ISSN 2321 3361 © 2016 IJESC

Research Article Volume 6 Issue No. 7

Hybrid Machine Learning Algorithm for Human Activity Recognition Using Decision Tree and Particle Swarm Optimization

Rajbir Singh[1], Divya Garg[2]
Department of Computer Science
Universal Group of Institutions, Lalru, Mohali, India

Abstract:
Activity recognition aims to recognize the actions and goals of one or more agents from a series of observations on the agents' actions and the environmental conditions. This paper presents different data mining algorithms for efficiently mine the human activity recognition using smart phones. The overall objective of this research work is to evaluate the performance of the J48, LAD tree and random forest tree based data mining algorithms. Then a new hybrid algorithm will be proposed for the human activity recognition using smart phones. The hybrid algorithm utilizes particle swarm optimization with decision tree algorithm to enhance the results further. The comparative analysis with existing techniques clearly indicates that the proposed technique outperforms over the available techniques.

Keywords: Smartphones, Activity Recognition, Core Techniques, Feature Computation, Data Mining Scheme, Methodology-Particle Swarm Optimization

1. Introduction

Smartphones are ubiquitous and becoming more and more advanced, with ever-growing computing, networking, and sensing powers. This has been adjusting the landscape of people's daily life and has exposed the opportunities for several interesting data mining purposes, which range from wellness and conditioning monitoring, personal biometric trademark, downtown computing, assistive engineering, and elder-care, to indoor localization and navigation, etc. Human task acceptance is really a core making stop behind these applications. It takes the raw senor reading as inputs and anticipates a user's movement activity. Many main flow smartphones are built with different devices, including accelerometers, GPS, light devices, temperature devices, gyroscope, measure, etc. These devices are becoming a rich data supply to evaluate different areas of a user's daily life. The typical actions include strolling, jogging, sitting, etc. Because unobtrusiveness, low/none installment price, and easy-to-use, smartphones are becoming the main software for human task recognition [1].

Figure 1 shows an average method for task acceptance with smartphone sensors. Task acceptance is essential in many true applications. Let's elaborate this utilizing the subsequent examples. To begin with, together branch of human pc connection, it makes the pc actually “smarter”, that's, it may give you the equivalent services based on what the consumer is doing. For example, guess that the telephone finds that the consumer is approximately to leave the space and its weather request suggests so it can water later, an indication can pop-up with a note “Provide an umbrella. It will probably water with a high probability”. Yet another important request of task acceptance methods is in equally indoor and outdoor localizations for making navigation or augmenting the accuracy of context-aware services. Finally, as smartphones become as necessary as tips and the budget for a user's pocket material in these days, the experience acceptance methods may help in supporting life in healthcare. It may help in the prevention of harmful actions, such as folk people's drop detection [2], youth Autism Spectrum Disorder (ASD) recognition in a classroom, etc. It could also aid in a aggressive way. For example, in order to support the consumer form a healthier conditioning habit, the smartphone can send an indication when it finds that she/he has been sitting too long.

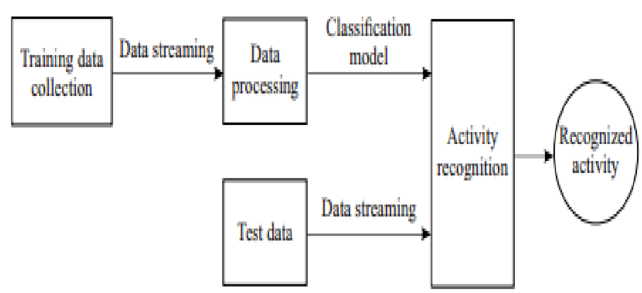

Figure 1 Activity recognition process.

Several recent popular fitness trackers such as Fitbit One [3] are built upon wearable sensors and activity recognition techniques. They track people’s steps taken, stairs climbed, calorie burned, hours slept, distance travelled, quality of sleep, etc.

An activity recognition application takes the raw sensor reading as inputs and predicts a user’s motion activity. Before we dive into the algorithmic details in the next section, let us review these basic concepts in this section.

A. Inputs: Sensors

Detectors are the origin for natural knowledge selection in task recognition. We categorize devices in to three categories:

movie devices, environmental-based devices, and wearable sensors. Video devices are generally cameras that are mounted in the repaired places like the entrance/exit of the public places (to find people's appearance and actions), or in the living areas or bedroom (to track the consumers 'everyday life). Cameras may also be stuck in robots for an even more effective visual knowledge capture. Aesthetic tracking for task acceptance is found in many purposes such as for instance investigator, anti-terrorists, and anti-crime securities along with living saving and assistance. Environmental-based devices are accustomed to identify the people 'relationship with the environment. They're radio based devices like WiFi, Wireless, and the infrared sensors. These devices are usually started in interior places such as company structures or homes. They passively check the look of people at a specific location, or the people relationship with things that are also designed with sensors.

Their restrictions are that
(1) They are able to only be applied to specific repaired places, and
(2) The price for the total deployment of such devices is often very high.

Wearable devices would be the mobile devices that are in small measurement and designed to be utilized on human human body in everyday activities. They are able to history people 'physiological states such as location improvements, moving guidelines, rate, etc. Such devices include accelerometers, microphones, GPS, barometers, etc. All of the mobile devices are equipped on smartphones.

B. Accelerometer

Accelerometer sensors sense the speed function of smartphones. The examining involves three axes whose instructions are predefined as in Fig. 2. The raw data supply from the accelerometer could be the speed of every axis in the devices of g-force. The raw data is represented in a set of vectors: Acci = < xi ; yi ; zi >; (i =1,2, 3,…). A time stamp can be delivered together with the three axes readings. Nearly all of active accelerometers give a consumer program to arrange the sampling frequency so your consumer can select a most useful sampling charge through experiments. Accelerometer has been applied greatly in smartphone sensors based task recognition [5].

C. Compass sensor

Compass is really a old-fashioned software to identify the direction regarding the north-south post of the planet earth by the usage of magnetism. The compass indicator on smartphones works together an identical functionality. Figure 3 shows the compass examining computer screen on a smartphone. The raw data examining from a compass indicator could be the float number between 0 ı and 360ı.

