International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org, editorijettcs@gmail.com
Volume 3, Issue 4, July-August 2014 **ISSN 2278-6856**

Artificial Neural Network: Approach for Classifying Trajectories on Road Network

Deepak S. Gaikwad[1] and Usha A. Jogalekar[2]

[1]PG Fellow, SKNCOE,Pune

[2]Assistant Professor, SKNCOE,Pune

Abstract

Classification is the classical problem in the process of machine learning and data mining. Classification is widely used for modeling various types of data sets, namely sets of items, text document, networks, signals and graphs. But, there is lack of study on Trajectory Classification. Trajectory classification is defined as the process of predicting vehicles class label based on its trajectory. Trajectory Data is the data which is collected from moving objects (vehicles). Large amount of trajectory (i.e. vehicles on road network) data is collected from technologies such as GPS, Sensors, Cameras, RFID etc., this data is useful for better Traffic prediction, activity recognition as well as transportation planning of the city. In this paper, Multi-Layer framework of Artificial Neural Network (ANN) is applying for classifying the trajectories on road network. ANN's are generally presented as systems of interconnected "neurons" which compute values from inputs. ANN is used for classification of items, signals etc. ANN is also employed in analysis of investment, stock exchange as well as verification of signature etc., but less employed on Trajectory data. In this paper, analyzing behavior of trajectories on road network, we have employed ANN on Trajectory data and projecting a model using sequential pattern. Sequential patterns are good feature candidate as it preserves order of visiting sequence of trajectories on road network.

Keywords: Trajectory Classification, Feature Extraction, Data Mining, Sequential Pattern, Artificial Neural Network

1. INTRODUCTION

In recent years, use of modern technologies such as Global Positioning System (GPS), Sensors, Radio Frequency Identification (RFID), a huge amount of trajectory data is gathered. Trajectory data is nothing but the path followed by an object to travel from one location to other. Analysis of Trajectory data is being widely acknowledged [1] [2] [3] [4] [5] [6] [7] [8]. Trajectory classification can be used for prediction of traffic, activity monitoring of vehicles and traffic planning of the city. This paper addresses the problem of classification of trajectories on road network using Artificial Neural Network. Trajectory classification means that the classification of moving object on the basis of their class label contains location where moving object (vehicle) visited as well as order of visited location. For Trajectory classification, first study the behavior of trajectories on the road network. Importantly, locations where vehicles are visited are very crucial. So, we use Sequential Patterns [9] [10] are better feature candidate because it preserves the order of information of trajectories [2].

In this paper, we propose to use sequential pattern as a feature candidate which is extracted from synthetic data set. There are six types of features namely Highway, Lanes, Maximum speed, One-way, Surface, Support etc. are extracted on the basis of sequential pattern, these Extracted Feature dataset can be fed to the classifier i.e. Multi-Layer Artificial Neural Network for the classification of patterns. The data set applied is Synthetic data, which is generated from data generator by Brinchoff [8].

This paper addresses the following goals:

- Focusing on the new kind of data i.e. Trajectory data.
- Sequential Pattern mining for building classification Model.
- Artificial Neural Network can be employed for building classification model as it well enough to achieve efficiency and accuracy.
- Aiming to demonstrate that Classifier i.e. ANN plays crucial role while building the classification model and which is best classifier in trajectory classification.

The application area of Trajectory Classification includes:

- Preparing Intelligent Transportation System which aiming towards city and transportation planning; traffic congestion recognition.
- Activity Recognition aiming towards detecting suspicious movement of vehicles in real time.
- Toll booth management on Highways.

The remaining paper managed as follows: Segment II talk about the Related Work. Segment III delivers our proposed Model and Mathematical Model. Segment IV states detailed Performance Analysis of proposed system with other system. At last, Segment IV concludes with Conclusion and Future Work.

