

Klasifikasi Aksara Bali Berbasis Suara Dengan Metode KNN dan FastDTW

I Putu Rama Anadya^{a1}, I Gusti Agung Gede Arya Kadyanan^{a2}, I Dewa Made Bayu Atmaja Darmawan^{b3}, Cokorda Rai Adi Pramatha^{b4}, Anak Agung Istri Ngurah Eka Karyawati^{b5}, Ngurah Agus Sanjaya ER^{b6}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana Badung, Bali, Indonesia

¹gusrama18@gmail.com

²gungde@unud.ac.id

³dewabayu@unud.ac.id

⁴cokorda@unud.ac.id

⁵eka.karyawati@unud.ac.id

⁶agus_sanjaya@unud.ac.id

Abstrak

Bahasa Bali merupakan bahasa yang digunakan masyarakat Bali untuk berkomunikasi. Untuk menulis bahasa Bali, mereka dapat menggunakan Aksara Bali. Pada era globalisasi ini, aksara Bali mulai di tinggalkan oleh masyarakat atau mulai menuju kepunahan. Saat ini aksara Bali membutuhkan pengembangan di dunia digital sehingga mampu menyesuaikan dengan perkembangan jaman, khususnya pada media pembelajaran seperti sebuah aplikasi. sebelum membangun aplikasi perlu dibuatkan sebuah sistem yang dapat mendukung aplikasi tersebut. Penulis membuat sistem klasifikasi dengan metode KNN dan FastDTW untuk mengenali suara aksara Bali khususnya aksara Wreasta. Hasil penelitian yang telah dilakukan mendapatkan akurasi yang baik yaitu 100% pada model klasifikasi dengan nilai k-3, k-5, k-7 dan k-9 serta akurasi terendah pada k-19 dengan akurasi 94,44%

Kata Kunci: Aksara Bali, pengenalan suara, KNN, FastDTW, suara

1. Pendahuluan

Bahasa Bali merupakan bahasa yang digunakan masyarakat Bali untuk berkomunikasi. Untuk menulis bahasa Bali, mereka dapat menggunakan Aksara Bali. Pada era globalisasi ini, aksara Bali mulai di tinggalkan oleh masyarakat atau mulai menuju kepunahan[2].

Untuk meningkatkan minat masyarakat dalam menggunakan aksara Bali, pemerintah mengeluarkan Perda No. 1 tahun 2018 Tentang Bahasa, Aksara, dan Sastra Bali juga Peraturan Gubernur No. 80 Tahun 2018 tentang Perlindungan dan Penggunaan Bahasa, Aksara dan Sastra Bali serta Penyelenggaraan Bulan Bahasa Bali. Saat ini aksara Bali membutuhkan pengembangan di dunia digital sehingga mampu menyesuaikan dengan perkembangan jaman, khususnya pada media pembelajaran.

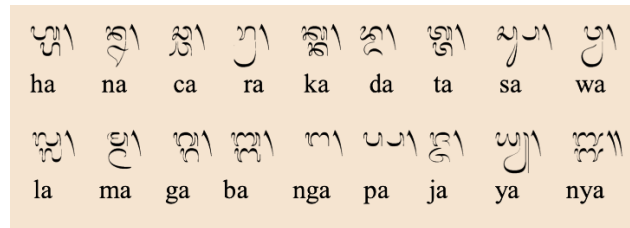
Salah satu pengembangan digital untuk media pembelajaran aksara Bali yaitu dengan membuat sebuah aplikasi yang dapat digunakan sebagai media pembelajaran baru. Dengan adanya aplikasi akan membuat masyarakat lebih tertarik untuk belajar aksara Bali. Akan tetapi sebelum membuat aplikasi, tentu perlu dikembangkan sebuah sistem untuk mendukung sebuah aplikasi tersebut. pada penelitian ini, penulis akan mengembangkan sebuah sistem klasifikasi aksara Bali berbasis suara dengan menggunakan metode KNN sebagai metode klasifikasi dan metode FastDTW sebagai metode untuk mengukur kedekatan antara dua data. Penulis akan menggunakan data berupa suara pengucapan aksara Bali khususnya aksara *Wreasta* sebanyak 18 karakter dan mengukur kemampuan metode KNN dan FastDTW dalam mengenali aksara Bali tersebut.

