

SKRIPSI

**KLASIFIKASI SUARA TANGISAN BAYI BERDASARKAN
PROSODIC FEATURES MENGGUNAKAN METODE DEK-
NEAREST NEIGHBOUR**



DISUSUN OLEH :

ADITYA SINGGI PRAYOGI

NIM : 04315037

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS NAROTAMA
SURABAYA
2019**

SKRIPSI

**KLASIFIKASI SUARA TANGISAN BAYI BERDASARKAN
PROSODIC FEATURES MENGGUNAKAN METODE
K-NEAREST NEIGHBOUR**

Dixusun Oleh :

ADITYA SINGGI PRAYOGI

NIM : 04315037

Diajukan Guna Memenuhi Persyaratan
untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Pada Program Studi Teknik Informatika
Fakultas Ilmu Komputer
Universitas Narotama Surabaya

Surabaya, 26 Juli 2019

Menyetujui
Dosen Pembimbing,



Maulana Rizqi, S.T., M.T., M.Sc.
NIDN. 0729078201

Surabaya, 26 Juli 2019

Menyetujui
Dosen Pembimbing,



Tresna Maulana Fahrudin, S.ST., M.T.
NIDN. 0701059301

SKRIPSI

KLASIFIKASI SUARA TANGISAN BAYI BERDASARKAN PROSODIC FEATURES MENGGUNAKAN METODE K-NEAREST NEIGHBOUR

Disusun Oleh :

ADITYA SINGGI PRAYOGI

NIM. 04315037

Dipertahankan di depan Penguji Skripsi
Program Studi Teknik Informatika
Fakultas Ilmu Komputer
Universitas Narotama Surabaya
Tanggal : 26 Juli 2019

Mengetahui/Menyetujui

Penguji

PRO PATRIA

Program Studi Teknik Informatika
Ketua,

1. M. Noor Azam, S.Kom., MMT
NIDN. 0701097001

Made Kamisutara, ST., M.Kom.
NIDN. 0706027501

2. Maulana Rizqi, S.T., M.T., M.Sc.
NIDN. 0729078201

Ilmu Komputer

3. Cahyo Darujati, ST, MT
NIDN. 0710097402

Aryo Nugroho, ST, S.Kom., MT
NIDN. 0721077001

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Skripsi ini tidak terdapat karya yang pernah di ajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat Karya atau Pendapat yang pernah ditulis oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar acuan atau Daftar Pustaka.

Apabila ditemukan suatu Jiplakan/Plagiat maka saya bersedia menerima akibat berupa sanksi Akademis dan sanksi lain yang diberikan oleh yang berwenang sesuai dengan ketentuan peraturan dan perundang-undangan yang berlaku.

Surabaya, 26 Juli 2019



Aditya Singgi Prayogi
NIM. 04314037

MOTTO DAN PERSEMBAHAN

MOTTO

“Barang siapa yang bersungguh-sungguh, sesungguhnya kesungguhan tersebut untuk kebaikan dirinya sendiri”

- QS.Al Ankabut/29:6 -

PERSEMBAHAN

Puji syukur kepada Allah S.W.T atas limpahan rahmat dan kasihnya, hingga penulisan skripsi ini dapat saya selesaikan. Skripsi ini saya persembahkan kepada kedua orang tua terkasih, yang senantiasa membimbing, mendoakan, dan melimpahkan kasih sayang kepada putra-putri mereka. Segala bentuk perhatian dan nasehat dari kedua orang tua memberikan saya dorongan untuk melakukan yang terbaik dalam menyelesaikan permasalahan yang saya hadapi selama proses penyusunan tugas akhir ini. Tak lupa juga ucapan terima kasih kepada para sahabat, staf, dan dosen Universitas Narotama yang terlibat dalam pembuatan skripsi ini dari awal hingga akhir.