2. LITERATURE SURVEY

There are various (frequent) Sequential patterns based classification methods have been developed for various fields. Jae-Gil Lee et al.[2] employed frequent pattern based classification framework for trajectory classification on road network. By examining the demeanor of the trajectories on road network, it is important that the order of visited vehicle should be preserved with the visited location. This method pretends that partial sequential pattern allows significantly improving high efficiency without losing accuracy as they preserve order as well as location of trajectory on the road network. So, the partial sequential pattern is good feature candidate for creating

model for trajectory classification. We employ this method to our work for selecting as well as classifying the features. Jae-Gil Lee [2] employed Support Vector Machine (SVM)[12] as classifier for building the classification model. Cheng et al. [3] have proposed frequent pattern based classification method for relational data, text document and graph. This method employs frequent pattern as combined features, but we use sequential pattern. According to our model we use set of single feature as well as set of combined features. This method has developed formal strategy for determining the minimum support threshold. We employ this method in our work for developing classification model. Mining Long, Sharable Pattern (LSP) are the methods presented by Gidofalvi and Pedersen [4] for trajectories on road network. Their main objective is ride-sharing application, and long patterns are mostly used for their purpose .So, these LSP are not suitable for traffic classification on road network since they are extremely long and having performance issue. Zheng et al. [7] presented the method for intersecting locations and travel sequence from vehicle's trajectories which enables GPS device. The main idea of this method is that to tap users travel experience. This work is important in trajectory pattern mining. There are various frequent patterns based classification methods have been developed for various fields. Lodhi et al. [11], Lesli et al. [12], Deshpande et al. [13] used phrases, substring, and subgraphs as a feature candidate for proposing classification models for text document, Protein sequence, and graph respectively. Hidden Markov Model (HMM) [14] is used by Fraile and Maybank [15] for classification of vehicles trajectory. This method depends upon vehicle's motion and the sequence of internal states. It is noticed that the behavior depend upon highest probability. This method is so similar to the method which is used for speech recognition. In this method, measurement sequence is divided into overlapping sections. In each section, the vehicle's trajectory estimated by smooth function and assign the any of four category i.e. ahead, left, right or stop which is having set {a, l, r, s}.This Set classified using HMM. That is HMM gives the motion of the vehicle which uses algorithm to find vehicles highest probability. Sun, G.Z [16] projected a model of a time warping recurrent neural network (TWRNN) to handle temporal pattern classification where severely time warped and deformed data may occur, and shows to have built-in time warping ability. TWRNN is beneficial for trajectory classification over several strategies such as dynamic programming, hidden Markov models, time-delayed neural networks, and neural network finite automata. Stephen Gang Wu [17] present framework for Probabilistic Neural Network (PNN) with image and data processing techniques to implement general purpose automated leaf recognition for classification of plant. The system automatically classifies 32 kinds of plants via the leaf images loaded from digital devices such as digital cameras or scanners. PNN is having fast speed as compare to others on training and having very simple structure. A. Fehske et al. [18] proposed A New Approach to Signal Classification Using Spectral Correlation and Neural Networks and suggested a method for classifying the communication signals based on cyclic spectral analysis and pattern recognition performed by a neural network. In this method two cases were studied i.e. Classification with as well as without prior knowledge of carries and bandwidth of the signal. Guobin Ou et al. [19] focused on Multiclass neural learning i.e. Multiclass Pattern Classification which involves determining appropriate neural network architecture, encoding schemes, learning algorithms. Discussion takes place on the major approaches used in neural networks for classifying multiple classes takes place and which architecture suitable for system of multiple neural networks or a single neural network. Database Mining is motivated by the decision problem faced due to large collection of data in various organizations such as retail, IT etc. and information collected through digital devices such as GPS, RFID, Signals, Sensors, Camera's etc. There are various classification techniques employed in the process of Database mining such as Decision Tree, Rule Based Classification, Memory Based Classification, Support Vector Machine and Neural Network etc. used for building various classification model which work on data sets such as text, graphs, networks, patterns and trajectories on road network. We employ Artificial Neural Network for building Trajectory classification model which is in less use as compared to SVM [12].

3. PROPOSED MODEL

3.1 Architecture of Trajectory Classification Process using Multi-Layer ANN

Figure 1 shows Architecture of Trajectory Classification Process using Neural Network which follows supervised learning process. Supervised learning is a problem solving strategy which is driven by real business problem of history data and the
Quality of result depend upon the quality of input data sets.

