teknik klasifikasi. Naïve Bayes adalah klasifikasi yang diusulkan oleh ilmuwan Thomas Bayes asal Inggris, yaitu meramalkan peluang masa depan berdasarkan pengalaman masa lalu yang disebut dengan teorema bayes.

Teorema ini digabungkan dengan kenafian yang mengasumsikan bahwa kondisi antara atribut tidak bergantung satu sama lain. Klasifikasi naïve bayes diasumsikan bahwa ada atau tidaknya karakteristik tertentu dari suatu kelas tidak ada sangkut pautnya dengan karakteristik kelas lainnya.

Oleh karena itu, penulis menggunakan algoritma naïve bayes untuk membuat prediksi status penjualan yang dapat digunakan untuk memperoleh informasi tentang prediksi penjualan produk terlaris pada perusahaan Seven Inc Jogja.

II. REFERENSI TERKAIT

Data mining (biasa dikenal *pattern recognition*) adalah metode pemrosesan yang digunakan untuk menemukan pola tersembunyi untuk diproses dari data dan hasil, untuk diolah menjadi ilmu pengetahuan baru dan informasi untuk pengambilan keputusan di masa depan [1].

Data mining juga disebut sebagai sistem pemrosesan data yang super besar, yaitu memberikan peran dalam banyak bidang di seluruh dunia, yaitu keuangan, industri, transportasi, cuaca, dan teknologi. Data mining dapat diartikan sebagai penyimpanan data dalam jumlah besar yang ada di dalam database dan *data warehouse* diandalkan untuk menyimpan serta menemukan pola. Adapula metode pada penambahan data seperti analisis data, pemrosesan sinyal, jaringan saraf, serta *pattern recognition* [1].

Suntoro, J (2019) berpendapat bahwa Teorema Bayes menghitungkan nilai berdasarkan kondisi (*posterior probability*) $P(H|X)$ menggunakan probabilitas $P(H)$, $P(X)$, dan $P(X|H)$. Nilai X adalah data. Nilai H adalah asumsi dari nilai X yang merupakan class yang lebih spesifik. Nilai $P(X|H)$ atau *likelihood* adalah probabilitas dari asumsi nilai X berdasarkan kondisi nilai H . Sedangkan, nilai $P(X)$ disebut *predictor prior probability*, yang merupakan *probabilitas prior data X* [20].

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

Keterangan :

$P(H|X)$ = Probabilitas asumsi berdasarkan kondisi (*posteriori probability*).

$P(X|H)$ = *likelihood* dari asumsi nilai X berdasarkan kondisi class H.

$P(H)$ = Probabilitas asumsi prior probability class H.

$P(X)$ = Probabilitas prior data X.

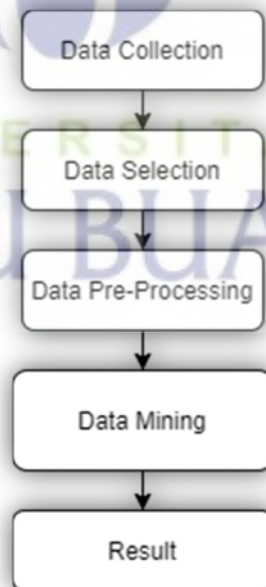
X = data

H = class

Teknik klasifikasi dilakukan untuk memprediksi label kategori yang tidak diketahui menggunakan data yang berbeda dengan label kategori yang diketahui. Klasifikasi didefinisikan untuk proses deklarasi objek data sebagai sebuah kategori kelas yang telah ditentukan sebelumnya [20].

III. METODOLOGI

Tahap penelitian yang dilakukan bertujuan sebagai bentuk lanjutan dari penelitian kualitatif. Tahapan-tahapan pada penelitian ini, yaitu *data collection*, *data selection*, *data pre-processing*, *data mining*, dan *result*. Gambar 1 merupakan tahap penelitian yang dilakukan, dengan adanya gambar untuk tahapan penelitian maka akan lebih mudah dipahami.



