HUMAN ACTIVITY RECOGNITION

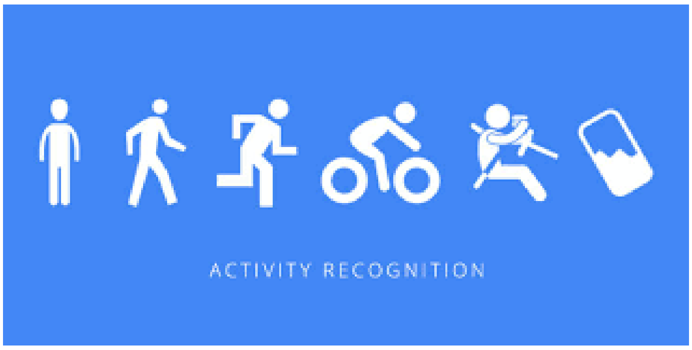
19CSE305 – MACHINE LEARNING

Asif P
Department of Computer Science Engineering
CH.EN.U4CSE20106
Amrita School of Computing, Chennai
ch.en.u4cse20106@ch.students.amrita.edu

Dr. Sangapu Sreenivasa Chakravarthi
Associate Professor
Department of Computer Science Engineering
Amrita School of Computing, Chennai
ss.chakravarthi@ch.amrita.edu

ABSTRACT:

Human activity recognition is one of the active research areas in machine learning and computer vision for various contexts like security, healthcare, and human-computer interaction. It has been a challenging problem yet it needs to be solved. It will mainly be used for eldercare and healthcare as an assistive technology when ensemble with other technologies like the Internet of Things (IoT). HAR can be done with the help of sensors, smartphones, or images. In this paper, we present various state-of-the-art methods and describe each of them through a literature survey. Different datasets are used for each of the methods where in the data are collected by different means such as sensors, images, accelerometers, gyroscopes, etc., and the placement of these devices at various locations. The results obtained by each technique and the type of dataset are then compared. Machine learning techniques like Logistic Regression, Logistic Regression CV, and random forest classifier are also presented.



INTRODUCTION

Human activity recognition (HAR) plays an important role in people's daily lives because it has the ability to learn profound advanced knowledge about human activities from raw sensor data technology of HAR has grown in popularity as human-computer interaction applications have both domestically and internationally. By removing elements from routine activities, people might automatically categorize the many types of human motion and collect the information that the body needs to express, which in turn serves as the foundation for additional intelligent applications. Home behavior analysis, video surveillance, gait analysis, and gesture recognition have all made extensive use of this technology.

However, only real-time human detection can help achieve rich benefits. Estimating how many people are in a room just by looking at their silhouettes can be very helpful in understanding the complexity of human activity. The ability to reliably track people across time and space has applications in law enforcement, healthcare, athletic performance, workplace ergonomics, and more.

The proposed method introduces collaborative learning A framework for determining the spatial and temporal extent of action. so that you can easily find the relevant part of the plot Video makes it easier to watch videos. or A person's position is derived as a latent variable. Temporary Flexibility is strengthened while learning spatiality model. Extract and display trajectory from video action. Render video using dense trajectories It's pixel-accurate, so it's more granular. Single media based solution.

LITERATURE SURVEY :

Over the years, recognition of human behavior Well researched. Most action detection methods have Manually annotate the relevant parts of the action Interest in video. This has been studied in recent years You can find the relevant part of the plot you are interested in Automatically recognize actions. we can check them out Action recognition method.

Detection was based on the Mahalanobis distance between the occurrence time description and each of the known actions. Newer popular methods using machine learning techniques like SVM and AdaBoost offer the possibility to integrate the information contained in a set of training examples. introduces the Action MACH Filter, a template-based action detection method that can capture the action by synthesizing a single action.

Most of the methods require that the relevant portion of the video has to be annotated with bounding boxes. Human intervention was tedious. So to overcome the bounding box Brendel et al. divides the video into a number of subgroups and then a model was generated that identifies the relevant subgroup.

