Impementation And Comparative Analysis of Human Activity Recognition Recommender System

Mahak Bhatia, Dhiraj Khurana
Department of Computer Science and Engineering,UIET
Maharishi Dayanand Univerity, Rohtak, Haryana

Abstract-

With the rapid growth of digital activity formats, managing and searching of activities have become important. Designing a personalised human activity recommender is complicated, and it is challenging to thoroughly understand the users' needs and meet their requirements. This work introduces to the field of recommender system and represents popular approaches in recommending human activity. The ever-lasting demand of devices associated with accelerometers has given the chance to catch the semantic parts of human action and accordingly prompts enhance the client encounters with conduct based proposals. These capacities depend vigorously on the exactness of human movement acknowledgment, and in this manner genuine applications that utilization.

This paper also presents about the previous researches that have been done by the researchers in the field of recommendation.

Keywords- Recommendation System, Human Activity Learning, MATLAB etc.

I. INTRODUCTION

Individuals find articulating what they need hard, however they are great at remembering it when they see it. This understanding has prompted the usage of importance input, where individuals rate site pages as fascinating or not intriguing and the framework endeavors to discover pages that match the "intriguing", positive cases and don't coordinate the "not intriguing", negative illustrations. With adequate positive and negative cases, present day machine-learning systems can group new pages with great exactness; at times message arrangement precision surpassing human capacity has been illustrated. Catching client inclinations is a hazardous undertaking. Just asking the clients what they need is excessively meddling and inclined, making it impossible to blunder, yet checking conduct unpretentiously and afterward finding important examples is troublesome and computationally tedious. Catching exact client inclinations is in any case, a basic undertaking if the data frameworks of tomorrow are to react powerfully to the changing needs of their clients.

Add up to data over-burden turns out to be progressively extreme in the cutting edge times of ubiquitous broad communications and worldwide correspondence offices, surpassing the human observation's capacity to dismember important data from insignificant. Thus, since over 64 years huge research endeavours have been endeavouring to imagine mechanized separating frameworks that give people attractive and important data as it were. Amid the most recent multi decade, recommender frameworks have been picking up energy as another proficient methods for decreasing many-sided quality while hunting down important data. Recommenders plan to furnish individuals with recommendations of items they will acknowledge, in light of their past inclinations, history of procurement, or statistic data or different sorts of data.

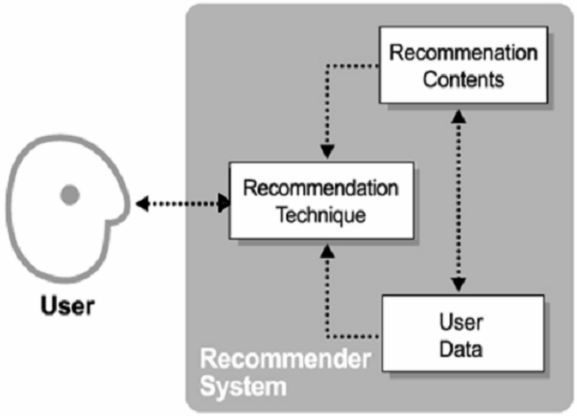

Figure 1: Constitution of Recommendation System[13]

The goal of gathering client data is to manufacture a profile that depicts a client interests, part in an association, privileges, and buys or other data. A recommender framework comprises of three components as appeared in figure (1). Numerous suggestion substance which are displayed to clients must be made. At that point, clients' inclinations or social information on these substance must be assembled. At long last, it needs to pick sort of prescribe at particle strategy about how to examination these client information and select the ideal substance to every client.

In this paper, it contemplates the idea of proposal framework. Further, in area II, it gives the related work of different analysts. In Section III, It characterizes rudiments of system utilized. Results are clarified in area IV. At last, conclusion is clarified in Section V.

II. RELATED WORK

Diego Sánchez-Moreno et.al [1] exhibited a suggestion strategy in light of playing coefficients was utilized to decide how much a client is dark sheep that is when substance and rating data isn't accessible or when next to no data is accessible. This technique took certain data to acquire client inclinations and to describe clients and things to be suggested. K-closest neighbor technique was utilized. Client based and thing based K-NN strategies were tried utilizing both cosine and pearson comparability measures. They demonstrated that the proposed strategy was superior to anything other CF approaches. Be that as it may, the strategy did not give any answer for the versatility issue.

