

Klasifikasi Suara Tangisan Bayi Berdasarkan Prosodic Features Menggunakan Metode Moments of Distribution dan K-Nearest Neighbours

Aditya Singgi Prayogi

Program Studi Teknik Informatika
Universitas Narotama
adityasinggiprayogi.15@fasilkom.na
rotama.ac.id

Maulana Rizqi

Program Studi Sistem Komputer
Universitas Narotama
maulana.rizqi@narotama.ac.id

Tresna Maulana Fahrudin

Program Studi Teknik Informatika
Universitas Narotama
tresna.maulana@narotama.ac.id

Abstrak – Bagi sebagian orang, suara tangisan bayi terdengar sangat mengganggu, apalagi jika tangisannya berlarut-larut. Sulit untuk dimengerti arti dari suara tangisan bayi. Di era teknologi informasi, pengenalan suara tangisan bayi dapat dilakukan secara otomatis menggunakan komputer. Hal tersebut tentu dapat membantu bagi orang tua untuk mengenali kebutuhan bayi agar dapat segera tenang. Untuk mengidentifikasi suara tangisan bayi dapat menggunakan salah satu algoritma klasifikasi di bidang Machine Learning, salah satunya adalah algoritma *K-Nearest Neighbour*. Langkah pertama untuk melakukan klasifikasi suara tangisan bayi, yakni data audio suara tangisan bayi diubah menjadi data numerik yang disebut proses ekstraksi fitur yang menghasilkan *Prosodic Features*. Setelah melewati proses ekstraksi fitur perlu dilakukan identifikasi pola untuk mendapatkan perbedaan pola antara satu data suara tangisan bayi dengan data suara tangisan bayi yang lain menggunakan Metode *Moment of Distribution*. Pengenalan suara tangisan bayi dilakukan dengan menerapkan algoritma klasifikasi menggunakan *K-Nearest Neighbour*. Akurasi terbaik pada proses klasifikasi menggunakan *data sampling Percentage Rate* yaitu 76% dimana nilai *K* yang digunakan adalah 9. Sedangkan akurasi terbaik pada proses klasifikasi menggunakan *data sampling Leave One Out* yaitu 42% dengan nilai *K* yang digunakan adalah 5.

Kata Kunci: *K-Nearest Neighbour*, *Moments of Distribution*, *Prosodic Feature*, *Suara Tangisan Bayi*, *Validation Sampling*

I. PENDAHULUAN

Suara tangisan bayi adalah salah satu cara berkomunikasi untuk menyampaikan suatu keadaan yang dialaminya pada saat menangis. Bayi dapat mengeluarkan suara tangisan yang berbeda tergantung dengan kondisinya. Pada tahun 2006 suara tangisan bayi telah diteliti oleh Priscillia Dunstan yang sekarang telah diakui oleh dunia dengan nama *Dunstan Baby Language*[1]. Ada 5 suara bayi yang telah diteliti sampai sekarang, yaitu lapar, lelah atau mengantuk, ingin sendawa, ingin buang angin dan sakit. Suara tangisan yang dikeluarkan oleh bayi antara lain "Neh" yang berarti Lapar, "Eairh" yang berarti ingin buang angin, "Eh" yang berarti ingin sendawa, "Heh" yang berarti bayi mengalami

kondisi tidak nyaman dan "Owh" yang berarti bayi mengantuk. Bagi orang tua yang mengasuh bayi tentu sulit untuk menentukan apa yang diinginkan oleh bayi, karena suara tangisan yang dikeluarkan hampir sama. Padahal ada perbedaan arti tertentu dari suara tangisan bayi tersebut. Pada penelitian ini digunakan 5 kategori sebagai label klasifikasi suara tangisan bayi. Kategori suara tangisan bayi yang diteliti didapatkan dari CD *Dunstan Baby Language* yaitu suara bayi berumur 0-3 bulan. Pada penelitian ini bertujuan untuk membuat model klasifikasi arti suara tangisan bayi dan mengevaluasi seberapa tinggi tingkat akurasi metode yang telah diterapkan pada suara tangisan bayi.

