A Novel Technique for Fingerprint Classification based on Naive Bayes Classifier and Support Vector Machine

Ashish Mishra
AISECT University
Bhopal, M.P.

Preeti Maheshwary
AISECT University
Bhopal, M.P

ABSTRACT
Fingerprint classification decreases the number of possible matches in automated fingerprint identification systems by categorizing fingerprints into predefined classes. Support vector machines are widely used in pattern classification and have produced high accuracy when performing fingerprint classification. In order to effectively apply Support vector machines to multi-class fingerprint classification systems.It is proposed a novel method in which the fingerprint classification can be done by the classifier used Naïve Bayes and Support vector machines efficiently reduce the search time by restricting the subsequent searching stage to either left hand thumb and right hand thumb databases.

Keywords
Fingerprint classification, Support vector machine; FingerCode; Naïve Bayes classifier; classifier combination, directional image, feature selection, subspace classifiers.

1. INTRODUCTION
Fingerprint classification means that assignment every fingerprint to a category in a consistent and reliable method, such that an unknown fingerprint to be searched, has to be compared only to the set of fingerprints within the information belonging to a similar class. Whereas fingerprint matching is typically performed according to fingerprint micro-features, like ridge terminations and bifurcations, fingerprint classification is usually based on macro-features, like global ridge structure.The recognition of an individual needs the comparison of his/her fingerprint with all the fingerprints during a database. a typical strategy to reduce the number of comparisons throughout fingerprint retrieval and, consequently, to enhance the time interval of the recognition method, is to divide the fingerprints into some predefined categories[1].The actual control of human idea is based on recognizing patterns. The enhanced the computers get at pattern recognition. Classification is an example of pattern recognition which effort to assign each input value to one of the individual classes. Its final goal is to optimally extract patterns support on certain conditions and to separate one class from the others. The performance of the classifier is essential for good decision making. But the performance of the classifier depends greatly on the individuality of data to be categorized.

Bayes method to reduce the search space for automatic and manual recognition techniques.

In current times, more and more civilian and commercial applications are either using or actively considering using fingerprint-based recognition because of the availability of inexpensive and compact solid state scanners as well as its superior and proven recognition performance over other biometric technologies.

1.1 Fingerprint Recognition
Fingerprint recognition is more important in many situations like rate of decreasing the thief identification, in the accident cases, in the computer vision, and to overcome the time in any type of applications. So for recognition of fingerprint has done in so many ways in which by using different biometrics like face, fingers, facial patterns, voice patterns, etc., So it is possible to use fingerprints for fingerprint recognition. [2] Fingerprint recognition can be done for two purposes one for identification of fingerprint based on ridge count, ridge density, ridge thickness to valley thickness ration, ridge dimension and fingerprint patterns used for fingerprint identification. And second, for classification permits an input fingerprint to be matched against only by a subset of a database and is critical in speeding-up fingerprint identification[11].

1.2 Fingerprint Classification
Classification consists of predicting a particular outcome supported a given input. So as to predict the results the algorithmic rule processes a training set containing a collection of attributes and also the respective outcome, typically referred to as goal or prediction attribute. It is used Naïve Bayes classifier for all tools.[3] A Naïve Bayes classifier could be a simple probabilistic classifier supported applying Bayes' theorem with robust independence assumptions. Fingerprint classification supported Fingerprint recognition using association rule mining and classification approach. Fingerprint recognition uses a data Mining technique recognize the standard global representations of fingerprints[12].

1.2.1 Different Methods of Fingerprint Classification
1.2.1.1 Rule based mostly Approach
The rule based approach codifies the human skilled knowledge of manual classification like the singular point removal and therefore the geometrical shape of ridge lines . This is the simplest technique than the others. This method depends on the acquired number and positions of the extracted singular points to classify fingerprints [5].

Repetitive regularization is completed by a valid range of singular points are detected so the noise may be reduced and also improves the accuracy.

1.2.1.2 Multiple Classifier Approach
Different parameters potentially present additional information about the patterns to be classified, which can be exploited to improve performance of the algorithm. It applies orientation field to detect core and data points. Stage classification is completed, firstly K nearest neighbor to find most likely categories and second neural network for further classification. It has affect core and delta points and ridge flow as feature vectors. Using singular point,the ridge is

traced in opposite directions to get the turn number. It then uses rules for classification [4].

