

# Aplikasi Berbasis Web Untuk Deteksi Bahasa Daerah Di Indonesia Dengan Naïve Bayes

**Guruh Agung Prasetyo<sup>1)</sup>, M.Taufik, ST.MIT<sup>2)</sup>, Sam Farisa c,ST.M.Kom<sup>3)</sup>**

<sup>1)</sup>Mahasiswa Program Studi Teknik Informatika, UNISSULA

<sup>2)</sup>Dosen Pembimbing Program Studi Teknik Informatika, UNISSULA

<sup>3)</sup>Dosen Pembimbing Program Studi Teknik Informatika, UNISSULA

**Abstrak** - Deteksi bahasa daerah merupakan pendeteksi arti bahasa daerah ke bahasa lainnya. Adapun Bahasa yang akan di artikan daerah Sulawesi (Dialek Bugis), Jawa (Dialek Tegal-Banyumas), Kalimantan (Dialek Melayu) yang nantinya akan di artikan ke Bahasa Indonesia atau Inggris begitupun sebaliknya. Pengembangan aplikasi ini menggunakan metode Naive Bayes yaitu dengan model Web Engineering (Rekayasa Web). Tahapan aktifitas Web Engineering meliputi Customer Communication, Planning, Modeling, Construction, Delivery & Feedback. Naive Bayes memiliki tahapan klasifikasi dengan metode probabilitas dan statistik, sehingga jika diasumsikan kondisi antar atribut saling bebas bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. Hasil yang diperoleh pada pengembangan Aplikasi Deteksi Bahasa berbasis Website menggunakan metode Naive Bayes ini terbukti dapat mengatasi masyarakat yang tidak paham bahasa daerah dialek Bugis,dialek Tegal-Banyumas,dialek Melayu dan membantu komunikasi saat berpergian ke daerah itu.

**Kata Kunci** : Deteksi bahasa, Naïve bayes, Website

**Abstract** - Detection of regional languages is a detection of the meaning of regional languages into other languages. The languages that will be interpreted as Sulawesi (Bugis Dialect), Java (Tegal-Banyumas Dialect), Kalimantan (Malay Dialect) which will later be interpreted into Indonesian or English and vice versa. The development of this application uses the Naive Bayes method, namely the Web Engineering model. The stages of Web Engineering activities include Customer Communication, Planning, Modeling, Construction, Delivery & Feedback. Naive Bayes has a classification stage with probability and statistical methods, so if it is assumed that the conditions between attributes are independent of whether or not certain features of a class have nothing to do with the characteristics of other classes. The results obtained in developing a Website-based Language Detection Application using the Naive Bayes method have proven to be able to overcome people who do not understand the language of the Bugis dialect, the Tegal-Banyumas dialect, the Malay dialect and assist in communication when traveling to the area.

**Key words:** Language detection, Naïve Bayes, Website

## I. PENDAHULUAN

Google translate merupakan salah satu platform layanan yang berhubungan dengan terjemahan bahasa, namun tidak terdapat layanan untuk menerjemahkan bahasa daerah di tiap negara. Terdapat lebih dari 100 bahasa yang ada pada terjemahan google. Dalam platform google translate pun beberapa bahasa daerah di berbagai Negara juga belum tercakup. Jika ada turis asing yang datang ke sebuah kota di Negara tertentu dan tidak paham mengenai logat bahasa daerah tersebut maka akan kesulitan untuk berbicara ataupun melakukan sosialisasi.

Di Indonesia terdapat 5 pulau besar dan tiap tiap pulau memiliki dialeg yang berbeda beda. Beberapa contoh dialeg bahasa yang umum digunakan di tiap pulau misalnya Sulawesi ( Dialek Bugis), Jawa (Dialek Tegal – Banyumas.), Kalimantan ( Dialek Melayu ).