D. Gyroscope

Gyroscope measures the phone's rotation charge by sensing the move, pitch, and yaw movements of the smartphones along the x, ymca, and z axis, respectively.. The raw data supply from a gyroscope indicator could be the charge of the rotation in rad/s (radian per second) around all the three physical axes: Rotationi = < xi ; yi ; zi >; (i =1,2, 3,…). Gyroscope is effective in the navigation programs along with some smartphone games which use the rotation data. In task recognition study, gyroscope is employed to assist the portable direction detection.

E. Barometer

Barometer is among the newest sensors prepared on some advanced smartphones (e.g., Samsung Universe S4 and Google Nexus 4/10). It measures the atmospheric stress of the surroundings that the indicator is put in. The air stress differs with different altitude. Three axes of gyroscope on smartphones. even with areas of the same altitude but having different structures (e.g., thin and large hallways) inside a building. Ergo, measure examining can be used to point the user's place modify in localization related task recognition [6].

2. Outputs: Activities

Actions recognized by the sensor's data can be labeled in numerous ways. For instance, they may be labeled with regards to the complexity of activities. An easy locomotion could possibly be strolling, jogging, strolling downstairs, getting elevator, etc. The complex activities are generally related to a variety of a lengthier period of activities (e.g., getting coach and driving). The activities may possibly just match the movements of certain elements of your body (e.g., writing and waving hand). There are numerous healthcare connected activities, such as slipping, workout, rehabilitations, etc. Location-based activities contain eating, searching, watching movies, etc. Vision-based activities contain causing or entering a place. Actions recognized by an infra-red indicator could be a consumer moving or being still, and the activities recognized by a house helping software could possibly be resting, getting drugs, or performing cleaning. The newest types of Android and iOS equally provide an API to detect a user's current task in among the four activities:

3. Core Techniques

In this part, we review the key data mining processes for task acceptance, including fresh data variety, data pre-processing, feature computation, design teaching, and classification. Accordingly, they're the main steps in the game acceptance process.

A. Raw data collection

How you can collect the fresh data may directly impact the precision in the acceptance period, as well as the adaptively of the classification models. The acceptance design qualified in one subject's data includes a lower precision in realizing yet another subject's activities, and the indicator position and orientation on your body, if distinctive from the way the design is qualified, may decrease the accuracy. The number of devices and all of the devices also impact the acceptance results [7], so does the location where the game is taken.

B. Preprocessing: De-noising and segmentation

Following collecting the fresh data from various devices, the next phase is always to preprocess it before performing any more calculation. One purpose of the info preprocessing is to cut back the sound from the people and the devices themselves. The two filters are employed for data preprocessing. The band-pass filtration is employed to remove the low-frequency acceleration (gravity) that catches the information in regards to the orientation of the indicator regarding the ground data, and the high-frequency indicate components developed by noise. Thus it keeps the medium frequency indicate components developed by powerful human motion. The low-pass filtration aims to remove the sound developed by the powerful human action and to preserve the low-frequency components. Another essential preprocessing stage is data segmentation, which can be to divide the (preprocessed) continuous data streaming in to little segments for feature removal and design training. The segmentation can be labeled in to two types: (a) segmentation with overlapping, and (b) segmentation without overlapping. The fixed-size no-data-overlapping screen segmentation strategy is commonly found in most task acceptance systems. It reduces the computation complexity of segmentation and thus is an excellent method when data is constantly recovered over time. However, the choice of the screen measurement could have a big affect the final acceptance accuracy [8].

4. Feature computation

As in any other data mining responsibilities, extracting the “right” characteristics is critical to the final acceptance performance. For task acceptance, we could get characteristics in equally time and volume domains.

A. Time-domain features

Time-domain features contain the fundamental data of each data part and these of different segments.

- Mean. The mean value of each part in each dimension.
- Maximum, Min. The maximum and minimal prices of each part in each dimension. Common deviation, Variance. The variance (and typical deviation) of each segment.
- Correlation. Correlation is determined between each set of axes of the speed data.
- Signal-Magnitude Place (SMA). SMA is determined because the amount of the magnitude of the three axes speed within the part window[9]. Besides SMA, there occur similar features to combine the three axes readings. Normal Resultant Acceleration is the common of the square origin of the amount of the prices of each axis. Still another similar feature could be the deviation of the amount of the square of speed along three axes.

B. Frequency-domain features

Frequency-domain features explain the periodicity of the signal, which are generally determined on the basis of the FFT.

- Energy. The energy feature is determined because the amount of the squared distinct FFT component magnitudes.
- Entropy. The entropy feature is determined because the normalized data entropy of the distinct FFT parts and it will help in discriminating the activities with the similar energy features [16].
- Time passed between peaks. That feature could be the time taken between the peaks in the sinusoidal waves [10].
- Binned distribution. That feature is essentially the historgram of the FFT. First, determine the range of prices for every single axis (e.g., maximum and minimum). Then, divide this selection in to 10 equal measured bins, and determine the fraction of the prices slipping within each of the bins.

5. Data Mining Scheme

As the dimensionality of features is very high (561 features), which can seriously affect the implementation in real-time on smart phone products, here propose a data theory based position of features because the preprocessing step for this purpose. In this method the features or features are ranked applying data get because the qualification and other simple features are discarded. It's worked remarkably well when compared with other attribute variety techniques, given that in this program context, we're dealing with very high-dimensional datasets, wherever we have to use about half the features to achieve the same level of acceptance performance. We moved out extensive studies with various features ranked by the information theory based position method, including standard conventional Naïve Bayes classifier, Decision tress, random woods, the classifiers predicated on attire learning (random committee), and lazy learning (IBk). Short information on a number of the classifiers examined for this performs is given below:

A. Naïve Bayes Classifier: This classifier is situated about Bayes'theorem and computes probabilities to be able to perform Bayesian inference. The simplest Bayesian strategy, Naive Bayes, is called a special situation of algorithm that requires number adaptation to data streams. The reason being it is founded on watched learning, and it's easy to train the design, and performs well when it comes to reliability and generalization, rendering it a great strategy for baseline comparison [11].

B. K-means Clustering: Clustering is an unsupervised learning strategy, and here the dataset doesn't must have branded data. The situations are arranged and if they're both the exact same or connected to each other they are placed in one group and those which will vary or un-related are positioned in still another group. K-Means is known to be the simplest and the most popular algorithm and predicated on some qualification (Euclidean distance or Ny distance) it analyses if the situations could be clustered without having any previous information about them. Because of simplicity, and their capacity to work on unlabeled data, it is a good choice for baseline reference for analyzing classifier performance. More facts of this classifier strategy are available from 13.