Saya menyadari bahwa Skripsi merupakan kewajiban yang harus dipenuhi sebagai syarat untuk mendapatkan gelar akademik. Akan tetapi lebih dari sebuah kewajiban, saya berharap Skripsi ini mampu memberikan manfaat kepada orang lain dan menjadi acuan serta referensi bagi para peneliti yang ingin melakukan penelitian dibidang yang sama. Dengan begitu penelitian yang saya lakukan tidak selesai sampai di sini. Akhir kata, saya sangat bersyukur pengerjaan Skripsi ini dapat

berjalan dengan lancar. Semoga semua yang saya kerjakan mendapatkan rahmat dan memberikan manfaat bagi saya pribadi maupun orang-orang di sekitar.



KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat dan hidayahnya sehingga penulis mampu menyelesaikan Skripsi ini, yang berjudul **“Klasifikasi Suara Tangisan Bayi Berdasarkan Prosodic Features Menggunakan Metode *K-Nearest Neighbours*”**. Skripsi atau dikenal sebagai tugas akhir bagi mahasiswa dikerjakan demi memenuhi syarat akademis untuk menyelesaikan program Sarjana (S1) jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Narotama. Lebih dari itu hasil dari penulisan tugas akhir ini juga diharapkan dapat memberikan manfaat bagi orang lain, terutama para peneliti yang berminat pada penelitian dibidang yang serupa.

Penyusunan tugas akhir ini tentunya sempat mengalami beberapa hambatan dan kendala, hingga akhirnya dapat diselesaikan dengan baik dan tepat waktu. Semua berkat dukungan orang-orang terkasih serta dosen pembimbing yang senantiasa mengarahkan dan memberikan solusi terhadap setiap permasalahan yang dihadapi. Oleh karena itu pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada :

1. Orang tua penulis, Ibu dan Ayah tercinta yang selalu mendukung, mendoakan dan melimpahkan kasih sayang mereka. Tak lupa pula selalu memberikan nasehat dan dukungan, hingga penulis mampu menyelesaikan penyusunan skripsi ini dengan baik.
2. Bapak Maulana Rizqi, S.T., M.T., M.Sc dan Bapak Tresna Maulana Fahrudin, S.ST., M.T selaku dosen pembimbing yang telah memberikan

nasehat dan bimbingan, hingga penulisan skripsi ini dapat berjalan dengan baik.

3. Bapak Made Kamisutara, S.T., M.Kom selaku ketua program studi Teknik Informatika.
4. Bapak Aryo Nugroho, S.T., M.Kom., M.T selaku dekan Fakultas Ilmu Komputer Universitas Narotama.
5. Para dosen penguji, Bapak Moh. Noor Al Azam, S.Kom., M.M.T, Bapak Cahyo Darujati, S.T., M.T.
6. Jajaran staf dan dosen Universitas Narotama yang tidak bisa saya sebutkan namanya satu per satu.
7. Sahabat-sahabat mahasiswa Universitas Narotama khususnya program studi Teknik Informatika angkatan 2015 dan rekan-rekan seperjuangan yang memberikan semangat dan motivasi kepada penulis.

Penulis memohon maaf yang sebesar-besarnya jika terdapat kesalahan, baik dalam penulisan maupun pemaparan materi yang kurang jelas. Semua saran dan kritik yang bersifat membangun sangat diharapkan, demi perbaikan dalam penulisan di masa mendatang.

ABSTRAK

Banyak orang tidak paham tentang arti dari suara tangisan bayi. Suara tangisan bayi merupakan sinyal dari si bayi untuk menyampaikan kondisi bayi saat ini. Dengan mengerti arti dari suara tangisan bayi kesalahan dalam merawat si bayi dapat diminimalisir. Bayi yang mendapatkan perawatan yang baik tentu akan merasa nyaman dan tidak gampang sakit. Untuk mengidentifikasi kebutuhan bayi berdasarkan suara tangisan bayi dapat dilakukan dengan menggunakan komputer. Komputer 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 antara satu data dengan data yang lain menggunakan Metode . Hasil ekstraksi fitur menggunakan *Prosodic Features* nanti yang akan menentukan hasil akurasi setelah melewati tahap klasifikasi. Setelah proses klasifikasi dilakukan dirasa kurang baik maka digunakanlah algoritma *Clustering* menggunakan algoritma *K-Means*. Penggunaan algoritma *K-Means* bertujuan untuk memfilter data mana yang memiliki pola yang sama kemudian dikelompokkan. Setelah melewati proses *Clustering* dilanjutkan kembali proses klasifikasi untuk mendapatkan akurasi. Akurasi data terbaik menggunakan proses klasifikasi yaitu 80% dimana nilai K yang digunakan adalah 9, sedangkan setelah melewati proses *Clustering* akurasi terbaik adalah 96% dimana nilai K yang digunakan adalah 9.

