MR Brain Image Classification: A Comparative Study on Machine Learning Methods

Shib Sankar Bhowmick[1,2,⋆], Indrajit Saha[1,⋆], Luis Rato[2], and Debotosh Bhattacharjee[1]

[1] Department of Computer Science and Engineering, Jadavpur University, Kolkata 700032, West Bengal, India.
[2] Department of Informatics, University of Evora, Evora 7004-516, Portugal.
shibsankar.ece@gmail.com, indra@icm.edu.pl, lmr@di.uevora.pt, debotosh@ieee.org

Abstract. The brain tissue classification from magnetic resonance images provides valuable insight in neurological research study. A significant number of computational methods have been developed for pixel classification of magnetic resonance brain images. Here, we have shown a comparative study of various machine learning methods for this. The results of the classifiers are evaluated through prediction error analysis and several other performance measures. It is noticed from the results that the Support Vector Machine outperformed other classifiers. The superiority of the results is also established through statistical tests called Friedman test.

Keywords: Machine Learning, Multi-spectral Magnetic Resonance Images, Supervised Classifiers, Statistical Test.

1 Introduction

Machine learning is a kind of data processing technique that deals with developing program to learn from past data. Machine learning techniques helps us to solve highly complicated problems in a efficient way by formulating programs to imitate some of the facets of human mind [1]. Thus, the application of machine learning in intelligent computer programs improves the efficiency and accuracy in decisions making situations. Classification is the most widely used machine learning technique which is capable of separating non-overlapping data in to different segments. Therefore, classification is a process of finding a set of models which distinguishes class labels of different data objects [2].

However, the design of classifier is a crucial task in machine learning research. For a given classification task, the classifier considers both the complexity in it, as well as the size of the training dataset. Theoretically the optimal classifiers are not necessarily the best practical choice if they are offering higher complexity. It has been noticed that, the performance of the classifiers depends on the application and the information available for the given problem. Here, an overview of the application of four classifiers which includes Support Vector Machine (SVM) [3], k-Nearest Neighbor (k-NN) [4], Decision Tree (DT) [5] and Naive Bayesian (NB) [6, 7] are provided in Table 1. It shows the importance of classification methods and its wide range of application areas.

To further describe the application of classifiers and compare the performance of the aforementioned classifiers among themselves, a comparative study is required. Since, the classification of Magnetic Resonance (MR) brain images into different tissue classes is very important in clinical study and neurological pathology. Also the MR images are inherently noisy and imprecise in nature, hence, classification is a challenging task for these types of images. Here, a comparative study of MR brain image classification is performed with the use of machine learning methods like Support Vector Machine, k-Nearest Neighbor, Decision Tree and Naive Bayesian classifiers. The

⋆ Correspondence should be addressed to: Shib Sankar Bhowmick (shibsankar.ece@gmail.com) and Indrajit Saha (indra@icm.edu.pl).

Table 1. Summary of Applications of Classifiers in Engineering Problems

Area	**Types of Classifier Applied with References**
Feature extraction	SVM: [8–10], k-NN: [11], DT: [12]
Micro-array data analysis	SVM: [13–15], k-NN: [16], NB: [17].
Multi-sensor data fusion	SVM: [18], k-NN: [19], NB: [20], DT: [21].
Optimal power flow	SVM: [22], k-NN: [23], NB: [24].
Parameter estimation of chemical process	SVM: [25–27], k-NN: [28], DT: [29], NB: [30].
Remote sensing	SVM: [31], k-NN: [32], DT: [33], NB: [34].
Sentiment analysis	SVM: [35], k-NN: [36], DT: [37], NB: [38].
Signal processing	SVM: [39], k-NN: [40], DT: [41], NB: [42].

performance of these classifiers is demonstrated on several normal and multiple sclerosis lesion MR brain images. Effectiveness of these classification results is established quantitatively, visually and statistically.

The paper is organized as follows. Section 2 briefly describes the background of various classification techniques along with the overview of datasets. In Section 3, the performance of the classifiers are shown on several normal and multiple sclerosis lesion magnetic resonance brain images. Finally, Section 4 concludes the paper.

