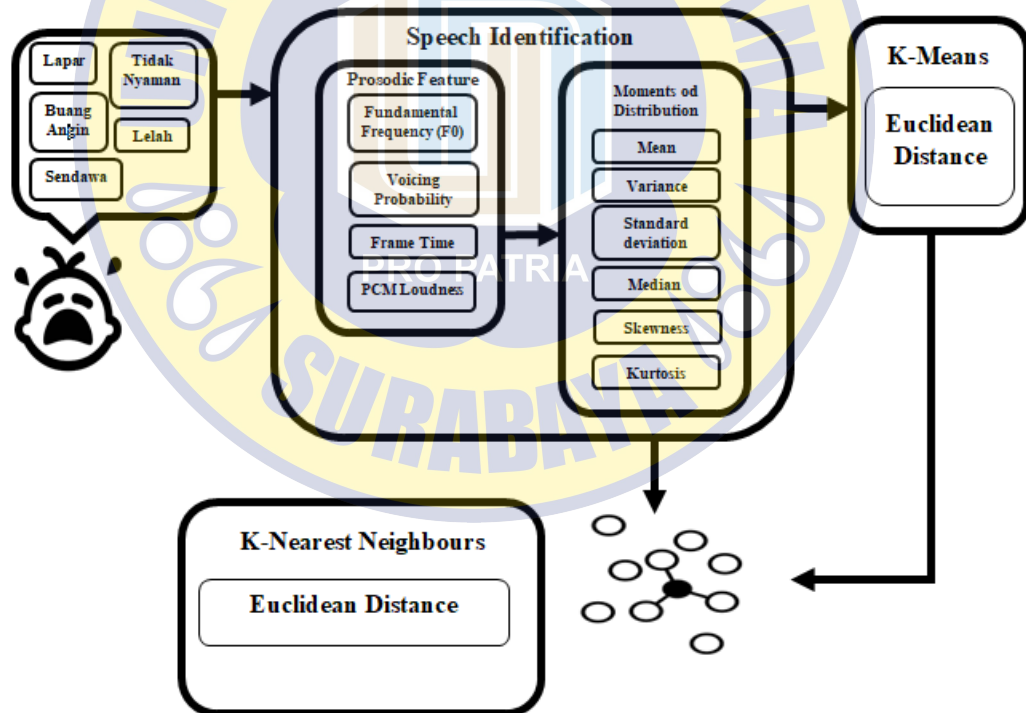


BAB III

METODOLOGI PENELITIAN

3.1.Rancangan Penelitian

Pembuatan model identifikasi suara tangisan bayi yang diterapkan berdasarkan kategori *Dunstan Baby Language* menggunakan metode ekstraksi fitur suara *Prosodic Feature* dan algoritma *K-Nearest Neighbours*. Data yang digunakan berasal dari bayi berumur 0-3 bulan yang telah diberikan label sesuai kategori *Dunstan Baby Language* sebanyak 150 file suara dalam format .wav. Berikut gambar desain sistem untuk mengilustrasikan penelitian ini.