Di Indonesia ada beberapa penelitian yang meneliti suara tangisan bayi seperti penelitian milik Irma Amelia Dewi, dkk yang berjudul "Identifikasi Suara Tangisan Bayi Menggunakan Metode *Linear Predictive Coding* (LPC) dan *Euclidean Distance*" [2] dan Welly Setiawan Limantoro, dkk yang berjudul "Rancang Bangun Aplikasi Pendeteksi Tangisan Suara Bayi" [3], mereka menggunakan *Dunstan Baby Language* sebagai acuan untuk label klasifikasi pada penelitian suara tangisan bayi.

Pada penelitian ini menguji kecocokan label suara tangisan bayi berdasarkan label *Dunstan Baby Language* dengan menggunakan 150 suara tangisan bayi dan mencocokkannya berdasarkan label prediksi yang dihasilkan oleh *Machine Learning*. 150 suara tangisan bayi tersebut melewati proses *preprocessing* menjadi 5 versi. Versi 1 yaitu hasil potongan audio dari video CD *Dunstan Baby Language* yang memiliki durasi antara 0,8 detik – 14 detik. Versi 2 yaitu file suara dari versi 1 yang durasinya hanya diambil selama 1 detik. Versi 3 yaitu file suara dari video 2 yang volumenya disamaratakan atau disetarakan dengan durasi 1 detik. Versi 4 yaitu file suara yang diambil hanya saat suara pertama kali keluar dengan durasi hanya 0,2 – 1 detik. Versi 5 yaitu file suara dari versi 1 yang diambil saat suara pertama kali keluar dengan durasi 0,2 – 1 detik.

Metode *Linear Predictive Coding* (LPC) dan *Mel-Frequency Cepstral Coefficients* (MFCC) adalah metode

yang banyak digunakan pada penelitian ekstraksi fitur suara [4]. Sehingga pada penelitian ini digunakan metode yang berbeda yakni menggunakan *Prosodic Features* yang menghasilkan 3 parameter yaitu *pitch*, durasi dan *energy* [5]. Metode ini cukup mewakili untuk menghasilkan parameter dari hasil ekstraksi fitur. Selanjutnya untuk mengidentifikasi pola dari setiap sampel suara tangisan bayi berdasarkan nilai rata-rata, varian, standar deviasi, kemiringan dan kecondongan, maka diterapkan metode *Moments of Distribution*. Lalu dilanjutkan pengenalan suara menggunakan algoritma klasifikasi menggunakan metode *K-Nearest Neighbour* [6] serta menghitung tingkat akurasi prediksi yakni dengan cara mencocokkan antara label aktual suara tangisan bayi dengan label prediksinya yang dihasilkan oleh algoritma klasifikasi.

II. METODOLOGI PENELITIAN

Metode yang digunakan dalam penelitian ini terdiri dari *preprocessing* suara, yaitu ekstraksi fitur menggunakan *Prosodic Features*, lalu untuk mengidentifikasi pola dari suatu sampel suara tangisan bayi menggunakan *Moments of Distribution* dan menerapkan algoritma klasifikasi suara tangisan bayi menggunakan metode *K-Nearest Neighbour* dimana *validation sampling* yang digunakan adalah *Percentage Rate* dan *Leave One Out* [7].

A. Subjek Penelitian

Ruang lingkup penelitian yang adalah bagaimana mengidentifikasi pola suara tangisan bayi yang berbeda di setiap label dan bagaimana cara sistem dapat mengenali label suara tangisan bayi berdasarkan data uji yang diberikan. Sampel yang digunakan pada penelitian ini adalah sebanyak 5 versi suara dari CD *Dunstan Baby Language*. Tiap versi berekstensi *.wav dan berjumlah 122-150 tiap versinya. Setiap sampel nantinya akan diproses menggunakan metode *Prosodic Feature* untuk mengetahui nilai *fundamental frequency (f0)*, *frame time*, *loudness*. Data ini selanjutnya diproses menggunakan metode *Moment of Distribution* agar mudah diidentifikasi pola yang berbeda tiap sampelnya. Selanjutnya pada proses pembagian data sampel, data akan dibagi menjadi data latih dan data uji dalam suatu porsi data, dan diterapkan metode *K-Nearest Neighbour* untuk mengetahui tingkat akurasi pengenalan suara tangisan bayi berdasarkan kecocokan label data aktual dan label data prediksinya.

B. Metode Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan cara, yaitu:

1. Mendapatkan file video suara bayi dari CD *Dunstan Baby Language* sebagai validasi data dan bukti eksistensi penelitian suara tangisan bayi serta digunakan sebagai acuan label suara tangisan bayi.
2. Studi pustaka yang dilakukan dengan membaca artikel ilmiah atau jurnal yang berkaitan dengan suara tangisan bayi maupun metode yang diusulkan dalam penelitian.

