

## IMPLEMENTASI METODE MEL-FREQUENCY CEPSTRAL COEFFICIENT DAN DTW PADA APLIKASI PENGENALAN SUARA TEMBANG SEKAR ALIT

Made Agung Raharja<sup>1</sup>, I Putu Gede Adiatmika<sup>2</sup>, I Nyoman Adiputra<sup>3</sup>,  
Susy Purnawati<sup>4</sup>, I Wayan Supriana<sup>5</sup>

<sup>1,5</sup>Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana  
Jl. Raya Kampus Unud Jimbaran, Kec. Kuta Sel., Kabupaten Badung, Bali, Indonesia

<sup>2,3,4</sup>Departemen Ilmu Faal, Fakultas Kedokteran, Universitas Udayana  
Jalan P.B. Sudirman, Denpasar, Bali, Indonesia

e-mail: [made.agung@unud.ac.id](mailto:made.agung@unud.ac.id)<sup>1</sup>, [ipgadiatmika@unud.ac.id](mailto:ipgadiatmika@unud.ac.id)<sup>2</sup>, [nadip2003@yahoo.com](mailto:nadip2003@yahoo.com)<sup>3</sup>,  
[susy\\_purnawati@unud.ac.id](mailto:susy_purnawati@unud.ac.id)<sup>4</sup>, [wayan.supriana@unud.ac.id](mailto:wayan.supriana@unud.ac.id)<sup>5</sup>

Received : April, 2022

Accepted : April, 2022

Published : April, 2022

### Abstract

*The province of Bali has various types of songs that have different structures and functions, one of which is the Sekar Alit song. Seeing current technological advances, conservation efforts should follow the development of existing technology. This study recognizes the singer's voice with the MeliFrequency iCepstrum Coefficients (MFCC) method used to perform feature extraction, i.e. to obtain a parameter and information regarding the characteristics of a person's voice and the training of voice pattern matching against the voice of the DTW Singer Alit (Dinamic) and the Time Warping Alit method. . So that the final result of this research is an android-based tembang sekar alit learning application that can be used by students to learn more easily and effectively which is called the SekARAI application. The results of usability testing on the application found that the average value was above 3, which means that the SekARAI software that has been implemented has met the Usability element and besides that the software is easy to use and understand by users. The Sekar Alit Song Voice Recognition application has been evaluated using the confusion matrix method, it is found that the MFCC algorithm test results in an accuracy of 76.6%.*

**Keywords:** Sekar Alit, Speech Recognition, and MFCC

### Abstrak

*Provinsi Bali terdapat berbagai jenis tembang yang mempunyai struktur dan fungsi yang berbeda-beda salah satunya adalah Tembang Sekar alit. Melihat kemajuan teknologi saat ini, upaya pelestarian hendaknya mengikuti perkembangan teknologi yang ada. Penelitian ini melakukan pengenalan suara penembang dengan metode Mel Frequency Cepstrum Coefficients (MFCC) digunakan untuk melakukan ekstraksi ciri, yaitu untuk mendapatkan suatu parameter dan informasi mengenai ciri dari suara seseorang dan pencocokan pola suara latih terhadap suara penembang Sekar Alit dan metode DTW (Dinamic Time Warping). Sehingga hasil akhir dari penelitian ini adalah aplikasi pembelajaran tembang sekar alit berbasis android yang dapat digunakan oleh para siswa untuk belajar lebih mudah dan efektif yang disebut Aplikasi SekARAI. Hasil Pengujian usability pada aplikasi didapatkan bahwa nilai rata-rata diatas 3, yang berarti bahwa perangkat lunak SekARAI yang telah diimplementasikan telah memenuhi*

unsur Usability dan selain itu perangkat lunak tersebut mudah digunakan serta dimengerti oleh para pengguna. Aplikasi Pengenalan Suara Tembang Sekar Alit telah dievaluasi dengan metode confusion matrix diperoleh bahwa pengujian algoritma MFCC mendapatkan hasil akurasi sebesar 76,6%.

