

## WRAPPER FEATURE SELECTION PADA EKSTRAKSI CIRI SINYAL ELEKTROKARDIOGRAM MENGGUNAKAN EMPIRICAL MODE DECOMPOSITION

Gilang Prakasa<sup>1</sup>, Achmad Rizal<sup>2</sup>, Rita Magdalena<sup>3</sup>

<sup>1</sup>Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

### Abstrak

Jantung adalah rongga berotot yang memompa darah lewat pembuluh darah oleh kontraksi berirama yang berulang. Dari aktifitas listrik otot jantung, dihasilkan suatu sinyal yang dinamakan elektrokardiogram. Elektrokardiogram (EKG) adalah gambaran sinyal hasil dari aktifitas impuls elektrik (kelistrikan) otot jantung selama periode waktu tertentu, yang direkam atau diinterpretasikan oleh perangkat atau alat bernama elektrokardiograf yang terhubung ke tubuh dengan prosedur non-invasif. Rekaman EKG digunakan oleh para dokter ahli untuk menentukan kondisi jantung seorang pasien.

Pada tugas akhir ini menggunakan metode EMD (Empirical Mode Decomposition) dengan WFS (Wrapper Feature Selection). Konsep dasar dari EMD adalah untuk mengidentifikasi skala waktu yang tepat yang dapat menunjukkan karakteristik fisik sinyal dan kemudian mengubah sinyal ke mode intrinsik dengan fungsi, yaitu Intrinsic Mode Function (IMF). Setelah itu dilakukan seleksi fitur menggunakan Wrapper Feature Selection (WFS). Wrapper adalah salah satu tipe seleksi fitur yang bertujuan untuk mendapatkan classifier pola yang mengevaluasi subset fitur dengan akurasi prediktif dengan menggunakan statistical resampling atau cross-validation. Untuk klasifikasi menggunakan K-Nearest Neighbor, suatu metode klasifikasi terhadap objek / data baru berdasarkan jarak data baru tersebut ke beberapa data / tetangga (neighbour) terdekat. Hasil akhir dari tugas akhir ini didapatkan nilai akurasi dari sistem yang dirancang dengan metode Wrapper Feature Selection dan K-Nearest Neighbor untuk melakukan klasifikasi tipe kondisi jantung seperti Normal Sinus Rhythm, Congestif Heart Failure dan Atrial Fibrillation melalui sinyal elektrokardiogram dengan menggunakan Feature Selection mencapai 84%. Sedangkan apabila tidak menggunakan Wrapper Feature Selection tetapi hanya menggunakan K-Nearest Neighbor, nilai akurasi terbaik hanya mencapai 70%.

**Kata Kunci :** Kelainan Jantung, Elektrokardiogram, Empirical Mode Decomposition, Wrapper Feature Selection, K-Nearest Neighbor



Telkom  
University

### Abstract

Heart is a muscle of cavity that pumps blood through blood vessels with repeated rhythmic contractions. Electrocardiogram (ECG) is an overview signal of the results of an electrical impulse activity (electrical) of heart muscle during a certain time period, which is recorded or interpreted by an electrocardiograph that connected to the body with a non-invasive procedure. ECG used by medical personnel to determine a patient's cardiac condition.

In this final project using EMD (Empirical Mode Decomposition) and WFS (Wrapper Feature Selection). The basic concept of the EMD is to identify the appropriate time scale to demonstrate the physical characteristics of the signal and then converts the signal into Intrinsic Mode Functions (IMF). Calculation process by reducing the number of signals analyzed by mean of the number of signals, and do repeatedly to obtain a stable signal conditions. For the feature selection using the Wrapper Feature Selection (WFS). wrapper is a type of feature selection that aims to get a pattern classifier with feature subset evaluate the predictive accuracy by using a statistical resampling or cross-validation. For classification using the K-Nearest Neighbor, a method using supervised algorithms wherein one of the methods of classification of objects / new data based on the distance of the new data to some nearby data / neighbour.

The result of this final project is an accuracy of the system designed by the method of Wrapper Feature Selection and K-Nearest Neighbor to classify types of heart conditions such as Normal Sinus Rhythm, congestive Heart Failure and Atrial Fibrillation through electrocardiogram reaches 84%. Whereas when not using Wrapper Feature Selection but only use K-Nearest Neighbor, the best accuracy just 70%.

**Keywords :** Heart Disease, Electrocardiogram, Empirical Mode Decomposition, Wrapper Feature Selection, K-Nearest Neighbor

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pada penelitian sebelumnya telah dilakukan pengenalan sinyal EKG menggunakan ekstraksi ciri *Empirical Mode Decomposition* (EMD) [1] yang menghasilkan nilai akurasi 81,33%. Penggabungan metode EMD dan *Wavelet* dilakukan pada [12] dengan nilai akurasi mencapai 90,4%. Pada [9] pemilihan fitur sinyal EKG yang dihasilkan melalui dekomposisi paket *wavelet* menggunakan *Wrapper Features Subset Selection* (WFFS) nilai akurasi dihasilkan oleh 12 ciri mencapai 100%. Pada penelitian ini menggunakan *Empirical Mode Decomposition* (EMD) sebagai metode ekstraksi ciri serta *Wrapper Feature Selection* sebagai metode tambahan untuk melakukan pemilihan atau seleksi fitur dan *K-Nearest Neighbor* sebagai metode klasifikasi.

