International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463
Vol. 3 Issue 7, July-2014, pp: (412-416), Impact Factor: 1.252, Available online at: www.erpublications.com

Improved Fuzzy Logic Classification Approach for Non Linear Time Series Analysis in Healthcare Data Set

Dr. Manish Pandey
Professor, School of Engineering and Technology, Shoolini University, Solan, Himanchal Pradesh, India

---

Abstract: Data mining techniques are frequently used to extract the disease related factors from the huge datasets. Data mining is the task of discovering formerly unknown, appropriate patterns and relationships in huge datasets. Generally, each data mining task differs in the type of knowledge it extracts and the kind of data demonstration it uses to convey the discovered information. Forecasting is a prediction of what will occur in the future, and it is an uncertain process. Because of the uncertainty, the accuracy of a forecast is as important as the outcome predicted by forecasting the independent variables. A forecast control must be used to find out if the accuracy of the forecast is within satisfactory limits. Two widely used methods of forecast control are a tracking signal, and statistical control limits. Incorporating seasonality in a forecast is useful when the time series has both trend and seasonal components. The final step in the forecast is to use the seasonal index to adjust the trend projection. One simple way to forecast using a seasonal adjustment is to use a seasonal factor in combination with an appropriate underlying trend of total value of cycles.

Keywords: Fuzzy Membership Function, Exponential Function, Robust Approach, Dimensionality, Regression

---

Introduction

Data mining techniques can be deployed for discovering the patterns of clinical sequences. Based on the patient record data, administrative data, clinical data and evidence based policy, mining process is functional. With their usage, we can detect the configuration of clinical paths and the pattern sequence among activities, which human beings could barely find. The development of clinical pathways is knowledge concentrated and it requires the cooperation among knowledge workers, clinicians, nurses and clinical management. The alternative to the linear regression technique of a dependent variable on explanatory variables would be the linear regression model of a fuzzy logic where dependent variable exists on the same explanatory variables. A dependent variable may be considered as some kind of membership function showing whether it is true that the observed value of the dependent variable varies into the given set of data.

In the same way, we can introduce into analysis a dependent variable as the case of fuzzy logic, which will be a fuzzy logical function denoting the degree the observed value of the dependent variable and a weighted function in the given category. In this paper we propose the use of exponential equation in place of linear equation, in which fuzzy membership function as a dependent variable. This fuzzy membership function denotes the dependant slope of the curve which is varying with the weight of dependant variable. In the paper, we show that the exponential model is more sensitive than the linear regression model of the fuzzy dependent variable. In the paper we present the results from the real study of medical data set where data of patients is a varying factor for the futuristic analysis of that dataset. We systematically test exponential regression model and the linear regression model with the fuzzy dependent variable. Using this fuzzy logic based function as a dependent variable in a linear regression model is another method to deal with uncertain categories in analysis of medical dataset.

To make a medical assessment that may have vital cost for the object under medical attention. The inherited sources of inaccuracy in the medical sector have long been a subject of debate between various parties, but here I propose to classify them according to the weighted slope depending to the variation of input variable that of the members of research group in such an interesting environment of medical engineering. This chapter surveys the use of fuzzy logic in clinical dataset, and it is based on search in variations in medical databases. In this, we tested the potency of the exponential non linear regression equation comparative to linear equation to small variations in the large dataset of medical dataset.

International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463
Vol. 3 Issue 7, July-2014, pp: (412-416), Impact Factor: 1.252, Available online at: www.erpublications.com

Literature Review

Neural networks can solve your prediction, classification, forecasting, and decision making problems accurately, quickly, and simply. The applications of neural networks are potentially limitless. Our customers have created an impressive suite of applications with our software tools. This is a brief discussion of some of the applications that have come to our attention. Many others exist and we would be happy to hear from any of our customers.[25] As an alternative robust approach to the problem of uncertain medical categories, we propose to use the linear regression model with the fuzzy membership function as a dependent variable. This fuzzy membership function denotes to what degree the value of the underlying (continuous) outcome falls below or above the dichotomization cut-off point. In the paper, we demonstrate that the linear regression model of the fuzzy dependent variable can be insensitive against the uncertainty in the cut-off point location.[1]

Machine learning classification techniques provide support for the decision-making process in many areas of health care, including screening, diagnosis, prognosis, monitoring, therapy, survival analysis, and hospital management [19]. Information about the patient given its subjectivity can be channelled towards a number of subsets, all of which include uncertainty. Also, as this information passes successively between patients, it becomes noisy and distorted, hence adding more uncertainty to medical history of patients, which is supplied.