This paper introduces a method that learns both spatial and temporal extents for detection improvement. The dense trajectory is used here as a local feature to represent human action.

PROBLEM STATEMENT:

Human Activity Recognition (HAR) is the problem of identifying a physical activity carried out by an individual dependent on a trace of movement within a certain environment. Activities such as walking, laying, sitting, standing, and climbing stairs are classified as regular physical movements and form our class of activities which is to be recognized. To record movement or change in movement, sensors such as triaxial accelerometers and gyroscopes, capture data while the activity is being performed.

LEARNING USED:

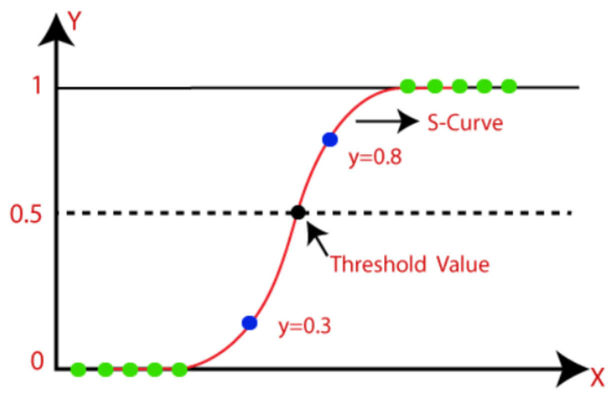
Supervised learning is the type of machine learning in which machines are trained using well "labelled" training data, and on basis of that data, machines predict the output. The labelled data means some input data is already tagged with the correct output. In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly. It applies the same concept as a student learns in the supervision of the teacher. Supervised learning is a process of providing input data as well as correct output data to the machine learning model. The aim of a supervised learning algorithm is to find a mapping function to map the input variable(x) with the output variable(y).

ALGORITHMIC MODELS USED:

LOGISTIC REGRESSION:

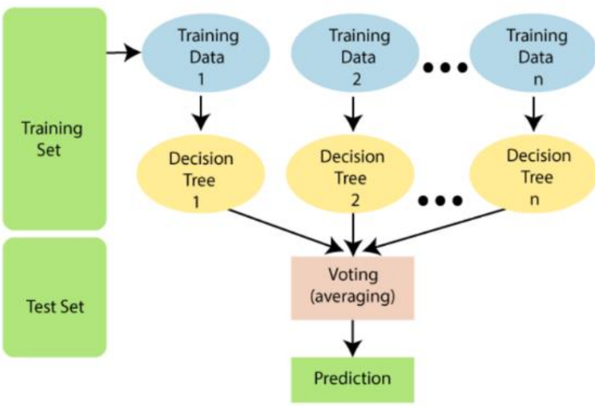
Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables.

Logistic regression predicts the output of a categorical dependent variable. Therefore, the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.



B.RANDOM FOREST:

Random forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both classification and regression problems in ml. It is based on the concept of ensemble learning, which is a process *of* combining multiple classifiers to solve a complex problem and improve the performance of the model. As the name suggests, "random forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.



PARAMETERS USED:

- mean(): Mean value
- std(): Standard deviation
- mad(): Median absolute deviation
- max(): Largest value in array
- min(): Smallest value in array
- sma(): Signal magnitude area
- 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
- maxInds(): index of the frequency component with largest magnitude
- meanFreq(): Weighted average of the frequency components to obtain a mean frequency
- skewness(): skewness of the frequency domain signal
- kurtosis(): kurtosis of the frequency domain signal
- bandsEnergy(): Energy of a frequency interval within the 64 bins of the FFT of each window.
- angle(): Angle between to vectors.

WHY DO YOU USE THIS MODEL?