2 Methods and Materials

2.1 Machine Learning Methods

Support Vector Machine: The Support Vector Machine (SVM) is a state-of-the-art classification method introduced in 1992 by Boser *et al.* [3]. For a binary classification training data problem, suppose a data set consists of N feature vectors (x_i, y_i), where $y_i \in \{+1, -1\}$, denotes the class label for the data point x_i. The problem of finding the weight vector ν can be formulated as minimizing the following function:

$$L(\nu) = \frac{1}{2} \|\nu\|^2 \tag{1}$$

subject to

$$y_i[\nu \cdot \phi(x_i) + b] \geq 1, i = 1, \ldots, N \tag{2}$$

Here, b is the bias and the function $\phi(x)$ maps the input vector to the feature vector. The SVM classifier for the case on linearly inseparable data is given by

$$f(x) = \sum_{i=1}^{N} y_i \beta_i \mathcal{K}(x_i, x) + b \tag{3}$$

where $\mathcal{K}$ is the kernel matrix, and N is the number of input patterns having nonzero values of the Langrangian multipliers β_i. These N input patterns are called support vectors, and hence the name SVM. The Langrangian multipliers β_i can be obtained by maximizing the following:

$$Q(\beta) = \sum_{i=1}^{N} \beta_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \beta_i \beta_j \mathcal{K}(x_i, x_j) \tag{4}$$

subject to

$$\sum_{i=1}^{N} y_i \beta_i = 0 \quad\quad 0 \leq \beta_i \geq \mathcal{C}, \quad i = 1, ..., N \tag{5}$$

where $\mathcal{C}$ is the cost parameter, which controls the number of non separable points. Increasing $\mathcal{C}$ will increase the number of support vectors thus allowing fewer errors, but making the boundary separating the two classes more complex. On the other hand, a low value of $\mathcal{C}$ allows more non separable points, and therefore, has a simpler boundary. Only a small fraction of the β_i coefficients are nonzero. The corresponding pairs of x_i entries are known as support vectors and they fully define the decision function. Geometrically, the support vectors are the points lying near the separating hyperplane. $\mathcal{K}(x_i, x_j) = \phi(x_i).\phi(x_j)$ is called the kernel function. The kernel function may be linear or nonlinear, like Polynomial, Sigmoidal, Radial Basis Functions (RBF), etc. RBF kernels are of the following form:

$$\mathcal{K}(x_i, x_j) = e^{-\gamma|x_i - x_j|^2} \tag{6}$$

where x_i denotes the ith data point and γ is the weight. In this paper, the above mentioned RBF kernel is used. In addition, the extended version of the two-class SVM that deals with multiclass classification problem by designing a number of one against all two-class SVMs, is used here.

Naive Bayesian Classifier: The Naive Bayes (NB) classifier [6, 7] is developed based on the Bayes' theorem. It assumes that the attributes or features are conditionally independent for the given class label y to compute the class-conditional probability. Therefore, the assumption of conditional independence is defined as follows:

$$P(X|Y = y) = \prod_{i=1}^{n} P(X_i|Y = y), \tag{7}$$

where attribute set $\{X_1, X_2, ..., X_n\}^T$ consists of n attributes. Thereafter, it uses to compute the conditional probability of each X_i for given Y. In order to classify a test data, the classifier computes the posterior probability for each class Y and it is defined as follows:

$$P(Y|X) = \frac{P(Y)\prod_{i=1}^{n} P(X_i|Y)}{P(X)}. \tag{8}$$

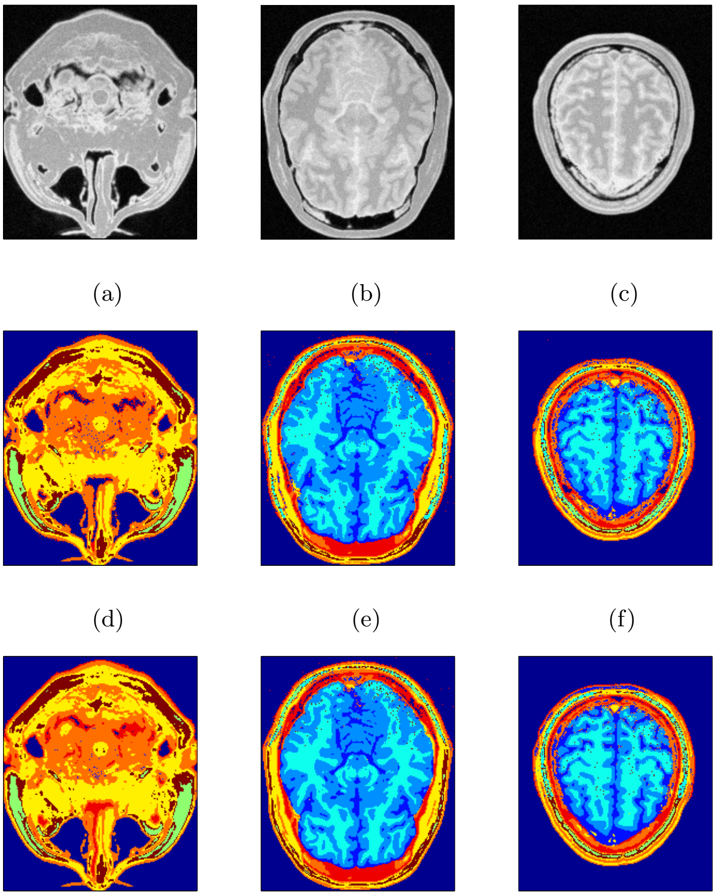
Here, the posterior probabilities are computed by multiplying the priori probabilities with the class-conditional probabilities. The priori probability of each class is calculated by the fraction of training points that belong to each class.

***k*-Nearest Neighbor:** The traditional k-NN algorithm is well-known and widely used for its simplicity and easy implementation [4]. In k-NN classifiers, each unlabeled data point is classified by the majority voting of its k-nearest neighbors in the training set. Its performance thus depends crucially on the distance metric used to identify nearest neighbors. In the absence of prior knowledge, most k-NN classifiers use simple Euclidean metric to measure the distance between data points represented as vector inputs [43]. The class label assigned to a test data point is determined by the majority voting of its k nearest neighbors. For example, for the test data points, if we consider 5-NN algorithm and found 3 nearest neighbor data points are belonging in class c_1 and other 2 data points in class c_2, then the test data point should belong to class c_1.

Decision Tree: C4.5 [5] is a widely used Decision Tree generating algorithm and the extended version of ID3 algorithm. Both the algorithms have been developed by Ross Quinlan. Moreover, the Decision Trees generated by C4.5 are often used for classification, hence, it is also known as statistical classifier. To classify the data points, C4.5 uses the concept of entropy to build the Decision Trees from a set of training data. For this purpose, at every step, the highest information gained attribute is considered. Based on that attribute, decision is taken to split the training set into one or two subsets. The process will continue recursively until all nodes are exhausted. Thereafter, depending on user given parameters, C4.5 prunes the generated tree in order to classify the test data points.

2.2 Datasets

The MR Brain Images of normal brain and multiple sclerosis lesions brain are obtained from the Brainweb database [44]. The images are available in three bands: T1-weighted, T2-weighted and proton density (pd)-weighted. In our experiment, all bands are considered together for classification. The images correspond to the 1 mm slice thickness, 3% noise (relative to the brightest tissue) and with 20% intensity nonuniformity. The images of size 217×181 are available in 181 different Z planes. For the normal brain image data, the images of the Z planes Z10, Z60 and Z130 are considered. Similarly for the multiple sclerosis lesions brain image data the images of Z planes Z40, Z90 and Z140 are used for our experiments. The ground truth information of these images is also available at the Brainweb website [44]. From the ground truth information, it is observed that each of the Z planes Z10, Z60 and Z130 for normal brain images contains nine classes and for brain images of multiple sclerosis lesions, Z planes Z40, Z90 and Z140 are having classes of nine, eleven and nine, respectively.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 1. (a), (b) and (c) are original T1-weighted MR images of the normal brains in Z10, Z60 and Z130 planes respectively, (d), (e) and (f) are MR images of the normal brains classified by SVM classifier in Z10, Z60 and Z130 planes respectively, and (g), (h) and (i) are MR images of the normal brain classified by k-NN classifier in Z10, Z60 and Z130 planes respectively.

