

# Klasifikasi Aktivitas Manusia Menggunakan *Extreme Learning Machine* dan Seleksi Fitur *Information Gain*

## (Classification of Human Activities Using *Extreme Learning Machine* and *Information Gain* Feature Selection)

Fitra A. Bachtiar<sup>1</sup>, Fajar Pradana<sup>2</sup>, Issa Arwani<sup>3</sup>

**Abstract**—Human activity recognition has various benefits in daily lives. However, research in this area is still facing problems which are unobtrusive data gathering, high dimensionality features, and the algorithm used to classify human activities. Those problems could impact the result of the developed model. This paper is a preliminary study in human activity recognition. Five common human activities will be recognized, namely, walking, walking upstairs, walking downstairs, sitting, and standing. The dataset used in this study consists of 1,500 data rows and 561 features. Feature selection is performed prior to the modeling step. Information Gain is used as the feature selection in which the percentile method is used to subset the number of features in the dataset. The features are then normalized and will be classified using ELM. A number of the optimal hidden neuron will be searched to yield high predictive accuracy. The results show 420 feature subsets return the higher accuracy. A number of 100 hidden neurons results in the highest predictive classification of human activity recognition. The classification result yields accuracy, precision, recall, and an F1-score of 0.85.

**Intisari**—Pengenalan aktivitas manusia dalam kehidupan sehari-hari dapat memberikan manfaat yang cukup besar. Akan tetapi, penelitian terkait pengenalan aktivitas manusia masih menghadapi beberapa permasalahan, yaitu pengambilan data yang bersifat *unobtrusive*, banyaknya fitur yang digunakan dalam pemodelan, dan penggunaan algoritme pemodelan untuk mengenali aktivitas manusia. Hal tersebut akan berdampak pada hasil penelitian dan hasil klasifikasi dari algoritme yang diusulkan. Pada makalah ini, dilakukan studi awal klasifikasi aktivitas manusia. Pengenalan aktivitas manusia dilakukan untuk memprediksi lima aktivitas manusia, yaitu berjalan, naik tangga, turun tangga, duduk, dan berdiri. Data yang digunakan adalah data sekunder dengan jumlah 1.500 baris data dan 561 fitur. Seleksi fitur dilakukan terlebih dahulu menggunakan *Information Gain* dengan pemilihan fitur menggunakan pendekatan persentil. *Subset* dari data tersebut kemudian dinormalisasi dan diklasifikasi menggunakan ELM. Parameter berupa jumlah *hidden neuron* optimum pada model ELM dicari terlebih dahulu untuk dapat menghasilkan nilai akurasi yang tinggi. Hasil seleksi fitur yang didapatkan yaitu sejumlah 420 *subset* fitur terbaik. Parameter sejumlah 100 *hidden neuron* menghasilkan akurasi tertinggi untuk pengenalan aktivitas manusia. Hasil pemodelan

menggunakan ELM memperoleh akurasi, presisi, *recall*, dan *F1-score* sebesar 0,85.

**Kata Kunci**—Klasifikasi, Aktivitas Manusia, Seleksi Fitur, IG, ELM.

### I. PENDAHULUAN

Perkembangan *ubiquitous sensing* membuat penelitian di bidang ini banyak dilakukan oleh para ilmuwan. *Ubiquitous sensing* adalah sebuah mekanisme untuk mengetahui sebuah keadaan atau lingkungan dengan penempatan sensor yang tertanam pada sebuah alat atau penempatan sensor yang tersebar. Pada dasarnya, bidang penelitian pada bidang *ubiquitous sensing* mencoba untuk mengekstraksi pengetahuan dari data yang diperoleh dari *pervasive sensors* [1], yang banyak diterapkan untuk *Human Activity Recognition* atau pengenalan aktivitas manusia. Penelitian pada bidang *ubiquitous sensing* telah banyak dilakukan untuk dapat mengembangkan suatu sistem yang dapat mengenali aktivitas manusia dengan model sederhana [2]. Selain peneliti, pihak komersial juga telah ikut bagian dalam penelitian tentang aktivitas manusia [3]. Pengenalan aktivitas manusia memiliki manfaat dalam mengenali keadaan psikologis, fisiologis, maupun interaksi antarmanusia. Pengenalan aktivitas manusia dapat memberikan informasi dari seseorang tersebut, misalnya kepribadian maupun kondisi psikologis seseorang [4]. Salah satu contoh penerapan pengenalan aktivitas manusia pada kehidupan sehari-hari adalah pendeteksian *Office Workers Syndrome* (OWS) [5]. Dalam bidang kesehatan, pengenalan aktivitas manusia juga dapat digunakan untuk memonitor pasien dalam perawatan [6], [7].

Meskipun bidang ini banyak diteliti, tetapi masih banyak tantangan-tantangan pada penelitian pengenalan aktivitas manusia. Tantangan-tantangan tersebut di antaranya adalah 1) pemilihan atribut yang akan diukur; 2) peralatan yang akan digunakan dalam mengukur aktivitas manusia tidak secara langsung diketahui oleh partisipan dan mudah dalam pengambilan data; 3) desain dalam seleksi fitur yang akan digunakan dan model inferensi yang digunakan untuk menentukan aktivitas manusia; 4) pengambilan data pada kondisi yang riil; 5) fleksibilitas dalam penambahan pengguna tanpa proses *training* kembali model inferensi; serta 6) penggunaan alat yang dapat hemat energi [8]. Selain itu, untuk mengenali aktivitas manusia tersebut, terdapat tantangan yang cukup tinggi untuk dapat mengekstraksi informasi [4].

