

LAPORAN PROJECT AKHIR MATA KULIAH PENGOLAHAN SINYAL DIGITAL

**ANALISIS DAN PERBANDINGAN FILTERING SINYAL AUDIO MENGGUNAKAN
FREKUENSI CUTOFF UNTUK PENGENALAN KATA HALO BERBASIS
ANALISIS SPEKTRAL DAN MFCC**



Disusun oleh:

Muslim Fazlur Rohman	24031554154
Muhammad Ramadhan Albaary Putra	24031554161
Kafka Praya Firmansyah	24031554182

Dosen Pengampu Mata Kuliah:

Dr. Atik Wintarti, M.Kom.
Kartika Chandra Dewi, S.Si., M.Si.

PROJECT PENGOLAHAN SINYAL DIGITAL

PROGRAM STUDI SAINS DATA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS NEGERI SURABAYA

2025

BAB I

PENDAHULUAN

A. Latar Belakang

Pengolahan sinyal audio merupakan bidang penting dalam pengolahan sinyal digital yang banyak diterapkan pada sistem pengenalan ucapan. Suara manusia memiliki karakteristik frekuensi yang berbeda pada setiap individu, meskipun kata yang diucapkan sama. Oleh karena itu, analisis sinyal suara dapat digunakan untuk mengenali pola ujaran berdasarkan ciri spektralnya. Kata “*halo*” sering digunakan sebagai sampel dasar karena memiliki struktur fonetik yang sederhana dan mudah direkam secara konsisten (Schwartz et al., 2018).

Tahap awal yang penting dalam pengolahan sinyal suara adalah filtering dengan menggunakan frekuensi cutoff. Filtering bertujuan untuk membatasi sinyal pada rentang frekuensi tertentu agar komponen yang tidak relevan, seperti noise di luar frekuensi suara manusia, dapat dikurangi. Penggunaan filter band-pass, khususnya filter Butterworth, terbukti efektif dalam mengisolasi sinyal suara manusia dan meningkatkan kualitas sinyal untuk proses analisis selanjutnya (Wu et al., 2024).

Setelah proses filtering, sinyal audio dianalisis menggunakan Fast Fourier Transform (FFT) untuk melihat distribusi frekuensi, serta diekstraksi cirinya menggunakan Mel Frequency Cepstral Coefficients (MFCC). MFCC mampu merepresentasikan karakteristik spektral suara manusia sesuai dengan persepsi pendengaran manusia, sehingga banyak digunakan dalam sistem pengenalan ucapan (Moondra & Chahal, 2023). Oleh karena itu, penelitian ini dilakukan untuk menganalisis dan membandingkan filtering sinyal audio menggunakan frekuensi cutoff pada pengenalan kata “*halo*” berbasis analisis spektral dan MFCC.

B. Tujuan Penelitian

1. Menganalisis pengaruh penerapan filtering sinyal audio menggunakan frekuensi cutoff terhadap kualitas sinyal suara pada pengenalan kata “*halo*”.
2. Membandingkan karakteristik spektral sinyal suara sebelum dan sesudah proses filtering menggunakan analisis Fast Fourier Transform (FFT).
3. Mengidentifikasi ciri suara berdasarkan fitur Mel Frequency Cepstral Coefficients (MFCC) untuk mendukung proses pengenalan kata “*halo*” pada sinyal suara manusia.

C. Manfaat Penelitian

1. Menghasilkan analisis yang menunjukkan pengaruh filtering sinyal audio menggunakan frekuensi cutoff terhadap kualitas sinyal suara pada pengenalan kata *"halo"*.
2. Memberikan perbandingan karakteristik spektral sinyal suara sebelum dan sesudah proses filtering menggunakan analisis FFT dan fitur MFCC sebagai dasar pengenalan suara.
3. Mendorong pemanfaatan metode analisis spektral dan MFCC dalam pengolahan sinyal suara manusia sebagai contoh penerapan pengolahan sinyal digital pada sistem pengenalan ucapan sederhana.































BAB II

METODE

A. Dataset

Dataset yang digunakan dalam project ini berupa rekaman suara yang dibuat secara mandiri oleh anggota kelompok. Setiap anggota kelompok merekam pengucapan kata “*halo*” sebanyak 10 kali dengan kondisi perekaman yang relatif beragam. Dengan demikian, total dataset yang digunakan dalam project ini berjumlah 30 sampel rekaman suara. Seluruh rekaman disimpan dalam format audio dan dikelompokkan berdasarkan masing-masing pembicara untuk memudahkan proses analisis. kami juga membuat suara non halo yaitu seperti suara *noise* (berisik) dan suara non halo yang totalnya berjumlah 7.