Issac Caswell et. al. [2] clarified that melodic structure is unique in relation to melodic classifications. Two bunching techniques were utilized: the k-Means calculation as an unsupervised learning strategy for seeing how to group chaotic music and a Markov Chain Model which decides relative probabilities for the following note's in all likelihood esteem, and assess our calculations' exactness in anticipating right class. The highlights were extricated in light of melodic expressions or melodic figures of speech or melodic notes. The highlights were separated in view of the normal successive highlights in every one of the classes and a few highlights which are exceptionally visit in one classification while rare in other class. Trials demonstrated that the k-Means approach is unassumingly effective for isolating out most kinds, while the Markov Chain Model has a tendency to be extremely precise for music classification.

Heung-Nam Kim et. al. [3] have utilized community oriented labeling to channel the client's inclinations for various things. They thought about three unique calculations User CF, Item CF, Tag CF. Client CF calculation deals with likeness between two client's appraisals. Thing CF calculation takes a shot at closeness between two things e.g. In the event that one thing is appraised by one gathering of the clients, at that point what are the odds that other thing will likewise be evaluated by a similar gathering of the clients. Label CF calculation deals with similitude between labels e.g. on the off chance that some client bookmarked some page at that point there are high shots that he will get a kick out of the chance to visit other comparative pages too. The outcomes demonstrated that Tag CF works superior to other two calculations if there should arise an occurrence of cool begin clients for prescribing new things to the objective client. However, it didn't perform well in recognizing the comparable clients as chilly begin client would have labeled less things.

Jin Halee et. al. [4] considered different famous music services,along with the investigation of the way these administrations are utilized and what characteristics are esteemed. Research demonstrated that Earlier clients utilize Yahoo, MTV, and VH1, and now clients utilize youtube for a similar reason. In 2012, Pandora was found as the significant application that serves the music needs. This Literature work was done to discover the quantity of substantial scale client contemplates exist in the MIR area.

Qing LI et. al. [5] depicted a CMRS (community oriented music recommender framework) for online cell phone ringtones. This framework depended on thing based probabilistic model, where things are characterized into gatherings and forecasts are made for clients thinking about the Gaussian dispersion of client appraisals. Utilization of sound highlights was the purpose behind the arrangement of three issues identified with inadequacy in community oriented recommender frameworks that are client inclination, non-affiliation and chilly begin. The main focal point of the creators was on the numerical appraisals, anyway the paired evaluations were generally utilized on the web.

Marius Kmaninskas et. al. [6] portrayed different apparatuses and strategies that can be utilized for tending to the exploration challenges postured by setting mindful music recovery and suggestion. Additionally depicts the subjects that may convey MIR closer to the clients and therefore help creating setting mindful frameworks incorporate intellectual brain science that is the investigations of human impression of music full of feeling figuring especially feeling acknowledgment in music, social processing that is misusing client produced setting for MIR.

Al Mamunur Rashid et. al. [7] displayed a calculation free technique(NUPD) for impacting the rating based proposal framework. NUPD distinguishes the gathering of clients whose evaluations got affected by a limit esteem on the off chance that we keep a specific client out while building the suggestion demonstrate. Be that as it may, this approach is tedious as we need to rehash a similar advance for every client in preparing informational index.

III. TYPES OF RECOMMENDATION SYSTEM

From an algorithmic perspective recommender frameworks fall into three general classes:

1. Information based frameworks.
2. Content separating frameworks.
3. Collective separating frameworks.

Progressed recommender frameworks tend to consolidate shared and content-based sifting, endeavoring to moderate the downsides of either approach and abusing synergetic impacts. These frameworks have been begat "Half and half Systems". Recommender frameworks have (I) foundation information, the data that the framework has before the proposal procedure starts, (ii) input information, the data that client must impart to the framework with a specific end goal to create a suggestion, and (iii) a calculation that joins foundation and info information to land at its recommendations. On this premise, it can be recognize three diverse proposal strategies.

