



**KLASIFIKASI *GENDER* BERBASIS ELEKTROKARDIOGRAM
MENGUNAKAN *PREPROCESSING* SINYAL DAN *MACHINE*
*LEARNING***

SKRIPSI

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

SELA SEPTIANA

NIM 2011016220002

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

APRIL 2024



**KLASIFIKASI *GENDER* BERBASIS ELEKTROKARDIOGRAM
MENGUNAKAN *PREPROCESSING* SINYAL DAN *MACHINE*
*LEARNING***

SKRIPSI

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

SELA SEPTIANA

NIM 2011016220002

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

APRIL 2024

SKRIPSI

KLASIFIKASI *GENDER* BERBASIS ELEKTROKARDIOGRAM MENGUNAKAN *PREPROCESSING* SINYAL DAN *MACHINE* *LEARNING*

Oleh:

SELA SEPTIANA
NIM. 2011016220002

Telah dipertahankan di depan Dosen Penguji pada tanggal 02 April 2024.

Susunan Dosen Penguji:

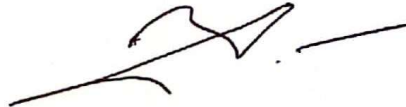
Pembimbing I



Mohammad Reza Faisal, S.T., M.T., Ph.D.

NIP. 197612202008121001

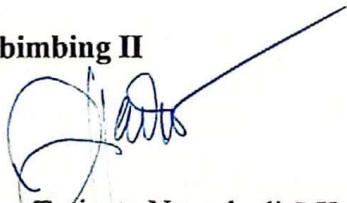
Dosen Penguji I



Muliadi, S.Kom, M.Cs.

NIP. 197804222010121002

Pembimbing II



Dodon Turianto Nugrahadi S.Kom., M.Eng.

NIP. 198001122009121002

Dosen Penguji II



Irwan Budiman, S.T., M.Kom

NIP. 197703252008121001



Perintah, 02 April 2024

Koordinator Program Studi Ilmu Komputer

Irwan Budiman, S.T., M.Kom

NIP. 197703252008121001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 02 April 2024

Yang Menyatakan,



Sela Septiana

NIM.2011016220002

ABSTRAK

KLASIFIKASI GENDER BERBASIS ELEKTROKARDIOGRAM MENGGUNAKAN *PREPROCESSING* SINYAL DAN *MACHINE LEARNING*

(Oleh : Sela Septiana; Pembimbing: Mohammad Reza Faisal., S.T., M.T., Ph.D. dan Dodon Turianto Nugrahadi S.Kom., M.Eng.; 2024; 68 halaman)

Penelitian ini fokus pada identifikasi jenis kelamin melalui sinyal Elektrokardiogram (ECG) dan pentingnya *preprocessing* sinyal untuk meningkatkan akurasi dan menangani gangguan seperti baseline wonder, artifact, dan noise otot. ECG merekam aktivitas listrik jantung dan hasilnya dapat memberikan identifikasi individu berdasarkan ciri-ciri unik sinyal tersebut. Metode klasifikasi yang digunakan adalah Support Vector Machine (SVM), *K-Nearest Neighbor* (KNN), dan Random Forest. Hasil penelitian menunjukkan bahwa *preprocessing* sinyal sebagian besar dapat meningkatkan akurasi model klasifikasi. Sebelum *preprocessing*, metode SVM mendapatkan akurasi masing-masing untuk data Raw dan Filtered sebesar 48,70% dan 61,61% meningkat menjadi 51,30% dan 66,13% setelah *preprocessing*. Metode Random Forest dengan $n_estimator=60$ dan max_depth default memberikan akurasi 50,00% untuk data Raw meningkat menjadi 54,22% setelah *preprocessing*. Sedangkan pada Random Forest dengan $n_estimator=40$ dan $max_depth=10$ memberikan akurasi 48,70% pada data Raw, lalu meningkat menjadi 51,62%. Selain itu, metode KNN dengan $k=1$ mendapatkan akurasi masing-masing untuk data Raw dan Filtered sebesar 50,65% dan 50,00% naik menjadi 53,90 % dan 51,61%. Metode KNN dengan $k=3$ akurasi sebesar 50,32% untuk data Filtered meningkat menjadi 54,84% setelah *preprocessing*. Selain itu, metode KNN dengan $k=5$ akurasi 49,03% untuk data Filtered meningkat menjadi 52,26% setelah *preprocessing*. Kesimpulannya, *preprocessing* sinyal, terutama dengan mencari titik puncak R pertama, memberikan kontribusi positif terhadap akurasi klasifikasi gender berbasis ECG. Metode SVM dan KNN dengan parameter tertentu memberikan hasil akurasi yang signifikan setelah *preprocessing* sinyal, memperkuat relevansi penggunaan ECG dalam klasifikasi jenis kelamin.