Penggunaan metode *Machine Learning* juga telah digunakan pada bidang ini. Penelitian sebelumnya telah dilakukan

<sup>1</sup> Jurusan Teknik Informatika, Universitas Brawijaya Malang, Jl. Veteran Malang, 65145 (telp: 0341-577911; fax: 0341-577911; e-mail: <sup>1</sup>fitra.bachtiar@ub.ac.id)

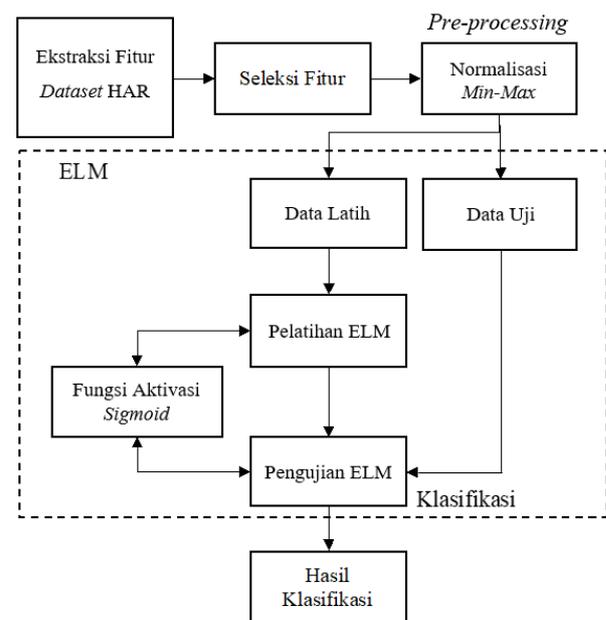
<sup>2,3</sup> Jurusan Sistem Informasi, Universitas Brawijaya Malang, Jl. Veteran Malang, 65145 (telp: 0341-577911; fax: 0341-577911; e-mail: <sup>2</sup>fajar.p@ub.ac.id, <sup>3</sup>issa.arwani@ub.ac.id)

menggunakan algoritme *Naïve Bayes* dan *Decision Tree* [9]. Dua algoritme tersebut digunakan untuk mengenali enam aktivitas manusia dengan durasi waktu yang cukup lama. Hasil penggunaan algoritme tersebut menunjukkan bahwa *Decision Tree* lebih unggul dibandingkan dengan *Naïve Bayes*, dengan akurasi 95,12%. Kemudian, telah dibandingkan juga tiga algoritme untuk mengenali enam aktivitas manusia, yaitu berdiri, berjalan, lari, naik tangga, turun tangga, dan berbaring [10]. Algoritme yang digunakan adalah *J48*, *Logistic Regression*, dan *Support Vector Machine (SVM)*, dengan parameter algoritme yang belum dicari nilai parameter terbaiknya. Hasil eksperimen tersebut menunjukkan bahwa *Regressi Logistic* memberikan hasil yang lebih tinggi dibandingkan dengan algoritme lainnya, dengan nilai akurasi 96%. Meskipun sudah menghasilkan akurasi yang tinggi, penelitian-penelitian ini masih memiliki kekurangan, yaitu parameter algoritme yang belum dicari nilai parameter terbaiknya. Permasalahan lainnya adalah penggunaan sensor pada bidang ini cenderung memiliki fitur dengan dimensi yang tinggi. Hal ini berdampak pada penurunan akurasi. Fitur-fitur dengan dimensi tinggi akan berdampak pada model inferensi yang tidak dapat maksimal dalam proses klasifikasi [11]. Selain akurasi, pemilihan metode klasifikasi juga perlu mempertimbangkan segi waktu komputasi. Dari penelitian sebelumnya, diketahui bahwa *k-NN* mudah diimplementasikan, tetapi memiliki kebutuhan penyimpanan yang tinggi dan waktu komputasi yang lama [12]. Di sisi lain, pada penelitian sebelumnya belum ada peneliti yang menggunakan algoritme berbasis *Neural Network*. Dari penelitian-penelitian tersebut belum didapatkan hasil yang maksimal tentang klasifikasi aktivitas manusia, padahal penelitian aktivitas manusia memiliki kebermanfaatan dalam kehidupan sehari-hari.

Dari uraian yang telah disampaikan di atas tentang permasalahan-permasalahan yang ada, makalah ini mengusulkan penelitian untuk mengenali aktivitas manusia dengan menggunakan seleksi fitur dan eksplorasi penggunaan algoritme *Neural Network*. Algoritme seleksi fitur yang digunakan pada makalah ini adalah *Information Gain* dan algoritme klasifikasinya adalah *Extreme Learning Machine (ELM)*. Data yang digunakan adalah data sensor yang didapatkan dari sensor *smartphone*. Tujuan makalah ini adalah membuat model klasifikasi yang dapat memprediksi aktivitas manusia. Studi dan pemodelan awal ini diharapkan dapat digunakan sebagai basis awal untuk pengembangan perangkat lunak yang dapat digunakan untuk mengenali aktivitas manusia yang dapat diterapkan untuk *monitoring* maupun personalisasi dalam berbagai bidang, khususnya dalam bidang psikologi dan kesehatan.