Learning based recommenders (KB)

This kind of proposal endeavors to recommend objects in light of deductions about a client's needs and inclinations. Learning based methodologies know about how a specific thing meets a specific need of the client, and can accordingly reason about the connection between a need and a conceivable suggestion. Learning construct recommenders depend either in light of express space information about

the things or information about the clients, (for example, statistic qualities) to determine applicable suggestions.
Numerous such frameworks depend on physically or consequently produced learning based choice decides that are utilized to prescribe things to clients who fulfill imperatives indicated by the put away standards

Content-based Filtering (CB)

For over three decades, PC researchers have been tending to the issue of data over-burden by outlining programming innovation that naturally perceives and sorts data. Such programming consequently produces depictions of everything' s substance, and afterward thinks about the portrayal of every-thing to a depiction of the client' s data need to decide whether the thing is pertinent to the client's need.

The depictions of the client's advantage needs are either provided by the client, for example, in a question, or gained from watching the substance of things the client devours. These procedures called content-based in light of the fact that the product performs sifting in view of programming investigation of the substance of the things broke down. Content web search tools are a prime case of substance based separating. Numerous content web indexes utilize a procedure called term-recurrence ordering in term recurrence ordering, archives and client data needs are depicted by vectors in a space with one measurement for each word that happens in the database. Every segment of the vector is the recurrence that the particular word happens in the archive or the client question. The archive vectors that are observed to be the nearest to the question vectors (registered utilizing the dab item) are viewed as the destined to be applicable to the client's inquiry. Most data separating and data recovery frameworks today are fabricated utilizing altogether content-based data recovery innovation. Different cases of substance based separating are Boolean pursuit lists, where the question is an arrangement of catchphrases consolidated by Boolean administrators; probabilistic recovery frameworks, where probabilistic thinking is utilized to decide the likelihood that a record meets a client's data require.
Content-based recommender frameworks experience the ill effects of numerous constraints:
• New client issue: for the framework to comprehend and precisely coordinate a client' s inclinations, the client needs to rate an adequate number of items;
• Limited substance examination: because of the restricted highlights that are unequivocally connected with items in the suggestion framework;
• Over-specialization: the framework can't prescribe items that are not quite the same as anything the client has evaluated previously, since the framework can just discover items that score profoundly against a client inclinations

Cooperative sifting frameworks (CF)

Content-based sifting just works when managing areas where include extraction is doable and property data promptly accessible. Community oriented sifting (CF), then again, utilizes content less portrayals and does not confront that same restriction. Communitarian sifting (CF) has been produced to address territories where content-based separating is frail. CF frameworks are not the same as conventional electronic data sifting frameworks in that they don't require modernized comprehension or acknowledgment of substance. In a CF framework, things are sifted in light of client assessments of those things rather than the substance of those things. For the most part, in the cooperative sifting, the substance examination is disregarded and just other client's feelings on the considered substance are viewed as significant. In this way, the shared separating approach is particularly fascinating for content for which content examination is frail or unimaginable. Be that as it may, the execution of the community oriented sifting approach depends on the accessible client inclination information for the considered substance and in this way falls flat when few or no feelings are known. Figure 2 shows the schematic diagram of the synergistic isolating procedure. CF figuring's address the entire $m \times n$ customer thing data as an assessments matrix, A. Each segment aij in An address the tendency score (evaluations) of the ith customer on the jth thing. Each individual examinations is inside a numerical scale and it can be 0 too to an unprecedented case showing that the customer has not yet assessed that thing. The CF Ingredients (input data) are:

1. List of m Users U=[u1, u2,.......um] and a rundown of n Items I=[i1, i2.......].

2. Each client I has a rundown of things I he/she communicated their assessment about (can be an invalid set)

3. Explicit assessment a rating score (numerical scale)

4. Sometime the rating is certainly

5. Active client for whom the CF forecast undertaking is performed

6. A metric for estimating closeness between clients

7. A technique for choosing a subset of neighbors for foreseeing a rating for things not right now evaluated by the dynamic client

