

## **BAB II**

### **TINJAUAN PUSTAKA**

Dalam bab ini akan menjelaskan tentang seluruh teori, bahan penelitian lain yang diarahkan untuk menyusun konsep yang berkaitan dengan penelitian dan terdiri dari penjelasan studi-studi sebelumnya dan dasar-dasar teori yang digunakan.

#### **2.1 Penelitian Terdahulu**

Penelitian terdahulu ini menjadi salah satu acuan penulis dalam melakukan penelitian sehingga penulis dapat memperkaya teori yang digunakan dalam mengkaji penelitian yang dilakukan. Dari penelitian terdahulu, penulis tidak menemukan penelitian dengan judul yang sama seperti judul penelitian penulis. Namun penulis mengangkat beberapa penelitian sebagai referensi dalam memperkaya bahan kajian pada penelitian penulis. Berikut merupakan penelitian terdahulu berupa beberapa jurnal terkait dengan penelitian yang dilakukan penulis.

##### **2.1.1 Penelitian I-Teguh Yulianto (2019)**

Penelitian terdahulu pertama yang dilakukan oleh teguh yulianto pada tahun 2019 dengan mengambil judul “Prediksi Penjualan Produk Menggunakan Algoritma Naïve Bayes”.

Tujuan dari penelitian ini bertujuan untuk mengimplementasikan Algoritma Naïve Bayes pada sistem penjualan produk di Couple Store. Sedangkan hasil dari penelitian ini adalah diimplementasikan Algoritma Naïve Bayes untuk memprediksi penjualan produk pada Couple Store supaya dapat membantu serangkaian proses penjualan produk sesuai yang dibutuhkan masyarakat. Selain itu juga bisa membantu pengguna dalam proses peningkatan produk.

### **2.1.2 Penelitian II-Fakhrizal Rizki (2020)**

Penelitian terdahulu kedua yang dilakukan oleh Fakhrizal rizki pada tahun 2020 dengan mengambil judul “Penerapan Metode Naïve Bayes Untuk Penjualan Pada UD. Hikmah Pasuruan Berbasis Web”.

Tujuan penelitian ini adalah menerapkan metode naïve bayes pada sistem penjualan perlengkapan sekolah. Berdasarkan hasil pengujian dengan metode Black box bahwa perangkat lunak dapat mengetahui fungsi-fungsi yang tidak benar atau hilang, kesalahan *interface*, kesalahan dalam struktur data atau akses database eksternal, kesalahan kinerja, inisialisasi, kesalahan terminasi dan secara fungsional mengeluarkan hasil yang sesuai dengan yang diharapkan.

### **2.1.3 Penelitian III-Yusuf Sulisty Nugroho (2017)**

Penelitian terdahulu ketiga yang dilakukan oleh Yusuf Sulisty Nugroho pada tahun 2017 dengan mengambil judul “Penerapan Data Mining Untuk Prediksi Rating Penjualan Buku Menggunakan Metode Naive Bayes”.

Tujuan penelitian ini adalah untuk memprediksi rating penjualan buku. Sedangkan hasil dari penelitian ini adalah sistem data mining menggunakan metode naïve bayes untuk memprediksi rating penjualan buku memiliki kinerja yang baik dari segi akurasi, presisi dan *recall*. Dengan nilai akurasi sebesar 74,60%.

### **2.1.4 Penelitian IV-Dini Rezekika (2020)**

Penelitian terdahulu keempat yang dilakukan oleh Siska Delima pada tahun 2020 dengan mengambil judul “Prediksi Penyediaan SparePart Kendaraan Roda Dua Dengan Metode Naïve Bayes ”.

Tujuan penelitian ini adalah menguji keakuratan dari Metode Bayes dalam penyetokan sparepart. Sedangkan hasil dari penelitian ini menghasilkan sebuah sistem yang dapat membantu karyawan dalam menentukan persediaan stok sparepart pada Toko dewi Motor.