Testing the strength of Exponential Regression Model over Linear Regression Model

In this problem we consider a medical dataset containing observations of patients in 12 months of the continuous dependent variable Y and independent variable X. Table 1 shows the medical dataset of patients over three years and forth year predicted value through Time Series Analysis. Let y_j denote the value of the variable y for observations, j (j = 1,…, N) where N is number of months in a year, and let x_i be the observed value of the independent variable x for observation of number of patients. Suppose we have constants ‘a’ and ‘b’ for an exponential function

$y_j = a\, e^{bxi}$ (1)

Where number of patients ‘y’ depends on the exponential function of ‘x’ with constants ‘a’ and ‘b’. Now taking logarithmic analysis in both sides for the calculation of the two constants.

$\log_e y = \log_e a\, e^{bx}$ (2)

$\log_e y = \log_e a + \log_e e^{bx}$ (3)

$\log_e y = \log_e a + bx$ (4)

Now consider $Y = \log_e y$; $A = \log_e a$ and $X = x – M$

Where M is the mean of N observations.

We get the equation

Y = A + bX (which is a linear equation of exponential values of a dependant variable X)

To calculate the values of A and b we have the following equations:

$$A = \frac{\sum Y - b \sum X}{n} \qquad (5)$$

and

$$b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} \qquad (6)$$

now

X = x - M

X = x – 6.5 (6.5 is the mean value of 12 observations of x)

$\sum X = \sum x - \sum 6.5$

We get,

$\sum X = 0$

Putting $\sum X = 0$ in above values of constants 'A' and 'b' we get,

$$A = \frac{\sum Y}{n} \quad (7)$$

and

$$b = \frac{\sum X\,Y}{\sum X^2} \quad (8)$$

for the year 2007, after putting the values of X, Y and n (as table 2) we get the values of the constants:

b = 0.000753496 and A = 1.67
Taking antilog of A we get
a = antilog A = 5.29

for the year 2008, after putting the values of X, Y and n (as table 3) we get the values of the constants:

b = 0.005326188 and A = 1.67
Taking antilog of A we get
a = antilog A = 5.29

for the year 2009, after putting the values of X, Y and n (as table 4) we get the values of the constants:

b = 0.007793100 and A = 1.67
Taking antilog of A we get
a = antilog A = 5.29

for the predicted year 2010, after putting the values of X, Y and n (as table 5) we get the values of the constants:

b = 0.004031928 a\d A = 1.67
Taking antilog of A we get
a = antilog A = 5.29

Fuzzy sets are known for their ability to introduce notions of continuity into deductive thinking [14]. Many practical applications of fuzzy logic in medicine are known to use its continuous subset features such as fuzzy scores, continuous version of conventional scoring systems. We describe a novel approach for the comparison of actual data and the tentative predicted data of number of patients. The approach is demonstrated on three standard medical data sets of three different years and the forth year's predicted value. The approach should be beneficial in situations categorized by high data dimensionality. Table 1 shows the dataset for three year's number of patients and 4^{th} year predicted data using linear regression. After applying exponential functional equation we get the exponential coefficients. Table 2 shows the year and month wise dataset of fuzzy coefficients. In table 3 we clip the higher values to get a non linear graph for the number of patients. After clipping the higher values in fuzzy coefficients we add the average value of different years in the fuzzy coefficients and get the fuzzy data for the number of patients. Table 4 shows the fuzzy dataset of number of patients for different years. Figure 1 shows the fuzz patients dataset graph for different years.

Conclusion

This technique provides support for the decision-making process in many areas of health care and hospital management. This Tool includes fuzzy approach based on the exponential series. Medical applications of a larger datasets include modelling of data to analyse health surveys. Accuracy is very important in the medical applications. Many medical applications are characterized by a large number of medical data sets and relatively few training examples. Moreover, using the full set of description to define weighted slop in the predicted value may lead to overfitting due to the dimensionality problem. This paper proposes the approach for a formalized paradigm for multi-phase exponential regression capable of producing a high-degree polynomial model in effective predictors.