## II. PENGENALAN AKTIVITAS MANUSIA

Pengenalan aktivitas manusia adalah suatu bidang yang sedang marak dikembangkan dalam berbagai bidang. Sistem pengenalan aktivitas manusia berfokus pada sekumpulan aktivitas yang dilakukan oleh manusia tertentu [12]. Aktivitas manusia dapat dikelompokkan menjadi beberapa jenis, yaitu aktivitas sederhana, aktivitas kompleks, aktivitas kesehatan, aktivitas hidup, dan aktivitas kerja. Aktivitas sederhana yang



Gbr. 1 Tahapan penelitian.

dilakukan manusia di antaranya berjalan, duduk, berdiri, dan jogging, sedangkan aktivitas yang kompleks pada umumnya berhubungan dengan aktivitas lain dalam waktu yang cukup lama, misalnya berbelanja dan mengemudi [13]. Pengenalan aktivitas manusia dimanfaatkan di antaranya dalam bidang kesehatan, yaitu untuk memonitor aktivitas pasien yang sedang menjalani terapi [14]. Pengenalan aktivitas manusia juga mencakup transisi perpindahan aktivitas yang dilakukan, contohnya pengenalan aktivitas manusia dari gerakan berjalan ke duduk atau berdiri ke duduk atau sebaliknya [15].

## III. KLASIFIKASI AKTIVITAS MANUSIA

Terdapat beberapa tahapan yang dilakukan untuk membangun klasifikasi aktivitas manusia dengan seleksi fitur *Information Gain* dan ELM. Tahapan tersebut dilakukan secara berurutan, dimulai dari studi literatur, pengumpulan data dan ekstraksi fitur, seleksi fitur, normalisasi, klasifikasi aktivitas manusia, analisis hasil, dan yang terakhir adalah penarikan kesimpulan. Alur penelitian diilustrasikan pada Gbr. 1.

### A. Studi Literatur

Tahapan ini dilakukan untuk memahami kajian secara teoretis dan kajian secara empiris yang telah dilakukan sebelumnya. Dasar teori tentang aktivitas manusia, sensor giroskop, sensor akselerometer, ekstraksi fitur, seleksi fitur, dan ELM ditelaah untuk dipelajari tahapan-tahapannya. Selain itu, kajian-kajian penelitian sebelumnya juga dilakukan untuk mengetahui permasalahan pada penelitian tentang pengenalan aktivitas manusia serta metode-metode yang sudah pernah diusulkan untuk mengatasi permasalahan tersebut.

### B. Pengumpulan Data

Data yang digunakan dalam makalah ini adalah data sekunder dari penelitian yang telah dilakukan [16]. Total data yang ada pada penelitian ini berjumlah 10.299 baris data dan

terdiri atas 561 fitur. *Dataset* ini didapatkan dari eksperimen yang melibatkan tiga puluh responden dengan rentang umur 19-48 tahun. Setiap responden melakukan enam aktivitas, yaitu berjalan, naik tangga, turun tangga, duduk, berdiri, dan berbaring. Data aktivitas tersebut direkam menggunakan perangkat Samsung Galaxy S II yang merekam menggunakan sensor akselerometer dan giroskop pada *3-axial linear acceleration* dan *3-axial angular velocity* pada frekuensi 50 Hz. Perangkat tersebut diletakkan pada pinggang responden. Data yang digunakan dari penelitian tersebut kemudian disampling secara acak sejumlah 1.500 data. Dari data tersebut, 1.000 data akan digunakan untuk pelatihan dan 500 data akan digunakan untuk pengujian. Kelas label yang digunakan pada penelitian ini berjumlah lima kelas label, yaitu berjalan, naik tangga, turun tangga, duduk, dan berdiri. Jumlah fitur yang digunakan sejumlah 561 fitur.

C. Ekstraksi Fitur

Data sekunder yang digunakan telah melalui ekstraksi fitur dari pengambilan data awal. Fitur utama yang didapatkan dari sensor *smartphone*, yaitu: *body acc*, *gravity acc*, *body acc jerk*, *body angular speed*, *body angular acc*, *body acc magnitude*, *gravity acc mag*, *body acc jerk mag*, *body angular speed mag*, *body angular acc mag*. Dari fitur data tersebut, kemudian dilakukan pemetaan fitur dengan fungsi pengukuran, yaitu *mean value*, *standard deviation*, *median absolute value*, *largest value in array*, *smallest value in array*, *signal magnitude area*, *average sum of the squares*, *interquartile range*, *signal entropy*, *autoregression coefficients*, *correlation coefficient*, *largest frequency component*, *frequency signal weighted average*, *frequency signal skewness*, *frequency signal kurtosis*, *energy of a frequency interval*, dan *angel between two vectors*. Dari pemetaan fitur tersebut, dihasilkan total 561 fitur data.

D. Seleksi Fitur

Seleksi fitur dilakukan setelah tahapan pengumpulan data sekunder dan ekstraksi data selesai. Seleksi fitur yang digunakan pada makalah ini adalah seleksi fitur *Information Gain*. Persamaan seleksi fitur *Information Gain* dapat dilihat pada (1).

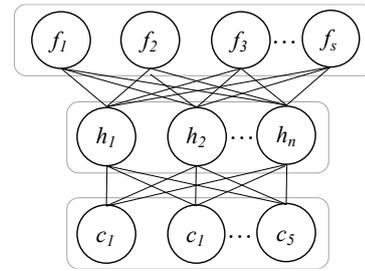
$$IG(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{S} \times Entropy(S_i) \quad (1)$$

dengan  $S$  adalah himpunan data pengenalan aktivitas manusia,  $n$  adalah jumlah partisi dari atribut  $A$ ,  $|S_i|$  adalah jumlah kasus partisi ke- $i$ , dan  $S$  adalah jumlah kasus dari data pengenalan aktivitas manusia.