Metode *Naive Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya

## II. TINJAUAN PUSTAKA

### 2.1 Metodologi

Metodologi merupakan suatu formula dalam penerapan penelitian dimana dalam melakukan penelitian tersebut terdapat langkah-langkah dan juga hasil penelitian. Sedangkan metodologi penelitian dalam ilmu komputer atau sistem informasi atau teknologi informasi merupakan langkah – langkah atau tahapan perencanaan dengan

bantuan beberapa metode ,teknik ,alat (*tools*) dan dokumentasi dengan tujuan untuk membantu peneliti dalam meminimalkan resiko kegagalan dan menekankan pada proses atau sasaran penelitian [1].

## **2.2 Sistem Pendukung Keputusan (SPK)**

Sistem Pendukung Keputusan (SPK) atau sering disebut Decision Support System (DSS) adalah Sistem berbasis model yang terdiri dari prosedur - prosedur dalam pemrosesan data dan pertimbangannya untuk membantu dalam mengambil keputusan. SPK bertujuan untuk menyediakan informasi, membimbing, memberikan prediksi serta mengarahkan kepada pengguna informasi agar dapat melakukan pengambilan keputusan dengan lebih baik [2].

## **2.3 Sistem informasi**

Sistem informasi adalah suatu sistem di dalam suatu organisasi yang mempertemukan kebutuhan pengolahan transaksi harian yang mendukung fungsi organisasi yang bersifat manajerial dalam kegiatan strategi dari suatu organisasi untuk dapat menyediakan kepada pihak luar tertentu dengan laporan – laporan yang diperlukan [3].

## **2.4 Naïve Bayes**

Metode Naive Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya [4].

### **III. METODE PENELITIAN**

#### **3.1 Metode ADDIE**

Teknik Analisa Data merupakan bagian yang sangat penting dalam penelitian. Pemecahan masalah penelitian dan penarikan kesimpulan dari suatu penelitian sangat tergantung dari hasil analisa data. Pada tahap ini akan dilakukan analisa dengan metode ADDIE ( *Analysis, Design, Development, Implementation and Evaluation* ).

1. Analisis

Dalam tahap ini akan dilakukan analisa untuk kebutuhan sistem yang akan dibuat. Dalam tahap ini pula telah ditentukan variabel variabel yang terkait dan pengguna yang dapat menggunakan sistem.

2. Design Sistem

Dalam tahap ini setelah kebutuhan sistem di analisa akan dibuat perancangan proses sistem dan tampilan atau *Graphic User Interface* untuk sistem yang akan digunakan.

3. Pengembangan / *Development* Sistem

Pada tahap ini bahan dan data – data yang diperlukan untuk membuat sistem seperti data data variabel, kriteria tiap variabel dan rumus rumus yang diperlukan untuk dirubah ke kode PHP.

4. Implementasi Sistem

Dalam tahap ini akan dilakukan *deployment* terhadap analisa kebutuhan dan data untuk deteksi bahasa.

5. Evaluasi Sistem

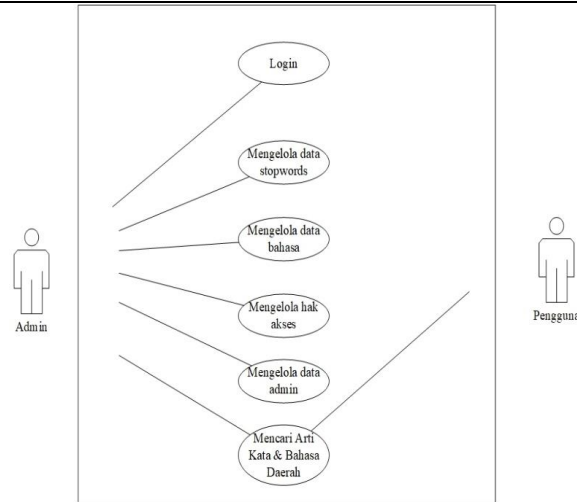
Tahap ini bertujuan untuk melihat sejauh mana aplikasi yang dibuat dapat mencapai sasaran dan tujuan yang sudah ditetapkan sebelumnya.

### **IV. HASIL DAN PEMBAHASAN**

#### **4.1 Membuat pemodelan data & perancangan database**

Pada tahap ini berisi mengenai perancangan, pembuatan atau perancangan pemodelan data, perancangan database.