C. Alat dan Bahan

Spesifikasi perangkat keras yang digunakan dalam penelitian ini, yaitu:

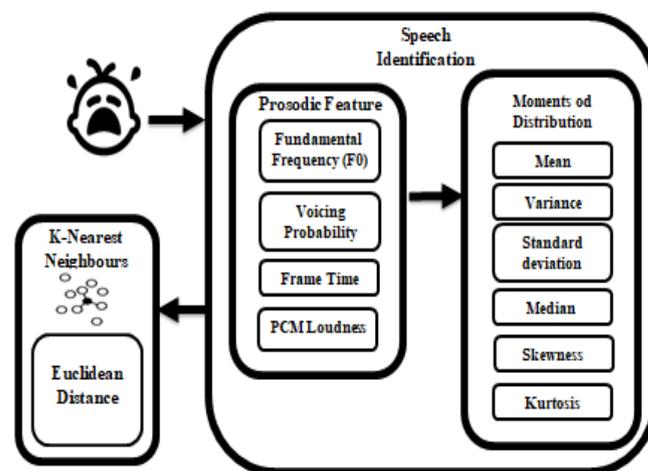
1. Prosesor Intel Core i5-8250U
2. Memori RAM DDR4 4GB
3. HDD 1TB

Perangkat lunak yang digunakan, yaitu:

1. Sistem Operasi Windows 10 Home
2. *Software* Jupyter dan Anaconda untuk *coding* Python
3. *Software* Excel untuk menyimpan data berekstensi *.csv dan *.xls
4. *Software* Format Factory untuk konversi video ke audio.

D. Desain Sistem

Desain sistem pada penelitian ini yang diilustrasikan pada Gambar 1 menjelaskan langkah apa saja yang dilakukan pada saat penelitian dimulai dari mengkoleksi file suara tangisan bayi, lalu dilakukan ekstraksi fitur menggunakan metode *Prosodic Features*. Kemudian mengidentifikasi pola suara tangisan bayi menggunakan metode *Moment of Distribution*, lalu menerapkan algoritma *K-Nearest Neighbours* untuk mengetahui akurasi pada model klasifikasi suara tangisan bayi yang telah dibangun.



Gambar 1. Desain Sistem Penelitian

E. Preprocessing

Data video dari CD *Dunstan Baby Language* dipotong menjadi 150 file dan formatnya diubah menjadi file berekstensi *.wav. 150 file tersebut diubah menjadi 5 versi suara tangisan bayi yang berbeda, dan hal ini merupakan tahapan demi tahapan dalam fase *preprocessing*. Untuk mendapatkan suara tangisan bayi yang memiliki panjang dan kualitas suara yang baik, maka diperlukan tahap *preprocessing* yang mendalam.

Flowchart dari tahap *preprocessing* dari file suara versi 1 yaitu ditunjukkan pada Gambar 2. Suara yang dihasilkan berasal dari hasil potongan video CD *Dunstan Baby Language*. Video yang berdurasi 24 menit tersebut dipotong menjadi 150 file. Kemudian potongan video tersebut dikonversikan menjadi file audio berekstensi *.wav.



Gambar 2. Tahapan *Preprocessing* Data Versi 1

Suara yang dihasilkan dari data versi 2 berasal dari suara versi 1. Pada data versi 2 semua file audio durasinya disamaratakan 1 detik. *Preprocessing* dari file suara versi 2 ditunjukkan pada Gambar 3.



Gambar 3. Tahapan *Preprocessing* Data Versi 2

Suara yang dihasilkan dari data versi 3 berasal dari versi 2. File audio yang berdurasi 1 detik dari versi 2 tiap sampelnya diubah dan ditingkatkan level audionya. Tiap sampel diubah level audionya sama seperti sampel pertama. *Preprocessing* dari file suara versi 3 ditunjukkan pada Gambar 4.



Gambar 4. Tahapan *Preprocessing* Data Versi 3

Suara data versi 4 berasal dari data suara versi 3. Setiap sampel dipilih berdasarkan suara yang paling jelas menurut kategori *Dunstan Baby Language*. Data suara versi 4 berjumlah 123 sampel dengan durasi adalah 0,2-1 detik. *Preprocessing* dari file suara versi 4 ditunjukkan pada Gambar 5.



