Networks in Smooth Estimates of Normal Mixtures
Trade Off Strategies In Evaluating Diagnostic Tests: Neural

R.Selvaraj and G.Gopal

Department of Statistics,
University of Madras,
Chennai 600 005, India
E-Mail Address: rselvarajin@yahoo.com; govgopal@indya.com

SUMMARY
The unbounded multimodal likelihood surface causes analysis of mixtures of normal distribution difficult. Although Bayesian procedures have theoretical advantage of demonstrating consistency of posterior distribution (Roeder and Wasserman), there are practical difficulties in estimating the likelihood surface and the posterior distribution becomes multimodal (Mengerson and Robert). Under these circumstances, standard methods for examining posterior distribution gets trapped with their mean placed in other data clouds producing small variance and providing inaccurate picture of the posterior distribution. This work aims at a prior adjustment using neural networks so that these modes at the extremes of the data clouds gets a faster descend, producing a smooth distance measure in the evaluation of diagnostic tests.

1.INTRODUCTION

Recently a large number of learning algorithms have been used for classification purposes(Desai and Chidananda gowda). A new proposed Neural Pattern Recognition Technique (NPRT) based clustering with varying probability $P(w_i)$ produced clusters that are not well isolated and are found to be overlapping(Selvaraj and Gopal).The bimodality of the test results are seen when drawn as graph for each category and for each antigen. These extend to several modes and components when more antigens are taken for analysis as revealed by (Richardson and Green). Besides, posterior distribution becomes multimodal (Mengerson and Robert) with Markov chain Monto carlo (MCMC) implementation.

2.MATERIALS AND METHODS:

Study Population: A total of 273 finger-prick specimens, were collected from selected categories of subjects from Trivellore (a place adjacent to Chennai, South India) BCG Trial area. The categories of the sample included with antigens used and design adopted are detailed in (Alamelu Raja and others). The data set forms a mixture of normal distributions with two modes Chart 1. These extend to several modes on inclusion of the third antigen (PPD).

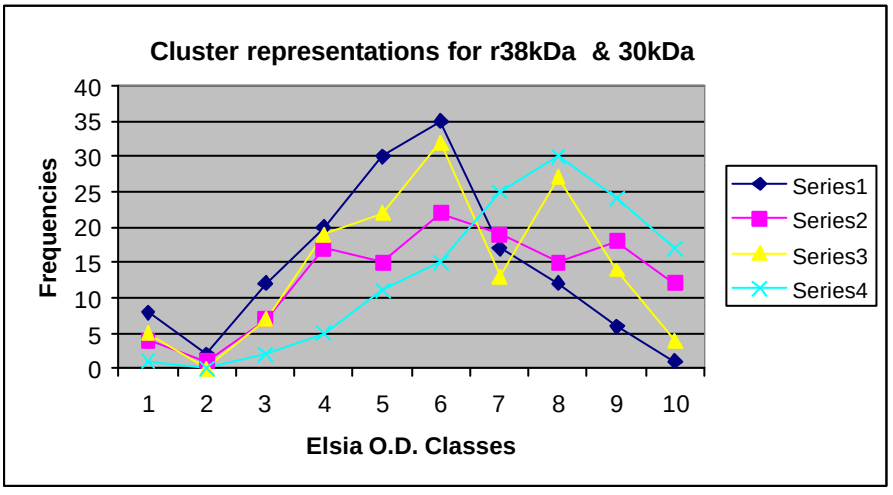

Chart 1 Series 1 – r38kDa Cluster 1 Series 2 – r38kDa Cluster 2

Series 3 – 30kDa Cluster 1 Series 4 – 30kDa Cluster 2

Let

$$f(y/\theta) = \sum_{j=1}^{k} P_j N(m_j, s_j^2)(y) \quad \ldots \quad (1)$$

where P=($p_{1,..,}p_k$) is a vector and $\sum P_j = 1 \ \forall \ p_j > 0$ denotes the normal density with mean m_j and variance s_j^2 computed at y. The likelihood surface is unbounded for values of s_j^2 when there is a small cluster of data with a s_j^2. In order to avoid such situation, we postulate

$$R = \log_{10}\left(\frac{\max s_j^2 - obs\, s_j^2}{obs\, s_j^2}\right) \quad (2)$$

Demonstrating faster tail descend at the rate of $\exp(-R^2)$

Majority of the modes occurs for large values of R. By incorporating (2) many of the modes are smoothened. Better visual evaluation of a distance measure is seen when there is less number of modes in the posterior predictive density. We have focused on two-component mixture and re parameterization as done in (Roeder and Wasserman). Fitting a mixture of two normal distributions (1) using the neural computation algorithm for R after the initial definition (2). The predictive density was calculated.

