

Pengembangan Transliterasi Offline Teks Aksara Bali dengan Pendekatan Metode Rule-Based Berdasarkan Unicode

Rolando Alex Richo¹, Ahmad Asroni², Gede Indrawan³, Luh Joni Erawati Dewi⁴

Universitas Pendidikan Ganesha
Jalan Udayana No.11, Singaraja, Buleleng, Bali, Indonesia

e-mail: rolando@undiksha.ac.id¹, ahmad.asroni@undiksha.ac.id², gindrawan@undiksha.ac.id³,
joni.erawati@undiksha.ac.id⁴

Received : November, 2023

Accepted : December, 2023r

Published : December, 2023

Abstract

Natural language and writing systems have evolved in response to the human communication needs, creating a variety of languages and writing systems that reflect global cultures and societies. One form of cultural preservation is through educational regulations that incorporate the Balinese script as a mandatory local subject. In this era of internet development, it greatly assists teachers in seeking references online. However, this becomes a challenge in some remote and underdeveloped areas (3T/Dasus) that experience limited internet access, impacting teachers and educational facilitators in carrying out their teaching duties, leading to suboptimal Balinese script learning.

The research proposes the use of a rule-based machine translation method for a mobile-based with offline transliteration application from Latin to Balinese script that can function even without internet connectivity. This application utilizes Unicode Bali as the standard font for the Balinese script, involving two stages of input extraction, whether in Latin or Balinese script. These stages employ a word boundary method to extract syllables that can be processed through a rule-based algorithm using a key-value method.

Keywords: rule based, word boundary, offline transliteration, aksara

Abstrak

Bahasa alami dan sistem penulisan berkembang sebagai respons kebutuhan komunikasi manusia, menciptakan beragam bahasa dan sistem penulisan yang mencerminkan budaya dan masyarakat global. Salah satu wujud pelestarian budaya yaitu melalui regulasi pendidikan yang memasukkan aksara Bali sebagai mata pelajaran lokal wajib. Di era perkembangan internet saat ini sangat membantu guru dan tenaga pengajar mencari referensi di internet namun hal ini menjadi kendala di beberapa daerah 3T/Dasus yang mengalami keterbatasan akses internet sehingga berdampak kepada guru dan tenaga kependidikan /penggerak dalam melakukan tugasnya dalam mengajar murid sehingga pembelajaran aksara bali tidak optimal .

Pada penelitian mengusulkan penggunaan metode penerjemahan mesin berbasis aturan untuk aplikasi pembelajaran transliterasi offline Latin ke Aksara Bali berbasis mobile dapat tetap berjalan dalam kondisi tidak terhubung internet. Aplikasi ini menggunakan Unicode Bali sebagai font standar Aksara Bali, dengan melakukan ekstraksi inputan/masukan baik berupa latin atau aksara yang kemudian di ekstraksi dengan dua tahapan dimana tahapan tersebut menggunakan metode word boundary dengan sehingga didapat suku kata yang dapat di proses ke dalam algoritama rule based menggunakan metode key-value.

Kata Kunci: rule based, word boundary, transliterasi offline, aksara

1. PENDAHULUAN

Bahasa alami dan sistem penulisan berkembang secara alamiah sebagai hasil dari kebutuhan komunikasi antara manusia. Seiring berjalannya waktu, manusia menciptakan beragam bahasa dan sistem penulisan yang sesuai dengan beragam budaya dan masyarakat di seluruh dunia. Bahasa dan aksara menjadi sarana penting dalam menyampaikan ide, berbagi informasi, serta menjaga keragaman warisan budaya.

Di Indonesia, keberagaman aksara adalah salah satu ciri khas yang mencirikan negara ini, dengan berbagai kelompok etnis dan budaya yang memiliki sistem penulisan khas mereka sendiri. Sebagai contoh konkret adalah aksara Bali, yang memiliki sejarah yang kaya dan mendalam dalam budaya dan komunikasi di Bali.