3 Empirical Results

In this section, the experimental results of the compared machine learning methods are analyzed. For this purpose, different measures of the classifiers, i.e., prediction error analysis, evaluation of other validity measures like *Kappa*-Index (KI) [45], Minkowski Score (MS) [46] and Adjusted Rand Index (ARI) [47] as well as statistical tests of the prediction errors are discussed in following subsections.

In this experiment, the parameters of SVM such as γ for kernel function and the soft margin $\mathcal{C}$ (cost parameter), are set to be 0.5 and 2.0, respectively. Note that, RBF (Radial Basis Function) kernel is used here for SVM. The k value for the k-NN classifier is chosen as 13 for the satisfactory operation of the classifier and for the case of DT, C4.5 classifier is used.

3.1 Results and Discussions

We compare the performance of Machine Learning Methods like SVM, k-NN, C4.5 or DT and NB, in this section. As there are no separate training and testing data for the aforementioned images, hence these image pixels are randomly divided into 70% training dataset and 30% testing dataset to compute the error rate of each classifier.

Table 2. Average values of Prediction error (In %) of different Classifiers for MR brain images

MR Image		**Machine Learning Method**			
		SVM	***k*-NN**	**DT**	**NB**
Normal	**Z10**	10.39	10.47	11.29	10.74
Brain	**Z60**	11.18	11.19	11.26	11.46
	Z130	10.11	10.47	11.16	11.26
Multiple	**Z40**	18.89	19.53	19.77	19.25
Sclerosis	**Z90**	10.71	10.99	11.21	11.55
Lesion Brain	**Z140**	09.09	09.79	10.92	10.63

Table 2 shows the average results of prediction error produced by different classifiers for MR images of the above mentioned Z planes of normal and multiple sclerosis lesions brains. It is evident from the table that for all the images, the SVM classifier produces better average prediction error values compared to that produced by the other classifiers. It also appears that k-NN classifier perform reasonably good in terms of predicting average error values. Figures 1(a) to (c) show the original MR normal brain images in T1 band projected on Z10, Z60 and Z130 planes, respectively. Figures 1(d) to (f) and (g) to (i) show the segmented images of MR normal brain for SVM and k-NN on Z10, Z60 and Z130 planes, respectively. It appears from these figures that the SVM classifier has identified the different tissue classes of the normal brain images reasonably well.

On the other hand, Table 3 reports the average values of KI, MS and ARI of different classifiers for MR brain images. The KI, MS and ARI values are also found better for SVM. Moreover, it is observed that the results of SVM and k-NN are superior in their corresponding groups while the SVM performs better than the k-NN. Figures 2(a) to (c) show the T1-weighted original images corresponding to the Z planes Z40, Z90 and Z140 for the multiple sclerosis lesions brain image data. The corresponding segmented images obtained by SVM and k-NN classifier are also shown in Figures 2(d) to (f) and (g) to (i), respectively. It is clear from the figures that the SVM classifier has identified the different homogeneous regions of the images very well. Hence, form the above quantitative and visual results for both the brain image datasets it is evident that the SVM classifier outperforms all its competitors.

