

# PREDIKSI CURAH HUJAN DI KOTA GORONTALO MENGUNAKAN ALGORITMA NAIVE BAYES

*by* Imelda Sintia Soleman T3116297

---

**Submission date:** 31-Jul-2020 07:21PM (UTC+0700)

**Submission ID:** 1364378717

**File name:** AH\_HUJAN\_DI\_KOTA\_GORONTALO\_MENGUNAKAN\_ALGORITMA\_NAIVE\_BAYES.pdf (2.79M)

**Word count:** 12084

**Character count:** 71573

**PREDIKSI CURAH HUJAN DI KOTA GORONTALO  
MENGUNAKAN ALGORITMA NAIVE BAYES**

Oleh

**IMELDA SINTIA SOLEMAN**

**T3116297**

**1  
SKRIPSI**

Untuk memenuhi salah satu syarat ujian  
guna memperoleh gelar sarjana



**PROGRAM SARJANA  
TEKNIK INFORMATIKA  
UNIVERSITAS ICHSAN GORONTALO  
GORONTALO  
2020**

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Iklim merupakan geografi yang sangat penting karena mempengaruhi aktivitas manusia di lingkungan, seperti transportasi, pertanian. Ada banyak iklim yang berbeda di dunia. Iklim adalah cuaca yang terjadi di area yang luas untuk waktu yang lama. Cuaca adalah udara di area tertentu dan relatif singkat. Cuaca yang ditandai siang dan malam dapat berubah dalam waktu singkat. Di antara faktor penting perubahan iklim Salah satunya adalah curah hujan, curah hujan adalah elemen iklim, curah hujan juga merupakan faktor langsung dalam perubahan iklim baik dan buruk, curah hujan adalah jumlah udara yang jatuh di atas permukaan datar menjadi milimeter [1].

Cuaca merupakan kondisi udara dalam kurun waktu tertentu serta pada daerah tertentu yang relative singkat dan pada jangka masa yang sempit. Formasi adalah kolaborasi antara elemen cuaca dan periode cuaca hanya dalam beberapa jam. Iklim adalah kondisi cuaca rata-rata selama periode waktu yang lama (minimal 30 tahun) selama survei dilakukan dan mencakup area yang luas [2].

Kota Gorontalo sama seperti kota atau wilayah di Secara umum, di Indonesia, dengan iklim tropis, intensitas curah hujan menjadi tidak stabil karena penguapan air yang sangat besar ke atmosfer. Intensitas curah hujan juga disebut sebagai curah hujan yang sangat deras atau ekstrim dan dapat menyebabkan banjir. Curah hujan ini terjadi ketika curah hujan mencapai lebih dari 100mm / hari (BMKG).

Intensitas curah hujan di kota Gorontalo berbeda-beda. Pada beberapa tempat terkadang hujan lebat dan dilain tempat sama sekali tidak terjadi hujan. Berdasarkan data yang diperoleh dari BMKG Stasiun Klimatologi Tilongkabila Gorontalo berikut ini ditampilkan data curah hujan pada beberapa wilayah di kota Gorontalo:

**Tabel 1.1.** Data Set prediksi curah hujan 2019

No	Bulan	suhu rata rata ( $^{\circ}C$ )	Kelembaban Udara (%)	Tekanan Udara (Mbs)	Penyinaran (%)	Kecepatan Angin (Knot)	Curah Hujan
1.	Jan	26,9	83	Rendah	Rendah	Tinggi	Hujan lebat
2.	Feb	27	83	Rendah	Rendah	Rendah	Hujan lebat
3.	Mar	27,5	81	Rendah	Sedang	Sedang	Hujan ringan
4.	Aprl	27	81	Sedang	Sedang	Sedang	Hujan ringan
5.	Mei	27,8	84	Tinggi	Rendah	Sedang	Hujan lebat
6.	Jun	27,5	83	Sedang	Rendah	Sedang	Hujan lebat
7.	Jul	27,5	80	Rendah	Sedang	Sedang	Hujan lebat
8.	Agt	28	76	Rendah	Sedang	Tinggi	Hujan ringan
9.	Sept	28,5	75	Rendah	Rendah	Tinggi	Tidak hujan
10.	Okt	28,9	74	Rendah	Tinggi	Tinggi	Hujan lebat
11.	Nov	28,4	80	Tinggi	Sedang	Sedang	Tidak hujan
12.	Des	28,6	78	Tinggi	Tinggi	Tinggi	Hujan lebat

(Sumber data: BMKG Stasiun Klimatologi Tilongkabila Gorontalo 2019)

Seperti halnya data tersebut diatas (diambil seampel Kecamatan Dumbo Raya), intensitas curah hujan berbeda-beda setiap bulannya, hal tersebut menurut penjelasan kepala BMKG Stasiun Klimatologi Tilongkabila Gorontalo diakibatkan oleh adanya perubahan suhu minimum, suhu maksimum, kelembaban rata-rata, penyinaran dan kecepatan angin. Seperti yang disebutkan sebelumnya adalah iklim tropis dan iklim yang menyebabkan penguapan air ke atmosfer sangat besar, mengakibatkan intensitas curah hujan. Itu menjadi tidak stabil.

Prediksi atau Peramalan curah hujan telah menjadi salah satu tantangan paling ilmiah dan teknologi, menantang dunia selama abad terakhir. Ini disebabkan oleh dua faktor utama yang digunakan untuk melakukan berbagai aktivitas manusia dan kemajuan komputasi. Untuk memprediksi jumlah curah hujan yang tepat maka dalam penelitian ini penulis mengusulkan penggunaan algoritma *Naïve Bayes* dan evaluasi model dengan melihat hasil klasifikasi menggunakan *confusion matriks*. Dipilihnya *Naïve Bayes*, disebabkan algoritma tersebut merupakan satu sistem diantara beragam sistem yang cocok dalam prediksi curah hujan, tanpa mengecilkan metode prediksi

yang lain. Data atau atribut inputan yang digunakan adalah penyinaran, suhu rata-rata, kecepatan angin, tekanan udara. Hasil atau output nantinya untuk menentukan prediksi curah hujan adalah tidak hujan, hujan ringan, hujan lebat.

Beberapa penelitian tentang prediksi menggunakan naïve bayes diantaranya yang dilakukan oleh Naisha Rahma Indraswari, 2018 [3] dengan judul “**Aplikasi Prediksi Usia Burung dengan Metode Naive Bayes**” Hasil penelitian ini adalah aplikasi yang dapat memprediksi usia kelahiran dengan nilai akurasi terapan tertinggi 78,69%, nilai presisi tertinggi 70,14% dan tingkat pemulihan tertinggi. Angka 63,64% Ada. Berikutnya penelitian oleh Syarli dkk [4] dengan judul Metode Naïve Bayes untuk Prediksi Kelulusan, 2016. Evaluasi hasil menunjukkan data yang diklasifikasikan dengan benar (mengoreksi instance yang diklasifikasikan) selaras pada pengelompokan alternatif pertama, alternatif kedua serta yang tidak berhasil dengan persentase 93,6288% dalam artian 338 data, tetapi kurang cocok dengan prediksi kelas (instance tidak diklasifikasikan dengan benar) yang seharusnya berupa kelompok dua atau opsi Lulus tetapi disertakan pada grup pertama sebesar 6,3712% yakni 23 data. Nilai Persentase akurasi menunjukkan keefektifan set data Penerimaan diterapkan pada metode Klasifikasi Naïve Bayes, yang mencapai 94%. Selaras dengan penjelasan diatas, dengan ini penulis mengangkat tema penelitian dengan Judul “**Prediksi Curah Hujan di Kota Gorontalo menggunakan Metode Naïve Bayes**”.

## **1.2. Identifikasi Masalah**

1. Perubahan cuaca dapat terjadi begitu cepat
2. Jumlah curah hujan setiap tahun di kota Gorontalo berbeda-beda

### 1.3. Rumusan Masalah

Berdasarkan identifikasi masalah diatas, maka permasalahannya dapat dirumuskan sebagai berikut:

1. Bagaimana ketepatan metode Naïve Bayes dalam <sup>5</sup> memprediksi curah Hujan di kota Gorontalo?
2. Seberapa besar tingkat akurasi metode Naïve Bayes dalam memprediksi curah hujan di Kota Gorontalo?

### 1.4. Tujuan Penelitian

Berdasarkan Rumusan permasalahan diatas, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk memperoleh model yang tepat dalam penggunaan metode Naïve Bayes <sup>5</sup> pada proses prediksi proses curah hujan di kota Gorontalo.
2. Untuk meningkatkan akurasi yang baik dalam metode Naïve Bayes <sup>5</sup> pada proses prediksi curah hujan di kota Gorontalo.

### 1.5. Manfaat Penelitian

Penelitian ini diharapkan mempunyai manfaat, yaitu:

1. Teoritis.

Penelitian ini diharapkan dapat memberikan masukan bagi pengembangan ilmu pengetahuan bidang kajian data mining tentang kemampuan metode Naïve Bayes dalam melakukan teknik prediksi.

2. PraktisiHasil

penelitian dapat digunakan sebagai salah satu alternatif atau masukan kepada instansi terkait guna mendukung dalam pengambilan keputusan dalam klasifikasi khususnya pada prediksi curah hujan di Kota Gorontalo.

## BAB II

### TINJAUAN PUSTAKA

#### 2.1. Tinjauan Studi

Prediksi tentang curah hujan menggunakan berbagai macam metode merupakan bidang penelitian yang sudah banyak dikembangkan saat ini.

Berikut penelitian yang menjadi referensi.

**Tabel 2.1.** Penelitian Tentang Prediksi Curah Hujan

Peneliti/Tahun	Judul	Keterangan
Diera Desmonda, Turzina & M Azhar Irwansyah, 2018 [5].	Prediksi Besaran Hujan Menggunakan Metode <i>Fuzzy Time Series</i> .	Hasil dari pengujian diperoleh nilai <i>MAPE</i> ( <i>Mean Average Percentage Error</i> ) bervariasi tergantung jumlah data dan jumlah interval yang digunakan. Nilai <i>MAPE</i> terbaik yang diperoleh adalah 0,151% pada penggunaan data curah hujan periode 2015 - 2017 dengan jumlah interval 401.
Satriogi Putramulyo & Siti Alaa, 2018 [6].	Prediksi Curah Hujan Bulanan di Kota Samarinda menggunakan Persamaan Regresi dengan Prediktor Data Suhu Udara dan Kelembapan Udara.	Hasil pengolahan data bahwa prediksi tahun 2016 curah hujan bulanan di kota Samarinda mengalami estimasi dan terbaik hasil prediksi terjadi pada bulan Oktober. Prediksi curah hujan bulanan dengan prediktor suhu udara dan kelembaban menghasilkan koefisien korelasi yang sangat kuat dan RMSE baik dimana $r = 0,93$ dan $RMSE = 46,26\%$ .
Yulia Fatmi, 2019. [7]	Prediksi Pembukaan Peluang Usaha di Kecamatan Medan Amplas menggunakan Naïve Bayes.	Pilihan algoritma Naïve Bayes dalam analisis ini adalah karena algoritma ini hanya membutuhkan sejumlah kecil data pelatihan untuk menentukan estimasi parameter yang diperlukan dalam proses klasifikasi. Dari hasil penelitian ini diharapkan dapat mengetahui prediksi peluang usaha bagi masyarakat Medan

## **2.2. Tinjauan Teori**

### **2.2.1. Definisi Curah Hujan**

Curah hujan adalah sedimen cair atau padat atau sedimen air yang berasal dari atmosfer. Untuk menentukan kemungkinan bencana air dan kemungkinan masalah dan bencana, perlu untuk "mengetahui" karakteristik curah hujan lokal. [1]. Curah hujan yang direkam dalam inci atau milimeter dapat diartikan sebagai ketinggian pengumpulan air pada permukaan yang datar, dengan asumsi bahwa itu tidak tenggelam, tidak mengalir, dan tidak menguap ke atmosfer [8]. Ketinggian curah hujan diasumsikan sama dengan daerah yang dicakup oleh alat pengukur hujan, tergantung pada homogenitas daerah tersebut. Intensitas air hujan adalah ketinggian atau kedalaman waktu air hujan. Sifat umum hujan adalah bahwa semakin pendek hujan cenderung meningkat kekuatannya, dan semakin lama periode kembalinya, semakin tinggi kekuatannya. Curah hujan dapat diukur dalam milimeter (mm). Gelas ukur adalah gelas ukur.

Curah hujan yang parah secara signifikan lebih tinggi atau lebih rendah dari kondisi stabil rata-rata. Secara umum, curah hujan ekstrem dapat dibagi menjadi curah hujan basah ekstrem yang menyebabkan banjir dan curah hujan ekstrem ekstrem yang menyebabkan kekeringan. Efek kekeringan. Menurut Supriatna dalam curah hujan BMKG, curah hujan dengan kekuatan > 100 mm / hari adalah parameter untuk hujan ekstrem. Hujan ekstrim peneliti juga mirip dengan kejadian hujan banjir pada 6-7 Januari dan 10-11 Februari 1996 di DAS Ciliwung [9].

**Tabel 2.2.** Data Set Prediksi Curah Hujan Tahun 2018

No.	Bulan	suhu rata rata ( $^{\circ}C$ )	Kelembaban Udara (%)	Tekanan Udara (Mbs)	Penyinaran (%)	Kecepatan Angin (Knot)	Curah Hujan
1.	Jan	27,9	81	Rendah	Tinggi	Tinggi	Hujan lebat
2.	Feb	28	83	Tinggi	Rendah	Rendah	Hujan lebat
3.	Mar	27,2	81	Rendah	Sedang	Sedang	Hujan lebat
4.	Aprl	27	80	Sedang	Sedang	Sedang	Hujan ringan
5.	Mei	27,9	84	Tinggi	Rendah	Tinggi	Hujan lebat
6.	Jun	27,5	83	Sedang	Rendah	Sedang	Hujan lebat
7.	Jul	27,4	80	Rendah	Sedang	Sedang	Hujan ringan
8.	Agt	27	79	Rendah	Tinggi	Tinggi	Hujan lebat
9.	Sept	28,3	75	Rendah	Rendah	Tinggi	Tidak hujan
10.	Okt	28,7	78	Rendah	Tinggi	Tinggi	Hujan ringan
11.	Nov	28,4	80	Tinggi	Sedang	Sedang	Hujan ringan
12.	Des	28,2	79	Tinggi	Tinggi	Tinggi	Hujan lebat

Sumber Data: BMKG Stasiun Klimatologi Tilongkabila Gorontalo 2019

**Tabel 2.3.** Data Set Prediksi Curah Hujan Tahun 2019

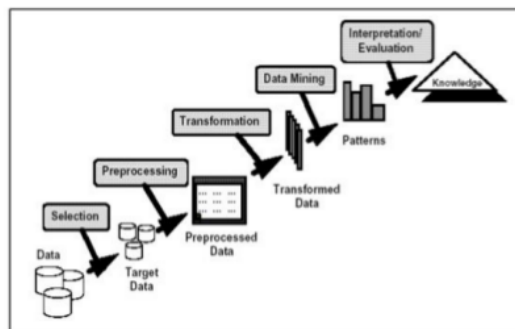
No	Bulan	suhu rata rata ( $^{\circ}\text{C}$ )	Kelembaban Udara (%)	Tekanan Udara (Mbs)	Penyinaran (%)	Kecepatan Angin ( <i>Knot</i> )	Curah Hujan
1.	Jan	26,9	83	Rendah	Rendah	Tinggi	Hujan lebat
2.	Feb	27	83	Rendah	Rendah	Rendah	Hujan lebat
3.	Mar	27,5	81	Rendah	Sedang	Sedang	Hujan ringan
4.	Aprl	27	81	Sedang	Sedang	Sedang	Hujan ringan
5.	Mei	27,8	84	Tinggi	Rendah	Sedang	Hujan lebat
6.	Jun	27,5	83	Sedang	Rendah	Sedang	Hujan lebat
7.	Jul	27,5	80	Rendah	Sedang	Sedang	Hujan lebat
8.	Agt	28	76	Rendah	Sedang	Tinggi	Hujan ringan
9.	Sept	28,5	75	Rendah	Rendah	Tinggi	Tidak hujan
10.	Okt	28,9	74	Rendah	Tinggi	Tinggi	Hujan lebat
11.	Nov	28,4	80	Tinggi	Sedang	Sedang	Tidak hujan
12.	Des	28,6	78	Tinggi	Tinggi	Tinggi	Hujan lebat

Sumber Data: BMKG Stasiun Klimatologi Tilongkabila Gorontalo 2019

### 2.2.2 Data Mining

Menurut Han dan Kamber (2011), penambangan data adalah proses menemukan pola dan pengetahuan yang menarik dari sejumlah besar data. Menurut Linoff dan Berry (2011), penambangan eData bertujuan untuk mencari dan menganalisis sejumlah besar data dan menemukan makna pola dan aturan. Menurut Connolly dan Begg (2010), Dan menurut Vercellis (2009), penambangan data adalah aktivitas yang menggambarkan proses analitis berulang dalam basis

data besar untuk mengekstrak informasi dan pengetahuan yang akurat dan berpotensi bermanfaat bagi para profesional pengetahuan yang terlibat dalam pengambilan keputusan dan penyelesaian masalah. Banyak orang menggunakan data mining sebagai istilah populer KDD. Data Knowledge Discovery (KDD) adalah keseluruhan proses pencarian dan identifikasi pola-pola sepele dalam data yang ditemukan sebagai pola yang valid, pola baru yang bermanfaat dan dapat dipahami [10].



**Gambar 2.1:** Proses *Knowledge Discovery in Database* (KDD)

(Sumber: Prasetyo, [11]).

"Menurut Han dan Kamber [11], penambangan data umumnya dapat dibagi menjadi dua kategori utama:

1. Predictive

*Predictive* Memprediksi adalah proses menemukan pola dalam data menggunakan beberapa variabel yang berbeda di masa depan. Salah satu teknik yang terlibat dalam penambangan prediktif adalah klasifikasi. Tujuan dari tugas prediksi adalah untuk memprediksi nilai properti tertentu berdasarkan nilai properti lain.

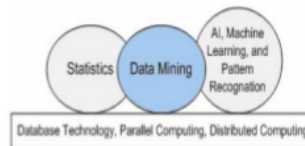
## 2.Descriptive

*Descriptive* Penambangan data adalah proses menemukan karakteristik penting data dalam database. Tujuan dari pekerjaan penjelas adalah untuk mengurangi pola (korelasi, tren, kelompok, wilayah, dan anomali) yang meringkas hubungan kunci dalam data.

Menurut Hoffer, Ramesh & Topi [12], tujuan dari penambangan data adalah:

1. *explanatory*, Ini untuk menggambarkan aktivitas atau kondisi pengamatan.
2. *confirmatory*, Yaitu, untuk mengkonfirmasi hipotesis yang ada.
3. *exploratory*, Ini untuk menganalisis data baru dalam hubungan yang aneh.

