Comparison of neural networks and discriminant analysis for prenatal factors in Kinshasa

Donking KIALANDA NSIDIOVOVA: donkingkialanda@gmail.com, Simon NTUMBA BADIBANGA, Eugène MBUYI MUKENDI
Mathematics and IT Department, Sciences Faculty, University of Kinshasa

Introduction: *Decisions are made based on data analysis. Most intelligent systems are developed to make good decision by using datamining methods. How to choose the right learning coefficient to have a good performance for these methods? More babies in the world die due to premature birth. The survival rate depends to each country. We apply these methods to analyze the prenatal factors in Kinshasa to appreciate trends.*

Materials and methods

- Neural network allow, from input information, determine wich cluster it belongs. The network must learn to produce such result. We use supervised learning in a network with three layers: input, hidden and output. MNN algorithm implemented by Microsoft SQL Server 2012 is used to assess our network. A low learning coefficient leads slow convergence and the high may produce oscillations.
- Discriminant analysis separates clusters with linear combinations descriptors. Decision depends on rules to produce a minimum error in future. To estimate the performance of discriminant functions, we have used test sample and bootstrap validation in SPAD 5.0.

Results and Discussion

We have considered 17 hospitals, 13 predictor variables and 1 to predict with 390 women. We have set up learning (289) and test (101) samples. We have established 10 different networks including 5 with test sample and other without sample. We vary the percentage of test data and obtain :

	% of test data	10	30	50	70	90
With test samples	Correct filling rate	53,85	68,32	72,28	72,28	69,31
	% of correct classified	66,67	68,32	72,28	72,28	69,31
Without test samples	Correct filling rate	**74,36**	64,10	65,64	63,74	57,55
	% of correct classified	**74,36**	64,10	65,64	64,10	57,55

The best network reaches 74.36% of correct-classified data. 25.64% of the sample data are misclassified. The assessment of our discriminant function by test sample gives 56.99% of correct classified and 30.69% misclassified. 12 gave birth prematurely while the discriminant function provides non-premature birth and 19 others are normal deliveries while they are not. The Discriminant function is: F = 0,9342X1-0,9357X2-1,4596X3 +0,0468X4+0,0317X5-0,0192X6-1,2741X7

With X1: total number of pregnancy, X2: number of pregnancy in term, X3: gestational age, X4: age of the baby, X5: expansion on arrival for delivery, X6: cervical effacement on arrival for delivery, and X7: fetal weight. These 7 variables are the cause of prematurity in Kinshasa.

Conclusion

we have used neural networks and discriminant analysis applied to the same data. We have interested by the performance of neural network by varying the percentage of test data and maximum number of cases in test samples. Neural network provides more reliability than discriminant analysis which gives a better rate (74.36%) then discriminant analysis (56.99%). However, Discriminant analysis gives details of causes (7 variables) of prematurity in Kinshasa.

References

1. Blencowe H, Cousens S, Oestergaard M, Chou D, Moller AB, Narwal R, Adler A, Garcia CV, Rohde S, Say L, Lawn JE. National, regional and worldwide estimates of preterm birth Estimates from 2010, The Lancet, 9; 379(9832): pages 2162-2172, June 2012.
2. Daniel T. Larose, Exploration de données : Method and datamining models, Vuibert, Paris, 2012.
3. Gilles Celeux and Jean-Pierre Nakache, discriminant analysis on quality variables, Polytechnica, Paris, 1994.
4. Pierre Borne, Mohamed Benrejeb and Joseph Haggège, Neural network : presentation and applications, Ophrys, Paris, 2007.
5. Phyllis Zelkowitz, Prematurity and its impact on psychosocial development and child emotion, 2nd ed., McGill University, Montreal, July 2008.
6. Romain Lucas Glèlè Kakaï, Règles de classement et taux d'erreur en Analyse discriminante, ed. european universities , October 2010.
7. Stéphane Tufféry, Case studies in decisional statistic, ed. Technip, Paris, 2009.
8. Witten Ian, Frank Eibe, Hall, Mark A, Datamining: Practical machine learning tools and techniques 3rd edition, Elsevier science and technology books, Amsterdam, 2011.