A SURVEY-CLASSIFIER FUSION

Kanchan Saxena[*1] and Vineet Richaria[2]

[*1]Computer Science, Rgpv/Lnct, Bhopal, M.P, India
Kanchansaxena4@gmail.com[1]

[2]Computer Science, Rgpv/Lnct, Bhopal, M.P, India
Vineet_rich@gmail.com[2]

Abstract: A number of classifier fusion methods have been recently developed opening an alternative approach leading to a potential improvement in the classification performance. As there is little theory of information fusion itself, currently we are faced with different methods designed for different problems and producing different results. This paper presents the survey of various classification technique for getting the optimal result applying with fusion technique. Classification is the example of supervised learning. Classification is a data mining function that assigns items in a collection to target categories or classes. The goal of classification is to accurately predict the target class for each case in the data .many classification technique are used for improving the accuracy of the classifier such as k-nearest neighbor (knn), support vector machine (svm), clustering etc. The growth of rate of data increses in current decade. The internet genrate huge amount of unstuctured data, the whole data contains of text, document, vedio and image. The gropping of data required the classification .The classification as a part of supervised learning, in this technique the gropping of data occur in a guided fashion. We insentively review various research and journal paper related to data classification used such different methodology such technique are knn(k-nearest neighbour), svm(support vector machine), clustering and classification. In recent research data mining evolved a new emerging technique such a technique is called DATA FUSION.

INTRODUCTION

Classification is a data mining function that assigns items in a collection to target categories or classes. The goal of classification is to accurately predict the target class for each case in the data. The classification has done on the basis of training set and testing set.

Suppose we have a set =[Apple,Mobile,paper,Bag,Pen,Coin]
TRAINING SET={ collection of attributes}
Apple=properties of apple={sweet,red,solid}
Mobile=properties of mobile={solid,ring,light}
Paper=properties of paper={read,write}
Bag= properties of bag={cloth,leather,solid,space}
Coin= properties of coin={solid,thin,circle}
Testset={solid,ring,light,read,wright,solid,cloth,leather,solid ,space,solid,thin, circle}.

Now we have to classify sets according to the common attributes that share
$c_1, c_2, c_3, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots cM$
cM= confusion matrix.

A confusion matrix displays the number of correct and incorrect predictions made by the model compared with the actual classifications in the test data. The matrix is *n*-by-*n*, where *n* is the number of classes. There is one important point is noted that for improving the performance of the classification must try to reducing the confusion matrix

ABOUT DATA FUSION

Merging the retrieval results of multiple systems A data fusion algorithm accepts two or more ranked lists and merges these lists into a single ranked list with the aim of providing better effectiveness than all systems used for data fusion.
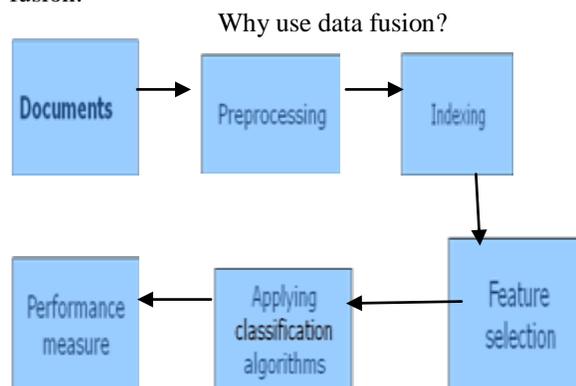
Why use data fusion?