C. Decision Trees: Decision Tree classifiers derive from predictive machine-learning types that determine the dependent variable, or the goal value of a fresh trial from the many attributes of the info available. Here, various attributes are denoted by the interior nodes of your decision tree, and the possible values that the attributes can have in the seen samples is denoted by the limbs involving the nodes. More, the last values (classification) of the dependent variable are represented by the final nodes. The dependent variable indicates the attribute that requires to be predicted, and their value is decided by values of most other attributes. The independent variables in the dataset then type the independent attributes, and they assist in predicting the value of the dependent variable [12].

D. Random Forests: Arbitrary Woods are an ensemble of decision woods, and derive from ensemble learning techniques for classification and regression. They're also looked at as form of a nearest friend predictor, that construct numerous decision woods at instruction time and result the method of the courses because the result class. Arbitrary Woods take to to reduce the problems with high bias and difference by processing a typical, and managing the two extremes. Moreover, Arbitrary Woods have hardly any parameters to song and the majority of the time work well by simply using them with parameter settings collection to default values [13].

E. Random Committee: Arbitrary committee is also a form of ensemble learning strategy and based on the assumption of increasing efficiency by combining classifiers. In this sort of classifier, a different arbitrary number seed is employed for each classifier structure; but, they are based on the same data. After that it computes an average of forecasts created by each one of these specific bottom classifiers, and outputs that normal because the result class [15].

F. Lazy IBk Classifier: Sluggish learners classifiers are based on the principle of learning on travel during classification time, and in reality store working out situations during instruction time. IBk classifier is much like k-nearest friend classifier. As the majority of the learning occurs during classification stage, they are generally slow, and it's possible to speed up the task of choosing the nearest neighbors, using a number of various search algorithms. A linear search approach was employed for that perform, nevertheless the efficiency can also be enhanced by utilizing kD-trees, or protect trees. The length purpose applied was Euclidean distance. The number of neighbors applied was 1, with no weighting predicated on distance from the test instance [16].

6. Literature Survey

Stikic et al. [1] created a multigraph-based semi-supervised learning approach which propagates labels by way of a data that contains equally marked and unlabeled data. Each node of the data corresponds to an instance while every side encodes the characteristics between a couple of nodes as a possibility value. The topology of the data is distributed by the k-nearest neighbors in the feature space. A possibility matrix Z is projected using equally Euclidean range in the feature space and temporal similarity. Once labels have been propagated throughout the data (i.e., all cases are labeled), classification is carried out with a Help Vector Unit classifier that utilizes a Gaussian radial foundation function kernel. The classifier also applied the possibility matrix Z to add knowledge on the amount of self-confidence of every label. The entire accuracy was as much as 89.1% and 96.5% following evaluating two community datasets and having labels for only 2.5% of working out data.

Guan et al. [2] planned an expansion of co-training which does not need the limitations of their predecessor. The device was tested with twenty ambulation actions and compared to three different completely supervised classifiers (the k-nearest neighbors, na¨ıve Bayes and a determination tree). The most problem rate improvement reached by en-co-training was from 17% to 14% —when 90% of working out information weren't labeled. If 20% or more of working out information are marked, the problem rate big difference between en-co-training and the best completely supervised classifier does not surpass 1.3%.

Ali et al. [3 executed a Multiple Eigenspaces (MES) approach on the basis of the Key Element Analysis coupled with Concealed Markov Models. The device was created to identify finger expressions with a laparoscopic gripper tool. The people wore an indicator glove with two biaxial accelerometers sampling at 50Hz. Five different turn and interpretation actions from the individual's hand were acknowledged with as much as 80% of accuracy. This technique becomes difficult to analyze since no details are presented on the total amount of marked information or the evaluation procedure.

Huynh et al. [4] combined Multiple Eigenspaces with Help Vector Products to identify seven ambulation and daily activities. Eleven accelerometers were added to people'legs, legs, elbows, shoulders, arms, and hip. The quantity of marked instruction information various from 5% to 80% and the general accuracy was 88% to 64%, respectively. Their approach also outperformed the completely supervised na¨ıve Bayes algorithm, which was applied as a baseline. Still, actions such as banging hands, ascending steps and descending steps were usually confused.

Blanke et al. [5] provide an overview with this topic and propose an answer through a few layers of inference.The actions we have observed so far are very simple. In fact, many can participate more complicated workouts or behaviors. Imagine, like, the situation of instantly knowing when a person is enjoying tennis. Such task is constructed by a few instances of walking, working, and sitting, among others, with specific rational collection and duration. The recognition of these blend actions from some atomic actions could certainly enrich situation recognition but, at once, provides additional uncertainty.

Helaoui et al. [6] the assumption that an individual just performs one task at a time is true for basic ambulation actions (e.g., walking, working, lying, etc.). Generally, individual actions are relatively overlapping and concurrent. An individual could possibly be walking while discovering their teeth, or watching TV while having lunch. Because just several performs have been noted in this area, we predict good research options in this field.

Ravi et al. [7] conducted a relative study in terms of classifying seven activities. They examined an inclusive group of classifiers: Enhancing, Bagging, Plurality voting, putting with OrdinaryDecision Woods (ODTs), and Putting with Meta-Decision tress (MDT). Most of the base-level classifiers (e.g., NB, SVM, KNN, Choice Tree, and Choice Table) and most of the meta-level classifiers mentioned previously were examined with four different test settings. The Plurality Voting classifier outperforms different classifiers generally and hence is advised as the best classifier for task recognition from an individual accelerometer.

P. Viola et al. [8] One of many meta-level voting classifier, Bag-of-Features (BoF). The BoFis applied to build task recognition versions using histograms of simple representations, and then validated experimentally the effectiveness of the BoF-based framework for knowing eight task classes. Their framework for a long-term task recognition program based on accelerometer data. Meta-level classifiers may also be utilized in feature selection.

A. Reiss et al.[9] An altered version of AdaBoost is useful for feature selection. Given the most amount of functions that the experience acceptance program seeks to utilize, it automatically prefers the absolute most discriminative sub-set of functions and employs them to learn an set of discriminative static classifiers for task recognition. Using the information how comfortable the poor learners are to calculate the type of situations, the algorithm allows the voting weights of the poor learners to alter in reaction (decrease or increase in the weight) to the confidence. Ergo the new instance is categorized predicated on measured voting.

