Performance Analysis on Human Activity Detection using KNN and Random Forest

Sai Narendra L, Samuel Kiran, Naga Brahmani K. Vamsidhar E

Abstract: : Human Activity Recognition is a promising area being able to profit the human culture by making assistive types of progress so as to help old, unendingly incapacitated and besides for individuals with phenomenal prerequisites. Exact improvement insistence is attempting since human movement is amazing and exceptionally different. Making study performed around there has uncovered information tunneling algorithm are utilized for solicitation of exercises. Hybrid mining frameworks, Naive Bayes with SVM and C4.5 with Neural Network are wound up being productive in portraying the accelerometers looking at information. These datasets are having wide arrangement of occasion with many proceeds with qualities. Working up a classifier that get-together such information is as of not long ago a troublesome errand. Sporadic woods is known for accomplishing high precision all together. It's quality in social occasion broad datasets is promising. This paper proposes a sporadic timberland based course of action display for social classifying/predicting the strategy for activities. Preparing information is pre-managed to accomplish consistency. Points of reference from preparing dataset are pulled in sporadic for n tests, and n choice tree are made. Thus, an emotional choice backwoods is worked for depicting begins based accelerometers information respects. To predict unlabeled exercise information, total of n trees is performed. Primer takes a gander at are composed to inspect the movement confirmation limit of the model the outcomes are separated and transcendent regulated solicitation structures. It is seen that the proposed model beat the other depiction methodologies in relative examination. The sorted out social event show is restricted to perform action confirmation regarding weight lifting works out. Human Activity insistence is can be related with some reality, human-driven issues.

Index Terms: KNN, Random Forests, Machine Learning.

I. INTRODUCTION

Mobile devices and specifically smartphones have beginning late utilized dumbfounding and assembled sensors. These sensors unite Accelerometer, Gyroscope, Light, Battery, Temperature, Magnetometer, and some sensibly, each with their particular utmost. In light of advancement in progressions, these PDAs are bending up consistently restricted, clear, and explicitly, on a very basic level smart PCs. The capacity to send, get information have made them all around open in our general populace. These contraptions, by ethics of normal enrolling point of confinement and marvelous sensors, have shown another examination space – Wireless Sensor Data mining. Sensors can be utilized to discover information relating to battery of gadget, to powers being related on the gadget.

One such sensor limit has been mauled in this paper, is, the accelerometer sensor. Accelerometer sensor assesses the growing rate powers being related on telephone, concerning three dimensional X, Y and Z form structure. The sensor surveys the powers in m/s2. This paper utilizes this making sense of how to depict the human action under 4 classes, that is, Walking, Running, Sitting and Standing. We will control speaking attempt the way that, while the client is playing out any action, the telephone kept in the pants pocket would encounter a power related (on Y turn) that tends to advancement structure. This model can be seen for various exercises. The Authors have picked Android-based PDAs as the stage for our WISDM experience in light of the manner in which that the Android working structure is free, open-source, simple to program, and expected to change into a stunning portion in the phone business center (this is without a doubt occurring) [5]. This errand can in addition be scaled to other versatile stages, for example, IOS and others in market. Accelerometers were at first combined into these gadgets to help induced redirection play and to connect with altered screen turn in any case they obviously have different various applications. [5] to be honest, there are different valuable applications that can be made whether accelerometers can be utilized to see a client's improvement. For instance, we can typically screen a client's improvement level and make every day, well ordered, and month to month action reports, which could be subsequently educated to the client [5]. These reports would demonstrate a general advancement level, which could be utilized to evaluate if the client is getting a satisfactory extent of development and check the measure of reliably calories used. These reports could be utilized to engage solid practices and may caution two or three clients to how lethargic they or their young people really are. To see the advancement through remote utilizing supervised learning, the producers initially collect the information while client plays out the action with an android based application assembling the required information. Once the dataset is made, we utilize different AI estimations to assemble the action under the as of late referenced groupings.