1.2.1.3 Correlation-based Approach
Two fingerprint images are placed over and the correlation among corresponding pixels is computed for different alignments.

1.2.1.4 Minutiae-based Approach
Minutiae recognition essentially consists of finding the alignment between the pattern and the input minutiae sets that result in the maximum number of minutiae pairings.

1.2.1.5 Non-minutiae feature-based Approach
Minutiae extraction is difficult in extremely low-quality fingerprint images, while other features of the fingerprint ridge pattern .

2. RELATED WORK
Based on our survey related to many fingerprint classification approaches have been proposed in literature and many researchers are still working in this field. Some of these works are briefly reported.

Anil K. Jain et al [2]: In this approach formulated a mathematical framework based on the Bayesian decision theory for integrating the soft biometric information with the output of the primary biometric system. They demonstrated that the utilization of ancillary user information like gender, height, and ethnicity can improve the performance of the traditional biometric systems like a fingerprint. Although these soft biometric characteristics are not as permanent and reliable as the traditional biometric identifiers like a fingerprint, they provide some information about the identity of the user that leads to higher accuracy in establishing the user identity. Experiments conducted on a database of 263 users show that the recognition performance of a fingerprint system can be improved significantly (5%) by using additional user information like gender, ethnicity, and height.

Ballan et al. in [4] first reduce image distortion and contrast, before computing the fingerprint directional image. Then, they extract the singular points from it and finally classify the fingerprint using topological and numerical considerations about these points.

Cho, Kim, Bae et al's Classification approach [5]: They described a novel fingerprint classification algorithm that apply only the information related to the core points. The algorithm detects core point(s) candidates roughly from the directional image and the type and the orientation of core point(s) are extracted for the classification step.

Yao, Marcialis, Pontil, et al's Classification approach [6]: Here, a new fingerprint classification algorithm is based on the two machine learning approaches: support vector machines and recursive neural networks . RNNs are trained on a structured representation of the fingerprint image. SVMs are combined with a new error correcting coding scheme, which, unlike earlier systems, can also exploit the information contained in ambiguous fingerprint images.

Wei, Yonghui, et al's Classification approach [7]: In this approach, a new fingerprint classification algorithm is introduced that uses some curve features of the ridge lines. The algorithm basically exploits the total direction changes of the ridge lines during classification. However, the sampling of the ridge lines still can be found to be time consuming.

3. PROPOSED SYSTEM
The goal of the proposed work is the implementation of a novel technique, based on Naïve Bayes classifier and Support vector machine for efficient training and test phases in fingerprint classification. The proposed system is based on the comparison results of classification   selected two classifiers Support vector machine and Naïve Bayes, classified on the Left and Right fingerprint classes.With more details, it is composed of two independent modules: Training Dataset Selection Module and Fingerprint Classification Module. Fig.1 shown that Flowchart of the implemented fingerprint Classification system.
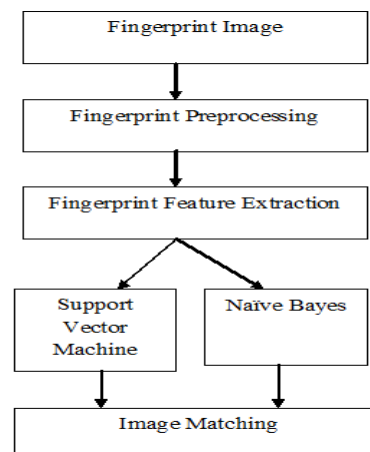

Fig. 1 Flowchart of the implemented fingerprint Classification system

3.1. Naive Bayes' Classifier
A Naive Bayes classifier could be an easy probabilistic classifier supported the Bayes theorem with robust independence assumptions. It assumes that the domain variables are independent, given the class, and every variable contains a finite number of values.

Despite their naive design and apparently simple assumptions, Naive Bayes classifiers work fine in several complicated real-world things and that they only need a small amount of training data. All model parameters (i.e., previous class possibilities and have probability distributions) are often approximated with relative frequencies from the training set.[8] These are most probability estimates of the probabilities. Previous class possibilities are also calculated using two methods:

1. By assumptive a similar category,the probability (i.e.,priors = 1 / number of classes), or

2. By calculate an estimate for the category likelihood from the training set (i.e., previous for a given category =number of samples within the category / total number of samples).

