

UAS DATA MINING II

KLASIFIKASI SUARA MANUSIA DAN TTS

KELAS SD-A1 KELOMPOK A



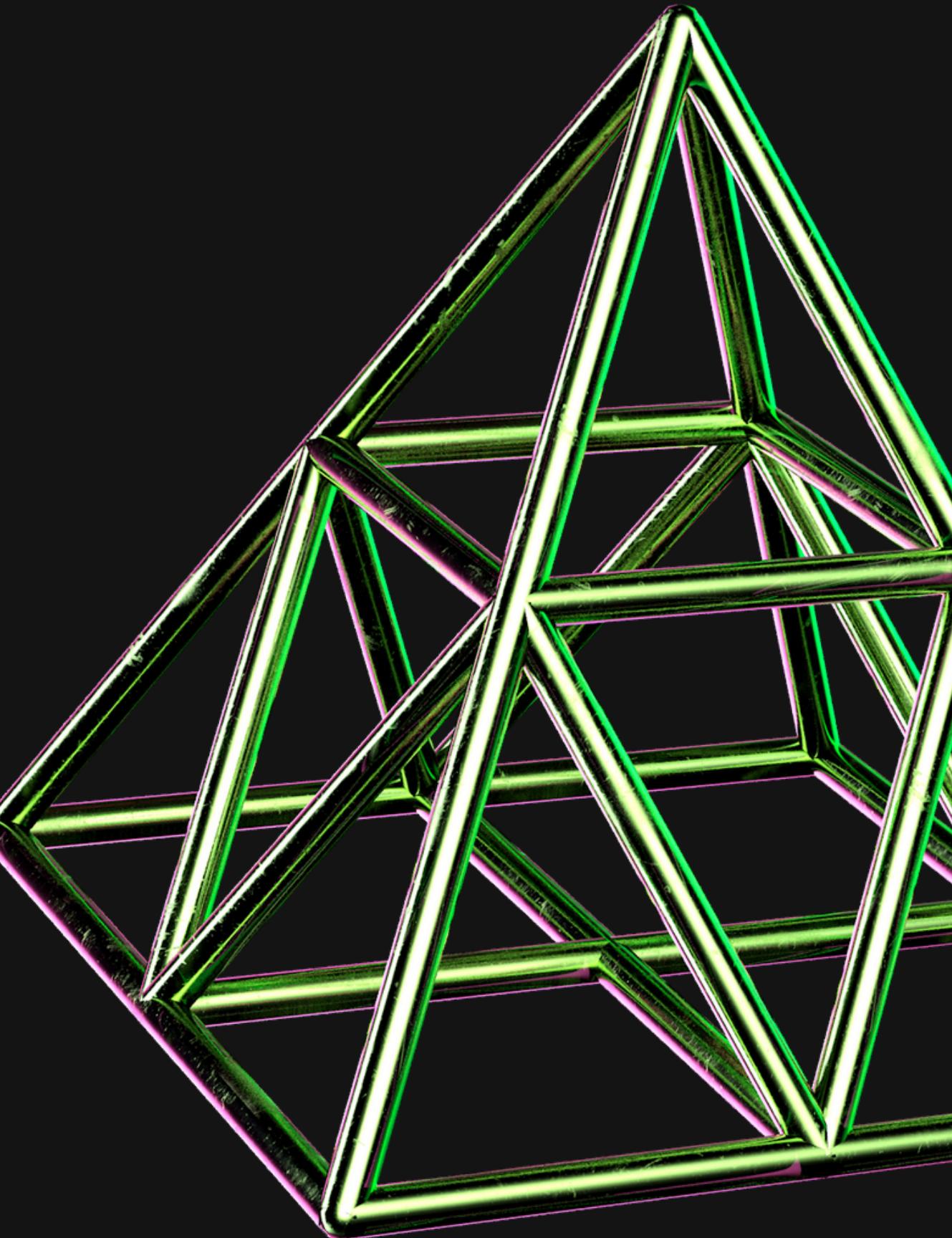
Anggota

162012133042 - Viola Rosa Indah Mulya

162012133087 - Fadya Amalia Zahra

162012133088 - Lukas Leonardo S

162012133092 - Muhammad Othman Lutfi



OUTLINE

latar belakang

deskripsi dataset

analisis fitur

klasifikasi dengan CNN

Latar Belakang



Perkembangan teknologi sekarang ini membuat orang-orang semakin cerdik dalam melakukan pemalsuan, mulai dari gambar, suara dan lain-lain. Suara merupakan hal yang dapat dilakukan salah satu pemalsuan suara adalah dengan menggunakan text-to-speech generator. Text-to-Speech (TTS) merupakan sistem otomatis yang dapat mengubah teks menjadi ucapan atau suara dengan penyusunan fonem-fonem untuk membentuk suatu ucapan, tujuannya yaitu untuk meniru gaya pronunciation suara manusia untuk mengucapkan teks secara alami dan natural. Banyak web yang dapat mentransformasi teks ke dalam bentuk audio. Teknologi saat ini memungkinkan AI untuk menghasilkan suara yang natural seperti suara manusia.

Oleh sebab itu, tujuan penelitian ini dilakukan untuk mengetahui perbedaan dari suara text-to-speech dan suara manusia dengan menggunakan klasifikasi secara otomatis pada sistem aplikasi atau web. Menggunakan speech recognition agar komputer dapat mengenali dan membedakan suara manusia dan suara text-to-speech, sehingga setelah pengklasifikasian berhasil dibuat dan diimplementasikan, pemalsuan terhadap terhadap suara text-to-speech dapat terhindarkan.

