HUMAN ACTIVITY RECOGNITION

WITH USING SMARTPHONE DATA

# TABLE OF Contents

1. Project Introduction 03

2.Project Progress 3-5

2.1 Understanding Data

2.2Feature Selecting

2.3 Algorithms

2.4 Create our dataset

3. Project Issues 5-6

4. Conclusion 07

1.Project ıntroductıon

In this project, we conduct an activity recognition experiment with phone data.we receive data from the smartphone as the only sensor that collects time series signals when the 3 properties (acceleration, gyroscope, gravity) are generated in the frequency domain.All activities are classified using classification algorithms. Dimensionality reduction is provided by feature selection. The results of the experiments should show the suitability of activity recognition by smartphones used by everyone according to the success rate of the algorithms.

2.Project Progress

2.1 Understanding Data

Firstly, We check an previously prepared data set (from UCI [1]).This dataset created with an experiment .This dataset will produced by experiments with group of 30 volunteers. Each person performed six activities (WALKING, WALKING\_UPSTAIRS, WALKING\_DOWNSTAIRS, SITTING, STANDING, LAYING) using a smartphone on their waist. Using smartphones embedded accelerometer and gyroscope with 50 Hertz frequency.

These signals were used to estimate variables of the feature vector for each pattern [2]:

'-XYZ' is used to denote 3-axial signals in the X, Y and Z directions.

tBodyAcc-XYZ

tGravityAcc-XYZ

tBodyAccJerk-XYZ

tBodyGyro-XYZ

tBodyGyroJerk-XYZ

fBodyAcc-XYZ

fBodyAccJerk-XYZ

fBodyGyro-XYZ

The set of variables that were estimated from these signals are:

mean(): Mean value

std(): Standard deviation

mad(): Median absolute deviation

max(): Largest value in array

min(): Smallest value in array

energy(): Energy measure. Sum of the squares divided by the number of values.

iqr(): Interquartile range

entropy(): Signal entropy

arCoeff(): Autorregresion coefficients with Burg order equal to 4.

Correlation(): correlation coefficient between two signals.

WeloaddatatoWeka in .csv formjustforchecking.Ourallattributesarenumericaland

data has notnullvalues. Data has nearlyequallyseperatedactivitylabels (Sitting, Laying, Walking, Standing,WalkingUpstairs, WalkingDownstairs)

2.2Feature Selecting

We use Weka also for selecting features.It is a technique of selecting a subset of most significant features from the original features.

*Weka -> Select attributes ->CfsSubsetEval (AttributeEvaluator)*

*->BestFirst (SearchingMethod)*

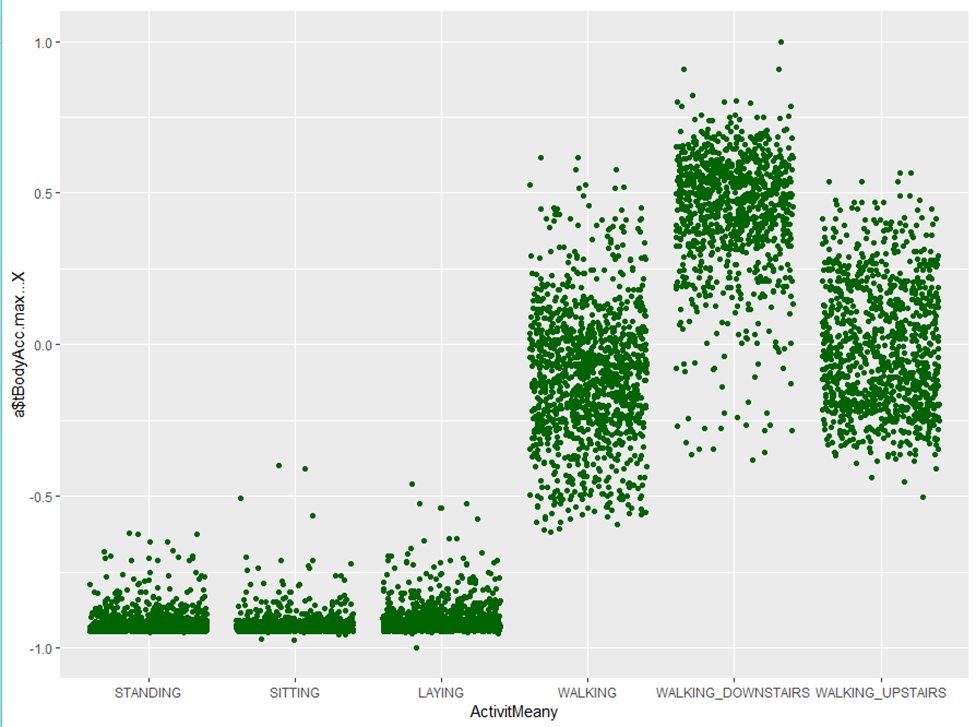
[3]: “CfsSubsetEval:Evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them”

We determine most 10 significant features with cfsSubsetEval which are:

tbody-AccMax()X,tbody-Acc-correlation()-XY,tGravityAcc-mean()-X,tGravityAcc-mean()-Y,tGravityAcc-std()-X,tGravityAcc-max()-X,tGravityAcc-max()-y,tGravityAcc-max()-z,tGravityAcc-min()-X,tGravityAcc-min()-y

We analyze a few features with using rStudio

## tbodyAccMax()X



*Thesignificantattribute is maximummeasure of acceleration .Becauseacceleration is the most important factor for the definition of motion. The maximum value of the accelerometer for classification of the six motions facilitates classification, because the value of the acceleration when sitting and lying is very close to zero or zero, but the magnitude of the acceleration is decisive when moving.*

**tGravityAcc-max()-X**

ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

Gravity is also decisive for classification

2.3 Algorithms

We apply algoritms in Weka and analyze successes of algorithms.

Random Forest algorithm is the most successful for UCI dataset.

2.4 Create our dataset

For this experiment, we created our own dataset. To do this we use an app named SensorRecorder[4].As in the UCI data, we receive accelerometer, gravity and gyroscope data at a fixed frequency of 50 Hertz. We collect 20-25 movement examples for each activity. This app give us 3- axial (one column for each axis) for accelerometer, gravity and gyroscope.These 9 column is not enough for classification.We write a script in rStudio which contains functions for calculating significant features (predetermined in UCI data).Thus, our own data composed of significant features.

After collecting data, we apply same algorithms on our data and analyze them.

The most successful algorithm is RandomForest as in UCI data.Every algorithm success increases because of measurement errors but 80% success is also successful for our aim.

2.Project ıssues

Our project problem is the misclassification of some movements in algorithms due to errors we made while receiving data. It increases our algorithm success

4.conclusıon

In this project, a smartphone data set that recognizes six human activities was produced: walk, stand, run, lie, go up and go down. In the experiment the phone has collected signals using the application, both time and frequency domain generated 9 features. A new data set was created using these features, a number of algorithms were tested: The best classification rate in our experiment was 82% obtained byRandomForest.As a conclusion, RandomForest is the optimal choice for our problem. Future planning can consider more activity and apply a real-time system on smartphones.

[1]:<https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones>

[2]:<https://archive.ics.uci.edu/ml/machine-learning-databases/00240/UCI%20HAR%20Dataset.names>

[3]:<http://weka.sourceforge.net/doc.dev/weka/attributeSelection/CfsSubsetEval.html>

[4]:<https://play.google.com/store/apps/details?id=de.martingolpashin.sensor_record>