**Kata Kunci:** Sekar Alit, Speech Recognition, dan MFCC

## 1. PENDAHULUAN

Tembang ialah seni suara yang dibentuk dari beragam laras dan nada sebagai materinya. Seiring dengan perkembangan zaman pembelajaran tembang telah terjadi pergeseran dari penggunaan media pembelajaran konvensional hingga saat ini berkembang menjadi media pembelajaran interaktif berbasis teknologi informasi khususnya pembelajaran tembang Bali. Terdapat berbagai jenis tembang pada umumnya di Bali yang memiliki struktur dan fungsi yang berbeda, yaitu *gegendingan*, tembang *mecepat (sekar alit)*, *kekidungan (sekar madya)*, dan *kekawin/wirama*. Tembang *sekar alit* sangat dekat dengan masyarakat Bali sehingga dapat dijadikan sebagai sarana media untuk memberikan pesan moral dan Pendidikan bagi masyarakat luas [1]. Tembang *sekar alit* pada beberapa dekade tahun lalu masih banyak ditembangkan oleh anak-anak maupun orang dewasa di luar sekolah formal. Hasil dari beberapa penelitian mendeskripsikan eksistensi tembang *macapat* atau *sekar alit* di kalangan generasi muda menurun dan menemukan upaya preventif dalam rangka alkuturasi budaya barat di kalangan generasi muda [2]. Penurunan kualitas dan minat generasi muda untuk belajar tembang Bali disebabkan dari tidak ada inovasi media pembelajaran yang mengikuti perkembangan jaman saat ini. Salah satu pemanfaatan teknologi informasi dan komunikasi yang ada saat ini telah banyak digunakan dalam bidang pendidikan, khususnya untuk perguruan tinggi dan sekolah adalah penggunaan media pembelajaran [3].

SekARAI adalah perangkat lunak yang dibuat dengan metode MFCC (*Mel Frequency Cepstrum Coefficients*) diperuntukan saat melakukan *feature extraction*, adalah untuk mendapatkan informasi dan suatu parameter mengenai ciri dari suara dan pencocokan pola suara latih terhadap suara yang akan di uji menggunakan metode *DTW (Dinamic Time Warping)*. Suara pengguna sebagai penembang tembang *sekar alit* akan secara otomatis

diverifikasi oleh SekARAI kebenarannya sesuai dengan suara acunya yang sudah ada pada sistem yang dapat di instal pada perangkat mobile yang tersambung pada suatu server pengolahan data audio. Fitur lainnya pada SekARAI adalah terdapat teks tembang yang dibagi dalam setiap baitnya dengan ukuran teks tembang yang bisa dibesar kecilkan sesuai dengan keinginan penggunaannya.

Permasalahan yang sering muncul dalam desain perangkat lunak pada interaksi antara manusia dengan komputer yaitu terjadinya salah persepsi oleh manusia (pengguna) perangkat lunak yang ada, sehingga tidak efektif dan efisiensi kerja yang didapat, melainkan menyebabkan kesulitan dalam menggunakan perangkat lunak tersebut. Masalah lainnya adalah perangkat lunak terlalu kompleks sehingga sulit dipelajari, perangkat lunak kurang sesuai berdasarkan kebutuhan *user* dan tidak mewakili kebutuhan yang penting bagi pengguna. Masalah yang terjadi akibat penerapan sistem dapat dihindari dengan menerapkan konsep HCI dengan baik. HCI berfokus pada desain sistem pengguna atau yang biasa disebut dengan *User Centered Design (UCD)*. Sehingga perlu dilakukan evaluasi terhadap aplikasi yang telah diimplementasikan. Evaluasi ini nantinya berguna untuk mengetahui seberapa efektif dan efisien *SekARAI* bagi pengguna [4]. Evaluasi ini juga akan sangat bermanfaat, sebagai salah satu dasar untuk mengembangkan aplikasi Android yang dimiliki, jika kedepannya akan menambah beberapa fitur. Evaluasi yang dapat dilakukan untuk mengetahui sejauh mana penggunaan suatu teknologi atau aplikasi dilakukan, salah satunya dengan menggunakan *usability analysis* [5]. Kegunaan adalah bagian dari ilmu Interaksi Komputer Manusia. yang berfokus mempelajari desain antarmuka dan interaksi antara manusia dan computer [6].

## 2. METODE PENELITIAN

Metode penelitian yang di gunakan dalam penelitian rancang bangun aplikasi pembelajaran Tembang Sekar Alit Menggunakan

Metode *Mel Frequency Cepstral Coefficients (MFCC)* dan *DTW*, disesuaikan dengan metode pengembangan perangkat lunak *System Development Life Cycle (SDLC)* [7]. Secara keseluruhan tahap-tahap penelitian dipaparkan sebagai berikut:

### 2.1 Identifikasi Kebutuhan Sistem

Bagian bagian indentifikasi kebutuhan sistem adalah melakukan indentifikasi kepada fungsionalitas sistem yang paling utama yaitu dengan kata lain adalah yang sangat dibutuhkan oleh *user* atau pengguna sistem. Bagian ini akan menjelaskan tentang analisis kebutuhan dalam pembangunan aplikasi pembelajaran Tembang Sekar Alit Menggunakan Metode *Mel Frequency Cepstral Coefficients (MFCC)*. Terdapat dua jenis analisis yang akan dilakukan pada tahap ini adalah analisis kebutuhan fungsional serta non fungsional.