2. Metodologi Penelitian

2.1. Aksara Bali

Aksara Bali merupakan sekumpulan karakter, lambang atau tanda yang digunakan oleh masyarakat Bali untuk menulis aksara Bali [4]. aksara Bali dapat dibagi menjadi empat bagian yaitu aksara *Wreasta*, aksara *Swalalita*, Aksara *Wicaksana*, dan Aksara *Modre*. Akan tetapi penelitian ini hanya akan berfokus pada aksara *Wreasta*.

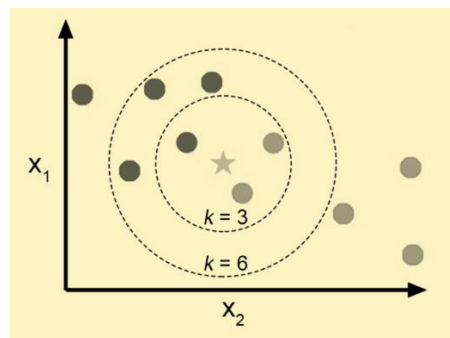
Aksara *Wreasta* merupakan aksara yang digunakan yang paling umum digunakan oleh masyarakat untuk menulis bahasa Bali[4]. Aksara *Wreasta* terdiri dari 18 karakter yang dapat dilihat pada Gambar 1.



Gambar 1. 18 Karakter dari Aksara *Wreasta*[5]

2.2. K-Nearest Neighbor

Dikutip dari penelitian yang dilakukan oleh Wu pada tahun 2009, metode *K-Nearest Neighbor* (KNN) merupakan suatu metode yang digunakan dalam proses klasifikasi yang mana dilakukan dengan mencari kelompok k objek dalam data latih yang paling mirip dengan data uji[3]. Klasifikasi ini didasarkan pada pembelajaran dengan analogi, yaitu mencari kelas terdekat dari objek yang dimiliki dengan *dataset* yang sudah ada dengan mengukur jarak dari setiap objek dengan *dataset* sebanyak jumlah *dataset* yang dimiliki. Untuk ilustrasi dari algoritma KNN dapat dilihat pada Gambar 2.



Gambar 2. Ilustrasi metode KNN

Adapun algoritma dari KNN secara umum adalah sebagai berikut.

1. Menentukan jumlah tetangga atau K yang digunakan untuk pertimbangan menentukan kelas dari data uji.
2. Hitung jarak dari fitur data uji dengan masing-masing fitur data dalam *dataset*. Dalam penelitian ini, metode pengukuran jarak yang digunakan adalah metode *fastDTW*.
3. Ambil sejumlah k data terdekat
4. Pilih kelas yang paling banyak dari k data yang diambil

2.3. Fast Dynamic Time Warping

Fast Dynamic Time Warping (*FastDTW*) adalah versi terbaru dari *Dynamic Time Warping* (*DTW*) yang dikembangkan oleh Salvador dan Chan pada tahun 2007 untuk mengatasi masalah kompleksitas kuadratik yang terjadi pada *DTW*[1]. Menurut [1] *FastDTW* menggunakan pendekatan multilevel dengan tiga operasi kunci yaitu *coarsening*, *projection* dan *refinement*

2.3.1 Coarsening

FastDTW melakukan pemangkasan atau mengubah ukuran data tetapi data masih dapat merepresentasikan kurva yang sama seakurat mungkin dengan data yang lebih sedikit.

2.3.2 Projection

FastDTW mencari *warp path* dengan jarak minimum pada resolusi yang lebih rendah, dan gunakan jalur lusi itu sebagai tebakan awal untuk jalur lusi jarak minimum resolusi yang lebih tinggi.