Bahasa dan aksara lokal yang terancam punah [1] merupakan bagian tak terpisahkan dari kekayaan budaya lokal, untuk itu di perlukan upaya pelestarian agar pengetahuan dan penggunaannya dapat terus bertahan dengan dampak yang lebih luas. Pemerintah Bali telah mengambil langkah-langkah pelestarian melalui peraturan Gubernur Bali [2][3][4]. Melalui semua regulasi pelestarian tersebut, Bahasa Bali (termasuk seluruh aspek Aksara Bali) telah dijadikan sebagai mata pelajaran lokal yang wajib diajarkan, dimulai dari sekolah dasar hingga sekolah menengah di Provinsi Bali Namun, penggunaannya terbatas dan cenderung tergantikan oleh huruf Latin dalam kehidupan sehari-hari. Hal ini menyebabkan banyak masyarakat, terutama anak-anak, tidak terampil menulis menggunakan aksara Bali, meskipun tinggal di Bali. Kebanyakan keluarga telah mengajarkan anak-anak mereka huruf Latin sejak kecil, sehingga aksara Bali semakin terpinggirkan dengan perkembangan zaman[5].

Salah satu cara dalam menjaga dan melestarikan kebudayaan lokal adalah dengan mengadopsi ke dalam teknologi dalam hal ini menggunakan penerjemahan mesin berbasis aturan yang mana salah satu pendekatan dalam penerjemahan bahasa alami bergantung pada pengetahuan tata bahasa. Kemampuan mesin penerjemah dalam metode ini bergantung pada aturan tata bahasa yang diterapkan dalam sistem yang di rancang. Implementasi aturan ini

tidak mengikuti pola tertentu, melainkan bergantung pada aturan bahasa asal dan bahasa sasaran.

Penerjemahan mesin (Machine Translation/MT) adalah salah satu aplikasi dari pengolahan bahasa alami [6]. Di era komputasi dan konektivitas internet yang semakin meluas, penelitian dan pengembangan di bidang ini terus berkembang dengan cepat untuk mendukung komunikasi multibahasa dalam komunitas global saat ini.

Permasalahan muncul di sisi teknologi, di mana banyak penerjemahan mesin menjadi sangat tergantung pada internet. Namun, dalam kondisi tertentu, beberapa wilayah sulit untuk mendapatkan koneksi internet yang stabil terutama yang menjadi kendala di beberapa daerah Tertinggal, Terdepan dan Terluar (3T/Dasus) sehingga berdampak kepada guru dan tenaga kependidikan/penggerak dalam melakukan tugasnya sebagai pengajar.

Penelitian ini mengusulkan metode penerjemahan mesin berbasis aturan (Rule-Based Machine Translation/RBMT) pada aplikasi pembelajaran berbasis transliterasi *offline* latin ke Aksara Bali menggunakan metode key-value pairing sehingga dapat berjalan tanpa harus terhubung dengan koneksi internet. Aplikasi ini menggunakan Unicode Bali [7][8] sebagai font standar Aksara Bali pada sistem perangkat mobile. Penelitian ini merupakan kontribusi penulis dalam pengembangan dari penelitian sebelumnya [9][10][11][12] dalam melakukan proses transliterasi *offline* dari tulisan Latin ke Aksara Bali dan sebaliknya.

2. METODE PENELITIAN

Untuk menerjemahkan teks, diperlukan lebih dari sekadar menukar frasa dari bahasa A ke bahasa B. Hal ini sangat penting untuk memahami aturan tata bahasa dari bahasa yang bersangkutan dan memperhatikan norma budaya yang berlaku di wilayah tempat bahasa tersebut digunakan. Hal lain yang juga perlu diperhatikan adalah proses terjemahan juga harus cepat, bebas dari kesalahan penafsiran, dan konsisten. Tantangannya adalah menemukan keseimbangan antara ketepatan dan kecepatan, serta antara otomatisasi dan intervensi manusia. Selama perkembangan sejarah terjemahan mesin, metodenya sebagian

