

# 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 Teknik Informatika  
Universitas Narotama  
maulana.rizqi@narotama.ac.id

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

**Abstrak** – Bagi orang dewasa suara tangisan bayi terdengar sangat mengganggu, apalagi jika tangisannya berlarut-larut. Karena tidak ada yang mengerti arti dari suara tangisan bayi. Dijaman sekarang yang serba otomatis pengenalan suara tangisan bayi dapat dilakukan secara otomatis menggunakan komputer. Jika suara tangisan bayi dapat diartikan secara otomatis oleh komputer maka dapat membantu orang dewasa mengenali kebutuhan si bayi dan si bayi bisa tenang dengan mudah. Pengenalan suara tangisan bayi secara otomatis dapat dilakukan dengan menggunakan sistem perangkat lunak komputer. Sistem tersebut dapat mengidentifikasi suara tangisan bayi menggunakan algoritma klasifikasi. Algoritma yang dapat digunakan untuk melakukan klasifikasi adalah *K-Nearest Neighbour*. Untuk melakukan klasifikasi suara bayi perlu diubah menjadi data numerik yang dapat digunakan pada proses klasifikasi proses tersebut dinamakan ekstraksi fitur. Ekstraksi fitur yang digunakan pada penelitian ini adalah *Prosodic Features*. Setelah melewati proses ekstraksi fitur perlu dilakukan pengenalan pola untuk mendapatkan perbedaan pola antara satu data suara tangisan bayi dengan data suara tangisan bayi yang lain menggunakan Metode *Moment of Dsitribution*. 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 bayi berkomunikasi untuk menyampaikan suatu keadaan yang dialami oleh bayi pada saat bayi menangis. Bayi dapat mengeluarkan suara yang berbeda tergantung 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/mengantuk, ingin sendawa, ingin buang angin, dan yang terakhir sakit. Suara yang dikeluarkan oleh si bayi

yaitu "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 dewasa yang mengasuh bayi tentu sulit untuk menentukan suara tangisan bayi apa yang keluar karena suaranya hampir sama. Padahal ada perbedaan suara ketika bayi menangis. Pada penelitian ini akan 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 akan dibuat model untuk meneliti 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 peneltian suara tangisan bayi mereka.

Pada penelitian ini ditujukan untuk memvalidasi bahwa suara tangisan bayi milik Dunstan *Baby Language* dapat digunakan sebagai acuan label suara tangisan bayi dengan menggunakan 150 suara tangisan bayi yang didapatkan dari CD Dunstan *Baby Language*. 150 suara tangisan bayi tadi akan 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]. Maka kali ini akan digunakan metode *Prosodic Features* yang menghasilkan 3 parameter yaitu *pitch*, durasi, dan energi [5]. Metode ini telah digunakan pada beberapa penelitian tentang suara. Setelah didapatkan hasil dari ekstraksi fitur kemudian diklasifikasikan menggunakan metode *K-Nearest Neighbour* [6].

## II. METODOLOGI PENELITIAN

Metode yang digunakan dalam penelitian ini terdiri dari *preprocessing* suara, yaitu ekstraksi fitur menggunakan *Prosodic Features*, lalu untuk mendapatkan inti pola dari suatu sampel suara tangisan bayi menggunakan *Moments of Distribution* dan menerapkan 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

Penelitian yang dibahas adalah klasifikasi suara tangisan bayi berdasarkan *prosodic features* dan *K-Nearest Neighbours*. Sampel yang digunakan pada penelitian ini adalah 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 diperoleh suatu pola yang berbeda tiap sampelnya. Selanjutnya data ini digunakan untuk proses klasifikasi menggunakan metode *K-Nearest Neighbour*.