Sumber Data

Table 1. Variabel Penelitian

Variabel	Keterangan
Manusia	Data suara manusia asli dari platform YouTube
TTS	Audio text-to-speech (AI)

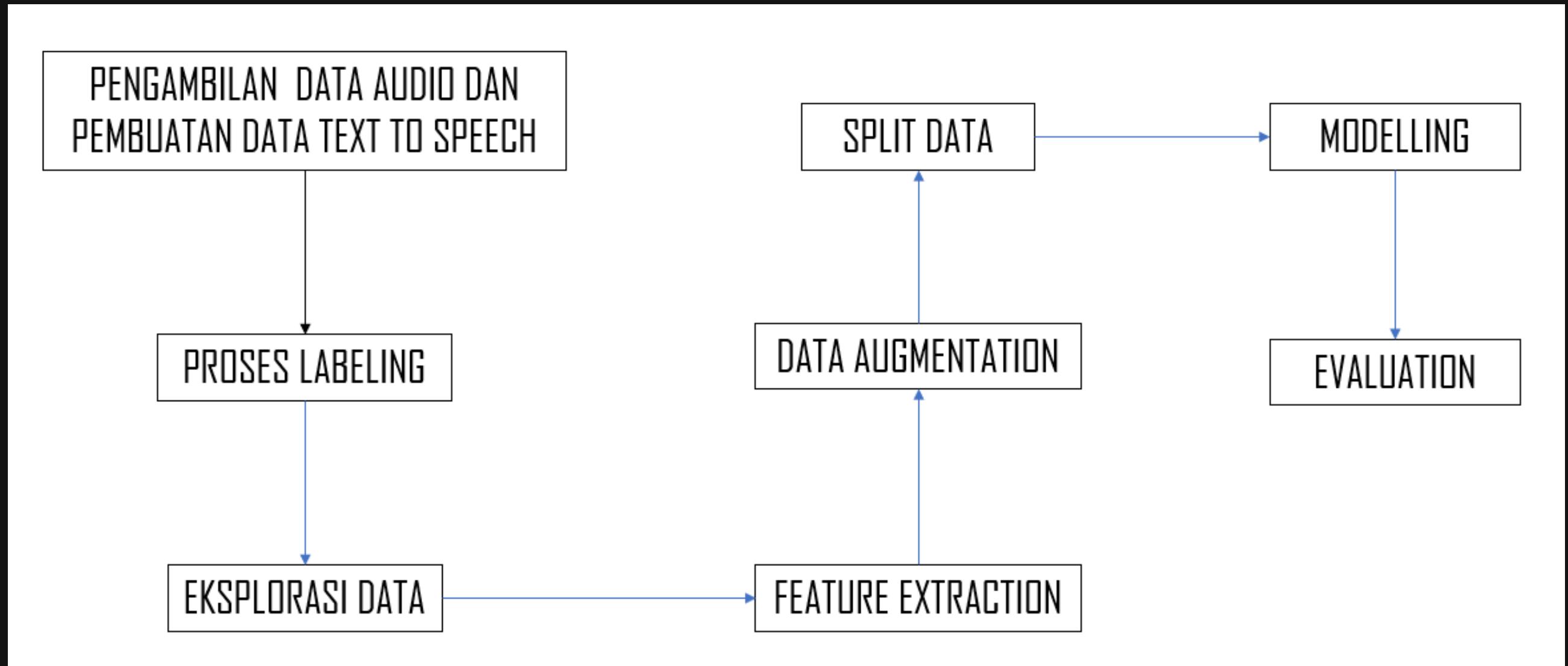