Fitur-fitur diseleksi menggunakan pendekatan pemilihan persentil. Hasil dari *Information Gain* pada setiap fitur diurutkan mulai dari fitur dengan nilai *Information Gain* tertinggi ke terendah. Nilai persentil yang digunakan yaitu persentil ke-25, persentil ke-50, dan persentil ke-75. Penggunaan seleksi fitur ini diharapkan dapat meningkatkan akurasi serta mempercepat waktu komputasi.

E. Normalisasi Min-Max

Normalisasi fitur data dilakukan sebelum data hasil seleksi fitur dengan persentil digunakan untuk memodelkan klasifikasi



Gbr. 2 Arsitektur ELM.

aktivitas manusia. Normalisasi dilakukan untuk menghindari fitur yang memiliki rentang nilai yang berbeda. Fitur data dengan rentang nilai yang besar akan berpengaruh pada hasil algoritme dibandingkan dengan nilai fitur yang memiliki rentang nilai yang kecil. Persamaan normalisasi *min-max* yang digunakan ditunjukkan pada (2).

$$x' = \frac{x - \min(X)}{\max(X) - \min(X)} \quad (2)$$

dengan  $x'$  adalah hasil normalisasi data,  $x$  adalah nilai data asli,  $\min(X)$  adalah nilai terkecil dari fitur  $X$ , dan  $\max(X)$  adalah nilai terbesar dari fitur  $X$ .

F. Model Klasifikasi Aktivitas Manusia

Arsitektur ELM yang diimplementasikan diperlihatkan pada Gbr. 2. ELM yang diusulkan berdasarkan algoritme yang diusulkan pada penelitian sebelumnya [17]. Masukan pada arsitektur ELM yang digunakan berupa sejumlah fitur yang telah diseleksi menggunakan seleksi fitur *Information Gain* yang diambil sebagian dengan persentil yang dapat ditunjukkan pada (3).

$$X_i^{P_j} = X'(X_i^{P_j}), (i = 1, \dots, 1.500, j = 25, 50, 75) \quad (3)$$

dengan  $X_i^{P_j}$  adalah data yang telah diambil sebagian dari keseluruhan dengan persentil dan  $X'(X_i^{P_j})$  merupakan normalisasi yang dilakukan pada setiap *subset* data persentil menggunakan (2).

Nilai bobot pada arsitektur ini dibangkitkan secara acak. Jumlah *hidden neuron* pada arsitektur ELM dicari terlebih dahulu dengan rentang  $\{n|10 \leq n \leq 700\}$  dan interval kelipatan nilai 10 dan 200. Fungsi aktivasi yang digunakan yaitu *sigmoid*. Untuk proses pelatihan ELM, digunakan sejumlah 1.000 data dan untuk proses pengujian digunakan 500 data. Total kelas yang diklasifikasikan berjumlah lima kelas. Secara singkat, algoritme ELM dapat ditunjukkan pada *pseudocode* ELM berikut.

Algoritme ELM

Input:

- $w_i$  (input random bobot)
- $b_i$  (input random bias  $[1 \times N]$ )
- $N$  (jumlah *hidden layer*)
- $g(x)$  (fungsi aktivasi sigmoid)
- $y$  (label kelas)

Output: Hasil prediksi kelas  $x$

1. Set nilai parameter *hidden node* secara acak ( $w_i, b_i$ )
2. Hitung matriks keluaran  $\mathbf{H}$ ,  
dimana  $\mathbf{H} = \{h_{ij}\}$  ( $i = 1, \dots, x$ , dan  $j = 1, \dots, N$ )  
dan  $h_{ij} = g(w_j \cdot x_i + b_j)$

3. Hitung matriks bobot keluaran sebagai  $\hat{\beta} = H^+T$ , dimana  $H^+$  adalah generalisasi dari invers matriks Moore-Penrose  $H$
4. Hitung keluaran actual  $o$  sesuai dengan  $y$
5. Label( $y$ ) =  $\text{argmax}(o)$

### G. Rancangan Pengujian

Pengujian dilakukan untuk mendapatkan hasil kemampuan prediksi klasifikasi secara maksimal. Pengujian pertama yang dilakukan adalah menentukan jumlah *subset* fitur yang digunakan untuk tahapan klasifikasi. *Subset* fitur persentil (25, 50, 75) terbaik ditentukan menggunakan parameter 100 *hidden neuron* dan fungsi aktivasi *sigmoid* dengan acuan nilai akurasi. Pengujian kedua dilakukan untuk menentukan jumlah *hidden neuron* yang memberikan nilai prediksi tertinggi. Jumlah *hidden neuron* yang diujikan adalah 10, 20, 30, 50, 70, 90, 100, 300, 500, dan 700. Pengujian keseluruhan dilakukan setelah didapatkan nilai-nilai parameter dari seleksi fitur dan jumlah *hidden neuron* terbaik. Pengujian ini dilakukan sebanyak dua kali untuk mendapatkan generalisasi rata-rata terbaik. Hal ini dilakukan karena bobot pada ELM dibangkitkan secara acak.

### H. Metode Evaluasi

Evaluasi dilakukan untuk mengetahui keandalan dari metode yang diusulkan. Evaluasi ini dilakukan pada tahap seleksi fitur dan evaluasi model klasifikasi. Persamaan evaluasi yang digunakan yaitu akurasi, presisi, *recall*, dan *F1-score* pada kasus klasifikasi multikelas, ditunjukkan pada (4) sampai (7).

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

$$F1 - \text{score} = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

## IV. HASIL DAN PEMBAHASAN

Hasil dari pengujian pada bagian ini dibagi menjadi dua bagian, yaitu pengujian seleksi fitur dan hasil klasifikasi aktivitas manusia. Penjelasan masing-masing bagian dijabarkan pada subbagian selanjutnya.