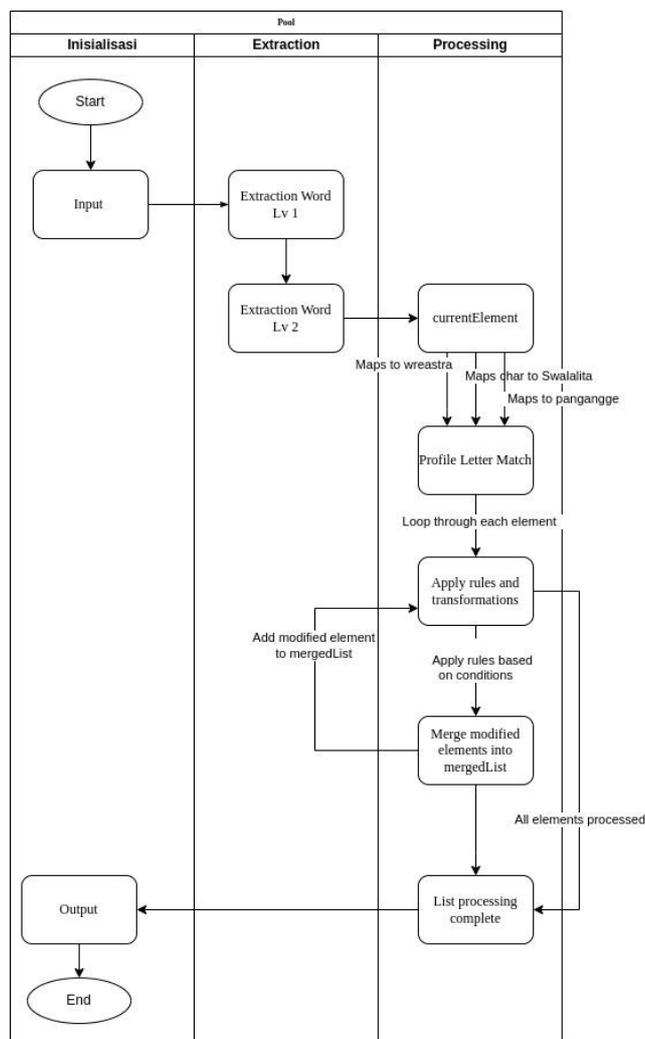
besar didasarkan pada aturan dan statistik, yang pada beberapa kasus terasa kaku dan tidak alami dalam menangkap bahasa alami. Terjemahan mesin berbasis aturan (RBMT) di dasarkan pada informasi yang diprogramkan

dengan menentukan *key-value* yang kemudian dihubungkan dengan aturan-aturan yang melekat pada bahasa target sebagaimana sebuah kata atau frasa dalam bahasa sumber harus dibaca dalam bahasa target seperti dalam penelitian ini adalah aksara bali.

2.1 Rancangan Penelitian

Rancangan penelitian ini bertujuan mengembangkan sistem terjemahan bahasa dengan pendekatan rule-based yang memahami batas kata (word boundary) [13] dan dapat beroperasi secara *offline*. Proses penelitian melibatkan pengembangan algoritma terjemahan, pengujian, evaluasi, perbaikan, dan analisis hasil. Kriteria evaluasi mencakup akurasi terjemahan, kecepatan operasi offline, dan kemudahan penggunaan. Diharapkan penelitian ini akan menghasilkan aplikasi terjemahan yang handal dan efisien.

Penelitian yang dilakukan dengan analisis kuantitatif kemudian dilanjutkan dengan analisis kualitatif deskriptif untuk memperkuat analisis data berdasarkan literatur atau jurnal yang berkaitan dengan subjek penelitian. Pada penelitian ini penulis membagi dalam tiga tahapan atau proses seperti Gambar 1 yaitu dalam proses inialisasi, merupakan input berupa huruf Latin dimasukkan ke dalam aplikasi, misalnya, "Aksara".



Gambar 1. Alur Aplikasi

Kemudian pada proses PreProcessing hasil inputan pada aplikasi di lakukan ekstraksi untuk di lakukan normalisasi hal ini akan berulang sebanyak dua kali proses normalisasi. Pada tahap pertama menghasilkan urutan dalam array yang belum sempurna dalam ejaan Latin, yaitu "[a, ks, a, ra]".

Format tersebut akan di lanjut dengan normalisasi pada tahap dua dilakukan untuk mendeteksi huruf yang terkait dengan konsonan dan vokal, menghasilkan urutan "[a, k, sa, ra]". Setelah proses ekstraksi selesai, kemudian di lanjut dengan proses pengolahan atau processing dengan cara melakukan perulangan Pada currentElement, pemetaan dilakukan menggunakan algoritma berbasis aturan[14][15] yang terdiri dari huruf Latin dan Unicode, menggunakan metode key-value pairs. Hal ini bertujuan untuk memastikan kinerja smartphone tidak terbebani, dan menghindari penggunaan ruang penyimpanan pada

perangkat jika sistem basis data diimplementasikan.