Deng et al.[10] proposed a cross-person task acceptance model to eliminate the effect of individual sensitivity. The model teaching stage consists of two components: The first model is experienced off-line and the adaptive model is current online. For new consumers in the online period, the algorithm selects those high comfortable acceptance benefits to be able to make the new teaching dataset. Based with this new teaching dataset, the algorithm may upgrade the acceptance model to alleviate the topic sensitivity.

Park et al. [11] revealed a tool build classification technique on the cornerstone of the regularized kernel algorithm. It provides a way of how exactly to determine the sma. The orientation sensitivity by utilizing still yet another warning: magnetometer. The magnetic subject warning materials the magnetic vector along three axes of the device's coordinate program in the orthogonal directions. Ergo, perhaps it's used to get the units ‘azimuth angle. Your accelerometer studying could be converted to the the world corresponding axes reading.

Guan et al. [12] prolonged the co-training method by set and proposed the en-co-learning method. For task acceptance, the collection of unlabelled data is straightforward and requires near zero consumers'effort. In place of using two different labelled datasets for the first teaching, the en-co-learning semi-supervised learning method employs only 1 labelled dataset and three different classifiers. In this manner, it bypasses the frustrating assurance computation and eliminates one labelled dataset.

Lockhart et al.[13 categorized the purposes of mobile task acceptance relating with their targeted helpful matters: (1) application for the end consumers such as conditioning tracking, wellness monitoring, fall recognition, behaviour-based context-awareness, home and function automation, and self-managing program; (2) purposes for the 3rd events such as targeted advertising, study platforms for the information series, corporate management, and accounting; and (3) purposes for the crowds and organizations such as cultural marketing and activity-based crowd-sourcing.

Si et al.[14] proposed a model program that gives aged persons the individualized advice to accomplish lifestyle actions by learning their living styles using mobile sensors. An structure on the smartpphone is produced with the purpose of consumers'fall detection. Activity acceptance and monitor devices could help elders in a positive way such as life schedule memory (e.g., using medicine), living task monitoring for a remote robotic assist.

Xing Su et al. [15] The ubiquity of smartphones along with their ever-growing research, marketing, and detecting powers have been changing the landscape of people's daily life. And others, task recoginition, which takes the fresh sensor reading as inputs and predicts a user's movement task, is now an energetic study place in recent years. It's the key developing block in many high-impact purposes, which range from wellness and conditioning monitoring, personal biometric signature, urban research, assistive technology, and elder-care, to indoor localization and navigation, etc.

Girija Chetty et al. [16] Computerized task acceptance programs intention to capture their state of the user and their setting by exploiting heterogeneous devices, and permit continuous monitoring of numerous physiological signs, wherever these devices are attached with the subject's body. This is immensely of good use in healthcare purposes, for automatic and wise daily task monitoring for aged people. In that paper, we provide book data analytic scheme for wise Human Activity Acceptance (AR) using smartphone inertial devices predicated on information theory centered feature rank algorithm and classifiers predicated on random woods, set learning and lazy learning. Intensive experiments with a publicly accessible database1 of individual task with cell phone inertial devices show that the proposed method may certainly result in development of wise and automatic realtime individual task monitoring for eHealth application situations for aged, disabled and individuals with particular needs.

7. Methodology

In this section we have mentioned the details about the implementation of the proposed protocol and the results found after the implementation.

Particle swarm optimization algorithm was first proposed by John Holland [1975], are nature-inspired optimization strategy that can be advantageously used for several optimization problems. Particle swarm optimization algorithm begins with an original population of individuals generated at random. Each individual in the population presents a possible solution to the issue under concern and every individual is considered for fitness with regards to the optimization task to be solved. The persons evolve through successive iterations, called generations. Throughout each era, every individual in the population is considered using fitness function. Variety within the population is performed in a fitness-proportionate way: the more fit an individual, the much more likely it will be picked for replica into these time [Back et aI., 2000], ergo increasing successive generations. The population of the following era is done through particle swarm optimization operators, specifically G-best and mutation. G-best needs two people and yields two new people while velocityalters one particular to make a simple new offspring. The process remains before termination situation is satisfied. For recommender systems, an individual designs are changing over amount of time in gentle of the most up-to-date ratings.