Figure: 1

a. Combining evidence from different systems leads to performance improvement
i. Use data fusion to achieve better performance than the individual systems involved in the process
b. Same idea is also used for different query representations
ii. Fuse the results of different query representations for the same request and obtain better results

Trends of Fusion in Data Mining:

In this paper review the main uses of information fusion techniques in the field of data mining.A classification of these uses is given into three rough classes-
a. 1.preprocessing
b. 2.Building models
c. 3.Information extraction

Now a days large amount of data are available in a companies industries and researchers. because gathering

data is easy and usually expensive .however most data is raw and to be useful relevant knowledge has to be extracted from it . data mining (DM) and knowledge discovery(KNN) in databases are fields that study and provide methods for extracting this knowledge .Data mining uses information fusion technique for improving the quality to the extracted knowledge.three main uses can be distinguished:-

a) Information fusion in preprocessing- fusion is used to improve the quality of the raw data prior to the applications of data mining methods.

b) Information fusion for building model-the model built from data uses some kind of information fusion technique(eg.a particular aggregation operator to fuse partial results)

c) Information fusion used to extract information- the knowledge extracted from the data is the result of particular information fusion technique.
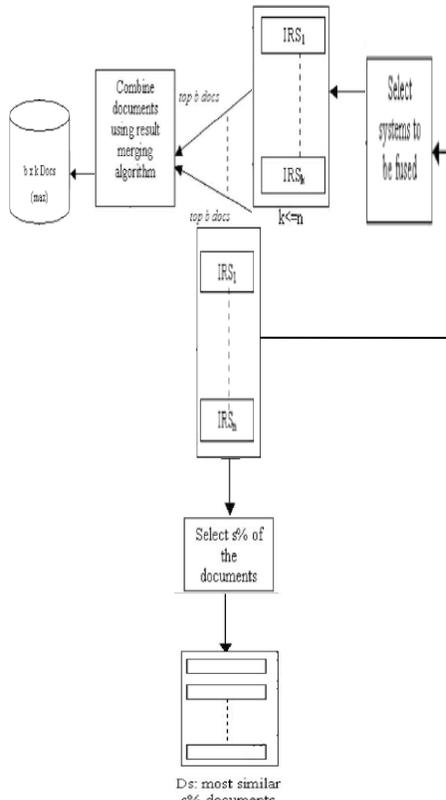(eg. An aggregated value computed from the data.)

Figure: 2 basic concepts of data fusion.

Litrature Survey and Related Work:

This section gives an extensive literature survey on the classifie rperfomence with fusion technique. We study various research paper and journal and know about the classifier performance gives better and accurate result if we use fusion. All methodology and process are not described here. But some related work in the field of classification discuss by the name of authors and their respective title.

a) BY Muhammad A.Khan, Zahoor Jan and Anwar M. Mirza ("performance analysis of classifier fusion model with minimum feature of subset and rotation of the dataset") in this paper which is in the field of the (Classification) and in this paper we investigated three aspects of classifier fusion system aaplied to the Gender Classification problem. we getting the result in this paper is that the "classification combination usually promises better performance in comparision to individual classifier but In the fusion classifier there are use diverse models like SINGLE,BEST,FIXED RULE COMBINER and Classifier Combiner model. Each of the model has its advantage. and we investigate in this paper that Fixed Combiner and Classifier Combiner produces better result than the Single Best Classifier. We use the fixed combiner as better performing model Model on the rotated data set with any number of features and using the classifier combiner model has better performance on the rotated dataset for using the minimum number of features.

b) BY Sampath Deegalla and Henrik Bostrom ("Improving Fusion of Dimentionality Reduction Methods for Nearest neighbour Classification") this paper which is in the field of (classification) ,we investigated in this paper " two novel methods for fusing features and classifiers in the conjuction with three dimentionality Reduction Methods for Nearest Neighbour classifier in high dimensions"

c) BY Norman Poh and Samy Bengio ("Using Chimeric to Construct Fusion Classifier in Biometric Authentication Tasks: AN Investication") this paper in the field of (classification), IN this paper we investigated that (" *how a model can be built using a chimeric database, an approachwhich to the best of our knowledge, has not been investigated before. One important conclusion from this preliminary study is that a fusion operator derived from a chimeric-user database does not improve nor degrade the generalization performance (on real users) with respect to training it on real users."*). The current study aims to answer the second question. Having tested on four classifiers and as many as 3380 face and speech bimodal fusion tasks (over 4 different protocols) on the BANCA database and four different fusion operators, this study shows that generating multiple chimeric databases *does not degrade nor improve* the performance of a fusion operator when tested on a real-user database with respect to using only a real-user database.