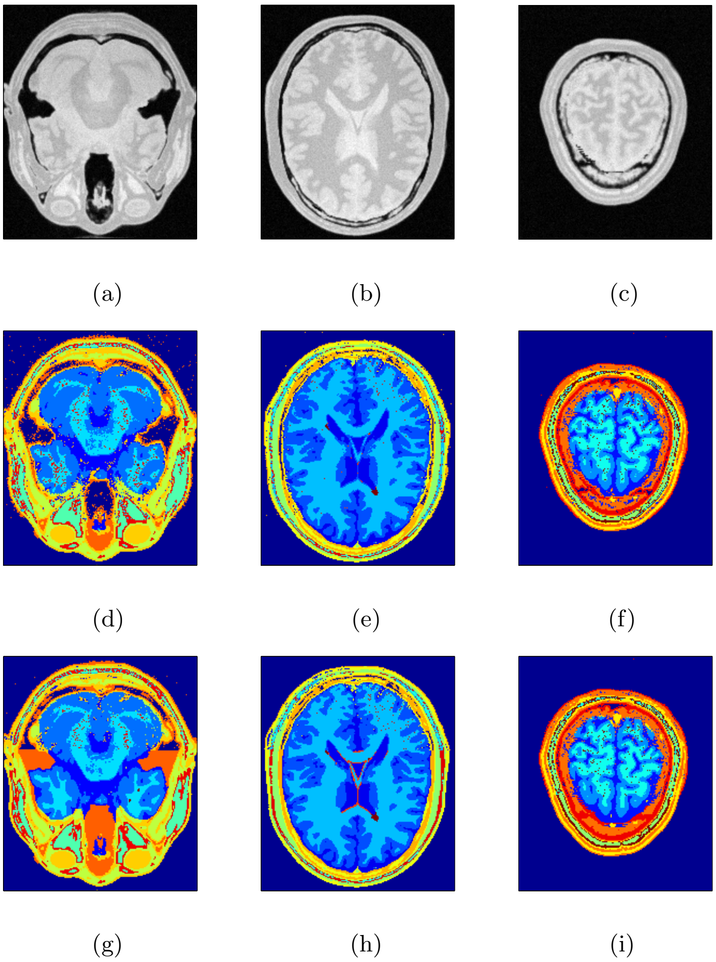


Fig. 2. (a), (b) and (c) are original T1-weighted MR images of the multiple sclerosis lesions brains in Z40, Z90 and Z140 planes respectively, (d), (e) and (f) are MR images of the multiple sclerosis lesions brains classified by SVM classifier in Z40, Z90 and Z140 planes respectively, and (g), (h) and (i) are MR images of the multiple sclerosis lesions brains classified by k-NN classifier in Z40, Z90 and Z140 planes respectively.

Table 3. Average values of KI, MS and ARI over 20 runs of different classifiers for MR brain images

Classifier	Normal Brain									Multiple Sclerosis Lesion Brain								
	Z10			Z60			Z130			Z40			Z90			Z140		
	KI	MS	ARI	KI	MS	ARI	KI	MS	ARI	KI	MS	ARI	KI	MS	ARI	KI	MS	ARI
SVM	0.83	0.45	0.67	0.86	0.39	0.80	0.85	0.45	0.67	0.77	0.48	0.61	0.85	0.44	0.71	0.87	0.39	0.80
***k*-NN**	0.87	0.39	0.80	0.87	0.37	0.80	0.87	0.37	0.80	0.81	0.44	0.71	0.89	0.36	0.81	0.08	0.39	0.80
DT	0.82	0.44	0.71	0.83	0.45	0.68	0.78	0.45	0.68	0.77	0.47	0.66	0.81	0.45	0.68	0.76	0.48	0.61
NB	0.89	0.34	0.85	0.86	0.39	0.80	0.84	0.45	0.67	0.82	0.44	0.71	0.87	0.39	0.80	0.86	0.39	0.80

3.2 Statistical Significance Test

Statistical significance of the results produced by different classifiers, are analyzed at here. For this purpose, Friedman test [48, 49] is conducted. Generally, Friedman test ranks the classifiers for each dataset separately. To compute the average rank $\mathcal{R}_j$, let r_i^j be the rank of the jth algorithm for ith dataset where the number of datasets and algorithms are N and Q respectively. Therefore the average rank $\mathcal{R}_j = \frac{1}{N}\sum_i r_i^j$.