References

[1]. Arkady Bolotin, "Fuzzy Logic Approach to Robust Regression Models of Uncertain Medical Categories", International Journal of Biological and Life Sciences 2:4 2006
[2]. A.J.C. Sharkey, N.E. Sharkey, S.S. Cross, Adapting an ensemble approach for the diagnosis of breast cancer, Proceedings of the International Conference on Artificial Neural Networks (1998) 281-286.
[3]. Bolotin A., "Fuzzification of Linear Regression Models with Indicator Variables in Medical Decision Making", In: Proceedings of the CIMCA 2005, IEEE, 2006, Vol. 1, pp. 572-577.
[4]. Bolotin A., "Replacing indicator variables by fuzzy membership func- tions in statistical regression models: Examples of epidemiological studies" In: Lecture Notes in Computer Science: Biological and Medical Data Analysis, Springer, 2004, pp. 251-258.
[5]. Bolotin A., "Uncertain categories in medical data analysis", In: Proceed- ings of the IPMU 2006, Paris, 2006, to be printed.
[6]. B.O. Peters, G. Pfurtscheller, H. Flyvbjerg, Automatic differentiation of multichannel EEG signals, IEEE Transactions on Biomedical Engineering 48 (2001) 111–116.
[7]. G.J. Montgomery, K.C. Drake, Abductive networks, Proceedings of the SPIE Conference on the Applications of Artificial Neural Networks, (1990) 56-64.
[8]. Gujarati, D., Basic Econometrics [Chapter 15: Regression on dummy- dependent variables]. McGraw-Hill, 2003.
[9]. Irwin J., Kirchner, J., "Anemia in Children", Am. Fam. Physician, 64 (2001), 1379-86.
[10]. J. Cheung, Z.Y. Lin, R.W. McCallum, J.D.Z. Chen, Screening of delayed gastric emptying using electrogastrography and abductive networks, Gastroenterology Suppl. S 112 (1997) A711.
[11]. J. Echauz, G. Vachtsevanos, Neural network detection of antiepileptic drugs from a single EEG trace, Proceedings of the Electro/94 International Conference (1994) 346-351.
[12]. Johnston, J., DiNardo, J., Econometric Methods [Chapter 13: Discrete and limited dependent variable models], McGraw-Hill. 1997.
[13]. Levy A., Fraser D., Rosen S., Dagan R., Deckelbaum R., Coles C., Nag- gan L., "Anemia as a risk factor for infectious diseases in infants and toddlers: results from a prospective study", Eur J Epidemiol., 2005; 20(3): 277-84
[14]. Martin, J. F., "Fuzzy Control in Anesthesia," J. Clin. Monit., Vol. 10, 1994, pp. 77–80.
[15]. Maronna R., Martin D., Yohai V., Robust Statistics: Theory and Meth- ods, Wiley, 2006.
[16]. Mazumdar M., Glassman R., "Categorizing a Prognostic Variable: Re- view of Methods, Code for Easy Implementation and Applications to Decision-Making about Cancer Treatments", Statist. Med. 19 (2000), 113-132.
[17]. P. Gopinath, N.P. Reddy, Toward intelligent Web monitoring: performance of committee neural networks vs single neural network, Proceedings of the IEEE International Conference on Information Technology Applications in Biomedicine (2000) 179-182.
[18]. Peracchi F., Econometrics [Chapter 15: M-Estimators], Wiley, 2001. [11] Preisser J., Koch G., "Categorical Data Analysis in Public Health", Ann. Rev. Public Health. 18 (1997), 51–82.
[19]. R. E. Abdal Aal, "Improved Classification of Medical Data Using Abductive Network Committees Trained on Different Feature Subsets" Department of Computer Engineering, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.
[20]. R.E. Abdel-Aal, A.M. Mangoud, Modeling obesity using abductive networks, Comput. Biomed. Res. 30 (1997) 451-471.
[21]. R.E. Abdel-Aal, A.M. Mangoud, Abductive machine learning for modeling and predicting the educational score in school health surveys, Methods Inf. Med. 35 (1996) 265-271.
[22]. S.J. Farlow, The GMDH algorithm, in: S.J. Farlow (Ed.), Self-Organizing Methods in Modeling: GMDH Type Algorithms, Marcel-Dekker, New York, 1984, pp. 1-24.
[23]. Siberry G., Iannone R., eds., The Harriet Lane handbook. 15th ed. St. Louis: Mosby, 2000.
[24]. T. Kondo, A.S. Pandya, J.M. Zurada, GMDH-type neural networks and their application to the medical image recognition of the lungs, Proceedings of the 38th IEEE SICE Annual Conference, (1999) 1181-1186.
[25]. http://www.wardsystems.com/apptalk.asp

Table 1: Number of patients in different months, for three years and forecast of third year using linear regression.