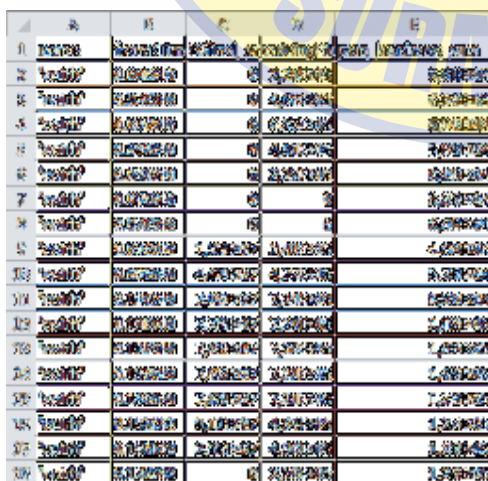


Gambar 3. 1Desain Sistem Proses Penelitian

Bayi yang berusia lebih dari 3 bulan memiliki suara tangisan yang berbeda sehingga suara yang dihasilkan menurut kategori *Dunstan Baby Language* tidak muncul. Penelitian milik Irma Amelia Dewi menggunakan 25 sampel data mencapai akurasi 76%[2], maka di penelitian kali ini akan dicoba menggunakan data sebanyak 150 sampel data. Setelah sampel data terkumpul maka diproses melalui *Speech Identification*, yaitu menggunakan ekstraksi fitur suara *Prosodic Feature*, mencari pola sinyal dengan *Moment of Distribution*.

3.2. Populasi dan Sampel

Terdapat 5 kategori suara pada data identifikasi suara tangisan bayi berdasarkan *Dunstan Baby Language*, yaitu “neh” berarti lapar, “owh” berarti lelah atau mengantuk, “eairh” berarti ingin buang angin, “eh” berarti ingin sendawa dan “heh” berarti tidak nyaman. Pada gambar berikut akan ditampilkan contoh hasil ekstraksi fitur menggunakan *Prosodic Feature* dari salah satu suara “Eairh” berarti ingin buang angin dan salah satu suara “Eh” berarti ingin sendawa.

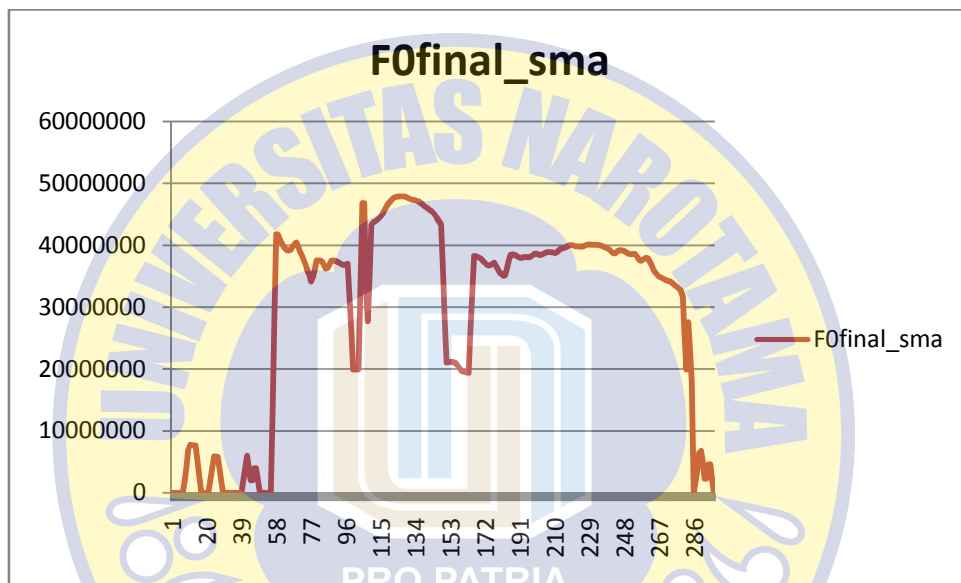


Gambar 3. 2 hasil ekstraksi fitur salah satu suara “Eairh”