Under the null-hypothesis, which states that all the algorithms are equivalent and so their ranks $\mathcal{R}_j$ should be equal. The Friedman statistic (chi square value) is computed as follows:

$$\chi_F^2 = \frac{12N}{Q(Q+1)}\left[\sum_j \mathcal{R}_j^2 - \frac{Q(Q+1)^2}{4}\right] \tag{9}$$

The Friedman statistic is distributed according to χ_F^2 with $Q-1$ degrees of freedom, when $N > 10$ and $Q > 5$. For a smaller number of algorithms and data sets, exact critical values are computed [50, 51].

Table 4. The Friedman ranks of all classifiers for MR brain images

MR Image		**Machine Learning Method**			
		SVM	***k*-NN**	**DT**	**NB**
Normal	**Z10**	3	2.5	4	3.5
Brain	**Z60**	3.5	2.5	3	4
	Z130	2	2.5	3.5	4
Multiple	**Z40**	2.5	3.5	4	3
Sclerosis	**Z90**	2.5	3	3.5	4
Lesion Brain	**Z140**	2	3	4	3.5
Average Rank		**2.583**	**2.833**	**3.666**	**3.666**

Table 4 reports the ranks of different classifiers for different images as well as average ranks for each classifier. From Friedman test the average rank for the classifiers SVM, k-NN, DT and NB are computed as 2.583, 2.833, 3.666 and 3.666, respectively. Moreover, from this average ranks using Equation 9, χ_F^2 is computed as 31.693. Therefore, its corresponding p value is 0.11×10^{-4} at $\alpha = 0.05$ significance level, which emphasize the acceptance of alternative hypothesis strongly. So, the results produced by the SVM are statistically significant.

4 Conclusion

In this paper, a comparative study of various machine learning methods for multispectral Magnetic Resonance brain images is conducted. For the machine learning methods, Support Vector Machine, k-Nearest Neighbor, Naive Bayesian and C4.5 or Decision Tree are used. The classification results reveals that the average values of prediction errors produced by the Support Vector Machine are better than the other classifiers. The investigation of *Kappa*-Index, Minkowski Score and Adjusted Rand Index indicates the same for Support Vector Machine. Furthermore, statistical test also shows that the average error values produced by Support Vector Machine are statistically significant. Finally, considering all conducted tests and statistics, it is established that the results of Support Vector Machine are quantitatively, visually and statistically superior than other three classifiers.

Acknowledgment

This work is partially supported by Erasmus Mundus Mobility with Asia (EMMA) grant 2012 from the European Union at the Department of Informatics, University of Evora in Portugal.

