

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bahasa adalah media komunikasi yang dimanfaatkan oleh manusia untuk mengungkapkan ide, pemikiran, konsep, atau emosi kepada orang lain (Apriani dkk., 2016). Bahasa memiliki fungsi khusus yang digunakan sesuai kebutuhan individu, seperti mengekspresikan diri, berkomunikasi, dan menyesuaikan diri dengan lingkungan sosial serta kondisi tertentu. Indonesia merupakan negara dengan keragaman luar biasa dalam hal suku, budaya, dan bahasa. Menurut Badan Bahasa Kementerian Pendidikan dan Kebudayaan RI, tercatat ada 718 bahasa daerah di seluruh wilayah Indonesia, dengan 90 persen di antaranya tersebar di kawasan timur. Di Papua sendiri dilaporkan terdapat 428 bahasa, sementara di Maluku ada 80 bahasa, di Nusa Tenggara Timur 72 bahasa, dan di Sulawesi sebanyak 62 bahasa. Perbedaan bahasa antar suku ini dapat menjadi hambatan dalam komunikasi, karena individu dari suku lain mungkin tidak dapat memahami pesan yang disampaikan.

Perbedaan bahasa ini mendorong banyak pihak untuk menciptakan sistem penerjemahan yang kemudian berkembang menjadi penerjemahan otomatis. Terjemahan otomatis adalah proses menerjemahkan kalimat dari satu bahasa ke bahasa lain menggunakan teknologi komputer. Teknologi ini memungkinkan komputer atau sistem AI untuk memahami dan menerjemahkan konten dari satu bahasa ke bahasa lain tanpa intervensi manusia.

Mesin penerjemah yang digunakan dalam penerjemahan otomatis adalah pemrosesan bahasa alami (Natural Language Processing atau NLP) dalam bidang

terjemahan mesin (Machine Translation atau MT). MT melakukan konversi dari bahasa sumber ke bahasa target. Dalam penerjemahan mesin terdapat beberapa pendekatan yang umum digunakan, yaitu Statistical Machine Translation (SMT), Rule-Based Machine Translation (RMT), Phrase-Based Statistical Machine Translation (PMT), serta Neural Machine Translation (NMT) (Shah & Bakrola, 2019). NMT adalah pendekatan yang paling sesuai untuk penerjemahan otomatis (Abidin, 2017). NMT merupakan metode yang relatif baru dalam teknologi penerjemahan mesin, yang menggabungkan penggunaan *encoder* jaringan saraf tiruan yang mengubah bahasa sumber menjadi vektor berdimensi tetap dan *decoder*, yaitu jaringan saraf tiruan yang bertugas menghasilkan terjemahan akhir dari vektor tersebut. Salah satu metode yang paling umum digunakan dalam NMT adalah *Arsitektur Transformer*, Arsitektur Transformer adalah arsitektur penerjemahan yang menerapkan mekanisme *attention* tanpa tambahan Recurrent Neural Network (RNN) ataupun Convolutional Neural Network (CNN) (Vaswani dkk., 2017).

Dengan latar belakang ini, penulis mengembangkan penerapan arsitektur Transformer untuk penerjemahan bahasa Indonesia ke bahasa daerah Papua (studi kasus daerah Kokas). Kokas merupakan salah satu distrik di Kabupaten Fakfak, Papua Barat, yang memiliki wilayah geografis yang kaya akan keanekaragaman hayati dan budaya, sehingga penelitian ini bertujuan untuk memfasilitasi komunikasi dan akses informasi bagi masyarakat yang berbicara dalam bahasa daerah tersebut. Hal ini juga mempermudah komunikasi di daerah tersebut, mengingat tidak sedikit masyarakat yang menggunakan bahasa Kokas. Dengan demikian, penulis ingin melestarikan komunikasi dalam bahasa daerah tersebut.

