

Identifikasi Genre Musik Menggunakan SVM dan Ekstraksi Ciri MFCC

Septian Yogi Yehezkiel^{*1}, Yohanes Suyanto²

¹Prodi Elektronika dan Instrumentasi, DIKE, FMIPA, UGM, Yogyakarta, Indonesia

²Departemen Ilmu Komputer dan Elektronika, FMIPA UGM, Yogyakarta, Indonesia

e-mail: ^{*1}septian.yogi.y@mail.ugm.ac.id, ²yanto@ugm.ac.id

Abstrak

Indonesia adalah negara yang kaya keberagaman karena wilayah yang luas dan terdiri dari berbagai suku. Setiap daerah memiliki budaya dan kesenian yang beragam, sehingga Indonesia memiliki banyak musik tradisional. Pada penelitian ini penulis menggunakan metode pengenalan pola Support Vector Machine (SVM) untuk mengidentifikasi genre musik tradisional Indonesia. Sistem identifikasi genre musik tradisional ini dapat menghasilkan akurasi sebesar 83% menggunakan ekstraksi ciri MFCC.

Kata kunci— identifikasi musik tradisional, Mel Frequency Cepstral Coefficient, Support Vector Machine.

Abstract

Indonesia is a very diverse country because it has a vast territory and is occupied by millions of people from various tribe. Therefore, traditional music in Indonesia is also diverse because each region has its own culture and art. In this study, the author used the Support Vector Machine(SVM) pattern recognition to identify the Indonesian traditional music genre. This genre identification system is able to produce an accuracy of 83% using MFCC.

Keywords—traditional music identification, Mel Frequency Cepstral Coefficient, Support Vector Machine.

1. PENDAHULUAN

Music Information Retrieval (MIR) merupakan salah satu Data Mining yang informasinya akan digali dari sumber data yang berupa musik. Banyak penelitian yang dilakukan pada bidang ini dengan berbagai latar belakang seperti untuk keperluan di bidang musik, psikologi, pemrosesan sinyal, machine learning, maupun kombinasi dari beberapa latar belakang tersebut. Sama seperti bidang data mining yang lainnya, pada MIR digunakan berbagai metode untuk melakukan klasifikasi musik seperti Support Vector Model (SVM), k-Nearest Neighbour (kNN), Gaussian Classifier, Hidden Markov Model (HMM), Neural Network, dan Fuzzy Model[1].

Perkembangan dan penyebaran musik secara digital saat ini sangat mudah, hal ini didukung oleh kemajuan teknologi yang pesat. Dengan bertambahnya jumlah musik dalam

bentuk digital, terjadi sebuah masalah yaitu bagaimana cara mengorganisasi musik-musik tersebut. Data musik dapat diorganisasi berdasarkan penyanyi, album, tahun rilis, atau berdasarkan genre[2].

Algoritme *Supervised Learning* adalah algoritme yang mampu mengidentifikasi masukan data baru berdasarkan data latih yang sudah tersimpan sebelumnya. kNN adalah salah satu dari *Supervised Learning* seperti yang dilakukan menggunakan Spotify API untuk ekstraksi ciri dan metode kNN untuk klasifikasi. Penelitian ini dilakukan dengan mengambil 11 fitur musik yaitu *speechiness, energy, danceability, loudness, tempo, mode, valence, instrumentalness, acousticness, dan liveliness*, kemudian menggunakan kNN untuk metode klasifikasi menjadi 10 genre dengan akurasi rata-rata 44,8% [3].