References

1. Lin, Y., Wu, M., Bloom, J.A., Cox, I.J., Miller, M.: Rotation, scale, and translation resilient public watermarking for images. IEEE Transactions on Image Processing **10**(5) (2001) 767–782
2. Chaudhary, A., Kolhe, S., Rajkamal: Machine learning techniques for mobile intelligent systems: A study. In: Proceedings of the IEEE International conference on Wireless and Optical Communications Networks (2012)
3. Boser, B.E., Guyon, I.M., N.Vapnik, V.: A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory (1992) 144–152
4. Sun, S.L.: Ensembles of feature subspaces for object detection. Lecture Notes in Computer Science **5552** (2009) 996–1004
5. Quinlan, J.R.: C4.5: Programs for machine learning. Morgan Kaufmann Publishers, San Francisco, USA (1993)
6. Duda, R.O., Hart, P.E.: Pattern Classification and Scene Analysis. John Wiley & Sons, New York (1973)
7. Langley, P., Iba, W., Thompson, K.: An analysis of bayesian classifiers. In: Proceedings of the International Conference on Artificial Intelligence (1992)
8. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: SVMeFC: SVM ensemble fuzzy clustering for satellite image segmentation. IEEE Geoscience and Remote Sensing Letters **9**(1) (2011) 52–55
9. Saha, I., Mukhopadhyay, A.: Improved crisp and fuzzy clustering techniques for categorical data. IAENG International Journal of Computer Science **35**(1) (2008) 438–450
10. Saha, I., Maulik, U., Plewczynski, D.: A new multi-objective technique for differential fuzzy clustering. Applied Soft Computing **11**(2) (2011) 2765–2776
11. Balamurugan, A.A., Rajaram, R., Pramala, S., Rajalakshmi, S., Jeyendran, C., Prakash, J.D.S.: Nb+: An improved nave bayesian algorithm. Knowledge-Based Systems **24**(5) (2011) 563–569
12. Suresh, S., Sundararajan, N., Saratchandran, P.: Risk-sensitive loss functions for sparse multi-category classification problems. Information Sciences **178**(12) (2008) 2621–2638
13. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: Improvement of new automatic differential fuzzy clustering using svm classifier for microarray analysis. Expert Systems with Applications **38**(12) (2011) 15122–15133
14. Saha, I., Plewczynski, D., Maulik, U., Bandyopadhyay, S.: Improved differential evolution for microarray analysis. International Journal of Data Mining and Bioinformatics **6**(1) (2012) 86–103
15. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: Unsupervised and supervised learning approaches together for microarray analysis. Fundamenta Informaticae **106**(1) (2011) 45–73
16. Liaw, Y.C., Leou, M.L., Wu, C.M.: Fast exact k nearest neighbors search using an orthogonal search tree. Pattern Recognition **43**(6) (2010) 2351–2358
17. Luo, H., Puthusserypady, S.: A sparse bayesian method for determination of flexible design matrix for fmri data analysis. IEEE Transactions on Circuits and Systems I: Regular Papers **52**(12) (2005) 2699–2706
18. Shaopeng, L., Gao, R.X., John, D., Staudenmayer, J., Freedson, P.S.: SVM-based multi-sensor fusion for free-living physical activity assessment. In: Proceedings of Annual International Conference of the IEEE for Engineering in Medicine and Biology Society (2011) 3188–3191
19. Liu, L., Yang, S., Wang, D.: Force-imitated particle swarm optimization using the near-neighbor effect for locating multiple optima. Information Sciences **182**(1) (2012) 139–155
20. Chair, Z., Varshney, P.K.: Distributed bayesian hypothesis testing with distributed data fusion. IEEE Transactions on Systems, Man and Cybernetics **18**(5) (1988) 695–699
21. Pal, M., Mather, P.M.: An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment **86**(4) (2003) 554–565
22. Omitaomu, O.A., Jeong, M.K., Badiru, A.B., Hines, J.W.: Online support vector regression approach for the monitoring of motor shaft misalignment and feedwater flow rate. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews **37**(5) (2007) 962–970
23. Guohui, L., Yanhong, L., Jianjun, L., Shu, L., Fumin, Y.: Continuous reverse k nearest neighbor monitoring on moving objects in road networks. Information Systems **35**(8) (2010) 860– 83
24. Jianchao, H., fu, T.Z., Bin, W.M., Haiyang, J.: The optimal bidding models for power energy producer based on static bayesian game theory. In: Proceedings of Conference of Control and Decision (2008) 3123–3128
25. Saha, I., Mazzocco, G., Plewczynski, D.: Consensus classification of human leukocyte antigens class ii proteins. Immunogenetics **65**(2) (2013) 97–105