Data hasil crapping
video youtube untuk
label manusia

Data pembuatan
suara text-to-speech
yang erasal dari
website dan python

Dataset

Guideline analisi dan klasifikasi

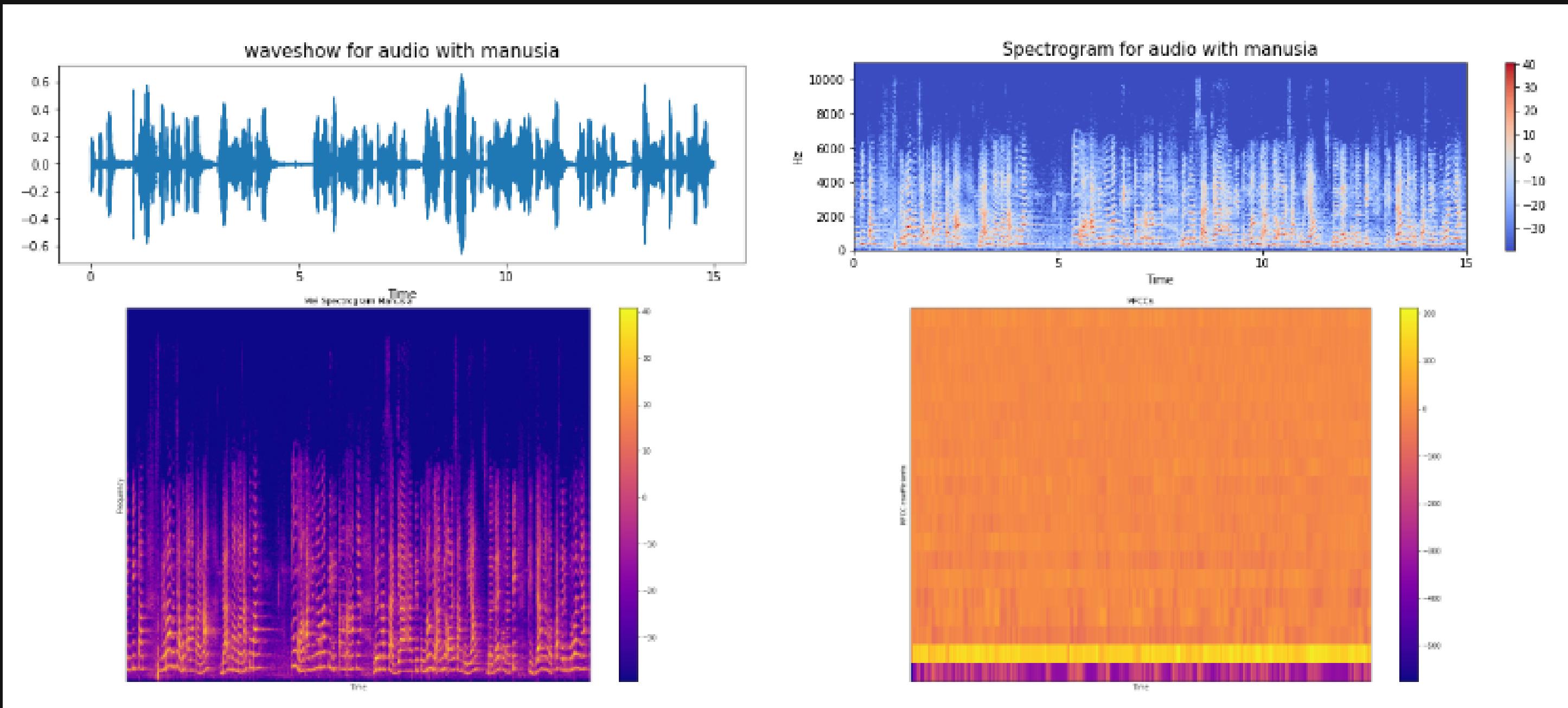