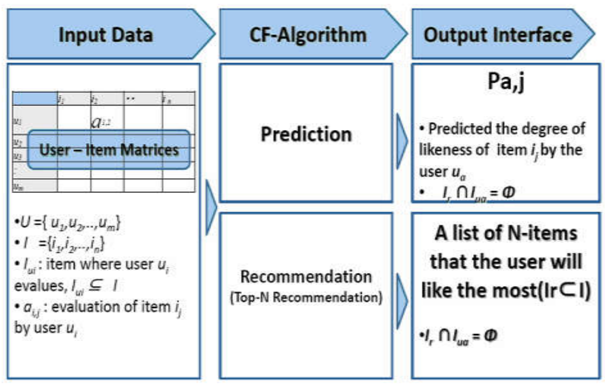

Figure 2: Collaborative Filtering Process[3]

The memory-based methodologies are among the most mainstream expectation procedures in collective sifting. The essential thought is to process the dynamic client's anticipated vote of a thing as a weighted normal of votes by other comparable clients or K closest neighbors (KNN). A few frameworks utilize factual procedures to give individual proposals of reports by finding a gathering of different clients, known as Neighbours that have a background marked by concurring with the objective client. More often than not, neighbourhoods are framed by applying nearness measures, for example, the Pearson relationship between's the suppositions of the clients. These are called closest neighbour strategies.
Synergistic sifting in view of k - closest Neighbour (kNN) approach includes looking at the dynamic record for an objective client with unquestionable records of various customers remembering the ultimate objective to find the best k customers who have practically identical tastes or interests.

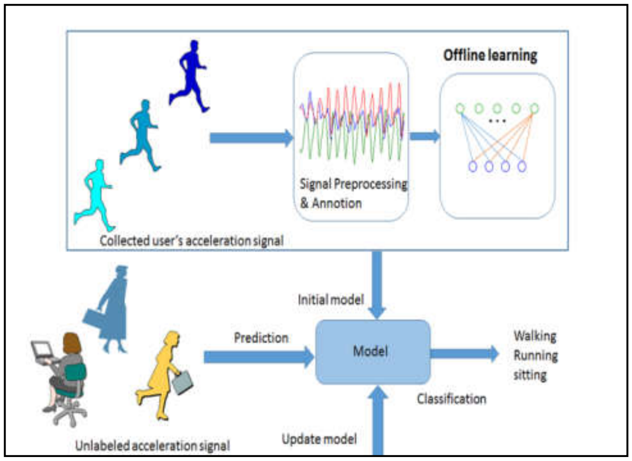

Figure 3: Human Activity Recommendation System[21]

IMPLEMENTATION

Accelerometer and Gyroscope sensors from the UCI Machine Learning Repository are used for collecting the data based on Human Activity Recognition that we use for monitoring people. Various Human Activities which are considered for the experiment are:

Laying
Sitting
Standing
Climbing Stairs
Walking

Indeed, even extraordinary people may complete diverse human exercises, there is as yet a lot of cover which can be utilized to get one of a kind depiction of a movement. For a human it is simple and easy to comprehend human movement in words. The issue is to make an interpretation of it into scientific capacities which we can use to create a signal.so, the arrangement is Feature Extraction. Feature Extraction endeavours to build what an action looks like to the spectator. We have performed include extraction utilizing PCA (Principal Component Analysis) calculation. It is a measurable strategy that uses a symmetrical change to change over an arrangement of perceptions of perhaps connected factors into an arrangement of estimations of directly uncorrelated factors called Principal Components.

PCA Algorithm using Co-Variance Method:

1. Organize the Dataset.
2. Calculate the Empirical Mean.
3. Calculate the deviations from the mean.
4. Find the Co-Variance Matrix.
5. Get the Eigen Vectors and Eigen Values of the co-variance matrix.
6. Rearrange the Eigen vectors and Eigen values.
7. Compute the Cumulative energy content for each Eigen vector.
8. Select the subset of the Eigen vectors as basic vectors.
9. Project the z-scores of the data onto the new basis.