### **2.1.5 Penelitian V-Arya Mukti Saputra (2020)**

Penelitian terdahulu kelima yang dilakukan oleh Arya Mukti Saputra pada tahun 2020 dengan mengambil judul “Sistem Prediksi Persediaan Obat Pada Apotek Menggunakan Metode Naïve Bayes”.

Tujuan penelitian ini adalah membangun sebuah sistem pengelolaan persediaan obat pada apotek dengan menggunakan metode naive bayes. Sedangkan hasil dari penelitian ini adalah Rancangan sistem yang dibuat dapat mengolah data detail obat dan transaksi keluar masuknya obat yang sedang berjalan serta menampilkan laporan harian.

## **2.2 Sistem Prediksi**

Dari jurnal yang saya baca prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi (Herdianto, 2013 : 8).

Pengertian Prediksi sama dengan ramalan atau perkiraan. Menurut kamus besar bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Prediksi menunjukkan apa yang akan terjadi pada suatu keadaan tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan. Prediksi bisa berdasarkan metode ilmiah ataupun subjektif belaka. Ambil contoh, prediksi cuaca selalu berdasarkan data dan informasi terbaru yang didasarkan pengamatan termasuk oleh satelit.

Pengertian Prediksi sama dengan ramalan atau perkiraan. Menurut kamus besar bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Prediksi menunjukkan apa yang akan terjadi

pada suatu keadaan tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan.

Kesimpulannya prediksi adalah memperkirakan atau meramal kejadian di masa depan berdasarkan data di masa lalu.

## 2.3 Data Mining

Data mining adalah suatu proses penambangan informasi penting dari suatu data. Informasi penting ini didapat dari suatu proses yang amat rumit seperti menggunakan *artificial intelligence*, teknik statistik, ilmu matematika, *machine learning*, dan lain sebagainya. Teknik-teknik rumit tersebut nantinya akan mengidentifikasi dan mengekstraksi informasi yang bermanfaat dari suatu database besar.

Pengertian Data Mining atau *Knowledge Discovery in Database* menurut William J. Frawley, Gregory Piatetsky-Shapiro dan Christopher J. Matheus [2010]: Data Mining atau *Knowledge Discovery in Database* (KDD) adalah penyaringan data secara implisit di mana sebelumnya tidak diketahui terdapatnya informasi yang potensial.

### 2.3.1 Fungsi dan Tugas Data Mining

Data mining menganalisis data menggunakan *tool* untuk menemukan pola dan aturan dalam himpunan data. Perangkat lunak bertugas untuk menemukan pola dengan mengidentifikasi aturan dan fitur pada data. *Tool* Data mining diharapkan mampu mengenal pola ini dalam data dengan input minimal dari *user*. Dalam penelitian ini pembahasan Data Mining diklasifikasikan dalam fungsi *Association*.

### 2.3.2 Kelebihan Data Mining

Data mining merupakan sebuah proses interaktif untuk mendapatkan sebuah pola baru yang menarik. Pola tersebut tentunya akan sangat bermanfaat. Model yang dihasilkan dari proses data mining biasanya sudah sempurna sehingga dapat digeneralisasi untuk kepentingan di masa depan. Karena prosesnya yang cukup panjang dan rumit, maka dari proses awal

biasanya akan menghasilkan sesuatu yang baru, yang tidak diketahui sebelumnya. Sesuatu yang baru ini akan menambah pengetahuan para pengguna ataupun peneliti dan tentunya akan sangat bermanfaat karena dapat digunakan untuk melakukan tindakan tertentu. Penggalan data juga sering dikatakan sebagai proses interaktif. Proses interaktif maksudnya yaitu proses yang masih memerlukan interaksi manusia agar bisa terlaksana. Sedangkan proses interaktif, maksudnya adalah proses yang tidak hanya dilakukan sekali, perlu proses yang berulang-ulang untuk mendapatkan data penting yang dimaksud.