Gambar 1 Tahap Penelitian

[Sumber: Dokumen Pribadi]

3.1 Data Collection

Pada tahap *data collection*, semua data dikumpulkan. Pengumpulan data sudah dilakukan dengan cara mendapatkan data dari sumbernya yaitu pemilik Seven Inc. Data yang dikumpulkan berupa data penjualan di bulan Januari 2020.

Tabel 1 Variabel Data Sebelum Diolah

No	Atribut
1	Nama Juragan
2	CS Utama
3	CS Closing
4	Nomor Resi
5	Nama Buyer
6	Jenis Pakaian
7	Kode
8	Size
9	Jumlah Transfer
10	Bank
11	Pembayaran Resi
12	Harga Setelah Dikurangi Resi
13	Harga Produk

Tabel 1 merupakan variabel data sebelum diolah dan belum dilakukan sampai kepada tahap penyeleksian. Data yang sudah dikumpulkan terdiri dari 507 records di bulan Januari 2020 yang terdiri dari 13 variabel yang terdiri dari Nama Juragan, CS Utama, CS Closing, Nomor Resi, Nama Buyer, Jenis Pakaian, Kode, Size, Jumlah Transfer, Bank, Pembayaran Resi, Harga Setelah Dikurangi Resi, dan Harga Produk yang sesuai dengan atribut yang dimilikinya.

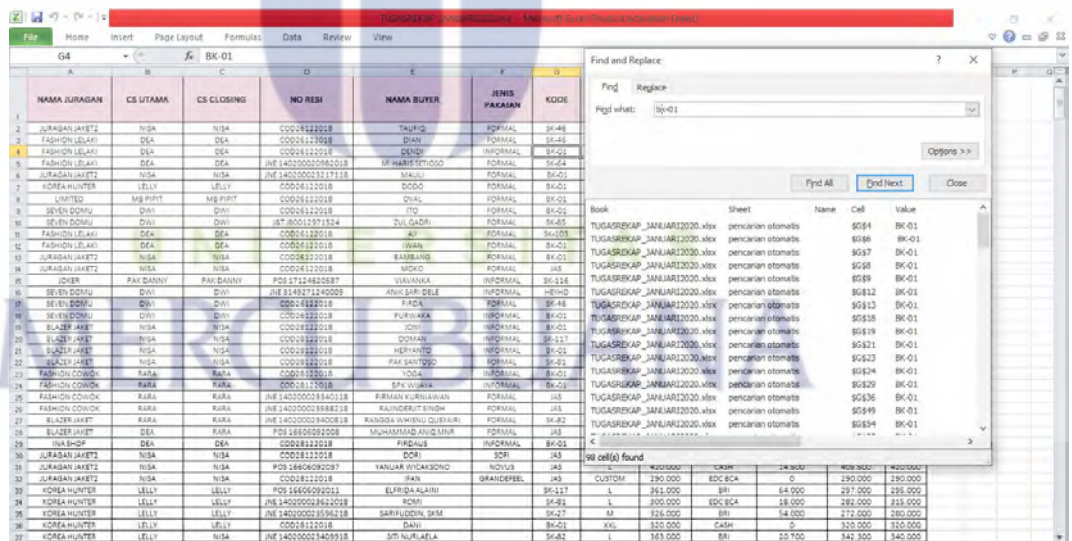
3.2 Data Selection

Pada tahap *data selection*, semua data diseleksi termasuk data yang tidak relevan seperti dengan menghilangkan data outlier atau data yang anomali. Dengan tujuan untuk mendapatkan hasil terbaik. Maka perlu dilakukan beberapa tahap penyeleksian seperti memilah beberapa variabel supaya mampu memenuhi tolok ukur pada saat proses pengklasifikasian.

