

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Indonesia merupakan negara yang terdiri dari banyak suku, ras, bahasa, agama dan budaya yang berbeda-beda. Keberagaman ini merupakan anugerah Tuhan Yang Maha Esa yang tidak dimiliki negara lain. Terkait dengan Bahasa Indonesia, penggunaan bahasa yang dilakukan seseorang atau sekelompok orang dapat mencerminkan perilaku berbahasa masyarakat Indonesia. Penggunaan bahasa Indonesia yang baik dan benar tergantung pada ragam penggunaan bahasa Indonesia dalam situasi, kondisi, dan bidang kehidupan tertentu. Keberagaman bahasa tidak dapat dipisahkan dari budaya karena bahasa merupakan bagian dari kebudayaan, termasuk Bahasa Lampung. Bahasa Lampung dikenal sebagai bahasa daerah atau bahasa daerah.

Menurut penelitian sebelumnya, Lampung adalah salah satu daerah di pulau Sumatera yang memiliki bahasa daerah dan aksara Lampung. Stemming kata bahasa Lampung dialek api dilakukan dengan pendekatan *Brute-force*. Pada bahasa Lampung terdapat verba infleksional (verba bentukan dari dasar berkategori verba) dan verba derivasional (verba bentukan dari dasar berkategori selain verba seperti nomina, adjektiva, adverbial, pronomina dan numeralia). Penelitian ini bertujuan untuk melakukan stemming kata dengan pendekatan *Brute-force*, menghasilkan aplikasi sebagai *Stemmer* kata bahasa Lampung dialek api menggunakan bahasa pemrograman C# dan basis data secara online menggunakan *Firebase*. Metode yang digunakan terdiri dari peneliti mencari, mengidentifikasi, mendata, menetikkan secara manual 2000 kata berimbuhan berikut kata dasar bahasa

Lampung dialek api, membuat algoritma stemming dengan pendekatan *Brute-force*, menguji coba aplikasi yang telah dibuat. Hasil penelitian yaitu aplikasi mampu melakukan stemming kata untuk kata-kata yang telah teridentifikasi pada 2000 kata berimbuhan dan jika tidak dapat dilakukan stemming maka disediakan fasilitas untuk memperbaharui basis data yang digunakan dalam aplikasi agar dapat digunakan untuk stemming karena aplikasi stemming sangat menunjang pada aplikasi mesin penerjemah berbasis kamus (Abidin et al., 2021).

Pada penelitian kata bahasa Lampung dialek api sebelumnya dengan menggunakan metode pendekatan *Brute-force*, proses stemming dilakukan dengan menggunakan 2000 kata berimbuhan berikut kata dasarnya dengan bahasa pemrograman C# ini memiliki kelebihan yaitu guna memperbaiki mesin penerjemah bahasa Lampung berbasis kamus ketika menemui kata yang tidak ada di dalam basis data aplikasi tersebut. Pendekatan *Brute-force* pada stemming bahasa Lampung ini dapat mencocokkan kata yang di-*input* dengan ketersediaan kata yang ada dalam basis data kata berimbuhan. Namun, pendekatan *Brute-force* ini juga memiliki kekurangan yaitu jika kata berimbuhan tidak ada diluar 2000 kata berimbuhan berikut kata dasarnya.

*Brute-force* dapat digunakan untuk membangun mesin penerjemah kalimat bahasa Lampung dialek api yang mampu mengatasi kata berimbuhan (Abidin et al., 2021), yang mana ini tepat untuk mencocokkan kata yang diinputkan dengan ketersediaan kata yang ada dalam basis data kata berimbuhan. Tetapi *Brute-force* memiliki kelemahan yaitu hanya dapat memproses 2000 kata berimbuhan yang disediakan berikut kata dasarnya.

Pada penelitian ini pendekatan Algoritma Nazief & Adriani diterapkan untuk melakukan proses stemming pada bahasa Lampung. Dikarenakan dalam Bahasa Lampung mempunyai banyak keunikan khususnya pada morfologi. Morfologi merupakan ilmu yang mempelajari satuan kebahasaan seperti satuan gramatikal. Proses morfologi adalah terbentuknya kata yang muncul dari bentuk dasarnya melalui imbuhan, pengulangan, dan pemajemukan. Banyaknya variasi morfologi suatu bahasa sangat mempengaruhinya. Oleh karena itu, kata-kata yang varian morfologinya berbeda harus diganti dengan bentuk dasarnya.

Tujuan penelitian ini adalah untuk stemming kata Bahasa Lampung dengan pendekatan modifikasi Algoritma Nazief & Adriani. Proses stemming digunakan untuk menormalkan teks dan memudahkan dalam pengolahannya. Ini adalah langkah penting dalam pra-pemrosesan teks. Program stemming biasa disebut dengan algoritma stemming atau *stemmer*. Hasil yang diharapkan dari pendekatan Nazief & Adriani bagi pengguna ialah dapat memberikan kemudahan dalam mencari kata dasar pada salah satu kata yang terdapat pada kata berimbuhan Bahasa Lampung, dengan melakukan dua eksperimen untuk melihat nilai *Gold Standard Assessment*. Eksperimen 2 diharapkan mampu untuk memperbaiki nilai akurasi eksperimen 1.

## **1.2 Rumusan Masalah**

Adapun rumusan masalah dari penelitian ini adalah bagaimana memodifikasi pendekatan Algoritma Nazief & Adriani untuk stemming kata Bahasa Lampung Dialek Api untuk melihat nilai *Gold Standard Assessment* dengan melakukan eksperimen 1 dan eksperimen 2?

### 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, adapun tujuan dari penelitian ini adalah memperbaiki nilai akurasi modifikasi Pendekatan algoritma Nazief & Adriani untuk stemming pada kata bahasa Lampung Dialek Api dengan melakukan dua eksperimen, eksperimen 1 dimulai dari urutan sufiks, prefiks, konfiks, dan infiks, eksperimen 2 dimulai dari urutan prefiks, sufiks, konfiks, dan infiks.

### 1.4 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini berfokus pada bahasa Lampung Dialek Api.
2. Kata dalam kamus sebanyak 7327 kata dasar sebagai acuan.
3. Kata uji yang digunakan berjumlah 210 kata berimbuhan yang digunakan sebagai kata uji.

### 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah salah satu *tools text preprocessing* bahasa Lampung, dan salah bentuk pelestarian bahasa, yaitu bahasa Lampung.