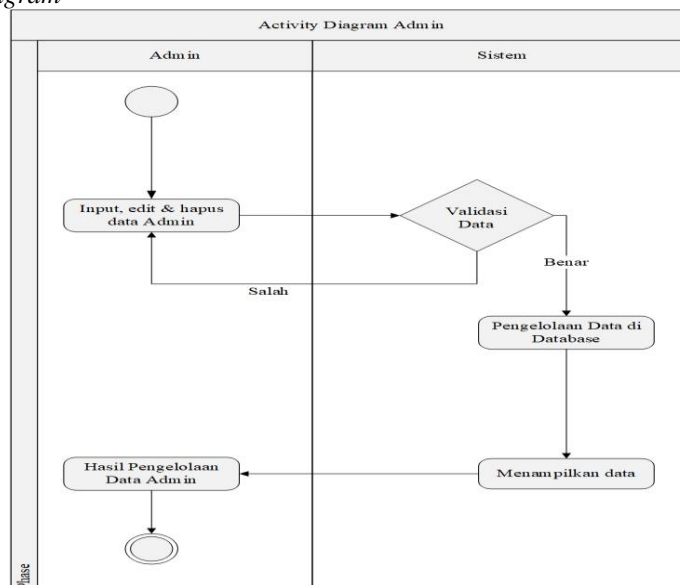
##### **4.1.1 Usecase Diagram**



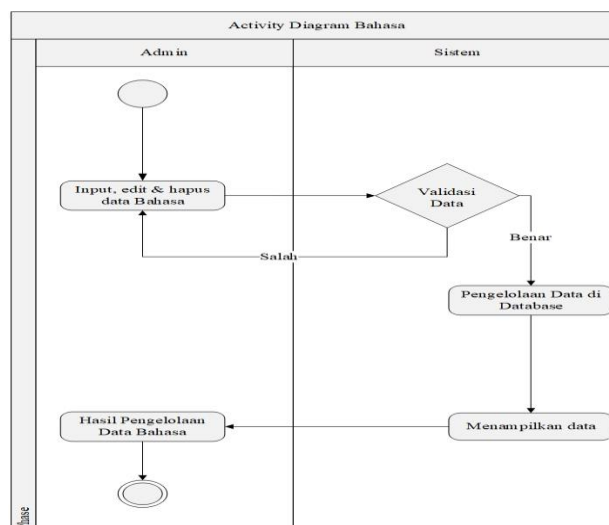
Gambar 4.1 use case diagram

Pada gambar 4.1 adalah gambar *use case diagram* yang menggambarkan tentang apa saja akses antara pengguna serta admin yang akan mengelola sistem.

