DWM EXP. 11
RESEARCH PAPER

Presented by group.....

9697	Sheldon Chettiar
9689	Abhijeet Basant
9696	Shivprasad Chavarattil

About the research paper

Title: Recognition Physical Activities with Optimal Number of Wearable Sensors using Data Mining Algorithms and Deep Belief Network

Authors: Ali H. Al-Fatlawi
Hayder K. Fatlawi
Sai Ho Ling

Data Mining

Data mining is the process of uncovering valuable insights from large datasets. It involves using algorithms and statistical models to discover hidden patterns and trends within the data. Data mining has diverse applications in healthcare, finance, marketing, and more, enabling better decision-making and process optimization

Significance of the paper

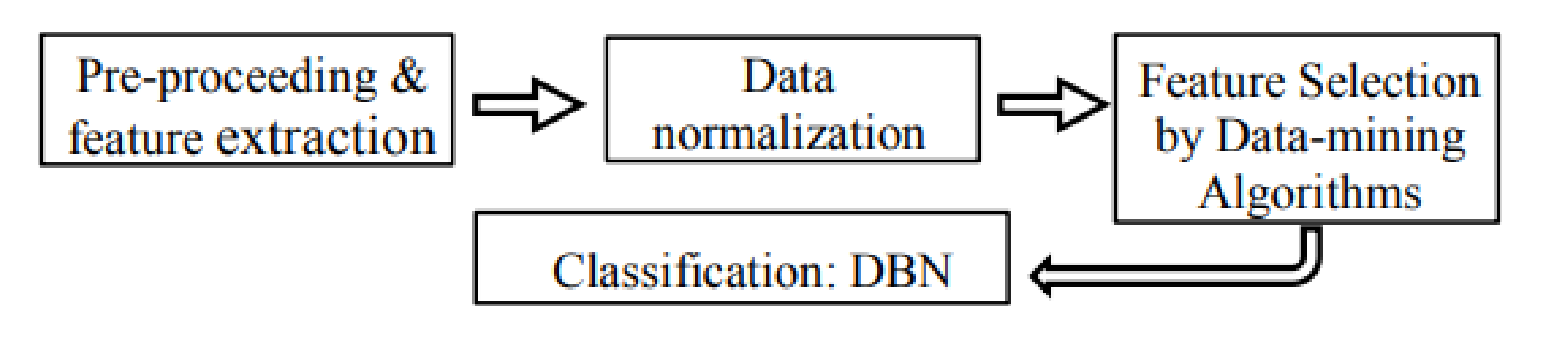
- Monitoring human physical activities plays a significant role in improving healthcare and treatment processes.
- Diseases like obstructive pulmonary disease and dementia are influenced by daily activities, making monitoring crucial.
- Monitoring daily exercises can save lives and enhance healthcare, especially when using wearable devices.
- Wearable sensors extend beyond healthcare, supporting various sectors, including sports, entertainment, and human-computer interaction (HCI).

OBJECTIVES

- The paper focuses on optimizing the recognition of human physical activities using data mining techniques.
- Data mining algorithms (Decision Tree, Random Forest, and PART Algorithm) are employed to assess the significance of sensors' features in the classification process.
- The primary objective is to reduce the number of sensors required for activity classification while maintaining performance.
- Additionally, a deep belief network (DBN) is developed to classify activities based on a reduced number of features.

METHODOLOGY

In this paper, the main purpose is hybridising the data mining algorithms and DBN to monitor 33 daily human physical activities by using an optimised number of wearable sensors. The figure below demonstrates the flow of the signals through the processing stages that are required to build an efficient system able to identify the activities with high accuracy rate.



A. Data description and normalisation

The first process is to extract and present the features of the collected raw data. Then data normalisation is processed to reduce data redundancy and improve data integrity. For this normalisation operation, "mean zero standard" method has been applied and given below. x is the original inputs and S is the normalised input vectors respectively, and n is the number of the elements in the input vector

$$S_i = \frac{x_i - \bar{x}_i}{\sigma} \qquad , i = 1, 2, \ldots n$$

where

$$\bar{x}_i = \frac{\sum_n^{i=1} x_i}{n}$$

$$\sigma = \sqrt{\frac{\sum_n^{i=1} x_i - \bar{x}_i}{n-1}}$$

B. Feature Selection Process

For the feature selection task, three data-mining algorithms (Decision Tree, Random Forest and PART Algorithm) are applied. The sensors generate signals in responding to each activity, and these signals will be transmitted to the processor which may be smartphones, tablets or computers. Optimal feature selection aims to reduce dimensionality for improved resource efficiency. Below are the data mining algorithms used.

1. Decision Tree
2. PART Rule Algorithm
3. Random Forest

C. Classification: Deep Belief Network (DBN)

- Develops a DBN to handle nonlinearity in selected features.
- DBN is a feed-forward neural network with a stack of Gaussian-Bernoulli Restricted Boltzmann Machines (GBRBM).
- Fine-tunes using the Backpropagation algorithm.

Experiment and Results

A. Experiment Setup

Using the REALDISP dataset, this study analyzes data from nine wearable sensors, tracking 33 daily human activities. Subjects contributed under different sensor placement conditions (ideal, self, mutual), with data divided into training and testing sets.

TABLE I. THE POSITIONS OF THE WEARABLE SENSORS

Symbol	Position	Symbol	Position	Symbol	Position
S1	left calf	S5	back	S9	right upper arm
S2	left thigh	S6	left lower arm		
S3	right calf	S7	left upper arm		
S4	right thigh	S8	right lower arm		

B. Features Selection results

The study evaluates feature importance for nine sensors using data mining techniques. Results indicate that S1, S4, and S5 have the lowest importance based on Gain Ratio and Information Gain metrics, with agreement between PART and Random Forest. The selected sensors for further analysis are S2, S3, S6, S7, S8, and S9.

C. Classification results

TABLE III. CLASSIFICATION RESULTS FOR THE CONSIDERED CASES

Case	# Considered Features	# Test Samples	Classification Results		Accuracy (%)
			Corrected	Fail	
Nine sensors (S1 to S9)	117	32891	32042	849	97%
Without S1	104	32891	32072	819	97.51%
Without S4	104	32891	32003	888	97.30%
Without S5	104	32891	31753	1138	96.54 %
Without S1 & S4 & S5	78	32891	31904	987	96.52%

Feature selection identified sensors with limited impact (S1, S4, S5). Eight classifiers were optimized for five selected cases, as shown in TABLE III using Ideal-displacement data.

The Deep Belief Network achieved a high recognition rate of 97% for all 33 activities with all nine sensors. Removing S4 while keeping the other eight sensors produced similar results. Eliminating S5 slightly reduced accuracy but remained acceptable. The optimal classifier requires only six sensors (S2, S3, S6, S7, S8, and S9), achieving an overall accuracy of 96.52% with ideal placement data.

Conclusion

This paper presents a cost-effective, flexible classifier for monitoring 33 daily activities with just six optimized wearable sensors. Utilizing data mining algorithms and a Deep Belief Network (DBN) classifier, the system achieves a reliable 96.52% accuracy with fewer sensors, reducing costs while maintaining high precision.

THANK YOU