II. RELATED WORK

The area of improvement certification isn't new. Amin Rasekh et al. sorted out [1] an improvement confirmation framework subject to a remote. The structure utilizes a3-dimentional remote accelerometer as the essential sensor to gather time strategy signals, from which 31 highlights are made in both time and rehash locale. Exercises are composed utilizing 4 distinctive kept learning systems, i.e., quadratic classifier, k-closest neighbor estimation, strengthen vector machine, and phony neural

structures. Results demonstrate that the social occasion rate of inactive learning achieves 84.4%.Akram Bayat et al. [2] propose an insistence framework in which another robotized low-pass channel is sorted out to pull back the bit of gravity quickening from that of body growing pace in the unpleasant information. The structure was organized and endeavored in a research assorted streets regarding different human subjects in real conditions. It was discovered that utilizing the common of probabilities as the mix method achieved a general precision rate of 91.15%. Sang and Vu Ngoc Thanh [3] tended to well ordered action watching utilizing accelerometer and whirligig sensors presented in a remote. Signs were recorded from accelerometer and spinner sensors while a client was wearing the PDA and performing unquestionable exercises. For isolated examination, the get-together checks with k-closest neighbor (kNN), fake neural structure (ANN) and fortify vector machine (SVM) were related with see client's exercises. The general exactness of seeing five exercises was 74% for kNN, 75.3% for ANN and 94.5% for SVM autonomously. Rao Fu et al. [4] focus around improving solicitation exactness and lessening computational multifaceted nature for human action attestation issue on open datasets. Straightforwardly off the bat they got grungy information from sensors. On account of PDA, the information was gotten from the accelerometer. This foul information preprocessed for solicitation preparing figuring. For telephone accelerometer, time course of action information were divided into parts. Highlights are then made and picked subject to grungy information. In the wake of giving enough information tests, show was mapped utilizing reasonable learning estimations. The general exactness accomplishes over 80%.

III. IMPLEMENTATION

A. Data Collection: For our examination work out, we have utilized the dataset accessible wholeheartedly on [1]. The dataset is a social affair of accelerometer readings from 4 sensors (on belt, on left thigh, on right lower leg, on right arm) worn by all of 4 sound subjects while playing out unequivocal exercises in 5 contrasting ways (taking a seat, standing up, strolling, standing, sitting) for totally of 8 hours. Every improvement was performed self-rulingly by the subjects. The result class to foresee was the path by which action was performed (sitting, standing, standing up, taking a seat, and walking).

B. Feature Selection: Following highlights were picked for model structure. The outline of best highlights to be utilized for model was gotten from . The last highlights utilized for model are as indicated by the going with: (1) Sensor on the Belt: discretization of the module of enlivening vector, change of pitch, and differentiation of move; (2) Sensor on the left thigh: module of quickening vector, discretization, and instability of pitch; (3) Sensor on the correct lower leg: difference of pitch, and refinement of move; (4) Sensor on the correct arm: discretization of the module of extending speed vector; From all sensors: run of the mill resuscitating and standard deviation of growing rate. Instability of various pitch and move centers and customary extending speed and standard deviation were settled utilizing moving run of the mill with window length of 9. Accelerometer examining can be considered as time course of action information. The length 9 on moving window was associated at watching the spikes in Auto-Correlation and Partial Autocorrelation of the information.

c. Proposed Work: There has been surprising change in the manner, information is verified, saw and dealt with. Enormous extent of information is made each second and if this information is utilized and examined proficiently, it can uncover essential bits of learning. Bundle of information mining techniques has advanced in isolating the gigantic extent of information. One fundamental piece of the figure is the choice of reasonable models. In our development, we have made models utilizing two or three AI methods and looked exactness of changed algorithms.

Random Forest Random forests is in like way the estimation which will when all is said in done union powerless understudies to improve exactness. It bootstraps specific markers and makes diverse weak trees from bootstrapped pointers. Bootstrapping of markers guarantees less related trees. In end it joins frail choice trees to predict the result. This calculation comparably yields much better game-plan precision over decision trees.