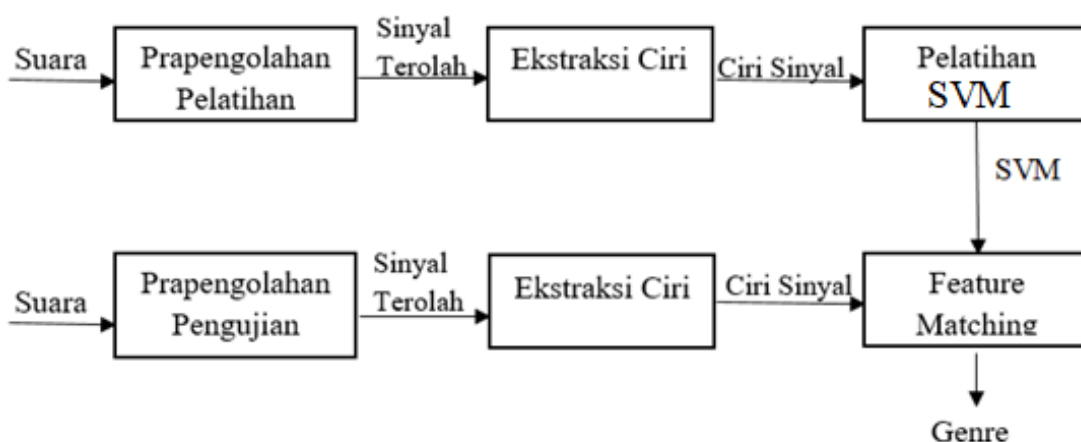
Penelitian yang dilakukan menggunakan MFCC untuk ekstraksi ciri dan menggunakan KNN untuk mengklasifikasi akor kibor berhasil diperoleh akurasi sebesar 99,03% namun masih belum bisa mengenali akor yang memiliki pembeda setengah nada dalam susunan nadanya secara akurat[4].

Penelitian yang dilakukan untuk membandingkan metode klasifikasi k-NN, *Naïve Bayes, Decision Tree, SVM, dan Random Forest* untuk melakukan klasifikasi 10 genre lagu. Dari penelitian yang dilakukan, diperoleh bahwa metode SVM dengan ekstraksi ciri MFCC memiliki akurasi terbesar yakni 69,90% [5].

2. METODE PENELITIAN

2.1 Analisis Sistem

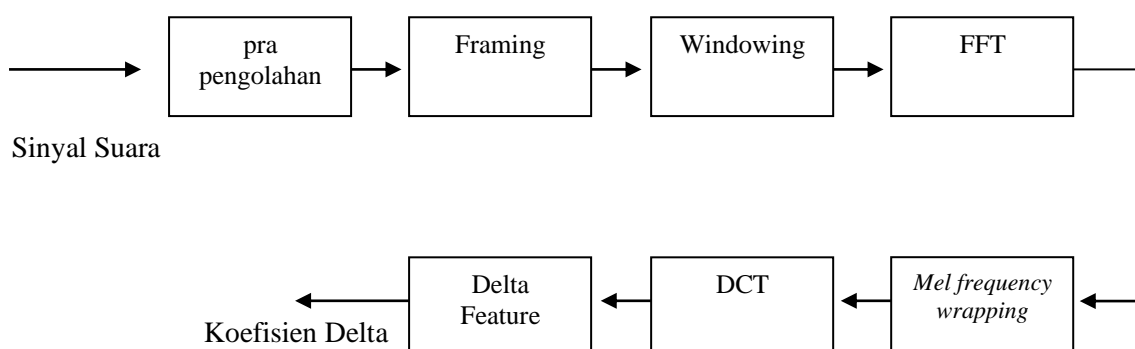
Pada penelitian ini sistem dirancang untuk dapat melakukan identifikasi genre musik tradisional Indonesia yakni cilokaq, gambang kromong, dan keroncong. Data musik diunduh dari platform youtube kemudian dilakukan ekstraksi ciri MFCC sehingga dihasilkan 15 koefisien MFCC. Klasifikasi genre musik dilakukan menggunakan metode Support Vector Machine (SVM) dengan fitur berupa nilai rata-rata dan standar deviasi dari 15 koefisien MFCC yang diperoleh. Hasil klasifikasi menggunakan SVM dengan ekstraksi ciri MFCC akan dibandingkan dengan klasifikasi menggunakan SVM dengan ekstraksi ciri ZCR. Bentuk diagram blok proses sistem klasifikasi ditunjukkan pada Gambar 1.