### A. Hasil Seleksi Fitur

Pada tahap seleksi fitur, tahap awal yang dilakukan adalah menghitung nilai *Information Gain* pada setiap fitur yang ada. Nilai *Information Gain* pada setiap fitur kemudian diurutkan dari nilai terbesar ke nilai terkecil. Semakin tinggi nilai *Information Gain* pada sebuah fitur menandakan fitur tersebut akan berpengaruh terhadap klasifikasi aktivitas manusia. Sebaliknya, semakin kecil nilai *Information Gain* pada sebuah fitur menandakan fitur tersebut tidak berpengaruh terhadap klasifikasi aktivitas manusia. Dari nilai fitur tersebut, didapatkan bahwa nilai fitur tertinggi bernilai 0,921, yaitu fitur *fBodyAccMag-std()*, sedangkan nilai fitur terendah yaitu 0, pada *tBodyGyroJerk-arCoeff()-X,4*, *tBodyGyroMag-arCoeff()4*, dan *tBodyGyroMag-arCoeff()3*. Cuplikan nilai *Information Gain* pada fitur yang ada disajikan pada Tabel I. Secara visual, distribusi nilai *Information Gain* pada setiap fitur diperlihatkan pada Gbr. L1 pada lampiran.

TABEL I  
HASIL *INFORMATION GAIN* PADA SETIAP FITUR YANG DIURUTKAN

No	Nama Fitur	IG
1	<i>fBodyAccMag-std()</i>	0,92127016
2	<i>fBodyAccMag-energy()</i>	0,91912127
3	<i>tGravityAccMag-std()</i>	0,91807319
4	<i>tBodyAccMag-std()</i>	0,91807319
5	<i>fBodyAcc-bandsEnergy()-1,16</i>	0,91037796
6	<i>fBodyAccMag-mad()</i>	0,90848717
7	<i>tBodyAcc-max()-X</i>	0,90116932
...	...	...
...	...	...
559	<i>tBodyGyroJerk-arCoeff()-X,4</i>	0
560	<i>tBodyGyroMag-arCoeff()4</i>	0
561	<i>tBodyGyroMag-arCoeff()3</i>	0

TABEL II  
INFORMASI FITUR BERDASARKAN PERSENTIL

	Persentil ke-		
	25	50	75
Jumlah fitur	140	280	420
<i>Min</i>	0,748	0,697	0,296
<i>Max</i>	0,919	0,919	0,919
<i>Average</i>	0,797	0,760	0,682
<i>Standard Dev.</i>	0,044	0,049	0,144

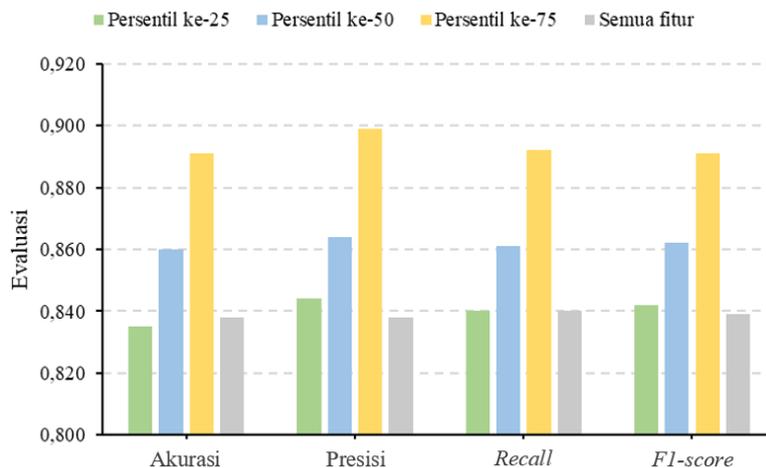
Pada tahap ini, seleksi fitur dilakukan dengan mengambil nilai persentil dari nilai *Information Gain* yang sudah didapatkan. Nilai persentil yang digunakan adalah persentil ke-25, persentil ke-50, dan persentil ke-75. Fitur diambil dan diurutkan dari fitur yang memiliki nilai *Information Gain* tertinggi sampai pada fitur yang memiliki nilai terendah. Informasi fitur yang terseleksi dengan nilai persentil tersebut ditunjukkan pada Tabel II.

Analisis awal pemodelan klasifikasi dilakukan untuk mengetahui jumlah fitur persentil yang terbaik yang akan digunakan ke tahapan pemodelan. Fitur yang dipilih dibagi menjadi empat *subset* fitur, yaitu persentil ke-25, persentil ke-50, persentil ke-75, dan semua fitur. Akurasi digunakan sebagai metode evaluasi untuk memilih *subset* fitur. Parameter yang digunakan pada pemodelan untuk pemilihan fitur adalah 100 *hidden neuron* dan fungsi aktivasi *sigmoid*.

