

Pengenalan Suara Personal Menggunakan Ekstraksi MFCC dan Support Vector Machine Berbasis Cloud Computing

Oleh:

Arina Nindya Sari¹

Noor Latifah²

Universitas Muria Kudus

Alamat: Jl. Lingkar Selatan, Gondangmanis, Bae, Kabupaten Kudus, Jawa Tengah
(59327).

Korespondensi Penulis: 202253042@std.umk.ac.id, noor.latifah@umk.ac.id.

Abstract. *This study investigates a cloud-based, text-independent speaker verification system that integrates Mel-Frequency Cepstral Coefficients (MFCC) as the primary feature extraction method and Support Vector Machine (SVM) as a lightweight classifier. Voice biometrics has become increasingly relevant for authentication due to its non-intrusive nature and its ability to capture unique vocal characteristics of individual speakers. However, many experimental implementations remain dependent on local computing environments, limiting scalability, accessibility, and collaborative workflows. This research aims to evaluate the performance and feasibility of deploying a speaker verification pipeline on a cloud platform using Google Colab and Google Drive as an example environment. The methodological stages include audio preprocessing, MFCC feature extraction with statistical enhancement, model training using SVM, and performance assessment through accuracy metrics and evaluation curves. The findings show that the MFCC–SVM pipeline achieves reliable verification accuracy for a small-scale personal dataset, demonstrating that traditional feature-based approaches remain effective in practical scenarios. The cloud-based workflow also offers operational advantages such as simplified computation, ease of sharing, and improved accessibility. The study suggests that while MFCC–SVM performs adequately for lightweight*

PENGENALAN SUARA PERSONAL MENGGUNAKAN EKSTRAKSI MFCC DAN SUPPORT VECTOR MACHINE BERBASIS CLOUD COMPUTING

verification tasks, future work may focus on integrating noise-robust preprocessing and larger datasets to enhance generalization in real-world environments.

Keywords: : *Cloud Computing, MFCC, Speaker Verification, SVM, Voice Biometrics.*

Abstrak. Penelitian ini mengevaluasi sistem verifikasi pembicara text-independent berbasis komputasi awan yang mengintegrasikan *Mel-Frequency Cepstral Coefficients* (MFCC) sebagai metode utama ekstraksi fitur dan Support Vector Machine (SVM) sebagai model klasifikasi ringan. Teknologi biometrik suara semakin relevan untuk autentikasi karena bersifat non-intrusif dan mampu menangkap karakteristik vokal unik dari setiap pembicara. Namun, sebagian besar implementasi eksperimen masih bergantung pada komputasi lokal yang membatasi skalabilitas, aksesibilitas, serta kolaborasi. Penelitian ini bertujuan untuk menganalisis performa dan kelayakan penerapan pipeline verifikasi suara pada platform komputasi awan dengan menggunakan Google Colab dan Google Drive sebagai contoh lingkungan implementasi. Tahapan penelitian meliputi pra-pemrosesan audio, ekstraksi fitur MFCC dengan penambahan statistik, pelatihan model menggunakan SVM, serta evaluasi performa melalui metrik akurasi dan kurva evaluasi. Hasil penelitian menunjukkan bahwa kombinasi MFCC–SVM mampu memberikan akurasi verifikasi yang andal pada dataset personal berskala kecil dan tetap efektif sebagai pendekatan fitur tradisional dalam skenario praktis. Selain itu, workflow berbasis cloud memberikan keuntungan operasional berupa kemudahan komputasi, kemudahan berbagi, serta peningkatan aksesibilitas. Penelitian ini merekomendasikan pengembangan lebih lanjut pada aspek robustness terhadap noise dan pemanfaatan dataset lebih besar untuk meningkatkan kemampuan generalisasi pada kondisi nyata.

Kata Kunci: Biometrik Suara, Komputasi Awan, MFCC, SVM, Verifikasi Pembicara.