Gambar 5. Tahapan *Preprocessing* Data Versi 4

Suara data versi 5 berasal dari data suara versi 1. Suara data versi 5 berbeda dengan alur data versi 1-4 dimana suara versi 1-4 digunakan untuk menentukan perbedaan tiap tahapnya. Sedangkan suara data versi 5 langsung menggunakan suara versi 1, lalu dikenali suara tangisan bayi mana yang terdengar jelas dan disamaratakan durasinya.



Gambar 6. Tahapan *Preprocessing* Data Versi 5.

Data versi 1 digunakan sebagai data rujukan utama karena data tersebut merupakan *raw data*. Pembuatan data suara versi 5 adalah filter setiap sampel yang dipilih berdasarkan suara yang paling jelas menurut kategori *Dunstan Baby Language*. Data suara versi 5 memiliki jumlah sampel data sebanyak 122 sampel. Durasi suaranya adalah 0,2-1 detik. *Preprocessing* dari file suara versi 5 ditunjukkan pada Gambar 6.

Tujuan dilakukan *preprocessing* data pada penelitian ini sehingga menghasilkan 5 versi data adalah rangkaian tahap perbaikan-perbaikan terhadap data suara tangisan bayi agar mencapai data bersih dan berkualitas, dimana sebelumnya data tersebut masih belum seragam dari sisi frekuensi suara, panjang pendeknya *frame time* dan *noise*.

F. Ekstraksi Fitur Berdasarkan Prosodic Features

Dalam *Prosodic Feature* ada tiga karakteristik utama yaitu, *fundamental frequency*(f_0), *frame range*, dan *loudness* yang diambil dari sinyal suara. Sebagaimana yang dibahas dalam banyak jurnal, *Prosodic Feature*, terkenal untuk memberikan informasi yang berguna tentang gaya bicara seseorang, dan dengan demikian, banyak digunakan dalam *Speech Recognition*. Fitur *pitch*, menangkap level *fundamental frequency*(f_0), sedangkan fitur energi mencerminkan pola naik dan turunnya level *loudness*. Fitur durasi (*frame time*) merepresentasikan setiap milidetik yang didapatkan dari *pitch* dan *loudness*.

G. Identifikasi Pola Sinyal Berdasarkan Metode *Moment of Distribution*

Setelah ekstraksi fitur dilakukan, hasil yang didapatkan masih belum maksimal. Banyaknya data yang diekstrak memungkinkan banyak kesalahan yang terjadi atau eror pada saat proses klasifikasi. Karena tingkat keunikan dari kategori suara tangisan bayi menurut *Dunstan Baby Language* belum terdeteksi. Maka untuk mencari dan mengidentifikasi pola sinyal suara berdasarkan distribusinya dari setiap suara

tangisan bayi dapat menggunakan *Moment of Distribution*. Rumus dari metode *Moment Of Distribution* dapat dilihat pada persamaan 1 sampai 5 berikut ini.

1. Mean

$$Mean = \frac{\sum Xi}{N} \tag{1}$$

Keterangan:

X_i = nilai data ke- i

N = banyaknya data

2. Median

$$Median = \frac{X_{(n+1)}}{2} \tag{2}$$

Keterangan:

X = urutan data

n = banyaknya data

3. Skewness

$$Skewness = \frac{X-M}{S} \tag{3}$$

Keterangan:

X = nilai rata-rata

M = median

S = standar deviasi

4. Kurtosis

$$Kurtosis = \frac{\sum_{i=1}^N \frac{(X_i - \bar{X})^2}{N}}{S^4} \tag{4}$$

Keterangan:

\bar{X} = nilai rata-rata

X_i = nilai data ke- i

S = standar deviasi

N = banyaknya data

5. Standar Deviasi

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}} \tag{5}$$

Keterangan:

s = standar deviasi

\bar{X} = nilai rata-rata

X_i = nilai data ke- i

n = banyaknya data

H. *Validation Sampling* Menggunakan *Percentage Rate* dan *Leave One Out*

Validation Sampling adalah teknik untuk menentukan bagaimana pembagian porsi antara data latih dan data uji.