3.2. Fingerprint Dataset
CASIA Fingerprint Image Database Version 5.0 or also known as CASIA-FingerprintV5 is a fingerprint database provided by the Biometrics Ideal Test. All the images were captured using an optical fingerprint sensor, i.e., URU4000. CASIA-FingerprintV5 contains 20000 fingerprint images of 500 subjects. Each subject contributed 40 fingerprint images are 8 bit gray-level BMP files and the image resolution is 328*356. CASIA database includes eight fingerprints, i.e., thumb, index, middle, and ring fingers from both hands. thumb of each subject is used of both hands and 4 angles. Therefore, the size of the dataset is of 4000 fingerprints.

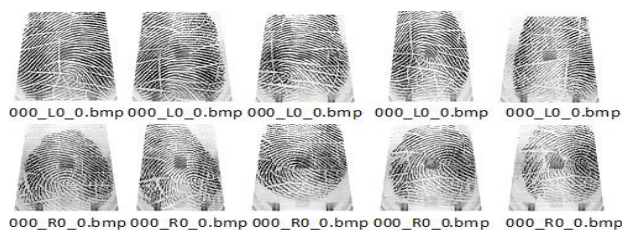Fig.2 shown that Sample Fingerprint Images in Left and Right Hand Thumb.

000_L0_0.bmp 000_L0_0.bmp 000_L0_0.bmp 000_L0_0.bmp 000_L0_0.bmp

000_R0_0.bmp 000_R0_0.bmp 000_R0_0.bmp 000_R0_0.bmp 000_R0_0.bmp

Fig. 2 SampleFingerprint Images in Left and Right Hand Thumb

The proposed approach, the classifier is trained with the Naïve Bayes, obtained from the fingerprint data set. 50% of the data set is used for training and the remaining 50% is used for testing.

The accuracy obtained is shown in Table 1.1. A similar iteration process is repeated for the three classes (W, X, Y) classified on the Left(1) and Right(2) fingerprint classes.These three classes is combination of termination X,termination Y,Angle,bifurcation X,bifurcation Y, Angle-1, Angle-2 and Angle-3.. The accuracy results obtained is shown in Table 1.2 , Table 1.3 and Table 1.4.

Table 1.1 Naïve Bayes classifiers fingerprint database

	W			X		Y			
termina tion-x	termina tion-y	Angl e	bifurc ation-x	bifurca tion-Y	Angle- 1	Angle- 2	Angle- 3	Class
72	15	1.05	198	48	2.62	0	0.52	1
60	17	0.79	91	51	-2.62	0	-1.05	1
127	27	-2.62	199	78	2.62	0.79	-0.79	1
126	33	0.52	169	81	3.14	1.57	0.52	1
178	34	3.14	188	93	3.14	-2.09	0	1
184	34	0	100	98	2.62	-2.09	0.52	1
62	37	-2.62	171	107	3.14	-2.09	0	1
126	40	0	155	114	-2.62	1.57	-0.52	1
191	40	0	171	126	2.62	-2.09	0	1
103	41	1.05	77	129	-2.62	1.57	0	1
185	41	3.14	102	130	2.36	-1.57	0.79	1
77	43	1.57	72	134	3.14	-1.57	1.05	1
159	43	0.52	58	144	-2.09	1.57	-0.52	1
96	44	-2.09	155	148	-2.62	1.57	0	1
97	28	0	208	110	2.62	-2.09	-0.52	2
170	28	3.14	204	116	2.62	1.05	-0.79	2
136	29	-0.52	140	120	2.62	-2.09	0.52	2
73	33	0.52	218	126	2.36	-2.36	-0.79	2
190	33	0	234	139	-2.62	2.09	-0.79	2
206	33	0	67	144	-2.36	1.57	0	2
220	34	0	100	164	2.09	-1.57	0.79	2
172	54	-0.52	211	176	2.62	-1.57	0	2
174	64	2.36	248	183	3.14	1.57	-0.52	2
189	159	-0.5	79	98	2.62	-1.57	0.52	1
83	160	-2.4	80	105	-2.62	1.57	0	1
205	161	2.09	NaN	0	0	0	0	1
155	166	2.36	95	111	3.14	-1.05	0.52	1
111	169	-2.6	73	116	3.14	-1.57	0.52	1
76	179	-2.4	95	117	-2.62	1.05	-0.52	1
201	181	2.09	191	118	2.36	-1.57	0	1
79	188	0.52	82	120	3.14	1.57	0	1
159	191	2.09	72	124	3.14	1.57	0	1
74	194	-2.4	187	124	2.62	0.79	-0.79	1
195	64	127	212	127	258	127	263	2
66	208	73	132	157	139	211	144	2
149	80	104	145	145	154	149	215	2
81	194	83	153	241	153	58	154	2
143	84	125	127	162	229	230	235	2
97	28	0	208	110	2.62	-2.09	-0.52	2
136	29	-0.52	140	120	2.62	-2.09	0.52	2
73	33	0.52	218	126	2.36	-2.36	-0.79	2
190	33	0	234	139	-2.62	2.09	-0.79	2