Tabel 2 Variabel Data Sesudah Diolah

No	Atribut
1	Kode Barang
2	Jenis Pakaian
3	Size
4	Harga
5	Status

Tabel 2 merupakan tabel data yang sudah diolah. Dengan variabel awal yaitu 13 variabel lalu diseleksi menjadi 5 variabel yang terdiri dari Kode Barang, Jenis Pakaian, Size, Harga, dan Status yang sesuai dengan atribut yang dimilikinya. Status Laris ditentukan dengan penjualan maksimal 15 pakaian per bulannya sesuai dengan kode barang. Cara menentukan status Laris yaitu secara manual di excel dengan menekan tombol CTRL+F, misalnya ingin mencari kode barang BK-01 dan terdapat produk yang terjual lebih dari 15 produk di bulan Januari 2020. Dengan menggunakan data sebelum diolah yang berjumlah 507 records.



Gambar 2 Data Penjualan 2020 Sebelum Diolah

3.3 Data Pre-Processing

Pada tahap *data pre-processing*, semua data diidentifikasi untuk mengetahui data yang tidak konsisten dan data yang mengalami ketidaklengkapan. Tahap *data pre-processing* dilakukan dengan tujuan untuk

melakukan perbaikan data dengan cara memilih karakteristik yang sesuai sebelum dilakukan kepada tahap pemodelan kemudian dilakukan tahap transformasi dengan tujuan mengurangi atribut yang tidak konsisten yang akan dilakukan menggunakan RapidMiner.

3.4 Data Mining

Pada tahap *data mining*, penulis menggunakan metode *naïve bayes* untuk melakukan pengklasifikasian dan melakukan analisis pada dataset yang penulis dapatkan dari sumbernya, yaitu pemilik Seven Inc. Tahapan ini merupakan langkah-langkah pemrosesan data latih yang akan digunakan dengan tujuan untuk melatih model yang akan digunakan pada data uji.

IV. HASIL DAN PEMBAHASAN

Pada tahap hasil dan pembahasan merupakan tahap hasil dimana seluruh data latih akan ditentukan hasil akurasi datanya dan juga melakukan uji pada model yang sudah ditentukan. Penulis menggunakan *confusion matrix* untuk menganalisis model dengan tujuan supaya data yang belum digunakan untuk pelatihan dapat digunakan sebagai data uji yang berguna untuk menguji model algoritma yang sudah ditentukan dan untuk menetapkan hasil akurasi pada data. *Confusion matrix* mampu memberikan hasil *accuracy*, *precision*, dan *recall*.

Confusion Matrix merupakan alat untuk menghitung hasil evaluasi yang membandingkan nilai aktual dengan nilai dari hasil prediksi model yang digunakan yang memiliki tabel berupa metrik pengukuran seperti *Accuracy*, *Precision*, dan *Recall* [15]. *Confusion Matrix* merupakan hasil kerja dalam bentuk tabel [18].

Tabel 3 Confusion Matrix

PREDICTED VALUES	1	0
1	True Positive	False Positive
0	False Negative	True Negative

Keterangan :

True Positive (TP) merupakan jumlah data dari *class 1 actual values* dan bernilai benar sebagai *class 1 predicted values*.

False Negative (FN) merupakan jumlah data dari *class 1 actual values* dan bernilai salah sebagai *class 0 predicted values*.

False Positive (FP) merupakan jumlah data dari *class 0 actual values* dan bernilai salah sebagai *class 1 predicted values*.

True Negative (TN) merupakan jumlah data dari *class 0 actual values* dan bernilai benar sebagai *class 0 predicted values*.

Pada keterangan tabel 2.2, berikut ini cara menghitung *accuracy*, *precision*, dan *recall* dengan *Confusion Matrix* :

1. *Accuracy*. Nilai *accuracy* yang didapatkan dari jumlah data yang bernilai benar.

$$\text{Rumus Accuracy} : \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

2. *Precision*. Nilai *precision* yang didapatkan dari hasil kejadian yang bernilai benar.

$$\text{Rumus Precision} : \frac{TP}{TP+FP} \times 100\% \quad (3)$$

3. *Recall*. Nilai *recall* merupakan nilai dengan kategori positif yang diprediksi dengan benar.

$$\text{Rumus Recall} : \frac{TP}{TP+FN} \times 100\% \quad (4)$$

Pada penelitian, penulis menggunakan perbandingan 80% untuk data latih dan 20% untuk data uji dari dataset penjualan.