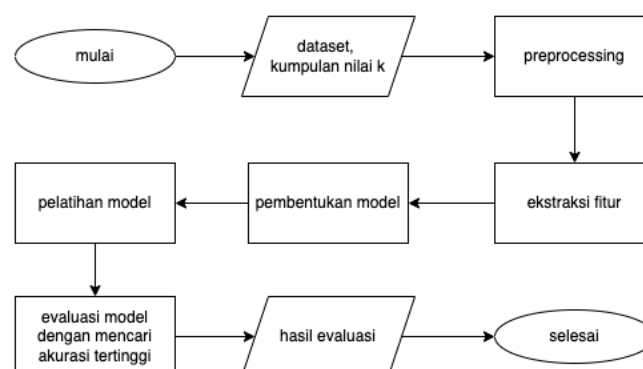
2.3.3 Refinement

FastDTW memperbaiki jalur lusi yang diproyeksikan dari resolusi yang lebih rendah melalui penyesuaian *warp path*.

3. Hasil dan Pembahasan

3.1. Alur Penelitian

Berdasarkan alur penelitian pada Gambar 3, proses pertama yaitu menginputkan dataset dan nilai k yang akan digunakan sebagai parameter pengujian. dataset yang digunakan merupakan data primer berupa *file* suara berformat *waveform* (.wav). kemudian setiap data akan melalui *preprocessing* dan ekstraksi fitur untuk mendapatkan fitur dari setiap audio. Setelah itu data akan dibagi menjadi data latih dan data uji. Kemudian membuat model klasifikasi dengan algoritma KNN dengan metode pengukur jarak fastDTW. Model akan dilatih dengan data latih kemudian akan dievaluasi dengan data uji. Kelas yang didapat akan dibandingkan dengan kelas sebenarnya dari data uji sehingga didapatkan akurasi. Hasil evaluasi akan menjadi luaran dari penelitian yang dilakukan.



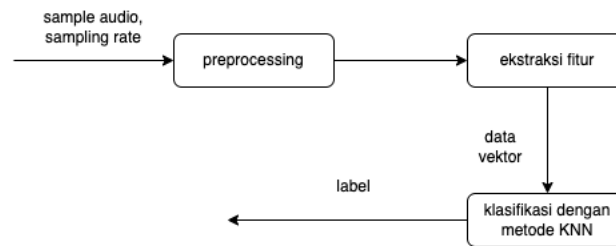
Gambar 3. Alur Penelitian

3.2. Data Penelitian

Data yang digunakan merupakan data primer yang berupa file suara dengan format *waveform* (.wav) yang diperoleh dengan menggunakan *voice recorder* bawaan smartphone. setiap orang akan direkam 27 kali untuk setiap labelnya. Label terdiri dari 18 karakter yang merupakan bagian dari aksara Wreasta. Total data yang dimiliki yaitu 486 data yang akan dibagi menjadi 70% data latih dan 30% data uji.

3.3. Desain Sistem

Pada penelitian ini penulis menggunakan metode KNN dan FastDTW untuk melakukan klasifikasi pada suara. sebelum suara di klasifikasi, diperlukan pra proses seperti penghilangan *noise* dan ekstraksi fitur untuk mengubah suara ke bentuk vektor agar suara dapat di klasifikasi. adapun alur dari program dapat dilihat pada Gambar 4.



Gambar 4. Alur klasifikasi dengan metode KNN

Dilihat pada Gambar 4, program menerima inputan berupa *sample audio* dan *sampling rate* dari data suara. Kemudian masuk ke pra proses untuk menghilangkan *noise* dari suara. setelah pra proses, *sample audio* tersebut akan di ekstraksi menggunakan metode ekstraksi fitur. Dalam penelitian ini penulis menggunakan metode MFCC dengan koefisien 13. Setelah didapatkan fitur dari audio, lanjut ke tahap klasifikasi dimana fitur tersebut dicocokkan dengan fitur data yang ada pada sistem. sesuai dengan algoritma KNN, sebanyak k data dengan jarak terdekat akan diambil dan label dari data yang paling banyak muncul akan digunakan sebagai label dari data suara yang baru.

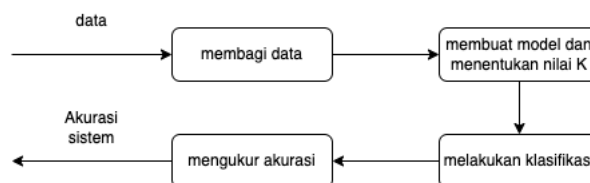
Sebelum sistem dapat melakukan klasifikasi, perlu dilakukan pelatihan terlebih dahulu untuk menghasilkan model yang paling optimal untuk melakukan klasifikasi. Pelatihan yang akan dilakukan pada penelitian ini yaitu dengan mencari nilai k terbaik dari model klasifikasi. Adapun alur dari pelatihan yang akan dilakukan dapat dilihat pada Gambar 5.