Mengacu kepada hasil ekstraksi tahap dua "[a, k, sa, ra]" sebagai contoh, setiap huruf melewati iterasi untuk menemukan karakter yang cocok. Pada iterasi pertama, huruf "a" dalam Pangangge Swara [16] adalah huruf vokal independen karena vokal ini berada di awal kata. Oleh karena itu, huruf "a" menggunakan "a kara" atau "1B05". Iterasi kedua untuk huruf "k" menggunakan aksara Wreastra "ka" + " " simbol khusus untuk meredam huruf "a", yang direpresentasikan dalam Unicode sebagai "1B13+1B44". Iterasi ketiga untuk huruf "sa" menggunakan aksara Swalalita, ditandai sebagai "sa sepa" dengan Unicode "1B44+1B31". Iterasi array terakhir untuk "ra" menggunakan aksara Wreastra dengan Unicode "1B2D". Setelah seluruh proses iterasi selesai, karakter Unicode kemudian digabungkan dan diatur menjadi bentuk kalimat, yang kemudian ditampilkan pada aplikasi sebagai output Aksara.

```
I/flutter (12364): [extraction] input ==> aksara
I/flutter (12364): [extraction] check currentPairs ==> [a, ks, a, ra]
I/flutter (12364): [extraction] letter r and a is consonant and vocal
I/flutter (12364): [extraction] groupConsonan ==> [a, k, sa, ra]
I/flutter (12364): [processing] last index 3
I/flutter (12364): [processing] currentElement ==> a ==>index 0
I/flutter (12364): [processing] follow all to swara is ==> false
I/flutter (12364): [processing] follow all to swalalita is ==> true
I/flutter (12364): [processing] currentElement ==> k ==>index 1
I/flutter (12364): [processing] follow all to swara is ==> false
I/flutter (12364): [processing] follow all to swalalita is ==> true
I/flutter (12364): [processing] currentElement ==> sa ==>index 2
I/flutter (12364): [processing] detect swalalita elemen 2 in []
I/flutter (12364): [processing] follow all to swara is ==> false
I/flutter (12364): [processing] follow all to swalalita is ==> false
I/flutter (12364): [processing] currentElement ==> ra ==>index 3
I/flutter (12364): [processing] checking wreastra 3 in []
I/flutter (12364): [processing] follow all to swara is ==> false
I/flutter (12364): [processing] follow all to swalalita is ==> false
```

Gambar 2. Proses iterasi

2.2 Instrument Penelitian

Instrumen yang akan digunakan dalam penelitian ini adalah kuesioner berdasarkan Model USE (Usefulness, Satisfaction, Ease of Use) dengan skala Likert [17]. Kuesioner tersebut akan disebar oleh penulis guna

mengumpulkan data yang akan menjadi dasar untuk menilai validitas dan reliabilitas aplikasi dalam mengadopsi algoritma terjemahan mesin berbasis aturan (*Rule-based Machine Translation*). Variabel yang diukur menggunakan skala Likert adalah *usefulness*, *satisfaction*, dan *ease of use*.

suku kata yang bermakna suci atau sehingga mampu menentukan output yang tepat. kemudian aplikasi ini juga dapat berjalan tanpa harus menggunakan atau terkoneksi dengan internet dan tanpa harus menggunakan basis data dikarenakan pada inputan langsung dilakukan iterasi terhadap suku kata untuk dikonversikan kedalam Unicode serta dapat mengkonversikan huruf latin tanpa harus bahasa tertentu. Kemudian metode *word boundary* sangat bermanfaat dalam memperbaiki atau

memanipulasi masukan string, sehingga membantu proses pembacaan algoritma berbasis aturan.

PERNYATAAN PENGHARGAAN

Penulis berterima kasih kepada Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi atas dukungannya dalam penelitian ini yang mengeksplorasi topik riset teknologi untuk data informasi salah satu bentuk kearifan lokal di Indonesia.

