

BAB I

PENDAHULUAN

1.1. Latar Belakang

Bahasa kawi merupakan bahasa yang dipergunakan pada naskah Hindu kuno di mana berkembang pada abad ke 9 sampai abad ke 15. Bahasa Kawi merupakan bahasa campuran antara bahasa Sanskerta dan bahasa Jawa Kuno (Surada, 2018). Bahasa Kawi berasal dari kata *kavya* yang berarti syair atau puisi sehingga bahasa Kawi bisa disebut juga sebagai bahasanya seorang pengarang atau penyair (Surada, 2018). Sayangnya bahasa Kawi tidak digunakan lagi sebagai bahasa komunikasi pada kehidupan sehari-hari (Surada, 2018). Pada kehidupan modern, Bahasa Kawi biasa digunakan pada prosesi upacara keagamaan Hindu. Bahasa Kawi biasa ditemukan pada naskah-naskah literatur kuno seperti naskah keagamaan (lontar Tattwa, Sasana, Niti), Naskah-naskah sastra (lontar Parwa, Kekawin, Kidung) serta pada naskah-naskah pengobatan seperti Lontar Usada (Yunairi & Bhattacharya, 2020).

Terdapat kekhawatiran terhadap potensi punahnya bahasa Kawi melihat media yang digunakan penulisan bahasa Kawi rentan rusak ataupun penutur bahasa Kawi saat ini sangat terbatas, terutama di kalangan linguis yang secara khusus mempelajari bahasa tersebut. Hal ini diperkuat juga dengan pemaparan dari (Purnami, 2019) bahwa tidak banyak keluarga pemilik lontar yang paham isi dari Lontar. Hal ini memungkinkan bahwa generasi ke depan akan kehilangan akses terhadap eksistensi dari bahasa Kawi di mana berpotensi hilangnya pengetahuan terdahulu yang termuat di Lontar.

Terdapat beberapa upaya dalam melestarikan bahasa Kawi. Hal ini dilakukan dengan melakukan penerjemahan terhadap beberapa manuskrip lontar ke bahasa yang umum digunakan seperti bahasa Indonesia ataupun bahasa Bali. Namun hal ini terbatas pada manuskrip-manuskrip yang dipelajari oleh para lingustik. Beberapa manuskrip belum dalam keadaan di terjemahkan melihat keberadaan dari manuskrip yang tersebar di beberapa tempat seperti di griya, puri/jero, rumah-rumah penduduk (Purnami, 2019) yang memungkinkan beberapa

lontar belum tergapai kepada para peneliti bidang linguistik tersebut. Permasalahannya lainnya adalah lamanya waktu yang dibutuhkan untuk melakukan alih aksara dan alih bahasa terhadap lontar. Pada Antara NEWS oleh (Adnyana, 2018), di Perpustakaan Gedong Kirtya, proses alih aksara dan alih bahasa pada 2 sampai 3 lontar memerlukan waktu sekitar setahun untuk penyelesaiannya. Setidaknya pada perpustakaan Gedong Kirtya memiliki sejumlah 1808 untuk naskah kuno serta 5200 salinan lontar yang didapatkan di masyarakat. Dari sekian banyak lontar, 80 diantara-Nya telah di alih aksara dan di alih bahasakan serta 40 diantara-Nya telah di bukukan.

Adapun upaya pembuatan sistem, dalam menanggulangi keterbatasan tersebut yaitu terbatasnya jumlah manuskrip yang telah di terjemahkan oleh para ahli linguistik. Beberapa penelitian tersebut adalah mengenai penelitian Sistem Transliterasi Lontar yaitu penggunaan model untuk alih huruf dari huruf Bali yang ada pada Lontar yang dikonversi menjadi huruf Latin seperti penelitian yang dilakukan (Kesiman & Dermawan, 2021b, 2021a; Suciati dkk., 2022). Ada juga upaya dalam memahami bahasa yang tertera pada Lontar yaitu sistem Translasi bahasa Kawi yang terdapat pada penelitian (Budaya dkk., 2022a, 2022b). Perkembangan pada penelitian transliterasi aksara pada lontar memiliki kemajuan yang pesat di mana pada penelitian terakhir pada (Suciati dkk., 2022) menggunakan model *object detection* untuk mendeteksi setiap aksara yang ada pada lontar, namun penelitian terkait translasi bahasa Kawi ke Bahasa Indonesia masih memiliki ruang untuk berkembang. Hal ini terlihat dari penggunaan model *Recurrent Neural Network* (RNN) pada penelitian (Budaya dkk., 2022a, 2022b) yang menunjukkan adanya peluang untuk mengeksplorasi pendekatan ke model yang lebih jauh.