Kelebihannya membuat analisa suatu data besar menjadi semakin mudah. Pencarian pola baru atau trend baru bisa dilakukan dengan mudah sehingga bisa membantu mengambil keputusan di masa yang akan datang atau bisa memprediksi data tertentu sehingga bisa menganalisis apa yang harus dilakukan.

## 2.4 Pengertian DFD

DFD adalah suatu diagram yang menggambarkan aliran data dari sebuah proses yang sering disebut dengan sistem informasi. Di dalam *data flow diagram* juga menyediakan informasi mengenai input dan output dari tiap entitas dan proses itu sendiri. Dalam diagram alir data juga tidak mempunyai kontrol terhadap *flow*-nya, sehingga tidak adanya aturan terkait keputusan atau pengulangan. Bentuk penggambaran berupa *data flowchart* dengan skema yang lebih spesifik. Menurut Kenneth Kozar, tujuan dari adanya DFD sendiri adalah sebagai penyedia atau menjembatani antara pengguna dengan sistem. *Data flow diagram* berbeda dengan UML (*Unified Modelling Language*), di mana hal mendasar yang menjadi pembeda antara kedua skema tersebut terletak pada *flow* dan objective penyampaian informasi di dalamnya.

Pengertian *Data Flow Diagram* atau DFD yang dikemukakan oleh Ladjamudin (2013), “Diagram Aliran Data merupakan model dari sistem untuk menggambarkan pembagian sistem ke modul yang lebih kecil”. Selain

pengertian DFD menurut Ladjamudin tersebut terdapat dua pengertian DFD menurut pendapat ahli yang lain yaitu Menurut Indrajani (2015:27), “*Data Flow Diagram* (DFD) adalah sebuah alat yang menggambarkan aliran data sampai sebuah sistem selesai, dan kerja atau proses dilakukan dalam sistem tersebut” dan menurut Saputra (2013:118) menjelaskan, “*Data Flow Diagram* atau yang disingkat DFD merupakan suatu diagram yang menggambarkan alir data dalam suatu entitas ke sistem atau sistem ke entitas”.

#### **2.4.1 Fungsi Data Flow Diagram**

- Menyampaikan rancangan sistem
- Menggambarkan suatu sistem
- Perancangan model

### **2.5 Pengertian WEB**

*Website* adalah kumpulan halaman dalam suatu domain yang memuat tentang berbagai informasi agar dapat dibaca dan dilihat oleh pengguna internet melalui sebuah mesin pencari. Informasi yang dapat dimuat dalam sebuah website umumnya berisi mengenai konten gambar, ilustrasi, video, dan teks untuk berbagai macam kepentingan. Biasanya untuk tampilan awal sebuah website dapat diakses melalui halaman utama (*homepage*) menggunakan browser dengan menuliskan URL yang tepat. Di dalam sebuah *homepage*, juga memuat beberapa halaman web turunan yang saling terhubung satu dengan yang lain.

#### **2.5.1 Fungsi WEB**

Terdapat beberapa fungsi *website* yang memiliki keunggulan dan kualitas yang berbeda sesuai dengan target pemasaran maupun bisnis. Berikut ini merupakan beberapa fungsi yang dikategorikan sesuai dengan tujuan bisnis.

- Sarana informasi
- Sebagai Blog
- Sarana *Ecommerce*

### 2.5.2 Sejarah WEB

Sejarah *website* pertama kali dimulai dari seorang ilmuwan yang berasal dari Inggris, bernama Tim Berners-Lee. Orang tua dari Berners juga merupakan ilmuwan komputer pada era awal dunia komputasi. Tujuan awal dari Tim Berners membuat sebuah *website* adalah supaya lebih memudahkan para peneliti di tempat kerjanya untuk mendapatkan dan bertukar informasi. Kemudian, pada tanggal 30 April 1993, secara resmi CERN yang merupakan laboratorium fisika di Swiss mengumumkan tentang perilisan *website* secara gratis.

Sebelum itu pada tahun 1990, Tim Berners-Lee juga menuliskan tentang tiga teknologi dasar web, antara lain:

➤ HTML (*HyperText Markup Language*)

Merupakan bahasa markup atau format untuk halaman web.