Tujuan dari data mining adalah untuk mengkategorikan pola-pola yang harus dicari oleh data mining. Secara umum, data mining dapat dikategorikan ke dalam dua kategori: teknis dan prediktif [13]. Ada juga pekerjaan dan keterampilan terkait; Gambar 2.2 menunjukkan hubungan antara data mining dan area lainnya.



**Gambar 2.2** Irisan Bidang Ilmu Data Mining

(Sumber: witten et al.,[14])

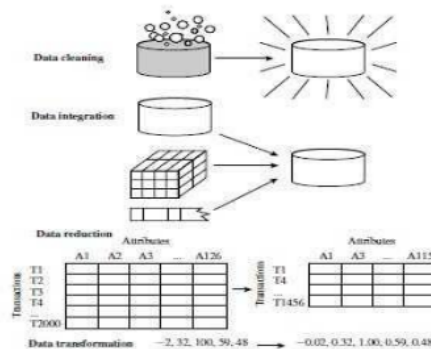
### 2.2.3 Proses Tahapan Data Mining

Menurut Han dan Kamber [11], tahapan data pra-pemrosesan dibagi menjadi:

#### 1. Data preprocessing:

gambaran umum Bagian ini memberikan tinjauan umum tentang preprocessing data. Jelaskan banyak faktor yang menentukan kualitas data di bagian Kualitas Data. *Major Tasks in Data Preprocessing*: Langkah-langkah utama yang terlibat dalam preprocessing data, yaitu pembersihan data, integrasi data, reduksi data

dan transformasi data. Pembersihan data mengisi nilai-nilai yang hilang, membersihkan data yang berisik, mengidentifikasi atau menghilangkan pencilan, dan menyelesaikan ketidakkonsistenan untuk membersihkan data. Langkah preprocessing yang berguna adalah menjalankan data dengan pembersihan data. Ini adalah format data yang sudah diproses



**Gambar 2.3:** Format data yang diproses

(Sumber: Han dan Kamber, [11]).

## 2. Data Cleaning

Pembersihan data (atau pembersihan data) berupaya untuk mengisi nilai yang hilang, menghaluskan data berisik, mengidentifikasi pencilan, dan memperbaiki perbedaan data.

- **Missing Values:** Banyak tuple yang tidak memiliki nilai tertulis ke properti.
- **Noisy Data:** Noise adalah kesalahan acak atau variasi dari variabel yang diukur
- **Data Cleaning as a Process:** Gunakan alat komersial untuk mendeteksi perbedaan metadata (domain, rentang, dependensi, distribusi), mendeteksi kelebihan, mendeteksi aturan keunikan, aturan kontinuitas, dan nol.

## 3. Data Integration

Konsolidasi data melibatkan penggabungan data dari berbagai basis data ke dalam satu basis data baru. Data yang diperlukan untuk penambahan data jarang diimpor dari banyak basis data atau file teks, bukan hanya satu basis data..

1

#### 4. Data Reduction

Data Menyusut data berguna untuk mengurangi representasi dari kumpulan data yang jauh lebih kecil dalam volume.

1

#### 5. Data Transformation and Data Discretization

Dalam transformasi data dan diskritisasi data, data dimodifikasi atau dikonsolidasikan untuk membuatnya lebih mudah untuk memahami pola yang ditemukan untuk membuat proses penambangan yang dihasilkan lebih efisien.

1

### 2.2.4 Teknik Data Mining

Teknologi data mining dibagi menjadi tiga kategori.

#### 1. Association Rule Mining

Menurut Olson dan Shi [14], Rule Mining Association digunakan untuk memprediksi pola, data mining untuk menemukan aturan yang mengaitkan kombinasi item atau objek tertentu dalam transaksi data dengan objek tertentu dalam transaksi data Teknologi. Menurut Han dan Kamber [11], asosiasi penambangan aturan terdiri dari serangkaian item yang sering terjadi. Asosiasi penambangan aturan adalah statistik antara item set A dan B untuk mengungkap aturan korelasi untuk menyampaikan korelasi item Dapat dianalisis.

1

#### 2. Classification

Menurut klasifikasi "Olson dan Shi [14], metode ini dimaksudkan untuk mempelajari berbagai fungsi memetakan setiap data yang dipilih ke salah satu kelompok kelas yang telah ditentukan". Menurut Han dan Kamber [11], "Klasifikasi adalah proses menemukan model atau fungsi yang menggambarkan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas suatu objek yang labelnya tidak diketahui. Metrik untuk mengukur kualitas pencarian teks.

### 2.2.2. Naïve Bayes

*Naive Bayes* adalah klasifikasi probabilitas sederhana yang menghitung probabilitas yang ditentukan dengan menambahkan kombinasi frekuensi dan nilai dari set data yang diberikan. Algoritma ini menggunakan teorema Bayes dan mengasumsikan semua properti independen atau ireversibel yang diberikan kepada variabel kelas yang nilainya [15]. Definisi lain adalah bahwa *naive Bayes* memprediksi peluang masa depan berdasarkan pengalaman masa lalu, diurutkan berdasarkan probabilitas dan metode statistik yang disajikan oleh mantan ilmuwan Inggris Thomas Bayes.

*Naive Bayes* didasarkan pada asumsi yang disederhanakan bahwa nilai properti independen secara kondisional dengan nilai output. Dengan kata lain, probabilitas mengamati bersama ketika mempertimbangkan nilai output adalah produk dari probabilitas individu [17]. Keuntungan menggunakan *Naive Bayes* adalah bahwa metode ini hanya membutuhkan sejumlah kecil data pelatihan untuk menentukan estimasi parameter yang diperlukan untuk proses klasifikasi. *Naive Bayes* bekerja jauh lebih baik dari yang diharapkan dalam sebagian besar situasi dunia nyata.

### 2.2.3. Persamaan Naïve Bayes

Bentuk umum model dari persamaan teorema bayes [16]:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (2.1)$$

Dimana:

$X$  : Data class yang belum di ketahui

$H$  : Hipotesis data adalah kelas spesifik

$P(H|X)$  : Probabilitas Probabilitas hipotesis  $H$  didasarkan pada kondisi  $X$  (posteriori probabilitas)

$P(X|H)$  : "Probabilitas  $X$  didasarkan pada hipotesis  $H$ "

$P(H)$  : Probabilitas hipotesis  $H$  (probabilitas kamus)

$P(X)$  : Probabilitas  $X$

Untuk menjelaskan metode Naive Bayes, Anda perlu mengetahui proses klasifikasi. Banyak petunjuk diperlukan untuk menentukan kelas yang tepat untuk sampel yang dianalisis. Jadi metode Naive Bayes di atas disesuaikan sebagai berikut:

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \quad (2.2)$$

Dimana Variabel  $C$  merepresentasikan kelas, sementara variabel  $F_1 \dots F_n$  Mencirikan pedoman yang diperlukan untuk klasifikasi. Rumus kemudian menggambarkan probabilitas memasukkan sampel karakteristik tertentu di kelas  $C$  (Posterior) menggambarkan probabilitas bahwa karakteristik sampel akan muncul di kelas  $C$  (juga dikenal sebagai probabilitas penggunaan) dalam probabilitas terjadinya kelas  $C$  (sering disebut kamus). Jika properti sampel cenderung muncul Seluruh Dunia (juga dikenal sebagai bukti). Jadi rumus di atas juga dapat ditulis hanya sebagai:

$$\text{Posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}} \quad (2.3)$$

Nilai-nilai bukti selalu ditetapkan untuk setiap kelas dalam satu sampel. Nilai posterior kemudian dibandingkan dengan nilai-nilai posteriors kelas lain untuk menentukan kelas mana sampel akan diklasifikasikan ke dalam. Penjelasan lebih lanjut dari formula Bayes dilakukan dengan menjelaskan:

$(C|F_1, \dots, F_n)$  menggunakan aturan perkalian sebagai berikut:

$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C)P(F_1, \dots, F_n|C) \\ &= P(C)P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3, \dots, F_n|C, F_1, F_2) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2)P(F_4, \dots, F_n|C, F_1, F_2, F_3) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2) \dots P(F_n|C, F_1, F_2, F_3, \dots, F_{n-1}) \end{aligned} \quad (2.4)$$

dapat melihat bahwa hasil dari kecanggihan menghasilkan kondisi faktor yang semakin kompleks yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisis satu per satu. Akibatnya, sulit untuk menghitung. Ini mengarah pada

asumsi independen yang sangat tinggi (naif). Di sinilah digunakan, dan setiap instruksi ( $F_1, F_2 \dots F_n$ ) tidak tergantung satu sama lain.

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_i)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i) \quad (2.5)$$

Untuk  $i \neq j$ , sehingga

$$P(F_i|C, F_j) = P(F_i|C)$$

Persamaan di atas adalah model teorema Naive Bayes yang akan digunakan dalam proses klasifikasi. Untuk klasifikasi menggunakan data kontinu, rumus kepadatan Gaussian digunakan.

$$P(X_i = x_i|Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}} \quad (2.6)$$

15

Dimana:

$P$  : Peluang

$X_i$  : Atribut ke  $I$

$x_i$  : Nilai atribut ke  $I$

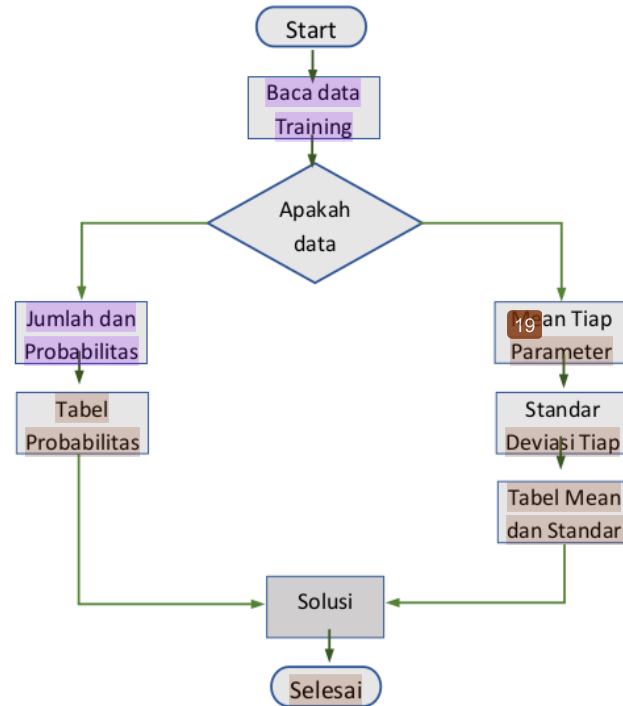
$Y$  : Kelas yang dicari

$y_i$  : Sub kelas  $Y$  yang dicari

$\mu$  : mean, menyatakan rata-rata dari seluruh atribut

$\sigma$  : Deviasi standar, menyatakan varian dari seluruh atribut

Alur dari metode Naïve Bayes sebagai berikut:



**Gambar 2.4.** Alur Metode Naïve Bayes

Sumber: [18]

Gambar di atas adalah sebagai berikut.

1. Baca data pelatihan
2. Hitung angka dan probabilitas jika data numeriknya adalah:
  - a. Temukan mean dan simpangan baku dari setiap parameter yang merupakan data numerik. Persamaan yang digunakan untuk menghitung nilai rata-rata (rata-rata) dari penghitungan adalah:

$$\mu = \frac{\sum_{i=1}^n x_i}{n}$$

Atau

(2.8)

$$\mu = \frac{x_1 + x_2 + x_3 + \cdots x_n}{n}$$

Dimana:

20

$\mu$  : rata-rata hitung (mean)

$x_1$  : nilai sampel ke  $i$

$n$  : jumlah sampel

Sedangkan untuk persamaan menghitung nilai simpangan baku (standar deviasi) adalah sebagai berikut:

(2.9)

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}}$$

Dimana:

$\sigma$  : standar deviasi

$x_i$  : nilai  $x$  ke  $i$

$\mu$  : rata-rata hitung

$n$  : jumlah sampel

b. Temukan nilai probabilitas dengan membagi jumlah data yang sesuai dalam kategori yang sama dengan jumlah data dalam kategori itu.

1. Dapatkan nilai dari tabel mean, standar deviasi dan probabilitas.
2. Lalu solusinya dibuat .

### 2.2.7. Penerapan Metode Naïve Bayes

Penelitian Naisha Rahma Indraswari [3], Melakukan Studi pada 2018. Penelitian ini bertujuan untuk memprediksi usia kelahiran menjadi tiga kategori berdasarkan faktor yang dialami oleh wanita hamil menggunakan aplikasi prediksi. Dalam penelitian ini, kategori usia saat lahir meliputi prematur, durasi normal atau penuh, dan pasca-tanggal.

penyelesaian. Berikut ini adalah perjanjian yang menerapkan Naif Bayes. Data yang diperoleh dalam penelitian ini dibagi menjadi dua bagian: data pelatihan dan data uji. Tabel 2.3 adalah data pelatihan yang digunakan untuk mendapatkan model, dan Tabel 2.4 adalah data uji untuk menguji model perhitungan data pelatihan.

**Tabel 2.3** Data Training

Usia	Tekanan Darah	Jml Bayi	Riwayat Persalinan	Riwayat Abortus	Nutrisi	Penyakit Lain	Masalah Saat Hamil	Usia Kelahiran
Lebih	Tinggi	1	Normal	Tidak	Normal	Tdk Ada	PEB	Posdate
Kurang	Normal	1	Normal	Tidak	Normal	Tdk Ada	Tdk Ada	Normal
Lebih	Normal	1	Prematur	Ya	Normal	Anemia	Tdk Ada	Prematur
Cukup	Tinggi	1	Anak Pertama	Tidak	Normal	Anemia	PEB	Postdate
Cukup	Normal	1	Prematur	Tidak	Normal	Tdk Ada	Tdk Ada	Normal
Cukup	Tinggi	1	Anak Pertama	Tidak	Normal	Hipertensi	PEB	Prematur
Lebih	Normal	1	Normal	Tidak	Normal	Tdk Ada	Tdk Ada	Normal
Lebih	Tinggi	1	Prematur	Ya	Normal	Asma	PEB	Prematur
Lebih	Normal	1	Anak Pertama	Tidak	Kurang	Asma	Tdk Ada	Prematur
Cukup	Normal	2	Normal	Tidak	Normal	Tdk Ada	Tdk Ada	Prematur

**Tabel 2.4** Data Testing

Usia	Tekanan Darah	Jml Bayi	Riwayat Persalinan	Riwayat Abortus	Nutrisi	Penyakit Lain	Masalah Saat Hamil	Usia Kelahiran
3								
Cukup	Tinggi	1	Anak Pertama	Tidak	Normal	Hipertensi	PEB	?

Dengan menggunakan hasil perhitungan pada Tabel 2.3, kita dapat menentukan kelas usia di Tabel Kelahiran 2.4 sebagai berikut: "

1. Menghitung Jumlah Probabilitas Variabel Y

$$P(Y = \text{Premature}) = \frac{5}{10} = 0,5$$

$$P(Y = \text{Normal}) = \frac{3}{10} = 0,3$$

$$P(Y = \text{Postdate}) = \frac{2}{10} = 0,2$$

2. Menghitung Probabilitas Variabel X terhadap variabel Y

$$P(X_1 = \text{Cukup} | Y = \text{Premature}) = \frac{2}{5} \quad P(X_5 = \text{Tidak} | Y = \text{Premature}) = \frac{3}{5}$$

$$P(X_1 = \text{Cukup} | Y = \text{Normal}) = \frac{1}{3} \quad P(X_5 = \text{Tidak} | Y = \text{Normal}) = \frac{3}{3}$$

$$P(X_1 = \text{Cukup} | Y = \text{Postdate}) = \frac{1}{2} \quad P(X_5 = \text{Tidak} | Y = \text{Postdate}) = \frac{2}{2}$$

$$P(X_2 = \text{Tinggi} | Y = \text{Premature}) = \frac{2}{5} \quad P(X_6 = \text{Normal} | Y = \text{Premature}) =$$

$$P(X_2 = \text{Tinggi} | Y = \text{Normal}) = \frac{0}{3} \quad \frac{4}{5}$$

$$P(X_2 = \text{Tinggi} | Y = \text{Postdate}) = \frac{2}{2} \quad P(X_6 = \text{Normal} | Y = \text{Normal}) = \frac{3}{3}$$

$$P(X_3 = 1 | Y = \text{Premature}) = \frac{4}{5} \quad P(X_6 = \text{Normal} | Y = \text{Postdate}) = \frac{2}{2}$$

$$P(X_3 = 1 | Y = \text{Normal}) = \frac{3}{3} \quad P(X_7 = \text{Hipertensi} | Y =$$

$$P(X_3 = 1 | Y = \text{Postdate}) = \frac{2}{2} \quad \text{Premature}) = \frac{1}{5}$$

$$P(X_7 = \text{Hipertensi} | Y = \text{Normal})$$

$$P(X_4 = \text{Anak 1} \mid Y = \text{Premature}) = \frac{2}{5}$$

$$P(X_4 = \text{Anak 1} \mid Y = \text{Normal}) = \frac{0}{3}$$

$$P(X_4 = \text{Anak 1} \mid Y = \text{Postdate}) = \frac{1}{2}$$

### 3. Membandingkan Hasil Probabilitas Tiap Kelas

$$P(X_1 = \text{Cukup}, X_2 = \text{Tinggi}, X_3 = 1, X_4 = \text{Anak 1}, X_5 = \text{Tidak}, X_6 = \text{Normal}, x_7 = \text{Hipertensi}, X_8 = \text{PEB} \mid Y = \text{Premature})$$

$$= \frac{2}{5} \times \frac{2}{5} \times \frac{4}{5} \times \frac{2}{5} \times \frac{3}{5} \times \frac{4}{5} \times \frac{1}{5} \times \frac{1}{5} = 0,00073728$$

$$P(X_1 = \text{Cukup}, X_2 = \text{Tinggi}, X_3 = 1, X_4 = \text{Anak 1}, X_5 = \text{Tidak}, X_6 = \text{Normal}, x_7 = \text{Hipertensi}, X_8 = \text{PEB} \mid Y = \text{Normal})$$

$$= \frac{1}{3} \times \frac{0}{3} \times \frac{3}{3} \times \frac{0}{3} \times \frac{3}{3} \times \frac{3}{3} \times \frac{0}{3} \times \frac{0}{3} = 0$$

$$P(X_1 = \text{Cukup}, X_2 = \text{Tinggi}, X_3 = 1, X_4 = \text{Anak 1}, X_5 = \text{Tidak}, X_6 = \text{Normal}, x_7 = \text{Hipertensi}, X_8 = \text{PEB} \mid Y = \text{Postdate})$$

$$= \frac{1}{2} \times \frac{2}{2} \times \frac{2}{2} \times \frac{1}{2} \times \frac{2}{2} \times \frac{2}{2} \times \frac{0}{2} \times \frac{1}{2} = 0$$

Probabilitas terbesar dalam perhitungan di atas dikenal sebagai P (Y = Premature), sehingga Anda dapat menyimpulkan bahwa data dari data uji berada di kelas Premature.