### 2.2 Data Penelitian

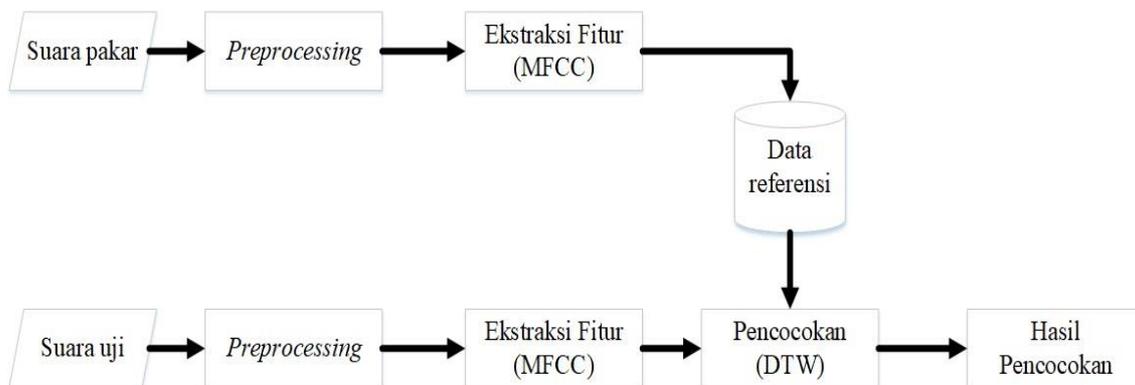
Bagian ini akan membahas mengenai data yang digunakan pada penelitian ini. Data yang digunakan merupakan data primer. Data diperoleh dengan merekam suara dari 2 orang pakar. Hasil rekaman suara dalam format *.wav*. Pada penelitian ini wirama yang dinyanyikan adalah Tembang Sekar Alit. Dimana seorang

pakar akan menyanyikan perbait kalimat sebanyak 4 bait kalimat.

Data hasil rekam suara pakar yang diperoleh akan digunakan pada proses pelatihan dan pengujian. Setiap pakar akan memberikan data suara untuk satu tembang sekar alit sebanyak 3 suara untuk masing-masing kalimat dalam satu tembang sekar alit.

### 2.3 Analisa Sistem

Pada tahap analisis sistem pada aplikasi pembelajaran Tembang Sekar Alit Menggunakan Metode *Mel Frequency Cepstral Coefficients (MFCC)* dan *DTW* adalah mencari informasi tertentu yang dapat dijadikan analisa sebagai kebutuhan pengguna dan mencoba menentukan kebutuhan proses informasi untuk masing-masing aktivitas sistem (*input, output* dan proses). Dalam sistem pengenalan pelafalan tembang Sekar Alit terdapat dua bagian proses yaitu, proses pelatihan dan pengenalan. Berikut adalah gambaran sistem secara garis besar. Pada tahap ini biasanya membutuhkan beberapa metode dalam melakukan pemodelan data serta proses dalam penelitian ini antara lain dapat dijabarkan dalam bagan alir pada gambar 3.2.



Gambar 1. Gambaran Umum Sistem Tembang Sekar Alit Menggunakan Metode MFCC [sumber : [2]]

Berdasarkan Gambar 3.2 diatas yang menunjukkan gambaran umum sistem yang akan di bangun. Nantinya suara pakar dan suara uji akan melalui tahapan preprocessing. Pada tahapan ini kedua file dengan ekstensi *.wav* akan melewati proses DC-Removal, Pre emphasize, Frame Blocking, Windowing, Fast Fourier Transform, Mel Frequency Warping,

Discrete Cosine Transform. Barulah akan masuk pada proses ekstraksi fitur dengan *MFCC (Mel Frequency Cepstrum Coefficients)* hasil dari ekstraksi akan melalui proses Cepstral liftering yang bertujuan untuk meminimalisasi sensitifitas hasil ekstraksi [8]. Fitur dari suara pakar akan disimpan dalam sebuah direktori untuk menjadi data referensi. Selanjutnya fitur-