Pada penelitian ini *validation sampling* yang digunakan adalah *Percentage Rate* dan *Leave One Out*. *Percentage Rate* adalah pengujian data dengan membagi porsi dari data latih dengan data uji menggunakan tingkat persentase. Pada penelitian ini akan digunakan perbandingan dengan rasio 70:30, 80:20 dan 90:10. Sedangkan *Leave One Out* (LOO) adalah pembagian porsi data dengan cara mengambil 1 sampel data sebagai data uji dari sekelompok sampel data, lalu mengujikannya ke dalam model data latih yang tersisa. Pengujian data menggunakan LOO dilakukan pada setiap individu data, sehingga individu data berperan menjadi data uji dan selanjutnya bergabung menjadi data latih dengan sekelompok data lainnya.

I. Klasifikasi dengan K-Nearest Neighbours

K-Nearest Neighbour merupakan algoritma *supervised learning* dimana hasil dari data yang baru diklasifikasikan terhadap objek yang jaraknya paling dekat dengan objek *K-Nearest* berdasarkan data pembelajaran. Tujuan dari algoritma ini adalah untuk mengklasifikasikan objek baru berdasarkan atribut dan sampel-sampel dari data latih. Rumus dari metode *K-Nearest Neighbour* dapat dilihat pada persamaan 6 dimana dihitung selisih suatu jarak antara individu data dengan individu data lainnya.

$$D(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{6}$$

III. HASIL DAN PEMBAHASAN

A. Hasil Preprocessing

Hasil *preprocessing* versi 1 sampai versi 5 berupa audio file akan diberikan pada Tabel 1.

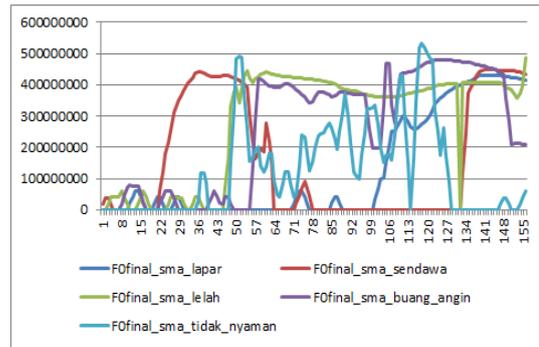
Tabel 1. Hasil *Preprocessing* Data Suara Tangisan Bayi

Versi	Jumlah
1	150
2	150
3	150
4	123
5	122

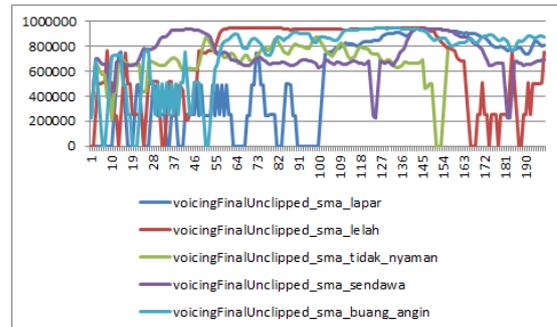
Data suara versi 1, 2, dan 3 berisi sampel sebanyak 150 sampel, data suara versi 4 berisi 123 sampel, dan data suara versi 5 berisi 122 sampel. Data suara versi 1 memiliki durasi antara 0,8 sampai 11 detik. Data suara versi 2 sampai 3 memiliki durasi 1 detik. Data suara versi 4 dan 5 memiliki durasi antara 0,2 sampai 0,8 detik.

B. Hasil Ekstraksi Fitur Prosodic Feature

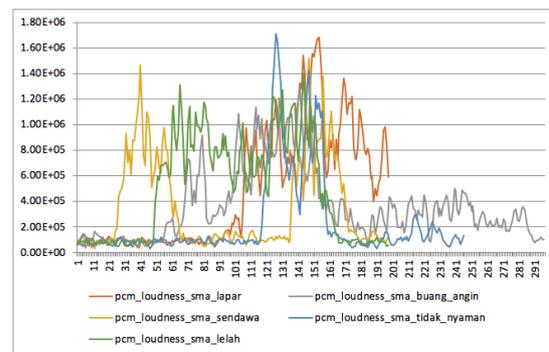
Hasil ekstraksi fitur akan ditunjukkan pada Gambar 7, 8, dan 9.