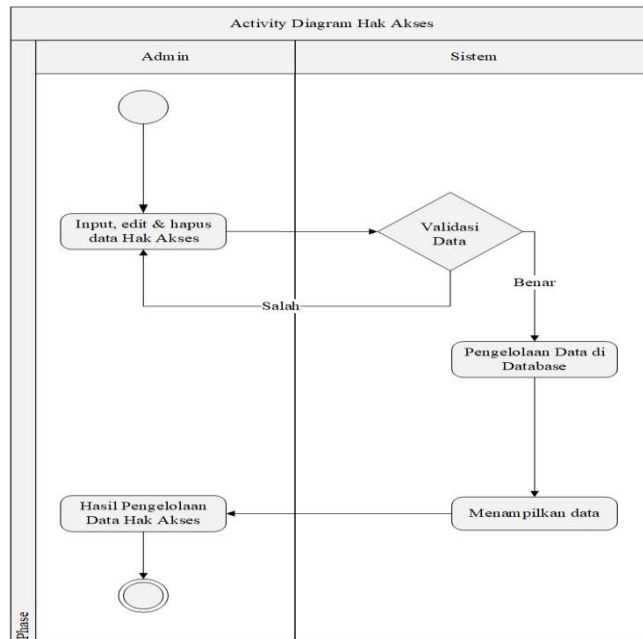
#### 4.1.2 Activity Diagram



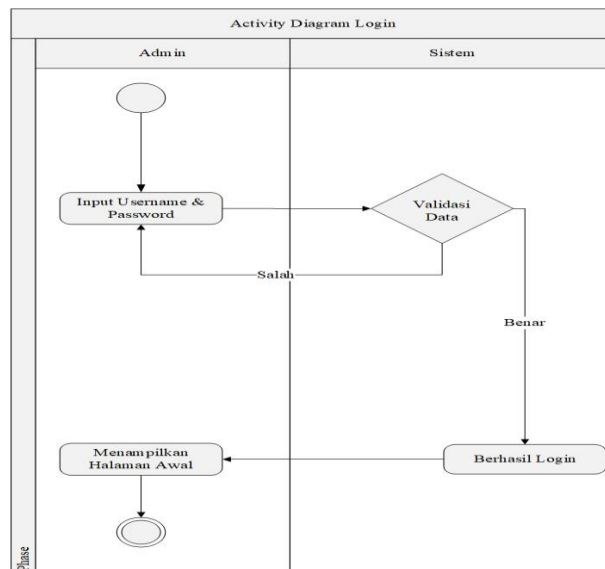
Gambar 4.2 Activity diagram pengelolaan data admin



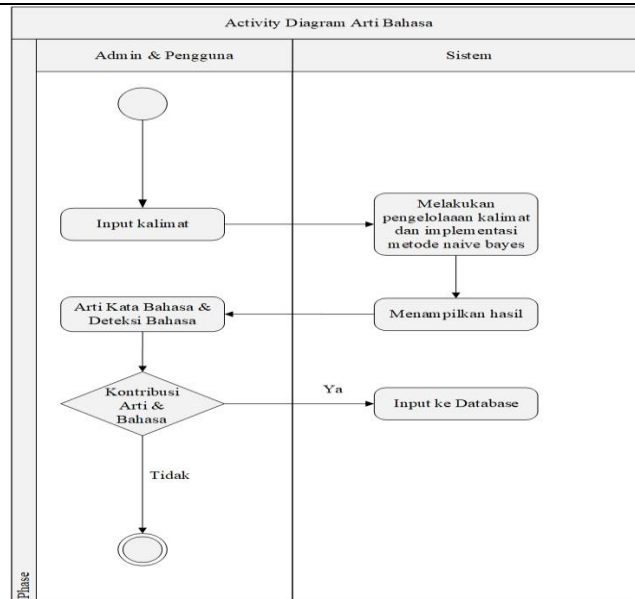
Gambar 4.3 Activity diagram pengelolaan data bahasa



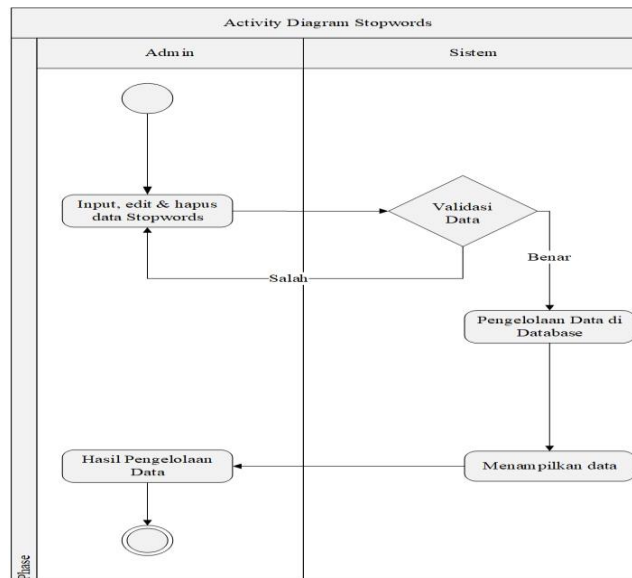
Gambar 4.4 Activity Diagram pengelolaan data hak akses