Kata Kunci: *Data Mining*, suara tangisan bayi, *K-Nearest Neighbours*, *K-Means*

ABSTRACT

Many people don't understand the meaning of the baby's crying sound. The sound of a baby's crying is a signal from the baby to deliver the baby's current condition. By understanding the meaning of the sound of a baby's cry, mistakes in caring for the baby can be minimized. Babies who get good care will certainly feel comfortable and not easily hurt. To identify the baby's needs based on the sound of a baby's crying can be done using a computer. Computers can identify the sound of crying babies using classification algorithms. The algorithm that can be used to classify is K-Nearest Neighbor. To classify the baby's voice it needs to be changed to numerical data which can be used in the classification process of the process called feature extraction. Feature extraction used in this study is Prosodic Features. After passing the feature extraction process, pattern recognition is needed to get the difference between one data and the other data. Using the Method. The feature extraction results will use Prosodic Features later which will determine the results of accuracy after passing the classification stage. After the classification process is done, it is considered that the Clustering algorithm uses K-Means algorithm. The use of the K-Means algorithm aims to filter out which data has the same pattern then grouped. After passing the Clustering process, the classification process is resumed to get accuracy. The best data accuracy uses a classification process that is 80% where the K value used is 9, while after passing the Clustering process the best accuracy is 96% where the K value used is 9.