### 2.2.8 Analisis Hasil Akurasi Prediksi

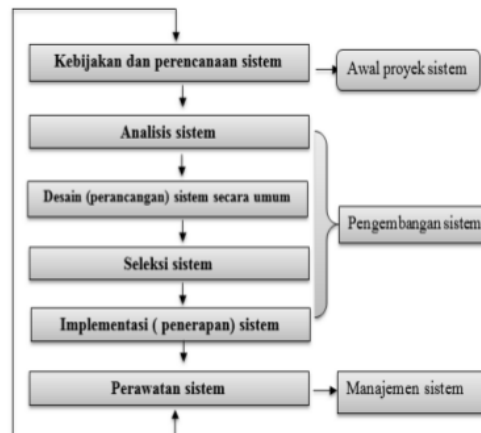
Pada tahap ini, penulis membagi total 561 total data <sup>3</sup> menjadi data pelatihan dan data uji secara berurutan. Ini bertujuan untuk menganalisis nilai presisi, daya ingat dan akurasi dalam menghitung algoritma Naive Bayes dalam aplikasi. Deskripsi tes presisi, daya ingat, dan akurasi dapat dilihat seperti yang ditunjukkan pada tabel berikut.

**Tabel 2.5.** Hasil Pengujian

Jumlah Data Training	Jumlah Data Testing	Precision	Recall	Accuracy
50	511	68.53%	52.94%	64.38%
100	461	66.50%	60.27%	65.94%
150	411	66.85%	60.40%	66.67%
200	361	70.14%	59.76%	69.25%
250	311	69.17%	56.85%	67.84%
300	261	66.34%	60.36%	70.11%
350	211	66.67%	57.78%	69.67%
400	161	60.38%	48.49%	65.84%
450	111	57.14%	43.24%	70.27%
500	61	43.75%	63.64%	78.69%

### 2.2.9 Siklus Hidup Pengembangan Sistem

Menurut Sutabri Tata [19], suatu bentuk yang digunakan untuk menggambarkan tahapan utama dan langkah-langkah pada tahapan tersebut dalam proses pengembangan sistem.



**Gambar 2.5:** Siklus pengembangan hidup

(Sumber: Sutabri Tata. [19])

### 2.2.10 Analisis Sistem

Analisis sistem (*system Analisa*) dapat didefinisikan sebagai penguraian sistem informasi yang lengkap menjadi bagian-bagian komponen untuk

mengidentifikasi dan mengevaluasi masalah, hambatan yang muncul, dan kebutuhan yang diantisipasi untuk perbaikan yang akan diusulkan.

Analisis sistem adalah pakar yang meninjau masalah dan kebutuhan organisasi untuk menentukan bagaimana orang, data, proses, dan teknologi informasi dapat mencapai kemajuan terbaik untuk bisnis.

Impak teknologi objek sangat berarti dalam dunia analisis dan desain sistem. Sebelum ada teknologi objek, kebanyakan bahasa pemrograman didasarkan pada apa yang disebut metode yang terstruktur (*structured method*). Contohnya COBOL bahasa yang domain 0, C, Fortran, Pascal, dan PL/i. Maka, metode analisis dan desain berorientasi objek telah muncul sebagai pendekatan terpilih untuk membangun kebanyakan sistem informasi saat ini.

Selain analisis sistem formal dan keterampilan desain, <sup>2</sup> analis harus mengembangkan atau memiliki keterampilan, pengetahuan, dan karakteristik lain untuk menyelesaikan pekerjaan mereka. Ini termasuk:

1. Pengalaman dan keahlian dalam pemrograman komputer Sulit membayangkan bagaimana seorang analis sistem dapat dengan benar mempersiapkan spesifikasi teknis dan bisnis untuk programmer yang tidak memiliki pengalaman pemrograman. Sebagian besar analis sistem perlu menguasai satu atau lebih bahasa pemrograman tingkat lanjut.

2. Pengetahuan umum tentang proses dan teknologi bisnis.

Analisis sistem perlu berkomunikasi dengan pakar bisnis untuk memahami masalah dan kebutuhan. Bagi analis, setidaknya sebagian dari pengetahuan ini berasal dari pengalaman. Pada saat yang sama, analis yang mengilhami harus memanfaatkan semua peluang bisnis untuk menyelesaikan mata kuliah teori bisnis inti Fase analisis sistem mencakup studi kelayakan untuk analisis permintaan.

Studi Kelayakan.

<sup>1</sup> Tugas-tugas yang tercakup dalam studi kelayakan meliputi:

1. Identifikasi masalah dan peluang dalam sistem.
2. Menetapkan tujuan sistem umum yang baru.

3. Identifikasi pengguna sistem.
4. Tetapkan ruang lingkup sistem.

Selain itu, selama dalam tahapan studi kelayakan sistem analisis juga melakukan tugas-tugas sebagai berikut:

1. Usulkan perangkat lunak dan perangkat keras untuk sistem baru.
2. Lakukan analisis untuk membuat atau membeli aplikasi.
3. Melakukan analisis biaya / manfaat.
4. Penilaian risiko proyek.

Studi kelayakan diukur dengan mempertimbangkan aspek teknologi, ekonomi, faktor organisasi, dan batasan hukum, etika, dan lainnya.[20].

a. Analisis kebutuhan.

Analisis kebutuhan dilakukan untuk menghasilkan spesifikasi kebutuhan (disebut juga spesifikasi fungsional). Spesifikasi kebutuhan adalah spesifikasi yang rinci tentang hal-hal yang akan dilakukan sistem ketika diimplementasikan. Spesifikasi ini sekaligus dipakai untuk membuat kesepakatan antara pengembang sistem, pemakai yang kelak akan menggunakan sistem, manajemen, dan mitra kerja yang lain (misalnya auditor internal). ada langkah-langkah dasar yang harus dilakukan untuk memindai sistem, yaitu sebagai berikut

1. Identifikasi, yang merupakan pemecahan masalah .
2. Memahami, yaitu, memahami cara kerja sistem yang ada
3. Menganalisis, yang memindai sistem tanpa laporan.
4. Laporan, yaitu membuat laporan tentang hasil analisis.

Tujuan utama dari pelaporan hasil dilakukan;

- a. Laporkan bahwa analisis telah selesai.
- b. Kesalahpahaman tentang apa yang ditemukan dan dianalisis oleh para analis sistem, tetapi tidak konsisten dengan manajemen

### 2.2.11 Desain Sistem

Setelah fase analisis sistem selesai, analisis sistem memiliki gambaran yang jelas tentang apa yang perlu dilakukan. Sudah waktunya untuk berpikir tentang bagaimana analisis sistem akan mengkonfigurasi sistem. Ini disebut desain sistem."

Desain sistem adalah spesifikasi atau pedoman solusi teknis dan berbasis komputer untuk kebutuhan bisnis yang diidentifikasi dalam analisis sistem.

- a. Untuk memenuhi kebutuhan pengguna sistem.
- b. Memberikan gambar yang jelas dan desain lengkap untuk programmer komputer dan pakar teknis lainnya.

### 2.2.12 Desain Sistem Secara Umum

Secara umum, tujuan perancangan sistem adalah untuk memberi pengguna gambaran umum tentang sistem baru. Ini adalah persiapan desain sistem yang terperinci. Untuk mengidentifikasi komponen sistem informasi yang akan dirancang secara rinci oleh pemrograman komputer dan insinyur lainnya, desain biasanya dilakukan melalui analisis sistem. Pada tahap ini, komponen sistem informasi dirancang untuk dikirimkan kepada pengguna. Komponen sistem informasi yang dirancang adalah model, input-output, database, teknologi, dan kontrol.

### 2.2.13 Desain sistem terinci

1. Tata letak keluaran terperinci

Desain output terperinci dimaksudkan untuk mengetahui bagaimana dan bagaimana output dari sistem baru akan terlihat. Desain keluaran terperinci dibagi menjadi dua jenis: desain laporan jenis kertas dan desain dialog layar terminal.

- a. Merancang hasil dalam bentuk laporan. Desain ini dimaksudkan untuk menghasilkan hasil dalam bentuk laporan kertas. Format laporan yang paling populer adalah tabel dan grafik atau tabel rancang output dalam bentuk dialog layar terminal.

b. Layout <sup>1</sup> ini adalah desain layout percakapan antara pengguna komputer dan sistem (pengguna). Percakapan ini dapat terdiri dari proses memasukkan data ke dalam sistem dan menampilkan informasi keluaran kepada pengguna, atau keduanya. Beberapa strategi untuk membuat layar dialog terminal:

1. Dialog tanya / jawab.

2. Menu.

Menu ini banyak digunakan karena merupakan jalur pengguna yang mudah dimengerti dan digunakan. Menu ini berisi berbagai alternatif atau opsi atau opsi yang disajikan kepada pengguna. Jika fitur dikelompokkan, opsi menu lebih baik.

2. Tata letak terperinci.

<sup>1</sup> Input adalah awal dari proses informasi. Bahan baku untuk informasi adalah data yang dihasilkan oleh transaksi yang dilakukan oleh organisasi. Data yang dihasilkan oleh transaksi dimasukkan ke dalam sistem informasi. Hasil dari sistem informasi tidak dapat dipisahkan dari data yang dimasukkan. Detail mulai dari tata letak default dokumen tata letak tidak dirancang dengan benar, sehingga kemungkinan item yang direkam lebih sedikit.

3. Desain detail dari database.

Basis data (basis data) adalah kumpulan data yang saling berhubungan satu sama lain dan disimpan di toko di luar komputer Anda dan menggunakan perangkat lunak khusus untuk manipulasi. <sup>2</sup> Database adalah salah satu komponen penting dalam sistem informasi. Ini karena sistem informasi berfungsi sebagian dengan memberikan informasi kepada pengguna. Menerapkan basis data ke sistem informasi disebut sistem basis data.

<sup>1</sup> Perspektifnya berbeda. Anda dapat melihat bagian kredit sebagai data penjualan, bagian karyawan sebagai data karyawan, dan bagian gudang data sebagai data inventaris. Semua terintegrasi ke dalam data umum.

4. Desain teknologi

Fase desain dibagi menjadi dua, desain teknis umum. Dalam fase ini, kami memutuskan teknologi mana yang akan kami gunakan untuk menerima input,

menjalankan model, menyimpan dan mengakses data, menghasilkan dan mengirimkan output, dan mengendalikan seluruh sistem.

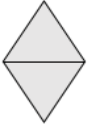

a) Perangkat keras (hardware) yang terdiri dari perangkat input, perangkat pemrosesan, perangkat output dan tangki eksternal.



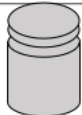



b) Sistem dapat berfungsi dengan baik, termasuk perangkat lunak sistem operasi (sistem operasi), perangkat lunak bahasa (perangkat lunak bahasa) dan perangkat lunak (perangkat lunak aplikasi), sistem perangkat lunak yang terdiri dari personel (brainware) (mis. Operator komputer, programmer, pakar telekomunikasi, analis) Desain teknis diperlukan selama fase implementasi dan pengujian untuk menunjukkan keberadaan.

#### 5. Fase desain

Fase desain dibagi menjadi dua kategori: desain umum dan detail model. Fase desain umum dari model ini adalah desain sistem fisik dan logis. Desain aktual dapat digambarkan sebagai bagian aliran sistem di bagian aliran dokumen, desain secara logis dijelaskan oleh diagram aliran data (DAD), dan model ditentukan secara rinci selama fase desain model terperinci. Urutan langkah-langkah dalam proses ini diwakili oleh program komputer. Bagian Aliran Sistem adalah grafik yang menunjukkan alur kerja sistem secara keseluruhan. Diagram alir sistem pada gambar dengan simbol berikut

**Tabel 2.6:** Bagan Alir Sistem

NO	NAMA SIMBOL	SIMBOL	KETERANGAN
1	Simbol Pengurutan Offline		menunjukkan operasi yang dilakukan di luar proses operasi komputer
2	Simbol Pita Magnetik		Menunjukkan input dan output menggunakan pita magnetic.

3	Simbol Kegiatan Manual		Menunjukkan pekerjaan manual
4	Simbol Diskette		Menunjukkan input dan output menggunakan diskette
5	Simbol Drum Magnetik		Menunjukkan input dan output menggunakan drum magnetik
6	Simbol Pita Kertas Berlubang		Menunjukkan input dan output menggunakan pita kertas berlubang
7	Simbol Keyboard		Menunjukkan input yang menggunakan on-line keyboard.
8	Simbol Display		Menunjukkan output yang ditampilkan di monitor

Sumber : Jogyanto, [21].

Untuk membuatnya lebih mudah untuk menggambarkan sistem yang ada atau baru yang akan dikembangkan secara logis, terlepas dari lingkungan fisik di mana data mengalir atau lingkungan fisik di mana data akan disimpan.

Data Flow Diagram (DAD) atau Data Flow Diagram (DFD). "Dalam menggambarkan sistem di mana pembentukan simbol perlu dilakukan, simbol berikut ini sering digunakan dalam DAD:

#### 1. External entity (kesatuan luar) atau boundary (batas sistem).

Setiap sistem harus memiliki batas sistem (boundary) yang memisahkannya dari lingkungan eksternal. Sistem menerima input dan menghasilkan output di lingkungan eksternal. "Entitas eksternal" adalah unit lingkungan eksternal sistem yang dapat berupa orang, organisasi, atau sistem lain di lingkungan eksternal, dan dapat menerima dan mengeluarkan input dari sistem [21].



**Gambar 2.6:** Notasi kesatuan luar di DAD

2. *Data flow* (arus data).

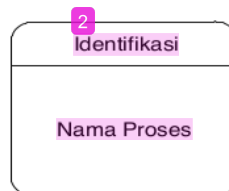
Aliran data ini menunjukkan aliran atau aliran data yang dapat dimasukkan ke dalam sistem atau menunjukkan hasil dari suatu proses sistem [21]. "



**Gambar 2.7:** Nama Arus Data di DAD

3. *Process* (proses).

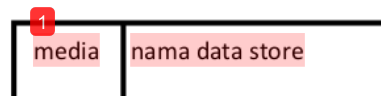
Suatu proses adalah kegiatan atau tugas yang dilakukan oleh seseorang, mesin, atau komputer dari hasil aliran data yang memasuki proses untuk membuat aliran data yang keluar dari proses [21].



**Gambar 2.8:** Notasi Proses di DAD

4. *Data store* (penyimpanan data).

Penyimpanan data DFD dapat dilambangkan dengan sepasang garis horizontal paralel di satu ujung. [21]



**Gambar 2.9:** Notasi penyimpanan data di DAD

#### **2.2.14 Pengujian**

Pada pendekatan berorientasi objek, pengujian merupakan suatu persoalan yang lebih kompleks dibanding dengan pendekatan konvensional, karena keberadaan pewarisan, polymorphism, dan pengkapsulan pada pengembangan sistem berorientasi objek menimbulkan suatu persoalan yang baru untuk perancangan kasus pengujian dan analisis hasil.

Hariyanto [22] mengungkapkan bahwa: fitur-fitur berikut berpengaruh dalam teknik-teknik pengujian yang perlu dilakukan:

- Pengkapsulan (*encapsulation*)
- Penyusunan objek-objek (*object composition*)
- Pewarisan (*inheritance*)
- Interaksi (*interaction*)
- *Polymorphism*
- Pengikatan dinamis (*dynamic binding*)
- Guna ulang (*reuse*)
- *Genericity* dan kelas abstrak”

Dari kompleksnya fitur –fitur yang mempengaruhi dalam pengujian sistem berorientasi objek maka strategi pengujian dilakukan pada:

1. Pengujian unit, dimana pengujian unit dilakukan hingga beberapa level dengan alasan adanya konsep pewarisan. Pengujian unit ini bertujuan untuk menjamin setiap unit memenuhi spesifikasi. Kelas-kelas merupakan sasaran pengujian unit.
2. Pengujian integrasi, pengujian ini dilakukan untuk memverifikasi implementasi dari satu use case yang telah bekerja seperti yang diharapkan. Pengujian validitas, pengujian ini dilakukan untuk menjamin fungsi-fungsi sistem/aplikasi telah dilakukan secara benar, pengujian dieksekusi ketika satu sistem (subsistem) yang lengkap telah dirakit. Pengujian validasi ini meliputi rincian-rincian objek yang tidak tampak, fokus pada masukan dan keluaran yang tampak oleh pemakai.

### 2.2.15 Implementasi Sistem

Tahapan implementasi merupakan tahap dimana dilakukan transformasi/penerjemahan dari bahasa modeling ke suatu bahasa pemrograman. hal ini merupakan tugas dari pemrogram, pada pengembangan sistem/perangkat lunak berorientasi objek penerjemahan dari setiap diagram-diagram DFD yang telah di rancang pada tahap analisis dan desain harus diterjemahkan ke dalam bahasa pemrograman sama persis dengan diagram-diagram yang ada guna menghindari terjadinya perubahan fungsi/tujuan dari pengembangan sistem/perangkat lunak

### 2.2.16 White Box Testing

*White Box* adalah metode desain kasus uji untuk mendapatkan kasus uji menggunakan struktur kontrol desain prosedural. Sistem yang menggunakan sistem analisis *White Box* yaitu

- a) Memastikan semua jalur independen dalam modul bekerja setidaknya satu kali.
- b) Buat semua keputusan logis
- c) Lakukan semua loop sesuai batas
- d) Bekerja dengan semua struktur data internal untuk memastikan validitas

Untuk melakukan proses uji kasus uji, konversi diagram alir ke notasi diagram alir (flow control) dilakukan terlebih dahulu.

Ada beberapa cara untuk membuat diagram alur:

2

1. Node adalah lingkaran diagram alur yang menggambarkan satu atau lebih perintah prosedural.

2. Tepi harus memiliki simpul target dengan panah yang menunjukkan aliran kontrol setiap simpul.

3. Wilayah adalah area yang dikelilingi oleh simpul dan tepi, dan area di luar grafik alir<sup>2</sup> juga harus dihitung.

4. Sebuah node conditioner adalah suatu kondisi yang ada pada sebuah node dan memiliki karakteristik dua atau lebih tepi yang berbeda.

5. Complex Complexity" adalah indikator perangkat lunak yang dapat mengukur kompleksitas logis suatu program secara kuantitatif dan dapat digunakan untuk menemukan jumlah jalur dalam diagram alur.