Gambar 4.5 Activity Diagram Login

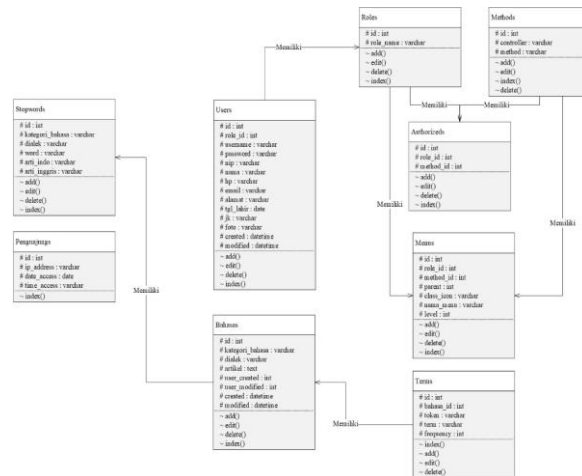


Gambar 4.6 Activity Diagram pencarian bahasa



Gambar 4.7 Activity Diagram pengelolaan Stopwords

### 4.1.3 Class Diagram



Gambar 4.8 Class Diagram

Pada gambar 4.8 merupakan gambar perancangan *class diagram* yang akan dibuat dalam sistem, serta pada gambar tersebut mewakili relasi *database* yang akan dibuat.

### 4.2 Desain tabel *database*

Tabel 4.1 Tabel Authorizeds

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	Primary Key
2	Role_id	Int(10)	Foreign Key
3	Method_id	Int(10)	Foreign Key

Pada tabel 4.1 yaitu tabel *authorizeds* berisi tentang hak akses dan siapa yang dapat melakukan akses pada halaman yang telah ditentukan.

Tabel 4.2 Tabel Bahasa

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	Primary Key
2	Kategori_bahasa	Varchar(20)	
3	dialek	Varchar(25)	
4	artikel	Text	
5	User_created	Int(10)	Foreign Key
6	User_modified	Int(10)	Foreign Key
7	created	Datetime	
8	modified	datetime	

Pada tabel 4.2 yaitu tabel bahasa berisi tentang dialek dan kategori yang perlu di simpan sesuai dengan kebutuhan awal.

Tabel 4.3 Tabel Menu

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	<i>Primary Key</i>
2	Role_id	Int(10)	<i>Foreign Key</i>
3	Method_id	Int(10)	<i>Foreign Key</i>
4	Parent	Int(10)	
5	Class_icon	Varchar(50)	
6	Icon_color	Varchar(15)	
7	Nama_menu	Varchar(100)	
8	Level	Int(3)	

Pada tabel 4.3 tabel menus merupakan tabel yang digunakan untuk menyimpan menu tiap admin.

Tabel 4.4 Tabel Method

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	<i>Primary Key</i>
2	Controller	Varchar(50)	
3	Method	Varchar(100)	

Pada tabel 4.4 merupakan tabel methods yaitu berisi tentang *controller* dan *function* dalam program.

Tabel 4.5 Tabel Pengunjung

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	<i>Primary Key</i>
2	Ip_address	Varchar(30)	
3	Date_access	datetime	
4	Time_access	Varchar(30)	

Pada table 4.5 merupakan tabel yang digunakan untuk menghitung jumlah pengunjung sistem.

Tabel 4.6 Tabel Roles

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	<i>Primary Key</i>
2	Role_name	Varchar(50)	

Tabel 4.6 akan digunakan untuk menyimpan data nama nama akses seperti contohnya admin, super admin atau yang lain.

Tabel 4.7 Tabel *Stopwords*

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	<i>Primary Key</i>
2	Kategori_bahasa	Varchar(20)	
3	Dialek	Varchar (30)	
4	Word	Varchar (30)	
5	Arti_indo	Varchar(50)	
6	Arti_inggris	Varchar(50)	

Pada Tabel 4.7 stopwords merupakan tabel yang akan digunakan sebagai penyimpanan arti bahasa daerah.

Tabel 4.8 Tabel *Terms*

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	<i>Primary Key</i>

2	Bahasa_id	Int(10)	<i>Foreign Key</i>
3	Token	Varchar (50)	
4	Term	Varchar (50)	
5	Frequency	Int(10)	

Pada table 4.8 merupakan tabel *terms* yang akan digunakan untuk menyimpan kata per kata dari artikel yang ada di tabel bahasa.