Feature Extraction is done to simplify the Classification Process. After extracting the features, data is trained using different classifiers. The data which is loaded from the individual files is trained so that the parameters belonging to the classifiers which are used is initially fit on a training dataset.

Classification in Supervised Learning is the problem of seeing to which category a sample belongs to by considering training set. In HAR this means we want the activity subject is performing in given time. There are lot of classification methods but there is no classifier that performs best on every given problem.

To check the accuracy of our HAR recommender system we consider two classifiers known as SVM (Support Vector Machine) and K-NN (k-nearest neighbour).

SVM

For a dataset, which consists of features set and labels set, a SVM classifier builds a model to predict classes for new examples. New example or data points are assigned to one of the classes, if it consists of two classes it is known as binary SVM classifier.

k-NN

The calculation utilizes the neighbor focuses data for the forecast of Target class.

K-closest neighbor (Knn) calculation pseudocode:

Let (Xi, Ci) where I = 1, 2… … ., n be information focuses. Xi indicates highlight esteems and Ci means marks for Xi for every I.

Expecting there are c no of classes:

Ci has a place with {1,2,3… … ..,c} for all estimations of I.

Give x a chance to be a point for which mark isn't known, and we might want to discover the name class utilizing k-closest neighbor calculations.

Knn Algorithm Pseudocode:

1. Calculate "d(x, xi)" I =1, 2, … .., n; where d means the Euclidean separation between the focuses.

2. Arrange the figured n Euclidean separations in non-diminishing request.

3. Let k be a +ve whole number, take the primary k separations from this arranged rundown.

4. Find those k-directs comparing toward these k-separations.

5. Let ki indicates the quantity of focuses having a place with the ith class among k focuses i.e. k ≥ 0

6. If ki >kj ∀ I ≠ j at that point place x in class I.

The purpose of the k Nearest Neighbor algorithm is to predict the classification of a new sample by using a database in which the data points are separated into several separate classes.

For predicting class Pearson Correlation Coefficient is used for the estimation of Relevance R of feature Fi

$$R(F_i;C) = \frac{cov(F_i, C)}{\sqrt{var(F_i) \cdot var(C)}}$$

where cov stands for the covariance and var the variance [21]

We next conducted experiments to evaluate the accuracy of our system using popular methods SVM and K-NN.

EXPERIMENTAL RESULTS:

In our Experiments, we select all examples to frame the preparation set, and the rest of the subjects are utilized as test sets. Trials are performed in MATLAB 2017 as a result of the immense libraries accessible there, capacities like Fast Fourier Transform, Principal Component Analysis and Covariance did not need to be composed another, which spares the time as the measure of code that must be composed is diminished. Likewise, MATLAB depends on its rich tool compartments, and its capacity to control vectors and networks. Comparison of two popular classifiers is done to evaluate the accuracy of the system.

Parameters which are considered for the accuracy evaluation are ROC (Receiver Operating Characteristic) and Training Response. A receiver operating characteristic (ROC) chart is a procedure for envisioning, arranging and choosing classifiers in view of their execution.

In our classification problem the yields of one layer are utilized as the contribution to the following layer and this procedure is reasonable for highlight choice.

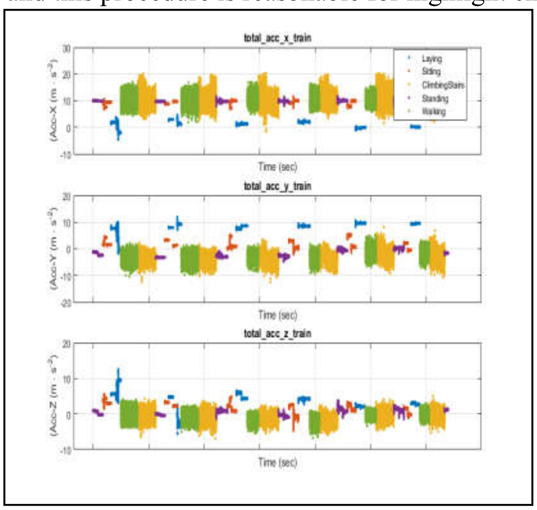

Fig 4: Human Activity Acceleration Data Results

After selecting the data from workspace, it uses cross validation method for validating the data. In this, it selects a number of folds for partition of data. Each fold is held out in turn for testing. After this, it trains a model for each fold using each data outside its fold. It can vary the cross validation folds.