### B. Metode Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan cara, yaitu:

1. Mendapatkan *file* video suara bayi dari CD *Dunstan Baby Language* untuk digunakan sebagai validasi bahwa suara tangisan bayi milik *Dunstan Baby Language* dapat 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:

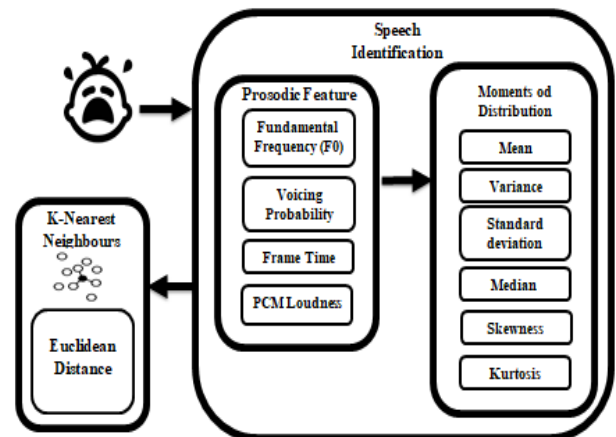
1. Prosesor Intel Core i5-8250U
2. Memori RAM DDR4 4GB
3. HDD 1TB

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 menceritakan langkah apa saja yang dilakukan pada saat penelitian mulai dari mendapatkan *file* suara tangisan bayi lalu diekstraksi fitur menggunakan metode *Prosodic Features* kemudian pengenalan pola menggunakan metode *Moment of Distribution* untuk mendapatkan persebaran data yang baik pada setiap kategori suara tangisan bayi, lalu terakhir menggunakan *K-Nearest Neighbours* untuk mendapatkan akurasi pada model suara tangisan bayi.



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* tadi akan diubah menjadi 5 versi. *Preprocessing* dari *file* suara versi 1 yaitu ditunjukkan pada Gambar 2.



Gambar 2. Tahapan *Preprocessing* Data Versi 1

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. *Preprocessing* dari file suara versi 2 ditunjukkan pada Gambar 3.



Gambar 3. Tahapan *Preprocessing* Data Versi 2

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 3 ditunjukkan pada Gambar 4.



Gambar 4. Tahapan *Preprocessing* Data Versi 3

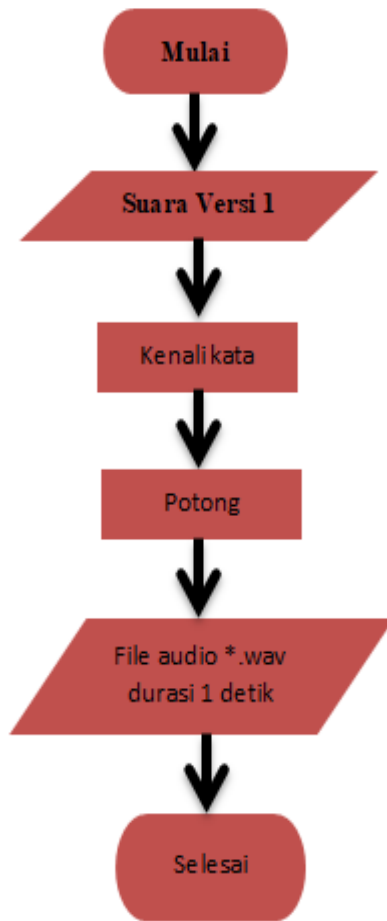
Suara yang dihasilkan dari data versi 3 berasal dari versi 2. File audio yang berdurasi 1 detik dari versi 2 tiap sampelnya diubah level audionya. Tiap sampel diubah level audionya sama seperti sampel pertama. *Preprocessing* dari file suara versi 4 ditunjukkan pada Gambar 5.



Gambar 5. Tahapan *Preprocessing* Data Versi 4

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 memiliki jumlah sampel data sebanyak 123 sampel. Durasi suaranya adalah 0,2-1 detik.

Suara data versi 5 berasal dari data suara versi 1. Suara data versi 5 berbeda dari alur versi 1-4 dimana suara versi 1-4 digunakan untuk menentukan perbedaan tiap tahapnya sedangkan suara data versi 5 langsung melalui tahap 1-4 menggunakan data suara versi 1 yaitu versi data suara yang pertama. Versi 1 digunakan sebagai data sumber karena datanya yang sama sekali belum bersih. Pembuatan data suara versi 5 memfilter setiap sampel 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.



Gambar 6. Tahapan *Preprocessing* Data Versi 5.