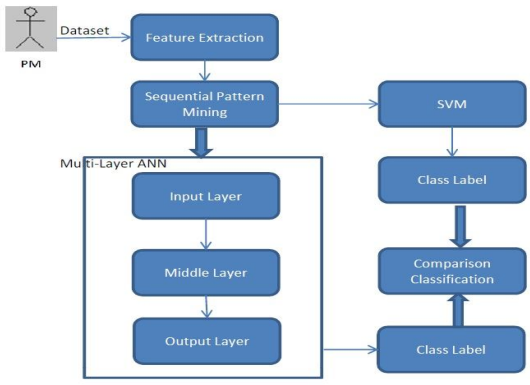

Figure 1 Architecture of Trajectory Classification

In Figure 1, Project Manager Provides input as a dataset to the classification system, which contains set of trajectories for Extraction features. Feature sets are mined using sequential patterns, which is input to the classifier. Artificial Neural network classifier plays important role in

classification process. For achieving high efficiency as well as accuracy we propose to use Multi-Layer neuron architecture of ANN. The class is assigned to the trajectory on the basis of output of ANN. Here Support Vector Machine (SVM) is used only for comparison purpose with ANN. In proposed Scheme, we employ Artificial Neural Network to construct classification model. We excerpt Feature Vectors from set of trajectories. In extraction process we use sequential patterns. The Figure 2 shows the Overall Learning Process of Multi-Layer Artificial Neural Network.

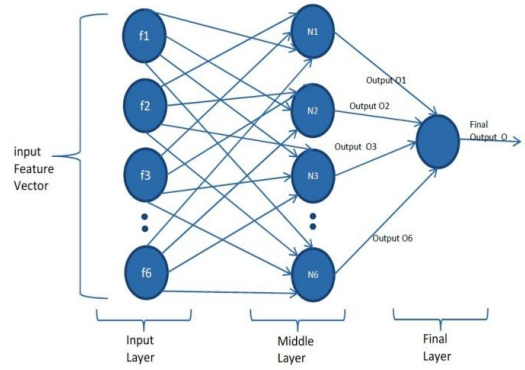

Figure 2 Learning Process of Multi-Layer Artificial Neural Network

In Figure 2 Input Layer contains the set of feature vector which are extracted from dataset. At the middle layer, ANN which produces the output on the basis of feature set. At Final Layer, the class Label is assigned to the trajectory. On the basis of class label the classification takes place. This classification can be used for various purposes like as city and transportation planning, Activity Monitoring etc..

Let set of traffic data, D ={d_0, d_1, ..., d_{n-1}, d_n} such that i < j < k then TV_i < TV_j < TV_k. Where TV is time of visit.

Let SC, set of support count of each candidate in D as

$$Sc_i = \sum_{i=0}^{n} f(i) \quad (1)$$

Where, f(i) = 1 if di is present 0 otherwise

$$D = D - di \ \exists \ di, Sci > ms \quad (2)$$

Where, ms is constant as minimum support.

Let L1= D for first iteration, for other iteration say i, Li is given by

$$Li = Li - 1 \ X \ Li - 1 \quad (3)$$

Reduce Li using

$$Li = Li - li \ \exists \ li, Sci > ms \quad (4)$$

Equation (3) and (4) are repeatedly applied m times, Where, m is a minimum length of sequential pattern. Let F = (f1, f2, ……, fn) Where F is the set of Feature Vector extracted from Traffic Data D. Let N = (N1, N2, …, Nn) Where N is the set of Neural Network Which resides on Middle Layer of Multilayer Architecture of Neural Network and produces output on the basis of Feature Vector. n = no of nodes in sequence, may vary for each sequence. Output of middle layer is given by

1 if $1/n \sum_{i=0}^{n} fi(k) > favg(k)$

Ok = -1otherwise Output of middle layer either +1 or – 1

Final Layer output $o = \sum_{i=0}^{6} oi$

. If output $O \geq$ { 0 sequence belong to class 1; Otherwise sequence belong to class 2

3.2 Algorithmic phases in Classification Process

- **Input:** Set of Feature Vector.
- **Output:** Classification.