Gambar 7. Hasil Ekstraksi Fitur F0 Menggunakan Prosodic Features Pada Data Suara Versi 1



Gambar 8. Hasil Ekstraksi Fitur Voicing Menggunakan Prosodic Features Pada Data Suara Versi 1



Gambar 9. Hasil Ekstraksi Fitur Loudness Menggunakan Prosodic Features Pada Data Suara Versi 1

Hasil ekstraksi fitur *Prosodic Features* terdapat 4 parameter utama yang dapat menentukan pola dari suatu kategori suara. Empat parameter utama tersebut adalah *frameTime*, *F0final_sma*, *voicingFinalUnclipped_sma* dan terakhir *pcm_loudness_sma*. Empat parameter tersebut memberikan pola yang berbeda dari tiap kategori suara tangisan bayi yang akan membantu dalam proses klasifikasi.

C. Hasil Moment of Distribution

Hasil dari pengenalan pola menggunakan metode *Moment of Distribution* akan ditunjukkan pada Gambar 10 dan 11.

name	frameTime	F0final_sma	voicingFinalUnclipped_sma	pcm_loudness_sma
lapar	0.010000	0	2,26E+05	8,81E+04
lapar	0.020000	0	4,60E+05	6,34E+04
lapar	0.030000	0	6,87E+05	6,54E+04
lapar	0.040000	0	4,61E+05	5,04E+04
lapar	0.050000	0	2,27E+05	6,86E+04
lapar	0.060000	0	0	3,35E+04
lapar	0.070000	0	0	6,66E+04
lapar	0.080000	1,85E+07	2,40E+05	4,68E+04
lapar	0.090000	4,42E+07	4,77E+05	5,75E+04
lapar	0.100000	7,00E+07	7,24E+05	8,58E+04
lapar	0.110000	7,73E+07	7,24E+05	1,09E+05
lapar	0.120000	7,69E+07	7,24E+05	1,62E+05
lapar	0.130000	7,66E+07	7,31E+05	1,43E+05
lapar	0.140000	7,63E+07	7,31E+05	1,34E+05

Gambar 10. Hasil Ekstraksi Fitur Menggunakan Prosodic Features Pada Data Suara Versi 1 Kategori Lapar

F0final_sma_mean, F0final_sma_median, F0final_sma_std, F0final_sma_skewness, F0final_sma_kurtosis, voicingFinalUnclipped_sma_mean, voicingFinalUnclipped_sma_median, voicingFinalUnclipped_sma_std, voicingFinalUnclipped_sma_skewness, voicingFinalUnclipped_sma_kurtosis, pcm_loudness_sma_mean, pcm_loudness_sma_median, pcm_loudness_sma_std, pcm_loudness_sma_skewness, pcm_loudness_sma_kurtosis

Gambar 11. Hasil Identifikasi Pola Suara menggunakan Metode Moment of Distribution pada Data Suara Versi 1 Kategori Lapar

Empat parameter hasil ekstraksi fitur diproses ke dalam identifikasi pola suara tangisan bayi menggunakan metode *Moment of Distribution*, dimana hanya 3 parameter yang dibutuhkan, yaitu *F0final_sma*, *voicingFinalUnclipped_sma* dan *pcm_loudness_sma*. Selanjutnya 3 parameter tersebut diproses menggunakan metode *Moment of Distribution* sehingga menghasilkan 15 parameter baru, yaitu:

1. *f0final_sma_mean*,
2. *f0final_sma_median*,
3. *f0final_sma_std*,
4. *f0final_sma_skewness*,
5. *f0final_sma_kurtosis*,
6. *voicingFinalUnclipped_sma_mean*,
7. *voicingFinalUnclipped_sma_median*,
8. *voicingFinalUnclipped_sma_std*,
9. *voicingFinalUnclipped_sma_skewness*,
10. *voicingFinalUnclipped_sma_kurtosis*,
11. *pcm_loudness_sma_mean*,
12. *pcm_loudness_sma_median*,
13. *pcm_loudness_sma_std*,
14. *pcm_loudness_sma_skewness*, dan
15. *pcm_loudness_sma_kurtosis*.

D. Hasil Klasifikasi K-Nearest Neighbours

Pada penelitian ini teknik *validation sampling* yang digunakan yaitu *Percentage Rate* dan *Leave One Out (LOO)*.