6. alur independen adalah jalur yang melalui atau melalui setidaknya proses perintah baru atau program dengan kondisi baru Rumus untuk menghitung jumlah rute independen dalam diagram alur adalah:

1. Jumlah *region flowgraph* mempunyai hubungan dengan *Cyclomatic Complexity (CC)*.

2.  $V(G)$  untuk *flowgraph* dapat dihitung dengan rumus :

a)  $V(G) = E - N + 2$

Dimana :

E = Jumlah *edge* pada *flowgraph*

N = Jumlah *node* pada *flowgraph*

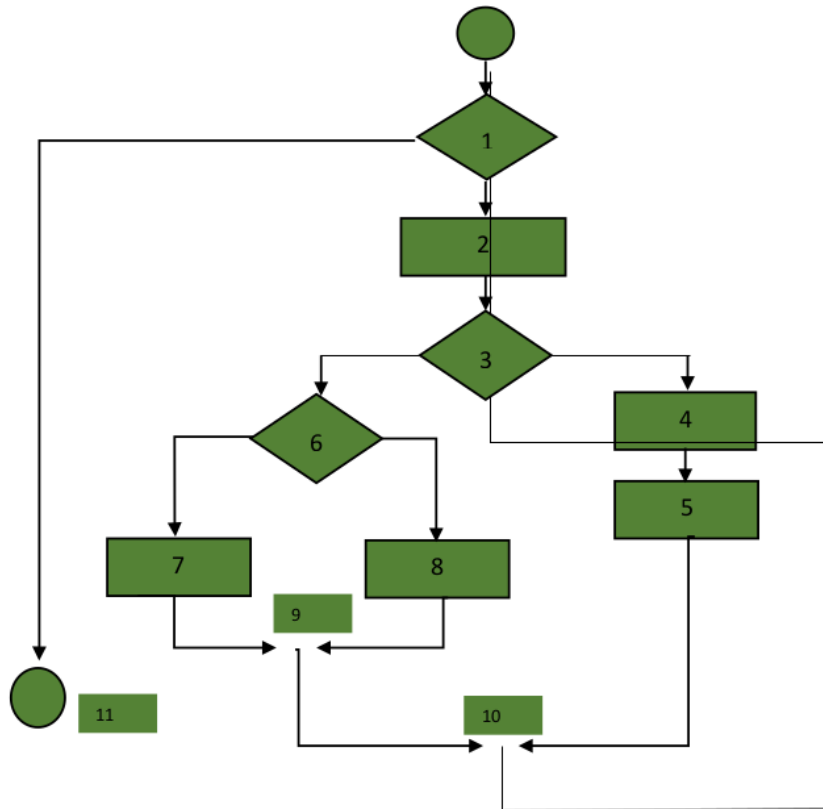
b)  $V(G) = P + 1$

Dimana :

P = Jumlah *predicate node* pada *flowgraph*

Teknik pelaksanaan pengujian *White Box* ini mempunyai tiga langkah yaitu:

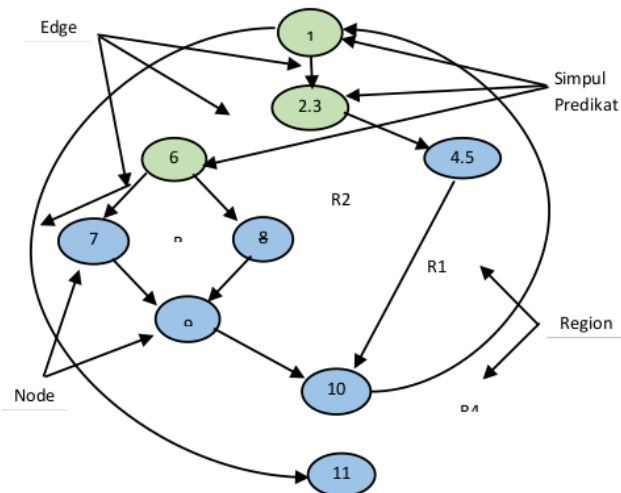
1. Menggambar *flowgraph* yang ditransfer oleh flowchart
2. Menghitung *Cyclomatic Complexity* untuk *flowgraph* yang telah dibuat
3. Menentukan jalur pengujian dari *flowgraph* yang berjumlah sesuai dengan *Cyclomatic Complexity* yang telah ditentukan.



**Gambar 2.10: Bagan Alir**

(Sumber: Roger S. Pressman, [23])

1 Bagan alir digunakan untuk menggambarkan struktur kontrol program dan untuk menggambar flowchart, dan perhatian harus diberikan pada representasi desain prosedural dalam diagram alir. Pada gambar di bawah, diagram alir memetakan diagram alir ke diagram alir yang sesuai (dengan asumsi bahwa kristal berlian pada diagram alir tidak mengandung kondisi kompleks). Merupakan satu atau lebih pernyataan prosedural. Proses pemesanan Agar dan keputusan permata dapat memetakan satu node. Panah ini, yang disebut tepi atau tautan, mewakili aliran kontrol dan mirip dengan panah diagram alir. Bahkan jika simpul "tidak menunjukkan deskripsi prosedural, ujungnya harus menggantung pada simpul tersebut.



**Gambar 2.3: Flowgraph**  
(Sumber: Roger S. Pressman, [23]).

Pada diagram *flowgraph* yang diperoleh di atas:

Rute 1 = 1-11

Rute 2 = 1-2-3-4-5-5-1-11

Rute 3 = 1-2-3-6-6-8-9-10-1-11

Rute 4 = 1-2-3-6-6-7-9-10-1-11

Rute 1,2,3,4 yang didefinisikan di atas adalah set dasar diagram alur.

*Cyclomatic complexity* digunakan untuk mencari jumlah path dalam satu flowgraph. Dapat dipergunakan rumusan sebagai berikut:

1. Jumlah region grafik alir sesuai dengan *cyclomatic complexity*.
2. *Cyclomatic complexity*  $V(G)$  untuk grafik alir dihitung dengan rumus:

$$V(G) = E - N + 2 \dots\dots\dots (1)$$

Dimana:

$E$  = jumlah edge pada grafik alir

$N$  = jumlah node pada grafik alir

1. *Cyclomatic complexity*  $V(G)$  juga dapat dihitung dengan rumus:

$$V(G) = P + 1 \dots\dots\dots (2)$$

Dimana  $P$  = jumlah *predicate node* pada grafik alir

Dari Gambar di atas dapat dihitung *cyclomatic complexity*:

1. *Flowgraph* mempunyai 4 region

2.  $V(G) = 11 \text{ edge} - 9\text{node} + 2 = 4$

3.  $V(G) = 3 \text{ predicate node} + 1 = 4$

Jadi *cyclomatic complexity* untuk *flowgraph* adalah 4

*Cyclomatic Complexity* yang tinggi menunjukkan prosedur kompleks yang sulit untuk dipahami, diuji dan dipelihara. Ada hubungan antara *Cyclomatic Complexity* dan resiko dalam suatu prosedur.

### 2.2.17 Black Box Testing

Menurut Pressman, tes Black-Box berfokus pada persyaratan fungsional perangkat lunak, yang memungkinkan para insinyur untuk mendapatkan serangkaian kondisi input untuk sepenuhnya mengimplementasikan persyaratan fungsional program. Tes kotak hitam mencari kesalahan dalam kategori berikut:

1. Fungsi yang tidak benar atau fungsi yang hilang
2. Kesalahan antar muka
3. Kesalahan dalam struktur data atau akses *database* eksternal
4. Kesalahan perilaku (*behavior*) atau kesalahan kinerja
5. Inisialisasi dan pemutusan kesalahan

Tes ini dirancang untuk menjawab beberapa pertanyaan-pertanyaan berikut ini:

- a. Bagaimana validitas fungsional diuji?
- b. Bagaimana perilaku dan kinerja sistem diuji?
- c. Apa kelas *input* akan membuat kasus uji yang baik?
- d. Apakah sistem *sensitive* terhadap nilai input tertentu?
- e. Bagaimana batas-batas kelas data yang terisolasi?
- f. Kecepatan dan volume data seperti apa yang dapat ditolerir sistem?

g. Efek apakah yang akan menspesifikasikan kombinasi data dalam sistem operasi?

6

1. Ciri-Ciri Black Box Testing

- a. *Black box testing* berfokus pada kebutuhan fungsional pada *software*, berdasarkan pada spesifikasi kebutuhan dari *software*.
- b. *Black box testing* bukan teknik alternatif daripada *white box testing*. Lebih daripada itu, ia merupakan pendekatan pelengkap dalam mencakup *error* dengan kelas yang berbeda dari metode *white box testing*.
- c. *Black box testing* melakukan pengujian tanpa pengetahuan detail struktur internal dari sistem atau komponen yang dites. juga disebut sebagai *behavioral testing*, *specification-based testing*, *input/output testing* atau *functional testing*

11

2. Jenis teknik *design tes* yang dapat dipilih berdasarkan pada tipe testing yang akan digunakan.

- a. *Equivalence Class Partitioning*
- b. *Boundary Value Analysis*
- c. *State Transitions Testing*
- d. *Cause-Effect Graphing*

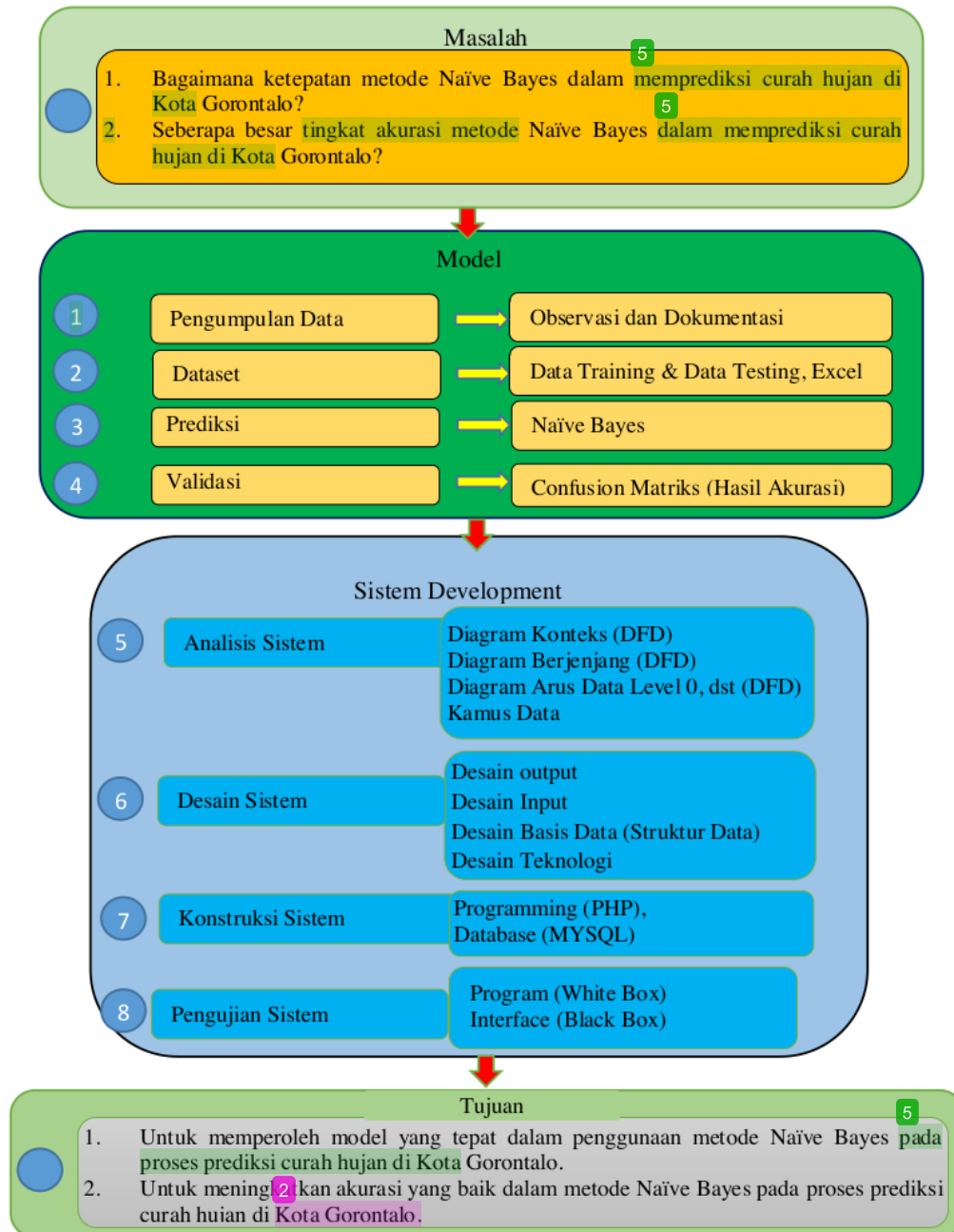
3. Kategori *error* yang akan diketahui melalui *black box testing*

- a. Fungsi yang hilang atau tak benar
- b. *Error* dari antar-muka
- c. *Error* dari struktur data atau akses eksternal database
- d. *Error* dari kinerja atau tingkah laku
- e. *Error* dari inisialisasi dan terminasi

### 3. Perangkat Lunak Pendukung

Perangkat lunak pendukung yang digunakan oleh penulis ketika membangun sistem ini adalah PHP dan MySQL.

#### 4. Kerangka Pikir



Gambar 2.5 Kerangka pikir

## BAB III

### METODE PENELITIAN

#### 3.1. Jenis, Metode, Subjek, Waktu dan Lokasi Penelitian

Berdasarkan dari tingkat penerapan maka, penelitian ini merupakan penelitian terapan. Dipandang dari jenis informasi yang diolah maka, penelitian ini merupakan penelitian kuantitatif. sedangkan dari perilaku terhadap data, maka penelitian ini merupakan penelitian konfirmatori.

#### 3.2. Pengumpulan Data

Untuk mengumpulkan data digunakan 2 (dua) jenis data, data primer dan data sekunder. Data primer adalah data dari penelitian lapangan dan data sekunder dari penelitian kepustakaan

1. Penelitian Data Primer (Bidang/lapangan) Untuk memperoleh data primer yang merupakan data dari objek penelitian yaitu bertempat di BMKG Stasiun Kkimatologi Tilongkabila Gorontalo maka dilakukan dengan teknik:

- a. Pengamatan, metode ini memungkinkan analisis sistem untuk secara langsung mengamati atau meninjau Adapun pada penelitian ini dilakukan dengan cara mengumpulkan data suhu rata-rata, kecepatan angin, kelembaban udara, tekanan udara, lama penyinaran pada BMKG Stasiun Klimatologi Tilongkabila Gorontalo.
- b. Melakukan Tanya jawab dengan pihak BMKG Stasiun Klimatologi Tilongkabila Gorontalo untuk proses prediksi curah hujan di Kota Gorontalo.

**Tabel 3. 1** Atribut data

No	Name	Type	Value	Keterangan
1.	Suhu rata-rata	Varchar	0 - 255	Parameter Input
2.	Kelembaban Udara	Varchar	0 - 255	Parameter Input
3.	Tekanan Udara	Varchar	0 - 255	Parameter Input
4.	Penyinaran	Varchar	0 - 255	Parameter Input
5.	Kecepatan Angin	Varchar	0 – 255	Parameter Input
6.	Curah Hujan	Varchar	0 – 255	Parameter Output

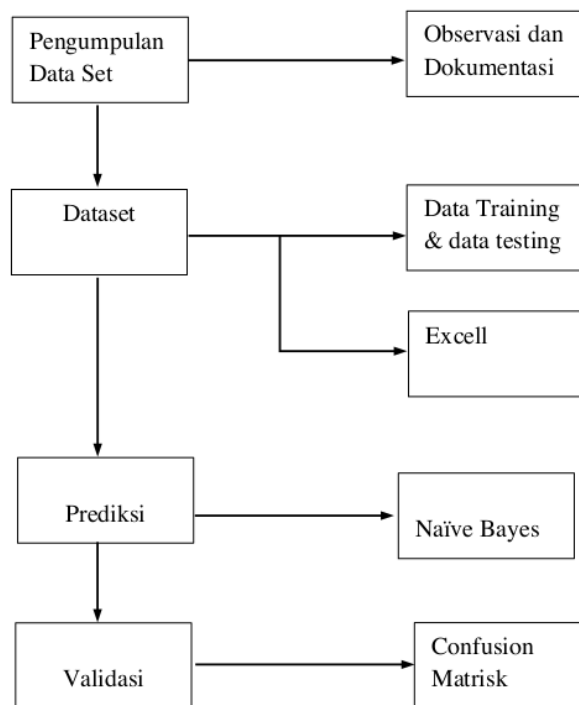
## 2. Penelitian data sekunder (literatur/kepuustakaan)

Diperlukan metode literatur untuk membawa data tambahan untuk tujuan melengkapi data primer. Data sekunder diperoleh dari studi literatur, termasuk landasan teori. Metode perpustakaan digunakan oleh analis sistem sebagai contoh dokumen yang terkait dengan bahan penelitian. Selain itu, analis sistem mencari data tentang masalah atau parameter dalam bentuk catatan, buku, majalah, dll yang terkait dengan penelitian ini.

### 3.3. Pemodelan / Abstraksi

#### 3.3.1. Pengembangan Model

Prosedur atau langkah-langkah pokok dalam Prediksi Curah Hujan di Kota Gorontalo menggunakan Metode Naïve Bayes yaitu dengan alat bantu tools PHP, Database MySQL serta *White Box Testing* dan *Black Box Testing* untuk menguji kinerja sistemnya.



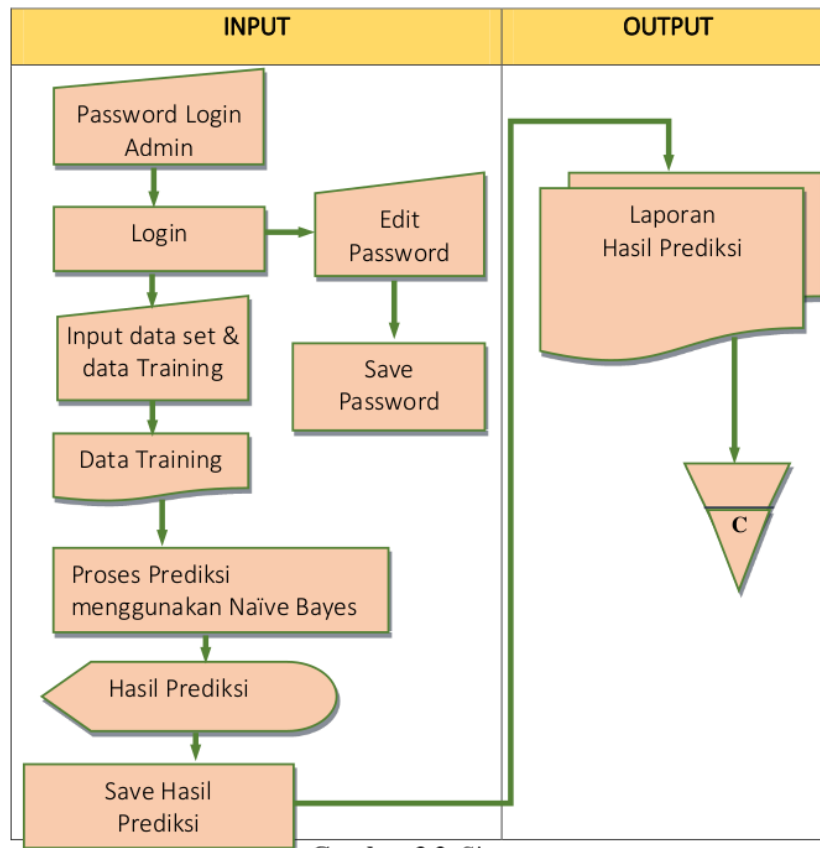
Gambar 3. 1. Model

### 3.3.2. Evaluasi Model

Model yang telah dihasilkan kemudian dievaluasi dengan menggunakan *Confusion Matrics* untuk mengetahui *Error*.

