**Performance evaluation of decision tree methods in human activity recognition**

**NR Nurwulan1 and G Selamaj2**

1Department of Industrial Engineering, Sampoerna University, Jl. Raya Pasar Minggu Kav. 46, Pancoran, Jakarta 12780, Indonesia

2Department of Economics, Universiteti Ismail Qemali, Skelë, Rruga Kosova, Vlorë 9401, Albania

E-mail: nurul.nurwulan@sampoernauniversity.ac.id

**Abstract.** Decision tree is a supervised classifier that is easy to understand. There are various decision tree methods. This study aimed to compare the performance of decision tree methods in human activity recognition using acceleration and jerk data. The subjects performed human activity daily living, namely walking on a flat surface, walking upstairs, walking downstairs, sitting, standing, and lying down. The features were grouped into three categories: acceleration features, jerk features, and combined features of acceleration and jerk. The evaluation was done using Decision Stump, Hoeffding Tree, J48, Logistic Model Tree, Random Forest, Random Tree, and Reduced Error Pruning Tree. The results showed that Random Forest outperformed the other classifiers with acceleration features performed better than the jerk features. However, the combined acceleration and jerk features yielded the highest accuracy. In conclusion, Random Forest is the best decision tree technique in recognizing the pattern in human activity.

1. **Introduction**

Machine learning is a common method in human activity recognition. There are various machine learning techniques and each technique has its advantages and disadvantages. Decision tree is one of the most common machine learning methods due to its ability in classifying the dependent variables into distinct groups [1]. It examines the data and identifies important independent variables by considering many potential variables to make predictions [1, 2]. Decision tree saves modeling time for large datasets because it does not need a long training process. Most importantly, the classification model is easy to understand and there is no a priori assumption of the nature of the data.

Decision tree is a supervised approach in the classification. It is a simple structure with non-terminal nodes to represent tests on one or more attributes and terminal nodes reflect decision outcomes [2]. Past studies reported promising results in using decision tree classifiers for human activity recognition [3, 4, 5, 6]. Several common decision tree techniques are Decision Stump (DS), Hoeffding Tree (HT), J48, Logistic Model Tree (LMT), Random Forest (RF), Random Tree (RT), and Reduced Error Pruning Tree (RepT). The aim of this current study was to evaluate and compare the performance of the above-mentioned decision tree classifiers in recognizing human activity daily living. In addition, we also compared the acceleration and jerk based features.

The organization of this paper is as follows. Section 2 describes the methodology in obtaining the dataset and a brief description of the theoretical background behind the decision tree classifiers. Section 3 presents results and discussion. Ultimately, the conclusion is presented in section 4.

1. **Methodology**

The dataset used to evaluate the performance of decision tree methods using acceleration and jerk was taken from the study by Anguita et al. [7]. The data was collected using the triaxial accelerometer built-in Samsung smartphone with a sampling frequency rate was set to 50 Hz. Thirty subjects aged between nineteen and forty-eight were asked to perform walking, walking upstairs, walking downstairs, sitting, standing, and lying down. The dataset was pre-processed with a fifty percent overlapping window of 2.56s. The features used in this study are both the time and frequency domain features of acceleration and jerk data. Jerk data were obtained by deriving acceleration data in time.

* 1. *Decision Stump*

Decision Stump (DS) is a one-level decision tree where the split at the root level is based on a specific attribute pair. It is a decision tree with one internal node (the root) that is connected to the terminal node (leaves). This technique makes a prediction based on the value of a single input feature [8].

* 1. *Hoeffding Tree*

Hoeffding Tree (HT) is using the Hoeffding bound to construct and analyze the decision tree to determine the number of instances to be run. This method assumes that the distribution generating examples do not change over time. The HT algorithm compares attributes better than other algorithms and it consumes less memory but delivers enhanced utilization. However, it takes a longer time in inspecting if ties occur [9].

* 1. *J48*

J48 is the enhanced modification of C4.5 and has been developed to generate a pruned or un-pruned C4.5 decision tree [6, 10]. It chooses the attribute of the data that splits the set of samples into subsets with the normalized information gain as the splitting criterion. It creates a leaf of a node for the decision tree to choose the class when all the samples in the list belong to the same class [4]. For discrete attributes, the number of distinct values is considered. While for continuous attributes, binary tests in all distinct values of the attribute are considered [2]. The training dataset in consideration is sorted for the values of the continuous attribute and the entropy gains of the binary cut based on each distinct value are calculated in one scan [2].