d) By Abdul Majid ,Asifullah Khan and Anwar M.Mirza ("Gender Classification Using Discrete Cosine Transformation:- A Comparision of Different Classifiers") this paper in the field of (Classification) in which we investigated that (",*problem **of** gender classification **using** a libraty **of four** hundred standard /ramal facial images employing five classfiers, namely K-mum, K-nearest neighbors. Linear Discriminant Analysis (LBA).* Gender classification system **can** be divided into two **parts:** feature extraction and classification. The main idea is to apply **DCT** to reduce the information redundancy and to compare the performance of different classifiers in that domain under different conditions. For input face images system first computes and select the limited DCT Mefficients. feeds

them **as** input to the chosen classifier. Finally classifier output prediction about gender face. ")

e) BY Claude Tremblay and Pierre valin ("Experiment of individual classifier and on a Fusion of a Set Of Classifier") this is the paper which is in the field of (Classification) in which we investigated that ("*a new method for ship infrared imagery recognition based on the fusion of individual results in order to obtain a more reliable decision* .The results indicate that individual classifiers can be a good choice. In our particular case, the individuals **DS**classifiers perform better. An advantage of this method is that we use simple algorithms.")

f) By Fabien Scalzo1, George Bebis2, Mircea Nicolescu2, Leandro Loss2 ("Feature Fusion Hierarchies for Gender Classification") deals in the field of classification in which paper we investigated a hierarchical feature fusion model for image classification that is constructed by an evolutionary learning algorithm. The model has the ability to combine local patches whose location, width and height are automatically determined during learning. The representational framework takes the form of a two-level hierarchy which combines feature fusion and decision fusion into a unified model. The structure of the hierarchy itself is constructed automatically during learning to produce optimal local feature combinations. A comparative evaluation of different classifiers is provided on a challenging gender classification image database. It demonstrates the effectiveness of these Feature Fusion Hierarchies (FFH).

g) By *Ming Li and Ronan Sleep("*IMPROVING MELODY CLASSIFICATION BY DISCRIMINANT FEATURE EXTRACTION AND FUSION ") deals in the field of classification in which paper we investigated a general approach to discriminant featureextraction and fusion, built on an optimal feature transformation for discriminant analysis. this experiments indicate that our approach can dramatically reduce the dimensionality of original feature space whilst improving its discriminant power. Our feature fusion method can be carried out in the reduced lowerdimensional subspace, resulting in a further improvement in accuracy. Our experiments concern the classification of music styles based only on the pitch sequence derived from monophonic melodies.

h) By Jiang Dong *, Dafang Zhuang, Yaohuan Huang and Jingying Fu (**"**Advances in Multi-Sensor Data Fusion: Algorithms and Applications") Describes an overview of recent advances in multi-sensor satellite image fusion. Firstly, the most popular existing fusion algorithms are introduced, with emphasis on their recent improvements. Advances in main applications fields in remote sensing, including object identification, classification, change detection and maneuvering targets tracking, are described. Both advantages and limitations of those applications are then discussed. Recommendations are addressed, including: (1) Improvements of fusion algorithms; (2) Development of "algorithm fusion" methods; (3) Establishment of an automatic quality assessment scheme.

CONCLUSION

This paper produce the survey of various paper in which we investigated that classifier combination produses the better result isteed the performing the single classifier.

REFERENCES

[1]. A. M. M. Abdul Majid, Asifullah Khan. Gender classification using discrete cosine transformation: A comparison of different classifiers. International IEEE Multi topic Conference (INMIC 2003), Islamabad, PAKISTAN, December.

[2]. D. R. B.Gabrys. Genetic algorithms in classifier fusion. Applied soft computing.

[3]. W. Y. D. Patridge. Engineering multi version neural-net systems. Neural Computation, (8), August 1996.

[4]. P. R. T. P. for Matlab implemented by R.W.P.Duin.