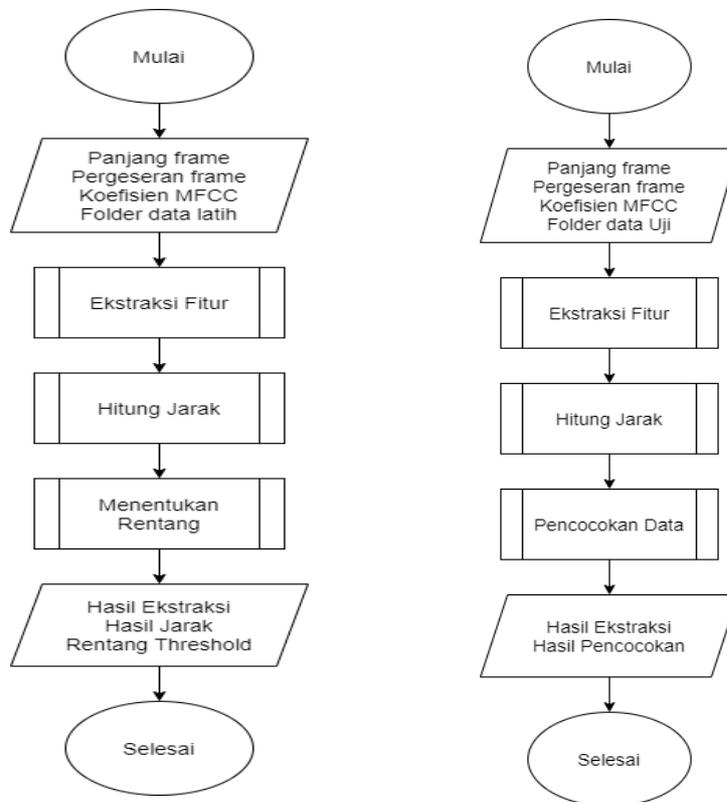
fitur yang dihasilkan oleh suara uji akan dicocokkan dengan data referensi dengan metode DTW.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Deskripsi Data

Pada sub bab berikut dilakukan perancangan sistem dengan diawali proses training. Proses *training* merupakan proses untuk melatih sistem mengenali pola data dengan cara menentukan rentang nilai jarak data benar. Masukkan pada proses *thresholding* berupa panjang *frame* ( $N$ ), panjang pergeseran *frame* ( $M$ ), jumlah koefisien MFCC, dan path letak folder data latih

disimpan. Data latih memiliki tipe data (.wav) dan telah dikelompokkan berdasarkan pada *wirama*. Selanjutnya file suara akan melalui proses ekstraksi fitur [9]. Proses ekstraksi fitur merupakan proses untuk mengekstraksi fitur setiap file latih yang terdapat didalam folder data latih sehingga data dapat melalui proses selanjutnya. Hasil dari proses ekstraksi fitur memuat fitur vektor yang berbentuk list dengan ukuran  $N \times K$ , dimana  $N$  adalah jumlah *frame* yang terbentuk dan  $K$  merupakan jumlah koefisien MFCC yang ditentukan.



Gambar 2. Flowchart sistem proses thresholding dan proses pencocokan

Berdasarkan pada Gambar 2 menjelaskan bahwa hasil dari proses ekstraksi data latih akan digunakan sebagai masukan pada proses menghitung jarak latih. Proses menghitung jarak latih merupakan menghitung jarak hasil ekstraksi data latih dengan hasil ekstraksi suara pakar menggunakan Dynamic Time Warping (DTW). Hasil dari setiap jarak data latih dengan suara pakar akan digunakan sebagai masukan

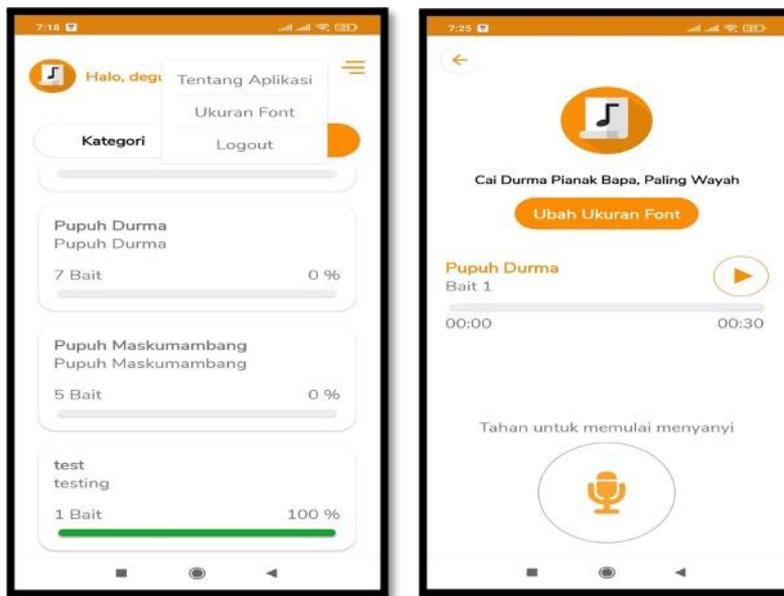
dalam proses perhitungan rentang threshold. Proses ini bertujuan untuk menentukan nilai batas atas dan batas bawah jarak data yang benar, sehingga pada proses testing nanti jika data yang dilakukan uji berada di dalam rentang maka data dianggap benar dan data yang berada diluar rentang maka data dianggap salah.