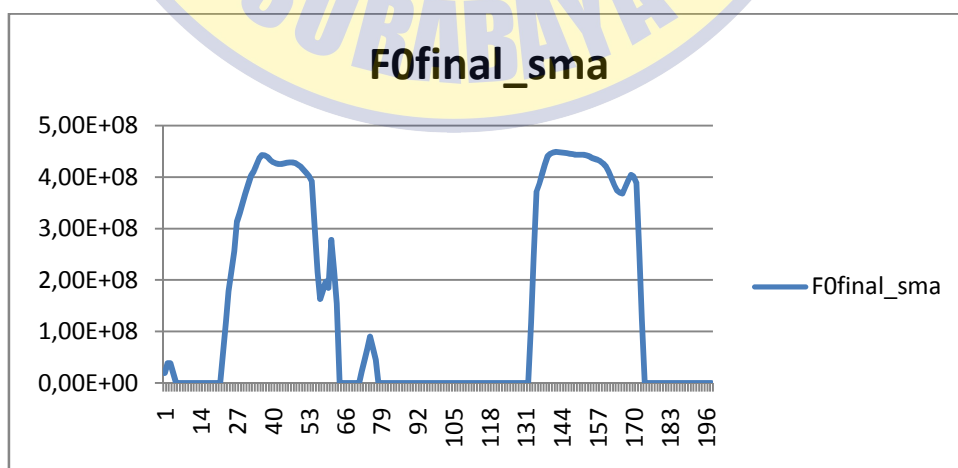
	A	B	C	D	E
1	name	frameTim	F0final_sr	voicingFir	pcm_loudness_sma
2	'eh'	0.010000	1,93E+07	2,35E+05	7,97E+04
3	'eh'	0.020000	3,84E+07	4,73E+05	1,03E+05
4	'eh'	0.030000	3,84E+07	7,04E+05	1,13E+05
5	'eh'	0.040000	1,92E+07	6,97E+05	1,39E+05
6	'eh'	0.050000	0	6,70E+05	1,06E+05
7	'eh'	0.060000	0	6,59E+05	9,02E+04
8	'eh'	0.070000	0	6,52E+05	1,17E+05
9	'eh'	0.080000	0	6,61E+05	9,67E+04
10	'eh'	0.090000	0	4,40E+05	9,90E+04
11	'eh'	0.100000	0	4,38E+05	3,60E+04
12	'eh'	0.110000	0	4,46E+05	5,88E+04
13	'eh'	0.120000	0	6,73E+05	3,09E+04
14	'eh'	0.130000	0	6,80E+05	9,24E+04
15	'eh'	0.140000	0	6,78E+05	8,62E+04

Gambar 3. 3 hasil ekstraksi fitur salah satu suara “Eh”

Dari hasil ekstraksi fitur dapat dilihat hasil yang dikeluarkan adalah *Name*, *Frame Time*, *Fundamental Frequency (f_0)*, *Voicing Probability* dan *PCM Loudness*. Bila ditampilkan dalam bentuk grafik maka akan terlihat perbedaan setiap suara. Berikut gambar hasil *Fundamental Frequency (f_0)* dari salah satu suara “Eairh” dan “Eh”.

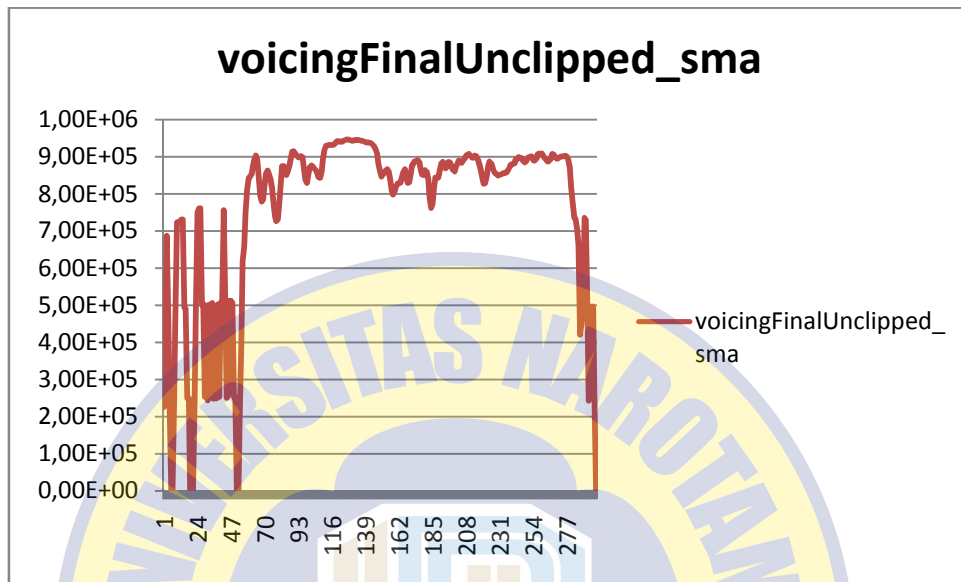


Gambar 3. 4 Grafik Fundamental Frequency dari salah satu suara “Eairh”.