[5]. G.Rogova. Combining the result of several neural network classifiers. Neural Networks.

[6]. http://white.stanford.edu/dilaro/ee3684/code/male.zip

[7]. L. Kuncheva. Combining Pattern Classifiers: Methods and Algorithms. John Wiley and Sons Inc.

[8]. L. Kuncheva and C. Whitaker. Ten measures of diversity in classifier ensembles: limits for two classifiers. Proceedings of the IEE Workshop on Intelligent Sensor Processing, Birmingham, UK.

[9]. C. S. R. D. L.I. Kuncheva, C.J Whitaker. Limits on the majority voting accuracy in classifier fusion. Pattern Analysis and Applications, (6), June 2003.

[10]. C. W. L.I. Kuncheva. Feature subset for classifier combination: an enumerative experiment. Proceedings of the 2nd International Workshop on Multi Classifier Systems, Cambridge, UK, Lecture Notes in Computer Science, LNCS 2096, Springer-Verlag.

[11]. R. Z.Pan and H.Bolouri. Dimensionality reduction for face images using dct for recognition. Technical Report, Science and Technology Research center (STRC).

[12]. R. Z.Pan and H.Bolouri. Image recognition using discrete cosine transformation as dimensionality reduction. IEEEEURASIP Workshop on Nonlinear Signals and Image Processing (NISP01), June.

[13]. M. E. Aladjem, "Combined discriminant analysis with binary features" (in Bulgarian), Biocybernetics, vol. 8, pp. 57–62, 1991.

[14]. E. Alpaydin and M. I. Jordan, "Local linear perceptrons for classification," IEEE Trans. Neural Networks, vol. 7, pp. 788–792, May 1996.

[15]. J. C. Bezdek, J. M. Keller, R. Krishnapuram, and N. R. Pal, Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. Norwell, MA: Kluwer, 1999.

[16]. B. V. Dasarathy and B. V. Sheela, "A composite classifier system design: Concepts and methodology," Proc. IEEE, vol. 67, pp. 708–713, 1978.

[17]. R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis. New York: Wiley, 1973.

[18]. R. P.W. Duin, "A note on comparing classifiers," Pattern Recognit. Lett., vol. 17, pp. 529–536, 1996.

[19]. T. K. Ho, J. J. Hull, and S. N. Srihari, "Decision combination in multiple classifier systems," IEEE Trans. Pattern Anal. Machine Intell., vol. 16, pp. 66–75, Jan. 1994.

[20]. N. S. V. Rao. On Design and Performance of Metafusers, Proceedings of the Workshop on Estimation, Tracking and Fusion: A tribute to Yaakov Bar-Shalom, Monterey, CA, May 200 1.

[21]. Y. Park and J. Sklansky. Automated Design of Linear Tree Classifiers, Pattern Recognition, Vol. 23, No. 12, pp.1393-1412, 1990.

[22]. M. K. Hu. Visual Pattern Recognition by moment invariant, IRE Trans. Inform. Theory, IT-8, pp.179-187, 1962.

[23]. Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.

[24]. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[25]. I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[26]. K. Elissa, "Title of paper if known," unpublished.

[27]. R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[28]. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[29]. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[30]. Electronic Publication: Digital Object Identifiers (DOIs):

[31]. D. Kornack and P. Rakic, "Cell Proliferation without Neurogenesis in Adult Primate Neocortex," Science, vol. 294, Dec. 2001, pp. 2127-2130, doi:10.1126/ science. 1065467. (Article in a journal)

[32]. H. Goto, Y. Hasegawa, and M. Tanaka, "Efficient Scheduling Focusing on the Duality of MPL Representatives," Proc. IEEE Symp. Computational Intelligence in Scheduling (SCIS 07), IEEE Press, Dec. 2007, pp. 57-64, doi:10.1109/SCIS.2007.357670. (Article in a conference proceedings)

Short Bio Data for the Author

Kanchan Saxena m.tech student in lnct bhopal.
Mr.Vineet Richaria hod of computer science in lnct bhopal