Gambar 1 Diagram blok proses sistem

2. 2 Rancangan Ekstraksi Ciri

MFCC adalah salah satu metode ekstraksi ciri yang didasarkan pada perilaku pendengaran manusia yang tidak dapat mengenali frekuensi lebih dari 1Khz. MFCC didasarkan pada perbedaan frekuensi yang dapat didengar oleh telinga manusia. Sinyal dinyatakan dalam skala MEL, skala ini didasarkan pada persepsi *itches* dalam jarak interval yang sama yang dinilai oleh pengamat. Skala ini menggunakan filter yang ditempatkan secara linier pada frekuensi di bawah 1000 Hz dan jarak *logarithmic* di atas 1000Hz [6]. Bentuk diagram alir proses ekstraksi ciri MFCC ditunjukkan pada Gambar 2.

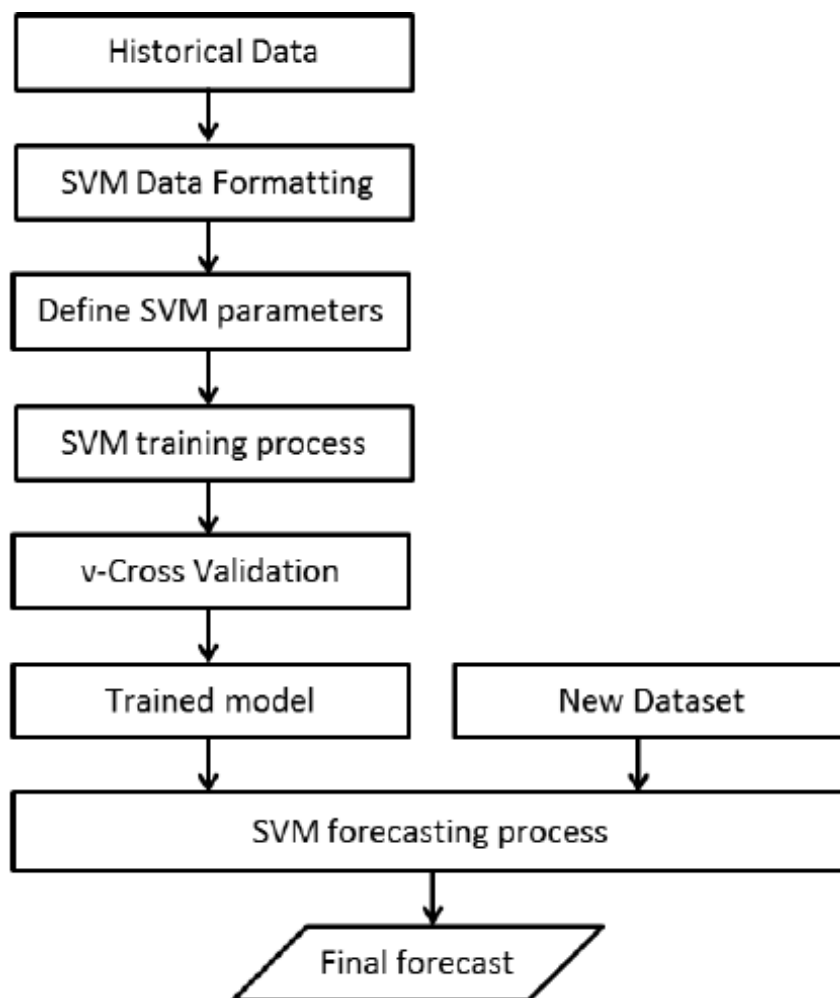


Gambar 2 Proses ekstraksi ciri MFCC

Pada Gambar 2 dapat dilihat urutan proses MFCC. Pada tahap awal sinyal suara yang berupa dataset dilakukan *pre-processing* yang bertujuan untuk memperbaiki kualitas sinyal suara. Setelah dilakukan *pre-processing*, kemudian sinyal dilakukan proses *framing* yang akan menganalisis sinyal suara ke dalam bentuk *frame*[7]. Kemudian sinyal suara akan dilakukan proses *windowing*, yang bertujuan untuk mengurangi efek *discontinue* pada ujung-ujung *frame* yang dihasilkan oleh proses *framing*. Setelah proses *windowing*, kemudian sinyal dilakukan proses *fast fourier transform* (FFT) yang bertujuan mengubah sinyal digital pada ranah waktu ke ranah frekuensi. Sinyal suara akan melalui tahap *mel frequency wrapping* yang bertujuan untuk mengetahui ukuran energi dari *frequency band* tertentu dalam sinyal suara. Sinyal kemudian melalui proses *discrete cosine transform* (DCT) yang mengubah bentuk perkalian hasil FFT terhadap *mel frequency wrapping* ke dalam bentuk penjumlahan. Kemudian dilakukan proses *delta-feature* yang digunakan untuk mendapatkan informasi dari ciri yang dinamis[8].

2. 3 Rancangan Support Vector Machine

Support Vector Machine adalah algoritme yang berasal dari teori pembelajaran statistik. Ide utama dari SVM adalah untuk metransformasi input asli ke dimensi fitur yang lebih tinggi dengan menggunakan fungsi kernel dan untuk mencapai tingkat klasifikasi optimum di *feature space* baru yang terdapat pemisah yang jelas antar fitur yang didapatkan dari penempatan optimal *hyperplane* pemisah[9]. Algoritme SVM dapat digunakan untuk klasifikasi dan regresi. Pada klasifikasi, algoritme ini memerlukan keluaran berupa optimal *hyperplane* yang mengkategorikan sampel set set baru [10]. Diagram alir proses SVM ditunjukkan pada Gambar 3.

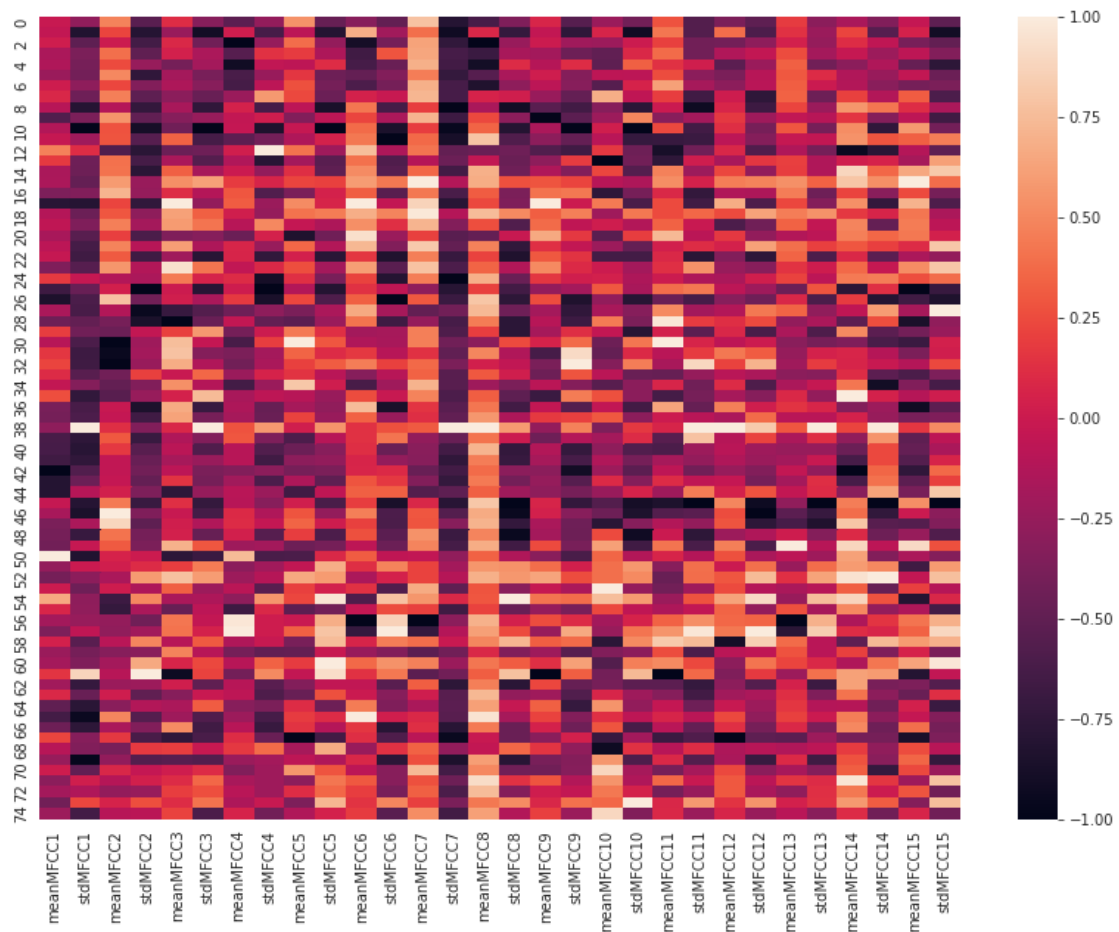


Gambar 3 Proses SVM

3. HASIL DAN PEMBAHASAN

3.1 Hasil Ekstraksi Ciri MFCC

Sinyal suara dibagi menjadi frame-frame sehingga setiap frame memiliki ukuran 512 sample, yang bertujuan untuk mendapatkan karakteristik ciri yang stabil. Setiap data penelitian akan menghasilkan 2.584 frame, kemudian dari setiap frame sampel data diambil 15 koefisien MFCC. Tiap genre musik dikelompokkan berdasarkan genrenya yakni musik 0 sampai 24 memiliki genre cilokaq, musik 25 sampai 49 memiliki genre gambang kromong, dan musik 50 sampai 74 memiliki genre keroncong. Gambar 4 menunjukkan hasil ekstraksi ciri MFCC dari 75 data musik.



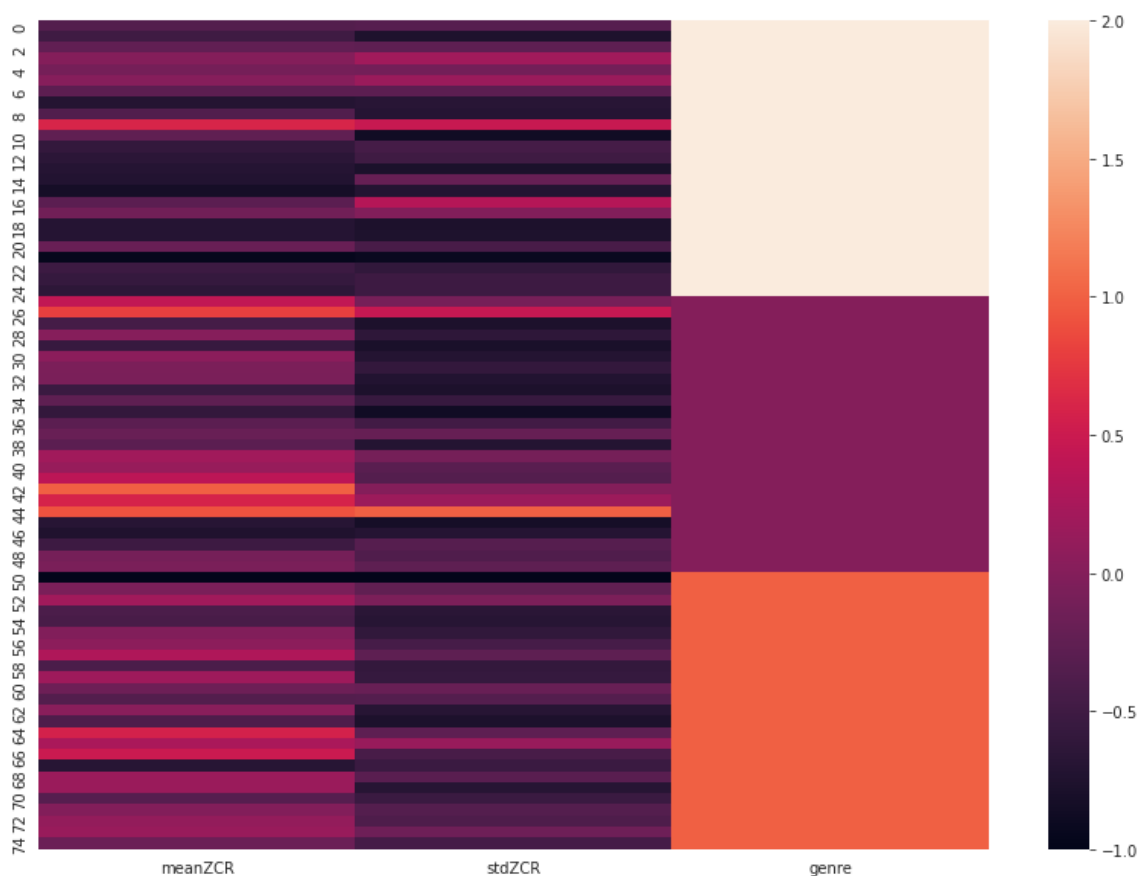
Gambar 4 Hasil ekstraksi ciri MFCC

Gambar 4 menunjukkan bahwa fitur berupa nilai rata-rata dari setiap koefisien MFCC memiliki kecerahan yang cukup beragam, hal ini menggambarkan nilai tiap koefisien beragam tergantung frekuensi tiap musik. Gambar diatas menunjukkan bahwa fitur berupa standar deviasi dari setiap koefisien MFCC didominasi berwarna gelap, hal ini dikarenakan penyimpangan yang terjadi pada tiap koefisien MFCC relatif bernilai kecil. Dari gambar tersebut juga terlihat bahwa koefisien yang memiliki penyimpangan nilai terkecil berada pada koefisien 7, hal ini terlihat pada fitur stdMFCC7 yang didominasi warna gelap karena memiliki nilai yang rendah.

3.2 Hasil Ekstraksi Ciri ZCR

Hasil dari proses ini berupa nilai rata-rata dan standar deviasi frekuensi dasar dari setiap sinyal suara. Nilai frekuensi dasar tersebut diperoleh dari perhitungan ZCR dari sinyal suara dibagi dengan durasi waktu sinyal suara. Tiap genre musik dikelompokkan berdasarkan genrenya yakni musik 0 sampai 24 memiliki genre cilokaq, musik 25 sampai 49 memiliki genre

gambang kromong, dan musik 50 sampai 74 memiliki genre keroncong. Hasil ekstraksi ciri ZCR beberapa musik ditunjukkan pada Gambar 5.



Gambar 5 Hasil Ekstraksi ciri ZCR

3.3 Hasil Pengujian SVM dengan ekstraksi ciri ZCR