Year/ Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Sum
2007	174	426	534	1137	1180	11	386	393	851	318	488	768	6666
2008	212	16	404	500	1358	1072	1564	720	688	224	520	792	8070
2009	317	655	602	303	1433	134	831	1624	2075	1386	1088	1100	11548
Mean	234.33	365.67	513.33	646.67	1323.67	405.67	927.00	912.33	1204.67	642.67	698.67	886.67	8761.33
Index	0.32	0.50	0.70	0.89	1.81	0.56	1.27	1.25	1.65	0.88	0.96	1.21	12.00
Expected 2010	102	328	423	268	2598	74	1055	2029	3424	1220	1041	1336	13899

Table 2: Year and month wise fuzzy coefficient

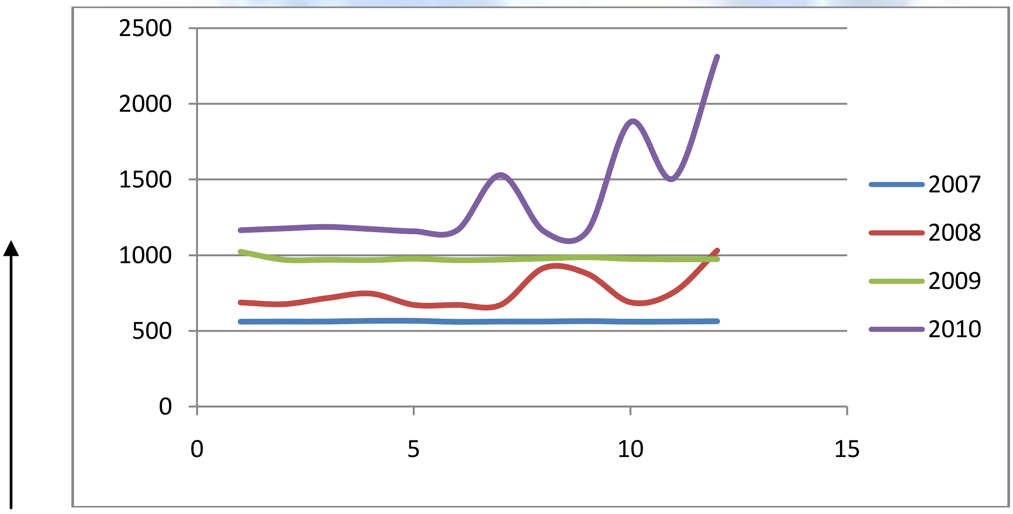
Year/ Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2007	6.03	7.29	7.91	12.46	12.87	5.33	7.08	7.11	10.04	6.72	7.64	9.44
2008	16.36	5.76	45.49	75.86	7323.05	1596.38	21938.02	244.86	206.49	17.44	84.39	359.30
2009	62.57	8.65	8.31	6.64	15.50	5.85	9.87	17.88	25.08	14.96	11.96	12.07
2010	7.98	19.85	29.12	15.59	187356.25	7.13	372.23	18894.01	5236373.40	723.98	351.80	1155.71

Table 3: Year and month wise fuzzy coefficients after clipping higher values

Year/ Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2007	6.03	7.29	7.91	12.46	12.87	5.33	7.08	7.11	10.04	6.72	7.64	9.44
2008	16.36	5.76	45.49	75.86				244.86	206.49	17.44	84.39	359.30
2009	62.57	8.65	8.31	6.64	15.50	5.85	9.87	17.88	25.08	14.96	11.96	12.07
2010	7.98	19.85	29.12	15.59		7.13	372.23			723.98	351.80	1155.71

Table 4: Number of patients after applying the fuzzy coefficients with the average value

Year/ Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2007	561	562	562	567	567	560	562	562	565	561	562	564
2008	688	677	717	747	672	672	672	916	878	689	756	1031
2009	1024	970	970	968	977	967	971	979	987	976	973	974
2010	1165	1177	1187	1173	1158	1165	1530	1158	1158	1881	1509	2313



No. of patients Months

Figure 1: Comparison of Graphs of four Years after Fuzzy Coefficient Application