ABSTRACT

In the report, Human Activity Recognition system has been presented on the basis of the date gathered by the smartphone sensors. Presently, there are many smartphone applications that already use that data to estimate basic fitness statistics, such as daily step count. But these applications produce a poor method of assessing person's fitness level. This problem has lead people to turn to wearable devices to track their fitness activities. As smartphones are quickly becoming ubiquitous part of the life in the modern world, their embedded sensors have the ability to record a significant amount of data about people's movement.

So in this paper, we are going to test different machine learning approaches to predict the human activity that he/she is performing according to the signals from the cell phone of the individual. Smartphones have inbuilt sensors like accelerometer and gyroscope whose input signals will be recorded and the output of each algorithm is a prediction of the type of activity performed. Finally it is going to recognize six different human activities: Lying, Sitting, Standing, Walking, Walking Downstairs and Walking Upstairs.

Contents

INTRODUCTION

In this fast developing modern world, we are coming across various technologies which are changing our lifestyle. Smartphone is playing a great role in this change. Since recent years, many new different kinds of sensors of much smaller size had been developed which can fit inside smartphones and it had revolutionized the way in which we use them. Even though these recent development in sensors had led into creation of new devices, still our smartphone has enough capabilities to cater majority of our demands.

By using our smartphone and it's sensors, we can easily track various kinds of physical activities done by the user. This information can be used for the well-being of the user. By using machine learning and data science algorithms, one can track the physically activities carried out by human with nearly same accuracy as that of other special purpose devices like fitness bands.

Applications

It can play a great role in the day to day life of a common human being. As today in this modern world the use of smartphone is growing day by day so this project can be used on a large scale by common person to track day to day activities. It can also reduce the use of external devices like smart watch etc. The tracking of activities can further help the individual to know about how much physical work one needs to do everyday to lead a healthy life. These days people are more focused on electronic gadgets and social media which is also affecting their personal health. So by using this project one can know to keep himself fit and maintain the personal hygiene.

Motivation

As the size of various electronic sensors is shrinking, more and more different kinds of devices are emerging into the market which is targeted to fulfill specific demands. Wearable devices like fitness band had recently emerged as a health device which senses human physical activities and provides daily status related to the user's health. The smartphone has nearly same type of sensors which are also available in other fitness devices for tracking human activity. The smartphone had not just become cheap; its market penetration is also very high due to which using it as a device to detect human physical activities can be useful to cater the demands of majority of the people in developed as well as developing countries.

Objective

Human activity recognition (HAR) is a highly challenging research topic. It aims at determining the activities of a person based on smartphone sensor. The objective of this project is to describe and evaluate a system that uses smartphone based accelerometer and gyroscope to track the user's physical activity recognition, a task which involves identifying the physical activity a user is performing based on that the person can improve physical activities for better health. It aims in reducing the use of expensive devices like smartwatch.

LITERATURE SURVEY

Principal Component Analysis

Principal Component Analysis is being used to identify the directions in feature space also called as “Principal Components” along which data varies the most. It uses orthogonal transformation to convert a set of observations of correlated variables into set of values of linearly uncorrelated variables which is called principal component. This transformation is defined in such a way that the first principal component has largest possible variance. It is basically used for the purpose of filtering the data by removing those data which have low or zero variance. Removing those data sets will not affect the results due to low variance.

PCA is basically an orthogonal linear transformation that transforms the data of one coordinate system into another coordinate system. Consider we have data containing m features and n observations and it is given into the form of $m \times n$ matrix, let this matrix be A. Let $\vec{\mu}$ be a vector of length m where μ_i is the mean of all the elements of i^{th} column $\vec{x_i}$.

$$\therefore \ \vec{\mu} = \frac{1}{n}(\vec{x_1} + \ldots + \vec{x_n})$$

Let B be $m \times n$ matrix whose i^{th} column is $\vec{x_i} - \vec{\mu}$. Hence $B = [\vec{x_1} - \vec{\mu} \mid \ldots . \mid \vec{x_n} - \vec{\mu}]$. Hence, the covariance matrix S will be

$$S = \frac{1}{n-1} BB^T$$

In covariance matrix S, every element S_{ij} gives covariance between i^{th} and j^{th} feature of the given data matrix. Hence S is a symmetric matrix and can be orthogonally diagonalized. Hence, by performing eigendecomposition of the matrix S, we get eigenvalues $\lambda_1 \geq \lambda_2 \geq \ \ldots \ \geq \lambda_m \geq 0$ and eigenvectors $\vec{u_1}, \vec{u_2}, \ldots ., \vec{u_m}$. These eigenvectors are called principal components of data set. In this basis, the largest eigenvalues correspond to the principal components that are associated with most of the covariability among a number of observed data. Hence, the dimensions which are having comparatively small eigenvalues have very negligible impact on the result of the data analysis and hence they can be removed from the matrix S. This helps in reducing the dimensions of the data to be processed and hence reduces the noise of the data.