Tabel 2. Hasil Klasifikasi Dengan Teknik Validation Sampling-Percentage Rate 90:10

Versi	Akurasi		
	K = 1	K = 3	K = 5
1	0,53	0,73	0,66

2	0,73	0,66	0,66
3	0,6	0,53	0,66
4	0,53	0,69	0,69
5	0,61	0,76	0,76

Tabel 2 menunjukkan pembagian data latih banding data uji 90:10 menghasilkan performa K=1 → data versi 2 lebih unggul dengan akurasi 73%, K=3 → data versi 5 lebih unggul dengan akurasi 76%, K=5 → data versi 5 lebih unggul dengan akurasi 76%.

Tabel 3. Hasil Klasifikasi Dengan Teknik Validation Sampling-Percentage Rate 80:20

Versi	Akurasi		
	K = 1	K = 3	K = 5
1	0,46	0,6	0,56
2	0,56	0,6	0,6
3	0,43	0,4	0,5
4	0,52	0,56	0,6
5	0,52	0,6	0,64

Tabel 3 menunjukkan pembagian data latih banding data uji 80:20 menghasilkan performa K=1 → data versi 2 lebih unggul dengan akurasi 56%, K=3 → data versi 1, 2 dan 5 lebih unggul dengan akurasi 60%, K=5 → data versi 5 lebih unggul dengan akurasi 64%.

Tabel 4. Hasil Klasifikasi Dengan Teknik Validation Sampling-Percentage Rate 70:30

Versi	Akurasi		
	K = 1	K = 3	K = 5
1	0,42	0,51	0,53
2	0,511	0,55	0,62
3	0,4	0,42	0,48
4	0,48	0,48	0,51
5	0,513	0,54	0,56

Tabel 4 menunjukkan pembagian data latih banding data uji 70:30 menunjukkan performa K=1 → data versi 5 lebih unggul dengan akurasi 51,3%, K=3 → data versi 2 lebih unggul dengan akurasi 55%, K=5 → data versi 2 lebih unggul dengan akurasi 62%.

Tabel 5. Hasil Klasifikasi Dengan Teknik Validation Sampling-Leave One Out (LOO)

Versi	Akurasi				
	K=1	K=3	K=5	K=7	K=9
1	26%	34,6%	42,6%	34%	30,6%
2	38,6%	35%	39%	36,6%	38,6%
3	28%	26%	31,3%	32%	34,6%
4	30,6%	31,3%	31,4%	30,4%	32,5%
5	28,9%	23%	38,9%	36,2%	34,5%

Tabel 5 menunjukkan *Leave One Out (LOO)* dengan akurasi terbaik didapatkan pada data versi 5 dengan akurasi 35%.

Dari 5 versi data suara tangisan bayi didapatkan hasil akurasi terbaik menggunakan teknik *Validation Sampling Percentage* yaitu 76% dimana nilai $K = 5$ pada data suara versi 5 menggunakan teknik *Validation Sampling Percentage Rate* 90:10. Sedangkan akurasi terburuk adalah 40% dimana nilai $K = 3$ pada data suara versi 3 menggunakan teknik *Validation Sampling Percentage Rate* 80:20 dan juga akurasi terburuk mencapai 40% dimana nilai $K = 1$ pada data suara versi 3 menggunakan teknik *Validation Sampling Percentage Rate* 70:30. Sedangkan dari 5 versi data suara tangisan bayi didapatkan hasil akurasi terbaik menggunakan teknik *Leave One Out* yaitu 42,6% dimana nilai $K = 5$ pada data suara versi 1.

Setelah melakukan berbagai percobaan, didapatkan beberapa hasil evaluasi yang perlu dijabarkan. Salah satunya, pertama, diperlukan data suara tangisan bayi yang konsisten untuk mendapatkan hasil yang maksimal, terkadang suara tangisan bayi pada sampel data terdengar mirip antara suara "eh" dan "heh" padahal terdapat arti yang berbeda bagi kondisi bayi. Kedua, 150 sampel data suara tangisan bayi masih dirasa kurang mewakili sebagai sampel data yang beragam untuk membuat model *machine learning*, perlu penambahan sampel data. Selain itu, kesulitan yang dihadapi adalah ketika mengolah data dan banyaknya nilai (*F0final_sma*, *voicingFinalUncilpped_sma* dan *pcm_loudness_sma*) pada tiap data suara tangisan bayi sebaran nilainya hamper mendekati. Nilai-nilai pada tiap data sangat mempengaruhi perhitungan jarak pada algoritma *K-Nearest Neighbours* yang berdampak pada prediksi kedekatan label.