3.3. Support Vector Machines Classifier

The support-vector Classifier could be a new learning machine for two-group classification problems. The machine,abstractly implements the subsequent initiative input vectors are non-linearly mapped to a very high-dimensional feature space. During this feature space a linear decision surface is constructed. This algorithmic rule works well for both the linearly separable and non-separable classes by searching for the hyper-plane that separates the categories optimally within the feature space [10]. B. Boser suggests a training algorithmic rule that maximizes the margin between the training patterns and therefore the decision boundary is presented. The approach consists of merging the filtering method for feature space extraction and information gain .

3.4. Fingerprint Dataset

Using same dataset Table 1.1 for SVM classifiers fingerprint database.

4.   RESULTS & PERFORMANCE
4.1 Naive Bayes classifier

Table 1.2 Obtained Results of Naïve Bayes Classification applied on 3 classes (W, X, Y) on the Fingerprint CASIA DB

	W			X			Y		
	termina tion-x	termina tion-y	Angle θ	bifurca tion-x	bifurca tion-y	Angle-θ1	Angle-θ2	Angle-θ3
Accurac y (%)	83.7								
Elapsed time (seconds)	9.808167								

Table 1.3 Obtained Results of Naïve Bayes Classification applied on 3 classes (W, X, Y) on the Fingerprint CASIA DB

	W			X			Y		
	termina tion-x	termina tion-y	Angle θ	bifurca tion-x	bifurca tion-y	Angle-θ1	Angle-θ2	Angle-θ3
Accurac y (%)	82.9								
Elapsed time (seconds)	1.55315								

Table 1.4 Obtained Results of Naïve Bayes Classification applied on 3 classes (W, X, Y) on the Fingerprint CASIA DB

	W			X			Y		
	termina tion-x	termina tion-y	Angle θ	bifurca tion-x	bifurca tion-y	Angle-θ1	Angle-θ2	Angle-θ3
Accurac y (%)	87.4								
Elapsed time (seconds)	1.167786								

Accuracy is defined as the ratio of correctly classified instances of total number of instances.Fig.3 shown that Classification Accuracy of fingerprint using Naïve classifier.
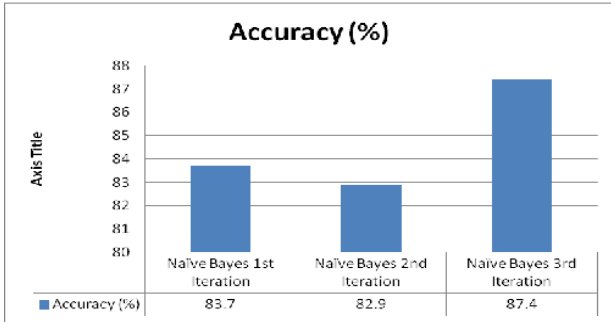
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

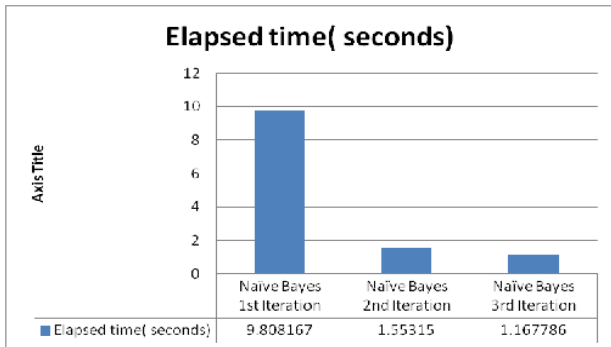**Fig.3 Classification Accuracy of fingerprint using the Naive Bayes classifier**

Fig.4 Classification Elapsed time of fingerprint using the Naive Bayes classifier