F. Ekstraksi Fitur *Prosodic Features*

Dalam *Prosodic Feature* ada tiga karakteristik utama yaitu, *fundamental frequency* (f0), *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*. Fitur *pitch* yang paling penting adalah menangkap level *fundamental frequency* (f0), sedangkan fitur energi yang paling penting mencerminkan pola naik dan turunnya level *loudness*. Untuk fitur durasi, yang paling penting adalah durasi setiap milidetik yang didapatkan dari *pitch* dan *loudness*.

G. Mencari Pola Sinyal *Moment of Distribution*

Setelah ekstraksi fitur dilakukan, hasil yang didapatkan masih belum maksimal. Banyaknya data yang diekstrak 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*. Rumus dari

metode *Moment Of Distribution* dapat dilihat pada persamaan 1 sampai 5.

1. *Mean*

$$Mean = \frac{\sum fi.xi}{\sum fi} \tag{1}$$

Keterangan:

$\bar{x}$  = rata-rata hitung dari data kelompok

$fi$  = frekuensi kelas ke-i

$xi$  = nilai tengah kelas ke-i

2. *Median*

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

Keterangan:

Me = *Median*

X = Urutan data

N = Banyaknya data

3. *Skewness*

$$SK = \frac{x-M}{s} \tag{3}$$

Keterangan:

SK = *Skewness*

X = Nilai rata-rata

M = *Median*

S = *Simpangan baku*

4. *Kurtosis*

$$a_4 = \frac{1}{n} \frac{(x-X)^4}{s^4} \tag{4}$$

Keterangan:

$a_4$  = *Kurtosis*

N = Jumlah data

(x-X) = Frekuensi

S = *Simpangan baku*

5. *Standard Deviation*

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

Keterangan:

s = *Standard deviation* (*simpangan baku*)

$x_i$  = Nilai x ke-i

n = Ukuran sampel

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

*Validation Sampling* adalah teknik untuk menentukan bagaimana data akan diuji. Pada penelitian ini *validation sampling* yang akan 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 pengujian data dengan mengambil 1 data latih dari semua data untuk diujikan dengan semua sisa dari data yang diambil, dimana sisa dari data adalah data latih. Pengujian data menggunakan LOO di lakukan pada setiap data, jadi tidak ada data yang tidak diuji.

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 adan sampel-sampel dari data latih. Rumus dari metode *K-Nearest Neighbour* dapat dilihat pada persamaan 6.

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

III. HASIL DAN PEMBAHASAN

A. Hasil *Preprocessing*

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

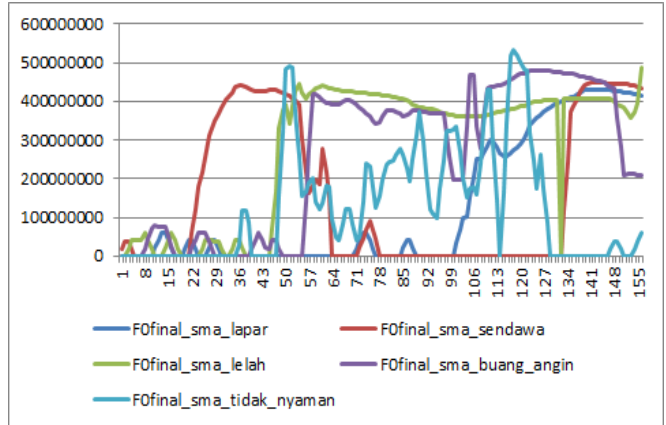
Tabel 1. Hasil *Preprocessing* Data Suara Tangisan Bayi