Kata kunci: Elektrokardiogram, Klasifikasi Gender, SVM, Random Forest, KNN

ABSTRACT

ELECTROCARDIOGRAM-BASED GENDER CLASSIFICATION USING SIGNAL *PREPROCESSING* AND MACHINE LEARNING

(By : Sela Septiana; *Supervisor*: Mohammad Reza Faisal., S.T., M.T., Ph.D. and Dodon Turianto Nugrahadi S.Kom.; 2024; 68 pages)

This research focuses on gender identification through Electrocardiogram (ECG) signals and the importance of signal preprocessing to increase accuracy and deal with interference such as baseline wander, artifacts, and muscle noise. ECG records the heart's electrical activity and the results can provide individual identification based on the unique characteristics of the signal. The classification methods used are Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest. The results show that signal preprocessing can largely improve the accuracy of the classification model. Before preprocessing, the SVM method obtained accuracy for Raw and Filtered data of 48.70% and 61.61% respectively, increasing to 51.30% and 66.13% after preprocessing. The Random Forest method with $n_estimator=60$ and default max_depth provides an accuracy of 50.00% for Raw data increasing to 54.22% after preprocessing. Meanwhile, Random Forest with $n_estimator=40$ and $max_depth=10$ gives an accuracy of 48.70% on Raw data, then increases to 51.62%. Apart from that, the KNN method with $k=1$ obtained accuracy for Raw and Filtered data of 50.65% and 50.00% respectively, increasing to 53.90% and 51.61%. The KNN method with $k=3$ has an accuracy of 50.32% for filtered data, increasing to 54.84% after preprocessing. In addition, the KNN method with $k=5$ accuracy of 49.03% for Filtered data increased to 52.26% after preprocessing. In conclusion, signal preprocessing, especially by searching for the first R peak point, makes a positive contribution to the accuracy of ECG-based gender classification. Random Forest and KNN methods with certain parameters provide significant accuracy results after signal preprocessing, strengthening the relevance of using ECG in gender classification.

Keywords: *Electrocardiogram, Gender Classification, SVM, Random Forest, KNN*

PRAKATA

Puji syukur penulis panjatkan ke Tuhan kita Yang Maha Esa karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “KLASIFIKASI *GENDER* BERBASIS ELEKTROKARDIOGRAM MENGGUNAKAN *PREPROCESSING* SINYAL DAN *MACHINE LEARNING*” untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Pada halaman ini penulis ingin menyampaikan ucapan terimakasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut :

1. Allah SWT, karena atas karunia-Nya lah penulis dapat menyelesaikan skripsi ini dengan tepat waktu.
2. Bapak Warnadi, Ibu Dartik, Kakak Sari Yuliana, Kakak Dwi Wulan Aprilia, dan Adek Handa Ajeng Wardani yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian skripsi ini.
3. Bapak Mohammad Reza Faisal., S.T., M.T., Ph.D. selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
4. Bapak Dodon Turianto Nugrahadi S.Kom., M.Eng. selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
5. Bapak Irwan Budiman S.T., M.Kom selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
6. Seluruh Dosen dan staff Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
7. Momo sebagai kucing kesayangan saya yang selalu menghibur saya dan seterusnya akan menghibur meskipun sudah lama hilang.

8. Alfi yang selalu membantu dalam menyelesaikan kendala dalam penyusunan skripsi ini dan selalu memberi dukungan dalam semua situasi.
9. Ica, Vania, Difa, Yra, dan Putri sebagai teman yang selalu memberi dukungan, hiburan dan tempat berbagi informasi selama masa perkuliahan.
10. Bang meftah, Bang Thor dan Kak Ridha yang membantu memberikan masukan dan saran serta memberi informasi penting.
11. Kakak – kakak dan adik – adik tingkat serta keluarga Ilmu Komputer angkatan 2020 yang memberikan dukungan dan selalu mengingatkan serta mendoakan dalam proses mengerjakan skripsi.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna, namun penulis mengharapkan bantuan serupa berupa saran dan kritik yang membangun dari semua pihak demi kesempurnaan dan mutu penulisan skripsi ini.

Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, 02 April 2024



Sela Septiana

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
PERNYATAAN.....	iii
ABSTRAK.....	iv
<i>ABSTRACT</i>	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Kajian Terdahulu.....	6
2.2 Landasan Teori.....	13
2.2.1 Elektrokardiogram.....	13
2.2.2 Klasifikasi <i>Gender</i>	14
2.2.3 Machine Learning.....	15
2.2.4 <i>Preprocessing</i> Sinyal.....	15
2.2.5 Data Mining.....	16
2.2.6 SVM.....	16

2.2.7	KNN.....	18
2.2.8	<i>Random Forest</i>	19
2.2.9	Cross Validation.....	20
2.2.10	Confusion Matrix.....	21
BAB III METODE PENELITIAN.....		23
3.1	Bahan Penelitian.....	23
3.2	Alat Penelitian.....	24
3.3	Prosedur Penelitian.....	25
BAB IV HASIL DAN PEMBAHASAN.....		28
4.1	Hasil.....	28
4.1.1.	Pengumpulan Dataset.....	28
4.1.2.	<i>Preprocessing</i> Sinyal.....	31
4.1.3.	<i>Preprocessing</i> Data.....	35
4.1.4.	Stratified K-Fold 5 Cross Validation.....	35
4.1.5.	Klasifikasi SVM.....	36
4.1.6.	Klasifikasi <i>Random Forest</i>	39
4.1.7.	Klasifikasi KNN.....	46
4.1.8.	Evaluasi.....	54
BAB V PENUTUP.....		67
5.1.	Kesimpulan.....	67
5.2.	Saran.....	68
DAFTAR PUSTAKA.....		69
LAMPIRAN.....		73
ISI LAMPIRAN.....		73

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Keaslian Penelitian.....	9
Tabel 2. Perencanaan Penelitian	12
Tabel 3. Confusion Matrix	21
Tabel 4. Dataset <i>Raw</i> awal	28
Tabel 5. Jumlah Distribusi Kelas data Raw	29
Tabel 6. Dataset Filtered Awal	29
Tabel 7. Jumlah Distribusi Kelas data Filtered	30
Tabel 8. Dataset Raw Processed	34
Tabel 9. Dataset Filtered Processed	34
Tabel 10. Detail Parameter SVM.....	36
Tabel 11. Detail Parameter Random Forest dengan Max Depth Default	40
Tabel 12. Hasil Confusion Matrix dari Random Forest dengan Max Depth Default Data Raw	40
Tabel 13. Hasil Confusion Matrix dari Random Forest dengan Max Depth Default Data Filtered.....	41
Tabel 14. Hasil Confusion Matrix dari Random Forest dengan Max Depth Default Data Raw Processed.....	41
Tabel 15. Hasil Confusion Matrix dari Random Forest dengan Max Depth Default Data Filtered Processed.....	42
Tabel 16. Detail Parameter Random Forest dengan Konfigurasi Max Depth	43
Tabel 17. Hasil Confusion Matrix dari Random Forest dengan Konfigurasi Max Depth Data Raw	44
Tabel 18. Hasil Confusion Matrix dari Random Forest dengan Konfigurasi Max Depth Data Filtered	44
Tabel 19. Hasil Confusion Matrix dari Random Forest dengan Konfigurasi Max Depth Data Raw Processed	45
Tabel 20. Hasil Confusion Matrix dari Random Forest dengan Konfigurasi Max Depth Data Filtered Processed	46

Tabel 21. Detail Parameter KNN dengan Konfigurasi n_neighbors.....	47
Tabel 22. Hasil Kinerja Klasifikasi SVM	54
Tabel 23. Hasil Kinerja Klasifikasi Random Forest dengan Parameter Max Depth Default.....	54
Tabel 24. Hasil Kinerja Klasifikasi Random Forest dengan Konfigurasi Max Depth	56
Tabel 25. Hasil Kinerja Klasifikasi KNN	58

