

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Besemah adalah nama suku yang mendiami daerah lembah Gunung Dempo di Kota Pagar Alam, Sumatera Selatan. Wilayah itu merupakan bagian dari Pulau Sumatera Selatan yang luas. Daerah yang didiami oleh suku Besemah itu kemudian disebut juga dengan tanah Besemah. Mereka menyebut dirinya sebagai jeme (orang) Besemah.(Bloom & Reenen, 2013)

Asal-usul kata "Besemah" dapat ditelusuri dari istilah "semah," yang merupakan jenis ikan sungai. Dengan penambahan awalan "be-" yang berarti "ada" atau "mempunyai," maka Besemah dapat diartikan sebagai daerah yang memiliki ikan semah atau sungai yang kaya akan ikan semah. Wilayah sekitar sungai tersebut disebut sebagai ayek Besemah, sementara daerah sekitarnya dikenal sebagai daerah Besemah. Penggunaan kata Besemah tidak hanya terbatas sebagai nama geografis, melainkan juga merujuk pada persatuan masyarakat di sekitar kabupaten Lahat dan Kota Pagar Alam, seperti Himpunan Keluarga Besemah. "Pesemah."(Wardiah, 2021)

Bahasa Besemah, sebagai salah satu bahasa daerah, menjadi bahasa sehari-hari bagi masyarakat di kota Pagar Alam dalam interaksi dengan penduduk desa. Saat terlibat dalam pertemuan resmi, warga Kota Pagar Alam beralih ke penggunaan bahasa Indonesia. Pemeliharaan dan penggunaan bahasa daerah ini mencerminkan rasa cinta dan upaya untuk melestarikan serta mempertahankan keberlanjutan bahasa tersebut. Kekayaan kearifan lokal dalam bahasa menjadi bagian integral dari identitas setiap daerah. Dalam bahasa daerah, juga terdapat variasi bentuk kata yang menjadi ciri khas yang digunakan oleh masyarakat..(Wardiah, 2021)

Mesin penerjemah merupakan perangkat lunak penerjemah bahasa yang proses penerjemahannya dilakukan oleh mesin (komputer) secara otomatis dan hasilnya dapat langsung disajikan kepada pengguna. Opsi penggunaan mesin penerjemah lebih umum dipilih karena cepat dan praktis dalam melakukan penerjemahan teks, berbanding terbalik dengan penerjemahan oleh manusia yang memerlukan waktu lama. Meskipun penerjemah manusia yang berkompeten akan menghasilkan penerjemahan yang lebih baik dan dalam beberapa keadaan tak tergantikan oleh penerjemah otomatis, namun ada banyak kasus di mana cukup untuk diterjemahkan menggunakan mesin penerjemah. Selain itu, hasil penerjemahan oleh mesin pun tidak serta merta digunakan begitu saja, namun disunting kembali oleh manusia. Sehingga penerjemah manusia berkualifikasi tinggi pun juga menggunakan mesin penerjemah untuk mempercepat proses penerjemahan. (Wismoyo, 2018)

Sebelumnya ada beberapa peneliti terdahulu Seperti Mesin Penerjemah Bahasa Indonesia -Bahasa Sunda Menggunakan *Recurrent Neural Networks* yang dirancang dengan Yustiana Fauziah, Ridwan Ilyas dan Fatan Kasyidi, dan hasil yang didapatkan arsitektur model variasi dari *RNN* didapatkan nilai optimal oleh *GRU* akurasi sebesar 99,17% dan ada juga Mesin Penerjemah Bahasa Komerling ke Bahasa Indonesia Menggunakan Algoritma *RNN* yang dirancang dengan Irfan Henuarianto dan hasil yang didapatkan dari model algoritma *RNN Embedding* ini berjalan dengan pelatihan, akan tetapi dari beberapa algoritma tersebut, *Algoritma Encoder-Decoder* lah yang bisa dikatakan pas untuk Pembuatan mesin Penerjemah Bahasa Besemah ke Indonesia tersebut. Oleh karena itu, untuk mengatasi permasalahan yang ada, maka diperlukan penerapan algoritma *Encoder-Decoder*. Arsitektur algoritma *Encoder-Decoder* adalah salah satu bentuk pendekatan dalam *machine learning* yang umumnya digunakan. Metode ini terdiri dari dua komponen utama, yakni *encoder* yang bertugas memahami dokumen input dan