Training and tuning HAR information utilizing unusual timberland philosophy and association with two other social occasion methodologies. Abstract Forests (RF) includes a mix of choice trees. It improves the social affair execution of a solitary tree classifier by joining the bootstrap gathering (bagging) system and randomization in the choice of isolating server farm focuses in the progression of choice tree. The errand of another acknowledgment vector to a class depends upon a greater part vote of the unquestionable choices given by each tree developing the woodland. Not with standing, RF needs gigantic extent of checked information to accomplish unimaginable introductions.

k-Nearest Neighbors k-Nearest Neighbors (k-NN) is a directed solicitation approach that can be viewed as a brief game-plan framework since it doesn't require a learning technique. It just requires the breaking point of the entire information. To organize another acknowledgment, the K-NN algorithm utilizes the standard of likeness (remove) between the status set and new perception to plan. The new acknowledgment is doled out to the most for the most part seen class through a lion's offer vote of its k closest neighbors. The unit of the neighbors of a perception is settled utilizing a division estimation called likeness limit, for example, Euclidean parcel. In addition, one ought to see that while utilizing the K-NN approach and another point of reference is doled out to a class, the check of parcels (i.e., the calculation time) increases as a section of the present models in the dataset.

Foerster et al. were the first to apply the k-NN strategy to confine between nine human exercises utilizing time-space highlights picked up from three uniaxial accelerometers. In Foerster and Fahrenberg joined k-NN with an alternate leveled choice way to deal with oversee see nine exercises utilizing rehash locale highlights. This method has had every one of the reserves of being ceaselessly proficient, to the degree demand exactness, showed up distinctively in

connection to the k-NN. Assorted examinations subject to k-NN for human action confirmation have additionally appeared to be bizarre state of exactness and satisfactory segmentation results.

Random Forests Random Forests (RF) incorporates a combination of decision trees. It improves the social event execution of a solitary tree classifier by combining the bootstrap totaling (bagging) technique and randomization in the choice of dividing server farm nodes in the headway of decision tree. The task of another observation vector to a class depends upon a lion's offer vote of the specific choices given by each tree including the forested territories. In any case, RF needs tremendous extent of named information to achieve good performances.

IV. PROPOSED SYSTEM

Evaluation : The accuracy measure is utilized to overview the classifiers appears. Truly, this metric measures the level of decisively composed models. As a result of twofold strategy, the accuracy can be granted as looks for after:

Classifier

A program or a activity which maps from unlabeled occasions to classes is known as a classifier.

Confusion Matrix

A confusion matrix, in like way called a contingeny table or blunder cross area, is utilized to envision the execution of a classifier.

The pieces of the matrix address the occasions of the anticipated classes and the fragments address the occasions of the veritable class. (Note: It can be the other path around additionally.)

By uprightness of binary classification the table has 2 rows and 2 columns.

Example:

Confusion Matrix		Predicted classes	
		male	Female
Actual classes	male	42	8
	Female	18	32

This proposes the classifier exactly anticipated a male individual in 42 cases and it wrongly predicted 8 male models as female. It effectively predicted 32 occasions as female. 18 cases had been wrongly predicted as male instead of female.

Accuracy (error rate)

Accuracy is a valid measure which is depicted as the remainder of right wants made by a classifier isolated by the total of predictions made by the classifier.

The classifier in our past perspective predicted effectively anticipated 42 male models and 32 female case

Therefore, the accuracy can be calculated by: accuracy = (42+32)/(42+8+18+32)   which is 0.74

Let's assume we have a classifier, which always predicts "female". We have an accuracy of 50 % in this case.

Confusion Matrix		Predicted classes	
		male	female
Actual classes	Male	0	50
	Female	0	50

We will demonstrate the so-called accuracy paradox.
A spam recognition classifier is described by the following confusion matrix:

Confusion Matrix		Predicted classes	
		spam	Ham
Actual classes	Spam	4	1
	Ham	4	91

The accuracy of this classifier is (4 + 91) / 100, i.e. 95 %.
The following classifier predicts solely "ham" and has the same accuracy

Confusion Matrix		Predicted classes	
		spam	ham
Actual classes	spam	0	5
	ham	0	95

The accuracy of this classifier is 95%, even though it is not capable of recognizing any spam at all.