➤ URI (*Uniform Resource Identifier*)

Merupakan sebuah alamat unik untuk membuka halaman situs. Fungsinya adalah mengidentifikasi setiap sumber daya yang ada pada web. Saat ini sering disebut dengan URL (*Uniform Resource Locator*).

➤ HTTP (*HyperText Transfer Protocol*)

Teknologi ini memungkinkan seseorang untuk mengambil kembali sumber daya yang terkoneksi dengan semua situs web.

### 2.6 Pengertian Penjualan

Menurut Reeve, Warren, dan Durhac pengertian dari penjualan adalah sejumlah total yang dikenakan kepada pelanggan untuk barang dagangan yang dijual, termasuk penjualan tunai dan kredit. Penjualan bisa diartikan dengan proses pemenuhan kebutuhan penjual dan pembeli baik secara tunai maupun kredit.

Penjualan sudah pasti merupakan bagian dari suatu bisnis. Baik itu penjualan barang maupun jasa. Proses penjualan menjadi salah satu tolak ukur apakah bisnis bisa berjalan lancar atau tidak. Jika penjualan memiliki

angka yang tinggi, berarti pelanggan banyak yang membutuhkan barang atau jasa yang dijual. Jika yang terjadi adalah kebalikannya, berarti ada kesalahan terhadap produk atau mungkin bisnis tidak menasar target pasar yang tepat. Maka dari itu, penjualan juga bisa menjadi tolak ukur untuk evaluasi bisnis. Karena dari penjualan bisa diketahui data yang akurat tentang kondisi produk atau jasa terhadap pasar.

### **2.6.1 Manfaat dan Tujuan Penjualan**

Proses penjualan adalah proses penting dalam semua bisnis. Alasan mengapa penjualan adalah hal yang penting karena bisnis akan mendapatkan banyak manfaat dari penjualan. Berikut ini adalah manfaat dari penjualan yang bisa didapat bisnis:

- Mendapatkan Laba tertentu
- Mendapatkan Volume Penjualan
- Pertumbuhan Bisnis

## **2.7 Algoritma Naïve Bayes**

Secara umum algoritma naïve bayes untuk membangun sebuah hasil kesimpulan adalah sebagai berikut “Naïve Bayes *Classifier* merupakan salah satu algoritma dalam teknik data mining yang menerapkan teori bayes dalam klasifikasi. Teorema keputusan bayes adalah pendekatan statistik yang fundamental dalam pengenalan pola (*Pattern Recognition*), Naïve bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Algoritma Naive Bayes merupakan suatu bentuk klasifikasi data dengan menggunakan metode probabilitas dan statistik. Algoritma Naive Bayes dapat diartikan sebagai sebuah metode yang tidak memiliki aturan, Naive Bayes menggunakan cabang matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi dengan cara melihat frekuensi tiap klasifikasi pada data training.

Naive Bayes *Classifier* bekerja sangat baik dibanding dengan model *classifier* lainnya. Hal ini dibuktikan pada penelitian Xhemali tahun 2009

yang membandingkan beberapa algoritma seperti *Neural Networks*, Naive Bayes, dan *Decision Tree* untuk analisis dan klasifikasi otomatis terhadap data atribut dari *training course web pages*, Naive Bayes merupakan algoritma terbaik dibandingkan *Neural Network* dan *Decision Tree* dengan nilai *F-Measure* lebih dari 97%. Metode Algoritma Naive Bayes Classifier lebih mudah untuk digunakan karena memiliki alur perhitungan yang tidak panjang sedangkan pada Metode Algoritma *Decision Tree* jika data diubah atau ditambah maka perhitungan akan memerlukan waktu yang lebih lama lagi. Persamaan dari teorema bayes adalah :

$$P(H|X) = \frac{P(X|H) \cdot (P|H)}{P(X)} \dots\dots\dots (2.1)$$

Keterangan :