ANALISIS FITUR

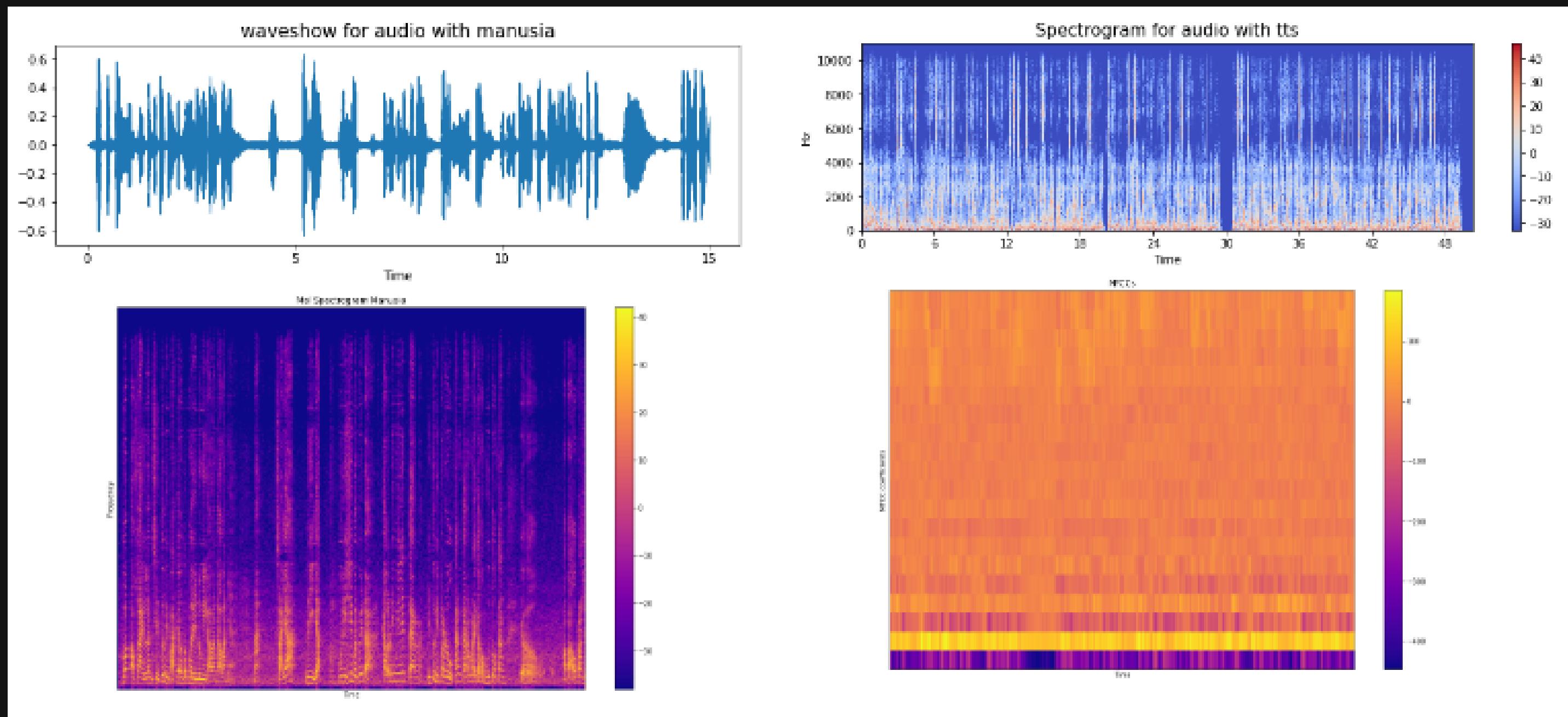
Visualisasi

Suara Manusia

- Sample 1

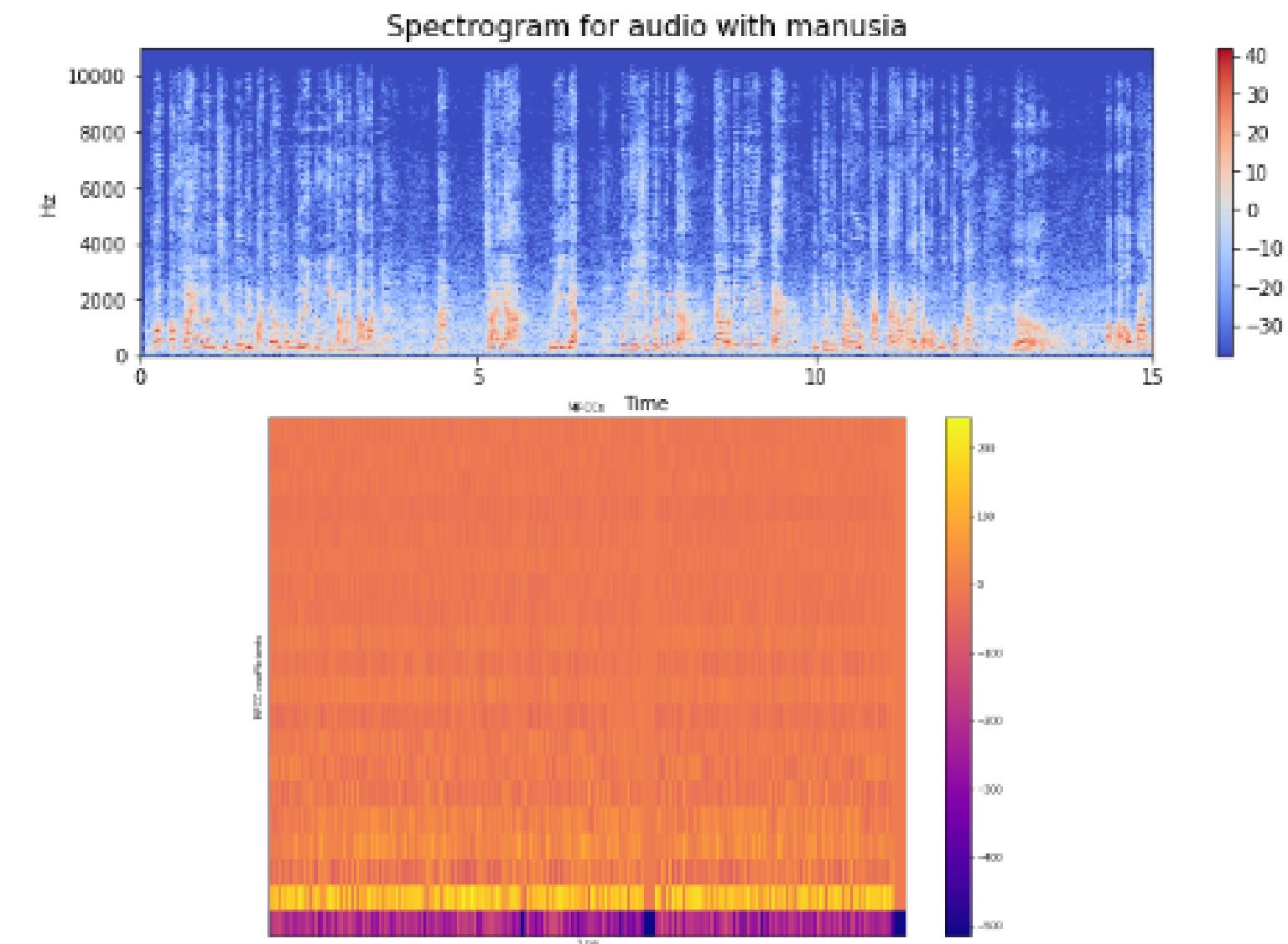
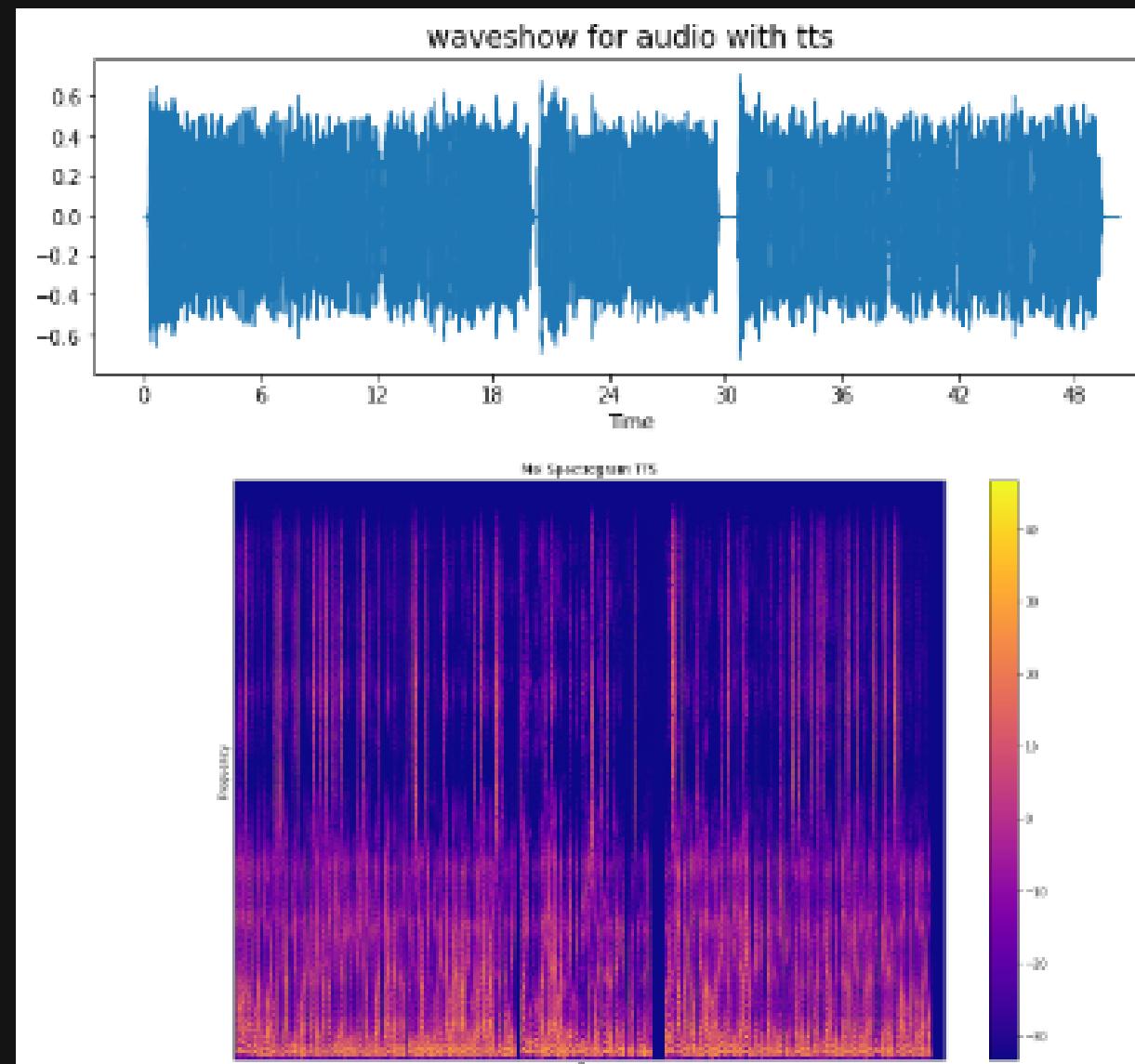