### 3.4. Pengembangan Sistem

Sistem yang diusulkan dapat digambarkan menggunakan *flowchart* berikut ini:



Gambar 3.2. Sistem

### 3.5. Analisa Sistem

Analisis sistem menggunakan pendekatan berorientasi *procedural/structural*:

- Diagram Konteks, menggunakan alat bantu DFD
- Diagram Berjenjang, menggunakan alat bantu DFD
- Diagram Arus Data Level 0,1,dst. menggunakan alat bantu DFD
- Kamus Data menggunakan alat bantu Ms. Word.

### 3.6. Desain Sistem

Desain sistem menggunakan pendekatan berorientasi objek yang digambarkan dalam bentuk :

#### a. Desain Output

Desain *output* dimaksudkan untuk mengetahui bagaimana bentuk *output-output* dari sistem yang akan di buat. Desain keluaran terperinci dibagi menjadi dua jenis: desain keluaran dalam format laporan untuk kertas dan desain keluaran dalam format dialog pada layar terminal (monitor).

#### b. Desain input

Desain *Input* merupakan awal dari pemrosesan informasi. Bahan baku untuk informasi adalah data yang berasal dari transaksi konsumen. Data yang dihasilkan dari transaksi tidak dapat dipisahkan dari data yang dimasukkan. Desain input terperinci dimulai dari *input* pertama dalam desain dokumen dasar. Jika dokumen yang mendasarinya tidak dirancang dengan benar, kemungkinan input yang direkam mungkin tidak akurat atau tidak cukup.

#### c. Desain database

Basis data (*database*) merupakan kumpulan data yang terhubung satu sama lain dan disimpan dalam simpanan di luar komputer dan mengumpulkan data menggunakan perangkat lunak tertentu. *Database* adalah komponen penting dalam sistem informasi karena mereka digunakan sebagai dasar untuk memberikan informasi kepada pengguna. Dalam aplikasi, aplikasi *basis data* disebut *sistem basis data*.

#### d. Desain teknis

Pada tahap ini, Anda memutuskan teknik mana yang akan digunakan untuk menerima input, menjalankan model, menyimpan dan mengakses data, menghasilkan dan mengirimkan output, dan mengendalikan seluruh sistem.

#### e. Desain Program

Pada tahap ini menggunakan alat bantu PHP dalam bentuk Pseudoce program pada untuk Prediksi Curah Hujan di Kota Gorontalo menggunakan Metode Naïve Bayes.

### 3.7. Konstruksi Sistem

Pada tahap ini dilakukan pembuatan sistem menggunakan *tools PHP* dan Database *MySQL* serta *White Box Testing* dan *Black Box Testing* untuk menguji kinerja sistem dan pengukuran akurasi menggunakan *Confusion Matrix*. Pada langkah ini, kami melakukan tahap produksi dari hasil sistem untuk analisis desain sistem sebelumnya. Ini termasuk menginstal paket tambahan untuk menjalankan program, membuat daftar program, dan membangun dalam bentuk formulir, antarmuka dan integrasi sistem-sistem program yang terdiri dari input, proses, dan output yang tersusun dalam sebuah sistem menu sehingga dapat dijalankan oleh pengguna sistem

### 3.8. Pengujian Sistem

Setelah analisis sistem, desain, dan fase produksi telah dilakukan, kami menguji semua perangkat lunak, program tambahan, dan semua program yang terlibat dalam membangun sistem untuk melakukan fase uji untuk memastikan bahwa sistem berfungsi dengan baik. Fokus pada menemukan kesalahan, meninjau dan mengevaluasi sistem yang dikembangkan pada tahap ini, terlepas dari apakah mereka dilakukan sesuai dengan desain atau tidak. Pengujian yang dilakukan dengan menggunakan teknik pengujian perangkat lunak yaitu:

#### a. Pengujian *White Box*

*Software* yang sudah direvisi kemudian diuji dengan metode *white box testing* pada kode program proses penerapan metodenya/modelnya. Kode program tersebut kemudian dipetakan kedalam bentuk *flowgraph* (bagan alir kontrol) yang tersusun dari beberapa node dan *edge*. Berdasarkan *flowgraph*, ditentukan jumlah *region* dan *Cyclomatic Complexity* (CC). Apabila *Independent Path* =  $V(G) = (CC) = \text{Region}$ , di mana setiap *Path* hanya dieksekusi sekali dan sudah benar, maka sistem dinyatakan efisien dari segi kelayakan logika pemrograman.

#### b. Pengujian *Black Box*

Pengujian *Black Box* melalui program *PHP* dan Database *MySQL*. Selanjutnya *software* diuji pula dengan metode *black box testing* yang fokus pada keperluan fungsional dari *software* dan berusaha untuk menemukan kesalahan dalam beberapa kategori, diantaranya: (1). Fungsi-fungsi yang salah atau hilang;

(2). kesalahan *interface*; (3). kesalahan dalam struktur data atau akses basis data eksternal; (4). kesalahan performa; (5). kesalahan inisialisasi dan terminasi. Jika sudah tidak ada kesalahan-kesalahan tersebut, maka sistem dinyatakan efisien dari segi kesalahan komponen-komponen sistem.

## 1 BAB IV

### ANALISA DAN DESAIN SISTEM

#### 4.1. Hasil Pengumpulan Data

Berikut ini adalah hasil pengumpulan data di Bmkg stasiun klimatologi tilongkabila Gorontalo yang didapatkan dan dikumpulkan dengan menggunakan teknik observasi dan wawancara melalui beberapa pihak yang terkait.

**Tabel 4.1** Hasil Pengumpulan Data

No.	Bulan	Suhu rata-rata ( $^{\circ}\text{C}$ )	Kelembaban Udara (%)	Tekanan Udara (Mbs)	Penyinaran (%)	Kecepatan Angin ( <i>Knot</i> )	Curah Hujan
1	Jan	26,9	83	Rendah	Rendah	Tinggi	hujan lebat
2	Feb	27	83	Rendah	Rendah	Rendah	hujan lebat
3	Mar	27,5	81	Rendah	Sedang	Sedang	hujan lebat
4	Aprl	27	81	Sedang	Sedang	Sedang	hujan ringan
5	Mei	27,8	84	Tinggi	Rendah	Sedang	hujan lebat
6	Jun	27,5	83	Sedang	Rendah	Sedang	hujan lebat
7	Jul	27,5	80	Rendah	Sedang	Sedang	tidak hujan
8	Agt	28	76	Rendah	Sedang	Tinggi	tidak hujan
9	Sept	28,5	75	Rendah	Rendah	Tinggi	tidak hujan
10	Okt	28,9	74	Rendah	Tinggi	Tinggi	hujan ringan
11	Nov	28,4	80	Tinggi	Sedang	Sedang	hujan ringan
12	Des	28,6	78	Tinggi	Tinggi	Tinggi	hujan lebat
13	Jan	26,9	83	Rendah	Rendah	Tinggi	hujan

							lebat
14	Feb	27	83	Rendah	Rendah	Rendah	hujan ringan
15	Mar	27,5	81	Rendah	Sedang	Sedang	hujan ringan
16	Aprl	27	81	Sedang	Sedang	Sedang	hujan ringan
17	Mei	27,8	84	Tinggi	Rendah	Sedang	hujan lebat
18	Jun	27,5	83	Sedang	Rendah	Sedang	hujan lebat
19	Jul	27,5	80	Rendah	Sedang	Sedang	hujan ringan
20	Agt	28	76	Rendah	Sedang	Tinggi	hujan ringan
21	Sept	28,5	75	Rendah	Rendah	Tinggi	hujan lebat
22	Okt	28,9	74	Rendah	Tinggi	Tinggi	hujan lebat
23	Nov	28,4	80	Tinggi	Sedang	Sedang	hujan lebat
24	Des	28,6	78	Tinggi	Tinggi	Tinggi	hujan ringan
25	Jan	26,9	83	Rendah	Rendah	Tinggi	hujan lebat
26	Feb	27	83	Rendah	Rendah	Rendah	hujan lebat
27	Mar	27,5	81	Rendah	Sedang	Sedang	hujan lebat
28	Aprl	27	81	Sedang	Sedang	Sedang	hujan ringan
29	Mei	27,8	84	Tinggi	Rendah	Sedang	hujan lebat
30	Jun	27,5	83	Sedang	Rendah	Sedang	hujan lebat
31	Jul	27,5	80	Rendah	Sedang	Sedang	hujan ringan
32	Agt	28	76	Rendah	Sedang	Tinggi	hujan lebat
33	Sept	28,5	75	Rendah	Rendah	Tinggi	tidak hujan
34	Okt	28,9	74	Rendah	Tinggi	Tinggi	hujan ringan
35	Nov	28,4	80	Tinggi	Sedang	Sedang	hujan

							ringan
36	Des	28,6	78	Tinggi	Tinggi	Tinggi	hujan lebat
37	Jan	26,9	83	Rendah	Rendah	Tinggi	hujan lebat
38	Feb	27	83	Rendah	Rendah	Rendah	hujan lebat
39	Mar	27,5	81	Rendah	Sedang	Sedang	hujan lebat
40	Aprl	27	81	Sedang	Sedang	Sedang	hujan ringan
41	Mei	27,8	84	Tinggi	Rendah	Sedang	hujan lebat
42	Jun	27,5	83	Sedang	Rendah	Sedang	hujan lebat
43	Jul	27,5	80	Rendah	Sedang	Sedang	hujan ringan
44	Agt	28	76	Rendah	Sedang	Tinggi	hujan lebat
45	Sept	28,5	75	Rendah	Rendah	Tinggi	tidak hujan
46	Okt	28,9	74	Rendah	Tinggi	Tinggi	hujan ringan
47	Nov	28,4	80	Tinggi	Sedang	Sedang	hujan ringan
48	Des	28,6	78	Tinggi	Tinggi	Tinggi	hujan lebat
49	Jan	26,9	83	Rendah	Rendah	Tinggi	hujan lebat
50	Feb	27	83	Rendah	Rendah	Rendah	hujan lebat
51	Mar	27,5	81	Rendah	Sedang	Sedang	hujan ringan
52	Aprl	27	81	Sedang	Sedang	Sedang	hujan ringan
53	Mei	27,8	84	Tinggi	Rendah	Sedang	hujan lebat
54	Jun	27,5	83	Sedang	Rendah	Sedang	hujan lebat
55	Jul	27,5	80	Rendah	Sedang	Sedang	hujan lebat
56	Agt	28	76	Rendah	Sedang	Tinggi	hujan ringan
57	Sept	28,5	75	Rendah	Rendah	Tinggi	tidak

							hujan
58	Okt	28,9	74	Rendah	Tinggi	Tinggi	hujan lebat
59	Nov	28,4	80	Tinggi	Sedang	Sedang	tidak hujan
60	Des	28,6	78	Tinggi	Tinggi	Tinggi	hujan lebat

#### 4.2. Hasil Penerapan Model Naïve Bayes

Hasil pemodelan naïve bayes ditunjukkan pada tabel dibawah ini dengan sebagai berikut:

##### 4.2.1. Data Training

**Table 4.2** Penerapan metode naïve bayes

No.	Bulan	Suhu rata-rata (°C)	Kelembaban Udara (%)	Tekanan Udara (Mbs)	Penyinaran (%)	Kecepatan Angin (Knot)	Curah Hujan
1	Jan	>26	>80	Rendah	Rendah	Tinggi	hujan lebat
2	Feb	<26	>80	Rendah	Rendah	Rendah	hujan lebat
3	Mar	<26	>80	Rendah	Sedang	Sedang	hujan lebat
4	Aprl	>26	>80	Sedang	Sedang	Sedang	hujan ringan
5	Mei	>26	>80	Tinggi	Rendah	Sedang	hujan lebat
6	Jun	>26	>80	Sedang	Rendah	Sedang	hujan lebat
7	Jul	>26	>80	Rendah	Sedang	Sedang	tidak hujan
8	Agt	>26	<80	Rendah	Sedang	Tinggi	tidak hujan
9	Sept	>26	<80	Rendah	Rendah	Tinggi	tidak hujan
10	Okt	>26	<80	Rendah	Tinggi	Tinggi	hujan ringan
...	....	.....	....	.....	.....	.....	.....
57	Sept	>26	<80	Rendah	Rendah	Tinggi	tidak hujan
58	Okt	>26	<80	Rendah	Tinggi	Tinggi	hujan lebat
59	Nov	>26	>80	Tinggi	Sedang	Sedang	tidak hujan
60	Des	>26	<80	Tinggi	Tinggi	Tinggi	hujan lebat

#### 4.2.2. Data uji

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

Suhu rata-rata = >26

Kelembaban = >80

Tekanan udara = tinggi

Penyinaran = tinggi

Kecepatan angin = Tinggi

Status ?

Perhitungan metode naïve bayes

Probabilitas kelas (status)

$P(\text{Hujan lebat}) = 33/60 = 0,55$

$P(\text{Hujan ringan}) = 20/60 = 0,33333$

$P(\text{Tidak hujan}) = 7/60 = 0,11667$

Menghitung Kemungkinan Termasuk kategori Hujan Lebat

$P(\text{Class Hujan Lebat/ suhu rata-rata} = >26) = 31/33 = 0,93939394$

$P(\text{Class Hujan Lebat/ kelembaban udara} = >80) = 25/33 = 0,75757576$

$P(\text{Class Hujan Lebat/ Tekanan udara} = \text{Tinggi}) = 10/33 = 0,3030303$

$P(\text{Class Hujan Lebat/ penyinaran} = \text{Tinggi}) = 6/33 = 0,18181818$

$P(\text{Class Hujan Lebat/ kecepatan angin} = \text{Tinggi}) = 14/33 = 0,42424242$

Class Hujan Lebat =

$0,55 * 0,93939394 * 0,75757576 * 0,3030303 * 0,18181818 * 0,42424242$

Class hujan Lebat = 0,00914904

Menghitung Kemungkinan Termasuk kategori Hujan Ringan

$P(\text{Class Hujan Ringan / suhu rata-rata} = >26) = 20/20 = 1$

$P(\text{Class Hujan Ringan / kelembaban udara} = >80) = 15/20 = 0,75$

$P(\text{Class Hujan Ringan / Tekanan udara} = \text{Tinggi}) = 4/20 = 0,2$

$$P(\text{Class Hujan Ringan} / \text{penyinaran} = \text{Tinggi}) = 4/20 = 0,2$$

$$P(\text{Class Hujan Ringan} / \text{kecepatan angin} = \text{Tinggi}) = 6/20 = 0,3$$

$$\text{Class Hujan Ringan} = 0,333333 * 1 * 0,75 * 0,2 * 0,3$$

$$\text{Class hujan Ringan} = 0,00299997$$

Menghitung Kemungkinan Termasuk kategori Tidak Hujan

$$P(\text{Class Tidak Hujan} / \text{suhu rata-rata} = >26) = 7/7 = 1$$

$$P(\text{Class Tidak Hujan} / \text{kelembaban udara} = >80) = 2/7 = 0,28571429$$

$$P(\text{Class Tidak Hujan} / \text{Tekanan udara} = \text{Tinggi}) = 1/7 = 0,14285714$$

$$P(\text{Class Tidak Hujan} / \text{penyinaran} = \text{Tinggi}) = 0/7 = 0$$

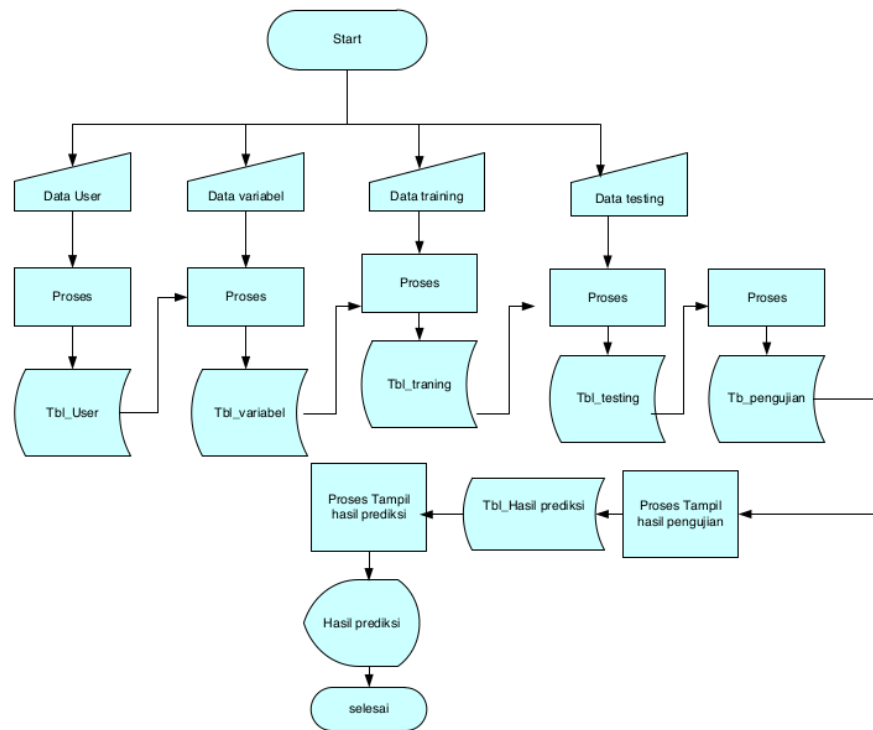
$$P(\text{Class Tidak Hujan} / \text{kecepatan angin} = \text{Tinggi}) = 5/7 = 0,71428571$$

$$\text{Class Tidak Hujan} = 0,11667 * 1 * 0,28571429 * 0,14285714 * 0 * 0,71428571$$

$$\text{Class Tidak Hujan} = 0$$

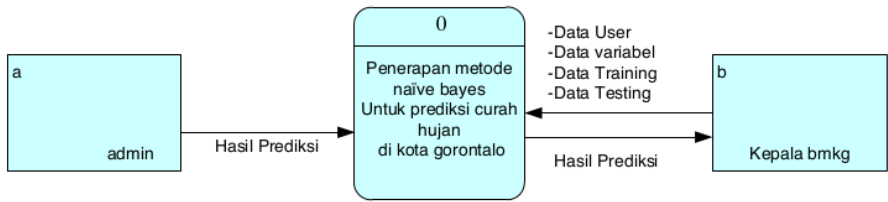
## 4.3 Hasil pengembangan sistem

### 4.3.1 Desain sistem



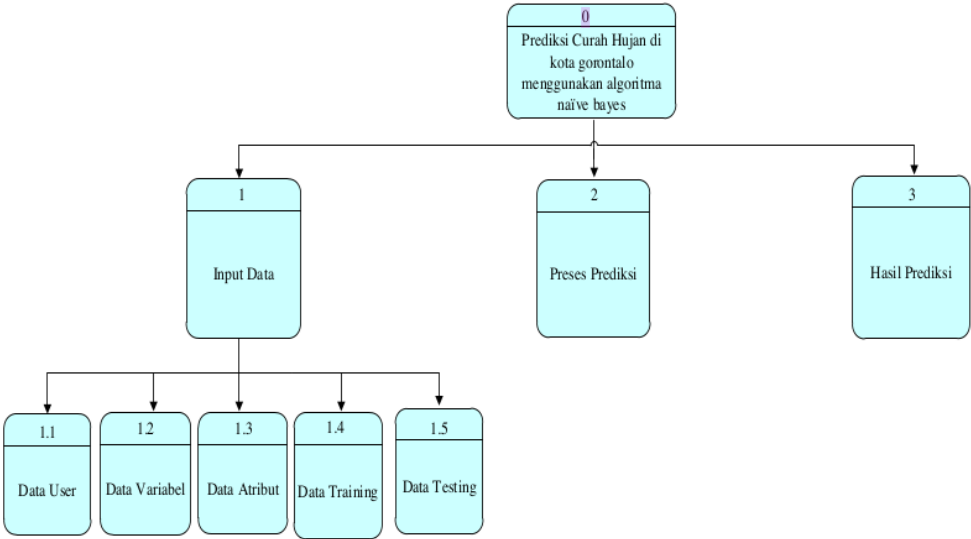
Gambar 4.1 Bagan Alir System yang Diusulkan