Penelitian terkait mesin penerjemah adalah hal yang menarik untuk dilakukan. Hal ini berbuah dari pertanyaan yaitu bagaimana cara agar komputer atau suatu algoritma mampu memahami bahasa dan melakukan alih bahasa dari satu bahasa ke bahasa lain. Demi menjawab pertanyaan ini, terdapat salah satu pertemuan yang dilakukan oleh kumpulan peneliti dari seluruh dunia demi bisa mengatasi permasalahan linguistik tersebut yang bernama WMT (*Workshop Machine Translation*). *Workshop* ini melahirkan beberapa model revolusioner dalam

perkembangan mesin dalam melakukan alih bahasa, model yang mampu memperhitungkan konteks dari kalimat dengan menggunakan *neural network*.

Pendekatan menerjemahkan bahasa dengan memperhitungkan konteks dari kalimat adalah kelebihan dari penerjemahan bahasa menggunakan *neural network*. Hal ini ditunjukkan dari beberapa kesuksesan model basis *neural network* dalam melakukan penerjemahan bahasa seperti pendekatan basis *Recurrent Neural Network* (Bahdanau dkk., 2016; Cho dkk., 2014; Sutskever dkk., 2014; Wu dkk., 2016), basis *Convolution Neural Network* pada penelitian (Gehring dkk., 2017; Kalchbrenner dkk., 2017) sampai pengembangannya menggunakan *Attention Network* yang disebut sebagai *Transformer* pada jurnal “*Attention is All You Need*” (Vaswani dkk., 2017). Setiap model tersebut dilatih pada dataset WMT dengan domain bahasa yang coba diselesaikan adalah penerjemahan bahasa Jerman dan bahasa Prancis ke bahasa Inggris. Kemampuan model dalam melakukan penerjemahan dapat dilihat pada Tabel 1.1 yang menunjukkan skor BLEU masing-masing model.

Tabel 1.1 Skor BLEU setiap model dalam penerjemahan bahasa Jerman dan Prancis ke bahasa Inggris

Peneliti	Model	BLEU Jerman - Inggris	BLEU Prancis - Inggris
(Cho dkk., 2014)	RNN	-	33.87
(Sutskever dkk., 2014)	LSTM	-	34.81
(Bahdanau dkk., 2016)	RNN + Attention	-	36.15
(Wu dkk., 2016)	LSTM + Attention (GNMT)	24.6	39.92
(Kalchbrenner dkk., 2017)	CNN (ByteNet)	23.75	-
(Gehring dkk., 2017)	CNN + Attention	26.36	41.29
(Vaswani dkk., 2017)	Attention (Transformer)	28.4	41.8

Berdasarkan Tabel 1.1 model dari Transformer memiliki BLEU skor lebih tinggi dari pada beberapa jenis model yang mengadaptasi model jenis *Recurrent Neural Network* ataupun *Convolutional Neural Network* dalam melakukan penyelesaian kasus penerjemahan bahasa Jerman dan Prancis ke bahasa Indonesia.

Hal ini didukung pada beberapa penelitian penggunaan Transformer pada beberapa bahasa yang dianggap sebagai *low resource* seperti penelitian pengembangan model penerjemahan dengan menggunakan Transformer dan RNN pada penelitian (Afra dkk., 2024; Lankford dkk., 2022; Primandhika dkk., 2021; Wangchuk dkk., 2023) di mana pada hasil penelitian tersebut ditunjukkan model Transformer memiliki keunggulan dalam melakukan penerjemahan yang ditunjukkan pada *Automatic Evaluation* skor BLEU. Beberapa penelitian seperti (Baniata dkk., 2021; Nongbri dkk., 2023; Nyoni & Bassett, 2021) menggunakan Transformer sebagai *base* model dalam penyelesaian masalah penerjemahan bahasa.