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

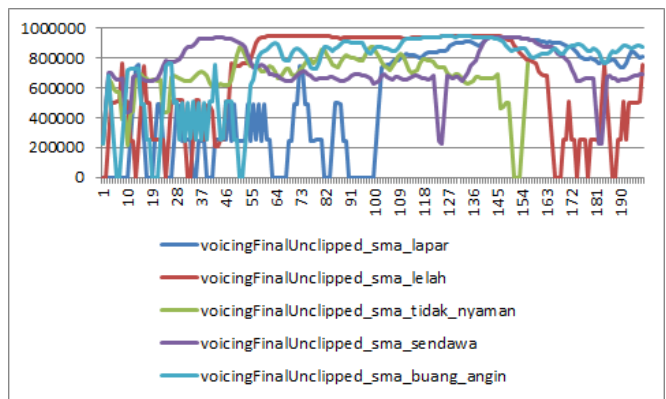
Data suara versi 1, 2, dan 3 berisi sampel sebanyak 150 sampel, data suara versi 4 berisi 123 sampel, dan data suara versi 5 berisi 122 sampel. Data suara versi 1 memiliki durasi antara 0,8 sampai 11 detik. Data suara versi 2 sampai 3 memiliki durasi 1 detik. Data suara versi 4 dan 5 memiliki durasi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

B. Hasil Ekstraksi Fitur *Prosodic Feature*

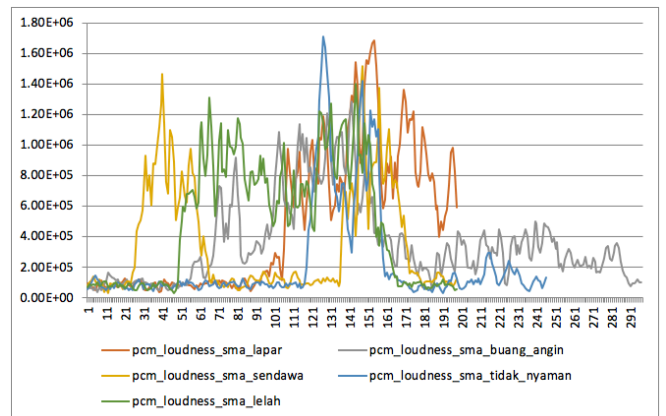
Hasil ekstraksi fitur akan ditunjukkan pada Gambar 7, 8, dan 9.



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* ada 4 parameter utama yang dapat menentukan pola dari suatu kategori suara. 4 parameter utama tersebut adalah frameTime, F0final\_sma, voicingFinalUnclipped\_sma dan terakhir pcm\_loudness\_sma. 4 parameter tadi 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\_std, voicingFinalUnclipped\_sma\_median, voicingFinalUnclipped\_sma\_skewness, voicingFinalUnclipped\_sma\_kurtosis, pcm\_loudness\_sma\_mean, pcm\_loudness\_sma\_std, pcm\_loudness\_sma\_median, pcm\_loudness\_sma\_skewness, pcm\_loudness\_sma\_kurtosis

293.5598154,373.0436,159.3381435,-0.008717967,-0.749386071,0.765233502,0.8627669,0.228125721,-1.894751767,2.67021638,0.360193316,0.2813234,0.279728247,1.2289.6453185,371.2539,180.6584942,-0.636680024,-1.113778173,0.789235058,0.8661138,0.197731632,-2.637784793,7.913546321,0.218757761,0.1696964,0.205981778,0.281.4642325,370.65795,150.8740388,-1.055026095,-0.397970279,0.761250962,0.88947085,0.285134569,-2.039031945,2.799322076,0.295991522,0.2984696,0.215115601,305.2617705,365.42635,154.0230125,-0.914250424,-0.346474311,0.826118147,0.89669495,0.178214005,-3.16076119,11.53603937,0.243147959,0.22437725,0.184291611,409.5241695,418.1655,45.91316566,-3.704839654,24.47530204,0.893374616,0.9216965,0.067532008,-4.463387547,37.9094229,0.302437183,0.290232,0.157513427,0.388.319.1934588,371.88175,148.2406863,-0.73906909,-0.498295932,0.816700843,0.8375461,0.133712075,-3.315120012,17.22000767,0.291523342,0.28261835,0.148868908,345.2044873,431.0321,190.0561364,-0.806392352,-0.883645342,0.838316591,0.8877962,0.114347503,-0.528384324,-0.88954151,0.229124399,0.2077847,0.19726573,0.9.287.240896,325.1784,153.1044613,-0.631974688,-0.442207692,0.781322944,0.8981124,0.254887745,-2.339911887,4.454996835,0.289165163,0.3048884,0.199762301,0.2.368.48315,415.4079,170.9864013,-1.131884791,0.312139867,0.70190668,0.7723112,0.253659479,-2.128668407,3.271983932,0.261185476,0.2533329,0.197481537,0.2415.386.8889525,411.4906,115.5869029,-1.371607667,2.09179091,0.890120772,0.929553,0.082297518,-2.138715447,6.017268061,0.321575852,0.3048191,0.193743693,0.491.335.7585781,418.3024,191.8816368,-0.911438542,-0.752241366,0.823985075,0.8941261,0.135399688,-1.045686189,1.4268837318,0.318683602,0.2207946,0.296059408,1.344.5556361,375.351,159.2039895,-1.30004028,0.378094006,0.796245157,0.88069775,0.208861505,-2.339254236,4.694912529,0.389654482,0.3308016,0.254826673,0.7.416.9660949,459.53075,129.4582002,-2.106991357,4.12687046,0.865089458,0.90919395,0.09986839,-1.651055064,4.036461706,0.341613807,0.2998016,0.267592054,1.1.328.0241745,383.7645,167.0242491,-1.136901874,-0.116787625,0.793726061,0.86666755,0.187005786,-2.127362787,5.464324411,0.245824609,0.24553605,0.202029644