1. This model is mainly based on classification. So we use Supervised learning to get the output.
2. We use Logistic Regression as we need to classify the given 6 types of activities performed as this calculates or predicts the probability of a binary yes/no occurring.
3. We use a random forest classifier as it is often used for both classification and regression problems in the concept of machine learning on the concept of ensemble learning. which is a process of combining multiple classifiers to solve a complex problem and improve the performance of the model.
4. We also use Logistic Regression CV as it cross validates the total models of Logistic Regression and implies the best one.

HOW DO YOU USE THIS MODEL?

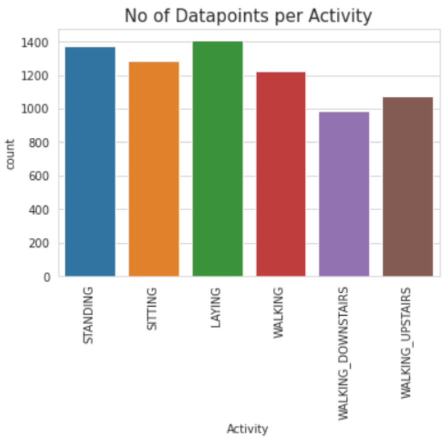
1) Import the necessary Libraries
2) Load and analyse the data
3) Find Correlations among the features
4) Split the data into train and test data (validation data)
5) Predict the activity using Logistic Regression and Logistic Regression CV
6) Calculate the Classification error matrix
7) Feature selection to pick the best features for the better prediction
8) Calculate the new classification error matrix
9)

DATASET

The data set used is activities of daily life (ADL). This project is to build a model that predicts human activities such as Walking, walking upstairs, walking downstairs, Sitting, Standing or Laying. This dataset is collected from 30 persons (referred to as subjects in this dataset), performing different activities with a smartphone to their waists. The data is recorded with the help of sensors (accelerometer and Gyroscope) in that smartphone. This experiment was video recorded to label the data manually. By using the sensors (Gyroscope and accelerometer) in a smartphone, they have captured '3-axial linear acceleration'(*tAcc-XYZ*) from the accelerometer and '3-axial angular velocity' (*tGyro-XYZ*) from Gyroscope with several variations.

-prefix 't' in those metrics denotes time.

-suffix 'XYZ' represents 3-axial signals in X , Y, and Z directions.



Total number of counts is given by:

```
LAYING                1407
STANDING              1374
SITTING               1286
WALKING               1226
WALKING_UPSTAIRS      1073
WALKING_DOWNSTAIRS     986
Name: Activity, dtype: int64
```

Total number of records

7352 in train data and 2947 in test data.

```
train.shape      test.shape

(7352, 563)      (2947, 563)
```

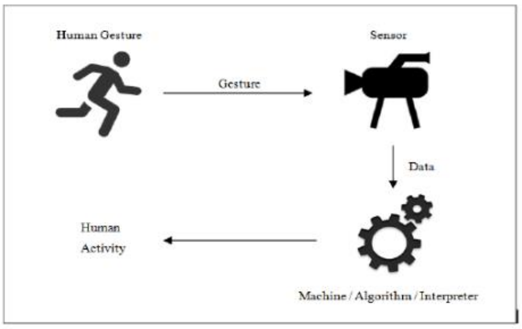
DESCRIPTION:

The experiments have been carried out with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed six activities (WALKING, WALKING UPSTAIRS, WALKING DOWNSTAIRS, SITTING, STANDING, LAYING) wearing a smartphone (Samsung Galaxy S II) on the waist. Using its embedded accelerometer and gyroscope, we captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The experiments have been video-recorded to label the data manually. The obtained dataset has been randomly partitioned into two sets, where 70% of the volunteers was selected for generating the training data and 30% the test data.

The sensor signals (accelerometer and gyroscope) were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window). The sensor acceleration signal, which has gravitational and body motion components, was separated using a Butterworth low-pass filter into body acceleration and gravity. The gravitational force is assumed to have only low frequency components, therefore a filter with 0.3 Hz cutoff frequency was used. From each window, a vector of features was obtained by calculating variables from the time and frequency domain.