Tabel 4.9 Tabel Users

No	Kolom	Tipe Data	Keterangan
1	Id	Int(10)	<i>Primary Key</i>
2	Role_id	Int(10)	<i>Foreign Key</i>
3	Username	Varchar(50)	
4	Password	Varchar(100)	
5	Nip	Varchar(25)	
6	Nama	Varchar(40)	
7	Hp	Varchar(14)	
8	Email	Varchar(40)	
9	Alamat	Varchar(255)	
10	Tgl_lahir	Date	
11	Jk	Varchar(15)	
12	Foto	Varchar(50)	
13	Created	Datetime	
14	Modified	Datetime	

Pada tabel 4.9 yaitu tabel users digunakan untuk menyimpan user admin dan data yang diperlukan.

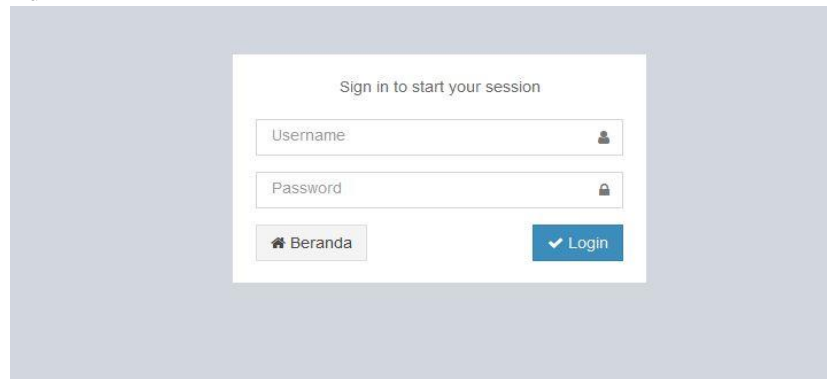
### 4.3 Implementasi Sistem

1. Halaman Beranda



Gambar 4.9 halaman beranda

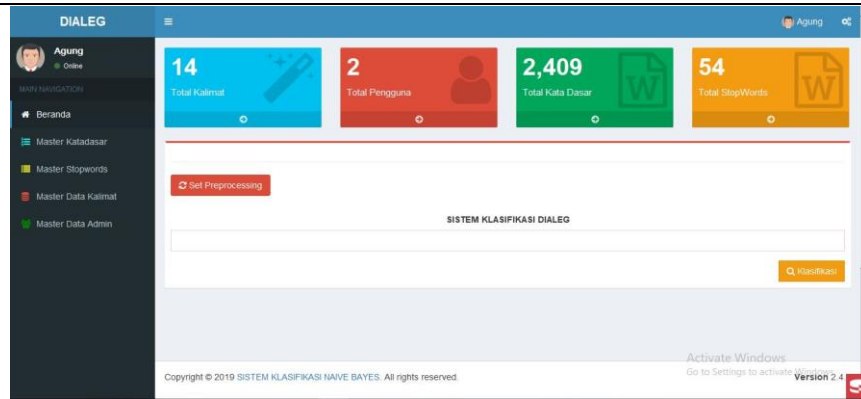
2. Login Admin



Gambar 4.5.2 login admin

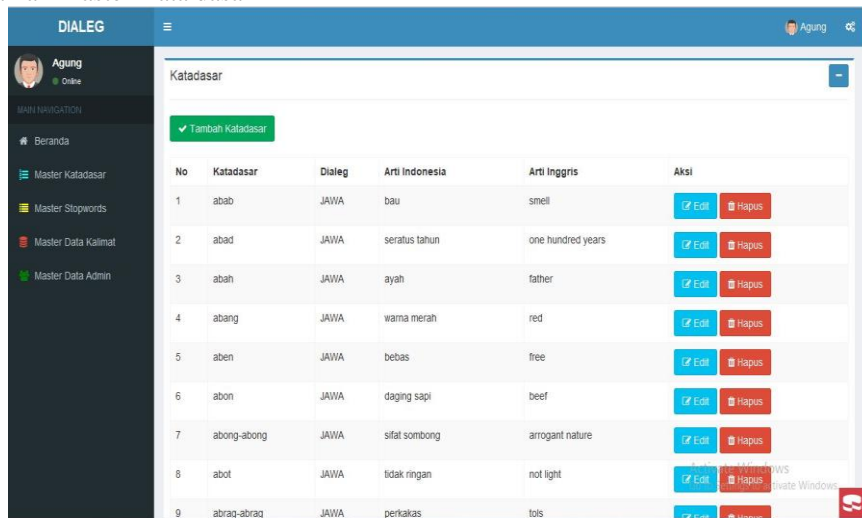
3. Halaman beranda Admin





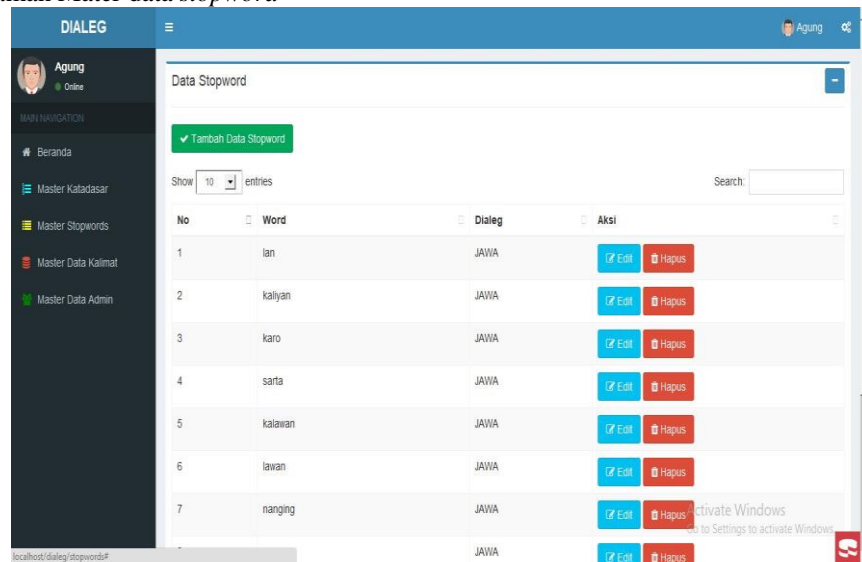
Gambar 4.5.3 Beranda Admin

4. Halaman Master Kata dasar



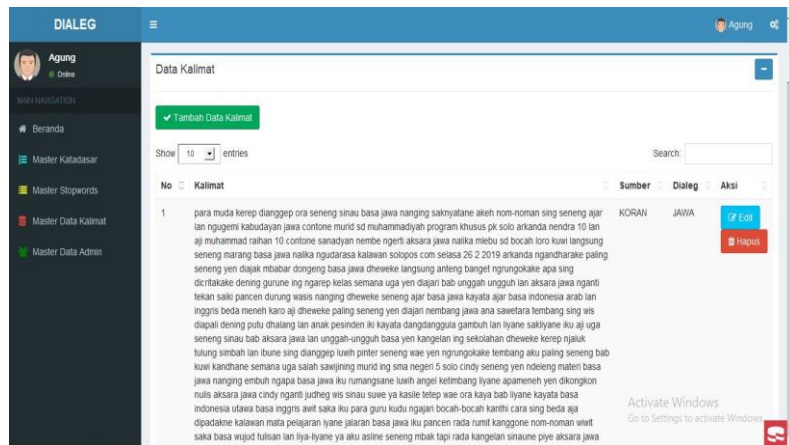
Gambar 4.5.4 Master Kata dasar

5. Halaman Mater data *stopword*



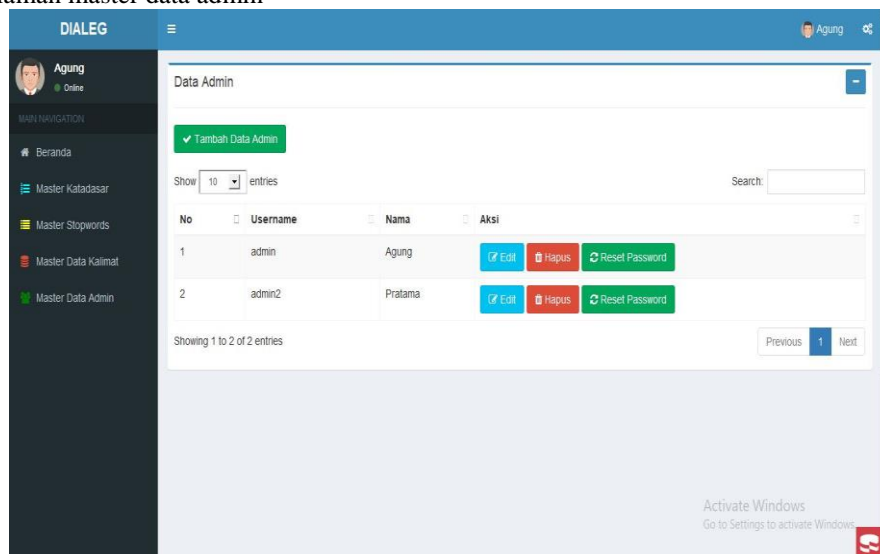
Gambar 4. 1.5 Halaman master data *stopword*