Dapat dilihat pada Gambar 5. Sistem menerima inputan berupa kumpulan data. Kemudian data tersebut akan dibagi menjadi data uji dan data latih dengan perbandingan 3:7. Kemudian sistem membuat model klasifikasi dengan metode KNN dan menentukan nilai K. setelah itu sistem akan melakukan klasifikasi dengan mengukur setiap data uji dengan data latih yang dimiliki. Kemudian sistem akan mengukur akurasi dari model dengan membandingkan label data uji dari hasil klasifikasi dengan label data uji sebenarnya. Dari hal tersebut akan didapatkan nilai akurasi dari siste.

Proses ini akan diulangi sebanyak 9 kali dengan nilai k yaitu k-3, k-5, k-7, k-9, k-11, k-13, k-15, k-17 dan k-19. Nilai K yang memiliki akurasi tertinggi akan digunakan sebagai model terbaik

3.4. Pengujian Sistem

Sistem yang dibangun oleh penulis menggunakan bahasa pemrograman *Python* pada aplikasi *Jupyter Notebook*. Adapun hasil pengujian dari penelitian ini dapat dilihat pada Tabel 1.



Gambar 5. Alur Pelatihan Sistem

Dari hasil pengujian yang telah dilakukan didapat hasil bahwa model terbaik memiliki nilai k-3,k-5,k-7 dan k-9 dengan akurasi 100% dan model dengan akurasi terendah ada pada model dengan nilai k-19 dengan akurasi 94,44%.

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dihasilkan sebuah sistem klasifikasi dengan menggunakan metode KNN sebagai metode klasifikasi yang didalamnya terdapat metode FastDTW untuk mengukur kedekatan antara kedua data. melalui pengujian yang telah dilakukan didapat hasil bahwa model terbaik memiliki nilai K-3,K-5,K-7 dan K-9 dengan akurasi 100% dan model dengan akurasi terendah ada pada model dengan nilai k-19 dengan akurasi 94,44%. Dapat disimpulkan bahwa

penggunaan metode KNN dan FastDTW dapat digunakan dalam pengenalan aksara Bali khususnya aksara Wreasta berbasis suara.

Tabel 1. Hasil Pengujian

No	Nilai K	Akurasi	Waktu Komputasi
1	K-3	100%	6,2 detik
2	K-5	100%	6,5 detik
3	K-7	100%	6,5 detik
4	K-9	100%	6,5 detik
5	K-11	98,15%	6,4 detik
6	K-13	98,15%	6,5 detik
7	K-15	98,15%	6,4 detik
8	K-17	96,30%	6,5 detik
9	K-19	94,44%	6,4 detik

References

- [1] Haviana, S.F.C. 2015. Sistem Gesture Accelerometer dengan Metode Fast Dynamic Time Warping (FastDTW). *Jurnal Sistem Informasi Bisnis*. 5, 10-21456.
- [2] Indrawan, G. Paramarta, I.K., Agustini, K., Sariyasa. 2018. Latin-to-Balinese Script Transliteration Method on Mobile Application: A Comparison. *Indonesian Journal of Electrical Engineering and Computer Science*. No. 3 (Vol. 10). Hal 1331-1342
- [3] Leidiyana, H. 2013. Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor. *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic*. No. 1 (Vol. 1). Hal 65-76
- [4] Suasta, I.B.M, Mayun, I.B., Rupa, W. 1996. *Modernisasi dan Pelestarian : Perkembangan Metode dan Teknik Penulisan Aksara Bali*. Jakarta: Proyek Pengkajian dan Pembinaan Nilai-nilai Budaya Direktorat Sejarah dan Nilai Tradisional Direktorat Jenderal Kebudayaan Departemen Pendidikan dan Kebudayaan.
- [5] Suwija, I.N. 2015. *Menulis Bahasa Bali 2*. Diklat. Denpasar: Fakultas Pendidikan Bahasa dan Seni IKIP PGRI Bali

This page is intentionally left blank.