DAFTAR PUSTAKA

- [1] N. K. S. Pratiwi and P. S. Oktarina, "Pentingnya Pelestarian Bahasa Bali Pada Pendidikan Formal," 2018.
- [2] Bali Government, "Bali Government Regulation No. 3 on Balinese Language, Script, and Literature [Peraturan Pemerintah Bali No. 3 tentang Bahasa, Aksara, dan Sastra Bali]," no. 379, 1992. https://bphn.go.id/data/documents/perda_3_1992.pdf
- [3] B. Government, "Bali Governor Circular Letter No. 3172 about The Usage of Balinese Traditional Clothing and Balinese Scripts [Penggunaan Busana Adat Bali dan Aksara Bali]." 2019. <https://jdih.baliprov.go.id/produk-hukum/peraturan/abstrak/24741>
- [4] Bali Government, "Peraturan Pemerintah Bali 80 / 2018 mengenai Perlindungan Bahasa, Aksara, dan Sastra Bali," 2018. <https://jdih.baliprov.go.id/produk-hukum/peraturan/abstrak/24665>
- [5] I. P. A. Sulaksana, K. Agustini, and I. M. A. Pradnyana, "Pengaruh Aplikasi Pembelajaran Pengenal Aksara Bali Terhadap Prestasi Belajar Siswa Kelas III pada Mata Pelajaran Bahasa Bali di SD Negeri 1 Kawan," *Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika (KARMAPATI)*, vol. 10, no. 2, pp. 150–159, 2021.
- [6] Z. Z. Linn, M. Moh Aung, M. S. Hlaing, H. E. Hlaing, P. E. Phyu, and K. K. Htay, "Rule-based Machine Translation from Myanmar to Kayah-Li Language," 2017.
- [7] Unicode Consortium, *The Unicode® Standard*. Unicode Consortium, 2021. <https://www.unicode.org/reporting.html>.
- [8] I. Habibi, "Pemrosesan Teks Berbasis Standar Unicode Aksara Bali." <http://www.unicode.org>
- [9] G. Indrawan, I. K. Paramarta, K. Agustini, and Sariyasa, "Latin-to-Balinese script transliteration method on mobile application: A comparison," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 10, no. 3, pp. 1331–1342, 2018, doi: 10.11591/ijeecs.v10.i3.pp1331-1342.
- [10] I. N. Jampel, G. Indrawan, and I. W. Widiana, "Accuracy analysis of Latin-to-Balinese script transliteration method," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 3, pp. 1788–1797, 2018, doi: 10.11591/ijece.v8i3.pp1788-1797.
- [11] D. A. P. P. Sanjani, G. Indrawan, and I. G. A. Gunadi, "Pengembangan Metode Pemilahan Suku Kata Pada Transliterasi Teks Latin Ke Aksara Bali Berbasis Finite State Machine Dengan Font Noto Serif Balinese," *Jurnal Ilmu Komputer Indonesia (JIK)*, vol. 6, no. 2, 2021.
- [12] G. Indrawan, I. G. A. Gunadi, and I. K. Paramarta, "Towards Ubiquitous Learning of Balinese-to-Latin Script Transliteration as Part of Balinese Language Education," vol. 438, no. 80, pp. 112–115, 2020, doi: 10.2991/assehr.k.200513.025.
- [13] T. A. Nguyen, M. de Seyssel, R. Algayres, P. Roze, E. Dunbar, and E. Dupoux, "Are word boundaries useful for unsupervised language learning?," 2022. Available: <http://arxiv.org/abs/2210.02956>
- [14] P. W. Pratama, A. Aranta, and F. Bimantoro, "Rancang Bangun Aplikasi

- Transliterasi Aksara Latin menjadi Aksara Sasak Menggunakan Algoritma Rule Based Berbasis Android," *Jurnal Teknologi Informasi, Komputer dan Aplikasinya (JTika)*, vol. 3, no. 2, pp. 232–243, 2021. <http://jtika.if.unram.ac.id/index.php/JTika/article/view/157>
- [15] A. Rahman, D. T. Murdiansyah, and K. M. Lhaksana, "Silabifikasi Menggunakan Metode Rule-Based dalam Bahasa Indonesia," 2021.
- [16] I. B. Made Suasta, I. B. Mayun, and W. Rupa, "Modernisasi dan pelestarian perkembangan metode dan teknik penulisan aksara Bali," 1996.
- [17] Tugiman, Herman, and A. Yudhana, "Uji Validitas Dan Reliabilitas Kuesioner Model Utaut Untuk Evaluasi Sistem Pendaftaran Online Rumah Sakit," 2022. <http://jurnal.mdp.ac.id>