### 3.3 Implementasi Aplikasi

Pembelajaran Tembang Sekar Alit (SekARAI) Aplikasi pencocokan suara tembang sekar alit menggunakan metode *Mel Frequency Cepstral Coefficients (MFCC)* aplikasi berbasis mobile android dan sisi server berbasis web dibangun dengan menggunakan bahasa pemrograman Python. Sub bab proses pembuatan aplikasi pencocokan suara tembang sekar alit ini dibahas ke dalam empat poin bahasan yang terdiri dari tampilan awal aplikasi, proses login, registrasi, menu utama yang terdiri dari : pemilihan kategori tembang sekar alit, mengubah ukuran font, proses input tembang

oleh user untuk dicocokkan, dan output aplikasi berupa kecocokan suara tembang.

Menu utama dari aplikasi pencocokan suara tembang sekar alit menggunakan (SekARAI) metode *Mel Frequency Cepstral Coefficients (MFCC)* terdiri dari beberapa menu, diantaranya adalah menu login awal, registrasi, menu utama yang terdiri dari : pemilihan kategori tembang sekar alit, ubah ukuran font, menu input suara tembang oleh user untuk dicocokkan, dan menu output hasil kecocokan dapat dilihat pada Gambar 3.



Gambar 3. Judul Gambar Tampilan Menu Ketegori Tembang Sekar Alit

### 3.4 Pengujian dan Evaluasi Sistem

Setelah memperoleh model klasifikasi algoritma MFCC, diperlukan pengujian untuk mengukur performa atau kinerja algoritma MFCC dalam mengklasifikasikan *tembang sekar alit atau pupuh*. Pengujian dilakukan terhadap

data uji sebesar 50 observasi dengan menggunakan metode *confusion matrix* [10]. Diperoleh hasil precision, recall, dan akurasi berdasarkan tabel *confusion matrix* di bawah ini.

Tabel 1. Hasil *confusion matrix*

Klasifikasi		Hasil Kelas Klasifikasi	
Data Aktual	Kelas	Negatif	Positif
	Negatif	22	6
	Positif	6	16

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\%$$

$$= 72.72\%$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\%$$

$$= 72.72\%$$

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$= 76.6\%$$

Dari Tabel 1 confusion matrix, diperoleh bahwa pengujian algoritma MFCC pada 50 observasi data uji mendapatkan hasil akurasi sebesar 76,6%, precision sebesar 72.72%, dan recall sebesar 72.72%.

#### 4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dari proses perancangan, implementasi dan evaluasi dari Implementasi Metode Mel-Frequency Cepstral Coefficient dan DTW Pada Aplikasi Pengenalan Suara Tembang Sekar Alit adalah sebagai berikut

- 1) Proses identifikasi suara *tembang sekar alit* menggunakan metode *Mel Frequency Cepstral Coefficients (MFCC)* dan *DTW* telah berhasil dilakukan dengan menggunakan 50 data set. Pengujian usability pada aplikasi didapatkan bahwa indikator *learnability* merupakan aspek yang mendapat penilaian paling tinggi, dengan nilai 4,21, serta aspek *memorability* menjadi aspek yang memiliki penilaian dengan nilai 4,18, satisfaction indikator ketiga dengan penilaian sebesar 4,17, Efficiency adalah indikator keempat dengan penilaian sebesar 4,03 dan indikator errors dengan nilai sebesar 3,93. Berdasarkan hasil rekapitulasi penilaian Usability yang menunjukkan keseluruhan atribut memiliki nilai penerimaan usability oleh user, rata-rata diatas nilai 3, sehingga dapat disimpulkan bahwa perangkat lunak pembelajaran *tembang Sekar Alit* yang telah dibuat telah memiliki nilai dari aspek

*Usability* dan mudah dipelajari serta mudah dimengerti oleh pengguna atau *user*.