* 1. *Logistic Model Tree*

Logistic Model Tree (LMT) is a combination of a decision tree and logistic regression functions [11]. The logistic regression functions are at the leaves. The leaf node has two child nodes that are branched to right and left, the left branch is for the attributes with smaller value and the right branch is for attributes with greater value than the threshold. The LMT is a promising method considering its robustness in giving explicit class probability estimates [12].

* 1. *Random Forest*

Random Forest (RF) is an ensemble of unpruned decision trees that can improve the classification performance by combining the bootstrap aggregating method and randomization in the selection of segmenting data nodes [5, 13]. Prediction is made by the majority vote or averaging for the regression of the ensemble. RF yields a generalization error rate but more robust to noise [2]. In a comparative study for human activity recognition, RF outperformed the other classifiers [5].

* 1. *Random Tree*

Random Tree (RT) is a decision tree drawn at random from a set of possible trees. In this algorithm, the trees are uniform, and each tree has an equal chance of being sampled. RT can be generated efficiently and the combination of large sets of RTs generally increases the accuracy [2].

* 1. *Reduced Error Pruning Tree*

Reduced Error Pruning Tree (RepT) is a technique that builds a decision tree from information gain as the splitting criterion, then prunes it using reduced-error pruning [2]. This technique sorts numeric attribute values once and the missing values are then searched using fractional instances of the C4.5 technique.

1. **Results and Discussion**

This study evaluated the performance of seven decision tree methods: DS, HT, J48, LMT, RF, RT, and RepT using acceleration and jerk data in order to see which decision tree method is the best for recognizing human activity daily living. Additionally, we also wanted to see the performance of jerk compared to acceleration in human activity recognition. The evaluation was conducted for three different groups: acceleration, jerk, and combination of both acceleration and jerk. Then, these groups were evaluated using the decision tree methods, as can be seen in Table 1. The RF outperformed the other classifiers except for jerk-based features. In jerk-based features, the LMT yielded the highest accuracy. Overall, RF and LMT performed better than other techniques, and DS is the worst technique in this study. RF is an effective method to rank the importance of variables. Unlike the other decision tree classifiers, each tree in the RF can only select a random subset of features that makes it possible to increase the variation among the trees in the model. Thus, the classification will result in higher accuracy considering the low correlation across trees [5, 14].

**Table 1.** Performance evaluation of the classifiers (in %).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Attribute | DS | HT | J48 | LMT | RF | RT | RepT |
| Acceleration | 78.47 | 82.65 | 82.65 | 87.00 | 87.45 | 85.72 | 85.80 |
| Jerk | 78.51 | 81.43 | 82.30 | 83.71 | 83.52 | 81.64 | 83.03 |
| Combined | 78.51 | 84.52 | 84.52 | 88.38 | 89.32 | 86.39 | 87.51 |

The performance of a combination of both acceleration and jerk yielded in highest accuracy. When the acceleration and jerk were evaluated as individual groups, acceleration resulted in higher accuracy than jerk. Jerk is the derivative of acceleration and it is supposed to detect slight changes in the activities [15]. However, jerk performed worse than acceleration in this study. This might happen because the past study used jerk to evaluate the activity recognition in animals with less-to-none instructions from the experimenters.

**Table 2.** Confusion matrix of RF for acceleration.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Labeled | Recognized results | | | | | |
| Standing | Sitting | Lying | Walking | Downstairs | Upstairs | |
| Standing | 323 | 174 | 171 | 0 | 0 | 0 | |
| Sitting | 231 | 218 | 174 | 0 | 0 | 0 | |
| Lying | 200 | 151 | 328 | 0 | 1 | 1 | |
| Walking | 0 | 0 | 0 | 508 | 14 | 81 | |
| Downstairs | 0 | 0 | 0 | 25 | 431 | 37 | |
| Upstairs | 0 | 0 | 1 | 106 | 24 | 410 | |