Phase 1: Sequential Pattern mining

1. Traffic data D should be sorted based on time and order of visit to location as Litemset (Large Itemset) L.
2. Scan Traffic data D for count of each candidate as support count.
   - Compare candidate support count with minimum support count.
3. If candidate support count < minimum support count then
4. Delete candidate from D.
5. In the first iteration L1 (i.e. Li with i=1) of the algorithm, each remaining item in D is a member of the set of candidate.
6. Discover the set of frequent i-item sets, Li, using L i-1 Join L i-1 to generate a candidate set of Li-itemsets, Ci.
7. Compare candidate support count with minimum support count.
8. If candidate support count < minimum support count then
   - Delete candidate from D.
9. Repeat step 6 to 8 for some fix number of times (i.e. Minimum sequential pattern length).

Phase 2: Classification of Patterns using NN

1. Let F be the set of Feature Vector, F = {f1, f2, f3, f4, f5, f6} where f1,f2,..,f6 be the feature vector for the respective features.
2. Let n = no. of nodes in system under consideration.
3. Length of f1, f2,.., f6 = n.
4. Let Avgf1, Avgf2, Avgf3, Avgf4, Avgf5, Avgf6=0.
5. Let Sumf1, Sumf2, Sumf3, Sumf4, Sumf5, Sumf6=0.
6. For i = 0 to n repeat,

- Sumf1 = Sumf1 + f1 (i)
- Sumf2 = Sumf2 + f2 (i)
- Sumf3 = Sumf3 + f3 (i)
- Sumf4 = Sumf4 + f4 (i)
- Sumf5 = Sumf5 + f5 (i)
- Sumf6 = Sumf6 + f6 (i)

7. Avgf1 = Sumf1 / n

- Avgf2 = Sumf2 / n
- Avgf3 = Sumf3 / n
- Avgf4 = Sumf4 / n
- .Avgf5 = Sumf5 / n
- Avgf6 = Sumf6 / n

8. Let m be the no of sequences.

- For i = 0 to n repeat
- a. let k be the no of nodes in this sequence
- b. for i = 0 to k repeat

9. Finding Summation

- Sumf1 = Sumf1 + f1 (i , k)
- Sumf2 = Sumf2 + f2 (i , k)
- Sumf3 = Sumf3 + f3 (i , k)
- Sumf4 = Sumf4 + f4 (i , k)
- Sumf5 = Sumf5 + f5 (i , k)
- Sumf6 = Sumf6 + f6 (i , k)

10. O1 = O2 = O3 = O4 = O5 = O6 = 0.

11. if (Sumf1 / k) > Avgf1

then O1 = +1 else O1 = -1

12. Repeat (11) for O2, O3, O4, O5 and O6.

13. if (O1 + O2 + O3 + O4 + O5 + O6) >= 0 then

assign class1
else assign class 2

14. END

3.3 Platform

The proposed system is implemented on the Windows 7. The platform for the implementation of this system is java

4. PERFORMANCE ANALYSIS

This section deals with the experimental performance evaluation of the Classification Technique. In order to achieve high Accuracy and Efficiency the sequential patterns are mined from the dataset with the help of features. The feature associated with edge is extracted from the synthetic data. The Trajectory Classification implemented using both ANN and SVM. The comparison of ANN and SVM is done on basis of following. Input: It specifies the input module to browse particular file from host system considered for Classification. In this scenario Dataset file D1,D2,D3,D4,D5,D6,D7,D8,D9,D10 are considered as input to system. Dataset file contains the information about trajectories on the basis of parameters such as motion and edge in the form of latitude and longitude.

Table 1: Classification Efficiency

Sr. No.	Dataset	SVM	Proposed System
1	D1	10624	10605
2	D2	10232	10148
3	D3	10589	10529
4	D4	10967	10440
5	D5	10481	10495
6	D6	10444	10387
7	D7	10888	10031
8	D8	10449	10347
9	D9	10421	10399
10	D10	10431	10311

Classification Efficiency:

- The Efficiency Results in the form of milliseconds (msec)
- The Efficiency of Classifier depends upon the Classifier and the length of sequential Pattern
- The result Table 1 shows details of comparison parameters.