- Sample 2



visualisasi

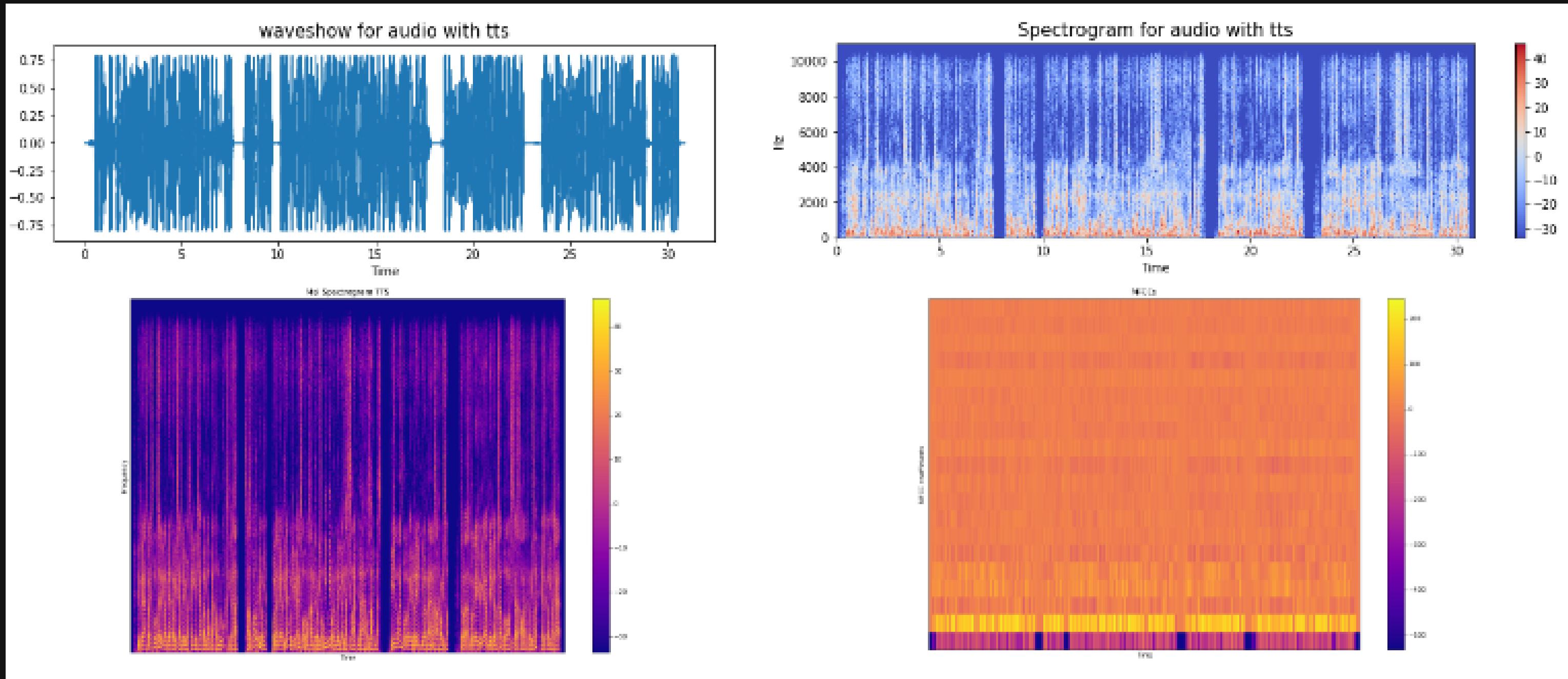
- Sample 1



Suara TTS

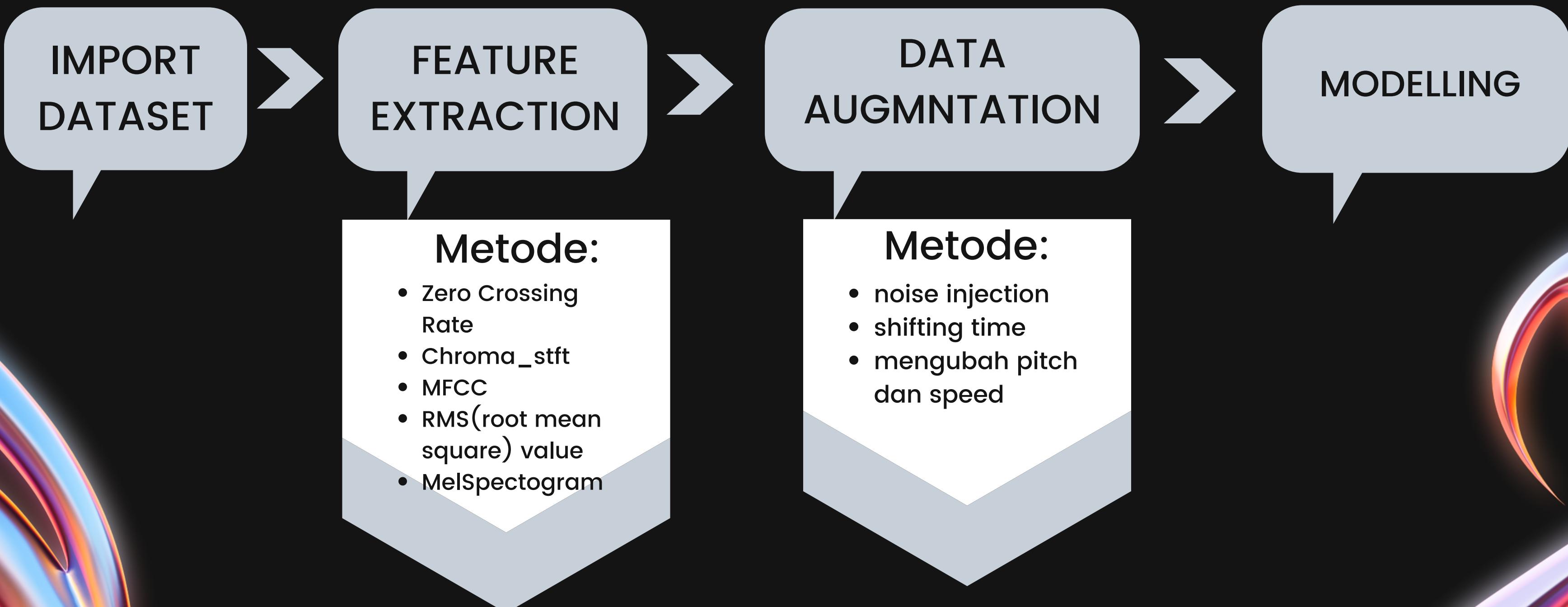


- Sample 2



KLASIFIKASI

ALUR PEMBUATAN MODEL



FITUR-FITUR

	0	1	2	3	4	5	6	7	8	9	...	153	154	155	156	157	158	159	160	161	labels
0	0.072718	0.257500	0.233406	0.226188	0.280225	0.275721	0.352861	0.338557	0.441913	0.680447	...	0.000023	0.000034	0.000053	0.000081	0.000052	0.000039	0.000013	5.644286e-06	4.431140e-06	trimmed_manusia
1	0.098402	0.290333	0.260447	0.282406	0.323051	0.334948	0.385765	0.364106	0.541846	0.695168	...	0.000959	0.001009	0.000981	0.001018	0.000979	0.000969	0.000929	1.033226e-03	9.965028e-04	trimmed_manusia
2	0.085605	0.287791	0.257705	0.224446	0.284970	0.292955	0.338435	0.335316	0.436603	0.717311	...	0.000006	0.000005	0.000007	0.000008	0.000018	0.000017	0.000007	3.865219e-06	3.366940e-06	trimmed_manusia
3	0.096847	0.416755	0.426715	0.497403	0.550624	0.478473	0.476163	0.548842	0.706878	0.685144	...	0.000007	0.000008	0.000010	0.000022	0.000021	0.000021	0.000009	1.009505e-06	3.919322e-07	trimmed_manusia