6. Halaman Master data kalimat



Gambar 4. 5.6 Halaman master data kalimat

7. Halaman master data admin



Gambar 4. 5.7 Halaman master data admin

8. Halaman hasil deteksi bahasa



Gambar 4. 5.8 Halaman hasil deteksi bahasa

## V. PENUTUP

### 5.1. Kesimpulan

Dari hasil yang telah didapat pada penjelasan bab sebelumnya, terdapat beberapa hal yang dapat disimpulkan yaitu sebagai berikut :

1. Aplikasi yang telah dibuat dengan metode Naïve Bayes untuk membantu mendeteksi bahasa daerah menjadi bahasa Indonesia maupun Inggris.
2. Sistem atau aplikasi yang dibuat untuk mendeteksi dapat berjalan di metode Naïve bayes
3. hasil pengujian menunjukan semua fitur dapat bekerja dengan baik
4. Membantu masyarakat atau wisatawan belajar bahasa-bahasa daerah.

### 5.2. Saran

Dalam pembuatan Sistem Deteksi Bahasa Daerah Di Indonesia tentu tidak lepas dari kekurangan baik dari sisi *internal* maupun *eksternal*. Maka dari itu diambil beberapa saran yang nantinya dapat dijadikan masukan untuk pengembangan sistem tersebut. Yaitu sebagai berikut :