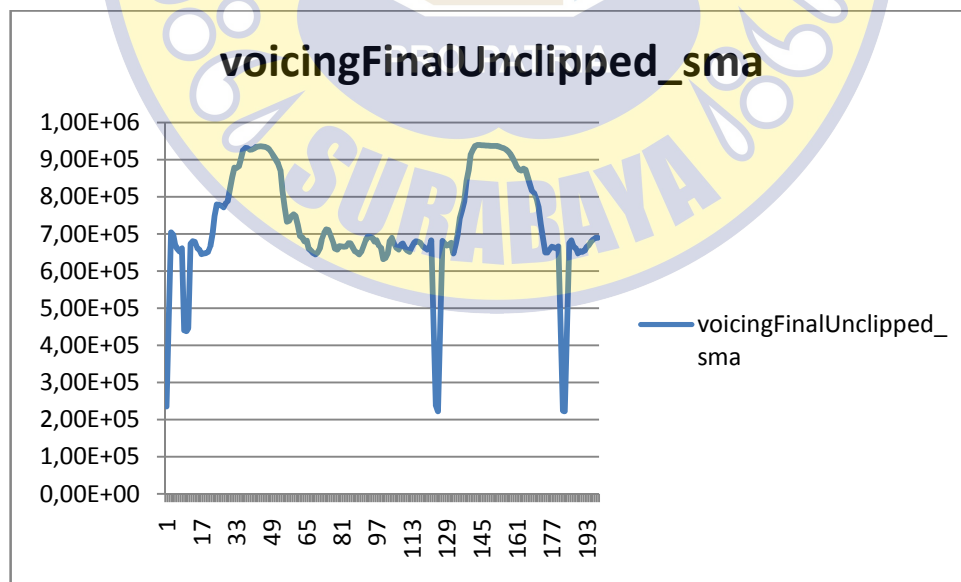


Gambar 3. 5 Grafik Fundamental Frequency dari salah satu suara “Eh”.

Berikut gambar hasil *Voicing Probability* dari salah satu suara “Eairh” dan suara “Eh”.

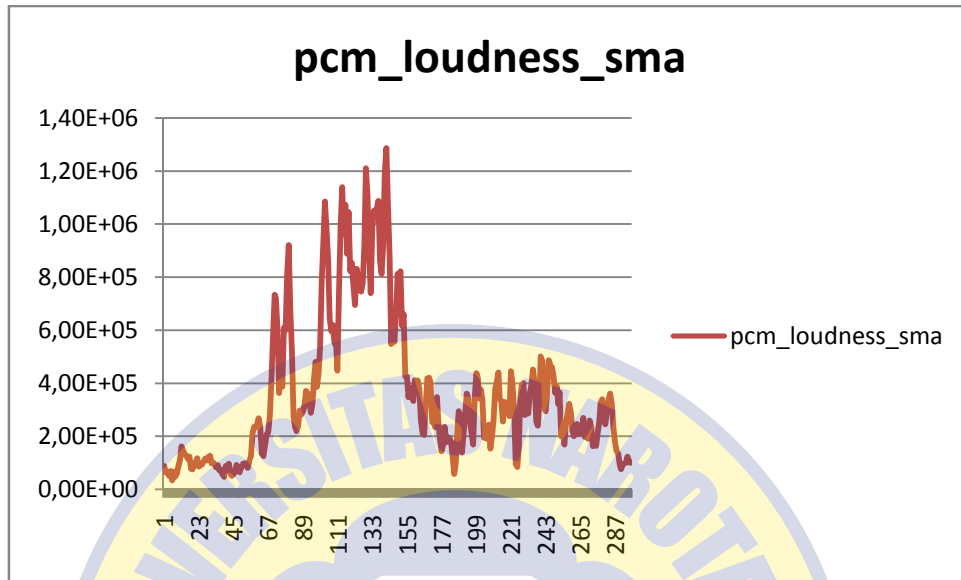


Gambar 3. 6 Grafik Voicing Probability dari salah satu suara “Eairh”