Kernelized Principal Component Analysis

Kernelized Principal Components Analysis (kPCA) is an extension of Principal Component Analysis. When the given data of n observations and m features are not linearly separable, then it is necessary to map those observations into N dimensions where $N > m$. Kernelized Principal Component Analysis (kPCA) attempts to identify similar variance-maximizing directions, but does so after the data have been mapped the generated covariance matrix of m dimension into some alternate space, via a feature mapping function of dimensions N. For eg. Consider the data points plotted into a 2-D feature space as shown below

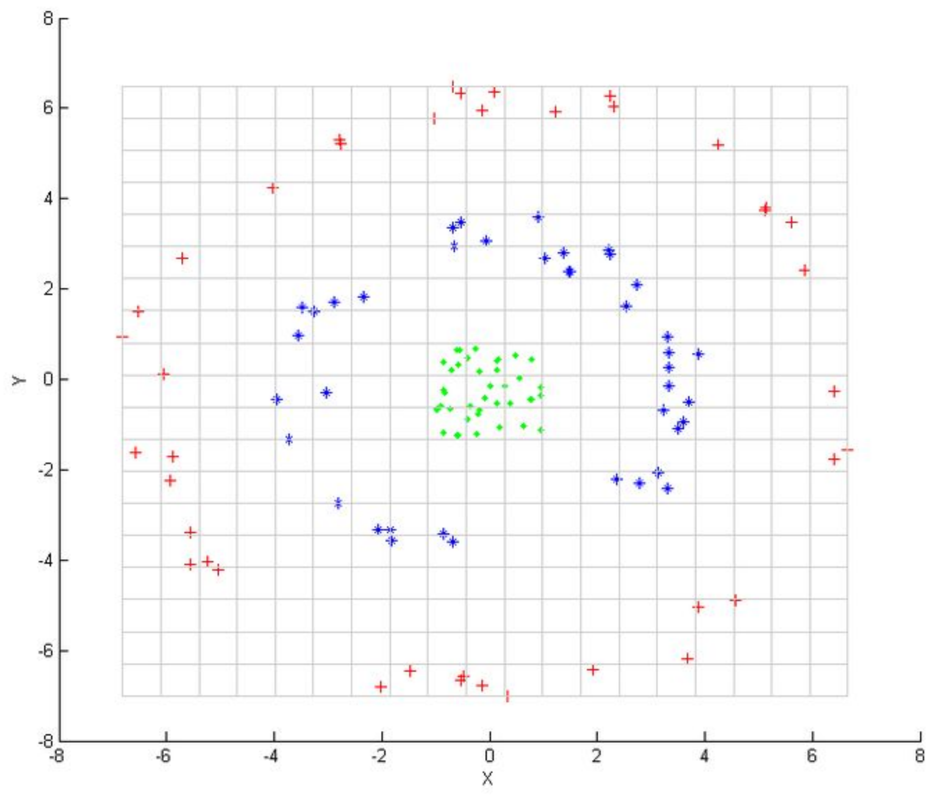

Fig. 1 kPCA input

It is very clear that there is no method to separate the red and blue points using through a single line. Hence the data points of features x_1, x_2 need to be mapped into another space of higher dimension.