Penggunaan dataset hasil rekaman sendiri bertujuan untuk memastikan keaslian data serta memberikan kontrol penuh terhadap sumber dan karakteristik sinyal suara yang dianalisis. Dataset ini kemudian digunakan sebagai input pada proses filtering, analisis spektral, dan ekstraksi ciri MFCC untuk mendukung pengenalan kata “*halo*” berbasis sinyal audio.

 farel1.wav	 Kafka1.wav	 Rama1.wav
 farel2.wav	 Kafka2.wav	 Rama2.wav
 farel3.wav	 Kafka3.wav	 Rama3.wav
 farel4.wav	 Kafka4.wav	 Rama4.wav
 farel5.wav	 Kafka5.wav	 Rama5.wav
 farel6.wav	 Kafka6.wav	 Rama6.wav
 farel7.wav	 Kafka7.wav	 Rama7.wav
 farel8.wav	 Kafka8.wav	 Rama8.wav
 farel9.wav	 Kafka9.wav	 Rama9.wav
 farel10.wav	 Kafka10.wav	 Rama10.wav

B. Ekstraksi Fitur Menggunakan MFCC

Mel-Frequency Cepstral Coefficients (MFCC). MFCC dipilih karena mampu merepresentasikan karakteristik suara manusia berdasarkan skala Mel yang menyerupai cara kerja sistem pendengaran manusia.

pada project ini, digunakan 50 koefisien MFCC diekstraksi dari setiap file audio. Nilai MFCC kemudian dirata-ratakan untuk menghasilkan satu vektor fitur representatif per file. Selanjutnya, fitur-fitur tersebut dirata-ratakan kembali untuk setiap pembicara. MFCC merupakan fitur yang paling umum dan efektif digunakan dalam sistem pengenalan dan identifikasi pembicara

```
for speaker in os.listdir(DATASET_PATH):
    speaker_path = os.path.join(DATASET_PATH, speaker)
    mfcc_list = []

    for file in os.listdir(speaker_path):
        if file.endswith(".wav"):
            audio_path = os.path.join(speaker_path, file)
            y, sr = librosa.load(audio_path, sr=16000)

            mfcc = librosa.feature.mfcc(y=y, sr=sr, n_mfcc=50)
            mfcc_mean = np.mean(mfcc, axis=1)
            mfcc_list.append(mfcc_mean.tolist())

    speaker_features[speaker] = np.mean(np.array(mfcc_list), axis=0).tolist()
```

```
def identify_speaker(mfcc_test):
    distances = {}
    for speaker, mfcc in db.items():
        dist = np.linalg.norm(np.array(mfcc_test) - np.array(mfcc))
        distances[speaker] = dist

    predicted = min(distances, key=distances.get)
    return predicted, distances

y, sr = librosa.load("/content/drive/Shared drives/Pengolahan Sinyal Data/Dataset/Farel/fareli.wav", sr=16000)
mfcc_test = np.mean(librosa.feature.mfcc(y=y, sr=sr, n_mfcc=50), axis=1)
predicted, distances = identify_speaker(mfcc_test)
```

C. Analisis Spektral Menggunakan FFT

Fast Fourier Transform (FFT) merupakan metode yang digunakan untuk mentransformasikan sinyal dari domain waktu ke domain frekuensi. Dalam project ini, FFT diterapkan pada sinyal audio pengucapan kata “halo” untuk menganalisis distribusi energi sinyal pada berbagai rentang frekuensi. Setiap sinyal suara diubah menjadi spektrum magnitudo frekuensi, kemudian dilakukan perhitungan spektrum rata-rata untuk masing-masing pembicara. Analisis spektrum harmonik suara terbukti efektif dalam membedakan karakteristik vokal individu karena setiap pembicara memiliki pola harmonik yang unik

```

DATASET_PATH = "/content/drive/SharedDrives/Pengolahan Sinyal Data/Dataset/Rama"

files = [f"Rama{i}.wav" for i in range(1, 10+1)]

plt.figure(figsize=(12,5))

for i, filename in enumerate(files):
    path = os.path.join(DATASET_PATH, filename)

    y, sr = librosa.load(path, sr=16000)

    # FFT dasar
    fft = np.abs(np.fft.rfft(y))
    freqs = np.fft.rfftfreq(len(y), 1/sr)

```

D. Penerapan Band-Pass Filter Berbasis Frekuensi Cutoff

Komponen yang tidak penting seperti noise lingkungan, getaran , maupun komponen frekuensi tinggi yang tidak diperlukan pada proses pengenalan suara dapat dihilangkan dengan menerapkan tahap filtering.