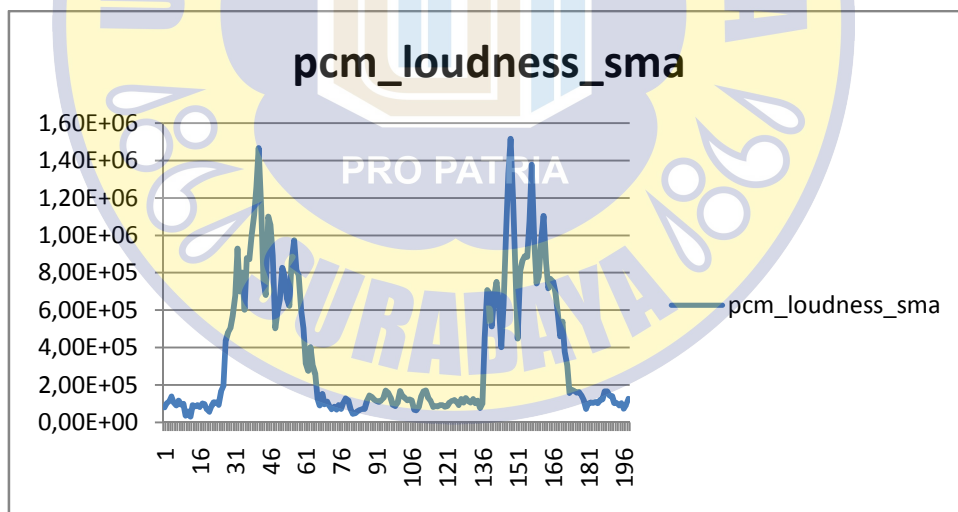


Gambar 3. 7 Grafik Voicing Probability dari salah satu suara “Eh”

Berikut gambar hasil *PCM Loudness* dari salah satu suara “Eairh” dan suara “Eh”.



Gambar 3. 8 Grafik PCM Loudness dari salah satu suara “Eairh”.



Gambar 3. 9 Grafik PCM Loudness dari salah satu suara “Eh”.

3.3. Teknik Pengumpulan Data dan Pengembangan Instrumen

Pada penelitian ini model data suara tangisan bayi merujuk pada model data *benchmark* milik *Dunstan Baby Language*. Priscillia Dunstan memaparkan bahwa

5 kategori *Dunstan Baby Language* yang ditelitinya selama 8 tahun dapat mewakili kondisi bayi. Oleh karena itu, penelitian ini menggunakan 5 kategori tersebut agar dapat memvalidasi kesesuaian antara penelitian Priscillia Dunstan dengan hasil komputasi melalui pendekatan *Speech Identification* dan *Speech Recognition*.