1.2 Rumusan Masalah

Berdasarkan kajian-kajian yang telah ada, rumusan masalah dapat disusun sebagai berikut: bagaimana cara mengimplementasikan Arsitektur Transformer pada penerjemahan Bahasa Indonesia ke Bahasa Papua Kokas.

1.3 Tujuan

Tujuan dari penelitian ini adalah mengimplementasikan Arsitektur Transformer pada penerjemahan Bahasa Indonesia ke Bahasa Papua Kokas serta mengetahui kualitas dan tingkat akurasi menggunakan algoritma BLEU.

1.4 Manfaat

Manfaat yang diharapkan pada peneliti

Memperluas wawasan dan keterampilan penulis dalam bidang *machine learning* serta berpotensi menjadi referensi bagi penelitian-penelitian selanjutnya yang memiliki kesamaan dengan studi ini.

Manfaat yang diharapkan pada pengguna

Penerjemahan otomatis dari Bahasa Indonesia ke bahasa kokas dapat memperluas jangkauan akses masyarakat setempat terhadap informasi yang tersedia dalam Bahasa Indonesia. Fenomena ini memungkinkan implementasi teknologi sebagai sarana pendukung dalam mendorong proses pembelajaran, fasilitasi komunikasi, serta meningkatkan tingkat partisipasi masyarakat dalam era digital yang terus berkembang.

1.5 Batasan Masalah

Batasan masalah yang ditetapkan lingkup pembahasan ini mencakup:

1. Penelitian ini hanya bisa menerjemahkan bahasa Indonesia ke bahasa Papua Kokas.
2. Data yang digunakan yaitu 2908 kalimat bahasa Indonesia dan 2908 kalimat bahasa Papua Kokas dan kalimat tersebut berupa kalimat sehari-hari, yang dimana jumlah data yang dimiliki masih terbatas.
3. Evaluasi yang digunakan untuk mengetahui skor yaitu algoritma BLEU.

1.6 Metodologi Penelitian

1.6.1. Bahan dan Alat Penelitian

Di dalam pembuatan sebuah sistem tentunya diperlukan alat dan bahan yang digunakan untuk menunjang proses pengerjaan sistem tersebut. Dalam hal ini penulis menggunakan software yang berhubungan dengan pemrograman web sebagai berikut:

1. Hardware
 - Laptop
2. Software
 - Sistem Operasi *Microsoft Windows 11*
 - Browser menggunakan *Chrome*
 - Desain *prototype* menggunakan *Figma*
 - Bahasa pemrograman *Python* dan menggunakan *Framework Django*
 - Code Editor menggunakan *Visual Studio Code*

1.6.2. Pengumpulan Data dan Informasi

Pengumpulan data penerjemahan Bahasa Indonesia dilakukan dengan mengambil data dari beberapa sumber, seperti Kaggle, korpus Bahasa Indonesia, dan ChatGPT. Kaggle digunakan untuk mencari kalimat sehari-hari dalam Bahasa Indonesia, demikian pula dengan korpus Bahasa Indonesia dan ChatGPT. Setelah data Bahasa Indonesia terkumpul, data tersebut disimpan dalam bentuk file Excel, kemudian file tersebut dikirimkan kepada masyarakat untuk diterjemahkan ke dalam Bahasa Kokas. Data yang telah diterjemahkan oleh masyarakat kemudian dikirimkan kembali untuk diproses. Dataset ini mencakup 2.908 kalimat paralel.