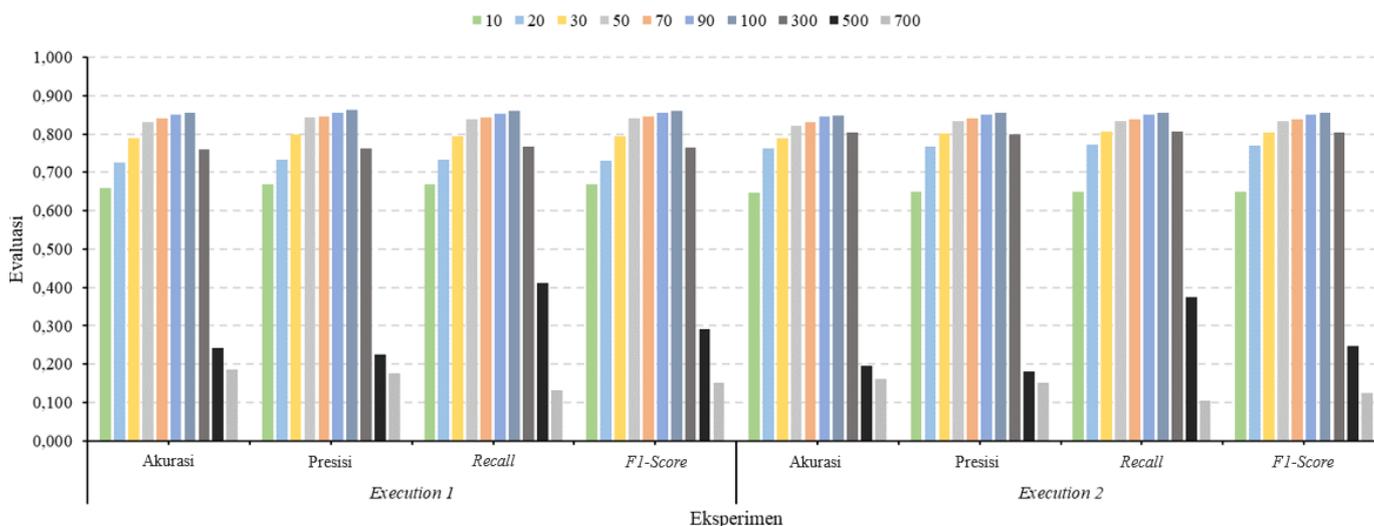
Hasil evaluasi pemilihan fitur diperlihatkan pada Gbr. 3. Dari gambar tersebut, secara umum terlihat bahwa nilai akurasi tertinggi didapatkan dari *subset* fitur persentil ke-75, yaitu 89,1%. Nilai ini lebih tinggi sekitar 4% daripada persentil ke-25, persentil ke-50, dan seluruh fitur yang digunakan. *Subset* fitur persentil ke-75 juga lebih tinggi daripada *subset* fitur yang lainnya, dilihat dari nilai evaluasi presisi, *recall*, dan *F1-score*. Hasil dari evaluasi ini adalah *subset* fitur persentil ke-75 memberikan nilai klasifikasi yang lebih tinggi dibandingkan *subset* fitur yang lainnya, sehingga *subset* fitur persentil ke-75 akan digunakan pada tahap pemodelan klasifikasi selanjutnya.

### B. Hasil Klasifikasi Aktivitas Manusia

ELM digunakan pada tahap ini untuk mengklasifikasikan aktivitas manusia. Jumlah *hidden neuron* terbaik sekaligus dicari pada tahapan ini. Jumlah *hidden neuron* yang dianalisis dimulai dari nilai rendah sampai nilai yang cukup tinggi.



Gbr. 3 Hasil evaluasi subset fitur.



Gbr. 4 Hasil eksekusi algoritme ELM dengan jumlah hidden neuron yang berbeda.

Masukan yang digunakan pada tahap ini adalah hasil seleksi fitur yang didapatkan di tahapan sebelumnya, yaitu sejumlah 420 fitur. Pemodelan dilakukan sebanyak dua kali eksekusi untuk mendapatkan generalisasi hasil klasifikasi pada ELM.

Hasil eksekusi evaluasi model dengan pencarian jumlah hidden neuron terbaik ditunjukkan pada Tabel III. Dari tabel tersebut dapat diketahui bahwa semakin besar jumlah hidden neuron, semakin tinggi pula akurasi yang didapatkan untuk mengklasifikasikan aktivitas manusia. Akan tetapi, nilai yang semakin besar juga tidak menjamin nilai akurasi semakin baik. Hal ini terlihat dari Tabel III, yaitu ketika hidden neuron berjumlah 300, akurasi yang didapatkan menurun sejumlah 10%. Tren semakin turunnya kemampuan model untuk mengklasifikasikan aktivitas berlanjut ketika jumlah hidden neuron dinaikkan. Secara visual, data pada Tabel III dapat diilustrasikan dalam Gbr. 4.

Dari hasil tersebut, rata-rata nilai dari eksekusi dapat dilihat pada Tabel IV. Tampak bahwa jumlah hidden neuron terbaik didapatkan pada nilai 100 hidden neuron dengan nilai akurasi, presisi, recall, dan F1-score sebesar, berturut-turut, 0,85, 0,85, 0,85, dan 0,85.

TABEL III  
HASIL KLASIFIKASI DENGAN JUMLAH HIDDEN NEURON YANG BERBEDA

HN	Eksekusi 1				Eksekusi 2			
	Ac	Pr	Rc	F1	Ac	Pr	Rc	F1
10	0,65	0,67	0,67	0,67	0,64	0,64	0,65	0,64
20	0,72	0,73	0,73	0,73	0,76	0,76	0,77	0,76
30	0,79	0,79	0,79	0,79	0,79	0,80	0,80	0,80
50	0,83	0,84	0,83	0,84	0,82	0,83	0,83	0,83
70	0,84	0,84	0,84	0,84	0,83	0,84	0,83	0,83
90	0,85	0,85	0,85	0,85	0,84	0,85	0,85	0,85
100	0,85	0,86	0,86	0,86	0,84	0,85	0,85	0,85
300	0,75	0,76	0,76	0,76	0,80	0,79	0,80	0,80
500	0,24	0,22	0,41	0,29	0,19	0,18	0,37	0,24
700	0,18	0,17	0,13	0,15	0,16	0,15	0,10	0,12