26. Saha, I., Maulik, U., Bandyopadhyay, S., Plewczynski, D.: Fuzzy clustering of physicochemical and biochemical properties of amino acids. Amino Acids **43**(2) (2012) 583–594
27. Plewczynski, D., Basu, S., Saha, I.: AMS 4.0: Consensus prediction of post-translational modifications in protein sequences. Amino Acids **43**(2) (2012) 573–582
28. Tran, T.N., Wehrens, R., Buydens, L.M.G.: Knn density-based clustering for high dimensional multispectral images. Second GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (2003) 147–151
29. Mugambi, E.M., Hunter, A., Oatley, G., Kennedy, L.: Polynomial-fuzzy decision tree structures for classifying medical data. Knowledge-Based Systems **17**(2-4) (2004) 81–87
30. Xu, W., Baggeroer, A.B., Richmond, C.D.: Bayesian bou nds for matched-field parameter estimation. IEEE Transactions on Signal Processing **52**(12) (2004) 3293–3305
31. Bruzzone, L., Mingmin, C., Marconcini, M.: A novel transductive svm for semisupervised classification of remote-sensing images. IEEE Transactions on Geoscience and Remote Sensing **44**(11) (2006) 3363–3373
32. Zhu, H., Basir, O.: An adaptive fuzzy evidential nearest neighbor formulation for classifying remote sensing images. IEEE Transactions on Geoscience and Remote Sensing **43**(8) (2005) 1874–1889
33. Moustakidis, S., Mallinis, G., Koutsias, N., B.Theocharis, J., Petridis, V.: SVM-based fuzzy decision trees for classification of high spatial resolution remote sensing images. IEEE Transactions on Geoscience and Remote Sensing **50**(1) (2012) 149–169
34. Davis, D.T., Chen, Z., Hwang, J.N., Tsang, L., Njoku, E.: Solving inverse problems by bayesian iterative inversion of a forward model with applications to parameter mapping using smmr remote sensing data. IEEE Transactions on Geoscience and Remote Sensing **33**(5) (1995) 1182–1193
35. Huang, Z., Chen, H., Hsu, C.J., Chen, W.H., Wu, S.: Credit rating analysis with support vector machines and neural networks: a market comparative study. Decision Support Systems **37**(4) (2004) 543–558
36. Fan, W., Sun, S.: Sentiment classification for online comments on chinese news. In: Proceedings of International Conference of Computer Application and System Modeling **4** (2010) 740–745
37. Sethi, I.K., Yoo, J.H.: Design of multicategory multifeature split decision trees using perceptron learning. Pattern Recognition **27**(7) (1994) 939–947
38. Ghorpade, T., Ragha, L.: Featured based sentiment classification for hotel reviews using nlp and bayesian classification. In: Proceedings of International Conference of Communication, Information Computing Technology (2012) 1–5
39. Tran, Q.A., Li, X., Duan, H.: Efficient performance estimate for one-class support vector machine. Pattern Recognition Letters **26**(8) (2005) 1174–1182
40. Wu, Y., Ianakiev, K., Govindaraju, V.: Improved k-nearest neighbor classification. Pattern Recognition **35**(10) (2002) 2311–2318
41. Rivera, F., Sanchez-Elez, M., Fernandez, M., Hermida, R., Bagherzadeh, N.: Efficient mapping of hierarchical trees on coarse-grain reconfigurable architectures. In: Proceedings of International Conference of Hardware/Software Codesign and System Synthesis (2004) 30–35
42. hua, T.J.: Research of vehicle video image recognition technology based on naive bayesian classification model. In: Proceedings of Third International Conference of Information and Computing **2** (2010) 17–20
43. Weinberger, K.Q., Saul, L.K.: Distance metric learning for large margin nearest neighbor classification. The Journal of Machine Learning Research **10** (2009) 207–244
44. Brainweb: Simulated brain database. `http://www.bic.mni.mcgill.ca/brainweb`
45. Cohen, J.A.: Coefficient of agreement for nominal scales. Educational and Psychological Measurement **20**(1) (1960) 37–46
46. Jardine, N., Sibson, R.: Mathematical Taxonomy. John Wiley and Sons, New Jersey, USA (1971)
47. Yeung, K.Y., Ruzzo, W.L.: An empirical study on principal component analysis for clustering gene expression data. Bioinformatics **17** (2001) 763–774
48. Friedman, M.: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the American Statistical Association **32** (1937) 675–701
49. Friedman, M.: A comparison of alternative tests of significance for the problem of m rankings. Annals of Mathematical Statistics **11** (1940) 86–92
50. Zar, J.H.: Biostatistical Analysis (4th Edition). Prentice Hall, New Jersey, USA (1998)
51. Sheskin, D.J.: Handbook of parametric and nonparametric statistical procedures. Chapman & Hall/CRC, London, UK (2000)