3.RESULTS:

The $r_{j\,is}$ defined to activate smoothness of predictive density is activated.

Table 1: Cluster 1 Precisions for the two antigens.

R_{1i}	R_{2i}	Group	R38kDa σ_{1j}	30kDa σ_{2j}	r_{1j}	$.r_{2j}$
0.833	-0.821	Category 1	.332	.427	9.072	5.485
-0.758	0	Category 2	.336	.398	8.858	6.313
0	-1.085	Category 3	.130	.414	10.406	5.834
0.149	0.603	Category 4	.481	.504	4.322	3.937
-0.199	-0.445	Overall	.396	.464	6.377	4.645

Table 2: Cluster 2 Precisions for the antigens

R_{1i}	R_{2i}	Group	R38kDaσ_{1j}	30kDa σ_{2j}	r_{1j}	r_{2j}
0.237	-0.723	Category 4	.512	.434	3.815	5.309
-0.078	-0.836	Category 5	.420	.426	5.669	5.510
0.193	-0.404	Category 6	.487	.470	4.065	4.527
0.176	-0.653	Overall	.490	.440	4.165	5.165

consolidating where few observations form a component and developing a distance measure.

$$R_{1i} = \log\left(\frac{\max r_j - obs\ r_j}{obs\ r_j}\right)$$

R_{1i}', R_{2i} 's are obtained for various r_{ij}' s. A quantity that measures the relative difference of the variances σ_j^2. Since R plays the central role in mixture inference, adjusting R becomes easy to implement and generalize components and mixtures. The priors show better results with neural computing implementation. Although we were able to estimate the priors, there were difficulties in estimating the posterior as explained by (Roeder and Wassermann). On implementing the basic neural computing techniques, the optimal weight from the various R_{1i} , R_{2i} 's were seen to be(R_{1i}= -0.199) for the first cluster and (R_{1i}= 0.193) for the second cluster.

4.DISCUSSION:

Kernal estimate though obtains a smooth parameter for the two components, it has its inability to deal with the tails of the distribution without over smoothing the main part of the density. Kernal density approach to estimation is resistant to the effect of outliers (Silvermann). The two other possibilities namely Explanatory projection pursuit density estimation procedure that have been considered in(Flick , Jones and others) and non parameteric estimation using group membership probabilities using multivariate thin plate splines (Villalobes and Wahba). Here, we have resorted to neural computing approach to obtain a smoothing prior structure for R. A issue becomes complicated when the group conditional densities are not unknown and neural network have adopted the following two sampling designs under which the training data are obtained. The first being mixture sampling and the second being separate sampling.

These are sample based parametric allocation rules. This implies that the estimated group conditional densities are available for well-separated groups. It is the fit in the tails that are more crucial than in the main part of the distributions. Further efforts have to be undertaken for smoothing the main part of the distribution using neural networks.

REFERENCE

1. Alamelu Raja, Acharyulu,G.S., Selvaraj, R. and Abdul Khudoos: Evaluation of antibody level to purified mycobacterial antigens for identification of tuberculosis infection. BioSciences;2001.21(1) pp:63-69.
2. Desai.P.V. and Chidananda gowda. K(1993): Efficient training and clustering using Self Organizing Maps- Proceedings of the national conference of networks,100-109.
3. Flick.T.E,Jones, L.K,Priest .R.G and Herman.C(1990): Pattern classification using projection pursuit . Pattern Recognition 23, 1367-1376.
4. Mengerson,K.L., and Robert C(1996) – Testing for mixtures : A Bayesian entropic approach . In Bayesian Statristics approach 5 (J.M. Bernade, J.O. Berger, A.P.David & Aims Smith edn.) Clarendon Press pp 255-276.
5. Richardson and Green (1997) : On Baysian analysis of mixture with an unknown number of components(with discussion). JRSS B 59: 731-792.
6. Roeder.K & Waserman,L.A., (1997) Practical Bayesian density estimations using mixtures of normal JASA 92, 894-902.
7. Selvaraj,R.,Gopal,G., Alamelu raja and Kumaraswamy,V., (2001): Pattern Recognition Technique in immunological antigenic test to identify M.tuberculous infection 2001(under revision Tubercle and Lung Disease.)
8. Silvermann,B.W.(1986):Density stimation for statistics and data analysis, Toronto :Chapman and Hall.
9. Villalobes and Wahba.G (1983): Multivariate thin plate spline estimates for the posterior probabilities in the classification problem. Commun. Statist-theory Meth 12,1449-1479.