International Journal of Data Mining Techniques and Applications
Volume: 04 Issue: 02 December 2015, Page No.92-96
ISSN: 2278-2419

Investigating Performance and Quality in Electronic Industry via Data Mining Techniques

Davoud GholamianGonabadi[1], Seyed Mohamad Hosseinioun[2], Jamal Shahrabi[3], Mohammad AliMoradi[4]
Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran
Email: Alimoradi@gmail.com, Dgholamian@gmail.com,M.moh.hoss@gmail.com,Jamalshahrabi@aut.ac.ir

Abstract-The organizations have always sought ways and methodologies to boost the quality level of their product. Obviously, economic aspect ought to be taken into account, eschewing irrational costs in the system. Data mining can be applied as one of the best technics at hand in case of analysis and prediction of performance and quality. In this work, we have used a new methodology based on data mining to predict and improve the quality level of the electronic parts. The results suggest better accuracy compared to the previous studies.

Keywords-Prediction of Quality; Data Mining; Electronic Industry

I. INTRODUCTION

High quality production has always been an issue, largely focused on by companies and organizations in their incessant attempt to gain advantage over their competitors, consequently, leading to a continuous quest for the right methodologies to boost the quality level of their products. On the other hand, the balancing the demanded quality level and the expenses of the desired change is an inevitable part of the decision making, as naturally, any change ought to be economically profitable at last.

From another viewpoint, defected products cause additional costs due to rework, amendments, reconstructions and wasted time, imposing unnecessary costs to the system. Obviously, as the probability of such failures decrease, and their causes are annihilated, we can expect the fall of respective costs. Therefore, studying the past data of defected parts and causes failures are invaluable part of production planning, chiefly by eliminating these causes and avoid the failures of the past in the future. Detecting the patterns and extracting the knowledge hidden in the data is the way to go, and data mining is an indispensable tool for such purpose that has been used and trusted for long [1]. Current methods of analyzing production parts data, mainly working based on annual statistical data or superficial checking have serious deficiencies regarding the prediction of quality, efficiency and accessibility of parts. Nowadays, consequently, analyzing experts try to exploit stronger methods for the abovementioned. Recently, data mining has been used widely to process and extract patterns in the data. Tools such as data warehousing, data mining, etc. have opened up new gates before industry and production to win noticeable advantage in the market [2]. Specifically, using data mining- the extraction of hidden information within large data bases- organization are enabled to predict the upcoming events and make informed decisions [3]. Data mining has largely helped productivity by increasing the opportunity of predicting and detecting error and inefficiency [4]. In this research we try to predict the quality and performance of electronic parts using data mining tools. It is our believe that, by following the offered methodology in the paper, we can boost the quality level of the electronic products. The literature review is presented in part 2 of the paper, the third part, offers the methodology employed, the fourth part covers a presentation of the data, and the application of the offered methodology on the case. And finally, in the fifth part, the results of the study are uncovered, analyzed and studied.

II. LITERATURE REVIEW

Data mining discovers the hidden patterns in the data, and is actually a part of a wider process of knowledge extraction [2]. Quality prediction is on the other hand one of the most essential issues in quality management and reliability. In this section, a number of papers corresponding to data mining are application in quality engineering and reliability, germane to the research at hand, is reviewed.

In the domain of detecting defected parts by data mining tools few papers are available [2]. Khediri et al. for instance, have applied support vector regression control charts for multivariable non-linear auto correlated processes-used to detect error- and process control charts. The results suggest that the employed control charts are capable of effectively monitoring the behavior of the processes, with considerable guarantee of limiting the prospective false errors [5]. Zhang et al. have applied data mining tools to control welding quality and management systems. In this research, the quality of welding is calculated using time series. The results show that gathering and analyzing data by data mining technics online would have considerable positive afflictions on the quality of future welding [6]. In the domain of quality management, Abbasi and Akhavan Niaki used neural network to predict the nonconforming and defected cases in high efficiency processes. The results indicate ease of application and high accuracy of the method [7]. Liu et al. in the control process based on principle component analysis in drying corn used neural network. As the drying process is often very complex due to the multi-variability and non-linearity, neural network was preferred to other methods for such prospect. The results suggest high accuracy and consistency [8].