4	0.279283	0.539433	0.553431	0.601811	0.686446	0.651020	0.604370	0.611997	0.723992	0.717314	...	0.005277	0.004882	0.005232	0.005337	0.005230	0.005176	0.005168	5.336315e-03	5.311970e-03	trimmed_manusia
...	
1429	0.043086	0.436114	0.725285	0.339924	0.252276	0.331246	0.412514	0.812879	0.397220	0.298305	...	0.000491	0.000460	0.000464	0.000447	0.000444	0.000407	0.000387	3.619737e-04	3.652386e-04	trimmed_manusia
1430	0.031641	0.313314	0.391307	0.200397	0.123008	0.160177	0.328964	0.799602	0.341298	0.214603	...	0.000064	0.000044	0.000032	0.000031	0.000034	0.000014	0.000004	5.266823e-07	2.892785e-07	trimmed_manusia
1431	0.122767	0.415636	0.225559	0.188088	0.214373	0.234979	0.390589	0.512587	0.260732	0.229040	...	0.036401	0.038818	0.034482	0.020014	0.016535	0.011759	0.002328	1.060531e-04	1.051278e-05	trimmed_manusia
1432	0.220866	0.455140	0.357468	0.352435	0.357070	0.409241	0.523165	0.532301	0.352785	0.370351	...	0.051884	0.055529	0.052115	0.034541	0.032516	0.029097	0.019473	1.671800e-02	1.723248e-02	trimmed_manusia
1433	0.122551	0.381840	0.216380	0.175989	0.189739	0.226402	0.391346	0.499145	0.238032	0.222614	...	0.003013	0.006989	0.008007	0.009076	0.005162	0.002385	0.000521	5.803028e-05	7.009388e-06	trimmed_manusia