Beberapa informasi yang didapatkan dari hasil percobaan ini adalah sebagai berikut. Dalam kasus pengenalan aktivitas manusia dengan menggunakan 1.500 subset data, fitur terbaik didapatkan pada fitur dengan persentil ke-75 atau sejumlah 420 fitur. Kemudian, jumlah hidden neuron terbaik untuk klasifikasi lima aktivitas manusia adalah 100 hidden neuron. Secara keseluruhan, model yang dibuat untuk

TABEL IV  
HASIL RATA-RATA EKSEKUSI MODEL KLASIFIKASI ELM

HN	Rata-rata Eksekusi			
	Ac	Pr	Rc	F1
10	0,65	0,66	0,66	0,66
20	0,74	0,75	0,75	0,75
30	0,79	0,80	0,79	0,79
50	0,82	0,83	0,83	0,83
70	0,83	0,84	0,84	0,84
90	0,84	0,85	0,85	0,85
100	0,85	0,85	0,85	0,85
300	0,78	0,78	0,78	0,78
500	0,21	0,20	0,39	0,26
700	0,17	0,16	0,11	0,13

mengklasifikasikan aktivitas manusia dapat memberikan nilai yang cukup tinggi, yaitu 0,85.

## V. KESIMPULAN

Dari hasil penelitian yang telah dilakukan, dapat diambil kesimpulan sebagai berikut. Pemodelan klasifikasi aktivitas manusia dilakukan dalam dua tahap. Tahap pertama adalah melakukan seleksi fitur data yang akan digunakan pada tahap pemodelan. Seleksi fitur yang digunakan adalah seleksi fitur *Information Gain*. Jumlah fitur yang diambil menggunakan pendekatan persentil ke-25, persentil ke-50, dan persentil ke-75. Dari hasil evaluasi jumlah fitur, didapatkan nilai persentil ke-75 memberikan hasil yang terbaik, dengan mengembalikan sejumlah 420 fitur. Fitur-fitur ini selanjutnya digunakan pada tahap pemodelan. Pada tahap pemodelan, parameter *hidden neuron* dari ELM dicari terlebih dahulu. Jumlah *hidden neuron* terbaik yang didapatkan yaitu sejumlah 100 *hidden neuron*. Fitur yang telah terseleksi dan parameter jumlah *hidden neuron* yang didapatkan digunakan untuk mengklasifikasikan aktivitas manusia. Model ELM yang dibuat menghasilkan nilai akurasi sebesar 0,85.

Penelitian yang dilakukan saat ini masih dalam tahap awal untuk pengembangan pengenalan aktivitas manusia selanjutnya. Penelitian ini masih memiliki beberapa kekurangan yang dapat digunakan sebagai dasar untuk memperbaiki hasil eksperimen, yaitu sebagai berikut. Penelitian ini masih menggunakan data sekunder dan menggunakan *subset* data karena keterbatasan alat. Seleksi fitur yang digunakan saat ini dalam proses pencarian fitur-fitur terbaik dilakukan secara bertahap dengan menghitung nilai *entropy*. Pada penelitian selanjutnya, metode heuristik dapat dilakukan untuk dapat mengevaluasi secara menyeluruh fitur-fitur yang ada, seperti fitur *fBodyAccMag-std()*, *fBodyAccMag-energy()*, dan fitur lainnya yang tercantum pada Tabel I yang memiliki nilai *Information Gain* yang cukup tinggi. Bobot ELM yang dibangkitkan pada penelitian ini masih berupa bobot acak. Oleh karena itu, penggunaan metode pembangkitan bobot yang tidak acak dapat dipertimbangkan pada penelitian selanjutnya untuk mempercepat waktu komputasi. Metode klasifikasi yang digunakan pada penelitian ini adalah metode berbasis *Neural Network* saja. Penelitian selanjutnya dapat mempertimbangkan untuk menggunakan pendekatan *machine learning*, seperti *Logistic Regression*, *XGBoost*, dan *k-NN*.