III. METHODOLOGY

In this study, in order to analyze and predict the performance and quality of the electronic parts, the following methodology is applied.

Volume: 04 Issue: 02 December 2015, Page No.92-96
ISSN: 2278-2419

A. Ordinary Least Square

Ordinary least squares (OLS) regression is arguably the most widely used method for fitting linear statistical models. An OLS regression model takes the familiar form:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ip} + \varepsilon_i \quad (1)$$

where Y_i is case i's value on the outcome variable, β_0 is the regression constant, X_{ij} is case i's score on the jth of p predictor variables in the model, β_j is predictor j's partial regression weight, and ε_i is the error for case i. Using matrix notation, Equation 1 can be represented as:

$$\boldsymbol{y} = \boldsymbol{X}\beta + \varepsilon \quad (2)$$

where $\boldsymbol{y}$ is an $n \times 1$ vector of outcome observations, X is a $n \times (p+1)$ matrix of predictor variable values (including a column of ones for the regression constant), and ε is an $n \times 1$ vector of errors, where n is the sample size and p is the number of predictor variables. The p partial regression coefficients in β provide information about each predictor variable's unique or partial relationship with the outcome variable. Researchers are often interested in testing the null hypothesis that a specific element in β is zero or constructing a confidence interval for that element using a sample-derived estimate combined with an estimate of the sampling variance of the estimate [9].

The validity of the hypothesis tests and confidence intervals as implemented in most statistical computing packages depends on the extent to which the model's assumptions are met. The assumptions of the OLS regression model include that (1) the Y_is are generated according to the model specified in Equation 1, (2) the $\boldsymbol{X}$ values are fixed (rather than random), (3) the errors are uncorrelated random variables with (4) zero means, and (5) constant variance, the latter assumption known as homoscedasticity. Assumptions (1) and (4) require book-length treatment but in effect state simply that the model is correctly specified (see, e.g., [10], [11]). When the predictors are random variables rather than fixed, assumption (2) can be relaxed without causing major problems by interpreting the partial regression coefficients as "conditional" on values of $\boldsymbol{X}$, provided that the conditional expectation of ε given the values of $\boldsymbol{X}$ is zero. Unfortunately, this conditional expectation may well be nonzero, such as when predictors contain measurement error. Measurement error in predictors can bias OLS regression estimates, a topic that is beyond the scope of the present article. A nice introduction can be found in [10].

B. Principal Component Analysis (PCA)

Principle Component Analysis is the most popular method of dimension reduction. The aim is to make an orthogonal transformation through which, using the linear combination of the principal components, we reach fewer number of new features. It is worth mentioning that chiefly, the method follows the same procedure for the numeric and categorical variables, but in some case, it has a whole conceptual difference. In this section, merely, the application of principle component analysis is explained on the data. To convert the initial features to the principle components, the following steps are to be executed [12].

1. Firstly, we standardize the data, in a way that all values are in one specific range. Plus, the average of each feature a_j will equal zero using the conversion below.

$$\tilde{x}_{i,j} = x_{i,j} - \frac{1}{m}\sum_{i=1}^{m} x_{i,j} \quad (3)$$

2. The matrix corresponding to the covariance of the data we have from the second step is calculated as below:

$$V = X'X \quad (4)$$

3. The eigenvalues and eigenvectors corresponding to the covariance matrix is calculated:

$$\det(V - \lambda I) = 0 \quad (5)$$

The eigenvalues are derived from the determinant shown in (5) in which it is the identity matrix. Solving the abovementioned determinant, we will have n eigenvalues at hand.
And to get to the eigenvectors, the following calculations ought to be done.

$$Vw_j = w_j \times \lambda_j; j = 1,2,\ldots,n \quad (6)$$

In which the w_j is the j^{th} eigenvector of covariance matrix.
4. Selecting the principle components.