Rahmadya Trias Handayanto dan Herlawati (2020:292), bahwa pada metode validasi silang terdapat saran untuk tidak membagi dataset menjadi 50% data latih dan 50% data uji melainkan beberapa partisi [12].

Dios Kurniawan, M.Sc (2020:15), mengatakan data uji jumlahnya lebih sedikit dan umumnya 20% dari total seluruh data [15].

V. KESIMPULAN

Pada penelitian yang dilakukan dengan menggunakan dataset penjualan yang sudah diolah sebanyak 255 records dengan dataset awal yang belum diolah sebanyak 507 records menghasilkan kesimpulan yang didapat diambil yaitu :

1. Berdasarkan hasil perhitungan dengan metode *Naïve Bayes* menggunakan RapidMiner maka mendapatkan nilai *accuracy* sebesar 90,20%, *precision* sebesar 87,10%, dan *recall* sebesar 96,43% untuk status produk laris. Nilai tidak akurasi terjadi pada saat nilai error dan persen error yang diperoleh terlalu tinggi.
2. Dengan menggunakan 80% data latih maka akan memberikan dampak terhadap hasil ujinya termasuk tingkat *accuracy*, *precision*, dan *recall*. 80% merupakan rasio perbandingan dataset yang terdiri dari data latih dan data uji. 80% merupakan data latih dan 20% merupakan data uji yang pada umumnya digunakan.
3. Berdasarkan hasil perhitungan menggunakan metode *Naïve Bayes*, produk terlaris dalam penjualan pakaian adalah produk dengan jenis pakaian formal dengan harga Rp.399.000 dan untuk produk dengan jenis pakaian informal terlaris di harga Rp.300.000. Dari hasil pengolahan data bulan Januari 2020 tersebut diharapkan dapat dimanfaatkan pelaku usaha untuk dapat menyusun strategi yang dapat meningkatkan penjualan produknya.
4. Dan metode *Naïve Bayes* tepat digunakan untuk mengolah data yang terbatas dalam rangka mendapatkan informasi yang berguna bagi pelaku usaha untuk meningkatkan pendapatannya.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada orang tua, dosen pembimbing, Seven Inc, serta kerabat atas dukungan dalam pelaksanaan penelitian.

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul “Prediksi Status Penjualan Dengan Metode Klasifikasi *Naïve Bayes*”. Kertas kerja berisi semua material hasil penelitian Tugas Akhir yang tidak dimuat/atau disertakan di artikel jurnal. Di dalam kertas kerja ini disajikan: literature review, analisis dan perancangan, source code, dataset yang digunakan, tahapan eksperimen dan hasil semua eksperimen secara keseluruhan.

Pada bagian I literature review, di dalam literature review ini disajikan hasil review atas literature yang terkait dengan penelitian yaitu: konsep data mining, *Naïve Bayes* dan klasifikasi. Literatur membantu penulis dalam mencari informasi yang dibutuhkan, membantu memperkuat informasi hasil dari suatu analisis atau hipotesa dan juga memberi tambahan informasi. Bagian II adalah analisis dan perancangan, pada bagian analisis dan perancangan diuraikan tahapan rancangan pengolahan data dan bagaimana rancangan dari awal hingga selesai. Bagian III adalah source code, pada bagian ini penulis memaparkan proses pengolahan data yaitu parameter yang digunakan dan keluaran yang dihasilkan yaitu prediksi dan akurasi menggunakan aplikasi Rapidminer. Bagian IV adalah dataset, bagian ini memaparkan dataset yang akan diolah yaitu data penjualan di bulan Januari 2020. Yang terdiri dari 5 variabel yaitu kode barang, jenis pakaian, size, harga, dan status yang sesuai dengan atribut yang dimilikinya. Data disimpan dalam format .xlsx kemudian diolah menggunakan aplikasi RapidMiner. Kemudian bagian V dan bagian VI yaitu tahapan eksperimen dan hasil eksperimen secara keseluruhan terhadap dataset yang digunakan menggunakan aplikasi RapidMiner.