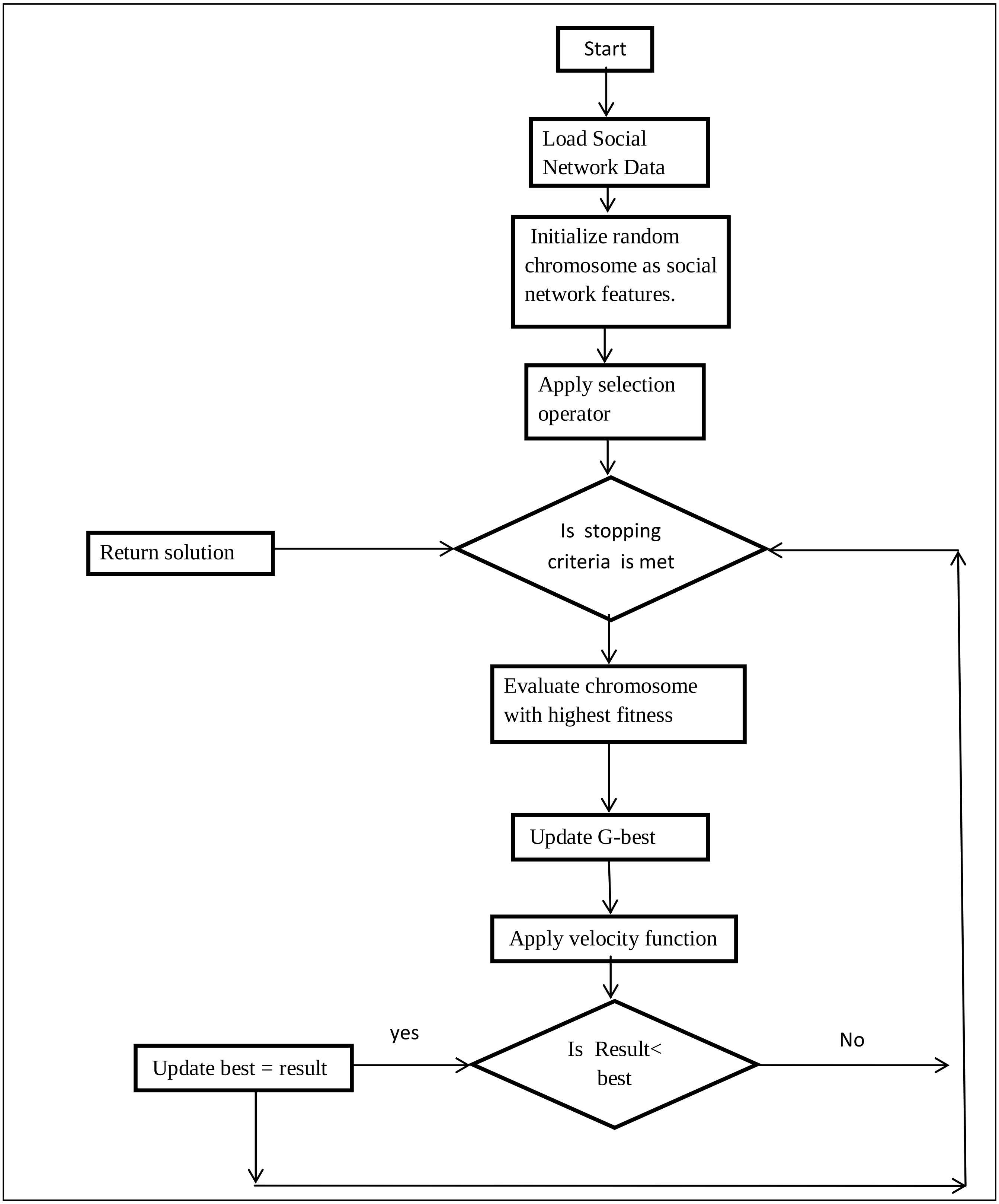

Figure 2. Flowchart of proposed technique

In that direction, individual choices ought to be discovered and prioritized to reveal the actual user's preferences. The particle swarm optimization algorithm has been successfully used to obtain the fat for every single item that tells its priority within an individual page [Min & Han, 2005; Ujjin & Bentley, 2004]. But, an individual page that is employed by Ujjin & Bentley [2004] adds category frequencies to each item. Thus an additional computation is necessary besides the large handling time of standard collaborative RSs. That additional computation is eliminated in the job of Min & Han [2005] by associating each item having its possess fat utilizing a modified Pearson link coefficient. But, the items of any collaborative RS are increasing day-to-day thus any energy to capture the weights of person choices at the item level be seemingly unrealistic and time consuming. Recently, a model-based RS for on the web looking industry is proposed by Betty &Ahn [2008]. The model is made using K-means clustering whose original seeds are optimized by GA.

Particle swarm optimization Algorithm Method of Recommender Programs The decision of similarity function is an essential issue for the device and depends heavily on the issue itself. The main difficulty in profile matching computation applying Formulae is due to the inclusion of all of the ranked films for a certain user in the profile directly. In line with the lightweight user product that difficulty gets eliminated by using just a fixed number of functions for the users regardless of the number of films they have rated. Now a suitable likeness function is required to match users. Euclidean range function [Han &Kamber, 2006] is utilized in our function that is given by:

$$\text{dis}(v_x, v_y) = \sqrt{\sum_{i=1}^{18} \left(\left(JIM_{v_x}(j_i), JIM_{v_y}(\llbracket J_i)) \rrbracket^2 \right) } \quad \dots (1)$$

Let us named the RS that uses Formula (4.l) for similarity computation because the Euclidean-based RS (EBRS). The planned person design contains 18 characteristics contributing similarly throughout likeness computation in the EBRS. That doesn't capture the real life situation where each person places various weights to various features. These weights are subject to alter with time and changing tastes of every user. An effective learning device to capture these weights is required. Ujjin& Bentley [2004] applied the major search to get person things for variety wavelengths tailoring the recommendation process to the tastes of each individual person but with a extended profile. Every person places an alternative concern on each feature, which can be known as feature weight. In order to apply really individualized RS, these weights need to be grabbed and fine-tuned to reveal each user's preference. By imposing characteristics'weights to Formula, particle swarm optimization algorithm can be utilized to get these weights resulting in a particle swarm optimization-algorithm-based RS (GBRS). For this process the weighted Euclidean range purpose [Han &Kamber, 2006] is used for likeness computation:

$$\text{dis}(v_x, v_y) = \sqrt{\sum_{i=1}^{18} N_i \times \left(\left(JIM_{v_x}(j_i), JIM_{v_y}(\llbracket J_i)) \rrbracket^2 \right) } \quad \dots (2)$$

GBRS usmg the small consumer design retains memory-based RS accuracy, and model-based RS scalability. The user design increases the internet procedure for generating a set of like-minded people within which a memory-based collaborative RS is moved out. These subsections offer a short introduction to GA and the conditioning function. Within our test, GA adapts the function weights to recapture an individual thing for different features [Ujjin& Bentley, 2004]. The function weights of client V a are displayed as a couple of weights, weight(va) = (NI '... , D I 8), wherever d is how many features. The genotype of Ni is just a real-valued numbers. When the fat for almost any function is zero, that function is ignored. That permits function variety to be flexible to each user's preference.

A particle swarm optimization algorithm (GA) procedures a population of competitive prospect alternatives represented by some chromosomes, 0. Each chromosome in the people represents a probable means to fix an optimization issue. An over-all platform of GA [Michalewicz, 1992] is defined below, wherever pet) indicates the people at era s:

Procedure: Particle swarm optimization Algorithm
begin
s = 0;
initialize M(s);
evaluate M(s);
while (not termination condition) do
begin
s = s+ 1;
select M(s) from M(s - 1);
alter M(s);
evaluate P(t);
end
end

Subsequent Houck et al. [1995], a chromosome representation describes every individual in the populace of interest that determines how the problem is structured in the GA and also determines the particle swarm optimization operators which can be used. Each chromosome is composed of a routine of genes from a certain alphabet. An alphabet could include binary numbers, flying stage figures, integers, designs, etc. In Holland's original design [Holland, 1975], the alphabet was limited to binary digits. However, it's been shown that more natural representations are more efficient and generate greater solutions. One of use representation of chromosome is just a real-valued representation which involves genes from an alphabet of flying stage figures with values within top of the and lower bounds of every variable. The real-valued representation moves the problem closer to the problem representation and offers larger detail with more regular benefits across replications. The chromosomes search place, fie, is:

$$\Omega_{\downarrow}\theta = \{\theta \epsilon M^{\uparrow}R \dashv | \, \theta_{\downarrow}1^{\uparrow}\mathbf{m}\text{in} \; \llbracket \le \theta \rrbracket \, {}_{\downarrow}1 \le \theta_{\downarrow}1^{\uparrow}max, \dots\dots, \theta_{\downarrow}r^{\uparrow}min \le \theta_{\downarrow}r \le \theta_{\downarrow}r^{\uparrow}max\} \quad \dots (3)$$