4.3.1.1 Diagram Konteks



2  
Gambar 4.2 Diagram Konteks

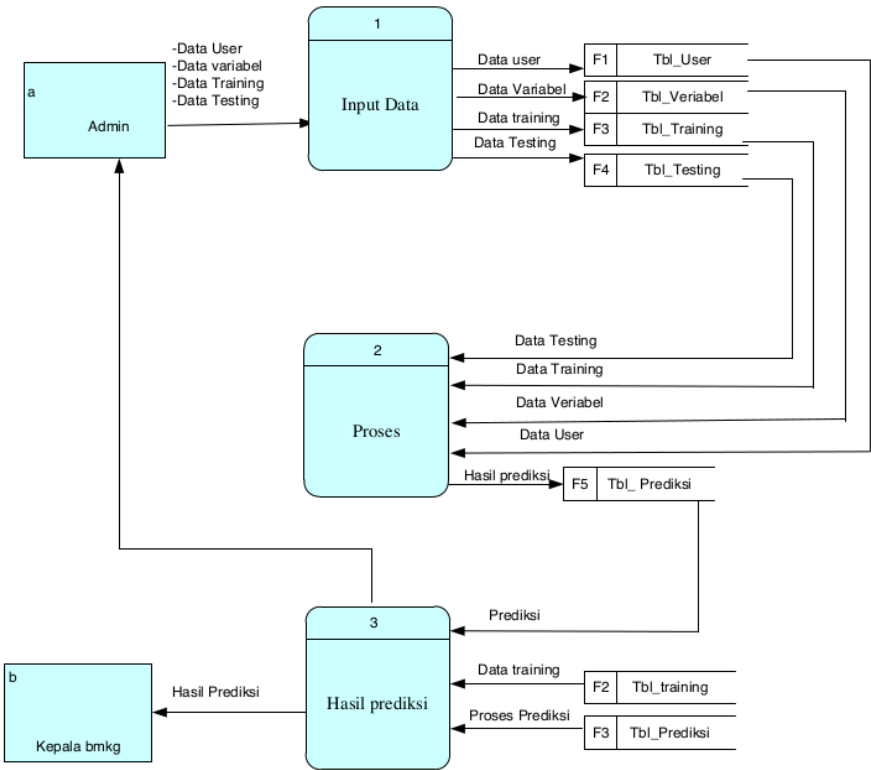
4.3.1.2 Diagram Berjenjang



2  
Gambar 4.3 Diagram Berjenjang

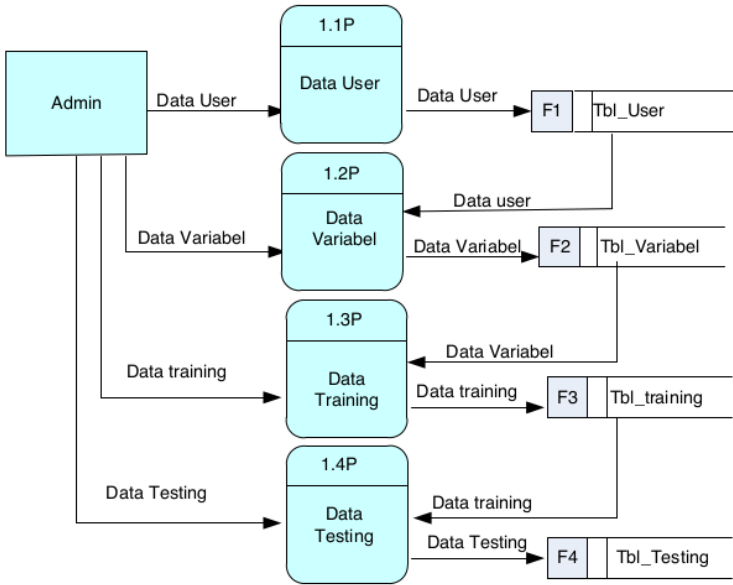
4.4 Diagram Arus Data

4.4.1 DAD Level 0



Gambar 4.4 DAD level 0

4.4.2 DAD Level 1



Gambar 4.5 DAD level 1

**1**  
**4.5. Kamus Data**

**Tabel 4.1. Kamus Data User**

Nama Arus Data : Data User

Penjelasan : Berisi data-data Users

Periode : Setiap ada penambahan data Users (non periodik)

Bentuk Data : Dokumen

Arus Data : a-1, 1-F1, a-1.1P, 1.1P-F1

25

No.	Nama Field	Type	Ukuran	Keterangan
1.	Id_User	N	3	Id User
3.	Username	C	100	Username User
4.	Password	C	100	Password User

**1**  
**Tabel 4.1. Kamus Data Variabel**

Kamus Data : Detail Variabel				
Nama Arus Data : Data Detail Variabel			Bentuk Data : Dokumen, Laporan	
Penjelasan : Berisi data-data detail Variabel untuk penginputan data			Arus Data : a-1-F1, F1-2, 2-F3, 1.1P, 1.1P-F1, F1-1.2P	
Periode : Non periodik				
Struktur Data sebagai Berikut :				
No	Nama Item Data	Type	Width	Keterangan
1.	id_Variabel	C	4	ID Variabel
2.	Nama_Variabel	C	100	Nama Variabel

**Tabel 4.3. Kamus Data Training**

Kamus Data : Data Training				
Nama Arus Data : Data Detail Training			Bentuk Data : Dokumen	
Penjelasan : Berisi data-data detail Training untuk penginputan data			Arus Data : a-1, 1-F2, F2-2, 2-F4, a-1.2P, 1.2P-F2, F2-1.3P, F2-2P, F2-3.1P	
Periode :Setiap ada penambahan data Training				
No	Field Name	Type	Size	Index
1	Id_training	int	5	Id training
2	Tahun	Int	10	Tahun
3	Bulan	Varchar	200	Bulan
4	Id_variabel	Varchar	4	Id_variabel
5	Id_atribut	Varchar	4	Id_atribut
6	Value	Varchar	100	Value
1	Kelas	Varchar	100	Kelas

**Tabel 4.3. Kamus Data Testing**

Kamus Data : Data Testing	
Nama Arus Data : Data Detail testing	Bentuk Data : Dokumen
Penjelasan : Berisi data-data detail testing untuk penginputan data	Arus Data : a-1, 1-F2, 2-F5, F5-3, a-1.3P, 1.3P-F2, a-2.1P, 2.1P-F2, a-2.1P, 2.1P-F3, F3-3.1P,
Periode : Setiap ada penambahan data Testing	

<sup>1</sup> No	Field Name	Type	Size	Index
<sup>1</sup>	Id_testing	int	5	Id testing
<sup>2</sup>	Tahun	Int	10	Tahun
<sup>3</sup>	Bulan	Varchar	100	Bulan
4	Id_variabel	Varchar	4	Id_variabel
5	Id_atribut	Varchar	4	Id_atribut
6	Value	Varchar	100	Value

<sup>1</sup>**Tabel 4.1. Kamus Data Atribut**

Kamus Data : Detail Atribut				
Nama Arus Data : Data Detail Atribut			Bentuk Data : Dokumen, Laporan	
Penjelasan : Berisi data-data detail Atribut untuk penginputan data			<sup>2</sup> Arus Data : a-1-F1, F1-2, 2-F3, 1.1P, 1.1P-F1, F1-1.2P	
Periode : Non periodik				
Struktur Data sebagai Berikut :				
No	Nama Item Data	Type	Width	Keterangan
1.	id_Atribut	Varchar	4	ID Atribut
2.	Nama_Atribut	Varchar	100	Nama Atribut

**Tabel 4.7** Kamus prediksi

<b>Kamus Data : Hasil Prediksi</b>				
Nama Arus Data : Prediksi				
Penjelasan : Berisi data-data Hasil Prediksi				
Periode : Setiap ada penambahan data Hasil Prediksi				
Struktur Data :				
No	Nama Item Data	Type	Width	Description
1.	Id_Testing	N	5	No id prediksi
2.	Probabilitas	Float	-	Probabilitas
3.	kelas	C	100	kelas

#### 4.6 Arsitektur Sistem

Spesifikasi hardware dan software yang di rekomendasikan yaitu :

- a. Prosesor : dual core atau lebih
  - b. RAM : 1 GB atau lebih
  - c. VGA : 256MB atau lebih
  - d. Harddisk : 40 GB atau lebih
  - e. Operating System : Windows 10
- Tools : Notepad ++

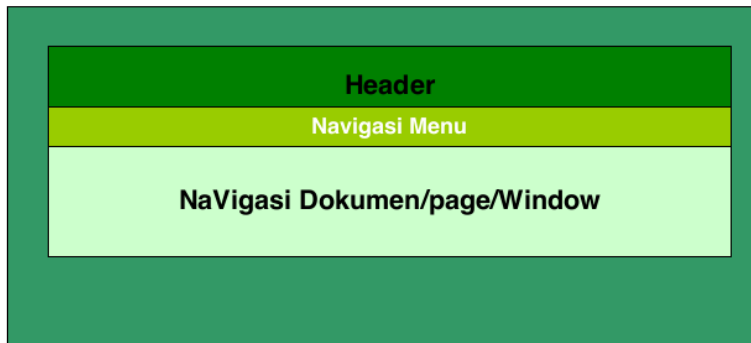
#### 4.7 Interface Desain

##### 4.7.1 Mekanisme User

**Tabel 4.3** Interface Desain

USER	KATEGORI	AKSES INPUT	AKSES OUTPUT
Admin	Administrator	All	All

#### 4.7.2 Mekanisme Navigasi



**Gambar 4.7** Interface Design Mekanisme Navigasi

#### 4.7.3 Interface Design Mekanisme Input

##### 4.7.3.1 Tampilan Input Halaman Login

The diagram shows the login form design. It is a gray rectangular box with the title 'login' at the top. Inside, there are two input fields: 'User Name' and 'Password', each with a corresponding label box to its left. Below these fields is a 'Login' button.

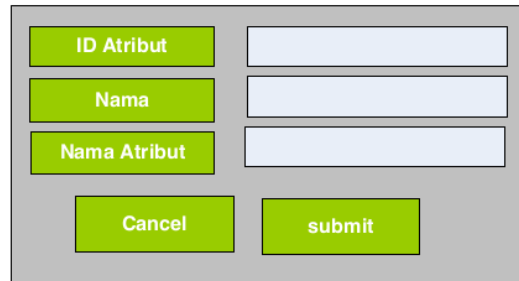
**Gambar 4.8** Desain Login Admin

##### 4.7.3.2. Tampilan Input Variabel

The diagram shows the input variable form design. It is a gray rectangular box. Inside, there are two input fields: 'ID' and 'Variabel', each with a corresponding label box to its left. Below these fields are two buttons: 'Cancel' and 'submit'.

**Gambar 4.9** Input Variabel

#### 4.7.3.3. Tampilan Data Atribut



The screenshot shows a form titled "Data Atribut" with a light gray background. It contains three input fields with light blue backgrounds and black borders, each preceded by a green label box. The labels are "ID Atribut", "Nama", and "Nama Atribut". Below the input fields are two green buttons: "Cancel" and "submit".

**Gambar 4.10** Data Atribut

#### 4.7.3.4 Tampilan Input Data Traning



The screenshot shows a form titled "Data Training" with a light green background. It has a green header bar with the title "Data Training". Below the header are seven input fields with light blue backgrounds and black borders, each preceded by a green label box. The labels are "ID", "Suhu rata rata", "Kelembaban udara", "Tekanan udara", "penyinaran", "Kecepatan angin", and "Status". The "Status" field is a dropdown menu. At the bottom are two green buttons: "Simpan Data" and "kembali".

**Gambar 4.11** Tampilan Data Training

#### 4.7.3.5. Tampilan input Halaman Data Testing

Data Testing	
ID	<input type="text"/>
Suhu rata rata	<input type="text"/>
Kelembaban udara	<input type="text"/>
Tekanan udara	<input type="text"/>
penyinaran	<input type="text"/>
Kecepatan angin	<input type="text"/>
↓	
<div>Proses</div> <div>kembali</div>	

**Gambar 4.12** Tampilan Data Testing

#### 4.7.3.6 Tampilan Input User

Data User	
ID	<input type="text"/>
User Name	<input type="text"/>
Password	<input type="text"/>
<div>Simpan</div> <div>Kembali</div>	

**Gambar 4.13** Input User

#### 4.7.4 Interface Output



**Gambar 4.14** Interface Output

#### 4.8. Alat Dan Bahan

Alat Dan Bahan yang di gunakan pada sistem prediksi Curah Hujan ini Menggunakan :

1. Notepad ++(txt) Sebagai Editor Web
2. Data Mysql server untuk Penyimpanan Database
3. Keduannya dihubungkan dan dimanipulasi dengan teknik disconnected data

#### 4.8. Struktur Data

**Tabel 4.4** Struktur Data User

Nama : tbl\_User

Type : Transaksi

Primary Key :Id\_User

ForegnKey : -

Media : Harddisk

Fungsi : Untuk Menambahkan Data User dan Login

Struktur Data :

21					
No	Field	Type	Size	Range	Keterangan
1	Id	Int	5	10	Id
2	Username	Char	100	200	Nama User
3	Password	Char	100	200	Password User

**Tabel 4.5** struktur Data Variabel

Nama : Variabel

Type : Transaksi

Primary Key :Id

ForegnKey : -

Media : Harddisk

Fungsi : Untuk Meyimpan Variabel

Struktur Data :

No	Field	Type	Size	Range	Keterangan
1	Id Variabel	Varchar	4	20	Id Variabel
2	Nama Variabel	Varchar	100	200	Nama Variabel

**Tabel 4.6** Struktur Data Training

Nama	: tbl_Data Training				
Type	: Transaksi				
Primary Key	:Id				
ForegnKey	: -				
Media	: Harddisk				
Fungsi	: Untuk Menginput Data Training				
Struktur Data	:				
No	Field	Type	Size	Range	Keterangan
1	Id Training	Int	5	10	Id Training
2	Tahun	Int	10	200	Tahun
3	Bulan	Varchar	200	200	Bulan
4	Id variabel	Varchar	4	200	Id variabel
5	Id atribut	Varchar	4	200	Id atribut
6	Value	Varchar	100	200	Value
7	Kelas	Varchar	100	200	Kelas

**Tabel 4.7**Struktur Data Testing

Nama	: tbl _Data Testing				
Type	: Transaksi				
Primary Key	:Id				
ForegnKey	: -				
Media	: Harddisk				
Fungsi	: Untuk Menginput Data Testing				
Struktur Data	:				
No	Field	Type	Size	Range	Keterangan
1	Id Testing	Int	5	10	Id Testing
2	Tahun	Varchar	10	200	Tahun
3	Bulan	Varchar	100	200	Bulan
4	Id Variabel	Varchar	4	200	Id Variabel
5	Id Atribut	Varchar	4	200	Id Atribut
6	Value	Varchar	100	200	Value

**Tabel 4.8** Struktur data Atribut

Nama	: Atribut				
Type	: Transaksi				
Primary Key	:Id				
ForegnKey	: -				
Media	: Harddisk				
Fungsi	: Untuk Meyimpan Atribut				
Struktur Data	:				
No	Field	Type	Size	Range	Keterangan
1	Id Atribut	Int	4	20	Id Attribut
2	Id Variabel	Varchar	4	200	Id Variabel
3	Nama Atribut	Varchar	100	200	Nama Attribut

**Tabel 4.9** Struktur Hasil prediksi

Nama : prediksi

Type : Transaksi

Primary Key :Id

ForegnKey : -

Media : Harddisk

Fungsi : Untuk Meyimpan Hasil prediksi

Struktur Data :

21					
No	Field	Type	Size	Range	Keterangan
1	Id Testing	Varchar	5	10	Id Testing
2	Probabilitas	Float			Probabilitas
3	Kelas	Varchar	100	200	Kelas

#### 4.9. Desain Output Secara Umum

**sistem** : Prediksi Curah Hujan Di Kota Gorontalo Menggunakan

Algoritma Naïve bayes

**Tahap** : Desain Output Secara Umum

**Tabel 4.10** Desain Output Secara Umum

Kode Input	Nama Input	Sumber	Tipe File	Periode
O-001	Hasil Prediksi curah hujan	Admin	Dokumen	Non Periodik

#### 4.10. Desain Input Secara Umum

**Sistem :** Prediksi Curah Hujan Di Kota Gorontalo Menggunakan  
Algoritma Naïve Bayes

**Tahap :** Desain Input Secara Umum

**Tabel 4.11** Desain Input Secara Umum

Kode Input	Nama Input	Sumber	Tipe File	Periode
I-001	Data User	Admin	Indeks	Non Periodik
I-002	Data variabel	Admin	Index	Non Periodik
I-003	Data Training	Admin	Index	Non Periodik
I-004	Data Testing	Admin	Index	Non Periodik