Pada project kali ini kami menggunakan band pass filter dengan rentang frekuensi cutoff antara 80 Hz dan 1100 Hz. rentang frekuensi ini kami pilih karena menurut penelitian, suara vokal manusia berada pada rentang 80-1100 Hz. Frekuensi di bawah 80 Hz umumnya berisi noise rendah, sedangkan frekuensi di atas 1100 Hz cenderung mengandung noise atau informasi yang tidak relevan.

```

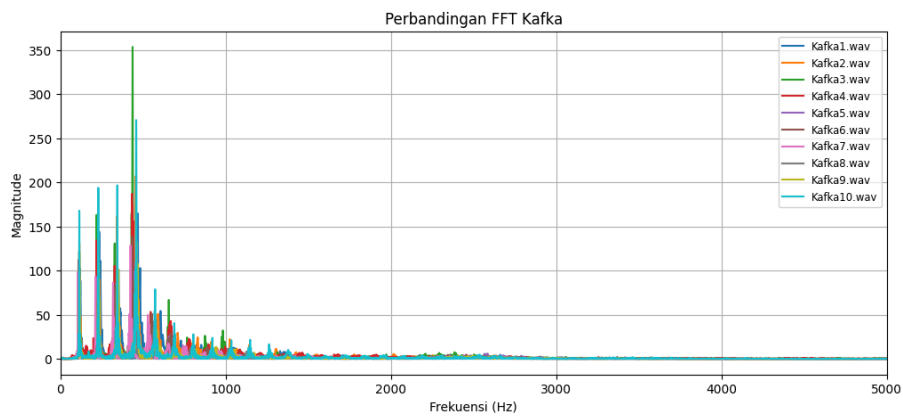
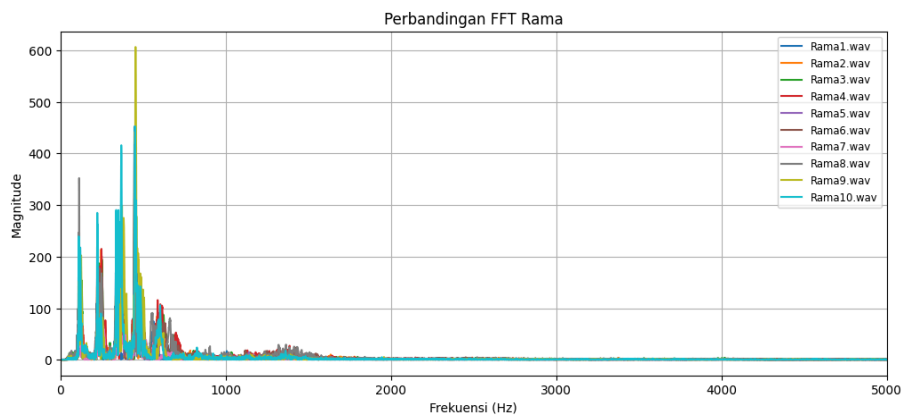
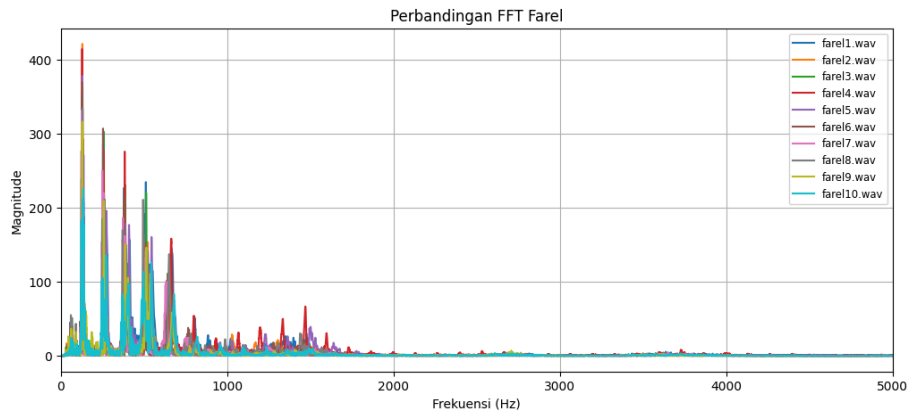
# Band-pass
def bandpass_filter(y, sr, low=80, high=1100):
    nyquist = sr / 2
    b, a = butter(4, [low/nyquist, high/nyquist], btype='band')
    return filtfilt(b, a, y)

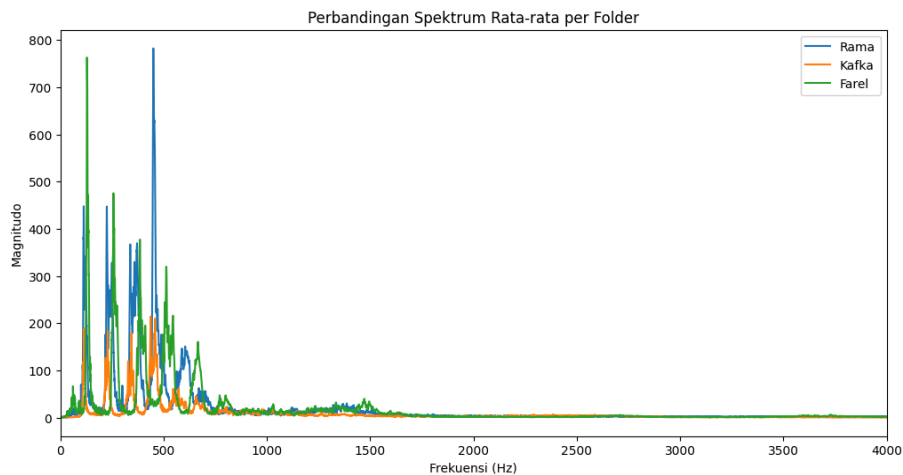
```