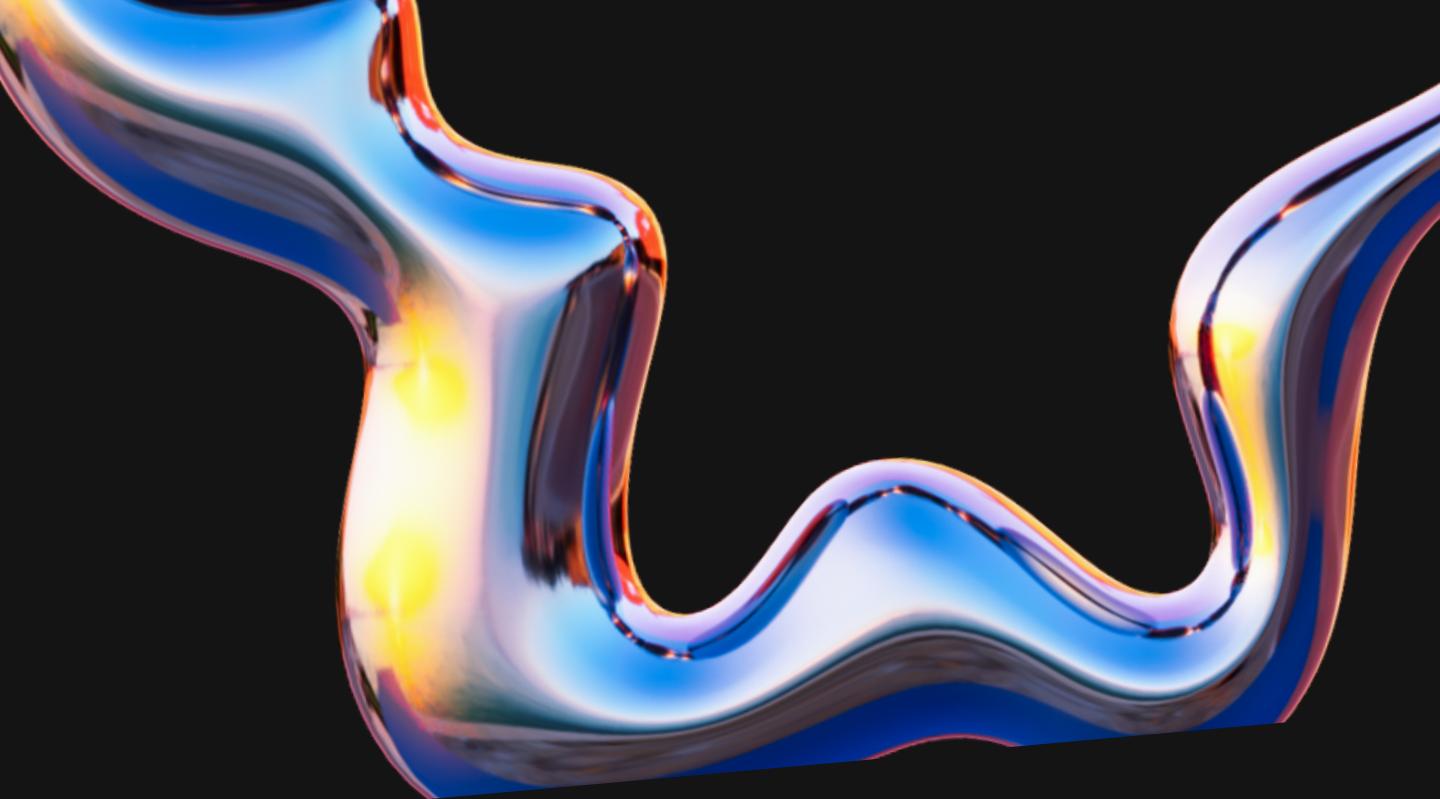
1434 rows × 163 columns

Model

Model: "sequential"