$$I_q = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_q}{\lambda_1 + \lambda_2 + \cdots + \lambda_n} \quad (7)$$

In the above relation we have: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.
I_q indicates the explicable changes of the q principle components.
To select the principle components, a threshold I_{min} is set.
Now, the number of principle components is the smallest q, for which the following relation applies.

$$I_q > I_{min} \quad (8)$$

5. Set the new features.

$$P_j = Xw_j; j = 1,2,\ldots,q \quad (9)$$

C. Multilayer Perceptron Neural Network

Multilayered Perceptron Neural Network is one of a kind of the leading neural networks that possess a more complex structure compared with the Perceptron as it encompasses all input, hidden and output nodes. The input nodes are designed to receive the descriptive features. Usually, the input nodes are equal to the descriptive variables. The hidden nodes are connected either to the input nodes or other hidden units by the input arches from one side and by the output arches to either other hidden nodes or the output layer from the other. In the output layer, the values received from the hidden layer are converted to the output values corresponding to the desired prediction in the form of output variables. It is usual to design only one output node in case of classification problems. The activation function is usually linear function, Sigmoid Function, or Hyperbolic Tangent. In this study we have applied the Levenberg-Marquardt method to determine the weights that we will elaborate in the following. The method has the following process for learning [12]:

1. Set a random variable for the weights and their threshold.
2. Make an input output pattern for the data in the form of (x^k, t^k), in which x^k is input value and t^k is target value.
3. The data received as the input are then transferred to the hidden layer using the weights and then activation functions

International Journal of Data Mining Techniques and Applications
Volume: 04 Issue: 02 December 2015, Page No.92-96
ISSN: 2278-2419

are applied.

4. The output of the network is calculated by the following formula (I is the input vector size)

$$o^k = f\left(\sum_{i=1}^{l} w_i x_i^k\right) \quad (10)$$

If the network output and the actual value are not consistent, using the Levenberg-Marquardt method, the weights are updated.

5. if the steps are taken and the difference of the weights in two consecutive iteration is not the equal to zero, go to the third step.

D. Levenberg-Marquardt

The application of Error Back Propagation 4 method has had considerable positive effects in the learning process of neural network. Despite the endeavors to speed up the learning process using Back Propagation, relatively little improvement has been achieved. Levenberg-Marquardt method is the expanded form of Back Prop algorithm. This method enjoys the speed of Newton algorithm and consistency of Gradient Descent Method.

In the Error Back Propagation algorithm, F(w) acts as the performance measure of difference between the output of the model and the real data and it is our aim to minimize this deficit.

$$F(w) = e^T e \quad (11)$$

In which $w = [w_1, w_2, \dots, w_N]$ corresponds to all the weights in the network. Also e is a vector representing the error for the training data. This way, the differences between weights in each step is calculated:

$$\Delta w = [J^T J + \mu I]^{-1} J^T e \quad (12)$$

In which J is the Jacobean Matrix, μ is the learning rate and is updated in accordance with β (depending on the output).

1. Assigning random weights and to μ (usually, it is considered as 0.1.)
2. Calculating the sum of squared error of the training data as F (w).
3. Calculating Δw and updating w.
4. If the F (w) in the previous iteration is smaller than that of the current iteration, then:

$$w = w + \Delta w \quad (13)$$

$$\mu = \mu \times \beta \quad (\beta = 0.1)$$

Go back to the 2nd step, otherwise:

$$\mu = \frac{\mu}{\beta} \quad (14)$$

And return to the 4th step.

The proposed approach in this paper is shown in Figure 1.

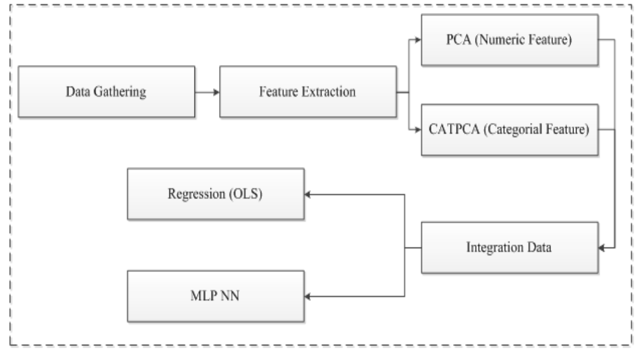

Figure 1: Proposed Approach

IV. Data and Result

The data is gathered from the website of California University Machine Learning. In this study, using the input variables, relative performance of the CPUs manufactured by numerous brands is analyzed. The summary of data is shown Table 1.

Table 1: Summary of Data

Variable Name	Variabl Type	Max	Min	Mean	Standard Deviation	PRP Correlation
vendor name	Nominal	----	----	----	----	----
Model Name (many unique symbols)	Nominal	----	----	----	----	----
machine cycle time in nanoseconds (MYCT)	Integer	1500	17	203.8	260.3	-0.3071
minimum main memory in kilobytes (MMIN)	Integer	32000	64	2868	3878.7	0.7949
maximum main memory in kilobytes (MMAX)	Integer	64000	64	11796.1	11726.6	0.863
cache memory in kilobytes (CACH)	Integer	256	0	25.2	40.6	0.6626
minimum	Integer	52	0	4.7	6.8	0.6089

Volume: 04 Issue: 02 December 2015, Page No.92-96
ISSN: 2278-2419

channels in units (CHMIN)						
maximum channels in units (CHMAX)	Integer	176	0	18.2	26	0.6052
published relative performance (PRP)	Integer	1150	6	105.6	160.8	1

E. The Results of Principle Component Analysis

As the variables take the form of either Nominal or Discrete, and no missing value is observed, we first perform feature extraction methods on the data. The PCA method is applied in case of discrete variables and CATPCA for the Nominal. The minimum amount of explicable change is selected to be 0.95. As the result, the Nominal variables were converted to two components and the Discrete variables to 5 components.

F. The OLS Regression Results

Using the OLS regression, the results are summarized in the Figure 2, Figure 3 and Table 2

TABLE 2: OLS REGRESSION RESULTS

Variables	Coefficient	t-Student	Std.Error
Constant	69.14**	17.55	3.93
Component 1	54.69*	1.87	35.50
Component 2	-53.91*	-1.82	35.44
Component 3	-86.02**	-6.45	13.32
Component 4	107.90**	3.38	31.82
Component 5	-202.22**	-7.70	26.25
Component 6	202.69**	3.81	53.08
Component 7	135.14**	4.65	29.01

Note: * and ** indicate 10 and 1 significant respectively.

In the 99 percent level, the components 1 and 2 are not sensible, also, the components 3 to 7 are sensible and the effect on variable y is as follows:

The components 3 and 5, have a negative effect on the variable y, causing it to decrease in case of their augment. Other components, 4, 6 and 7 are in positive relation. Among the components, 5 have the strongest effect (negative) and 6 the strongest positive effect on y.

Furthermore, considering the adjusted R-square, 83.6 percent of the changes in y are explicable by the input components mentioned, showing a considerable level of fitness. Mean Square Error and Mean Absolute Error in case of this data set are 2878.703 and 34.402 respectively. It is worth mentioning that the results achieved here are weaker compared to [13].

G. MLP Neural Network Results

With the purpose of predicting the Relative Performance of the CPUs, we have used the Multi-layer Perceptron Neural Network as well. After the determination of parameters, the results for MSE and MADare 820.83 and 22.24 respectively, as showed below is Figure 3. It is considerable that the results of this study are noticeably better than [13].

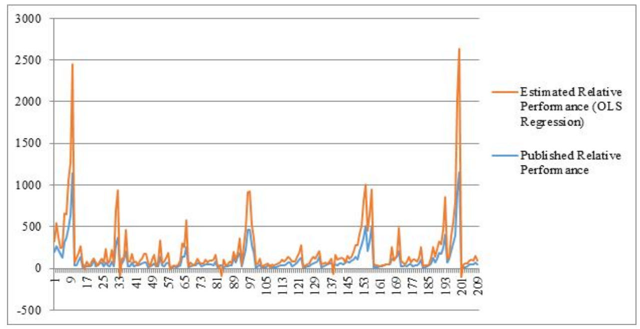

Figure 2: Results from OLS Regression and Published Regression

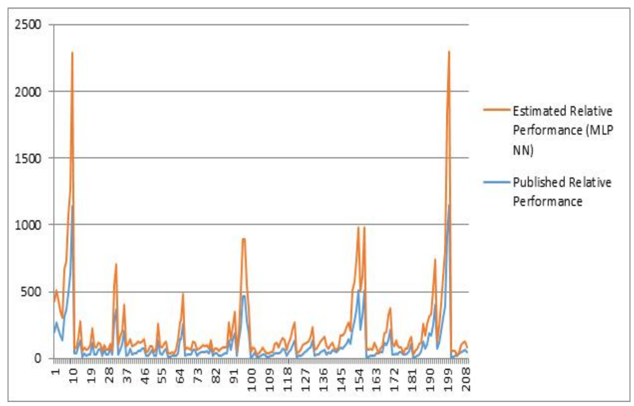

Figure 3: Results from MLP NN and Published Regression

V. CONCLUSION

In this research, with the purpose of analyzing and predicting the quality and performance of electronic parts, a new methodology has been exploited. To study the performance of this new methodology, data corresponding to the performance of CPUs are borrowed from the California University data base. The results suggesting a considerable improvement by applying the Multilayer Perceptron Neural Network compared to the OLS regression methods.

REFERENCES

[1] Ostadi, Bakhtiar; Rezai, Kamran; Aghdasi, Mohammad, 1386, The application of data mining for improving the product quality level, The proceeding of the first data mining conference in Iran, Tehran, Amirkabir University of Technology, The institute of Gita Data Processing. (In Persian)

[2] Amouie, Golriz; Bagheri Dehnavi, Malihe; Minaie Bigdeli, Behrouz, 1390, Proposing a model for detecting the product defects using data mining technics, The first conference of computer and IT students, Tabriz University, Tabriz. (In Persian)

International Journal of Data Mining Techniques and Applications
Volume: 04 Issue: 02 December 2015, Page No.92-96
ISSN: 2278-2419

[3] Chris Rygielski, Jyun-Cheng Wang , David C. Yen.” Data mining techniques for customer Relationship management” Technology in Society 24 (2002) 413–502.

[4] Kai-Ying Chen , Long-Sheng Chen, Mu-Chen Chen , Chia-Lung Lee.” Using SVM based method for equipment fault detection in a thermal power plant” Computers in Industry 62 (2011) 42–50.

[5] Khediri, I. B., Weihs, C., & Limam, M. (2010). Support vector regression control charts for multivariate nonlinear autocorrelated processes. Chemometrics and Intelligent Laboratory Systems, 103(1), 76-81.

[6] Zhang, C. H., Di, L., & An, Z. (2003, November). Welding quality monitoring and management system based on data mining technology. In Machine Learning and Cybernetics, 2003 International Conference on (Vol. 1, pp. 13-17). IEEE.

[7] Abbasi, B., & Akhavan Niaki, S. T. (2007). Monitoring high-yields processes with defects count in nonconforming items by artificial neural network. Applied mathematics and computation, 188(1), 262-270.

[8] Liu, X., Chen, X., Wu, W., & Zhang, Y. (2006). Process control based on principal component analysis for maize drying. Food control, 17(11), 894-899.

[9] A. F. Hayes and L. Cai, "Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation," Behavior Research Methods, vol. 39, pp. 709-722, 2007.

[10]W. D. Berry, Understanding regression assumptions vol. 92: Sage, 1993.

[11]H. White, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," Econometrica: Journal of the Econometric Society, pp. 817-838, 1980.

[12]V. Carlo, "Business intelligence :data mining and optimization for decision making," Editorial John Wiley and Sons, 2009.

[13]P. Ein-Dor and J. Feldmesser, "Attributes of the performance of central processing units: A relative performance prediction model," Communications of the ACM, vol .30 ,pp. 308-317, 19.