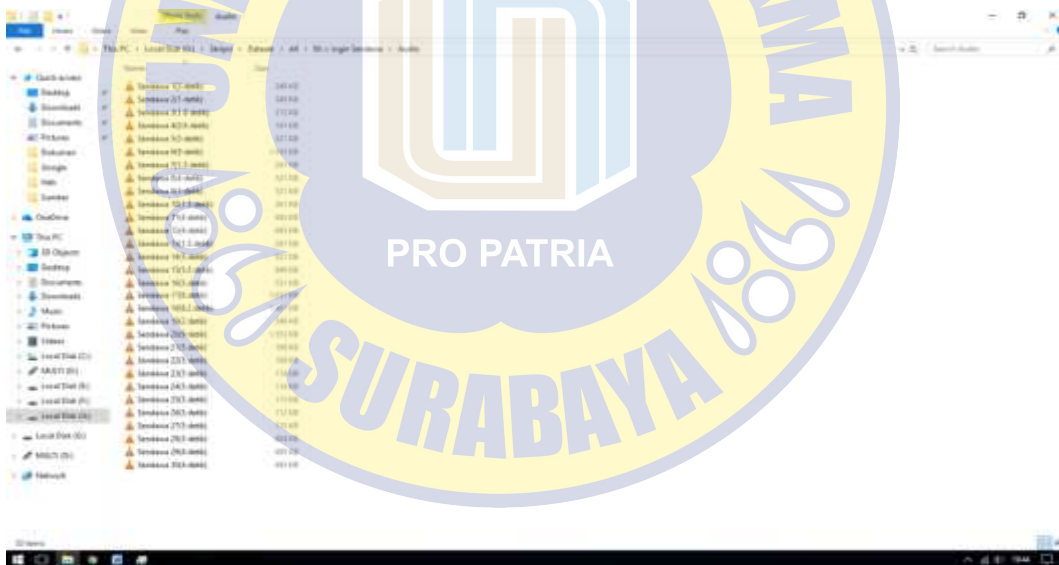
Dalam proses pengumpulan data didapatkan berbagai sampel data suara tangisan bayi dengan batasan bayi berumur 0-3 bulan sebagai berikut.

1. 5 sampel x 5 kategori (Cuplikan video tangisan bayi yang diputar pada acara My Mum Oprah, Nara Sumber: Priscillia Dunstan), model data ini dijadikan sebagai rujukan utama.
2. CD Dunstan Baby Language bayi (berjumlah 102 file).
3. Dataset dari akun github (berjumlah 8 data).
4. Dataset dari research.google.com (berjumlah 15 data).

Data dari CD *Dunstan Baby Language* berupa file video dengan format .mkv, video tersebut dipotong sesuai kebutuhan yaitu ketika ada contoh suara bayi lapar, buang angin, sendawa, tidak nyaman dan lelah. Data suara bayi dari youtube berupa video mp4, video tersebut dipotong sesuai kebutuhan yaitu ketika ada contoh suara bayi lapar, buang angin, sendawa, tidak nyaman dan lelah. Total data suara tangisan bayi yang telah terkumpul ada 150 file dalam format .wav. Berikut data audio yang telah di edit untuk di olah.



Gambar 3. 10 Hasil audio suara “Eairh” dari CD Dunstan Baby Language



Gambar 3. 11 Hasil audio suara “Eh” dari CD Dunstan Baby Language



Gambar 3. 12 Hasil audio suara “Heh” dari CD Dunstan Baby Language



Gambar 3. 13 Hasil audio suara “Neh” dari CD Dunstan Baby Language



Gambar 3. 14 Hasil audio suara “Owh” dari CD Dunstan Baby Language

3.4. Teknik Analisis Data

Teknik Analisis data yang digunakan pada penelitian ini, yaitu ekstraksi fitur *Prosodic Feature*, mencari pola sinyal suara *Moment of Distribution* dan algoritma *K-Nearest Neighbour*.

3.4.1. *Speech Identification*

Sebelum memasuki tahap *Speech Recognition*, tahap yang harus dilakukan adalah *Speech Identification*. *Speech Identification* meliputi ekstraksi fitur suara *Prosodic Feature*, mencari pola sinyal berdasarkan distribusinya *Moment of Distribution*.

1. Ekstraksi Fitur Suara *Prosodic Feature*

Dalam prosodik ada tiga karakteristik utama yaitu, *fundamental frequency*(f_0), *frame range*, dan loudness yang diambil dari sinyal suara.