Pengujian dilakukan untuk mengetahui apakah sistem identifikasi genre musik tradisional dapat mengidentifikasi dengan baik data uji yang disediakan. Dari pengujian diperoleh bahwa sistem memperoleh rata-rata akurasi sebesar 43%. Akurasi terbesar dari keseluruhan kelas diperoleh pada genre keroncong yakni 62%, sedangkan akurasi terkecil diperoleh genre cilokaq yakni 0%. Akurasi yang rendah pada pengujian SVM dengan ekstraksi ciri ZCR dapat disebabkan karena jumlah fitur yang digunakan terlalu sedikit untuk melakukan klasifikasi pada data yang memiliki kelas lebih dari dua serta komposisi instrumen musik yang mirip.

3.4 Hasil Pengujian SVM dengan ekstraksi ciri MFCC

Dari percobaan pada 23 data uji didapat rata-rata benar dari keseluruhan adalah 83%. Akurasi terbesar dari keseluruhan kelas diperoleh genre cilokaq yakni 88%. Akurasi terkecil diperoleh genre gambang kromong yakni 75%, hal ini dapat disebabkan karena adanya kemiripan dari instrumen yang digunakan pada genre cilokaq dan keroncong. Selain kemiripan instrumen pada genre gambang kromong dan keroncong, jumlah fitur yang terlalu banyak untuk

melakukan klasifikasi juga dapat mempengaruhi hasil klasifikasi. Performa pengujian menggunakan data uji model diperoleh dengan nilai rata-rata presisi 83%, *recall* 85%, *f1-score* 83%, dan akurasi uji 83%. Nilai performa model tersebut ditunjukkan pada Tabel 1.