1. Implementasi metode Naïve Bayes dapat dikembangkan lagi menggunakan metode metode yang lain agar dapat dilakukan perbandingan dan di dapat beberapa metode yang lebih baik dalam pengembangannya
2. Variabel kata dasar tiap daerah banyak yang sama dan beda artinya dan dapat mempersulit deteksi bahasa yang di artikan.

## DAFTAR PUSTAKA

- [1] Z. A. Hasibuan, *METODOLOGI PENELITIAN PADA BIDANG ILMU KOMPUTER DAN TEKNOLOGI INFORMASI*. Depok: Universitas Indonesia, 2007.
- [2] A. A. Khoiruddin, "Sistem Pendukung Keputusan Penentuan Kelayakan Calon Rintisan Sekolah Bertaraf Internasional Dengan Metode Fuzzy Associative Memory," *Semin. Nas. Apl. Teknol. Inf.*, vol. I, no. Snati, pp. 43 – 48, 2008.
- [3] G. Abdurrahman, "Penerapan Metode Tsukamoto (Logika Fuzzy) dalam Sistem Pendukung Keputusan untuk Menentukan Jumlah Produksi Barang Berdasarkan Data Persediaan dan Jumlah Permintaan," *Skripsi*, pp. 1–165, 2011.
- [4] L. Anggraini, "SISTEM PENDUKUNG KEPUTUSAN PENENTUAN KELAYAKAN PELANGGAN DALAM PENGAJUAN KREDIT MOBIL PADA PT. SPEEDLINE AUTO MEDAN DENGAN METODE NAÏVE BAYES," 2015.