BAB III

HASIL DAN DISKUSI

A. Perbandingan File Sebelum Filter

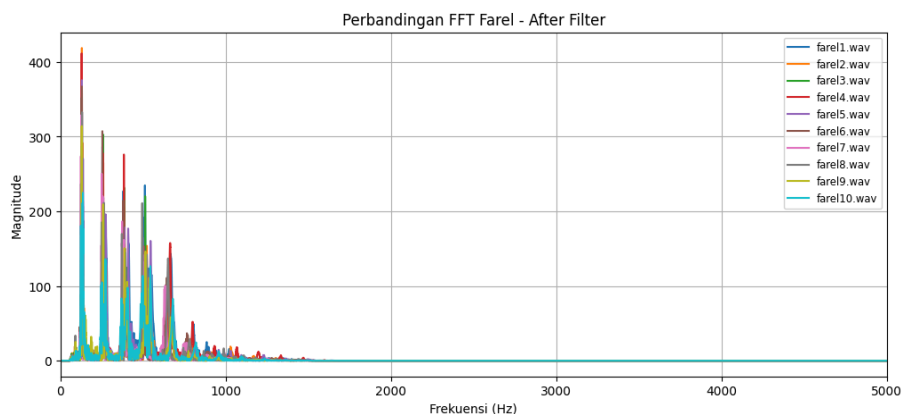


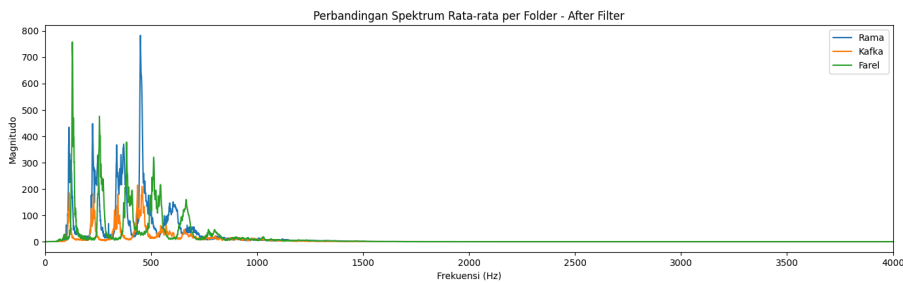
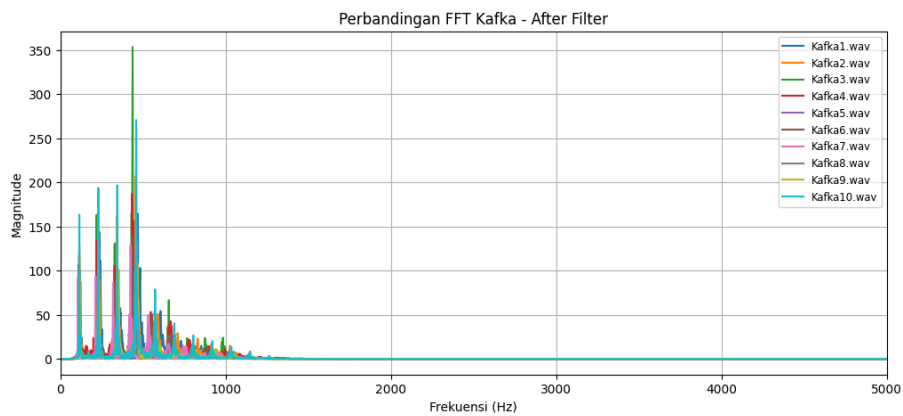
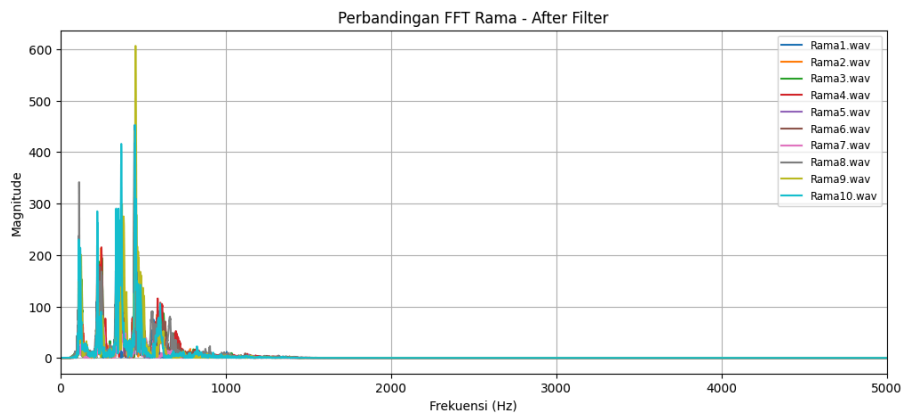


Rama memiliki energi spektrum tertinggi diantara Kafka dan Farel. Namun hal ini tidak bisa disimpulkan bahwa Rama memiliki suara yang lebih kuat karena faktor-faktor seperti jarak rekaman, level gain mikropon, dan lain-lain turut berpengaruh.

Ketiga grafik menunjukkan bahwa puncak spektrum dominan berada pada rentang frekuensi di bawah 1000 Hz. Puncak-puncak yang muncul merepresentasikan harmonik dan formant vokal. Kondisi ini sangat konsisten dengan karakteristik vokal manusia. Energi yang dominan pada rentang dibawah 1000 Hz menunjukkan bahwa sumber sinyal ini adalah suara vokal manusia.

B. Perbandingan File Setelah Bandpass Filter





Bandpass filter berfungsi sebagai penyaring frekuensi yang hanya mengizinkan gelombang dalam rentang frekuensi tertentu untuk lewat, dan meredam frekuensi yang berada di bawah atau di atas rentang tersebut. Penentuan rentang ini kami dapatkan jurnal, dimana rentang menunjukkan angka 800-1100 Hz.

Setelah dilakukan *bandpass filter*, energi spektrum yang tidak berada didalam rentang yang telah ditentukan akan diredam. Dari keempat grafik diatas, energi yang kurang dari 80 dan lebih dari 1100 hampir tidak ada kenaikan yang signifikan.

C. Prediksi Suara menggunakan Mel-Frequency Cepstral Coefficients (MFCC)

Sinyal suara yang telah direkam, diubah menjadi serangkaian angka (numerik) dimana kami menggunakan 50 fitur untuk setiap pembicara. Dari proses itu, setiap pembicara akan memiliki angka yang unik dan akan menjadi identitas. Untuk melakukan prediksi, kita mengubah suara yang akan diprediksi menjadi numerik dengan 50 fitur juga, kemudian dihitung jaraknya menggunakan euclidean. Hasil yang paling kecil menandakan hasil prediksi karena jarak terdekat kepada identitas setiap pembicara.

```
Pembicara terdeteksi: Farel  
  
Jarak Euclidean ke tiap pembicara:  
Rama: 165.9288  
Kafka: 114.9389  
Farel: 18.3877
```

Hasil menunjukkan farel memiliki jarak terkecil, maka hasil prediksi merupakan suara yang dimiliki oleh farel.