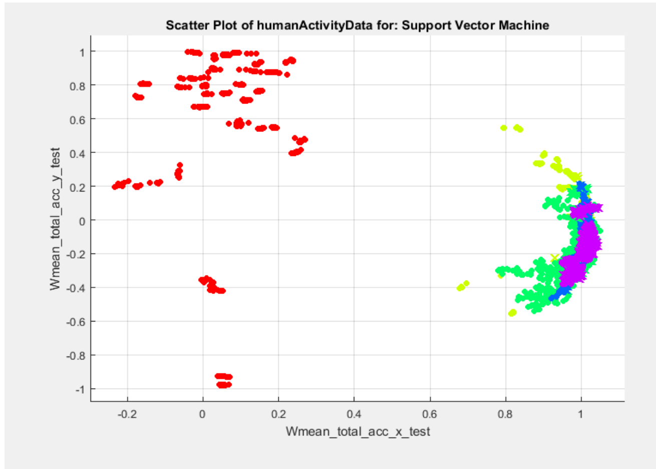

Figure 5: True Class Representation After Classification

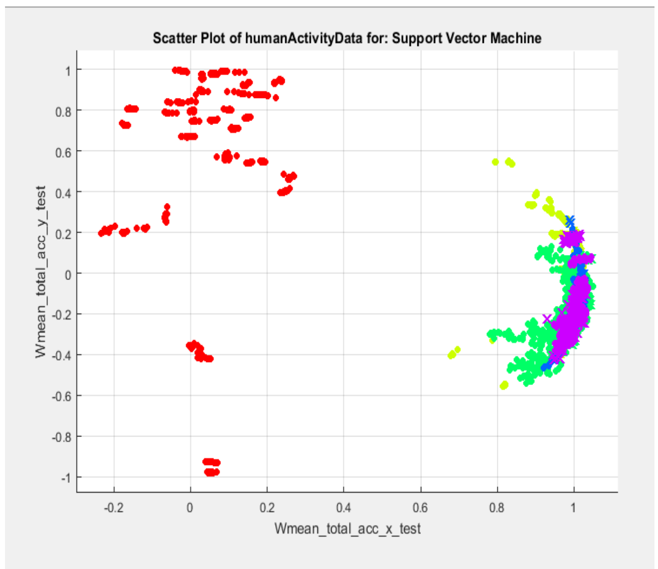

Figure6: Predicted Class Representation After Classification

To begin with we set up the applet as appeared in Figure 5 &6, with purposes of two unique hues in inverse corners of the screen, and a few purposes of arbitrary for each shading. Presently, these focuses will make up our preparation set, since we know their attributes (their x and y facilitates) and their characterization (their colour).

Activity	Area Under Curve (Linear SVM)	KNN (1 Neighbour)	KNN (2 N)	KNN (3N)
Laying	1	1	1	1
Sitting	0.9925	0.9749	0.9872	0.9906
Climbing Stairs	0.9731	0.9522	0.9734	0.9794
Standing	0.9930	0.9770	0.9880	0.9911
Walking	0.9643	0.9504	0.9664	0.9684

Table 1: ROC for Different Activities using Classifiers

Method	Training (%)
Complex Tree	92.6
SVM	88.7
KNN (1-N)	95.1
KNN (2-N)	94.5
KNN (3-N)	94.2