Model yang di pertunjukkan pada studi terbaru mengenai penerjemahan Bahasa Kawi mengadaptasi model *Recurrent Neural Network* sebagai fondasi pembangunan model penerjemahan bahasa Kawi. Beberapa penelitian serupa terkait penerjemahan bahasa pada penelitian (Afra dkk., 2024; Baniata dkk., 2021; Lankford dkk., 2022; Nongbri dkk., 2023; Nyoni & Bassett, 2021; Primandhika dkk., 2021; Wangchuk dkk., 2023) di mana pada studi masing-masing peneliti tersebut menemukan penggunaan *Transformer* yang meningkatkan hasil penerjemahan Bahasa dilihat dari Skor *Automatic Evaluation* skor BLEU. Melihat terdapatnya gap riset pada penerjemahan Bahasa Kawi ke Indonesia yaitu pada sisi model mesin penerjemah yang digunakan pada penelitian sebelumnya menggunakan *Recurrent Neural Network*, maka penelitian ini berfokus pada perancangan dan implementasi model *Transformer* sebagai riset lanjutan mengenai penerjemahan Bahasa Kawi ke Bahasa Indonesia.

1.2. Rumusan Masalah

Berdasarkan pemaparan persoalan yang di paparkan pada latar belakang masalah, maka dari permasalahan tersebut diambil rumusan masalah:

1. Bagaimana membuat rancangan dan implementasi Transformer untuk penerjemahan *low resource language* Bahasa Kawi ke Bahasa Indonesia?
2. Bagaimana hasil kinerja dari model Transformer untuk sistem penerjemahan *low resource language* Bahasa Kawi ke Bahasa Indonesia?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah yang di paparkan, maka tujuan penelitian dari penelitian ini yang sejalan dengan rumusan masalah dipaparkan sebagai berikut:

1. Untuk merancang dan mengimplementasi Transformer model untuk sistem penerjemahan *low resource language* Bahasa Kawi ke Bahasa Indonesia.
2. Mengetahui hasil kinerja dari Transformer model untuk sistem penerjemahan Bahasa Kawi ke Bahasa Indonesia.

1.4. Pembatasan Masalah

Pembatasan suatu masalah digunakan untuk menghindari adanya penyimpangan maupun pelebaran pokok masalah supaya penelitian tersebut lebih terarah dan memudahkan dalam pembahasan sehingga tujuan penelitian akan tercapai. Pembatasan masalah sebagai berikut :

1. Korpus diambil dari buku Musala Parwa, Phalawkya Kapi Parwa, Singhalangghyala Parwa.
2. Format penulisan Bahasa Kawi dan Bahasa Indonesia menggunakan skrip Latin.

1.5. Manfaat Penelitian

Berdasarkan tujuan penelitian yang telah disampaikan, adapun beberapa manfaat yang didapatkan dari penelitian ini :

1. Manfaat Teoritis

- a. Menambah wawasan mahasiswa mengenai penggunaan model *Deep Learning*, model Transformer untuk sistem penerjemahan Bahasa Kawi ke Bahasa Indonesia.
 - b. Menjadi bahan acuan atau sumber bacaan mengenai penerapan Transformer dalam penerjemahan Bahasa Kawi ke Bahasa Indonesia.
2. Manfaat Praktis

Bagi Peneliti

Dapat mengimplementasikan ilmu yang diperoleh di bangku perkuliahan mengenai penerjemahan Bahasa menggunakan *Deep Learning*.

Bagi Masyarakat

- a. Sebagai bentuk penelitian lanjutan dari penelitian-penelitian sebelumnya terkait penerjemahan Bahasa Kawi ke Bahasa Indonesia.
- b. Sebagai upaya pelestarian Bahasa Kawi.