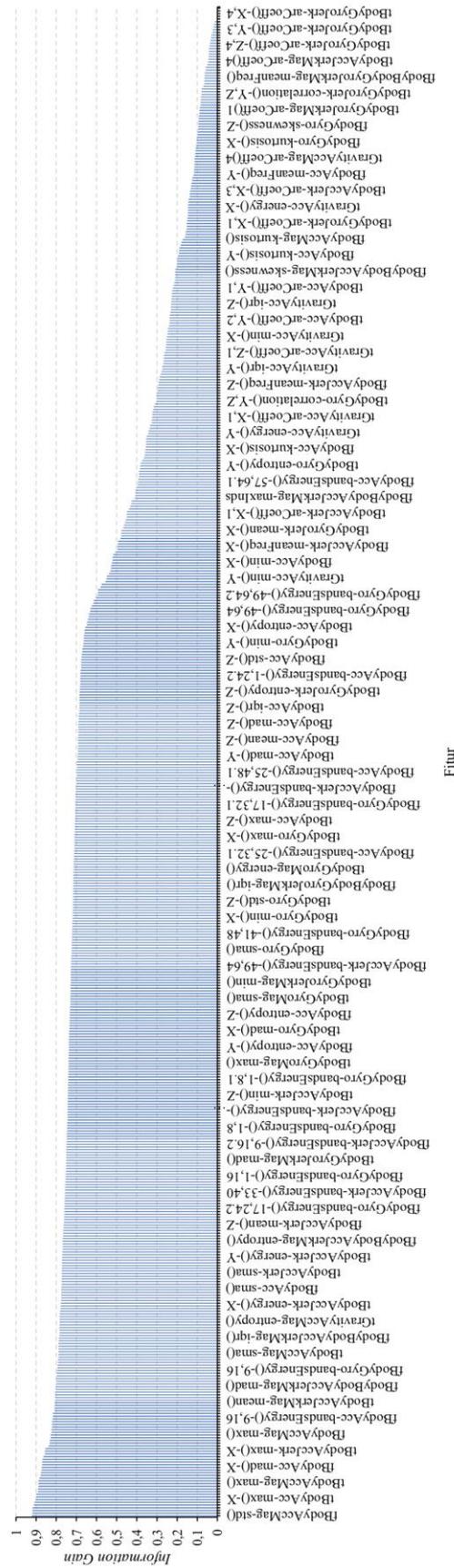
## UCAPAN TERIMA KASIH

Penelitian ini dapat terlaksana dengan bantuan Dana Penerimaan Negara Bukan Pajak (PNBP) Universitas Brawijaya, sesuai dengan Daftar Isian Pelaksanaan Anggaran (DIPA) Universitas Brawijaya Nomor DIPA-023.17.2.677512/2020.

## REFERENSI

- [1] A. Perez, M. Labrador, dan S. Barbeau, "G-Sense: A Scalable Architecture for Global Sensing and Monitoring," *IEEE Network*, Vol. 24, No. 4, hal. 57–64, 2010.
- [2] T. Stamer, B. Rhodes, J. Weaver, dan A. Pentland, "Everyday-user Wearable Computers," *Int. Symp. on Wearable Comput.*, 1999, hal. 1-12.
- [3] P. Dohnalek, P. Gajdos, T. Peterek, dan V. Snasel, "An Overview of Classification Techniques for Human Activity Recognition," *Vibroengineering PROCEDIA*, Vol. 2, hal. 117-122, 2013.
- [4] M. Vrigkas, C. Nikou, dan I.A. Kakadiaris, "A Review of Human Activity Recognition Methods," *Frontiers in Robotics and AI*, Vol. 2, hal. 1-28, 2015.
- [5] P. Paliyawan, C. Nukookit, dan P. Mongkolnam, "Prolonged Sitting Detection for Office Workers Syndrome Prevention Using Kinect," *2014 11th Int. Conf. on Elec. Eng./Electron., Comput., Telecom. and Inf. Technol. (ECTI-CON)*, 2014, hal. 1-6.
- [6] Y. Jia, "Dietetic and Exercise Therapy against Diabetes Mellitus," *2009 Second Int. Conf. on Intel. Net. and Intel. Sys.*, 2009, hal. 693-696.
- [7] J. Yin, Q. Yang, dan J. Pan, "Sensor-Based Abnormal Human-Activity Detection," *IEEE Trans. on Knowledge and Data Eng.*, Vol. 20, No. 8, hal. 1082–1090, 2008.
- [8] E. Kim, S. Helal, dan D. Cook, "Human Activity Recognition and Pattern Discovery," *IEEE Pervasive Comput.*, Vol. 9, No. 1, hal. 48–53, 2010.
- [9] T.D. Le dan C.V. Nguyen, "Human Activity Recognition by Smartphone," *2015 2nd Nat. Found. for Sci. and Technol. Dev. Conf. on Inf. and Comput. Sci. (NICS)*, 2015, hal. 219-224.
- [10] W.C. Hung, F. Shen, Y.L. Wu, M.K. Hor, dan C.Y. Tang, "Activity Recognition with Sensors on Mobile Devices," *2014 Int. Conf. on Mach. Learn. and Cybernetics*, 2014, hal. 449-454.
- [11] M. Zhang, dan A.A. Sawchuk, "A Feature Selection-based Framework for Human Activity Recognition Using Wearable Multimodal Sensors," *BodyNets 2011*, 2011, hal. 92-98.
- [12] G. De Leonardis, R. Samanta, B. Gabriella, A. Valentina, E. Panero, G. Laura, dan K. Marco, "Human Activity Recognition by Wearable Sensors: Comparison of Different Classifiers for Real-time Applications," *2018 IEEE Int. Symp. on Med. Meas. and Appl. (MeMeA)*, 2018, hal. 1-6.
- [13] X. Su, T. Hanghang, dan J. Ping, "Activity Recognition with Smartphone Sensors," *Tsinghua Sci. and Technol.*, Vol. 19, No. 3, hal. 235-249, 2014.
- [14] A. Jain dan V. Kanhangad, "Human Activity Classification in Smartphones Using Accelerometer and Gyroscope Sensors," *IEEE Sensors J.*, Vol. 18, No. 3, hal. 1169-1177, 2017.
- [15] H. Ghasemzadeh dan J. Roozbeh, "Physical Movement Monitoring Using Body Sensor Networks: A Phonological Approach to Construct Spatial Decision Trees," *IEEE Trans. on Ind. Inform.*, Vol. 7, No. 1, hal. 66-77, 2011.
- [16] D. Anguita, G. Alessandro, O. Luca, P. Xavier, dan J.L. Reyes-Ortiz, "A Public Domain Dataset for Human Activity Recognition Using Smartphones," *European Symp. on Artif. Neural Net., Comput. Intel. and Mach. Learning (ESANN)*, 2013, hal. 437-442.
- [17] G.B. Huang, Q.Y. Zhu, dan C.K. Siew, "Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks," *2004 IEEE Int. Joint Conf. on Neural Net. (IEEE Cat. No. 04CH37541)*, 2004, hal. 985-990.

Lampiran



Gbr. L.1 Visualisasi nilai Information Gain.