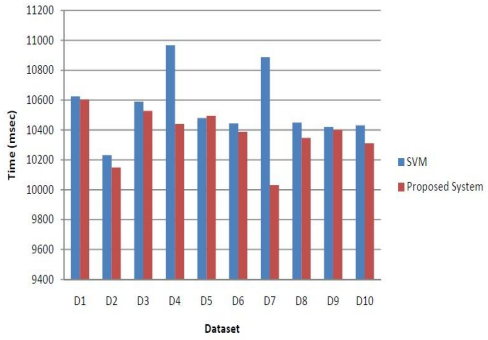

Figure 3 Classification Efficiency

Classification Accuracy (percent):

- The Accuracy Results in the form of percentage.
- The Accuracy of a Classifier is calculating by following Formula.
- Accuracy = (Tp + Tn) / Tpop

Where
Tp = Total Number of positive outcomes (sequences) from both classifier.
Tn = Total Number of positive outcomes (sequences) from both classifier.
Tpop = Number of population (sequences) of Classifier

- The result Table 2 shows details of comparison parameters.

Sr. No.	Dataset	SVM	Proposed System
1	D1	80	95
2	D2	80	90
3	D3	70	90
4	D4	70	95
5	D5	75	90
6	D6	70	85
7	D7	75	95
8	D8	70	90
9	D9	75	85
10	D10	80	95

All this analysis is shown by graph. Performance analysis shows that the proposed system has improved results.

- Improved the Classification Efficiency in Trajectory Classification: Ref Figure 3.
- Improved the Classification Accuracy in Trajectory Classification: Ref Figure 4.

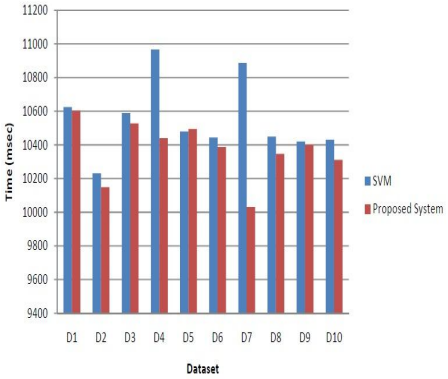

Figure 4 Classification Accuracy

5. CONCLUSION

In proposed Classification System, here we presented framework for classifying trajectories on road network using artificial neural network. Proposed Algorithm follows supervised learning process. Here, using Multi-Layer Artificial Neural Network (NN) classifier because ANN is a powerful tool for self-learning, and it can be generalized the characteristics of resource variations by proper training. In this paper, we have implemented a Classification technique of Data Mining for Trajectory Classification and analyzed the performance of the implemented system with existing system with following parameters namely efficiency and accuracy. In future work, plan to extend Artificial Neural Network approach to finding solution on Data Uncertainty and employing this framework on Real world Trajectory Data.

Acknowledgement
For proposing this model referred the IEEE Transaction paper under the title "Mining Discriminative Patterns for Classifying Trajectories on Road Networks", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 5, MAY 2011.
Also, the preliminary version of this model is accepted and published by International Journal of Computer Applications 81(13):14-16, November 2013 under the title "Classifying Trajectories on Road Network using Neural Network".

References

[1.] Deepak S Gaikwad and Usha A Jogalekar. Article: Classifying Trajectories on Road Network using Neural Network. International Journal of Computer Applications 81(13):14-16, November 2013. Published by Foundation of Computer Science, New York, USA.
[2.] Jae-Gil Lee, Jiawei Han, Fellow,Xiaolei Li, and Hong Cheng, "Mining Discriminative Patterns for Classifying Trajectories on Road Networks",IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 5, MAY 2011
[3.] H. Cheng, X. Yan, J. Han, and P.S. Yu, "Direct Discriminative Pattern Mining for Effective Classification," Proc. 24th Int'l Conf. Data Eng., pp. 169-178, Apr. 2008.
[4.] G. Gido falvi and T.B. Pedersen, Mining Long, Sharable Patterns in Trajectories of Moving Objects, GeoInformatica, vol. 13, no. 1, pp. 27-55, 2009. 724 IEEE TRANS-ACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 5, MAY 2011 TABLE 8 The Summary of the Notation Fig. 15. Estimation of the information gain as k varies.
[5.] J. Gudmundsson and M.J. Kreveld, Computing Longest Duration Flocks in Trajectory Data, Proc. 14th ACM Intl Symp. Geographic Information Systems, pp. 35-42, Nov. 2006.K. Elissa, "Title of paper if known," unpublished.
[6.] N. Mamoulis, H. Cao, G. Kollios, M. Hadjieleftheriou, Y. Tao, and D.W. Cheung, Mining, Indexing, and Querying Historical Spatiotemporal Data, Proc. ACM SIGKDD, pp. 236-245, Aug. 2004.
[7.] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, Mining Interesting Locations and Travel Sequences from GPS Trajectories, Proc. 18th Intl Conf. World Wide Web, pp. 791- 800, Apr. 2009.
[8.] T. Brinkhoff, A Framework for Generating Network-Based Moving Objects, GeoInformatica, vol. 6, no. 2, pp. 153-180, 2002.
[9.] R. Agrawal and R. Srikant, Mining Sequential Patterns, Proc. 11th Intl Conf. Data Eng., pp. 3-14, Mar. 1995.
[10.] X. Yan, J. Han, and R. Afshar, "CloSpan: Mining Closed Sequential Patterns in Large Databases," Proc. Third SIAM Int'l Conf. Data Mining, May 2003.
[11.] H. Lodhi, C. Saunders, J. Shawe-Taylor, N. Cristianini, and C. Watkins, "Text Classification