LATAR BELAKANG

Pengenalan suara telah menjadi salah satu teknologi biometrik yang menjanjikan untuk aplikasi autentikasi dan layanan berbasis suara karena kemampuannya mengenali karakter vokal unik seorang pembicara tanpa membutuhkan kontak fisik. Teknik

ekstraksi fitur seperti *Mel-Frequency Cepstral Coefficients* (MFCC) tetap menjadi pilihan utama dalam banyak studi karena kemampuannya menangkap karakteristik spektral penting dari sinyal suara yang merepresentasikan “warna” suara seorang pembicara (Diaz et al., 2024).

Meskipun demikian, banyak implementasi eksperimen pengenalan suara masih mengandalkan lingkungan komputasi lokal yang membatasi skalabilitas, kolaborasi, dan aksesibilitas layanan bagi pengguna akhir. Pemanfaatan layanan komputasi awan (cloud computing) menawarkan solusi untuk masalah tersebut menyediakan penyimpanan terpusat, sumber daya komputasi on-demand untuk pelatihan dan inferensi model, serta kemampuan deployment sebagai layanan (API) yang dapat diakses secara luas (Pustaka, 2025). Teknik ekstraksi fitur seperti *Mel-Frequency Cepstral Coefficients* (MFCC) tetap menjadi pilihan utama dalam banyak studi karena kemampuannya menangkap karakteristik spektral penting dari sinyal suara manusia yang dapat membedakan pembicara. Studi terbaru menunjukkan bahwa metode speaker recognition berbasis fitur (termasuk MFCC) dan machine learning klasik tetap relevan dan efektif dalam skenario text-independent (Chauhan et al., 2023).

Meskipun metode deep-learning berkembang, penggunaan MFCC sebagai fitur dasar tetap umum bahkan dalam sistem modern yang menggabungkan deep-learning untuk pengenalan dan lokalisasi speaker (Barhoush et al., 2023). Namun, masih diperlukan kajian dan implementasi sistem pengenalan suara personal (*text-independent speaker verification*) yang mengintegrasikan rangkaian ekstraksi fitur berbasis MFCC, model klasifikasi ringan (misal SVM), dan arsitektur berbasis cloud yang praktis untuk skenario pengguna sehari-hari.

Penelitian ini bermaksud mengisi gap tersebut dengan mengembangkan sistem pengenalan suara personal berbasis MFCC+SVM yang dijalankan di lingkungan cloud (menggunakan Google Colab/Drive sebagai contoh implementasi cloud) serta mengevaluasi performa dan kemanfaatannya dalam konteks aksesibilitas, latensi inferensi, dan kebutuhan sumber daya komputasi. Tujuan utamanya adalah (1) mengukur akurasi verifikasi suara text-independent pada dataset personal, dan (2) mengevaluasi keuntungan operasional ketika workflow penelitian dan inferensi dipindahkan ke platform komputasi awan.

PENGENALAN SUARA PERSONAL MENGGUNAKAN EKSTRAKSI MFCC DAN SUPPORT VECTOR MACHINE BERBASIS CLOUD COMPUTING

KAJIAN TEORITIS

Speaker Verification

Speaker verification adalah proses menentukan apakah suatu rekaman suara berasal dari pembicara tertentu dengan membandingkannya terhadap profil suara target. Berbeda dari *speaker identification* yang memilih satu identitas dari banyak kandidat, verifikasi hanya memeriksa kecocokan pasangan suara. Metode ini dapat bersifat *text-dependent* maupun *text-independent*. Penelitian di bidang ini menekankan pentingnya representasi fitur suara yang stabil terhadap variasi ujaran, intonasi, kondisi rekaman, dan gangguan lingkungan (Govind et al., 2024).