Keywords: *Data Mining, sound of baby crying, K-Nearest Neighbors, K-Mean*

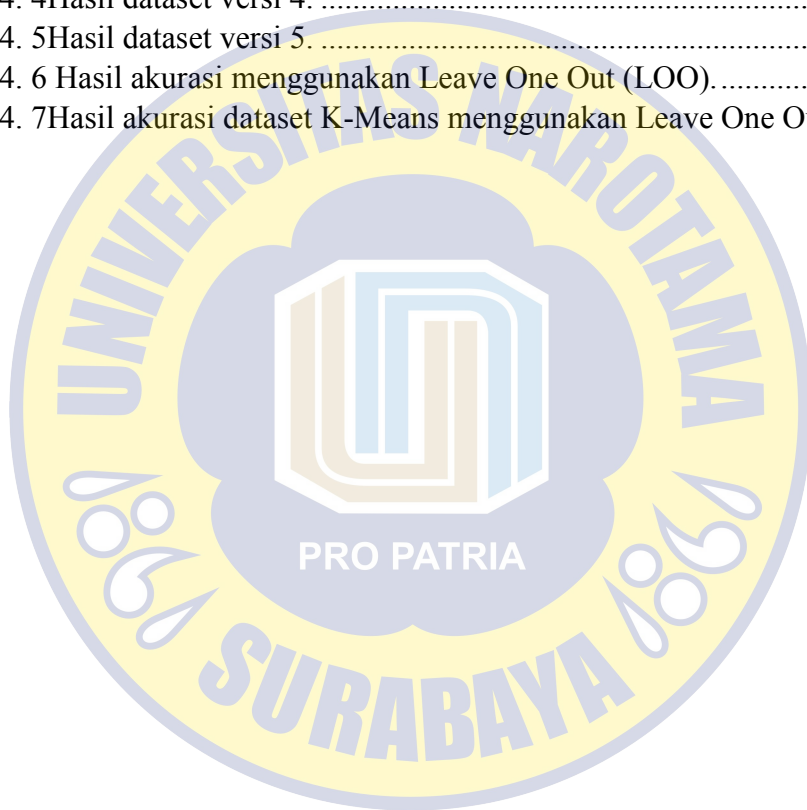
DAFTAR ISI

HALAMAN SAMBUNG	i
LEMBAR PERSETUJUAN	Error! Bookmark not defined.
LEMBAR PENGESAHAN	Error! Bookmark not defined.
SURAT PERNYATAAN	Error! Bookmark not defined.
MOTTO DAN PERSEMBAHAN	v
KATA PENGANTAR	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II	6
TINJAUAN PUSTAKA	6
2.1. Tinjauan Penelitian Terdahulu	6
2.2. Teori Dasar Yang Digunakan	15
2.2.1. <i>Dunstan Baby Language</i>	15
2.2.2. Suara Tangisan Bayi	18
2.2.3. Python	20
2.2.4. <i>Opensmile</i>	20
2.2.5. Pengenalan suara	21
2.2.6. <i>Speech Identification</i>	22
2.2.7. <i>Speech Recognition</i>	25
2.2.8. <i>Validation Sampling Menggunakan Percentage Rate.</i>	27

2.2.9. <i>Validation Sampling</i> Menggunakan <i>Leave One Out (LOO)</i>	27
2.2.10. Mengukur Kinerja Klasifikasi Menggunakan <i>Confusion Matrix</i>	27
BAB III	30
METODOLOGI PENELITIAN.....	30
3.1. Rancangan Penelitian	30
3.2. Populasi dan Sampel	31
3.3. Teknik Pengumpulan Data dan Pengembangan Instrumen.....	34
3.4. Teknik Analisis Data	38
3.4.1. <i>Speech Identification</i>	38
3.4.2. <i>Speech Recognition</i>	39
3.4.3. Mengukur Kinerja Klasifikasi Menggunakan <i>Confusion Matrix</i>	41
BAB IV	42
HASIL DAN PEMBAHASAN.....	42
4.1 <i>Preprocessing</i>	42
4.1.1. Ekstraksi Fitur.....	42
4.1.2 Mencari Pola Sinyal.....	49
4.2 Hasil Klasifikasi Dengan <i>K-Nearest Neighbour</i>	50
4.2.1 Testing Dataset Versi 1 dengan <i>Split Percentage</i>	51
4.2.2 Testing Dataset Versi 2 dengan <i>Split Percentage</i>	51
4.2.3 Testing Dataset Versi 3 dengan <i>Split Percentage</i>	52
4.2.4 Testing Dataset Versi 4 dengan <i>Split Percentage</i>	53
4.2.5 Testing Dataset Versi 5 dengan <i>Split Percentage</i>	54
4.2.6 Testing Dataset Dengan <i>Leave One Out (LOO)</i>	55
4.3 Hasil Clustering Dengan <i>K-Means</i>	56
4.4 Evaluasi Model.....	57
BAB V.....	62
PENUTUP.....	62
5.1 Kesimpulan.....	62
5.2 Saran.....	62
DAFTAR PUSTAKA	64

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	9
Tabel 2. 2 Tabel klasifikasi binary Confussion Matrix.....	28
Tabel 4. 1 Hasil Dataset Versi 1	51
Tabel 4. 2 Hasil dataset versi 2	52
Tabel 4. 3 Hasil dataset versi 3.....	53
Tabel 4. 4 Hasil dataset versi 4.....	54
Tabel 4. 5 Hasil dataset versi 5.....	55
Tabel 4. 6 Hasil akurasi menggunakan Leave One Out (LOO).....	56
Tabel 4. 7 Hasil akurasi dataset K-Means menggunakan Leave One Out (LOO).....	57