X : Data kelas yang belum diketahui

H : Hipotesis data X adalah suatu kelas spesifik

$P(H|X)$  : Probabilitas hipotesis H berdasarkan kondisi X (*posteriori probability*)

$P(H)$ : Probabilitas hipotesis H (*prior probability*)

$P(X|H)$  : Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$  : Probabilitas X

Contoh perhitungan dalam naïve bayes adalah sebagai berikut

Adapun alur dari metode naïve bayes adalah sebagai berikut :

- a. Mulai
- b. Baca data training
  1. Hitung  $P(H)$  untuk setiap kelas
  2. Hitung  $P(X|H)$  untuk setiap kriteria dan setiap kelas
  3. Cari  $P(X|H)$  yang paling besar menjadi kesimpulan
- c. Tampilkan hasil prediksi

Contoh kasus adalah prediksi penjualan produk menggunakan Algoritma Naïve Bayes yang akan dilakukan pengerjaan secara manual dengan mengambil 5 jenis kategori yang berbeda sebagai

sampel. Berikut ini data penjualan fashion pada Couple Store sesuai dengan jurnal (Teguh Yulianto, 2019)

Nilai probabilitas class bisa dilihat pada tabel 2.1 di bawah ini

**Tabel 2.1 Nilai Probabilitas Class**

Jumlah class		Nilai Probabilitas	
Laris	16	Laris	$16/41=0.39$
Tidak laris	25	Tidak laris	$25/41=0.61$

Dataset penjualan produk memiliki 2 masalah, sehingga dapat diketahui probabilitas (P) masing-masing kelas dengan cara membagi nilai masalah, dengan jumlah total nilai frekuensi/ jumlah data pada kelas tersebut.

Nilai probabilitas merk bisa dilihat pada tabel 2.2 di bawah ini

**Tabel 2.2 Nilai Probabilitas Merk**

	Jumlah merk		Nilai Probabilitas	
	Laris	Tidak Laris	Laris	Tidak Laris
Family	6	6	$6/16=0.735$	$6/25=0.24$
Couple hoodie	2	4	$2/16=0.125$	$4/25=0.16$
Reglan	5	8	$5/16=0.313$	$8/25=0.32$
Couple	3	7	$3/16=0.187$	$7/25=0.28$

Tabel di atas merupakan nilai probabilitas merk yang jumlah laris atau tidak larisnya dibagi dengan jumlah keseluruhan data tiap masing-masing masalah

Nilai probabilitas harga bisa dilihat pada tabel 2.3 di bawah ini

**Tabel 2.3 Nilai Probabilitas Harga**

	Jumlah harga		Nilai Probabilitas	
	Laris	Tidak Laris	Laris	Tidak Laris
150000	16	25	5.7117	5.6880
170000	16	25	7.8737	6.9905
200000	16	25	8.7717	7.9963
300000	16	25	4.9239	6.1251

Tabel di atas merupakan nilai probabilitas harga yang jumlah laris atau tidak larisnya dibagi dengan jumlah keseluruhan data tiap masing-masing masalah

Nilai probabilitas bahan bisa dilihat pada tabel 2.4 di bawah ini

**Tabel 2.4 Nilai Probabilitas bahan**

	Jumlah bahan		Nilai Probabilitas	
	Laris	Tidak Laris	Laris	Tidak Laris
Cotton 24s	14	21	$14/16=0.875$	$21/25=0.84$
Polyester	2	4	$2/16=0.125$	$4/25=0.16$

Tabel di atas merupakan nilai probabilitas bahan yang jumlah laris atau tidak larisnya dibagi dengan jumlah keseluruhan data tiap masing-masing masalah

Nilai probabilitas warna bisa dilihat pada tabel 2.5 di bawah ini

**Tabel 2.5 Nilai Probabilitas warna**

	Jumlah warna		Nilai Probabilitas	
	Laris	Tidak Laris	Laris	Tidak Laris
marun	2	7	$2/16=0.125$	$7/25=0.28$
dongker	5	7	$5/16=0.313$	$7/25=0.28$
biru	0	1	$0/16=0$	$1/25=0.04$
putih	5	8	$5/16=0.3125$	$8/25=0.32$
Hitam	4	1	$4/16=0.25$	$1/25=0.04$

