

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Mesin penerjemah telah menjadi pilihan utama bagi orang – orang untuk memahami pengetahuan di luar bahasa asli mereka, khususnya dalam hal pembelajaran. Studi menunjukkan bagaimana mesin translasi seperti *google translate* mampu membantu para pelajar untuk belajar bahasa asing dengan lebih efisien (Kai & Hua, 2021). Bahasa Kawi merupakan bahasa yang banyak digunakan dalam literatur kuno seperti naskah lontar, khususnya yang ada di provinsi Bali, Indonesia. Naskah lontar sendiri berisi berbagai pengetahuan dasar terkait agama, pengobatan (*usada*), dan catatan – catatan penting tentang peradaban di masa lalu (HINZLER, 1993).

Saat ini pembelajaran terhadap naskah lontar masih dilakukan oleh beberapa kelompok masyarakat sebagai pemenuhan tradisi maupun pembelajaran dan pengembangan kesusastraan yang dalam prosesnya didukung oleh para akademisi, komunitas lokal dan pemerintah (Ardiyasa, 2021). Pembelajaran terhadap naskah lontar ini memiliki tantangan tersendiri. Salah satu tantangan adalah proses pemahaman konteks melalui penerjemahan, dimana dari bahasa Kawi yang ada pada naskah lontar ke dalam bahasa yang lebih umum digunakan, seperti bahasa Indonesia. Pengembangan model mesin penerjemah dari bahasa Kawi ke dalam bahasa Indonesia dapat menjadi langkah awal dalam proses meningkatkan kemudahan akses belajar khususnya bagi pemula yang tertarik dengan naskah – naskah (lontar) berbahasa Kawi.

Beberapa penelitian telah dilaksanakan dalam pengembangan model Mesin penerjemah, yaitu mesin penerjemah berbasis aturan (*rule based*) (Forcada et al., 2011),

mesin penerjemah berbasis statistik (Koehn, 2005; Koehn et al., 2003) dan mesin penerjemah berbasis jaringan syaraf (*neural network*) (Bahdanau et al., 2015; Cho et al., 2015). Penelitian Vathsala & Holi (2020) dan Yang dkk (2020) membandingkan performa dari mesin penerjemah berbasis statistik dan mesin penerjemah berbasis *neural* dimana mesin penerjemah berbasis *neural*, khususnya dengan menggunakan *recurrent neural network* (RNN) memiliki performa dan akurasi yang lebih baik dari pada mesin penerjemah berbasis statistik. Penelitian

Sennrich & Zhang (2020) melaporkan hasil penelitian terkait mesin penerjemah berbasis RNN yang diimplementasikan untuk bahasa yang memiliki sedikit sumber dan pengguna aktif (*low resource language*) dimana hasilnya lebih baik ketimbang mesin penerjemah berbasis statistik. Berdasarkan penelitian – penelitian tersebut diketahui bahwa Mesin penerjemah berbasis RNN menghasilkan hasil terbaik, dan dapat diimplementasikan untuk pengembangan mesin penerjemah dari bahasa Kawi yang juga bersifat *low resource* ke dalam bahasa Indonesia.

Oleh karena itu dilaksanakan penelitian dengan judul “Mesin Penerjemah Hasil Transliterasi Lontar Bali dalam Bahasa Kawi ke dalam Bahasa Indonesia dengan Menggunakan *Deep Learning*”. Tujuan dari penelitian ini adalah untuk merancang model mesin penerjemah berbasis RNN yaitu RNN, *Bidirectional* RNN, GRU, *Bidirectional* GRU, LSTM, *Bidirectional* LSTM, LSTM dengan mekanisme *Attention*, dan *Bidirectional* LSTM dengan mekanisme *Attention* untuk proses penerjemahan dari bahasa Kawi ke dalam bahasa Indonesia. Beberapa skenario juga diterapkan yaitu dalam proses ekstraksi fitur untuk mengetahui manfaat penerapan *word vector* pada mesin penerjemah. Model – model dari mesin penerjemah tersebut lalu dibandingkan performanya berdasarkan parameter dari BLEU scores. Saat ini masih sedikit penelitian yang mengembangkan model mesin penerjemah dari bahasa lokal khususnya Kawi ke dalam bahasa Indonesia. Dalam penelitian ini juga menghasilkan dataset korpus paralel Kawi – Indonesia yang digunakan dalam kasus perancangan mesin penerjemah berbasis neural dan diimplementasikan secara sederhana menggunakan

microframework flask.

1.2 Identifikasi Masalah

Berdasarkan penjelasan latar belakang yang telah dipaparkan sebelumnya, dapat diidentifikasi beberapa masalah sebagai berikut.