- 2) Aplikasi pembelajaran *tembang sekar alit* menggunakan metode *Mel Frequency Cepstral Coefficients (MFCC)* telah berhasil diimplementasikan dengan masukan berupa file audio suara jenis .wav. Implementasi Metode Mel-Frequency Cepstral Coefficient dan DTW Pada Aplikasi Pengenalan Suara Tembang Sekar Alit telah dievaluasi dengan metode confusion matrix diperoleh bahwa pengujian algoritma MFCC pada 50 observasi data uji mendapatkan hasil akurasi sebesar 76,6%, precision sebesar 72.72%, dan recall sebesar 72.72%.

#### PERNYATAAN PENGHARGAAN

Merujuk dipublikasikannya penelitian ini, maka pada kesempatan ini kami peneliti mengucapkan terima kasih sebesar-besarnya kepada Lembaga Penelitian dan Pengabdian Masyarakat Universitas Udayana serta para anggota tim peneliti dari Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam serta Departemen Ilmu Faal, Fakultas Kedokteran Universitas Udayana.

#### DAFTAR PUSTAKA

- [1] I. W. Kotaniartha and A. Wijayanti, 'MAKNA PESAN MORAL LIRIK LAGU TRADISIONAL BALI (SEKAR ALIT) DALAM MEMBENTUK KARAKTER ANAK', in *Seminar Nasional INOBALI 2019 Inovasi Baru dalam Penelitian Sains, Teknologi dan Humaniora*, 2019, pp. 1178–1185.
- [2] I. D. G. B. D. Prabhawa, D. C. Khrisne, and M. Sudarma, 'Rancang Bangun Aplikasi Pengenalan Pupuh Bali Menggunakan Metode Mel Frequency Cepstral Coefficients', *Ranc. Bangun Apl. Pengenalan Pupuh Bali Menggunakan Metod. Mel Freq. Cepstral Coefficients*, vol. 3, no. 1, pp. 75–81, 2019, doi: 10.29303/jcosine.v3i1.237.
- [3] A. Setiawan and R. A. Widyanto, 'Evaluasi Website Perguruan Tinggi menggunakan Metode Usability Testing', *J. Inform. J. Pengemb. IT*, vol. 3, no. 3, pp. 295–299, Oct. 2018, doi: 10.30591/jpit.v3i3.912.
- [4] S. S. Zarish, 'Analyzing Usability of Educational Websites Using Automated Tools. 2019 International Conference on

- Computer and Information Sciences (ICCIS)', *2019 Int. Conf. Comput. Inf. Sci.*, pp. 1–4, 2019.
- [5] K. Al-omar, 'Evaluating the Usability and Learnability of the "Blackboard" LMS Using SUS and Data Mining', *2018 Second Int. Conf. Comput. Methodol. Commun.*, pp. 386–390, 2018.
- [6] J. Sauer, A. Sonderegger, K. Heyden, J. Biller, J. Klotz, and A. Uebelbacher, 'Extra-laboratorial usability tests : An empirical comparison of remote and classical field testing with lab testing', *Appl. Ergon.*, vol. 74, pp. 85–96, 2019.
- [7] M. A. Raharja, I. D. M. B. A. Darmawan, D. P. E. Nilakusumawati, and I. W. Supriana, 'Analysis of membership function in implementation of adaptive neuro fuzzy inference system (ANFIS) method for inflation prediction', *J. Phys. Conf. Ser.*, vol. 1722, no. 1, 2021, doi: 10.1088/1742-6596/1722/1/012005.
- [8] M. N. Rabbani, A. Rizal, and F. Y. Suratman, 'Implementasi Kunci Berbasis Suara Menggunakan Metode Mel Frequency Cepstral Coefficient (MFCC)', vol. 3, no. 3, pp. 3998–4007, 2016.
- [9] C. Asmita, T. Savitha, and K. Upadhya, 'Voice Recognition Using MFCC Algorithm', 2014. Accessed: Mar. 23, 2021. [Online]. Available: [www.ijirae.com](http://www.ijirae.com).
- [10] A. K. Santra and C. J. Christy, 'Genetic Algorithm and Confusion Matrix for Document Clustering', *Int. J. Comput. Sci. Issues*, vol. 9, no. 1, pp. 322–328, 2012.