BAB III

KESIMPULAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa penerapan filtering sinyal audio menggunakan frekuensi cutoff berpengaruh terhadap kualitas sinyal suara pada pengenalan kata "*halo*". Band-pass filter dengan rentang frekuensi 80–1100 Hz mampu meredam komponen frekuensi yang tidak relevan, khususnya noise pada frekuensi rendah dan tinggi, sehingga sinyal suara vokal manusia menjadi lebih jelas dan terfokus.

Hasil analisis spektral menggunakan Fast Fourier Transform (FFT) menunjukkan bahwa sebelum proses filtering, sinyal suara masih mengandung energi pada frekuensi di luar rentang vokal manusia. Setelah diterapkan band-pass filter, energi spektrum di luar rentang cutoff mengalami peredaman yang signifikan, sementara komponen utama suara tetap dipertahankan. Hal ini membuktikan bahwa filtering dengan frekuensi cutoff yang tepat efektif dalam meningkatkan kualitas representasi spektral sinyal suara.

Selain itu, proses ekstraksi fitur menggunakan Mel Frequency Cepstral Coefficients (MFCC) berhasil merepresentasikan karakteristik suara masing-masing pembicara dalam bentuk vektor numerik. Hasil prediksi menggunakan perhitungan jarak Euclidean menunjukkan bahwa sistem mampu mengidentifikasi pembicara dengan memilih jarak terkecil terhadap data latih. Dengan demikian, MFCC terbukti efektif digunakan sebagai fitur untuk mendukung pengenalan kata "*halo*" dan identifikasi pembicara berbasis sinyal audio.

Secara keseluruhan, penelitian ini menunjukkan bahwa kombinasi filtering berbasis frekuensi cutoff, analisis spektral FFT, dan ekstraksi fitur MFCC dapat digunakan secara efektif dalam pengenalan kata sederhana dan identifikasi pembicara. Hasil penelitian ini diharapkan dapat menjadi dasar pengembangan sistem pengenalan suara yang lebih lanjut dengan jumlah data, variasi kata, dan metode klasifikasi yang lebih kompleks.

DAFTAR PUSTAKA

- Schwartz, J. L., Basirat, A., Ménard, L., & Sato, M. (2018). Effects of signal bandwidth and noise on individual speaker identification. *The Journal of the Acoustical Society of America*, 144(5), EL447–EL452. <https://doi.org/10.1121/1.5078770>
- Wu, Y., Wang, Y., Zhang, Z., & Li, X. (2024). Human voice sensing through radio frequency technologies: A comprehensive review. *IEEE Transactions on Instrumentation and Measurement*, 73, 1–17. <https://doi.org/10.1109/TIM.2024.3354321>
- Moondra, R., & Chahal, P. (2023). Speaker recognition improvement for degraded human voice using modified MFCC with GMM. *International Journal of Advanced Computer Science and Applications*, 14(6), 298–305. <https://doi.org/10.14569/IJACSA.2023.0140634>
- Ranjan, R., & Thakur, A. (2019). Analysis of feature extraction techniques for speech recognition system. *International Journal of Innovative Technology and Exploring Engineering*, 8(7C2), 197–200. <https://www.ijitee.org/wp-content/uploads/papers/v8i7c2/G10460587C219.pdf>
- Sasilo, A. A., Saputra, R. A., & Ningrum, I. P. (2022). Sistem pengenalan suara menggunakan metode Mel Frequency Cepstral Coefficients dan Gaussian Mixture Model. *Komputika: Jurnal Sistem Komputer*, 11(2), 203–210. <https://ojs.unikom.ac.id/index.php/komputika/article/download/6655/3247>
- Vyas, G., & Kumari, B. (2013). Speaker recognition system based on MFCC and DCT. *International Journal of Engineering and Advanced Technology*, 2(5), 167–170. <https://www.ijeat.org/wp-content/uploads/papers/v2i5/E1736062513.pdf>

LAMPIRAN

Link Youtube: <https://youtu.be/ppCZlD3b1Lg>

Coding Colab:  PSD Final_Kelompok 5.ipynb