All the genes ()i in the chromosome will undoubtedly be changed in the confined place fie during the particle swarm optimization operations. Top of the and lower bounds of ()i must certainly be written by designer. After a developed chromosome by particle swarm optimization procedures moves beyond fie, then a original chromosome will undoubtedly be retained. Initially, a citizenry of people IS developed and then every individual is evaluated for fitness with regards to the optimization task to be solved. Selection within the population is performed in a fitness-proportionate way: the more fit a person, the more likely it is to be opted for reproduction into the next generation. The great freshly developed people (if exist) replace the present bad people to create the new citizenry for the following generation. GA terminates if the required price of fitness is reached or the most quantity of generations is elapsed.

Fitness Function

Locating a suitable Function purpose is often a complex problem for GA applications [Goldberg, 1989] and for the application being mentioned for the reason that chapter it's not a small job since every set of loads in the GA populace is used for similarity computation and so the RS needs to be re-run on the entire repository for every single new set of loads in order to computes their fitness [Ujjin& Bentley, 2004]. An unwanted (good) set of loads might (should) develop a bad (good) town set of men and women for the effective customer, and ergo bad (good) recommendations. One method to obtain the fitness purpose is by reformulating the issue as a administered understanding task. For this specific purpose the particular ratings of the effective customer are arbitrarily split into two disjoint versions, check always ratings selection (66%) and knowledge ratings selection (34%). To find the fitness record for the evolved set of loads, the RS should be function and the believed ratings for every single movie in exercising ratings selection should be computed. The average of the differences between the particular and believed ratings of all films in exercising ratings selection is used while the fitness record for that set of weights.

$$\text{Fitness}(v_a) = \frac{1}{|\delta_a^{PT}|} \sum_{j=1}^{|\delta_a^{PT}|} |s_{a,j} - ps_{a,j}| \quad \dots \ (4)$$

wherever $[\![|\delta]\!]_{\downarrow a}{}^{\uparrow PT}|$ could be the cardinality of working out ratings set equivalent to a given productive user.

8. Parameters Evaluation

A. Correlation Coefficient(r): The correlation coefficient is just a evaluate that decides the amount to which two variables movements are associated. The range of values for the correlation coefficient is -1.0 to 1.0. If a calculated correlation is greater than 1.0 or less than -1.0, an error has been made. A correlation of -1.0 suggests a perfect negative correlation, while a correlation of 1.0 suggests an ideal good correlation.

$$N_{yx} = \frac{\mathrm{cov}(p_y, p_x)}{\delta_y \delta_x} \quad \dots \ (5)$$

$$r = (s(\sum yx) - (\sum y)(\sum x))/\sqrt{([n\sum [y^{\uparrow}2 - (\sum y^{\uparrow}(.@)2)]][s\sum [x^{\uparrow}2 - (\sum [x)]^{\uparrow}2)]])}$$

.... (6)

While the relationship coefficient procedures a degree to which two parameters are connected, it just procedures the linear relationship involving the variables. Nonlinear associations between two parameters cannot be grabbed or expressed by the relationship coefficient.

A price of just 1.0 suggests there is a perfect good relationship between the two variables. For a confident increase in one variable, there is also a confident increase in the 2nd variable. A price of just -1.0 suggests there is a perfect bad relationship between the two variables. That reveals the parameters relocate opposite instructions; for a confident increase in one variable, there is a reduction in the 2nd variable. If the relationship is 0, that just suggests there is number relationship between the two variables. The effectiveness of the relationship differs in degree on the basis of the value of the relationship coefficient. For example, a value of 0.2 suggests there is a confident relationship between the two parameters, but it's weak.A correlation greater than 0.8 is generally identified as strong, whereas a correlation than 0.5 is generally identified as weak.

The most frequent formula is called the Pearson product-moment correlation. It is determined by first calculating the covariance of the 2 variables in question. Next, the standard deviations of each variable should be calculated. To find the link coefficient, get the covariance and split it by the product of the 2 variables typical deviations. Typical change is really a measure of the dispersion of information from its average. Covariance is really a measure of how two variables modify together, but its magnitude is unbounded so it's difficult to interpret. By dividing covariance by the product of the 2 typical deviations, a normalized version of the statistic is calculated. This is the correlation coefficient. These values can vary based upon the "type" of data being examined. A study utilizing scientific data may require a stronger correlation than a study using social science data.

B. Coefficient of determination, r ^{2}: The coefficient of determination of a linear regression model is the quotient of the variances of the fixed values and seen values of the dependent variable. If we denote z_i since the seen values of the dependent variable, $\bar{z}$ as its mean, $\hat{z}_i$ and since the fixed price, then a coefficient of perseverance is:

$$r^2 = \frac{\sum \overline{\hat{z}_i - (z)}^2}{\sum \overline{z_i - (z)}^2} \quad \dots (7)$$

The coefficient of determination, r ^{2}, is advantageous because it provides the percentage of the difference (fluctuation) of just one variable that's expected from one other variable. It is a evaluate that we can establish how certain you can be in creating predictions from a certain model/graph.

The coefficient of determination could be the relation of the discussed deviation to the full total variation. The coefficient of determination is such that $0 < r^2 < 1$, and denotes the energy of the linear association between *x* and *y* The coefficient of determination shows the percent of the data that's the best to the type of most readily useful fit. For example, if *page1=46* = 0.922, then r^2 = 0.850, meaning that 85% of the full total deviation in *y* may be discussed by the linear connection between *x* and *y* (as identified by the regression equation). One other hand 15% of the full total-deviatioin y remains-unexplained. The coefficient of determination is a measure of how effectively the regression point shows the data. If the regression point moves exactly through every place on the spread plot, it would manage to describe all the variation. The further the point is from the items, the less it can explain.