4.2 SVM classifier

Table 1.5 Obtained Results of SVM Classification applied on 3 classes (W, X, Y) on the Fingerprint CASIA DB

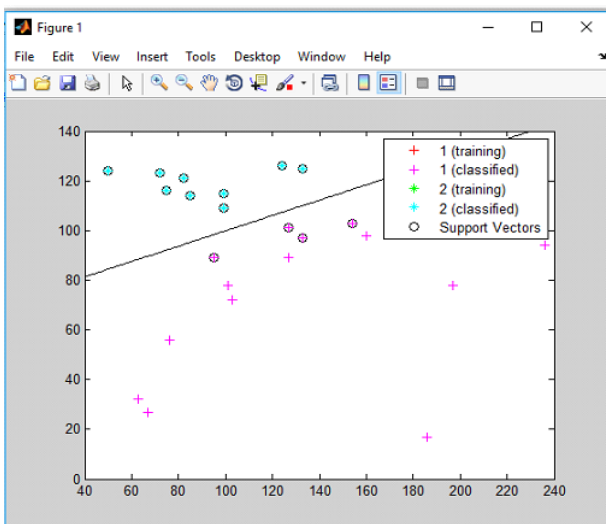|	W			X		Y		
	termina tion-x	termina tion-y	Angle θ	bifurca tion-x	bifurca tion-y	Angle-θ1	Angle-θ2	Angle-θ3
Accurac y (%)	76.06							
Elapsed time (seconds)	14.228505							

Fig.5 Plot the classification boundaries and visualize the support vector classifier

5. COMPARISON ACCURACY AND ELAPSED TIME WITH NAIVE BAYES AND SVM METHOD

In this Comparison tested the ability of classification using simple classifiers that are commonly used in image classification. In the experiments compared the results of classification selected two classifiers Naive Bayes and SVM. The best results are achieved using Naïve Bayes classifier as compared to SVM classifier. On reduced set of attributes and instances that Naïve Bayes classifier correctly classified 87.4% of instances and elapsed time is 1.167786 seconds. Based on the result, promising that with further data preprocessing and adjustment of the classifiers they can be improved. Comparison Accuracy and Elapsed Time with Naive Bayes and SVM method Shown in fig.6 and fig.7.

5.1 Confusion Matrix

The confusion matrix is used to measure the performance of two class problem for the given data set Table1.1. The right diagonal elements true positive and true negative correctly classify Instances as well as false positive and false negative incorrectly classify Instances.

Table 1.6 Confision Matrix

Actual Class	Predicted Class	
	Positive	Negative
Positive	True Positive(TP)	False Negative
Negative	False Positive(FP)	True Negative (TN)

Total number of instances = Correctly classified instance +Incorrectly classified instances
Correctly classified instance = TP + TN
Incorrectly classified instance = FP + FN

Tables 1.7 using Dataset have total no. of 135 instances. Fingerprint class has been chosen randomly from fingerprint dataset. When NaiveBayes classifier algorithm is applied to the dataset the confusion matrix is generated for Fingerprint class having two values, i.e. Left hand thumb and Right hand thumb.

Table 1.7 Naive Bayes Classifier confusion matrix

Actual Class	Predicted Class	
135	Positive	Negative
Positive	118	2
Negative	15	0

The above shown Table is about Confusion matrix that analyze that actual and predicted classification, i.e. Total no. of true positives for the class is 118 and Total no. of False positive for class 2. Total no. of true positives for the class is 15 and Total no. of false positive for the class is 0.

Tables 1.8 using Dataset have total no. of 447 instances. Fingerprint class has been chosen randomly from fingerprint dataset. When the SVM classifier algorithm is applied to the dataset the confusion matrix is generated for Fingerprint class having two values, i.e. Left hand thumb and Right hand thumb.