Table 2: Training Response using Different Classifiers

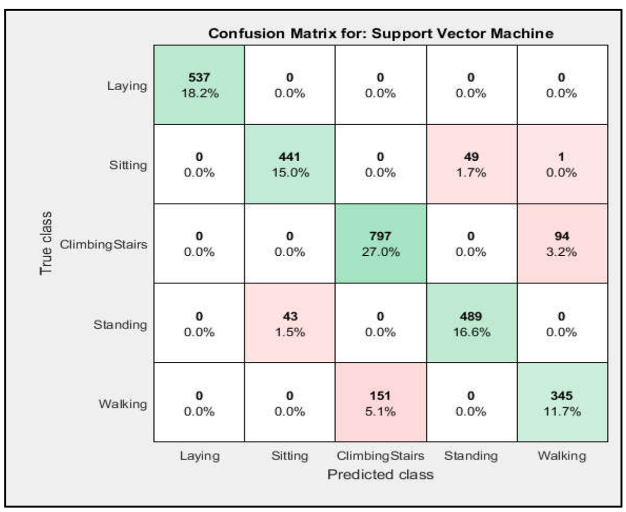

Figure 7: Confusion Matrix for Human Activities using SVM

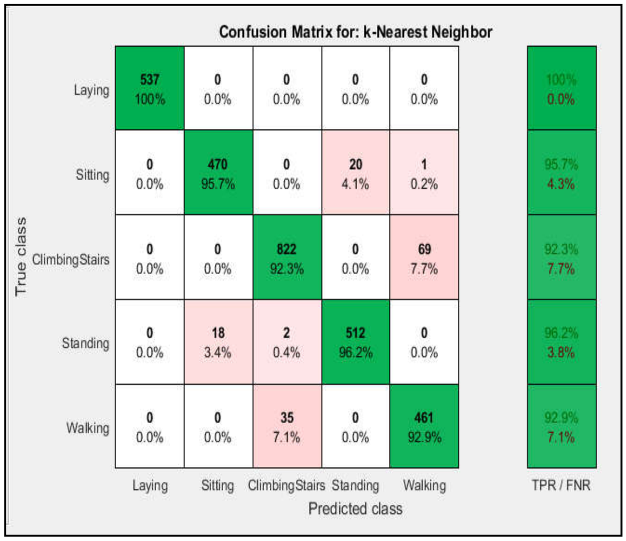

Figure 8: Confusion Matrix for Human Activities using KNN

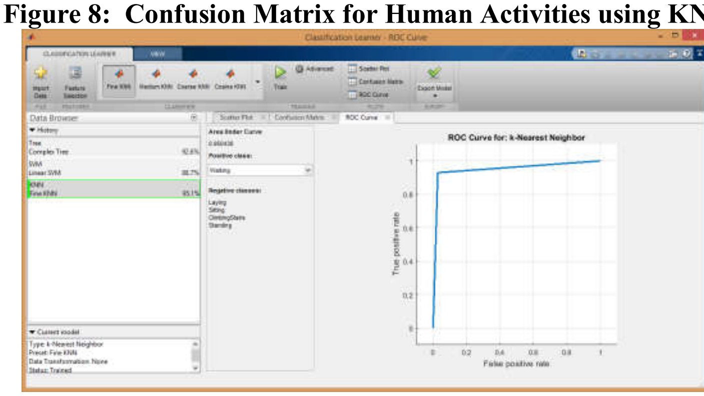

Fig 9: ROC curve for k-NN classifier

So, Experiment result shows that K-NN classifier gives the best accuracy both in terms of ROC curve and Training Response. Also on increasing the number of nearest neighbours k-NN classifier performs best recommendation with the highest accuracy.

V. CONCLUSIONS

The design of recommender system approach depends on the information sources and interest on various objects used by the system. Some of the sources are easily available and some of them are not easily available. An approach to user model for recommenders system based on the human activity factors is developed. Different attributes that capture activity features of a user is analysed and discussed. Advanced mobile phones furnished with three-pivot accelerometers are ending up progressively famous. Hence, cell phones based human movement acknowledgment is turning into a vital application, and has just added to improvements in web based life. Considering the multifaceted nature of the three-hub increasing speed flag, distinctive classifiers are utilized to exhibit human movement..In future scope, Introduce a machine learning approach to make the system more intelligent in re-computing the Impact Ratios and eliminating features or asking for other features, to make the system more stable.