Sebab sulit bagi algoritma *K-Nearest Neighbours* untuk membedakan karakteristik antara label. Untuk itu pada tahap *preprocessing* digunakan metode *Moment of Distribution* untuk mendapatkan pola yang berbeda pada tiap data. Secara keseluruhan, tahap *preprocessing* sangat menentukan kualitas dari data, yang mana merupakan faktor penting dalam keberhasilan klasifikasi. Sebab kualitas data yang rendah dapat berakibat pada kemampuan sistem dalam memprediksi dan pencapaian akurasi yang rendah.

IV. KESIMPULAN

Setelah melakukan penelitian terhadap suara tangisan bayi dapat ditarik beberapa kesimpulan. Pertama yakni pengujian menggunakan teknik *validate sampling Percentage Rate* akan mencapai akurasi yang lebih baik jika nilai K lebih tinggi dan porsi data latih lebih besar. Kedua yakni pengujian data suara tangisan bayi menggunakan data sampling *Leave One Out* menyatakan bahwa model klasifikasi data masih buruk karena akurasinya di bawah 50%. Ketiga yakni hasil klasifikasi menggunakan 2 teknik *validate sampling* penggunaan metode *Moment of Distribution* belum dirasa maksimal karena belum dapat menemukan pola berbeda dari hasil ekstraksi fitur *Prosodic Features* dimana seharusnya metode *Moment of Distribution*

dapat diperoleh nilai karakteristik persebaran data antar kategori suara tangisan bayi.

Pada penelitian selanjutnya diharapkan dapat membuat sebuah model dimana metode pengenalan pola yang digunakan dapat membuat nilai karakteristik persebaran data secara maksimal antar kategori. Penelitian ini adalah masih dalam tahap pengujian, sehingga masih diperlukan perbaikan model klasifikasi agar dapat memprediksi data uji lebih baik lagi. Saran untuk penelitian selanjutnya adalah perlu melakukan *preprocessing* lebih mendalam terhadap pola-pola tertentu pada kategori masing-masing tangisan bayi, karena ada beberapa suara tangisan bayi yang mirip namun memiliki arti yang berbeda yakni 'eh' dan 'heh'. Salah satu caranya adalah mempartisi satu suara tangisan bayi menjadi beberapa bagian, lalu ditransformasi menjadi fitur atau kolom data, sebagai contoh satu sampel suara tangisan bayi dibagi menjadi 1/3 dari tiap frame time, lalu nilai kolom (*pitch*, *voice*, dan *loudness*) diletakkan sejajar (horizontal) sebagai fitur baru.

REFERENSI

- [1] Limantoro, W.S., Faticah, C., & Yuhana, L. (2016). Rancang Bangun Aplikasi Pendeteksi Suara Tangisan Bayi. *Jurnal Teknik ITS*, Vol. 5, No. 2, pp. A476-A481.
- [2] Dewi, I.A., Zulkarnain, A., Lestari A.A. (2018). Identifikasi Suara Tangisan Bayi menggunakan Metode LPC dan Euclidean Distance. *Jurnal Elkomika*, Vol 6, No. 1, pp. 153-164.
- [3] Manfredi, C., Bandini, A., Melino, D., Viellevoe, R., Kalenga, M., & Orlandi, S. (2018). Biomedical Signal Processing and Control Automated Detection and Classification of Basic Shapes of Newborn Cry Melody. *Biomedical Signal Processing and Control*, 45, pp. 174–181, <https://doi.org/10.1016/j.bspc.2018.05.033>.
- [4] Singh, N., Khan, R.A. & Shree, R. (2012). MFCC and Prosodic Feature Extraction Techniques: A Comparative Study. *International Journal of Computer Applications*, Vol. 54, No. 1, pp. 9–13. <https://doi.org/10.5120/8529-2061>
- [5] Huang, Z., Chen, L. & Harper, M. (2006). An Open Source Prosodic Feature Extraction Tool. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, pp. 2116–2121.
- [6] Sumarlin (2015). Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM. *Jurnal Sistem Informasi Bisnis*, Vol. 01, pp. 52–62.
- [7] Rahagiyanto, A., Basuki, A., & Sigit, R. (2017). Moment Invariant Features Extraction for Hand Gesture Recognition of Sign Language Based on SIBI. *EMITTER International Journal of Engineering Technology*, Vol. 5, No. 1, pp. 119–138.