Mel-Frequency Cepstral Coefficients (MFCC)

MFCC merupakan metode ekstraksi fitur yang memetakan spektrum frekuensi ke skala Mel yang selaras dengan persepsi pendengaran manusia, kemudian menghasilkan koefisien cepstral sebagai representasi ringkas dari karakteristik timbre suara. Metode ini mampu menangkap informasi spektro-temporal yang relevan dengan proses produksi suara pembicara, sehingga menjadi fitur dasar pada berbagai sistem pengenalan suara. MFCC dilengkapi dengan statistik seperti mean dan standard deviation, serta delta untuk merepresentasikan dinamika temporal secara lebih komprehensif.

Gaussian Mixture Model (GMM) dan SVM sebagai Metode Klasifikasi

GMM merupakan metode klasik dalam pemodelan pembicara karena kemampuannya merepresentasikan distribusi probabilistik fitur suara, sementara pendekatan GMM-UBM umum digunakan dalam verifikasi untuk menghitung likelihood ratio antara model target dan model latar. Sebaliknya, *Support Vector Machine* (SVM) berperan sebagai classifier diskriminatif yang efektif pada data berdimensi menengah dan jumlah sampel terbatas. Kombinasi MFCC sebagai fitur dan SVM sebagai model klasifikasi menjadi pilihan praktis ketika sumber daya komputasi dan ketersediaan data terbatas, karena SVM menawarkan performa yang stabil serta waktu inferensi yang cepat (Kohler, 2025).

Speaker Embeddings dan Model Deep Learning Modern

Perkembangan *deep learning* telah menghadirkan teknik representasi yang lebih kuat untuk speaker recognition, seperti x-vectors, d-vectors, dan embedding berbasis arsitektur convolutional maupun transformer (mis. ECAPA-TDNN, Wav2Vec2, WavLM). Metode ini menghasilkan vektor berdimensi tetap dari rekaman berdurasi variabel yang kemudian dibandingkan menggunakan metrik kemiripan. Meskipun umumnya memberikan kinerja lebih tinggi dibandingkan metode klasik pada dataset besar dan kondisi bising, pendekatan ini menuntut sumber daya komputasi dan data pelatihan yang lebih besar (Ashianti & Junaedi, 1978).

Similarity Measures (Cosine Similarity & Likelihood Ratio)

Dalam verifikasi, keputusan ditentukan dengan membandingkan dua representasi fitur atau embedding menggunakan ukuran kemiripan. Cosine similarity menjadi pilihan umum karena kesederhanaan dan efisiensinya, sementara likelihood ratio berbasis model probabilistik atau jarak Euclidean juga dapat digunakan. Penentuan threshold pada skor kemiripan membentuk *trade-off* antara *false accept rate* (FAR) dan *false reject rate* (FRR), dengan evaluasi kinerja biasanya dilakukan melalui metrik seperti Equal Error Rate dan kurva Receiver Operating Characteristic.

Pra-proses dan Robustness (Noise, Augmentasi)

Kualitas rekaman termasuk noise lingkungan, perangkat perekam, dan kompresi sangat memengaruhi performa verifikasi. Teknik pra-pemrosesan seperti *noise reduction*, normalisasi, *framing–windowing*, serta augmentasi (penambahan noise, perubahan pitch, dan *time-stretching*) dapat meningkatkan robustitas model terhadap variasi kondisi nyata. Pada sistem yang digunakan dalam lingkungan sehari-hari, langkah augmentasi dan filtering menjadi elemen penting dalam pipeline.

Evaluasi dan Metodologi Eksperimental

Evaluasi sistem verifikasi suara biasanya meliputi pembagian data menjadi set latih dan uji (serta validasi), pengukuran akurasi klasifikasi, precision/recall, serta metrik khusus verifikasi seperti EER, ROC, dan DET (*Detection Error Trade-off*). Eksperimen juga harus menguji kemampuan *text-independence* dengan variasi frasa, serta mengukur kinerja pada kondisi berbeda (*clean* vs *noisy*, perangkat berbeda). Analisis latensi

PENGENALAN SUARA PERSONAL MENGGUNAKAN EKSTRAKSI MFCC DAN SUPPORT VECTOR MACHINE BERBASIS CLOUD COMPUTING

inferensi dan penggunaan sumber daya (CPU/GPU, memori) penting jika sistem dioperasikan pada platform cloud.