Precision and Recall

Confusion Matrix		Predicted classes	
		negative	positive
Actual classes	Negative	TN	FP
	Positive	FN	TP

Accuracy: (TN+TP)/(TN+TP+FN+FP)(TN+TP)/(TN+TP+FN+FP)

Precision: TP/(TP+FP)TP/(TP+FP)   Recall: TP/(TP+FN)TP/(TP+FN)

V. EXPERIMENTTAL RESULTS

A. KNN Classification:

```
Python 3.7.0 (v3.7.0:1bf9cc509, jun 27 2018, 04:06:47) [msc v.1914 32 bit (intel)] on win32
Type "copyright", "credits" or "license()" for m
ore information.
>>>
RESTART: E:\E-drive\java projects\PYTHON PROJECTS\Assigment_Australia\knn_final.py
```

S.NO	precision	Recall	F1-score	support
1	0.85	0.97	0.91	496
2	0.90	0.90	0.90	471
3	0.93	0.78	0.85	420
4	0.87	0.78	0.82	491
5	0.82	0.89	0.85	532
6	1.00	1.00	1.00	537
Avg total	0.89	0.89	0.89	2947

```
TP – true negative 481
FP – false positive 4
FN – false negative 36
TN– true positive 422
Accuracy rate : 0.30641330166270786
Misclassification rate: 0.013573125212080081
```

B. Random Forest:

```
Python 3.7.0 (v3.7.0:1bf9cc509, jun 27 2018, 04:06:47) [msc v.1914 32 bit (intel)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
RESTART: E:\E-drive\java projects\PYTHON PRO
JECTS\Assigment_Australia\knn_final.py
```

S.NO	precision	Recall	F1-score	Support
1	0.85	0.97	0.91	496
2	0.90	0.90	0.90	471
3	0.93	0.78	0.85	420
4	0.87	0.78	0.82	491
5	0.82	0.89	0.85	532
6	1.00	1.00	1.00	537
Avg /total	0.89	0.89	0.89	2947

```
TP – true negative 481
FP - false positive 4
FN – false negative 36
TN- true positive 422
Accuracy rate : 0.30641330166270786
Misclassification rate: 0.013573125212080081
```

C. Linear regression:

```
RESTART: E:\E-drive\java projects\PYTHON PROJECTS\Assigment_Australia\fina_rfc.py
Training Accuracy of decision tree : 0.9351196953210011
Testing Accuracy of Decision tree: 0.828978622327791
[[537    0    0    0    0    0]
 [  0  438  233    0    0    0]
 [  0   53  299    0    0    0]
 [  0    0    0  435    6   23]
 [  0    0    0   46  357   71]
 [  0    0    0   15   57  377]]
```

	0	1	2	3	4	5
0	537	0	0	0	0	0
1	0	438	233	0	0	0
2	0	53	299	0	0	0
3	0	0	0	435	6	23
4	0	0	0	46	357	71
5	0	0	0	15	57	377

```
TP – true negative 537
FP - false positive 0
FN – false negative 0
TN- true positive 438
Accuracy rate : 0.330844927044452
Misclassification rate: 0.0
```

Prediction over different activities by humans using linear regression is as follows:

	Walking	walking upstairs	....	Standing	laying
Walking	537	0	.....	0	0
Walking upstairs	0	424	......	0	0
Walking downstairs	0	64	....	0	0
Sitting	0	0	....	1	6
Standing	0	0	......	413	1
Laying	0	3	....	6	464