Tabel di atas merupakan nilai probabilitas warna yang jumlah laris atau tidak larisnya dibagi dengan jumlah keseluruhan data tiap masing-masing masalah

Nilai probabilitas ukuran bisa dilihat pada tabel 2.6 di bawah ini

**Tabel 2.6 Nilai Probabilitas ukuran**

	Jumlah ukuran		Nilai Probabilitas	
	Laris	Tidak Laris	Laris	Tidak Laris
S	0	0	$0/16=0$	$0/25=0$
M	3	7	$3/16=0.1875$	$7/25=0.28$
L	10	15	$10/16=0.625$	$15/25=0.06$
XL	3	3	$3/16=0.1875$	$3/25=0.12$
XXL	0	0	$0/16=0$	$0/25=0$

Tabel di atas merupakan nilai probabilitas ukuran yang jumlah laris atau tidak larisnya dibagi dengan jumlah keseluruhan data tiap masing-masing masalah

### 1.1 Menghitung jumlah laris dan tidak laris

$$P(\text{kelas}=\text{"laris"}) = 16/41=0.39$$

$$P(\text{kelas}=\text{"tidak laris"}) = 25/41=0.61$$

### 2.1 Jumlah kasus yang sama dengan kelas yang sama

#### a. Atribut Merk

$$P(\text{merk}=\text{MRK0001}|\text{laris}) = 6/16=0.375$$

$$P(\text{merk}=\text{MRK0001}|\text{tidak laris}) = 6/25=0.24$$

#### b. Atribut harga

$$\text{Miu laris} = 3060000/16=191250$$

$$\text{Gaus laris} = 47592.016$$

$$P(\text{harga} = 200000|\text{laris}) = 1/119319.499 \cdot \exp(-0.017) = 8.2595894457695\text{E-}6$$

$$\text{Mui tidak_laris} = 4790000/25=191600$$

$$\text{Gaus tidak_laris} = 51855.569$$

$$P(\text{harga} = 200000|\text{tidak_laris}) = 1/130008.792 \cdot \exp(-0.013) = 7.5924414021191\text{E-}6$$

#### c. Atribut Bahan

$$P(\text{bahan}=\text{cotton24s}|\text{laris}) = 14/16=0.875$$

$$P(\text{bahan}=\text{cotton24s}|\text{tidak_laris}) = 21/25=0.84$$

#### d. Atribut Warna

$$P(\text{warna}=\text{dongker}|\text{laris}) = 5/16=0.313$$

$$P(\text{warna}=\text{dongker}|\text{tidak_laris}) = 7/25=0.28$$

#### e. Atribut Ukuran

$$P(\text{ukuran}=\text{L}|\text{tidak_laris}) = 15/25=0.6$$

$$P(\text{ukuran}=\text{L}|\text{laris}) = 10/16=0.625$$

### 3.1 Kalikan semua hasil

$$P(\text{kelas}=\text{laris}(x)) = (0.375 \cdot 8.2595894457695\text{E-}6 \cdot 0.875 \cdot 0.313 \cdot 0.625 \cdot 0.39024990243962) = 2.063979475116\text{E-}7$$

$$P(\text{kelas=tidak\_laris}(x)) = (0.24 * 7.5924414021191E-6 * 0.84 * 0.28 * 0.6 * 0.60975609756096 = 1.5679687763539E-7$$

#### 4.1 Membandingkan hasil perkelas

Dari data testing merk =MRK0001, harga = 200000, bahan = cotton24s, warna = dongker, ukuran =L masuk dalam kelas laris dengan nilai probabilitas laris = 2.063979475116E-7 dan nilai probabilitas tidak laris = 1.5679687763539E-7