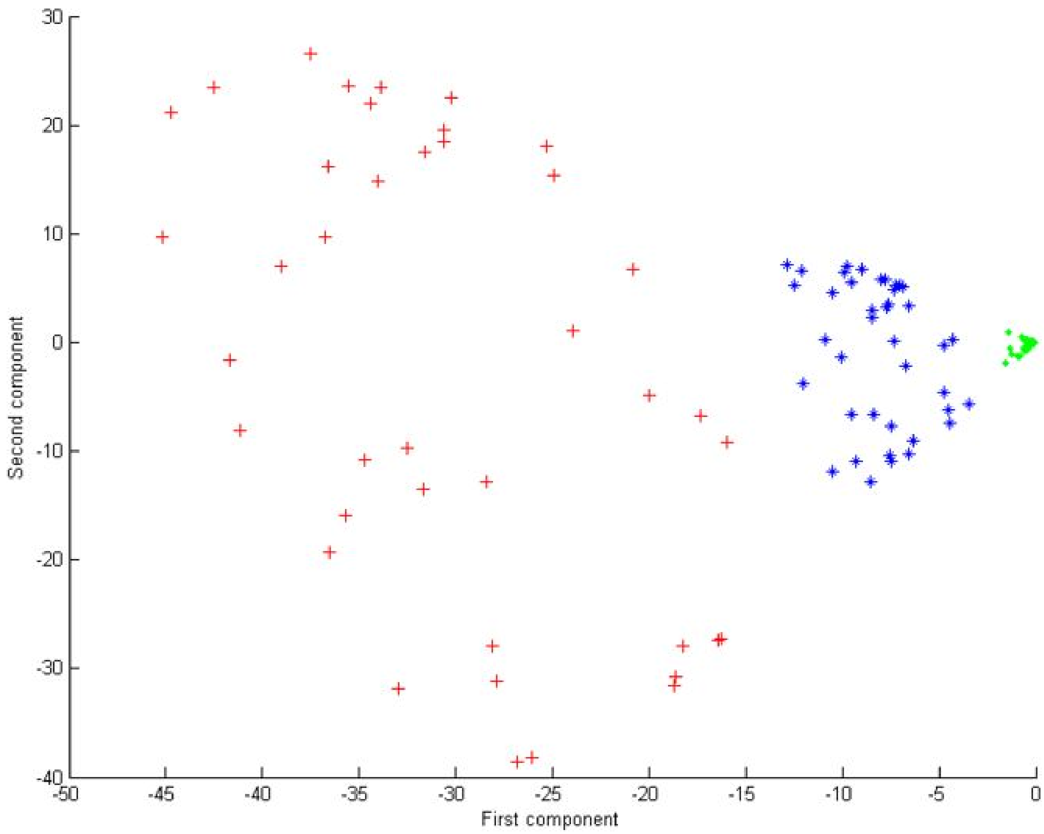

Fig. 2 kPCA output

Let ϕ be the function which maps the point into a space of dimension N.

$$\therefore \ \phi : R^m \to R^N; \ x \to X$$

Hence, covariance matrix $C_F = \frac{1}{n}\sum_{j=1}^{n} \phi(x_j)\phi(x_j)^T$. In covariance matrix C_F ,

$\sum_{i=1}^{n} \phi(x_i) = 0$. Hence, eigenvector $v = \sum_{i=1}^{n} \alpha_i \phi(x^i)$ where $\vec{\alpha} = [\alpha]^n{}_{i=1}$ is a normalized eigenvector of matrix $K = [\phi(x^{(i)})^T \phi(x^{(j)})]_{i,j=1}{}^m = [K(x^{(i)}, x^{(j)})]_{i,j=1}{}^m$. Here $K(x, y)$ is the kernel function associated with ϕ. Hence, score of an example x of the j^{th} kPC gives the projection $v_j{}^T \phi(x) = \sum_{i=1}^{n} \alpha^j{}_i \phi(x^{(i)})^T \phi(x) = \sum_{i=1}^{n} \alpha^j{}_i K(x^{(i)}, x)$. Thus, any kPC score can be computed without explicitly representing $\phi(x)$. Because ϕ can map the data to a high dimensional space, kPCA can often capture nonlinear relationships among the data points, which are missed by the standard, linear version of PCA, described previously.

k-Nearest-Neighbours

k-Nearest-Neighbours (kNN) algorithm classifies a point based on the points which are nearest to it. The predicted probability if a point falls into class j is proportional to the number of the k training points nearest to desired point which are in class j. Mathematically, the probability of a desired point to fall in class j is given as:

$$P(y^{(i)} = j|x = x^{(i)}) = \frac{1}{k} \sum_{q \in N} 1\{y^{(q)} = j\}$$

where, N - set of indices of the k points closest to $x^{(i)}$

Thus a desired point is predicted to belong to the same class as like the other nearest k training points. Further validation set is used to select the number of principal components or kernelized principal components and also the optimal value of k in kNN.

Support Vector Machine

Support Vector Machines are the binary classifiers which attempt to find a vector of parameters α to minimize the regularized loss function

$$J(\alpha)_\lambda = \frac{1}{m}\sum_{i=1}^{m}[1 - y^{(i)}K^{(i)^T}\alpha] + \frac{\lambda}{2}\alpha^T K\alpha$$

where, K = $[K(x^{(i)}, x^{(j)})]_{i,j=1}^{m} = [K^{(1)} \ldots K^{(m)}]$ $\exists$ kernel function $K(x, y) = x^T y$.

As Support Vector Machines work well for binary classifications, it can also be used in many-class classification problems in two ways:

1. One vs One

In this method, the $\binom{L}{2}$ SVMs are trained in such a way that each SVM serves to separate one pair of classes from each other. The desired point is predicted to be in the class that is chosen most frequently, when the point is evaluated using all of the SVMs.

2. One vs All

In this method, the L SVMs are trained in such a way that each SVM serves to separate one class from rest of the classes. The desired point is predicted to be in the class for which its signed distance to the decision boundary is greatest.