Linear regression representation graph :
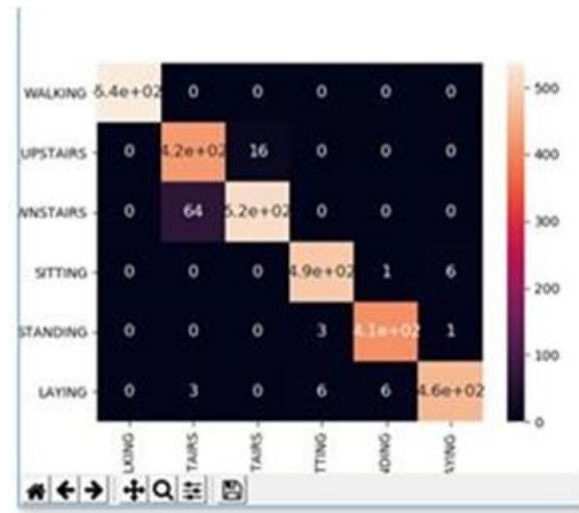
The graphical representation of the test accuracy and train accuracy is as follows:
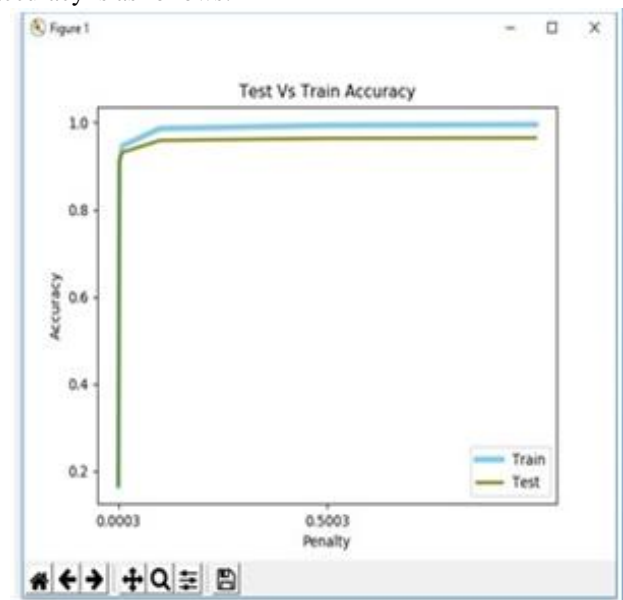
VI. CONCLUSION

The Random Tree show predicted over the test edifying amassing with a 99.97% accuracy. KNN show anticipated over the test educational collection with a 99.59% accuracy. This outcomes exhibits that Random Forest model was the right decision to isolate the information. The weight lifting preparing enlightening social event was utilized to make a model that predicted the manner by which a subset played out the weight lifting exercise. The sorted out depiction show is constrained to perform improvement certification regarding weight lifting works out.

FUTURE SCOPE: Not only the algorithms that used in this paper , there are good number of algorithms. so, this paper can further extended by using many other algorithms for suggesting algorithm which is more accurate than the algorithm proposed in this paper, this can also be extended for suggesting which algorithm is more effective in terms of space and time complexity.