Sebagaimana dibahas dalam banyak jurnal, *Prosodic Feature*, terkenal untuk memberikan informasi yang berguna tentang gaya bicara seseorang, dan dengan demikian, banyak digunakan dalam *Speech Recognition*[5]. Fitur *pitch* adalah menangkap level *fundamental frequency*(f_0), sedangkan fitur energi mencerminkan pola naik dan turunnya level *loudness*. Untuk fitur durasi adalah durasi setiap milidetik yang didapatkan dari *pitch* dan energi.

2. Mencari Pola Sinyal *Moment of Distribution*

Setelah ekstraksi fitur dilakukan, hasil yang didapatkan masih belum maksimal. Banyaknya baris data yang diekstrak berdasarkan *frame time* memungkinkan banyak kesalahan yang terjadi/error pada saat proses klasifikasi. Karena tingkat keunikan dari kategori suara tangisan bayi menurut *Dunstan Baby Language* belum terdeteksi. Maka untuk mencari pola sinyal suara berdasarkan distribusinya dari setiap suara tangisan bayi dapat menggunakan *Moment of Distribution*, Sehingga 1 suara tangisan bayi dapat terwakili dalam 1 sampel data saja.

3.4.2. *Speech Recognition*

Dalam penelitian kali ini pendekatan yang digunakan adalah pengenalan pola. Pengenalan pola yang akan digunakan adalah metode klasifikasi menggunakan *K-Nearest Neighbour*.

1. Metode Klasifikasi *K-Nearest Neighbour*

Pada tahap klasifikasi menggunakan algoritma *K-Nearest Neighbours* nilai *K* yang akan digunakan adalah 1,3,5,7 dan 9. Selanjutnya tiap data yang diklasifikasikan pada setiap nilai *K* yang berbeda akan diuji menggunakan data sampling *Percentage Rate* dan *Leave One Out (LOO)*. Data sampling *Percentage Rate* yang akan diklasifikasikan yaitu 70:30, 80:20 dan 90:10. Hasil dari tiap data sampling dan nilai *K* akan disajikan dalam tabel berisi akurasi. Lalu untuk menentukan mana klasifikasi terbaik berdasarkan nilai *K* maka hasil klasifikasi setiap nilai *K* akan di hitung rata-ratanya. Sama seperti sebelumnya untuk menentukan mana klasifikasi terbaik berdasarkan data sampling maka hasil klasifikasi setiap data sampling akan dihitung rata-ratanya

2. Metode *K-Means Clustering*

Metode ini dipilih sebagai alternatif, untuk melihat bagaimana data tangisan suara tersebut dikelompokkan secara *natural* oleh *K-means*. Metode ini dapat mengelompokkan data berdasarkan kemiripan dari karakteristik data dengan cara menentukan di awal dahulu berapa kelompok data yang ingin dibentuk. Dalam penelitian ini ditentukan kelompok data sebanyak 5 sesuai dengan kategori *Dunstan Baby Language*. Sehingga akan terbentuk kelompok atau *cluster* data yang belum diketahui nama-nama kategori datanya. Diterapkannya metode *K-means* pada penelitian ini untuk mengetahui seberapa cocok label *cluster* dengan label aktual pada data tangisan suara bayi tersebut. Biasanya,

kategori yang dihasilkan oleh label *cluster* milik K-means akan memiliki performa akurasi yang baik dibanding label aktualnya. Namun kekurangannya adalah dengan digunakannya label *cluster* sebagai model data, maka akan menyalahi aturan data label aktualnya yang sudah ditentukan di awal. Oleh karena itu, tujuan diterapkannya K-means clustering ini hanya untuk mengetahui seberapa cocok hasil label *cluster* milik K-means dengan label aktual pada data tangisan suara bayi.

3.4.3. Mengukur Kinerja Klasifikasi Menggunakan *Confusion Matrix*

Setelah mendapatkan hasil klasifikasi data baru terhadap data latih selanjutnya menentukan akurasi menggunakan *Confusion matrix* digambarkan dengan tabel yang menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan.