International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

A Neural Network based Approach to Predict Machine Status for Big Data using R

Swapnil Khobragade[†*] and Ganesh. K. Pakle[†]

[†]Department of Information Technology, Shri. Guru Gobind Singhji Institute of Engineering & Technology, Nanded, (An Autonomous Institute of Govt. of Maharashtra) Maharashtra-431606, India

Accepted 01 July 2015, Available online 06 July 2015, **Vol.5, No.4 (Aug 2015)**

Abstract

Big data studies recognize some of the major issues growing up in current market, like engaging new customers is quite difficult and not affordable than having legacy of old customers. Machine status (e.g. ON/OFF) prediction model to effectively manage the situation and to maintain efficient work flow in order to control various machinery are developed by academics and practitioners in past few years. Identifying machine status of operating machines is an important activity for mechanical industries, the caliber to properly predict machine status is exigent. As the automation in the mechanical industries era gets more emulative since 20s, machine status prediction management is a crucial task for machine industries where number of large scale machines operates. The article tenders a neural network founded outlook for prediction of machine status in association with various log records. Neural network associated nearness to the results of experiments more than 95% accuracy is being calculated with a machine data that can predict conditions. Further, it is in observation that medial sized Neural Networks performing excellent operations over the model designed for machine status to predict when various neural network's strategies taken & examined.

Keywords: *Neural Network, Machine status prediction, Predictive analytic, Machine management.*

1. Introduction

Competitive market enterprises mainly come from existing loyal customers, who rely on consistent profits. And hence, customer relationship management (CRM) eternally focuses on faithful clients which are the most productive, reliable and big source of data for decision making. The data used in this paper reflects various aspects of machine operation's behavior. Those kinds where behavioral data is being used for evaluation of machine's feasible life & scope time value (Hung, C., & Tsai, C.-F., 2008), to determine the risk of suspension of certain machine which was in working state, and to anticipate their future needs on the basis of which we can find patterns of machine status (Berry at al 2003). Now days selective or personalized decision making practices are replaced by traditional mass decision strategies for customer relationship management (CRM) systems (Burez, J. *et al* 2008).

All selective decision making practices involve identifying a sub-set of present patterns which may be likely responsible for machine operation. Some or all machine's aspects like Temperature, Pressure has been taken in consideration for prediction has received increasing consideration in the machine industries research studies and survey over the past few years. Summarizing those literatures indicate that little changes found in the assumption rate may refine result(s) in significant influence on business (Van d en Poel 2004). In caution to effectively management and value machine performance and standing, it's vital to make a more practical and correct machine standing prediction model. Applied math and data processing techniques are taken in consideration to construct the predictive models. The data mining techniques may be accustomed discover attention-grabbing patterns or relationships within the collected information, and predict or categories the behavior by mounting a model supported accessible knowledge. In different words, its associate degree knowledge domain space with a generic goal of predicting outputs & using subtle algorithms for processing to get in the main hidden patterns, associations, anomalies, and/or structure from in depth knowledge hold on in knowledge warehouses or different data repositories (Han J. 2001).

There are number of data mining proficiency that are proposed to predict potential customers that are most likely used to evaluate operation and condition of simple machine . Among the popular technique to predict machine status with respect to various scenes (i.e. Temperature, Pressure) are: neural systems,

*Corresponding author: **Swapnil Khobragade**

musical backup vector machines and logistic relapse models (Hung S.Y. 2006). Data mining examination writing recommends that for non-parametric datasets, machine learnedness system, for example, neural systems, regularly outflank conventional measurable and fundamentally prohibitive methods, for example, direct and quadratic separate investigative deduction approaches (Baesens *et al* 2002).As we have a tendency to contemplate from business intelligence approach. Machine management procedures under the customer relationship management (CRM) framework include two major analytical modeling actions. First action is predicting attributes that are directly responsible for machine status (ON/OFF) and second action includes assessing the extremely effective way on which machine operator can react.

This analysis supports the follow up task, suggests that intends as an example a way to apply data processing techniques which help machine status management. This article focuses artificial neural network technique to search out a most efficient model from keep machine data to predict status and to identify the machine's turnover. In the same manner, the machine business operators will enhance the competitive edge.

This article proposes a neural network (NNs) based perspective for prediction of machine status in subscription varied dependent aspects of running machine viz. Temperature, Pressure etc., the remainder of the article is organized as following sections. Section 2 pair of reviews this literature, associated with machine information and completely different data processing techniques accustomed predicts machine information in numerous studies. Section three describes this analysis operandi, and Section four shows the experimental results on real dataset used. Lastly, conclusion and future work is provided in Section 5.

2. Literature Review

2.1 Machine Data Operation and Management

Recent era of market is changing into additional emulative, several corporations has already started realizing the importance of customer-focused business strategies for sustaining the competitive edge which helps maintaining a rigid profit level at bottom line and prime line. That's corporations in the main trust the stable financial gain that derives from loyal customers. Whereas, making new infrastructure and holding loyal customers are troublesome and dear. As new client account is setup completely different expenses like credit looking, promotional and advertisement expenses square measure occurred. This kind of expenses square measure many times bigger than value of work force that may change the corporations to inhibit a client. So, it is evolved into associate industry wide trusts that the most effective core selling strategies used for the long run is to inhibit existing clients (Kin H.S. 2004).

Machine operation analysis literature has noted that 'Machine Status' may be a term employed in the machine business to point the status. Database normalization to lessen knowledge redundancy aspects and tables of a relational database is the method of organizing. Less fashioned is redundant (smaller) without losing understanding in the desk but desk accessories: historic refers back to the most important key of new folks to the desk to define the international key. Moreover to a forte cause, deletions, and changes made in just one desk and to redefine the database using foreign keys will also be propagated through the leisure of the order is to isolate information.

Standardization of data: The data and information within the data set is specific to the material to expose the ways in which clearly defined example values is a way of changing. Most models work well with the normalized data set. The measured values from -1 to +1 can be extended up to a limit. This method includes both decimal and standard deviation normalization techniques. For the purpose of this work, the latter is used. This method (column subtracting column data means) is established and the same column columns divided by standard deviation of the data column is meant to focus. It is usually the variability in the data set (dispersion) is used to reduce. This data column is meant to set the zero and column variance comes in, and it's showing up in every data model provides a similar opportunity to sample.

$$\boldsymbol{MC_i = x_i - \mu}$$

Where,
$\boldsymbol{MC_i}$ - Column mean (centering), having column means of 0.

$$\boldsymbol{SC_i = \frac{MC_i}{\sigma}}$$

Where,
$\boldsymbol{SC_i}$ - Column scaling having variance of 1 & column means of 0

2.2 Related Work

Designing an efficient machine status prediction model victimization varied techniques has become a determinative issue for business and lecturers in recent 20s. Hence to grasp however completely diverse studies have made some prediction models, this article focus a number of recent studies shown in following Table 1.

Table 1 Some Literature Related to Machine Data

Author	Data set	Prediction method
Lina Industry	Machine Data	Logistic regression & Neural Network
Burez and Poel, 2006	Pay-TV company	Logistic regression and Markov chains random forests
Hung et al., 2006	Wireless telecom. company	Classification (decision tree, neural network) clustering(Kmeans)
Buckinx and Poel, 2005	Retailing dataset	Neural networks, logistic regression
Coussement and Poel, 2008	Machine Data (2013)	Hazard model survival analysis

Coussement and Van cave Poel (2008) connected bolster vector machines (SVM) in an exceedingly daily paper membership beat setting in order to develop a stir model. The customer stir forecast execution of the bolster vector machine model is seat checked to supplying relapse and arbitrary gauges (Coussement, K., & Van den Poel, D. 2008).

Recent literature analysis concerning client churn, most of the connected work focuses on mistreatment just one data processing technique like classification or clump to mine the client retention knowledge. Few studies (Hung et al., 2006) applied over one technology that was supported cluster analysis and classification [6]. From the review of the literature will conclude that neural network can predict client churn in numerous domains such as Pay-TV [3], banking (Han, J., & Kamber, M. 2001), retail (Buckinx 2005) and finance (Chiang 2003.

This article proposes a neural network (NN) primarily oriented to predict machine status in subscription various dependent aspects viz. Temperature, pressure etc.

3. The Proposed Neural Network (NN) based Approach

3.1 Artificial Neural Network

An ANN could be an advanced network that includes an outsized set of straightforward nodes called neural cells. ANN was based on advanced biology analysis regarding human brain tissue and neural scheme, and might be accustomed simulate neural activities of data process within the human brain (C. L. Blake and C. J. Merz). Artificial Neural Network possesses some topological structures data knowledge process nodes which distribute information in analogous format. The map & reduce of inputs and calculable output responses are obtained via mixtures of non-linear transfer functions. We are able to create use of self-adaptive data pattern recognition protocols to research the coaching algorithms of the artificial neural networks victimization past expertise, neural cells, memory and association, to method fuzzy, nonlinear, and noise-containing knowledge while not developing any mathematical models. The assorted algorithms for neural networks coaching are Delta, Hebb, Kohonen, and BP computation. The principally used BP computation algorithmic rule is that the error back propagation algorithmic rule planned by the PDP cluster of Rumelhart in 1985. Neural networks are often differentiated into single layer perception and multilayer perception (MLP) network. The multilayer perception includes multiple layers of straightforward, two states, sigmoid transfer performs have process component or neurons act by mistreatment weighted node connections.

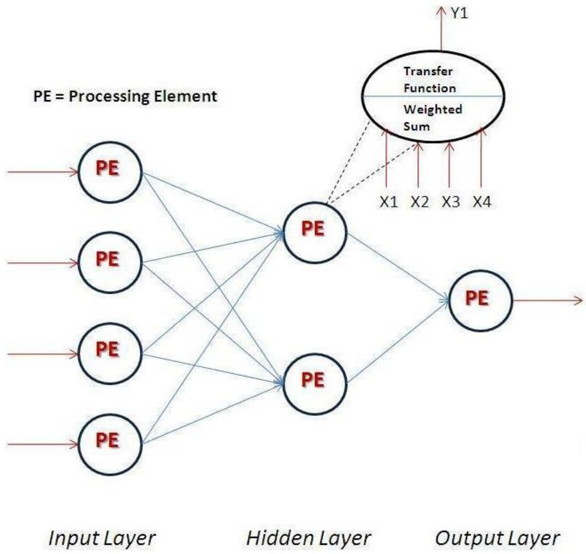

Figure 1: Neural Network with One Hidden Layer

Additionally, proposed neural network contains one or a lot of many subsequent hidden layers of neurons between the input and output layers. These types of mediator layers are called as hidden layers and nodes embedded in these layers are known as hidden nodes as these aren't accepting inputs directly from outsider nodes. The generic feed-forward multilayer perceptron neural network includes the input layer, hidden layer, and output layer (generic architecture is as shown in Figure 1). The neural network take in adoption the error back propagation training rules are known as BP networks, who's learning method includes each forward and backward propagation. Considering forward process, the sample signals can step by step lift through each layer with the sigmoid perform.

$F(x)=1/1+e^{-x}$

A neural network cell i.e. vegetative cell of every layer (hidden or input) solely responsible for the status of succeeding neural cell. Signals can't be obtained within output layer if the expected output, the load values of every layer of the cells should be changed to stick to minimum error. Inaccurate output signals are going in backward from the supply. Lastly, the signal error can arrive within sure area unit as with perennial propagation therefore the weights of the various layer neurons are changed.

Associated network is firm at $\boldsymbol{n}$ layers and y_j^n, such that y_j^n, represents the output with $\boldsymbol{n}$ layers and $\boldsymbol{j}$ nodes. When y_j^n is equal to $\boldsymbol{xj}$, and $\boldsymbol{j}$ represents the inputs. Let $\boldsymbol{Wij}$ *be* the connection weight between y_j^{n-1} and y_j^n, then here we get the threshold of θ_j^n of $\boldsymbol{n}$ layers and $\boldsymbol{j}$ nodes. Neural network learning algorithms consist of the following steps:

1. First of all, start initiating the node connection weights with random values and tune all other constraints.
2. Looking forward for desired output, read input signal vector. The signals progress at the networks with the following formula:

$$y_j^n = F(s_j^n) = F\left(\sum W_{ij}^n \; y_j^{n-1} + \theta_j^n\right) \tag{1}$$

With this formula network will start calculating nodes represented as $\boldsymbol{j}$ of every layer having the output y_j^n from the input layer to complete the calculation processing of nodes. $F(s)$ stands for one of the sigmoid transfer functions.
3. Tally the genuine output through the computations, working via layers.
4. Tally the error factor. The error matrix value assigned to node to the output layer will be obtained by various values which are in between the genuine output & the expected output (D_j^k) and it is written as:

$$\delta_j^n = y_j^n(1 - y_j^n)(D_j^k - y_j^n) \tag{2}$$

Error factor of every node for the last layer (hidden layer) rely on phenomenon of backward error propagation of every layer ***(n = n, n-1.....1)*** and is given by:

$$\delta_j^{n-1} = F(s_j^{n-1}) \sum W_{ij}^n \delta_j^n \tag{3}$$

5. Now start changing the node connection weights by starting backward from the output layer via hidden layers is calculated by using formula:

$$W_{ij}^n(p+1) = W_{ij}^n(p) + \eta\delta_j^n y_j^{n-1} + \alpha\left[W_{ij}^n(p) - W_{ij}^n(p-1)\right] \tag{4}$$

$$\theta_j^n(p+1) = \theta_j^n(p) + \eta\delta_j^n + \alpha\left[\theta_j^n(p) - \theta_j^n(p-1)\right] \tag{5}$$

Where,
$\boldsymbol{p}$ - Iterative times (epoch) of the layers.
$\boldsymbol{\eta}$ - Learning rate (constant)
$\boldsymbol{\alpha}$ - Momentum ($\mathbf{0}$ to $\mathbf{1}$)

3.2 Dataset

This article used the machine dataset from one of Machine Learning Database from Lina trade. The machine data-set deals with dependent aspects of running machine viz. Temperature, Pressure etc. knowledge in raw type (e.g., data warehouse) aren't continually the most effective to analyze, & particularly not for predictive data processing. Information should be preprocessed or ready and remodeled to induce the most effective mean type. Preparing Data/Knowledge is extremely vital as a result of completely different predictive data mining techniques behave otherwise reckoning on the preprocessing and transforming types.

Several numbers of techniques for knowledge/data preparation which will deliver the goods completely different data-mining goals.

The neural network is implemented on Lina Industry's machine data. In this paper the data set used contains around 42 columns worth of information about 8000 rows. Machine aspects like Temperature, Pressure, along with an indication of machine status (ON/OFF) is mentioned as final state of machine.

3.3 Training the data and testing it for Neural Network

The neural network used to perform predictive modeling, very first layer i.e. input layer consists of all possible input variables needed & included to predict the resultant variable i.e predicted values. Final layer i.e. output layer consist of artificial neural network's output field which is target of the prediction values. The input and output fields are symbolic or numeric.

The phase of data massaging and cleaning is major part of system where symbolic fields are reworked in to a numeric kind (binary encoding or dummy) just before processed by the network. Middle layer i.e. hidden layer consist of variety of neurons, which outputs from the previous layer mix. A network may have any range of hidden layers, though these layers are typically unbroken into minimum to change the predictive model which is finalized. All neurons in one layer of the network are joined to any or all possible neurons at intervals successive layer whereas the neural network gets expertise with the relationships between data and results, aforesaid to be training. The Figure 2 provides miscellaneous implementation of neural network trained machine data.

Machine status endues two totally different categories of organized neural networks, the Multi-Layer Perceptron (MLP) and also the Radial Basis Function Network (RBFN). The square measure of five totally variant algorithms offered inside the neural net node of trained input, however the article used the fast technique among all. The fast technique use a feed-forward back-propagation network whose topology relies on the quantity and kinds of the input and output fields. For avoiding overtraining issues which will exist among neural networks, willy-nilly elite proportion of the training data will employed to train the network. Because the information pass repeatedly via network, it is doable for the network to find out patterns which exist within the sample solely and therefore over train.

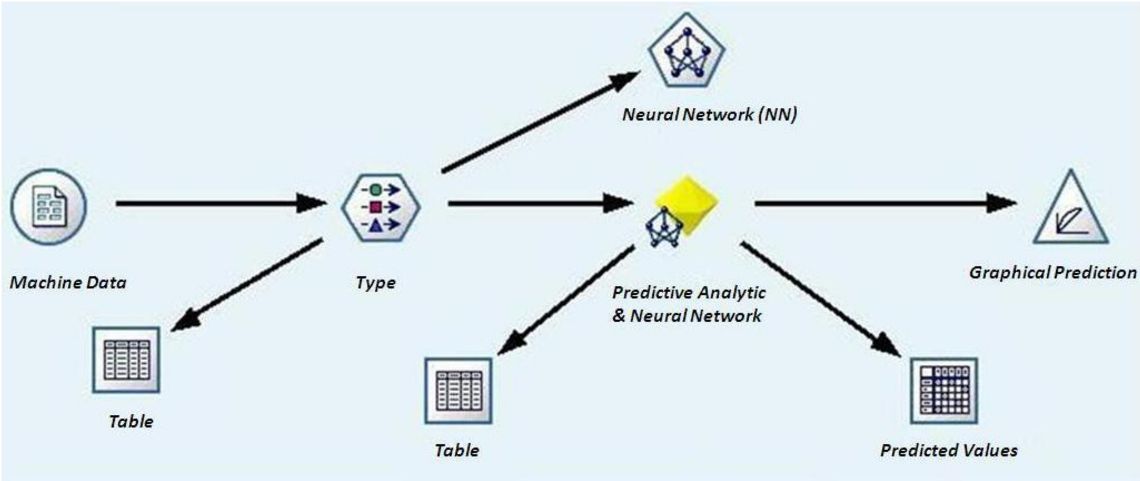

Figure 2: Neural Network implementation using machine data

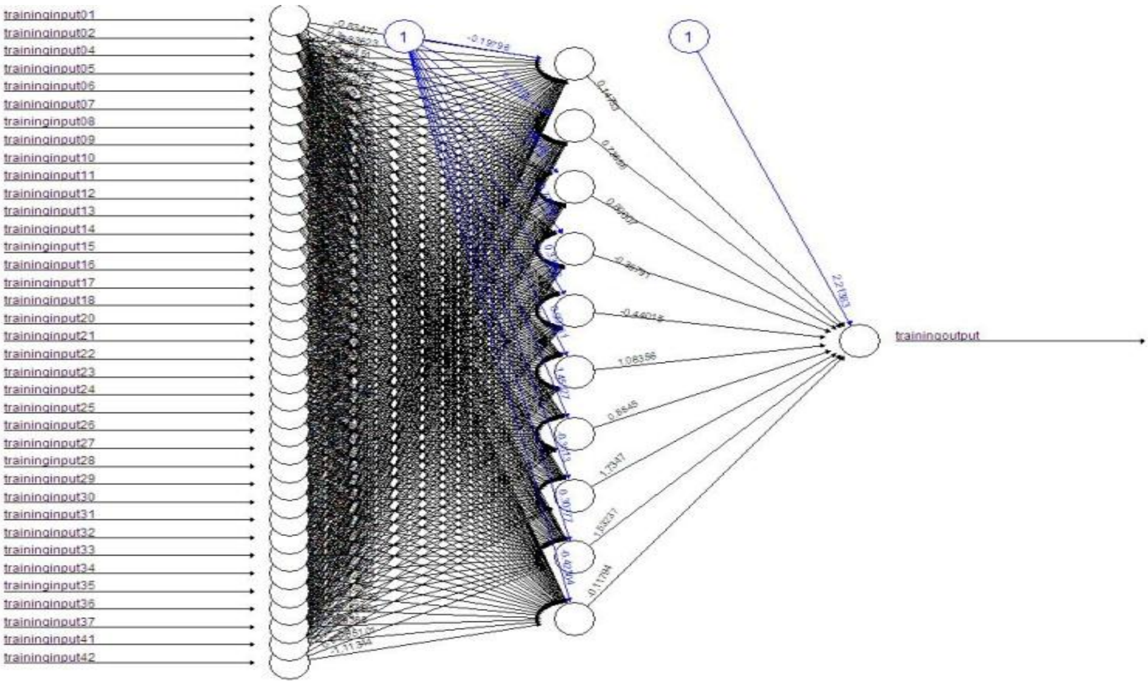

Figure 3: System Generated Model

Therefore the network might become typical to the training sample data and loose its ability to generalize the network. The willy-nilly elite proportion of the training data is employed to train the network & once this proportion of data is created a whole withstand the network; remainder is reserved as a take a look at (test data block) set to judge the performance of above neural specification.

4. Result & Discussion

Information regarding the neural network has been shown in analysis section where generated model is introduced. The ultimate model outline shown in Figure 3. Final predicted accuracy shown in confusion matrix for this neural network is approx. 99%, indicates the ratio of the test set & training set taken is properly predicted is shown in figure 4.

One neuron per numeric or flag is basic unit of input layer which is created in the system. Training set consist of inputs variables and every other aspect are verified wit final output status.

In this article while studying the data for designing neural network, number experiments have been done with multiple hidden layers containing two to ten neurons however the efficient and stable results were obtained having ten hidden layers & 3 neurons. One neurons reminiscent of the two values of the output field (ON/OFF) is present in output layer.

Figure 3 describes proportional importance of the input variables for performing sensitivity analysis of the system generated neural model. The input fields' area unit listed in descendent order of proportional importance. The values 0 and 1 are ultimate description for machine status, wherever, 1 indicates

OFF and 0 indicates ON. In application with this figure seldom goes 1 as machine stops operating in sure patterns of input. The generated Neural internet calculates 2 new fields, **$Machine_Status** (dataset) and **$Predicted_Machine_Status** for each record within the input data base. The primary represents the predicted status (ON/OFF) and also the next prior i.e. second confidence worth for the prediction. After most values ($) are simply applicable to symbolic values of outputs and can be 0 and 1, with the additional assured predictions where the values are nearer to 1. Once predicting the field contains symbols, it's good to opt for a data matrix of the new predicted values **($Predicted_Machine_Status)** and the real values **(Machine_Status)** also, so as to check however they compare and wherever the variations square measure. The matrix of actual (rows) and predicted (columns) values of the proposed approach is shown in Figure 5. The system designed is efficiently predicts over 95-99% accurately of the original data, properly however solely 66 of true machine failure patterns can be recognized within the data. If we tend to needed to properly predict the exact

ON/OFF status of machine where the patterns are identified as the opposite varieties this is able to seem to be an affordable model or system. On the opposite hand, if we tend to need to predict those machine statuses that were progressing to build the industrial loss by machine failure, proposed model would solely predict properly in concerning 75% of the instances.

Confusion Matrix (Predicted Value)			
Machine Data		0	1
	0	1635	0
	1	7	2
Total Rows		1644	
Accuracy		99.27%	

Figure 4: Data Matrix after prediction values

```
>
> confusionMatrix(x1,y1)
Confusion Matrix and Statistics

          Reference
Prediction    0    1
         0 1635    0
         1    7    2

               Accuracy : 0.9927229
                 95% CI : (0.9873228, 0.9962343)
    No Information Rate : 0.9987871
    P-Value [Acc > NIR] : 0.999999800

                  Kappa : 0.2484047
 Mcnemar's Test P-Value : 0.001496164

            Sensitivity : 0.9927140
            Specificity : 1.0000000
         Pos Pred Value : 1.0000000
         Neg Pred Value : 0.1428571
             Prevalence : 0.9987871
         Detection Rate : 0.9915100
   Detection Prevalence : 0.9915100
      Balanced Accuracy : 0.9963570

       'Positive' Class : 0
```

Figure 5: Confusion Matrix with Input Variables

Conclusion

Machine status prediction and management of the system is crucial task in mechanical industries in developing countries. In order to survive in competitive market, automation providers need to look forward to predict possible anomalies and take proactive actions to gain valuable results by means of which we can fine patterns to avoid any malfunctioning. Hence, to establish an efficient and incorruptible machine status prediction model, has become severe problem for researchers. In past few years, academicians and practitioners are working in the field of neural network based predictions.

This article proposes that data mining techniques can be promising solution for the machine automation and early-warning model for such non-steady-state systems can be established. Final section of this article where model summary is elaborated concludes that the proposed system gives more than 95% of overall accuracy for the prediction of the machine status with respect to various dependent aspects.

On account of future work, number of issues can be taken into consideration. Very first issue, as the data pre-processing stage in information mining is an extraordinarily primary step for the final prediction mannequin efficiency, the dimensionality discount or feature resolution step can also be involved moreover to knowledge discount. 2d, together with neural networks, other general prediction methods may also be applied in blend, similar to help vector machines, genetic algorithms, and many others to strengthen hybrid units. Eventually, the current operandi of computing device fame prediction can also be applied for different sectors like insurance, banking, or airline companies and comparisons may also be carried out for prediction accuracy.

References

Berry, M. J. A., & Linoff, G. S. (2003), *Data mining techniques: For marketing, sales, and customer support*. John Wiley & Sons.

Hung, C., & Tsai, C.-F. (2008). Segmentation based on hierarchical self-organizing map for markets of multimedia on demand. *Expert Systems with Applications,* 34(1), 780–787.

Burez, J., & Van den Poel, D. (2006). CRM at a Pay-TV Company: Using analytical models to reduce customer attrition by targeted marketing for subscription services. *Expert Systems with Applications,* 32(2), 277–288.

Han, J., & Kamber, M. (2001). *Data Mining: Concepts and Techniques.* Morgan Kaufmann

Van d en Poel, D., & Lariv ie're, B. (2004). Customer attrition analysis for financial services using proportional hazard models. *European Journal of Operational Research*, 157, 196–217.

Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecomm churn management. *Expert Systems with Applications*, 31(3), 515–524.

Baesens, B., Viaene, S., Van den Poel, D., Vanthienen, J., & Dedene, G. (2002). Bayesian neural network learning for repeat purchase modeling in direct marketing. *European Journal of Operational Research*, 138(1), 191–211.

Keaveney, S. M. (1995). Customer switching behavior in service industries: An exploratory study. *Journal of Marketing*, 59, 71–82.

Kim, H. S., & Yoon, C. H. (2004). Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. Telecommunications Policy, 28, 751–765.

Berson, A., Smith, S., & Thearling, K. (2000). Building data mining applications for CRM. New York, NY: *McGraw-Hill*.

Kentrias, S. (2001). Customer relationship management: *The SAS perspective*, www.cm2day.com.

Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameterselection techniques. *Expert Systems with Applications,* 34, 313–327.

Buckinx, W., & Van den Poel, D. (2005). Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting. *European Journal of Operational Research,* 164(1), 252–268.

Chiang, D. A., Wang, Y. F., Lee, S. L., & Lin, C. J. (2003). Goal-oriented sequential pattern for network banking churn analysis. *Expert Systems with Applications*, 25, 293–302.

Markopoulos, A. P., Manolakos, D. E., & Vaxevanidis, N. M. (2008). Artificial neural network models for the prediction of surface roughness in electrical discharge machining. *Journal of Intelligent Manufacturing,* 19, 283– 292.

C. L. Blake and C. J. Merz, Churn Data Set, UCI Repository of Machine Learning Databases, http://www.ics.uci.edu/~mlearn/MLRepository.html. University of California, *Department of Information and Computer Science, Irvine, CA,* 1998.

Clementine data mining workbench, *http:// www.spss.com/clementine*