Tabel 1 Nilai performa model SVM

Genre	Presisi	Recall	F1-score	Akurasi
Cilokaq	88%	88%	88%	83%
Gambang Kromong	75%	100%	86%	
Keroncong	86%	67%	75%	
Rata-rata	83%	85%	83%	

4. KESIMPULAN

Kesimpulan yang didapat dari penelitian ini adalah sistem klasifikasi genre musik tradisional Indonesia berdasarkan pengenalan pola suara menggunakan metode Support Vector Machine(SVM) telah berhasil dibuat dengan hasil analisis sistem memiliki rata-rata akurasi sebesar 83%, nilai presisi sebesar 83%, nilai recall sebesar 85%, dan f1-score sebesar 83%.

5. SARAN

Dari hasil pengamatan yang dilakukan terhadap sistem yang telah dibuat, didapatkan beberapa saran yang dapat dipertimbangkan untuk penelitian yang berkaitan dengan topik ini. Beberapa saran tersebut adalah antara lain perlu adanya penambahan jumlah dataset untuk dijadikan sebagai data latih dan data uji serta perlu adanya penambahan kelas untuk genre-genre musik tradisional Indonesia yang lain.

DAFTAR PUSTAKA

- [1] Rinaldi, A., Hendra dan Alamsyah, D., "Pengenalan Gender Melalui Suara dengan Algoritme Support Vector Machine", vol.2, no.1. pp.47-54,2016.
- [2] Gupta, M., Bharti, S.S. dan Agarwa, S., "Support Vector Machine Based Jenis Kelamin Identification Using Voiced Speech Frames", *Fourth International Conference on Parallel*, vol.3, no.2, pp. 16-20, 2016.
- [3] Setiawan, A. Hidayatno, and R. R. Isnanto, "Aplikasi Pengenalan Ucapan dengan Ekstraksi ciri MFCC Melalui Jaringan JST untuk Mengoperasikan Kursor Komputer", *Transmisi*, vol. 13, no 3, pp.82-86, Jun.2012[Online]. Available : <https://doi.org/10.12777/transmisi.13.3.82-86>.
- [4] Hidayat, Syahrani, "Speech Recognition of KV-Patterned Indonesian Syllable Using MFCC and Hmm", *Kursor*. 8.67.10.28961/kursor.v8i2.63.

-
- [5] Hasan, A.I., “Pembangkitan Warna Suara Saron Sintesis Berdasarkan Petikan Senar Gitar”. Available: <https://jurnal.ugm.ac.id/ijeis/article/view/15347>. [Accessed: 30-May-2021]
- [6] Rangga. P.W., “Klasifikasi Tingkat Kemurnian Bahan Bakar menggunakan algoritma K-Nearest Neighbor,” *IJEIS (Indonesian J. Electron. Instrum. Syst.*, Vol.9, No.2, pp. 161~172.2019[Online]. Available: <https://jurnal.ugm.ac.id/ijeis/article/view/49660/26017> [Accessed: 01-Dec-2021]
- [7] S.Faziludeen and P. Sankaran, “ECG Beat Classification Using Evidential K -Nearest Neighbours,” *Procedia Comput. Sci.*, vol. 89, pp. 499–505, 2016 [Online]. Available: <http://dx.doi.org/10.1016/j.procs.2016.06.106>
- [8] Y. F. Safri, R. Arifudin, and M. A. Muslim, “K-Nearest Neighbor and Naive Bayes Classifier Algorithm in Determining The Classification of Healthy Card Indonesia Giving to The Poor,” *Sci. J. Informatics*, vol. 5, no. 1, p. 18, 2018.
- [9] Martinez, J., Perez, dan Suzuki, M.M, “Speaker Recognition using MFCC and Vector Quantization techniques,” *International Conference on Electronics Communications and Computing*, vol. 6, no. 1, pp . 248-251 . Available : <https://dx.doi.org/10.1109/CONIELECOMP.2012.61890>. [Accessed: 22-June- 2021].
- [10] A. Mustofa, “Sistem Pengenalan Penutur dengan Metode-Mel-frequency Wrapping,” *J. Tek. Elektro*, Vol.7, no.2, pp.88-96, 2007[Accessed: 25-Nov-2021]