#### 4.11 Desain Database Secara Umum

**Sistem :** Prediksi Curah Hujan Di Kota Gorontalo Menggunakan  
Algoritma Naïve Bayes

**Tahap :** Desain Database Secara Umum

**Tabel 4.12** Daftar File Yang Didesain

Kode File	Nama File	Tipe File	Media File	Organisasi File	Field Kunci
I-001	Data User	Master	Hard Disk	Index	Id_user
I-002	Data variabel	Master	Hard Disk	Index	Id_variabel, Id_user, Id_pilihan
I-003	Data Training	Master	Hard Disk	Index	Id_training, Id_user, Id_pilihan
I-004	Data Testing	Master	Hard Disk	Index	Id_testing, Id_user, Id_pilihan

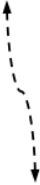
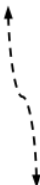
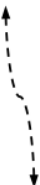
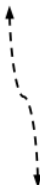
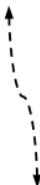
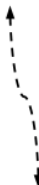
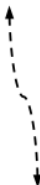
4.13. Desain Sistem Secara Terinci

Tabel 4.13 Rancangan Output Data Prediksi

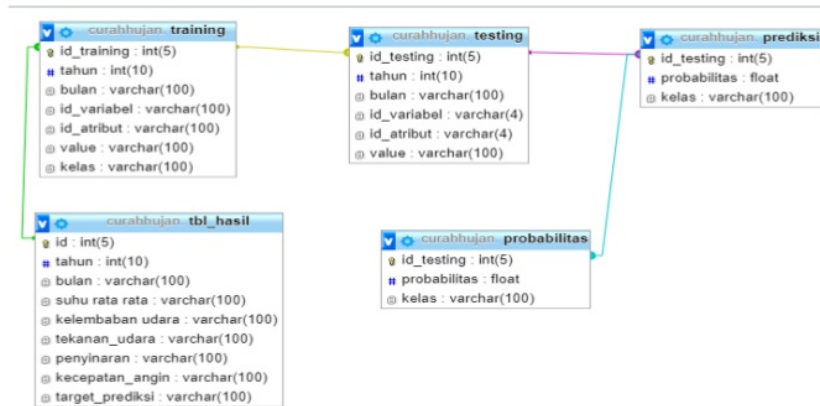


BMKG STASIUN KLIMATOLOGI TILONGKABILA

DATA PREDIKSI

ID Testing	Suhu Rata- Rata	Kelembaban udara	Kecepatan angin	Penyinaran	Tekanan udara	prediksi
X(4)	X(100)	X(100)	X(100)	X(100)	X(100)	X(100)
						

#### 4.14. Data Desain : Relasi Tabel



Gambar 4.15 relasi tabel

#### 4.15. Program Designs

**Tabel 4.14** Hasil Desain Sistem

Class Type	Attributes[Type]	Methods[Event or Type]
Menu Utama	Home[Menu]	Home[Click]
	Data[Menu]	Data[Click]
	Hasil[Menu]	Hasil[Click]
	Add[Toolbar]	Add[Click]
	Delete[Toolbar]	Delete[Click]
	Save[Toolbar]	Save[Click]
	Cancel[Toolbar]	Cancel[Click]
Login	Username[Textbox]	Username[Textbox]
	Password[Textbox]	Password[Textbox]
	Login[Button]	Login[Click]
Menu Input Data Traning	Item Data[Combobox]	Item Data[Click]
	View data[Grdview]	View data[Click]
Menu Input Data Testing	Item data[Combobox]	View data[Click]
Menu Hasil prediksi	View data[Gridview]	View Hasil prediksi[Click]

#### 4.16. Hasil Konstruksi Sistem

Pada tahap konstruksi sistem ini hasil dari analisis dan desain sistem kemudian diterjemahkan ke konstruksi sistem/*software* dengan menggunakan bahasa pemrograman PHP. Adapun alat bantu yang digunakan pada tahap ini adalah:

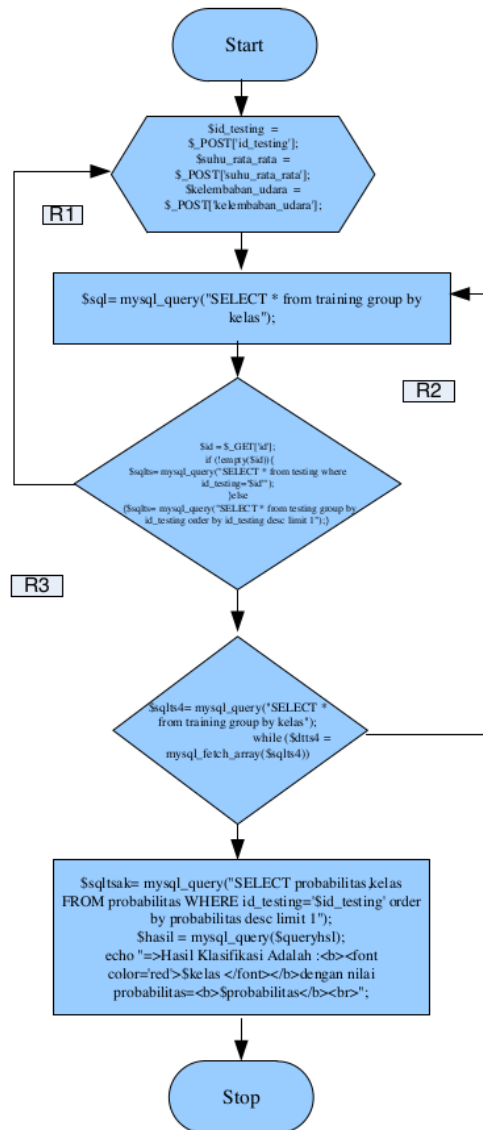
1. **PHP** untuk pemrogramannya;
2. **Mysql** untuk databasenya;
3. **Notepad ++** untuk editor webnya:

#### 4.17. Pengujian WhiteBox

```
$id_testing = $_POST['id_testing']; .....1
$jenis_produk = $_POST['jenis_produk']; .....1
$nama_produk = $_POST['nama_produk']; .....1
$query = "INSERT INTO testing (id_testing, jenis_produk, .....1
nama_produk, id_variabel, id_atribut, value) VALUES .....1
('$id_testing', '$jenis_produk', '$nama_produk', .....1
'id_variabel2', '$nilai32[$i]', '$nilai32[$i]')"; .....1
$sql= mysql_query("SELECT * from training group by kelas"); .....2
$id = $_GET['id']; .....3
if (!empty($id)){ .....3
    $sqlts= mysql_query("SELECT * from testing .....3
    where id_testing='$id'"); .....3
} else .....3
{ $sqlts= mysql_query("SELECT * from testing group by id_testing .....3
order by id_testing desc limit 1"); } .....3
$sqlts4= mysql_query("SELECT * from training group by kelas"); .....4
while ($dtts4 = mysql_fetch_array($sqlts4)) .....4
{ .....4
    $kelas=$dtts4['kelas']; .....4
    $hasil=$tprob1[$k]*$xyz[$k]; .....4
    echo ">$kelas(($tprob1[$k]*$xyz[$k])=$hasil<br>"; .....4
    $query = "INSERT INTO probabilitas .....4
    (id_testing, probabilitas, kelas) VALUES('$id_testing', '$hasil', '$kelas')"; .....4
    $hasil = mysql_query($query); .....4
    $k=$k+1; .....4
} .....4
$sqltsak= mysql_query("SELECT probabilitas, kelas .....5
FROM probabilitas WHERE id_testing='$id_testing' order .....5
by probabilitas desc limit 1"); .....5
$hasil = mysql_query($queryhs1); .....5
```

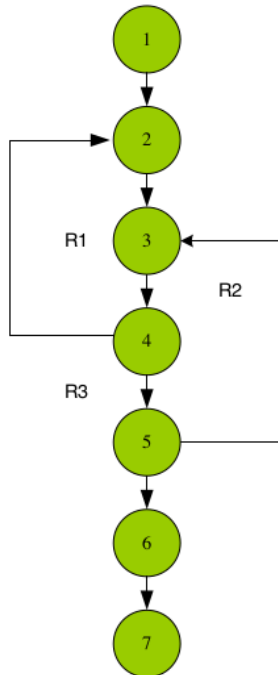
echo "=>Hasil Klasifikasi Adalah :<b><font color='red'>\$kelas .....5  
</font></b>dengan nilai probabilitas=<b>\$probabilitas</b><br>";.....5

#### 4.17.1 Flowchart Pengujian Whitebox



**Gambar 4.16** Flowchart program untuk pengujian White Box

#### 4.17.2. Flowgraph Pengujian Whitebox



**Gambar 4.17** Flowgraph pengujian White Box

Perhitungan CC pengujian White Box

Edge (E) : 8

Predikat Node (P) : 2 + 1

Rumus :  $V(G) = (E - N) + 2$

Atau  $V(G) = P + 1$

Penyelesaian :  $V(G)$  :  $(8 - 7) + 2 = 3$

$V(G)$  :  $2 + 1 = 3$

(R1,R2,R3)

#### 4.17.3. path pengujian White Box

**Tabel 4.15** Path pada pengujian White Box

NO	PATH	KET
1	1-2-3-4-3 .....7	OK
2	1-2-3-4-5-3 .....7	OK
3	1-2-3-4-5-6-7	OK

#### 4.18. Hasil pengujian black box

**Tabel 4.16** Hasil pengujian Black Box

NO	INPUT/ EVENT	FUNGSI	Hasil	Hasil
1	Login	Memasukan Username dan password	<ul style="list-style-type: none"><li>- Jika Password salah maka tidak masuk ke halaman menu utama</li><li>- Jika Password benar, maka tutup login dan masuk ke menu utama</li></ul>	Sesuai
2	Menu Data Variabel	Menampilkan Form data Variabel	Data Variabel berhasil ditampilkan	Sesuai
3	Menu Data Atribut	Menampilkan Form data Atribut	Data Atribut berhasil ditampilkan	Sesuai
4	Menu Data Training	Menampilkan Form data training	Data Training berhasil ditampilkan	Sesuai
5	Button Input Data Training	Memasukan Data Training	Data Training berhasil dimasukan	Sesuai
6	Button Update Data	Mengubah Data	Data Training berhasil	Sesuai

	Training	Training	diubah	
7	Menu Data Testing	Menampilkan Form Data Testing	Data Testing berhasil ditampilkan	Sesuai
8	Button Input Data Testing	Memasukkan Data testing	Data Testing berhasil dimasukan	Sesuai
9	Medenu Data User	Menampilkan Form User	Form user berhasil ditampilkan dan aktif	Sesuai
10	Button Input Data User	Memasukan Data User	Data User berhasil dimasukkan	Sesuai
11	Button Update Data User	Mengubah Data User	Data User berhasil diubah	Sesuai
12	Menu Hasil Prediksi	Menampilkan Form hasil prediksi curah hujan	Form hasil prediksi berhasil ditampilkan	Sesuai
13	Hapus Hasil Prediksi	Menghapus hasil prediksi curah hujan	Hasil prediksi berhasil dihapus	Sesuai

## BAB V

### PEMBAHASAN

#### 5.1 Pembahasan model

**Tabel 5.1** pengujian confusion matriks

No	Aktual	prediksi
1	hujan ringan	hujan ringan
2	hujan ringan	tidak hujan
3	hujan lebat	hujan ringan
4	hujan lebat	hujan lebat
5	hujan lebat	tidak hujan
6	hujan ringan	hujan ringan
7	tidak hujan	tidak hujan
8	hujan lebat	hujan lebat
9	tidak hujan	tidak hujan
10	hujan lebat	hujan ringan

**Confusion matriks:**

Hitung Akurasi :

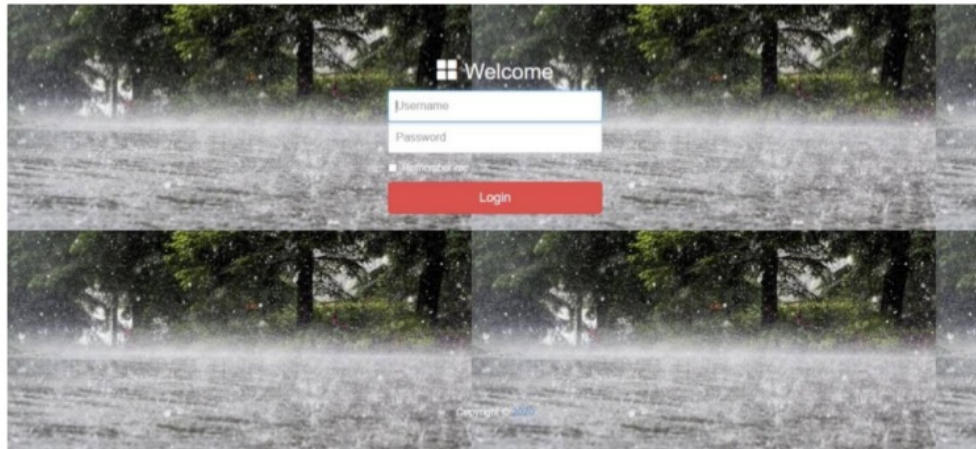
	Hujan Lebat	Hujan Ringan	Tidak Hujan
Hujan Lebat	a=2	b=1	c=1
Hujan Ringan	d=0	e=3	f=1
Tidak Hujan	g=0	h=0	i=2

**Gambar 5.1** Hitung Akurasi

$$\begin{aligned}
 \text{Hitung Akurasi} &= (a + e + i) / (a+b+c+d+e+f+g+h+i) \\
 &= (2+3+2) / (2+1+1+0+3+1+0+0+2) \\
 &= 70\%
 \end{aligned}$$

## 5.2 Pembahasan sistem

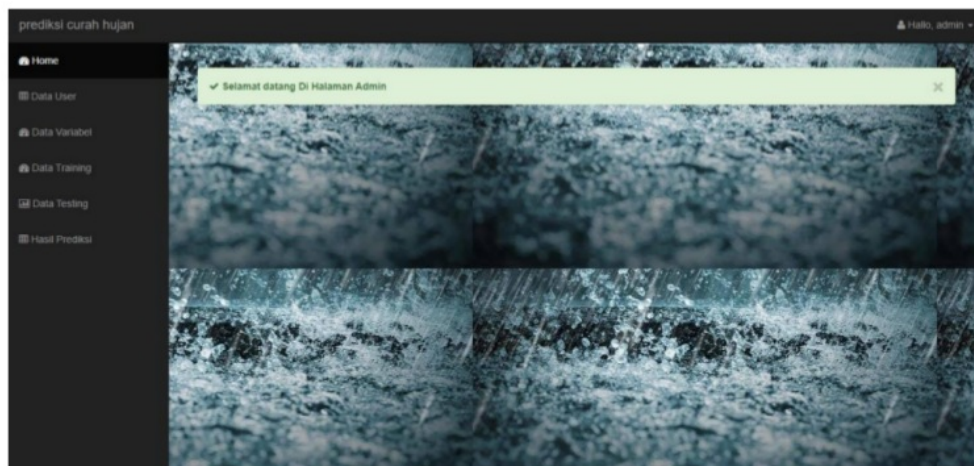
### 5.2.1. Tampilan login admin



**Gambar 5.2** Tampilan login admin

Tampilan login admin untuk masuk ke program prediksi curah hujan di input username dan password dan setelah di ketik username dan password lalu klik login.

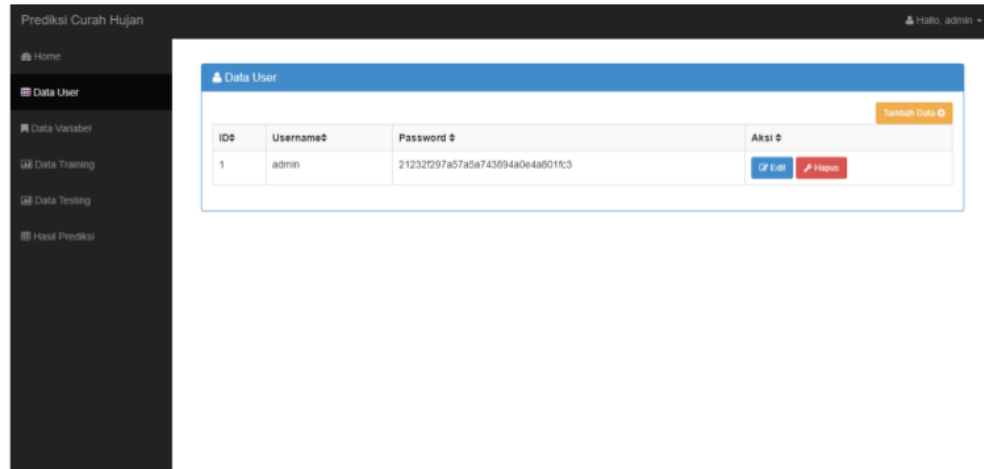
### 5.2.2 Tampilan index admin



**Gambar 5.3** Tampilan index admin

Tampilan index admin ini tampilan home prediksi curah hujan yang terdiri dari Data variabel, Data training, data testing, hasil prediksi dan di samping kanan atas ada halo admin.

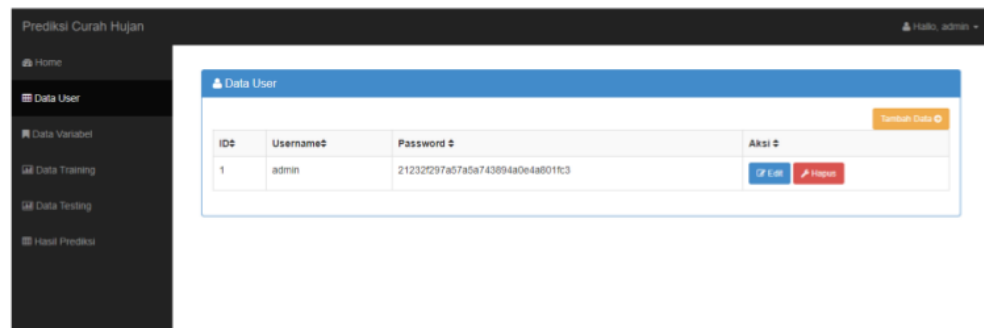
### 5.2.3 Tampilan Tabel user



**Gambar 5.4** Tampilan tabel user

Tampilan tabel user ini terdiri dari id , username ,password dan aksi jika kita ingin mengedit data bisa klik edit dan jika menghapus klik hapus.

### 5.2.4 Tampilan Tambah User



**Gambar 5.5** Tampilan Tambah User

Tampilan Tambah user untuk menambah user dengan cara mengisi id user,username,Password,lalu klik simpan data dan datangnya langsung tersimpan.