DAFTAR GAMBAR

Gambar 2. 1 Blok diagram pembelajaran pola.....	21
Gambar 2. 2 Blok diagram pengenalan suara.	22
Gambar 3. 1Desain Sistem Proses Penelitian	30
Gambar 3. 2 hasil ekstraksi fitur salah satu suara “Eairh”	31
Gambar 3. 3hasil ekstraksi fitur salah satu suara “Eh”	31
Gambar 3. 4 Grafik Fundamental Frequency dari salah satu suara “Eairh”.....	32
Gambar 3. 5 Grafik Fundamental Frequency dari salah satu suara “Eh”.....	32
Gambar 3. 6 Grafik Voicing Probability dari salah satu suara “Eairh”	33
Gambar 3. 7 Grafik Voicing Probability dari salah satu suara “Eh”	33
Gambar 3. 8 Grafik PCM Loudness dari salah satu suara “Eairh”.....	34
Gambar 3. 9 Grafik PCM Loudness dari salah satu suara “Eh”.....	34
Gambar 3. 10 Hasil audio suara “Eairh” dari CD Dunstan Baby Language.....	36
Gambar 3. 11 Hasil audio suara “Eh” dari CD Dunstan Baby Language.....	36
Gambar 3. 12 Hasil audio suara “Heh” dari CD Dunstan Baby Language.....	37
Gambar 3. 13 Hasil audio suara “Neh” dari CD Dunstan Baby Language.....	37
Gambar 3. 14 Hasil audio suara “Owh” dari CD Dunstan Baby Language	38
Gambar 4. 1Hasil ekstraksi fitur menggunakan prosodic feature kategori lapar (Neh) suara versi 1.	43
Gambar 4. 2 Grafik hasil ekstraksi fitur parameter F0final_sma kategori lapar (Neh) suara versi 1.	43
Gambar 4. 3Grafik hasil ekstraksi fitur parameter voicingFinalUnclipped_sma kategori lapar (Neh) suara versi 1.	44
Gambar 4. 4 Grafik hasil ekstraksi fitur parameter pcm_loudness_sma kategori lapar (Neh) suara versi 1.	44
Gambar 4. 5 Grafik hasil ekstraksi fitur parameter F0final_sma kategori Buang Angin (Eairh) suara versi 1.	44
Gambar 4. 6 Grafik hasil ekstraksi fitur parameter voicingFinalUnclipped_sma kategori Buang Angin (Eairh) suara versi 1.....	45
Gambar 4. 7Grafik hasil ekstraksi fitur parameter pcm_loudness_sma kategori Buang Angin (Eairh) suara versi 1.....	45
Gambar 4. 8 Grafik hasil ekstraksi fitur parameter F0final_sma kategori Sendawa (Eh) suara versi 1.	45
Gambar 4. 9 Grafik hasil ekstraksi fitur parameter voicingFinalUnclipped_sma kategori Sendawa (Eh) suara versi 1.....	46
Gambar 4. 10 Grafik hasil ekstraksi fitur parameter pcm_loudness_sma kategori Sendawa (Eh) suara versi 1.....	46
Gambar 4. 11 Grafik hasil ekstraksi fitur parameter F0final_sma kategori Tidak Nyaman (Heh) suara versi 1.....	46

Gambar 4. 12 Grafik hasil ekstraksi fitur parameter voicingFinalUnclipped_sma kategori Tidak Nyaman (Heh) suara versi 1.	47
Gambar 4. 13Grafik hasil ekstraksi fitur parameter pcm_loudness_sma kategori Tidak Nyaman (Heh) suara versi 1.	47
Gambar 4. 14 Grafik hasil ekstraksi fitur parameter F0final_sma kategori Lelah (Owh) suara versi 1.	47
Gambar 4. 15 Grafik hasil ekstraksi fitur parameter voicingFinalUnclipped_sma kategori Lelah (Owh) suara versi 1.	48
Gambar 4. 16 Grafik hasil ekstraksi fitur parameter pcm_loudness_sma kategori Lelah (Owh) suara versi 1.	48
Gambar 4. 17 Grafik hasil Moment of Distribution pada hasil ekstraksi fitur suara tangisan bayi.....	49
Gambar 4. 18Confusion Matrix Final model versi 5	60