Penggunaan metode *Wrapper Feature Selection* sebagai metode seleksi fitur diharapkan memiliki pengaruh yang cukup signifikan terhadap nilai akurasi sistem yang akan dihasilkan.

### 1.2 Tujuan dan Manfaat Penelitian

#### 1.2.1 Tujuan Penelitian

Tujuan dari tugas akhir ini adalah :

1. Mengetahui pengaruh dari penggunaan metode *Wrapper Feature Selection* sebagai seleksi fitur dari sistem yang dirancang.
2. Menganalisis performansi dari sistem yang dihasilkan.

#### 1.2.2 Manfaat Penelitian

Dengan adanya sistem ini menjadi suatu metode alternatif yang bisa diaplikasikan oleh tenaga medis maupun orang awam dalam mengenali kondisi beberapa tipe jantung berdasarkan input sinyal kardiogram.

### 1.3 Perumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, maka rumusan masalah pada tugas akhir kali ini yang akan diteliti dan dianalisis adalah :

1. Bagaimana proses kerja *Empirical Mode Decomposition* dan *Wrapper Feature Selection* dalam melakukan proses ekstraksi ciri dan pemilihan fitur klasifikasi,
2. Bagaimana pengaruh *Wrapper Feature Selection* apabila ditambahkan pada sistem yang dirancang, melihat dari ketepatan dan keakuratannya,
3. Bagaimana proses kerja *K-Nearest neighbor* sebagai *classifier* dan pengaruhnya terhadap sistem yang dirancang, melihat dari ketepatan dan keakuratannya.

### 1.4 Batasan Masalah

Dalam penyusunan tugas akhir ini, permasalahan dibatasi dalam beberapa hal yaitu :

1. Input adalah data sekunder sinyal elektrokardiogram,
2. Tipe kondisi jantung yang diklasifikasikan adalah *Normal Sinus Rhythm*, *Congestive Heart Failure* dan *Atrial Fibrillation*,
3. Ekstraksi ciri menggunakan *Empirical Mode Decomposition*,
4. Seleksi fitur menggunakan *Wrapper Feature Selection*,
5. Metode klasifikasi yang digunakan adalah *K-Nearest Neighbor*,
6. Simulasi sistem menggunakan Matlab R2009a pada proses awal hingga ekstraksi ciri (tidak ada perubahan sistem dan sintaks dari sistem sebelumnya) dan menggunakan WEKA 3.6.8 pada proses seleksi fitur dan pengklasifikasian,
7. Parameter yang dianalisis adalah tingkat ketepatan dan keakuratan sistem dalam melakukan klasifikasi tipe jantung.

## 1.5 Metodologi Penelitian

Metodologi yang digunakan dalam tugas akhir ini adalah :

1. Studi Literatur  
Mempelajari konsep tentang pengolahan sinyal serta metode yang digunakan seperti *Empirical Mode Decomposition*, *Wrapper Feature Selection* dan *K-Nearest Neighbor*. Selain itu juga mempelajari tentang konsep sistem Matlab dan WEKA.
2. Pengumpulan Data  
Melakukan pengambilan data sekunder.
3. Implementasi Perangkat Lunak  
Bertujuan untuk melakukan implementasi metode pada sistem sesuai dengan perancangan yang telah dilakukan.
4. Analisis Performansi  
Bertujuan untuk melakukan analisis performansi sistem yang telah dihasilkan.

## 1.6 Sistematika Penulisan

Tugas akhir ini disusun dengan sistematika pembahasan sebagai berikut :

### BAB I PENDAHULUAN

Bab ini menguraikan latar belakang penulisan, tujuan dan manfaat penulisan, perumusan masalah, batasan masalah, metodologi penelitian, dan sistematika penulisan.

### BAB II LANDASAN TEORI

Berisi tentang teori yang mendukung dan mendasari penulisan tugas akhir ini.

### BAB III PERANCANGAN SISTEM DAN SIMULASI

Bab ini menguraikan tentang proses perancangan tahap pemrosesan awal, proses ekstraksi ciri dan pemilihan fitur hingga proses pengklasifikasian tipe kondisi jantung.

#### BAB IV ANALISIS DAN HASIL SIMULASI

Berisi analisis terhadap hasil yang diperoleh dari tahapan perancangan sistem dan simulasi.