### 5.2.5 Tampilan Data Variabel

prediksi

Home

Data User

**Data Variabel**

Data Training

Data Testing

Hasil Prediksi

#### DATA VARIABEL

Aksi	Id Variabel	Nama Variabel	Nama Atribut
<input checked="" type="checkbox"/> <input type="checkbox"/>	VR01	Suhu Rata Rata	<25 >26 <a href="#">[?]</a> <a href="#">[tambah]</a>
<input checked="" type="checkbox"/> <input type="checkbox"/>	VR02	kelembaban udara	<75 >80 <a href="#">[?]</a> <a href="#">[tambah]</a>
<input checked="" type="checkbox"/> <input type="checkbox"/>	VR03	tekanan udara	rendah sedang tinggi <a href="#">[?]</a> <a href="#">[tambah]</a>
<input checked="" type="checkbox"/> <input type="checkbox"/>	VR04	penyinaran	rendah sedang tinggi <a href="#">[?]</a> <a href="#">[tambah]</a>
<input checked="" type="checkbox"/> <input type="checkbox"/>	VR05	kecepatan angin	rendah sedang tinggi <a href="#">[?]</a> <a href="#">[tambah]</a>

**Gambar 5.6** Tampilan Data Variabel

Tampilan Data Variabel prediksi curah hujan terdiri dari aksi ,id variabel, nama variabel dan nama atribut .

### 5.2.6 Tampilan Tambah Variabel

prediksi

Home

Data User

**Data Training**

Data Testing

Hasil Prediksi

id Variabel

VR06

Nama Variabel

Nama Variabel

Cancel

Simpan

**Gambar 5.7** Tampilan Tambah variabel

Tampilan tambah variabel terdiri dari id variabel , nama variabel, jika ingin menambah variabel di isi terlebih dahulu id dan nama variabel lalu klik simpan.

### 5.2.7 Tampilan Tabel Data Training

prediksi curah hujan

Home

Data User

Data Variabel

Input Training

**Data Training**

Data Testing

Hasil Prediksi

**Tabel Training**

Keterangan : VR01 = Suhu rata-rata VR02 = Kelembaban udara VR03 = Tekanan udara VR04 = Penyinaran VR05 = Kecepatan angin  
Tabel "Training"

NO	Tahun	Bulan	VR01	VR02	VR03	VR04	VR05	Kelas	Pilihan
1	2015	Januari	>25	>80	sedang	rendah	tinggi	hujan sedang	<a href="#">Edit</a> <a href="#">Hapus</a>
2	2015	februari	<25	>80	rendah	sedang	rendah	hujan sedang	<a href="#">Edit</a> <a href="#">Hapus</a>
3	2015	maret	<25	>80	rendah	sedang	sedang	hujan sedang	<a href="#">Edit</a> <a href="#">Hapus</a>
4	2015	april	>25	>80	sedang	sedang	tinggi	hujan tinggi	<a href="#">Edit</a> <a href="#">Hapus</a>
5	2015	mei	>25	>80	tinggi	rendah	sedang	hujan sedang	<a href="#">Edit</a> <a href="#">Hapus</a>

**Gambar 5.8** Tampilan Tabel Training

Tampilan data training terdiri dari No,Tahun,Bulan,VR01,VR02,VR03,VR04,VR05, Kelas,pilihan hapus dan edit .

### 5.2.8 Tampilan <sup>2</sup>Input Data Tranning

prediksi

Home

Data User

Data Variabel

**Input Training**

Data Training

Data Testing

Hasil prediksi

**Input Data Training**

Id Training

Tahun

Bulan

Suhu Rata Rata

kelembaban udara

tekanan udara

penyinaran

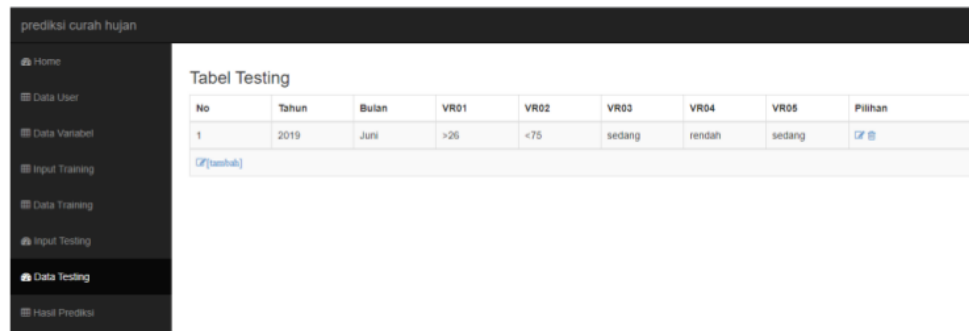
kecepatan angin

Class

**Gambar 5.9** Tampilan **Input Data** Tranning

Tampilan input data tranning, jika ingin membuat data tranning baru di isi,Id training, Tahun, Bulan, Suhu rata-rata, kelembaban udara, tekanan udara, penyinaran, kecepatan angin,class , lalu klik simpan dan data training langsung terinput.

### 5.2.9.Tampilan Tabel Data Testing



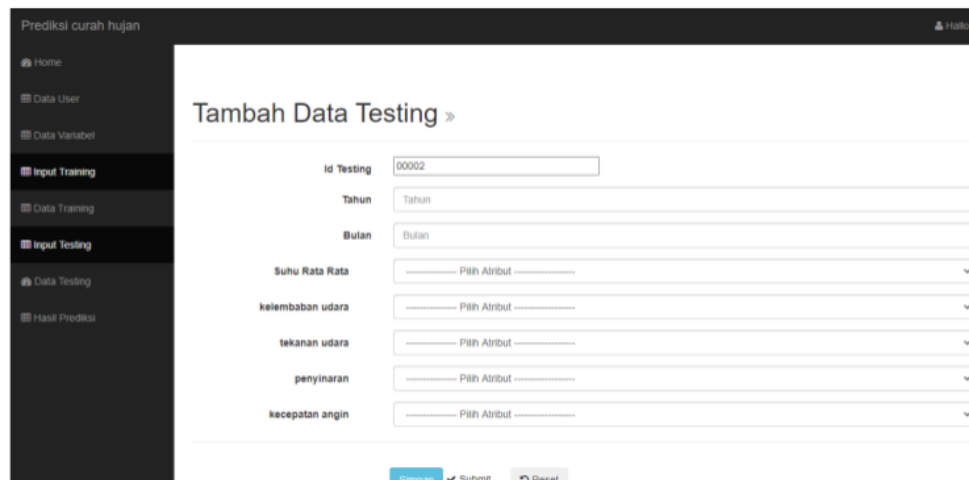
No	Tahun	Bulan	VR01	VR02	VR03	VR04	VR05	Pilihan
1	2019	Juni	>26	<75	sedang	rendah	sedang	

[\[Tambah\]](#)

**Gambar 5.10** Tampilan Tabel Data Testing

Tampilan Data testing terdiri dari , no ,tahun, bulan,vr01,vr02,vr03,vr04,vr05, dan Pilihan terdapat edit dan hapus.

### 5.2.10.Tampilan Tambah Data Testing



Prediksi curah hujan

Halo

### Tambah Data Testing »

Id Testing:

Tahun:

Bulan:

Suhu Rata Rata:

kelembaban udara:

tekanan udara:

penyinaran:

kecepatan angin:

**Gambar 5.11** Tampilan Tambah Data Testing

Tampilan Tambah Data Testing yang terdiri dari id testing,tahun,bulan,Suhu rata-rata, kelembaban udara,tekanan udara,penyinaran,kecepatan angin,jika Ingin menambah data testing di isi semua atribut lalu klik simpan langsung tersimpan ke data testing.

### 5.2.11. Tampilan Hasil Prediksi

NO	Tahun	Bulan	VR01	VR02	VR03	VR04	VR05	Hasil			
1	2019	maret	>26	>80	Rendah	Sedang	Sedang	Hasil hujan			
2	2019	april	<26	>80	Sedang	Sedang	Sedang	Hasil hujan			
3	2019	mei	>26	>80	Tinggi	Rendah	Sedang	Hasil hujan			
4	2019	Juni	>26	>80	Sedang	Rendah	Sedang	Hasil hujan			

**Gambar 5.12** Hasil Prediksi Curah Hujan

Tampilan di gambar 5.12 adalah tampil hasil prediksi curah hujan menggunakan algoritma naïve bayes.

### 5.2.12 Tampilan pengujian

prediksi curah hujan

Home

Data User

Data Variabel

Data Training

Data Testing

Hasil Prediksi

Pengujian

Pengujian

Nilai Aktual

Nilai Prediksi

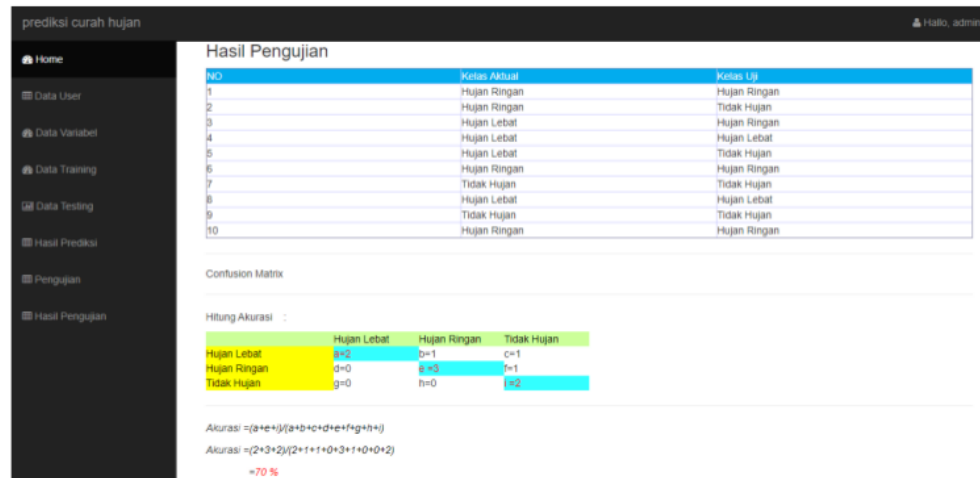
Hapus Form

Simpan Data

**Gambar 5.13** tampilan pengujian

Tampilan pengujian jika Ingin menguji data prediksi, di isi terlebih dahulu nilai aktual dan nilai prediksi lalu klik simpan data. Dan data yang di uji langsung terinput hasilnya di hasil pengujian.

### 5.2.13 Tampilan Hasil pengujian



**Gambar 5.14** Tampilan Hasil pengujian

5 Tampilan Hasil pengujian prediksi curah hujan dengan perhitungan confusion matriks

## BAB VI

### KESIMPULAN DAN SARAN

#### 6.1. Kesimpulan

Dari hasil pengujian yang telah dilakukan dapat disimpulkan sebagai berikut :

1. Keakuratan yang diperoleh untuk prediksi curah hujan di kota gorontalo menggunakan metode Naïve bayes dapat memperoleh hasil tingkat akurasi sebesar 70%.
2. Kinerja yang diperoleh dengan memprediksi curah hujan dapat diketahui dengan pengujian sistem white box  $V(G) = 3$ , sehingga system ini dapat di implementasikan pada prediksi curah hujan.

#### 6.2. Saran

Dari hasil penelitian pembahasan dan analisa pada penelitian, maka peneliti memberikan Saran untuk peneliti selanjutnya yang diharapkan bisa dijadikan dasar untuk mengembangkan penelitian ini yaitu:

1. perlu dilakukan pengujian ulang metode yang sama tapi dengan jumlah data yang lebih banyak pada penelitian ini, agar akurasi lebih akurat
2. Diharapkan adanya penelitian pada tahap selanjutnya yaitu menggunakan metode lain atau menggabungkan beberapa metode sehingga diharapkan dapat menutup kekurangan metode yang telah di gunakan sebelumnya .

## DAFTAR PUSTAKA

- [1] Prawirowardoyo, S. 1996. *Meteorologi*. Bandung: ITB.
- [2] Lakitan. Benyamin, 1997, *Dasar Dasar Klimatologi*. Raja Grafindo Persada, Jakarta.
- [3] Naisha Rahma Indraswari, 2018. *Aplikasi Prediksi Usia Kelahiran dengan Metode Naïve Bayes*. Skripsi Universitas Muhammadiyah Surakarta.
- [4] Syarli dan Asrul Ashari Muin, 2016. *Metode Naïve Bayes untuk Prediksi Kelulusan*. Makassar.
- [5] Diera Desmonda, Turzina & M Azhar Irwansyah, 2018. *Prediksi Besaran Hujan Menggunakan Metode Fuzzy Time Series*. Jurnal Sistem dan Teknologi Informasi Vol. 6, No. 4. Oktober 2018.
- [6] Satriogi Putramulyo & Siti Alaa, 2018. *Prediksi Curah Hujan Bulanan di Kota Samarinda menggunakan Persamaan Regresi dengan Prediktor Data Suhu Udara dan Kelembapan Udara*. Eigen Mathematic Journal. Vol. 1, No. 2. Desember 2018.
- [7] Yulia Fatmi, 2019. *Prediksi Pembukaan Peluang Usaha di Kecamatan Medan Amplas menggunakan Naïve Bayes*. Medan.
- [8] Soewarno. 2015. *Klimatologi, Pengukuran dan Pengolahan Data Curah Hujan, Contoh Aplikasi Hidrologi dalam Pengelolaan Sumber Daya Air*. Yogyakarta : Graha Ilmu.
- [9] Triadmojo, Bambang. 2008. *Hidrologi Terapan*. Yogyakarta : Beta Offset.
- [10] Rahmawati et all, 2014. *Identifikasi Curah Hujan Ekstrem Di Kota Semarang Menggunakan Estimasi Parameter Momen Probabilitas Terboboti Pada Nilai Ekstrem Terampat* (Studi Kasus Data Curah Hujan Dasarian Kota Semarang Tahun 1990-2013). Departemen Statistika FSM Undip, 2014.
- [11] Prasetyo, E., 2006, *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*, Andi Yogyakarta. Yogyakarta.

- [12] Han, J, Kamber M., 2006, *Data Mining: Concepts and Techniques*, Second Edition. Morgan Kaufman. California.
- [13] Hoffer, Jeffrey A., Ramesh, V., and Topi, Heikki. 2011. *Modern Database Management 10<sup>th</sup> Edition*. New Jersey: Pearson Education.
- [14] Witten, I.H. and Frank, E. 2005. *Data Mining Practical Machine Learning Tools and Techniques*. Second Edition. California: Morgan Kaufman.
- [15] David, Olson & Yong, Shi. *Introduction to Business Data Mining*. 2011. International Edition: Mc Graw Hill.
- [16] <sup>1</sup> Patil, T. R., Sherekar, M. S., 2013, Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification, *International Journal of Computer Science and Applications*, Vol. 6, No. 2, Hal 256-261.
- [17] Bustami., 2013, Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi, *TECHSI : Jurnal Penelitian Teknik Informatika*, Vol. 3, No.2, Hal. 127-146.
- [18] Pattekari, S. A., Parveen, A., 2012, Prediction System for Heart Disease Using Naïve Bayes, *International Journal of Advanced Computer and Mathematical Sciences*, ISSN 2230-9624, Vol. 3, No 3, Hal 290-294.
- [19] Nasari, F., 2014, *Analisa Faktor Penyebab Tingginya Pemakaian Listrik Rumah Tangga Menggunakan Algoritma C 4.5* (Studi Kasus di Kelurahan Tanjung Mulia), *Tesis*, Program Pasca Sarjana Ilmu Komputer, Universitas Putra Indonesia (YPTK), Padang.
- [20] Sutarbi, Tata. 2013. *Analisis Sistem Informasi*. Yogyakarta: Andi.
- [21] Witten, Jeffrey L, et all, *Metode Desain & Analisis Sistem*, Edisi 6, Edisi International, Mc Graw Hill, Andi, Yogyakarta: 2004.
- [22] Gujarati, N.D. 2003. *Basic Econometrics*. 4th ed. New York: McGraw-Hill Companies, Inc.
- [23] Kadir, Abdul. 2003. *Pengenalan Sistem Informasi*. Edisi I. Yogyakarta.Andi Yogyakarta.

- [24] Jogyanto, HM. 2005. *Analisis dan Desain Sistem Informasi : Pendekatan Terstruktur Teori dan Praktek Aplikasi Bisnis*. Edisi II. Yogyakarta : Andi Yogyakarta.
- [25] Hariyanto, Bambang, 2004. *Sistem Informasi Basis Data: Pemodelan, Perancangan, dan Terapannya*. Informatika, Bandung.
- [26] Pressman, R.S. 2002. *Rekayasa Perangkat Lunak : Pendekatan Praktis (Buku I)*. Yogyakarta : Andi Yogyakarta.

# PREDIKSI CURAH HUJAN DI KOTA GORONTALO MENGUNAKAN ALGORITMA NAIVE BAYES

## ORIGINALITY REPORT

25%

SIMILARITY INDEX

22%

INTERNET SOURCES

7%

PUBLICATIONS

20%

STUDENT PAPERS

## PRIMARY SOURCES

1

[www.scribd.com](http://www.scribd.com)

Internet Source

8%

2

Submitted to LL Dikti IX Turnitin Consortium

Student Paper

5%

3

[eprints.ums.ac.id](http://eprints.ums.ac.id)

Internet Source

2%

4

[rijjasihabuddin.blogspot.com](http://rijjasihabuddin.blogspot.com)

Internet Source

1%

5

Submitted to Sriwijaya University

Student Paper

1%

6

[repository.widyatama.ac.id](http://repository.widyatama.ac.id)

Internet Source

1%

7

[id.123dok.com](http://id.123dok.com)

Internet Source

1%

8

[kingarthur38.files.wordpress.com](http://kingarthur38.files.wordpress.com)

Internet Source

1%

9

[jurnal.untan.ac.id](http://jurnal.untan.ac.id)

Internet Source

1 %

10

Submitted to Universitas Islam Indonesia

Student Paper

1 %

11

widuri.raharja.info

Internet Source

1 %

12

eprints.umpo.ac.id

Internet Source

<1 %

13

titonkadir.blogspot.com

Internet Source

<1 %

14

lppm.unram.ac.id

Internet Source

<1 %

15

Submitted to UPN Veteran Yogyakarta

Student Paper

<1 %

16

pastebin.com

Internet Source

<1 %

17

Submitted to Lambung Mangkurat University

Student Paper

<1 %

18

library.binus.ac.id

Internet Source

<1 %

19

eltikom.poliban.ac.id

Internet Source

<1 %

20

Submitted to Universitas Brawijaya

Student Paper

<1 %

21	Submitted to Universitas Muria Kudus Student Paper	<1 %
22	eprints.undip.ac.id Internet Source	<1 %
23	Submitted to Universitas Dian Nuswantoro Student Paper	<1 %
24	ejournal.bsi.ac.id Internet Source	<1 %
25	eprints.akakom.ac.id Internet Source	<1 %
26	Budy Santoso, Azminuddin I. S. Azis, Andi Bode. "Pengendalian Lampu Lalu Lintas Cerdas di Persimpangan Empat Ruas yang Kompleks Menggunakan Algoritma Adaptive Neuro Fuzzy Inference System", Jurnal Edukasi dan Penelitian Informatika (JEPIN), 2020 Publication	<1 %

Exclude quotes Off  
Exclude bibliography On

Exclude matches < 25 words