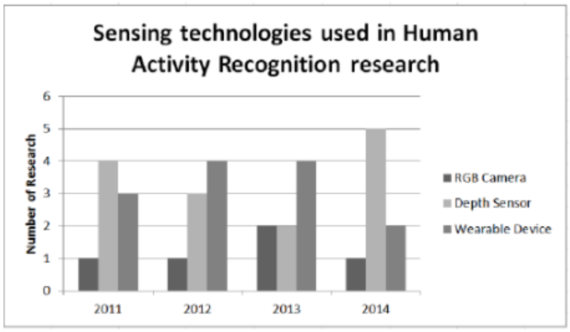
SENSING TECHNOLOGIES:

In general, conventional HAR sensors play an important role in detecting human activity. The given picture shows the process by which human activity is detected when body gestures are given as input. One or more sensors capture information captured by human body gestures, and a recognition engine analyzes that information to determine the type of activity being performed.



Depth sensors, also known as infrared sensors or infrared cameras, are employed in HAR systems to detect human activity. In other words, the depth sensor projects infrared rays into the scene and recaptures them with the infrared sensor to calculate and measure the depth or distance of each ray from the sensor. The review reveals that the Microsoft Kinect sensor is commonly used as the depth sensor for his HAR. Since the Kinect sensor can recognize 20 joints of the human body using real-world coordinates, many researchers have used coordinates to classify human activity.

Physical human activity can be identified easily through analyzing the data generated from various wearable sensors after being processed and determined by a classification
algorithm. Further to this, Kantoch and Augustyniak claim that GPS and temperature signal acquired from a smartphone can be further fed into machines for healthcare monitoring
purpose.



CLASS LABEL :

Class Label is defined as a predicted output for a given variable.

1. Walking
2. Walking Upstairs
3. Walking Downstairs
4. Sitting 5. Standing
6. Laying

FEATURES:

1. These sensor signals are preprocessed by applying noise filters and then sampled in fixed-width windows (sliding windows) of 2.56 seconds each with 50% overlap. ie., each window has 128 readings.

2. From Each window, a feature vector was obtianed by calculating variables from the time and frequency domain.

In our dataset, each datapoint represents a window with different readings

3. The accelertion signal was saperated into Body and Gravity acceleration signals(tBodyAcc-XYZ and tGravityAcc-XYZ) using some low pass filter with corner frequecy of 0.3Hz.
4. After that, the body linear acceleration and angular velocity were derived in time to obtian jerk signals (tBodyAccJerk-XYZ and tBodyGyroJerk-XYZ).
5. The magnitude of these 3-dimensional signals were calculated using the Euclidian norm. This magnitudes are represented as features with names like tBodyAccMag, tGravityAccMag, tBodyAccJerkMag, tBodyGyroMag and tBodyGyroJerkMag.
6. Finally, We've got frequency domain signals from some of the available signals by applying a FFT (Fast Fourier Transform). These signals obtained were labeled with prefix 'f' just like original signals with prefix 't'. These signals are labeled as fBodyAcc-XYZ, fBodyGyroMag etc.,.
7. These are the signals that we got so far.
    1. tBodyAcc-XYZ
    2. tGravityAcc-XYZ
    3. tBodyAccJerk-XYZ
    4. tBodyGyro-XYZ
    5. tBodyGyroJerk-XYZ
    6. tBodyAccMag
    7. tGravityAccMag
    8. tBodyAccJerkMag
    9. tBodyGyroMag
    10. tBodyGyroJerkMag
    11. fBodyAcc-XYZ
    12. fBodyAccJerk-XYZ
    13. fBodyGyro-XYZ
    14. fBodyAccMag
    15. fBodyAccJerkMag
    16. fBodyGyroMag
    17. fBodyGyroJerkMag
8. We can esitmate some set of variables from the above signals. ie., We will estimate the following properties on each and every signal that we recoreded so far.
    1. mean(): Mean value
    2. std(): Standard deviation
    3. mad(): Median absolute deviation
    4. max(): Largest value in array