REFERENCES

1. J. Perez, M. A. Labrador, and S. J. Barbeau, "G-sense: A scalable architecture for global sensing and monitoring," IEEE Network, vol. 24, pp. 57–64, 2010.
2. Ulf Blanke and BerntSchiele, "Remember and transfer what you have learned recognizing composite activities based on activity spotting ", Proceedings of ISWC, pp. 1– 8, 2010.
3. K. Partridge and P. Golle, " On using existing time-use study data for ubiquitous computing applications", Proceedings of UbiComp, ACM, pp. 144–153, 2008.
4. Bernd Tessendorf, Andreas Bulling, Daniel Roggen, Thomas Stiefmeier, Manuela Feilner, Peter Derleth, and Gerhard Tro ster. 2011a. Recognition of hearing needs from body and eye movements to improve hearing instruments. In Proceedings of the 9th International Conference on Pervasive Computing. Springer, 314–331.
5. DOI:http://dx.doi.org/10.1007/978-3-642-21726-5_20   Bernd Tessendorf, Franz Gravenhorst, Bert Arnrich, and Gerhard Tro ster. 2011b. An IMU-based sensor network to continuously monitor rowing technique on the water. In Proceedings of ISSNIP.
6. S. Katz, T. D. Downs, H. R. Cash, and R. C. Grotz. 1970. Progress in development of the index of ADL. The Gerontologist 10, 1 Part 1 (1970), 20.
7. K. Partridge and P. Golle. 2008. On using existing timeuse study data for ubiquitous computing applications. In Proceedings of UbiComp. ACM, 144–153.
8. Matthias Kranz, Andreas Mo ller, Nils Hammerla, Stefan Diewald, Thomas Plo tz, Patrick Olivier, and Luis Roalter. 2013. The mobile fitness coach: Towards individualized skill assessment using personalized mobile devices. Pervasive and Mobile Computing 9, 2 (2013), 203–215.
9. Eduardo Velloso, Andreas Bulling, Hans Gellersen, Wallace Ugulino, and Hugo Fuks. 2013b. Qualitative activity recognition of weight lifting exercises.In Proceedings of the 4th Augmented Human International Conference (AugmentedHuman   2013). 116–123.
10. DOI:http://dx.doi.org/10.1145/2459236.2459256   . [10] Taˆm Huynh, Mario Fritz, and BerntSchiele. 2008. Discovery of activity patterns using topic models. In Proceedings of UbiComp. 10–19.
11. F. Michahelles and B. Schiele, "Sensing and monitoring professional skiers ". IEEE Pervasive Computing, 4(3), pp. 40–46, 2005. a. M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. Information Technology in Biomedicine, IEEE Transactions on, 12(1):20 –26, jan. 2008.
12. K.-h. Chang, M. Y. Chen, and J. Canny. Tracking freeweight exercises. In Proceedings of the 9th ACM International Conference on Ubiquitous computing, volume 4717, pages 19–37. Springer, 2007.
13. T. Dutta. Evaluation of the kinect sensor for 3-d kinematic measurement in the workplace.Applied Ergonomics, 2011.
14. T. C. Havens, G. L. Alexander, C. Abbott, J. M. Keller, M. Skubic, and M. Rantz. Contour tracking of human exercises. In Proceedings of Computational Intelligence for Visual Intelligence, pages 22–28, 2009.
15. Moeller, L. Roalter, S. Diewald, M. Kranz, N. Hammerla, P. Olivier, and T. Ploetz. Gymskill: A personal trainer for physical exercises. In Proceedings of the International Conference on Pervasive Computing and Communications, 2012.
16. N. Y. Hammerla, T. Ploetz, P. Andras, and P. Olivier. Assessing motor performance with pca. In Proceedings of the International Workshop on Frontiers in Activity Recognition using Pervasive Sensing, 2011.
17. Strohrmann, H. Harms, G. Tro ster, S. Hensler, and R. Mu ller. Out of the lab and into the woods: Kinematic analysis in running using wearable sensors. In Proceedings of the 13th ACM International Conference on Ubiquitous Computing (UbiComp 2011). ACM,Sept. 2011.
18. Zinnen, U. Blanke, and B. Schiele. An analysis of sensor-oriented vs. model-based activity recognition. In Wearable Computers, 2009.ISWC '09. International Symposium on, pages 93 –100, sept. 2009.
19. Zinnen, C. Wojek, and B. Schiele. Multi activity recognition based on bodymodel-derived primitives. In Proceedings of the 4th International Symposium on Location and Context Awareness, LoCA '09, pages 1– 18, Berlin, Heidelberg, 2009. Springer-Verlag.
20. Reiss, G. Hendeby, G. Bleser, and D. Stricker. Activity recognition using biomechanical model based pose estimation. In Smart Sensing and Context, volume 6446 of Lecture Notes in Computer Science, pages 42–55. Springer, 2010.
21. M. Beetz, B. Kirchlechner, and M. Lames. Computerized real-time analysis of football games. IEEE Pervasive Computing, 4(3):33–39, 2005.
22. O. F. Standardisation. Iso 9000:2005 quality management systems – fundamentals and vocabulary. Management, 3, 2005.
23. P. B. Crosby. Quality is Free. McGraw-Hill, 1979. M. A. Hall. Correlation-based Feature Subset Selection for Machine Learning. PhD thesis, Department of Computer Science, University of Waikato, Hamilton, New Zealand, Apr. 1999.
24. L. B. Statistics and L. Breiman. Random forests. In Machine Learning, pages 5–32, 2001. L. Breiman. Bagging predictors. Machine Learning, 24:123–140, 1996. 10.10

AUTHORS PROFILE

SAI NARENDRA REDDY L is a student at the department of Computer Science and Engineering at K L Educational foundation, Deemed to be University, Vaddeswaram, Andhra Pradesh. He is doing his research work in knowledge engineering.

SAMUEL KIRAN K is a student at the department of Computer Science and Engineering at K L Educational foundation, Deemed to be University, Vaddeswaram, Andhra Pradesh. He is doing his research work in knowledge engineering.

NAGA BRAHMANI K   is a student at the department of Computer Science and Engineering at K L Educational foundation, Deemed to be University, Vaddeswaram, Andhra Pradesh. He is doing his research work in knowledge engineering.

DR,VAMSIDHAR E completed M.Tech from Andhra University in 2008 and Ph.D from JNTU Kakinada in 2017.The research Interests Includes Image preprocess- Ing,,Machine Learning and Bio-Informatics.