Table 1.8 SVM confusion matrix

Actual Class	Predicted Class	
447	Positive	Negative
Positive	332	23
Negative	84	8

The above shown Table is about Confusion matrix that calculates that actual and predicted classification, i.e. Total no. of true positives for the class is 332 and Total no. of False positive for class 23. Total no. of true positives for the class is 84 and Total no. of false positive for the class is 8.
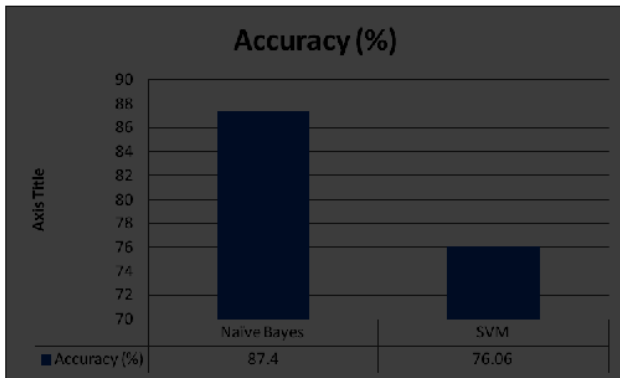

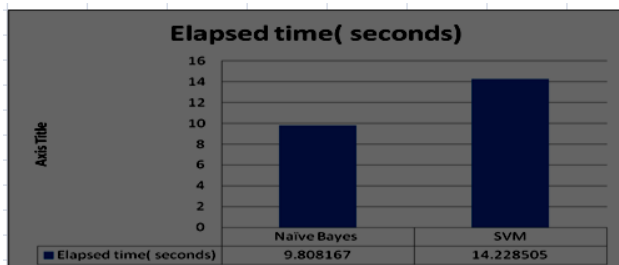**Fig.6 Comparison Accuracy with Naïve Bayes and SVMmethods**

Fig.7 Comparison Accuracy with Naïve Bayes and SVM methods

6. CONCLUSIONS AND FUTURE WORK

The fingerprint classification can be done by the classifier used Naïve Bayes and SVM efficiently reduce the search time by restricting the subsequent searching stage to either left hand thumb and right hand thumb databases. After classifying the left hand thumb and right hand thumb of the person, any of the appropriate biometric traits can be used for further classification. In this comparison the results of classification selected two classifiers Naïve Bayes and SVM. The best results are achieved using Naïve Bayes classifier as compared to SVM classifier.

Future work could contribute to performing the statistical experiments used in this project on a larger sample size, and to conduct a full analysis of the observed results. Further study into the statistical theory of fingerprint minutiae.

7. REFERENCES

[1] Biometric System Laboratory, University of Bologna.

[2] Anil K. Jain, Sarat C. Dass, and Karthik Nandakumar ,"Soft Biometric Traits for Personal Recognition Systems", Proceedings of International Conference on Biometric Authentication, LNCS 3072, pp. 731-738, Hong Kong, July 2004.

[3] Jin-Hyuk Hong, Jun-Ki Min, Ung-Keun Cho and Sung-Bae Cho," Fingerprint classification using one-vs-all support vector machines dynamically ordered with naïve Bayes classifiers", Published by Elsevier,0031-3203/$30.00 2007 Pattern Recognition Society.

[4] M. Ballan, F. A. Sakarya and B. L. Evans, "A Fingerprint Classification Technique Using Directional Images", proc. of the 31st Asilomar Conference on Signal, System and Computer,Vol. 1, pp. 101-104,1997

[5] Cho Byoung-Ho, Kim Jeung-Seop, Bae JaeHyung, Bae In-Gu, and Yoo Kee-Young. "Fingerprint Image Classification by Core Analysis," Proceedings of ICSP, 2000.

[6] Yao Y., et. al., "A new machine learning approach to fingerprint classification," 7th Congress of the Italian Association for Artificial Intelligence, pp. 57- 63, 2001.

[7] Wei L., Yonghui C., and Fang W., "Fingerprint Classification by Ridgeline and Singular Point Analysis," Congress on Image and Signal Processing, 2008.

[8] G. Vitello et Al.," A Novel Technique for Fingerprint Classification based on Fuzzy C-Means and Naive Bayes Classifier", Eighth International Conference on Complex, Intelligent and Software Intensive Systems, 2014.

[9] B. Boser, et Al, "A Training Algorithm for Optimal Margin Classifiers", Proceedings of the fifth annual workshop on Computational learning theory,1992

[10] C. Cortes et Al, "Support-Vector Networks", Machine Learning, issue 20, vol. 3, pp.273-297,1995

[11]Ashish Mishra et Al," An Efficient System for Fingerprint Finger Print Matching and Classification", ISSN 2277-128X, Volume 3, Issue 11, and Nov. 2013.

[12] Ashish Mishra et Al," A Review on Gender classification using association rule mining and classification based on Fingerprint", Fifth IEEE International Conference,page no 930-934, April 2015.