REFERENCES

1. Sánchez-Moreno, Diego, Ana B. Gil González, M. Dolores Muñoz Vicente, Vivian F. López Batista, and María N. Moreno García. "A collaborative filtering method for music recommendation using playing coefficients for artists and users." *Expert Systems with Applications* 66 (2016): 234-244.
2. Caswell, Isaac, and Erika Ji. "Analysis and Clustering of Musical Compositions using Melody-based Features." (2013).
3. Kim, Heung-Nam, Ae-Ttie Ji, and Geun-Sik Jo. "Enhanced prediction algorithm for item-based collaborative filtering recommendation." In *International Conference on Electronic Commerce and Web Technologies*, pp. 41-50. Springer, Berlin, Heidelberg, 2006.
4. Typke, Rainer, Frans Wiering, and Remco C. Veltkamp. "A survey of music information retrieval systems." In *Proc. 6th International Conference on Music Information Retrieval*, pp. 153-160. Queen Mary, University of London, 2005.
5. Lee, Jin Ha, and Nichole Maiman Waterman. "Understanding User Requirements for Music Information Services." In *ISMIR*, pp. 253-258. 2012.
6. Isinkaye, F. O., Y. O. Folajimi, and B. A. Ojokoh. "Recommendation systems: Principles, methods and evaluation." *Egyptian Informatics Journal* 16, no. 3 (2015): 261-273.
7. Min, Sung-Hwan, and Ingoo Han. "Detection of the customer time-variant pattern for improving recommender systems." *Expert Systems with Applications* 28, no. 2 (2005): 189-199.
8. Celma, Òscar, and Xavier Serra. "FOAFing the music: Bridging the semantic gap in music recommendation." Web Semantics: Science, Services and Agents on the World Wide Web 6, no. 4 (2008): 250-256.
9. Sarwar, Badrul, George Karypis, Joseph Konstan, and John Riedl. "Item-based collaborative filtering recommendation algorithms." In Proceedings of the 10th international conference on World Wide Web, pp. 285-295. ACM, 2001. [10] B. Logan, "Music recommendation from song sets," ISMIR, 2004, pp. 425–428.
10. Uitdenbogerd, Alexandra, and R. Schyndel. "A review of factors affecting music recommender success." In ISMIR 2002, 3rd International Conference on Music Information Retrieval, pp. 204-208. IRCAM-Centre Pompidou, 2002.
11. Yoshii, Kazuyoshi, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno. "Hybrid Collaborative and Content-based Music Recommendation Using Probabilistic Model with Latent User Preferences." In ISMIR, vol. 6, p. 7th. 2006.
12. Jain, Achin, Vanita Jain, and Nidhi Kapoor. "A literature survey on recommendation system based on sentimental analysis." Advanced Computational Intelligence 3, no. 1 (2016): 25-36.
13. Song, Yading, Simon Dixon, and Marcus Pearce. "A survey of music recommendation systems and future perspectives." In 9th International Symposium on Computer Music Modeling and Retrieval, vol. 4. 2012.
14. Haruna, Khalid, Maizatul Akmar Ismail, Damiasih Damiasih, Joko Sutopo, and Tutut Herawan. "A collaborative approach for research paper recommender system." PloS one 12, no. 10 (2017): e0184516.
15. MarkusSchedl.et.al,"current challenges and vision in music recommender system research",2017.
16. Ding, Yiwen, and Chang Liu. "Exploring drawbacks in music recommender systems: the Spotify Case."(2015).
17. Rashid, Al Mamunur, George Karypis, and John Riedl. "Influence in ratings-based recommender systems: An algorithm-independent approach." In Proceedings of the 2005 SIAM International Conference on Data Mining, pp. 556-560. Society for Industrial and Applied Mathematics, 2005.
18. Mortensen, Magnus. "Design and evaluation of a recommender system." Master's thesis, Universitetet i Tromsø, 2007.

19. Laplante, Audrey. "Improving music recommender systems: What can we learn from research on music tastes?." In ISMIR, pp. 451-456. 2014.
20. Kaminskas, Marius, and Francesco Ricci. "Contextual music information retrieval and recommendation: State of the art and challenges." Computer Science Review 6, no. 2-3 (2012): 89-119.
21. Multi-column Bi-directional Long Short-Term Memory for Mobile Devices-based Human Activity Recognition Dapeng Tao, Yonggang Wen, Senior Member, IEEE, and Richang Hong,2016.