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Pola Elektrokardiogram Normal	14
Gambar 2. Perbedaan Margin Dari Dua Hyperplane	17
Gambar 3. Ilustrasi Algoritma Random Forest	20
Gambar 4. Ilustrasi Dalam Pembagian Data	21
Gambar 5. Dataset Ecg <i>Raw</i>	23
Gambar 6. Gelombang Elektrokardiogram Dari Beberapa Record Pada Dataset <i>Raw</i>	23
Gambar 7. Dataset Ecg <i>Filtered</i>	24
Gambar 8. Gelombang Elektrokardiogram Dari Beberapa Record Pada Dataset <i>Filtered</i>	24
Gambar 9. Prosedur Penelitian.....	25
Gambar 10. Perbandingan presentase kelas female dan male data <i>Raw</i>	29
Gambar 11. Perbandingan presentase kelas female dan male data <i>Filtered</i>	30
Gambar 12. Contoh file pendukung setiap record sinyal elektrokardiogram	32
Gambar 13. Contoh Salah Satu Sinyal ECG pada Data <i>Raw</i>	32
Gambar 14. Contoh Salah Satu Sinyal ECG pada Data <i>Filtered</i>	33
Gambar 15. Contoh Hasil <i>Encoding</i> pada Data <i>Filtered Processed</i>	35
Gambar 16. Ilustrasi Proses 5-Fold Cross Validation.....	36
Gambar 17. Confusion Matriks Hasil Klasifikasi SVM pada Data <i>Raw</i>	37
Gambar 18. Confusion Matriks Hasil Klasifikasi SVM pada Data <i>Filtered</i>	38
Gambar 19. Confusion Matriks Hasil Klasifikasi SVM pada Data <i>Raw Processed</i>	38
Gambar 20. Confusion Matriks Hasil Klasifikasi SVM pada Data <i>Filtered</i> <i>Processed</i>	39
Gambar 21. Confusion Matriks Hasil Klasifikasi KNN dengan $n_neighbors=1$ pada Data <i>Raw</i>	47
Gambar 22. Confusion Matriks Hasil Klasifikasi KNN dengan $n_neighbors=1$ pada Data <i>Filtered</i>	48

Gambar 23. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=1 pada Data Raw Processed.....	48
Gambar 24. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=1 pada Data Filtered Processed.....	49
Gambar 25. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=3 pada Data Raw.....	50
Gambar 26. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=3 pada Data Filtered.....	50
Gambar 27. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=3 pada Data Raw Processed.....	51
Gambar 28. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=3 pada Data Filtered Processed.....	51
Gambar 29. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=5 pada Data Raw.....	52
Gambar 30. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=5 pada Data Filtered.....	52
Gambar 31. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=5 pada Data Raw Processed.....	53
Gambar 32. Confusion Matriks Hasil Klasifikasi KNN dengan n_neighbors=5 pada Data Filtered Processed.....	53
Gambar 33. Perbandingan Akurasi Pada Keempat Dataset ECG Dengan Menggunakan Algoritma SVM.....	60
Gambar 34. Perbandingan Akurasi Model Algoritma Random forest yang Tertinggi dari Setiap Dataset dengan Max Depth Default.....	61
Gambar 35. Perbandingan Akurasi Model Algoritma Random forest yang Tertinggi dari Setiap Dataset dengan Konfigurasi Max Depth.....	61
Gambar 36. Perbandingan Akurasi Pada Keempat Dataset ECG Menggunakan Algoritma KNN dengan n_neighbors=1.....	62
Gambar 37. Perbandingan Akurasi Pada Keempat Dataset ECG Menggunakan Algoritma KNN dengan n_neighbors=3.....	63

Gambar 38. Perbandingan Akurasi Pada Keempat Dataset ECG Menggunakan Algoritma KNN dengan $n_neighbors=5$	63
Gambar 39. Perbandingan Akurasi untuk Seluruh Klasifikasi	64
Gambar 40. Perbandingan Rata-rata Akurasi pada data Raw dan Filtered Sebelum dan Sesudah <i>Preprocessing</i> sinyal	65
Gambar 41. Perbandingan Rata-rata Akurasi pada Seluruh Dataset Sebelum dan Sesudah <i>Preprocessing</i> Sinyal	66

DAFTAR LAMPIRAN

Lampiran

Lampiran 1. Hasil Kinerja Klasifikasi Random Forest dengan Konfigurasi Parameter Max Depth pada Data Raw

Lampiran 2. Hasil Kinerja Klasifikasi Random Forest dengan Konfigurasi Parameter Max Depth pada Data Filtered

Lampiran 3. Hasil Kinerja Klasifikasi Random Forest dengan Konfigurasi Parameter Max Depth pada Data Raw Processed

Lampiran 4. Hasil Kinerja Klasifikasi Random Forest dengan Konfigurasi Parameter Max Depth pada Data Filtered Processed

Lampiran 5. Source Code Import Library

Lampiran 6. Source Code Import Dataset

Lampiran 7. Source Code Data Preparation

Lampiran 8. Source Code Data Encoding

Lampiran 9. Source Code Preprocessing Sinyal

Lampiran 10. Source Code Klasifikasi SVM

Lampiran 11. Source Code Klasifikasi Random Forest dengan Max_depth Default

Lampiran 12. Source Code Klasifikasi Random Forest dengan Konfigurasi Parameter Max_depth

Lampiran 13. Source Code Klasifikasi KNN