1.6.3. Analisis Data

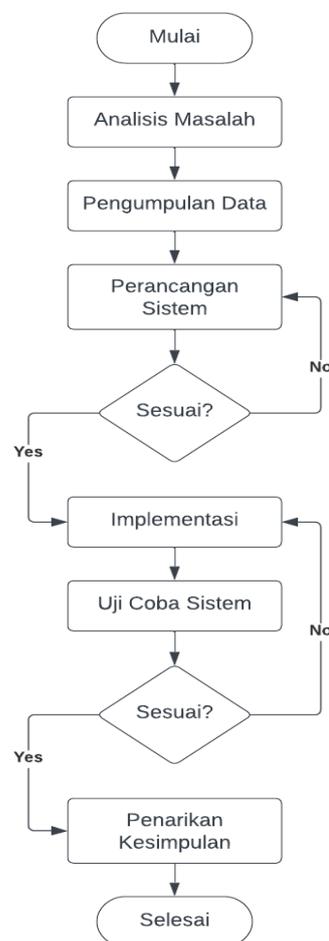
Dalam penelitian berjudul "Implementasi Arsitektur Transformers Penerjemahan Otomatis Bahasa Indonesia ke Bahasa Papua (Studi Kasus Daerah Kokas)," metode analisis data yang digunakan adalah metode analisis tekstual. Metode ini melibatkan proses sistematis untuk memahami, menafsirkan, dan mengekstrak makna dari data teks yang dikumpulkan, khususnya teks dalam bahasa Indonesia dan bahasa Papua Kokas.

Analisis tekstual dipilih karena metode ini sesuai untuk mengidentifikasi pola-pola bahasa dan struktur kalimat yang unik dalam data paralel kedua bahasa. Dengan menggunakan metode ini, penelitian dapat memeriksa bagaimana model Transformer menerjemahkan teks dari bahasa sumber (Indonesia) ke bahasa target (Papua Kokas). Selain itu, analisis ini membantu mengevaluasi kinerja model penerjemahan otomatis dengan melihat akurasi, konsistensi, dan relevansi hasil terjemahan dalam konteks bahasa Kokas.

Pada akhirnya, hasil analisis ini dirumuskan dalam kesimpulan yang menjawab masalah penelitian, yaitu sejauh mana arsitektur Transformer mampu melakukan penerjemahan otomatis dari bahasa Indonesia ke bahasa Papua Kokas dan mengatasi tantangan linguistik yang ada dalam bahasa daerah tersebut.

1.6.4 Prosedur Penelitian

Prosedur penelitian yang dilakukan sebagai berikut:



Gambar 1.1 Diagram Alir Penelitian

Analisis masalah

Pada tahap ini penulis perlu memahami beberapa masalah yang akan dihadapi dalam pembuatan sistem rekomendasi dan membuat pemecahan masalah agar memudahkan penulis dalam merancang solusi yang efektif saat pembuatan sistem.

Pengumpulan Data

Pada tahap ini penulis mengumpulkan data korpus teks dengan metode wawancara bersama masyarakat.

Perancangan Sistem

Pada tahap ini perancangan sistem dibuat untuk memastikan bahwa penelitian dilakukan secara sistematis dan sesuai dengan tujuan yang ditetapkan. Ini juga memberikan kerangka kerja yang memandu penelitian dari awal hingga akhir, sehingga mempermudah penulis untuk mengelola langkah-langkah penelitian dengan efisien dan efektif.

Implementasi

Setelah perancangan sistem maka penulis selanjutnya akan melakukan implementasi terhadap masalah yang dialami.

Uji Coba Sistem

Setelah melakukan implementasi, penulis melakukan uji coba terhadap hasil dari implementasi. Jika sistem belum sesuai dengan hasil perancangan, maka akan dilakukan revisi atau implementasi ulang.

Penarikan Kesimpulan

Setelah melakukan uji coba dan memastikan sistem berjalan dengan baik dan sesuai dengan perancangan, maka penulis dapat menarik kesimpulan.

1.7 Sistematika Penulisan

BAB I: PENDAHULUAN

Berisi tentang latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, metodologi penelitian dan sistematika penulisan.

BAB II: TINJAUAN PUSTAKA

Berisi tentang penelitian terdahulu dan teori terkait.

BAB III: ANALISIS DAN PERANCANGAN

Bagian ini mencakup analisis, perancangan, dan rancangan pengujian

BAB IV: PEMBAHASAN

Berisi tentang gambar umum penelitian, implementasi, dan uji coba

BAB V: PENUTUP

Berisi tentang kesimpulan dan saran.