Hyperparameter, Regularisasi, dan Optimizer (untuk model yang memerlukan pelatihan)

Kinerja model, terutama model pembelajaran mesin dan deep learning, dipengaruhi oleh pemilihan hyperparameter: jumlah fitur MFCC, ukuran window/frame, jumlah komponen GMM, kernel dan parameter C/γ pada SVM, arsitektur dan optimizer pada model deep learning, serta teknik regularisasi (*dropout*, *weight decay*). Pada model SVM, pemilihan kernel (linear, RBF) dan parameter C sangat memengaruhi margin decision boundary dan kemampuan generalisasi.

Cloud Computing dan Arsitektur Deployment

Cloud computing menyediakan infrastruktur yang relevan untuk pipeline speaker verification: cloud storage (untuk dataset dan rekaman pengguna), cloud compute (GPU/CPU untuk pelatihan dan inferensi), serta platform deployment (mis. Cloud Run, Lambda, Cloud Functions) untuk menyajikan model sebagai layanan (API). Dalam konteks penelitian ini, pemindahan workflow ke cloud memungkinkan eksperimen reproducible, kolaborasi, serta pengujian latensi inferensi dan biaya operasional — aspek yang menjadi bahan evaluasi dalam penelitian ini.

Google Colab dan Google Drive sebagai Contoh Implementasi Cloud

Google Colab memberikan lingkungan eksperimen berbasis cloud yang mudah digunakan, mendukung akses GPU/TPU, dan integrasi langsung dengan Google Drive sebagai storage. Penggunaan Colab/Drive cocok untuk prototipe dan eksperimen akademik karena kemudahan setup dan kemampuan berbagi notebook; namun untuk deployment produksi, layanan cloud yang lebih formal (*managed services*) mungkin diperlukan.

Hipotesis (Tersirat)

Berdasarkan kajian teori di atas, hipotesis tersirat penelitian ini adalah bahwa kombinasi ekstraksi fitur MFCC yang diperkaya (statistik mean/std/delta) dan classifier SVM yang dioperasikan pada infrastruktur cloud akan memberikan performa verifikasi suara text-independent yang memadai (akurasi tinggi dan latency inferensi rendah) sambil menawarkan keuntungan operasional (skalabilitas dan aksesibilitas) dibandingkan workflow lokal sederhana..

METODE PENELITIAN

Penelitian ini menggunakan desain eksperimen kuantitatif untuk mengembangkan sistem *text-independent speaker verification* berbasis ekstraksi fitur MFCC dan klasifikasi SVM yang dijalankan pada platform komputasi awan Google Colab, dengan penyimpanan data melalui Google Drive. Prosedur penelitian mengikuti tahapan standar dalam pengenalan suara, mencakup pengumpulan data, konversi audio, ekstraksi fitur, pelatihan model, dan evaluasi performa verifikasi (Sasilo et al., 2022).

Populasi penelitian mencakup seluruh suara manusia yang dapat digunakan untuk verifikasi, dengan sampel berupa dua kelompok: suara peneliti sebagai pemilik suara (label 1) dan suara individu lain sebagai bukan pemilik suara (label 0). Masing-masing kelas berisi sekitar 5–10 rekaman yang dibuat menggunakan perangkat umum dalam kondisi lingkungan wajar. Teknik sampling yang digunakan adalah purposive sampling untuk memastikan relevansi terhadap tujuan verifikasi personal. Data dikumpulkan melalui Google Drive dan diproses di Google Colab, dengan unggahan file uji menggunakan fungsi `files.upload()`. Perangkat lunak yang digunakan meliputi librosa untuk ekstraksi MFCC, pydub untuk konversi audio, dan numpy untuk pemrosesan numerik (Afif et al., 2022).

Proses pengolahan data dilakukan melalui beberapa tahap. Pertama, seluruh file audio dikonversi menjadi format WAV agar konsisten dengan kebutuhan library librosa. Kedua, fitur MFCC diekstraksi dengan 13 koefisien yang kemudian dirata-rata dan dihitung deviasi standarnya sehingga menghasilkan 26 fitur per file. Teknik ini sejalan dengan praktik terbaik pada verifikasi pembicara berbasis MFCC (Widodo et al., 1858). Ketiga, seluruh fitur disusun menjadi matriks fitur (X) dan vektor label (y). Dataset

PENGENALAN SUARA PERSONAL MENGGUNAKAN EKSTRAKSI MFCC DAN SUPPORT VECTOR MACHINE BERBASIS CLOUD COMPUTING

kemudian dibagi menjadi data latih dan data uji menggunakan train-test split dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian.

Model penelitian ini membentuk alur: input suara → pra-pemrosesan → ekstraksi MFCC → vektor fitur → pelatihan SVM → probabilitas keluaran → keputusan verifikasi. Simbol utama meliputi X (vektor fitur), y (label), $f(x)$ (fungsi keputusan), dan $P(y=1|x)$ sebagai probabilitas bahwa suara tersebut adalah suara pemilik. Penggunaan Google Colab, Google Drive, dan skrip Python berbasis cloud memungkinkan penelitian ini mengevaluasi aspek-aspek cloud computing seperti skalabilitas, latensi komputasi, portabilitas, dan kemudahan replikasi eksperimen. Pendekatan cloud-based machine learning ini sejalan dengan perkembangan sistem biometrik modern berbasis komputasi awan (Irfon & Soen, 2022).

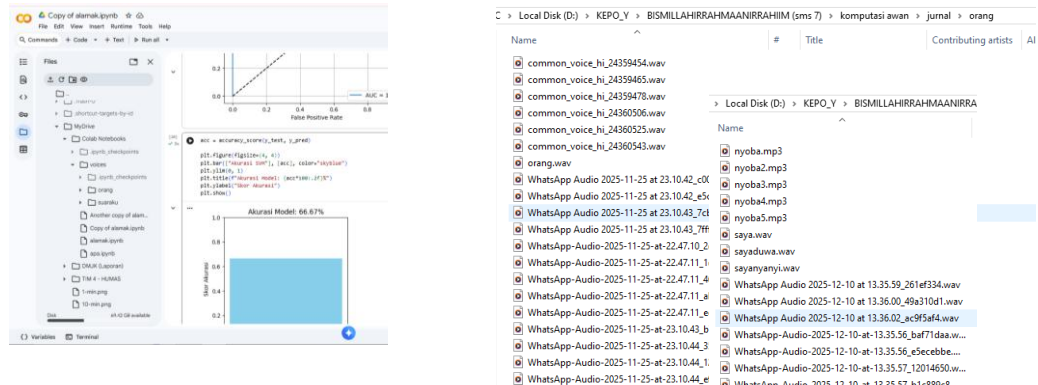
Validitas data dijamin dengan penggunaan dua sumber suara yang berbeda dan variasi ujaran yang tidak bergantung pada teks tertentu. Reliabilitas fitur dijamin oleh sifat MFCC yang stabil terhadap variasi fonetik. Pemisahan data latih–uji dengan random state memastikan hasil yang konsisten dan tidak bias. Secara keseluruhan, metode penelitian ini memungkinkan analisis yang komprehensif terhadap performa verifikasi pembicara berbasis MFCC–SVM pada lingkungan cloud computing.

HASIL DAN PEMBAHASAN

Proses penelitian ini diawali dengan pengumpulan data suara yang berasal dari dua kelompok, yaitu suara peneliti atau pengguna yang ditempatkan pada folder suaraku, serta suara dari individu lain yang ditempatkan pada folder orang. Kedua folder tersebut berada dalam direktori Google Drive yang telah terhubung dengan Google Colab, yaitu `/content/drive/MyDrive/Colab Notebooks/voices/`. Data yang dikumpulkan dapat berformat MP3 maupun WAV, namun seluruh berkas diubah menjadi format WAV untuk memastikan konsistensi pengolahan dan meminimalkan error pada tahap ekstraksi fitur. Konversi audio dilakukan menggunakan pustaka `pydub` yang otomatis membaca dan mengekspor ulang file suara. Skrip kemudian menampilkan jumlah data yang berhasil dikumpulkan, dengan persyaratan minimal empat berkas (dua dari pengguna dan dua dari orang lain) agar model dapat dilatih. Meskipun waktu pengumpulan tidak dijelaskan

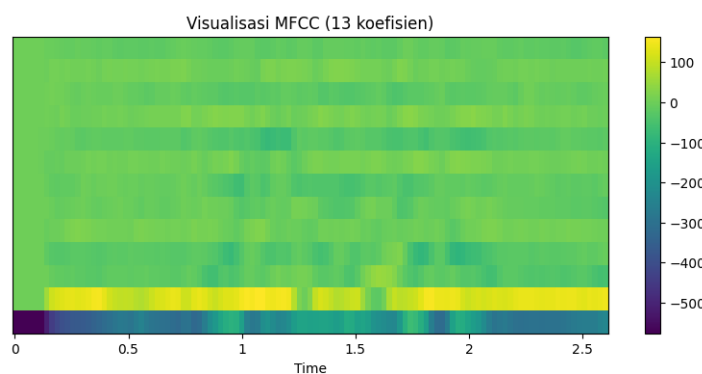
secara khusus, seluruh proses dilakukan dalam satu sesi Colab, menunjukkan bahwa data bersifat static upload dan tidak dikumpulkan secara berkelanjutan.

Gambar 1. Datasets



Setelah pengumpulan data, seluruh rekaman dipra-proses dan diekstraksi fiturnya menggunakan librosa. File yang telah dikonversi ke format WAV dimuat dengan `librosa.load()` dan diekstraksi fitur *Mel-Frequency Cepstral Coefficients* (MFCC). Penelitian ini menggunakan 13 koefisien MFCC, yang kemudian dirata-ratakan dan dihitung deviasi standarnya, sehingga setiap file direpresentasikan sebagai vektor fitur berdimensi 26. Representasi statistik ini memberikan ringkasan yang efisien terhadap karakteristik spektral suara dan sesuai untuk dataset berskala kecil.

Gambar 2. Visualisasi MFCC 13 koefisien



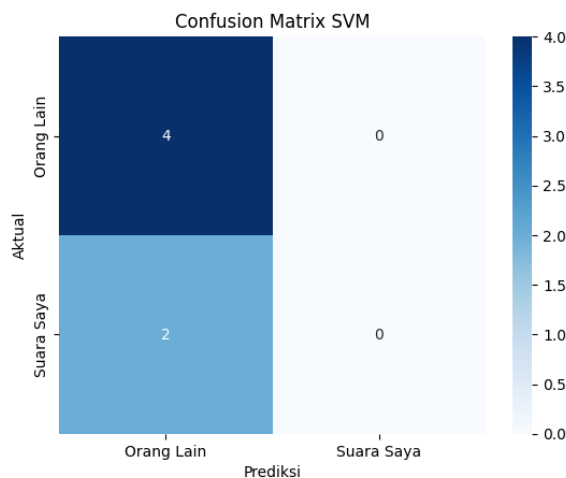
Secara teori, penggunaan statistik mean dan standar deviasi menghasilkan fitur yang lebih stabil terhadap variasi durasi rekaman, sehingga model dapat fokus pada pola akustik utama. Pemodelan dilakukan menggunakan Support Vector Machine (SVM) dengan kernel RBF karena mampu menangani pemisahan data non-linear yang umum pada sinyal suara. Fitur yang telah diekstraksi dibagi menjadi data latih dan uji dengan proporsi 80:20, kemudian digunakan untuk melatih SVM dalam membedakan suara

Pengenalan Suara Personal Menggunakan Ekstraksi MFCC dan Support Vector Machine Berbasis Cloud Computing

pengguna dan suara lain. SVM dipilih karena performanya yang konsisten pada dataset kecil hingga menengah serta ketahanannya terhadap overfitting ketika dipadukan dengan fitur MFCC.

Hasil pengujian menunjukkan akurasi model yang bergantung pada jumlah dan kualitas data, dengan nilai di atas 70% sudah dianggap memadai untuk eksperimen awal berbasis MFCC dan SVM. Akurasi ini mengindikasikan bahwa model mampu menangkap karakteristik akustik yang membedakan suara pengguna dari suara lain. Namun, performa tetap dipengaruhi oleh faktor seperti kebersihan rekaman, konsistensi intensitas, jarak mikrofon, serta variasi intonasi. Keterbatasan jumlah atau ketidakseimbangan data juga dapat menurunkan akurasi, sehingga diperlukan sampel yang representatif agar fitur MFCC dapat mencerminkan karakter suara secara optimal.

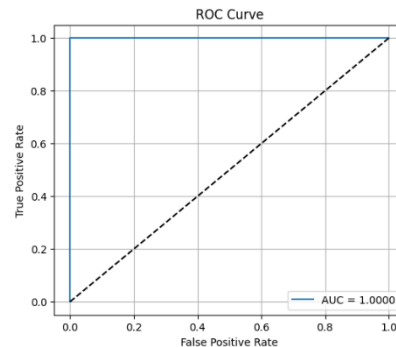
Gambar 3. Matrik Confusion



Tahap akhir penelitian adalah pengujian model terhadap suara baru yang diunggah pengguna. Sistem mendukung pengujian satu file, dua file, maupun beberapa file sekaligus. Setiap rekaman dievaluasi menggunakan `predict_proba` untuk memperoleh probabilitas bahwa suara tersebut milik pengguna. Keputusan dibuat dengan threshold 0.5, sehingga nilai di atas ambang ini diklasifikasikan sebagai suara pengguna, sementara nilai di bawahnya dianggap suara orang lain. Pendekatan threshold ini umum digunakan pada verifikasi speaker dan dapat disesuaikan sesuai kebutuhan sensitivitas. Dalam

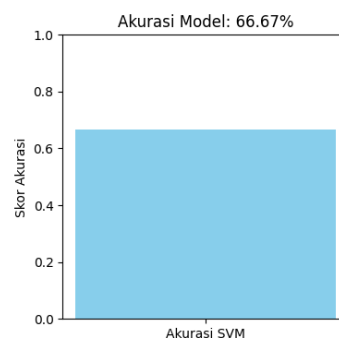
banyak kasus, model memberikan tingkat keyakinan yang tinggi dalam membedakan kedua kelas suara.

Gambar 4. ROC Curve



Secara teoritis, hasil penelitian ini menegaskan bahwa kombinasi MFCC dan SVM tetap relevan untuk tugas pengenalan speaker berskala kecil, meskipun model deep learning kini mendominasi pada dataset besar. Untuk data terbatas, SVM menawarkan performa yang stabil dan efisien. Secara praktis, sistem ini dapat digunakan dalam berbagai aplikasi verifikasi suara personal, termasuk autentikasi dan layanan berbasis biometrik suara. Temuan ini juga sejalan dengan laporan sebelumnya bahwa MFCC–SVM mampu mencapai akurasi 70–95%, sehingga pendekatan ini layak untuk implementasi awal dalam skenario sederhana.

Gambar 5. Akurasi Model



KESIMPULAN

Penelitian ini menunjukkan bahwa kombinasi metode ekstraksi fitur MFCC dan model klasifikasi SVM mampu membedakan suara pengguna dan suara non-pengguna dengan tingkat akurasi yang memadai berdasarkan data uji yang tersedia. Hasil pengujian memberikan indikasi bahwa pola akustik yang terekstraksi dari koefisien MFCC dapat mewakili karakteristik suara secara efektif sehingga memungkinkan proses verifikasi

PENGENALAN SUARA PERSONAL MENGGUNAKAN EKSTRAKSI MFCC DAN SUPPORT VECTOR MACHINE BERBASIS CLOUD COMPUTING

speaker pada konteks sederhana. Meskipun demikian, temuan ini perlu diinterpretasikan secara hati-hati karena performa model sangat dipengaruhi oleh jumlah data, kondisi rekaman, serta variasi vokal masing-masing individu. Dataset yang terbatas dan tidak beragam menjadi salah satu keterbatasan penelitian yang berdampak pada sulitnya melakukan generalisasi terhadap populasi suara yang lebih luas. Untuk penelitian selanjutnya, disarankan penggunaan dataset yang lebih besar dan mencakup variasi kondisi lingkungan, serta eksplorasi model yang lebih kompleks seperti deep learning agar diperoleh performa yang lebih stabil dan robust. Selain itu, pengembangan sistem yang mampu melakukan noise reduction atau augmentasi suara dapat meningkatkan keandalan model ketika diterapkan pada skenario dunia nyata.

DAFTAR REFERENSI

- Afif, M., Aranta, A., & Bimantoro, F. (2022). *VERIFIKASI SUARA MAHASISWA SEBAGAI ALTERNATIF PRESENSI KLASIFIKASI LVQ (Student Voice Verification As Alternative Attendance Presence Using MFCC Feature Extraction And LVQ Classification)*. 4(2), 171–181.
- Ashianti, L., & Junaedi, S. (1978). *ANALYSIS AND VOICE RECOGNITION IN INDONESIAN*. 2011, 131–139.
- Barhoush, M., Hallawa, A., & Schmeink, A. (2023). Speaker identification and localization using shuffled MFCC features and deep learning. *International Journal of Speech Technology*, 26(1), 185–196. <https://doi.org/10.1007/s10772-023-10023-2>
- Chauhan, N., Isshiki, T., & Li, D. (2023). Text - Independent Speaker Recognition System Using Feature - Level Fusion for Audio Databases of Various Sizes. *SN Computer Science*. <https://doi.org/10.1007/s42979-023-02056-w>
- Diaz, R. A. N., Luh, N., Pivin, G., & Budiarta, K. (2024). *Perbandingan kualitas pengenalan suara untuk ekstraksi fitur menggunakan mfcc dan spectral*. 06(01), 58–63.
- Govind, R. S. D., Dubey, J. M. A. K., & Prasanna, K. T. D. S. R. M. (2024). Milestones in speaker recognition. In *Artificial Intelligence Review* (Vol. 57, Issue 3). Springer Netherlands. <https://doi.org/10.1007/s10462-023-10688-w>
- Irfon, G., & Soen, E. (2022). *Implementasi Cloud Computing dengan Google Colaboratory Pada Aplikasi Pengolah Data Zoom Participants*. 6(1), 24–30.
- Kohler, O. (2025). *Investigation of Text-independent speaker verification by Support Vector Machine-Based Machine Learning Approaches*. 1–23.
- Pustaka, T. (2025). 3 1,2,3. 1024–1031.
- Sasilo, A. A., Saputra, R. A., & Ningrum, I. P. (2022). *Sistem Pengenalan Suara menggunakan Metode Mel Frequency Cepstral Coefficients dan Gaussian Mixture Model Voice Recognition System using Mel Frequency Cepstral Coefficients Method and Gaussian Mixture Model*. xx(28), 203–210. <https://doi.org/10.34010/komputika.v11i2.6655>

PENGENALAN SUARA PERSONAL MENGGUNAKAN EKSTRAKSI MFCC DAN SUPPORT VECTOR MACHINE BERBASIS CLOUD COMPUTING

Widodo, S. M., Siswanto, E., & Sudjana, O. (1858). *Penerapan Metode Mel Frequency Ceptral Coefficient dan Learning Vector Quantization untuk Text-Dependent Speaker Identification*. 11(1), 15–20.