Using String Kernels," J. Machine Learning Research, vol. 2, pp. 419-444, 2002.
[12.] C.S. Leslie, E. Eskin, and W.S. Noble, "The Spectrum Kernel: A String Kernel for SVM Protein classification," Proc. Seventh Pacific Symp. Biocomputing, pp. 566-575, Jan. 2002.
[13.] M. Deshpande, M. Kuramochi, N. Wale, and G. Karypis, "Frequent Substructure-Based Approaches for Classifying Chemical Compounds," IEEE Trans. Knowledge and Data Eng., vol. 17, no. 8, pp. 1036-1050, Aug. 2005.
[14.] L.R. Rabiner and B.H. Juang, "An Introduction to Hidden Markov Models," IEEE ASSP Magazine, vol. 3, no. 1, pp. 4-16, Jan. 1986.
[15.] R. Fraile and S.J. Maybank, "Vehicle Trajectory Approximation and Classification," Proc. Ninth British Machine Vision Conf., pp. 832-840, Sept. 1998.
[16.] Sun, G.Z.; Chen, H.H.; Lee, Y.C.; Liu, Y.D. Time warping recurrent neural networks and trajectory classification, Neural Networks, 1992. IJCNN. International Joint Conference. Volume: 1 Publication Year: 1992 , Page(s): 431 - 436 vol.1
[17.] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Qiao-Liang Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network", Signal Processing and Information Technology, 2007 IEEE International Symposium, pp. 11-16 Dec 2007.
[18] Fehske, J. Gaeddert and J. H. Reed, "A New Approach to Signal Classification Using Spectral Correlation and Neural Networks", New Frontiers in Dynamic Spectrum Access Networks, 2005, DySPAN 2005. 2005 First IEEE International Symposium on, Publication Year: 2005, Page(s): 144 - 150.
[19] Guobin Ou, Yi Lu Murphey, and Lee Feldkamp, "Multiclass Pattern Classification Using Neural Networks", Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on (Volume: 4), pp. 585 - 588 Vol.4.
[20] C.-C. Chang and C.-J. Lin, LIBSVM: A Library for Support Vector Machines, http://www.csie.ntu.edu.tw/ cjlin/libsvm, 2001.

AUTHOR

Deepak S. Gaikwad is a PG Fellow of 2012 batch in Computer Engineering of Smt. Kashibai Navale College of Engineering, Vadgaon (Bk), Pune-41(MH), India. This author received Bachelor of Technology degree in Computer Science and Engineering from Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded. He is currently doing stipendiary internship at the Smt. Kashibai Navale College of Engineering, Vadgaon (Bk), Pune-41. He got the 80 percentile in the GATE 2012. His research interest includes the Data Mining and ANN. He has one journal paper on his research work of PG project as "Classifying Trajectories on Road Network using Neural Network", INTERNATIONAL JOURNAL OF COMPUTER APPLICATIONS, VOL 81-NO.13, NOV. 2013.

Usha A. Joglekar is an Asst. Professor in Smt. Kashibai Navale College of Engineering, Vadgaon (Bk), Pune-41(MH), India.