**Table 3.** Confusion matrix of RF for jerk.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Labeled | Recognized results | | | | | |
| Standing | Sitting | Lying | Walking | Downstairs | Upstairs | |
| Standing | 240 | 197 | 231 | 0 | 0 | 0 | |
| Sitting | 226 | 198 | 199 | 0 | 0 | 0 | |
| Lying | 218 | 196 | 265 | 0 | 0 | 2 | |
| Walking | 0 | 0 | 0 | 369 | 146 | 88 | |
| Downstairs | 0 | 0 | 0 | 114 | 329 | 50 | |
| Upstairs | 0 | 0 | 1 | 96 | 20 | 424 | |

Since the RF outperformed the other classifiers, we focus on the RF to compare the performance of acceleration and jerk. Table 2 and Table 3 show the confusion matrix of the RF for acceleration and jerk. The standing, sitting, and lying are misclassified with one another. However, acceleration classified better than jerk. In contrast to the past study [15], jerk did not perform well in recognizing human activities. This might happen because there were little changes in forces during standing, sitting, and lying. Jerk is felt when there is a change in force [16]. As for walking on a flat surface, walking downstairs, and walking upstairs, jerk is also the worst in classifying the activities correctly. The acceleration group is able to discriminate against the walking activities. This might happen because the subjects in this study were healthy young adults [17]. They have better postural control in handling both increasing and decreasing forces during walking downstairs and upstairs.

Based on the results in this study, it can be said that RF is the best classifier for human activity recognition. In addition, jerk does not perform better than acceleration in human activity recognition when the change of force in the activities is negligible. Further works with activities with an obvious change of force are needed.

1. **Conclusion**

This paper evaluated the performance evaluation of decision tree methods in human activity daily living using acceleration and jerk. Supervised decision tree techniques, namely Decision Stump (DS), Hoeffding Tree (HT), J48, Logistic Model Tree (LMT), Random Forest (RF), Random Tree (RT), and Reduced Error Pruning Tree (RepT). Based on the evaluation results, it can be concluded that the RF outperformed the other machine learning techniques in terms of accuracy and jerk did not perform better than acceleration in recognizing human activities. This might happen due to only slight force changes in acceleration in the activities being evaluated. Further works evaluating more various activities are needed to evaluate the performance of jerk in human activity recognition.

1. **References**
2. Bresferean VP 2007 *Proc. 29th Int. Conf. on Information Technology Interfaces (Cavtat)*
3. Zhao Y and Zhang Y 2007 *Adv. Space Res.* vol 41 p 1955-1959
4. McGlynn TA, Suchkov AA, Winter EL, et al. 2004 *Ap. J.* 616-1284
5. Nurwulan NR and Jiang BC 2020 *Proc. 2nd Asia Pacific Information Technology Conf. (Bali)*

p 73-77

1. Nurwulan NR 2020 *Revue d’Intelligence Artificielle* vol 34 p 45-50
2. Nurwulan NR and Jiang BC 2020 *International Journal of Innovative Technology and*

*Interdisciplinary Sciences* vol 3 p 381-394

1. Anguita D, Ghio A, Oneto L, Parra X and Reyes-Ortiz JL 2013 *Proc. 21st European Symposium*

*on Artificial Neural Networks (Bruges)*

1. Iba W and Langley P 1992 *Proc. 9th Int. Conf. on Machine Learning (Aberdeen)* p 233-240
2. Kumar A, Kaur P and Sharma P 2015 *CPUH-Research Journal* vol 1 p 28-32
3. Ian HW and Frank E 2005 *Data Mining: Practical Machine Learning Tools and* Techniques

(San Francisco: Morgan Kaufmann)

1. Landwehr N, Hall M and Frank E 2005 *Machine Learning* vol 59 p 161-205
2. Mahesh V, Kandaswamy A, Vimal C and Sathish B 2009 *J. Biomed. Sci. Eng.* vol 2 p 405-411
3. Breiman L 2001 *Machine Learning* vol 45 p 5-32
4. Maroco J, Silva D, Rodrigues A, Guerreiro M, Santana I and de Mendonca A 2011 *BMC*

*Research Notes* vol 4 p 299

1. Hamalainen W, Jarvinen M, Martiskainen P and Mononen J *2011 Proc. 11th Int. Conf.*

*on Intelligent Systems Design and Applications (Cordoba)* p 831-836

1. Eager D, Pendrill AM, Reistad N 2016 *Eur. J. Phys.* vol 37 p 065008
2. Selamaj G 2020 *Indonesian Journal of Computing, Engineering, and Design* vol 2 p 32-37