mentransformasikannya menjadi terjemahan untuk setiap kata dalam bahasa Besemah. Sementara itu, *decoder* memiliki peran untuk mengambil representasi data input yang dihasilkan oleh *encoder* dan merangkainya menjadi hasil terjemahan dalam bahasa Indonesia. Kelebihan menggunakan Model *Encoder-Decoder* adalah untuk memperingkas teks menggunakan *pointer-generator dan coverage* untuk mengatasi dua masalah utama : ketergantungan terhadap kosakata dan repetisi kata. *Pointer-generator* memungkinkan model untuk meniru kata dari dokumen masukan, sehingga mengurangi ketergantungan terhadap kosakata. *Coverage* membantu mengurangi repetisi kata dengan melacak kata-kata yang telah digunakan dalam ringkasan. (AKBAR,2021).

Kekurangan model *Encoder-Decoder*, terletak pada peningkatan kemampuan *encoder* dan model secara keseluruhan memahami konteks *coverage*. *Pointer-Generator* model untuk meniru kata-kata langsung dari dokumen input. (AKBAR,2021).

Oleh karena itu, untuk menyelesaikan permasalahan yang ada maka diperlukan Model *encoder decoder*, merupakan algoritma yang dapat membantu pengembangan mesin penerjemah bahasa daerah. Pada bahasa daerah Besemah, pengembangan mesin penerjemah dengan memanfaatkan teknologi informasi belum terlaksana dengan baik. Untuk itu diperlukan teknologi penerjemahan otomatis dari Besemah ke Bahasa Indonesia agar bahasa Besemah semakin dikenal dan tidak ketinggalan dengan budaya bahasa asing. Melihat penjelasan masalah, Peneliti melakukan penelitian dengan judul "**Mesin Penerjemahan Bahasa Besemah Berbasis *Machine Learning* dengan Algoritma Model *Encoder-Decoder***".

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dijelaskan diatas, rumusan masalah dari penelitian ini yaitu Bagaimana menerjemahkan dari Bahasa Besemah ke Bahasa Indonesia berbasis *machine learning* menggunakan model Encoder-Decoder.

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk membuat mesin penerjemahan bahasa Besemah ke Indonesia dan sebaliknya. Menggunakan algoritma Encoder-Decoder berbasis *Machine Learning*.

1.4 Batasan Masalah

Adapun batasan-batasan masalah yang ditetapkan dalam penelitian ini antara lain sebagai berikut :

1. Penerjemahan yang dilakukan dan diterapkan batas 1 kalimat saja. Jika kalimatnya jadi banyak sampai terbentuk satu paragraf, maka penerjemahan model algoritma encoder-decoder dalam penelitian ini belum bisa melakukannya.
2. Penelitian ini dilakukan hanya sampai pada fase pemodelan algoritma encoder-decoder saja.

1.5 Manfaat Penelitian

Melalui pelaksanaan penelitian ini, diharapkan terdapat beberapa manfaat bagi pengguna, yaitu:

1. Dapat memudahkan bagi pengguna yang berkeinginan untuk menerjemahkan teks dari bahasa daerah Besemah ke bahasa Indonesia atau sebaliknya.
2. Dapat menjadi referensi pengetahuan bagi peneliti berikutnya yang tertarik dalam pengembangan Mesin Penerjemah bahasa Besemah.
3. Dapat Mengenalkan bahasa Besemah ke seluruh dunia.