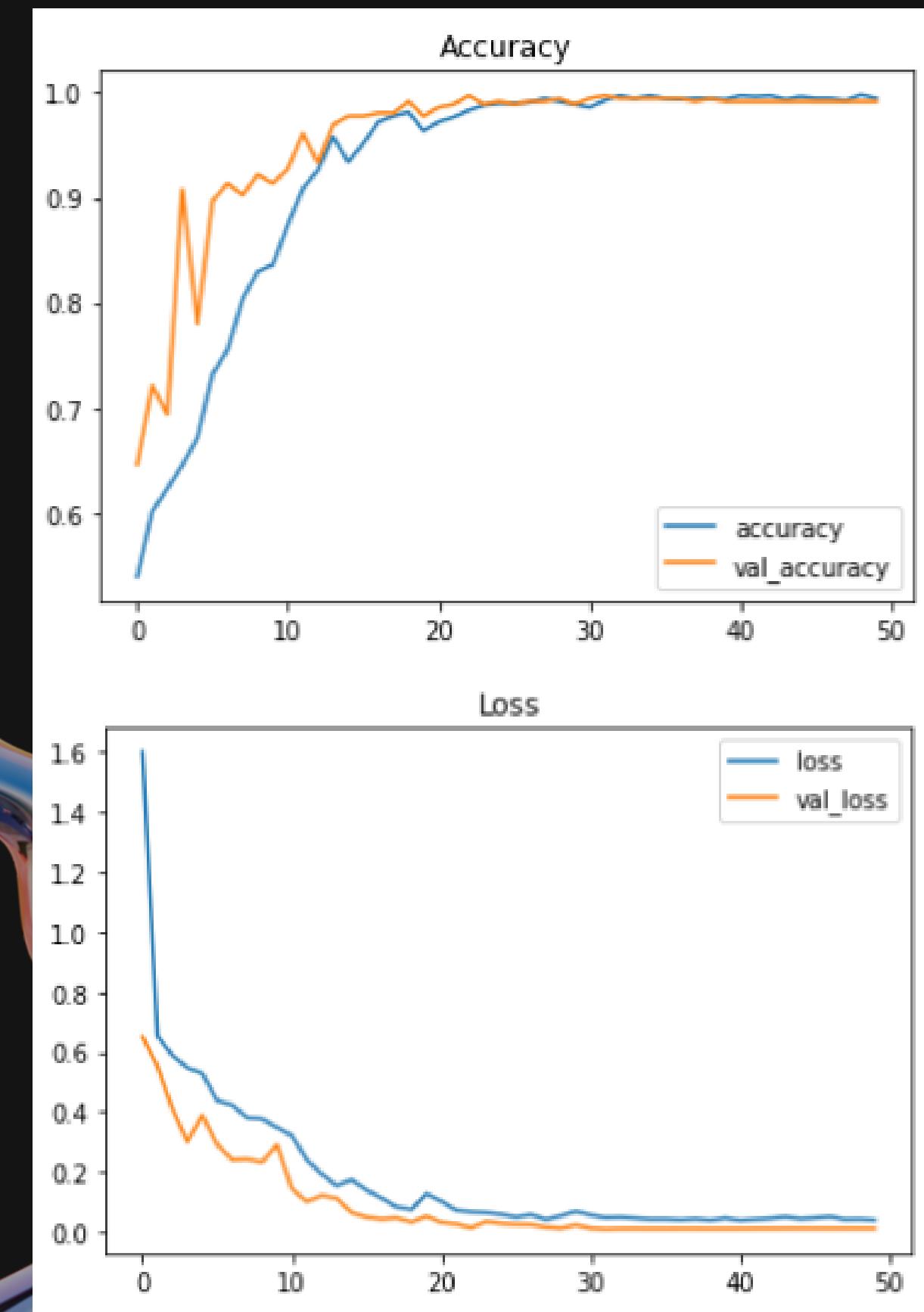
Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 162, 256)	1536
max_pooling1d (MaxPooling1D)	(None, 81, 256)	0
)		
conv1d_1 (Conv1D)	(None, 81, 256)	327936
max_pooling1d_1 (MaxPooling1D)	(None, 41, 256)	0
conv1d_2 (Conv1D)	(None, 41, 128)	163968
max_pooling1d_2 (MaxPooling1D)	(None, 21, 128)	0
dropout (Dropout)	(None, 21, 128)	0
conv1d_3 (Conv1D)	(None, 21, 64)	41024
max_pooling1d_3 (MaxPooling1D)	(None, 11, 64)	0

flatten (Flatten)	(None, 704)	0
dense (Dense)	(None, 32)	22560
dropout_1 (Dropout)	(None, 32)	0
dense_1 (Dense)	(None, 2)	66
<hr/>		
Total params: 557,090		
Trainable params: 557,090		
Non-trainable params: 0		
<hr/>		



HASIL

	precision	recall	f1-score	support
trimmed_manusia	0.98	0.98	0.98	180
trimmed_tts	0.98	0.98	0.98	179
accuracy			0.98	359
macro avg	0.98	0.98	0.98	359
weighted avg	0.98	0.98	0.98	359



KESIMPULAN

- Terdapat perbedaan fitur antara suara yang asli dihasilkan manusia dengan suara yang dibuat dengan text-to-speech.
- visualisasi ini menunjukan bahwa suara yang dihasilkan manusia cenderung lebih berbentuk tidak teratur, dilihat berdasarkan tinggi rendahnya amplitudo dan kerapatan antara gelombang-gelombang suara.
- Suara yang dihasilkan oleh text-to-speech di sisi lain berbentuk lebih rapi, teratur dan konsisten.
- Dapat dilakuakn klasifikasi
- mendapat akurasi sampai dengan 99% menggunakan CNN

Referensi

- [1] Sudirman Melangi. (2018). Text To Speech Bahasa Indonesia Menggunakan Synthesizer Concatenation Berbasis Fonem. *Jurnal Teknik Elektro CosPhi*, 2(2), 31–36.
- [2] <https://cloud.google.com/text-to-speech/docs/voices>
- [3] Oktaviari, E. A. (2019). Journal of Chemical Information and Modeling, 53(9), 1689.
- [4] Dimmita, Nagajyothi & Siddaiah, P.. (2018). Speech Recognition Using Convolutional Neural Networks. *International Journal of Engineering and Technology(UAE)*. 7. 133–137.



TERIMA KASIH!