1. Dalam hal meningkatkan kemudahan dalam pola belajar bahasa Kawi diperlukan pemanfaatan mesin penerjemah.
2. Masyarakat sudah semakin paham tentang penggunaan teknologi sehingga mesin penerjemah menjadi pilihan utama dalam pembelajaran khususnya dalam hal penerjemahan antar bahasa.
3. Mesin penerjemah dapat menjadi pilihan utama dalam proses penerjemah ketimbang menggunakan metode konvensional dengan kamus dwibahasa.
4. Belum adanya dataset berbasis kalimat dwibahasa Kawi – Indonesia sebagai data latih model mesin penerjemah *deep learning*.
5. Referensi penerapan mesin penerjemah hasil transliterasi lontar Bali dalam Bahasa Kawi ke dalam Bahasa Indonesia menggunakan *deep learning* belum ditemukan peneliti yang membahasnya.

1.3 Pembatasan Masalah

Berikut merupakan ruang lingkup masalah yang menjadi pokok penelitian, yakni:

1. Penelitian ini menggunakan dataset latih yang disusun dari hasil transliterasi lontar Bali dalam bahasa Kawi ke dalam Bahasa Indonesia sehingga membentuk korpus paralel dwibahasa yang terdiri dari 1088 baris kalimat bahasa Kawi dan hasil terjemahan dalam bahasa Indonesia dari 3 sumber korpus paralel, yaitu *Musala Parwa*, *Singhalangghyala Parwa*, dan *Phalawakya Kapi Parwa*.
2. Model mesin penerjemah *deep learning* yang digunakan adalah RNN, *Bidirectional* RNN, GRU, *Bidirectional* GRU, LSTM, *Bidirectional*

LSTM, LSTM dengan mekanisme *Attention*, dan *Bidirectional LSTM* dengan mekanisme *Attention*.

3. Evaluasi menggunakan matriks nilai *loss*, dan *BLEU Scores*.
4. *Input* berasal dari kalimat bahasa Kawi yang berasal dari hasil transliterasi lontar Bali.

1.4 Rumusan Masalah

Berdasarkan latar belakang tersebut, dapat dirumuskan permasalahan yang perlu dicarikan solusi adalah sebagai berikut:

1. Bagaimana hasil terjemahan dengan menggunakan model mesin penerjemah RNN, *Bidirectional RNN*, GRU, *Bidirectional GRU*, LSTM, *Bidirectional LSTM*, LSTM dengan mekanisme *Attention*, dan *Bidirectional LSTM* dengan mekanisme *Attention* pada penerapannya di mesin penerjemah hasil transliterasi lontar Bali dalam bahasa Kawi ke dalam Bahasa Indonesia?
2. Bagaimana perbandingan hasil terjemahan pada model mesin penerjemah RNN, *Bidirectional RNN*, GRU, *Bidirectional GRU*, LSTM, *Bidirectional LSTM*, LSTM dengan mekanisme *Attention*, dan *Bidirectional LSTM* dengan mekanisme *Attention* pada penerapan mesin penerjemah hasil transliterasi lontar Bali dalam bahasa Kawi ke dalam Bahasa Indonesia berdasarkan penerapan *word vector*, matriks nilai *loss*, dan *BLEU Scores*?

1.5 Tujuan

Berdasarkan permasalahan yang sudah dirumuskan, adapun tujuan yang ingin dicapai adalah sebagai berikut:

1. Untuk mengetahui hasil terjemahan pada mesin penerjemah hasil transliterasi lontar Bali dalam bahasa Kawi ke dalam Bahasa Indonesia pada model mesin penerjemah RNN, *Bidirectional RNN*, GRU, *Bidirectional*

GRU, LSTM, *Bidirectional LSTM*, LSTM dengan mekanisme *Attention*, dan *Bidirectional LSTM* dengan mekanisme *Attention*.

2. Untuk mengetahui perbandingan terjemahan pada mesin penerjemah hasil transliterasi lontar Bali dalam bahasa Kawi ke dalam Bahasa Indonesia pada model mesin penerjemah RNN, *Bidirectional RNN*, GRU, *Bidirectional GRU*, LSTM, *Bidirectional LSTM*, LSTM dengan mekanisme *Attention*, dan *Bidirectional LSTM* dengan mekanisme *Attention*.

1.6 Manfaat Penelitian

Penelitian ini diharapkan mampu memberikan manfaat diantaranya:

1. Manfaat Akademik

Penelitian ini diharapkan mampu menjadi acuan untuk pemilihan model mesin penerjemah *deep learning* yang tepat untuk kasus pada mesin penerjemah hasil transliterasi lontar Bali dalam bahasa Kawi ke dalam Bahasa Indonesia

2. Manfaat Praktis

Manfaat praktis dari penelitian ini adalah diharapkan dapat membantu proses penerjemah Bahasa Kawi ke Bahasa Indonesia dan dapat dijadikan referensi sehingga bisa mempercepat proses penerjemahan hasil transliterasi dokumen – dokumen lontar Bali. Seiring berjalannya waktu mesin penerjemah ini dapat terus dikembangkan melalui penambahan dataset paralel korpus dan pengembangan model mesin penerjemah.