C. RMSE (Root Mean Square Error): The root-mean-square error (RMSD) or root-mean-square error (RMSE) is really a commonly used way of measuring the differences between prices (sample and population values) believed by a type or an estimator and the prices really observed. The RMSD presents the taste normal change of the differences between believed prices and observed values. These specific differences are called residuals when the calculations are conducted over the data taste that has been useful for opinion, and are called prediction mistakes when computed out-of-sample. The RMSD provides to blend the magnitudes of the mistakes in forecasts for different situations in to a simple way of measuring predictive power. RMSD is a great calculate of accuracy, but simply to examine forecasting mistakes of various types for a specific variable and perhaps not between parameters, since it is scale-dependent.

The RMSE is right interpretable when it comes to rating models, and therefore is just a greater measure of goodness of fit than a co-relation coefficient.The RMSD of an estimator $\hat{\delta}$ with respect to an estimated parameter δ is explained whilst the square root of the mean sq problem:

$$\text{RMSE}(\hat{\delta}) = \sqrt{MSE(\hat{\delta})} = \sqrt{H\left((\hat{\delta} - [\delta])^2\right)} \quad \text{... (8)}$$

For an unbiased estimator, the RMSD could be the square root of the variance, known as the standard deviation.The RMSD of predicted values $\hat{l}_t$ for times t of a regression's dependent variable l_t is computed for d various forecasts while the square root of the suggest of the sections of the deviations:

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^{t}(\hat{l}_t - [l_t])^2}{t}} \quad \text{... (9)}$$

D. Accuracy: "Accuracy" blows here. It is not to be confused with the music. Detail is an explanation of arbitrary mistakes, a calculate of mathematical variability.**Accuracy** has two meanings:

1. more typically, it is an explanation of *systematic mistakes*, a calculate of mathematical error;
2. instead, the ISO describes accuracy as describing equally types of observational problem above (preferring the term *trueness* for the most popular meaning of accuracy).

$$\text{Accuracy} = \frac{\text{number of true positives} + \text{number of true negatives}}{\text{number of true positives} + \text{false positive} + \text{false negatives} + \text{true negatives}}$$

A dimension program may be appropriate although not accurate, accurate although not appropriate, neither, or both. For example, if a test includes a systematic problem, then increasing the test measurement typically increases detail but does not improve accuracy. The result would have been a regular yet inaccurate chain of benefits from the problematic experiment. Reducing the systematic problem improves accuracy but does not modify precision.

A dimension program is considered legitimate when it is both appropriate and accurate. Related terms include error (non-arbitrary or focused consequences the effect of a factor or facets unrelated to the independent variable) and problem (random variability).In mathematical evaluation, accuracy can be the nearness of a formula to the actual value; while precision could be the resolution of the representation, generally described by the amount of decimal or binary digits.

E. Execution time:

it represents the overall time which algorithm takes to do job successfully. But analysis shown that proposed techniques takes little more time than Decision tree algorithm.

9. Experimental Results

Table 1 and Figure 3 showing the analysis of Root mean square error (RMSE) as it is known that needs to be minimized. It is clearly shown in the graph that the RMSE of the proposed technique is lower among other algorithms.. Therefore proposed technique is effective than Decision tree, Linear model and neural network.

Table 1: RMSE readings

Iteration	DT	LM	NN	Proposed
1	13467.83	16341.62	16378.60	13141.08
2	13486.02	16324.79	16354.42	13182.22
3	13483.3	16332.70	16339.65	13340.65
4	13480.03	16330.80	16341.21	13201.65
5	13477.93	16345.88	16340.25	13139.94
6	13479.21	16348.25	16354.20	13143.18
7	13486.33	16330.23	16364.22	13209.09
8	13482.37	16349.20	16342.63	13067.58
9	13470.9	16350.02	16350.23	13098.56
10	13480.3	16347.30	16349.27	13184.21
11	13451.69	16356.77	16378.83	13337.51
12	13460.25	16352.32	16380.20	13117.19
13	13455.23	16350.27	16379.24	13206.41
14	13486.14	16354.30	16376.55	13234.65
15	13480.84	16341.71	16382.03	13190.67

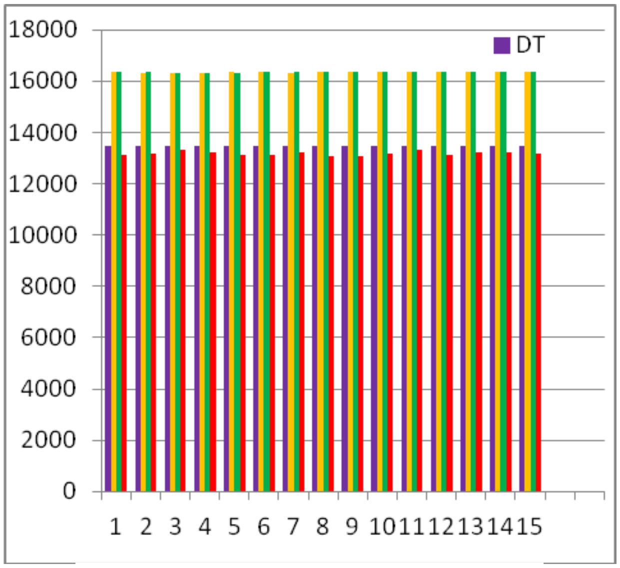

Figure 3. Comparative analysis of RMSE

Table 2 and Figure 4 showing the analysis of accuracy, correlation coefficient and coefficient of determination (R). Each parameter needs to be maximized . It is clearly shown in the fig that the parameters accuracy, correlation and R of the proposed technique are higher among other algorithms. Therefore proposed technique is effective than Decision tree, Linear model, and neural network.

Table 2. Correlation, Accuracy and R readings

Parameters	DT	LM	NN	Proposed
Correlation	0.93	0.91	0.91	0.97
Accuracy	0.89	0.77	0.76	0.963
R	0.86	0.83	0.83	0.94

Table 3 and Figure 5 showing the comparative analysis of execution time among the proposed and other techniques. The graph clearly shows that the proposed technique is not so effective w.r.t execution time.