#### BAB V PENUTUP

Berisi kesimpulan dari analisis yang telah dilakukan dan saran untuk pengembangan lebih lanjut.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

1. Tugas akhir ini telah berhasil merancang sebuah sistem pengklasifikasian beberapa tipe kondisi jantung dengan menggunakan metode *Empirical Mode Decomposition* sebagai metode ekstraksi ciri, serta *Wrapper Feature Selection* dan *K-Nearest Neighbor* sebagai metode seleksi fitur dan *classifier* dengan tingkat akurasi yang terbaik mencapai 84 %,
2. Penggunaan *Wrapper Feature Selection* sebagai penyeleksi fitur dinilai memiliki pengaruh yang cukup signifikan terhadap hasil akurasi dimana apabila menggunakan *Wrapper Feature Selection* nilai akurasi bisa mencapai 84% dengan metode algoritma pencarian *Best First* sedangkan apabila tidak menggunakan *Wrapper Feature Selection* tingkat akurasi terbaiknya hanya mencapai 70%,
3. Penggunaan metode *K-Nearest Neighbor* sebagai metode klasifikasi dinilai cukup tepat untuk jumlah data atau atribut yang besar. Namun masih memiliki sejumlah kelemahan terutama dari nilai akurasi yang dinilai masih belum mencapai hasil yang terbaik.

#### 5.2 Saran

1. Mencoba menggunakan metode ekstraksi ciri, metode seleksi fitur, dan metode klasifikasi yang lain untuk mendapatkan tingkat akurasi yang lebih baik dalam waktu yang cepat,
2. Mencoba pengujian untuk tipe data yang berbeda, baik dari segi jumlah maupun atribut, dan
3. Mencoba mengembangkan metode klasifikasi secara *realtime* bila memungkinkan.

## DAFTAR PUSTAKA

- [1] Dewi, Rosita. 2011. *Sistem Deteksi Kelainan Jantung Menggunakan Sinyal Elektrokardiogram dengan Metode Empirical Mode Decomposition*. IT Telkom Bandung
- [2] Gutierrez-Osuna, Ricardo. *Introduction to Pattern Analysis*. Texas A&M University
- [3] Huang, N.E., et al. 1998. *The Empirical Mode Decomposition and The Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis*. Proceedings of the Royal Society of London A 454 (1971): 903–995.
- [4] Karimifard, S., A. Ahmadian, M. Khoshnevisan, M.S. Nambakhsh. 2006. *Morphological Heart Arrhythmia Detection Using Hermitian Basis Functions and kNN Classifier*. Proceeding of the 28th IEEE EMBS Annual International Conference New York City, USA, Aug 30-Sept 3, 2006.
- [5] Ladha, L., et al. 2011. *Feature Selection Methods and Algorithms*. International Journal on Computer Science and Engineering (IJCSSE) Vol. 3 No. 5 Mei 2011.
- [6] Rizal, Achmad. 2011. *K-Nearest Neighbor(K-NN)*. Dari <http://arl.blog.itelkom.ac.id/blog/2011/07/k-nearest-neighbor-k-nn> (diakses 3 Desember 2012).
- [7] Susanto, Erdi. 2012. *Data Mining Menggunakan Weka*. Available : <http://erdi-susanto.blogspot.com/2012/06/data-mining-menggunakan-weka.html/> accessed 03/12/2012.
- [8] Hermawati, Fajar Astuti. 2013. *Data Mining*. Yogyakarta : CV Andi Offset.
- [9] Rizal, Achmad. 2012. *Wrapper Features Subset Selection Pada Ekstrasi Ciri Sinyal EKG Menggunakan Metode Dekomposisi Paket Wavelet*.



Jurnal Informatika Universitas Kristen Maranatha Bandung, Vol.8  
Tahun 2012.

- [10] Physionet, 2008. ECG data Bases. MIT-BIH Data Bases. Available :  
<http://physionet.org/physiobank/database/> accessed 05/01/2013.
- [11] Halomoan, Junarto. 2013. *Analisa Sinyal EKG dengan Metoda HRV (Heart Rate Variability) pada Domain Waktu Aktivitas Berdiri dan Terlentang*. Preceeing Seminar Nasional Aplikasi Teknologi Informasi (SNATI 2013), Universitas Islam Indonesia, Yogyakarta.
- [12] Rizal, A., Jondri., Hadiyoso, S. 2011. *Pengenalan Sinyal EKG Menggunakan Empirical Mode Decomposition (EMD), Dekomposisi Paket Wavelet dan K-Mean Clustering*. Preceeding Konferensi Nasional Sistem Informatika 2011, STMIK STIKOM Bali, Denpasar.
- [13] Leng, Jinsong., Valli, Craig., Amstrong, Leisa. 2010. *A Wrapper-based Feature Selection for Analysis of Large Data Sets*. Preceeding International Conference on Computer and Electrical Engineering (ICCEE 2010), 3rd Edition, Australia.
- [14] Guetlein, Martin. 2006. *Large Scale Attribute Selection Using Wrappers*. Freiburg. Germany.

Telkom  
University