5. min(): Smallest value in array
6. sma(): Signal magnitude area
7. energy(): Energy measure. Sum of the squares divided by the number of values.
8. iqr(): Interquartile range
9. entropy(): Signal entropy
10. arCoeff(): Autorregresion coefficients with Burg order equal to 4
11. correlation(): correlation coefficient between two signals
12. maxInds(): index of the frequency component with largest magnitude
13. meanFreq(): Weighted average of the frequency components to obtain a mean frequency
14. skewness(): skewness of the frequency domain signal
15. kurtosis(): kurtosis of the frequency domain signal
16. bandsEnergy(): Energy of a frequency interval within the 64 bins of the FFT of each window.
17. angle(): Angle between to vectors.

9. We can obtain some other vectors by taking the average of signals in a single window sample. These are used on the angle() variable' `
    1. gravityMean
    2. tBodyAccMean
    3. tBodyAccJerkMean
    4. tBodyGyroMean
    5. tBodyGyroJerkMean
10. In the dataset, Y_labels are represented as numbers from 1 to 6 as their identifiers.
    1. WALKING as 1
    2. WALKING_UPSTAIRS as 2
    3. WALKING_DOWNSTAIRS as 3
    4. SITTING as 4
    5. STANDING as 5
    6. LAYING as 6

Train and test data were separated:

- The readings from 70% of the volunteers were taken as trianing data and remaining 30% subjects recordings were taken for test data

Why did you choose this problem?

I choose this problem as Human activity recognition (HAR) plays an important role in people’s daily lives because it has the ability to learn profound advanced knowledge about human activities from raw sensor data technology of HAR has grown in popularity as human-computer interaction applications have both domestically and internationally. By removing elements from routine activities, people might automatically categorize the many types of human motion and collect the information that the body needs to express, which in turn serves as the foundation for additional intelligent applications. Home behaviour analysis, video surveillance, gait analysis, and gesture recognition have all made extensive use of this technology.

APPROACHMENT:

1. First we import the necessary libraries .
2. We load and analyse the data.
3. Check for null values as we start data cleaning as a part of pre-processing.
4. There are no null values in either the test and the train datasets
5. The subject column is not going to be usefull here so i will drop it from both data sets.
6. Scaling a dataset usually produces better dataset and more accurate predictions. First we check the range(the min and the max) for each of the datasets. Lets try using the .describe() method and lets exclude the activity column which is the last column..
7. Since They have the same data types. That is, mostly floats and one object feature.
8. Lets see what the object feature is abd extract it from the rest.
9. We need to encode the Activity column becasue sklearn won't accept sparse matrix as prediction columns . We will use LabelEncoder to encode the Activities.

```
le = LabelEncoder()
for x in [train, test]:
    x['Activity'] = le.fit_transform(x.Activity)
```

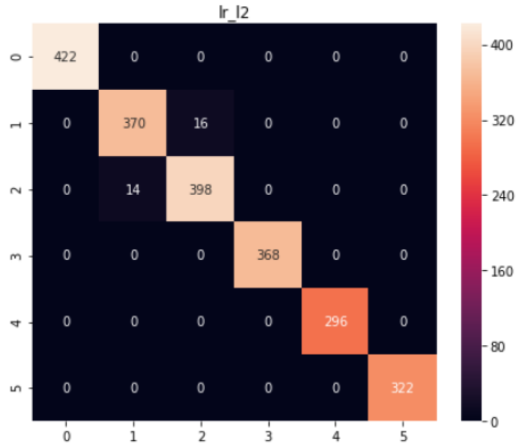
10. Correlation refers to the mutual relationship and association between quantities and it is generaly used to express one quantity in terns of its relationship with other quantities. The can either be Positive(variables change in the same direction), negative(variables change in opposite direction or neutral(No correlation).
11. Variable within a dataset can be related in lots of ways and for lost of reasons:

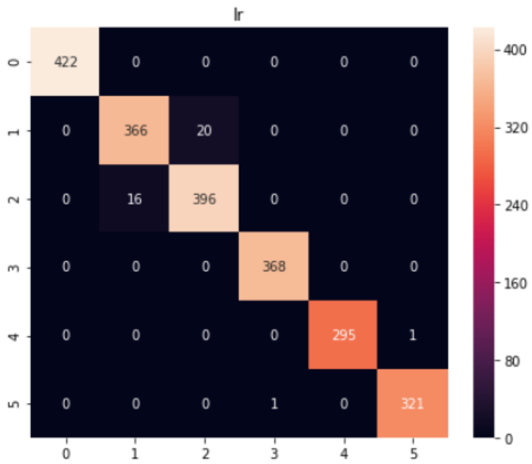
    - They could depend on values of ot her variable

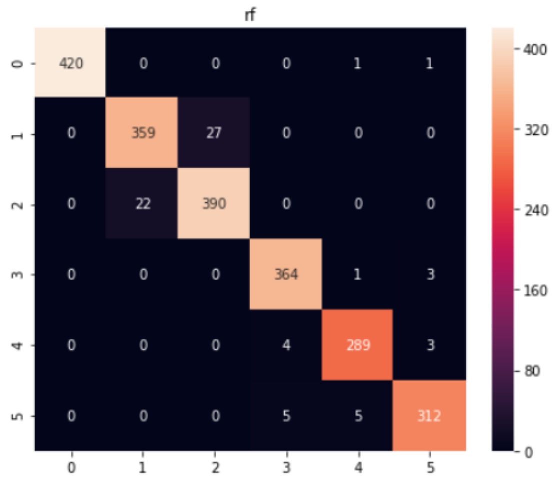
    - They could be associated to each o ther

    - They could both depend on a third variable.

12. In this project, we will be using the pandas method .corr() for calculating correlation between dataframe columns.
13. Split the dataset in train(70%) and test (30%) and perform the models.
14. We will do predictive modelling.
15. Finally calculate the error matrixes to determine the performance of classification model.







OUTPUT:

ACCURACY- The base metric used for model evaluation is often *Accuracy*, describing the number of correct predictions over all predictions

PRECISION- Precision is a measure of how many of the positive predictions made are correct (true positives)

RECALL- Recall is a measure of how many of the positive cases the classifier correctly predicted, over all the positive cases in the data

F1_SCORE- F1-Score is a measure combining both precision and recall. It is generally described as the harmonic mean of the two

	lr	lr_l2	rf
Precision	0.982787	0.986403	0.967402
Recall	0.982774	0.986401	0.967362
F_score	0.982771	0.986400	0.967357
Accuracy	0.982774	0.986401	0.967362

CONCLUSION:

We can see that the Logistic regression with L2 regularization gives slightly better error metric than the other models.

FUTURE SCOPE:

1. For HAR systems to reach their full potential, more research is required. Comparison between HAR systems is hindered and becomes unquantifiable as each researcher uses a different dataset for activity recognition.

2. Recognition of composite activities can enrich context awareness.

3. convolutional neural networks, hybrids of convolutional networks, and LSTMs should be further studied to determine their suitability to solve the problem of human activity recognition from raw signal data.

4. Existing HAR systems are mainly focused on individual activities but could be extended further towards recognizing patterns and activity trends for a group of people with the use of social networks.

BIBLIOGRAPHY:

https://www.kaggle.com/code/abheeshthmishra/eda-of-human-activity-recognition

https://www.neuraldesigner.com/solutions/activity-recognition

https://thecleverprogrammer.com/2021/01/10/human-activity-recognition-with-machine-learning/

https://www.sciencedirect.com/science/article/pii/S1877050919310166

https://www.researchgate.net/publication/339566682_Human_Activity_Recognition_using_Machine_Learning_Classification_Techniques