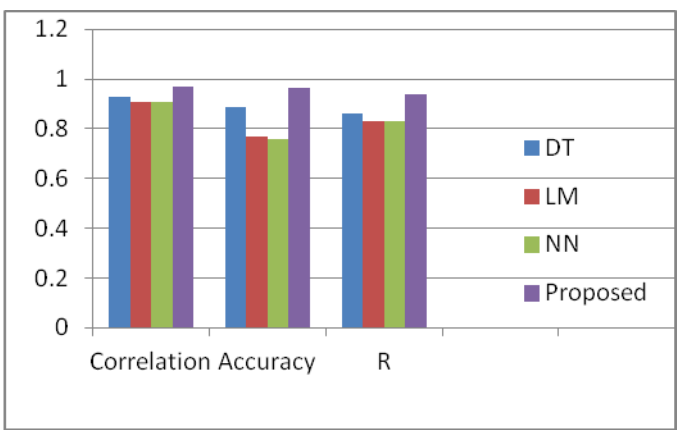

Figure 4. Comparasion analysis of Correlation, Accuracy and R parameters of techniques

Table 3. Execution time readings

Iteration	DT	LM	NN	Proposed
1	3.28	1.31	3.78	2.28
2	2.17	1.53	3.14	2.25
3	2.40	1.83	3.20	2.22
4	3.10	1.86	3.74	2.31
5	2.89	1.59	3.66	2.41
6	2.92	1.36	3.59	2.27
7	3.16	1.63	3.67	2.33
8	3.02	1.77	3.47	2.53
9	2.69	1.80	3.30	2.27
10	2.72	1.82	3.21	2.16
11	2.19	1.50	3.38	2.17
12	2.80	1.55	3.40	2.12
13	2.20	1.52	3.37	2.25
14	2.35	1.70	3.41	2.23
15	2.41	1.43	3.27	2.33

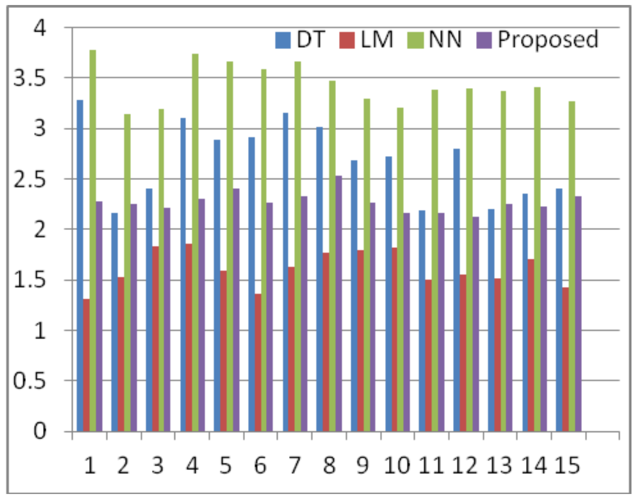

Figure 5. Execution time analysis

Conclusion

Sensor-based activity recognition integrates the emerging area of sensor networks with novel data mining and machine learning techniques to model a wide range of human activities. Mobile devices (e.g. smart phones) provide sufficient sensor data and calculation power to enable physical activity recognition to provide an estimation of the energy consumption during everyday life. Sensor-based activity recognition researchers believe that by empowering ubiquitous computers and sensors to monitor the behavior of agents (under consent), these computers will be better suited to act on our behalf.

This paper presents different data mining algorithms for efficiently mine the human activity recognition using smart phones. This research work evaluates the performance of the J48, LAD tree and random forest tree based data mining algorithms. The new proposed hybrid algorithm for the human activity recognition using smart phones utilizes particle swarm optimization with decision tree algorithm to enhance the results further. The comparative analysis with existing techniques clearly indicates that the proposed technique outperforms over the available techniques.

References

[1] M. Stikic, D. Larlus, S. Ebert, and B. Schiele, “Weakly supervised recognition of daily life activities with wearable sensors,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 12, pp. 2521– 2537, 2011.

[2] A. Ali, R. King, and G.-Z. Yang, “Semi-supervised segmentation for activity recognition with multiple eigenspaces,” in International Summer School and Symposium on Medical Devices and Biosensors, pp. 314–317, 2008.

[3] D. Guan, W. Yuan, Y.-K. Lee, A. Gavrilov, and S. Lee, “Activity recognition based on semi-supervised learning,” in 13th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, pp. 469–475, 2007.

[4] T. Huynh and B. Schiele, “Towards less supervision in activity recognition from wearable sensors,” in 10th IEEE International Symposium on Wearable Computers, pp. 3–10, 2006.

[5] R. Helaoui, M. Niepert, and H. Stuckenschmidt, “Recognizing interleaved and concurrent activities: A statistical-relational approach,” in IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 1–9, 2011.

[6] C. Elkan, “The foundations of cost-sensitive learning,” in Seventeenth International Joint Conference on Artificial Intelligence, pp. 973–978, 2001.

[7] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, Activity recognition from accelerometer data, AAAI, vol. 5, pp. 1541-1546, 2005.

[8] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, in Proc. 17th Int. Computer Vision and Pattern Recognition Conf., Kauai, USA, 2001.

[9] A. Reiss, Personalized mobile physical activity monitoring for everyday life, PhD dissertation. Dept. CS. Technical University of Kaiserslautern, Kaiserslautern, German, 2014.

[10] W.-Y. Deng, Q.-H. Zheng, and Z.-M. Wang, Cross-person activity recognition using reduced kernel extreme learning machine, Neural Networks, 2014.

[11] J.-G. Park, A. Patel, D. Curtis, S. Teller, and J. Ledlie, Online pose classification and walking speed estimation using handheld devices, in Proc. 14th Int. Ubiquitous Computing Conf., Pittsburgh, USA, 2012, pp. 113-122.

[12] D. Guan, W. Yuan, Y.-K. Lee, A. Gavrilov, and S. Lee, Activity recognition based on semisupervised learning, in Proc. 13th Int. Embedded and Real-Time Computing Systems and Applications Conf., Daegu, Korea, 2007, pp. 469-475.

[13] J. W. Lockhart, T. Pulickal, and G. M. Weiss, Applications of mobile activity recognition, in Proc. 14th Int. Ubiquitous Computing Conf., Seattle, USA, 2012, pp. 1054-1058.

[14] H. Si, S. J. Kim, N. Kawanishi, and H. Morikawa, A context-aware reminding system for daily activities of dementia patients, in Proc. 27th Int. Distributed Computing Systems Conf., Toroto, Canada, 2007, pp. 50- 50.

[15] Su, Xing, Hanghang Tong, and Ping Ji. "Activity recognition with smartphone sensors." *Tsinghua Science and Technology* 19.3 (2014): 235-249.

[16] Chetty, Girija, Matthew White, and Farnaz Akther. "Smart Phone Based Data Mining for Human Activity Recognition." *Procedia Computer Science* 46 (2015): 1181-1187.