Gambar 11. Hasil Pengenalan Pola Menggunakan Metode *Moment of Distribution* Pada Data Suara Versi 1 Kategori Lapar

Dari 4 parameter setelah ekstraksi fitur tadi yang akan dihitung untuk melewati proses pengenalan pola menggunakan metode *Moment of Distribution* hanya 3 parameter yang dibutuhkan, yaitu F0final\_sma, pcm\_loudness\_sma, dan pc\_loudness\_sma. Selanjutnya 3 parameter tiap kategori dan sampel suara tangisan bayi yang keluar setelah ekstraksi fitur akan dihitung tiap parameternya menggunakan metode *Moment of Distribution*. Maka didapatkan 15 parameter baru, yaitu 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, dan pcm\_loudness\_sma\_skewness, 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 dengan data uji banding data latih 1:9 hasilnya dimana 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 dengan data uji banding data latih 2:8 hasilnya dimana 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 dengan data uji banding data latih 3:7 hasilnya dimana 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 *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 dan yang terburuk adalah 0,4 dimana nilai  $K = 3$  pada data suara versi 3 menggunakan teknik *Validation Sampling Percentage Rate* 80:20 dan 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 banyak hasil yang diperoleh. Salah satunya, pertama perlu data suara tangisan bayi yang konsisten untuk mendapatkan hasil yang maksimal, kedua 150 data suara tangisan bayi masih dirasa kurang untuk membuat model *machine learning*. Kesulitan yang dihadapi adalah ketika mengolah data dan banyaknya nilai ( $F0_{final\_sma}$ ,  $voicing_{FinalUncilpped\_sma}$  dan  $pem\_loudness\_sma$ ) pada tiap data suara tangisan bayi. Nilai-nilai pada tiap data sangat mempengaruhi akurasi algoritma *K-Nearest Neighbours*. Sebab sulit bagi metode *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 nilai akurasi yang rendah juga.

#### IV. KESIMPULAN

Setelah melakukan penelitian terhadap suara tangisan bayi dapat ditarik beberapa kesimpulan. Yang pertama yaitu pengujian menggunakan teknik *Validation Sampling Percentage Rate* akan mencapai akurasi yang lebih baik jika nilai  $K$  lebih tinggi dan porsi data latih lebih besar daripada data uji. Yang kedua adalah pengujian data suara tangisan bayi menggunakan data sampling *Leave One Out* mengatakan semua data masih buruk karena akurasinya dibawah 50%. Yang ketiga yaitu dari 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 memisahkan persebaran data antar kategori suara tangisan bayi. Pada penelitian selanjutnya diharapkan dapat membuat sebuah model dimana metode pengenalan pola yang digunakan bisa membuat persebaran data secara maksimal antar kategori.

#### 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., Viellevoye, 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.