ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 9(44), DOI: 10.17485/ijst/2016/v9i44/98964, November 2016* ISSN (Online) : 0974-5645

Adaptive Data Mining Approach for PCB Defect Detection and Classification

P. K. Srimani and Vaddin Prathiba*

Department of Computer Science and Mathematics, Bangalore University, Bangalore - 560056, Karnataka, India; profsrimanipk@gmail.com, prathibavk@gmail.com

Abstract

Objective: To develop a model for PCB defect detection and classification with the help of soft computing technique. **Methodology:** To improve the performance of the prediction and classification we propose a hybrid approach for feature reduction and classification. The proposed approach is divided into three main stages: (i) data pre-processing (ii) feature selection and reduction and (iii) Classification. In this approach, pre-processing, feature selection and reduction is carried out by measuring of confidence with the adaptive genetic algorithm. Prediction and classification is carried out by using neural network classifier. A genetic algorithm is used for data preprocessing to achieve the feature reduction and confidence measurement. **Findings:** The system is implemented using MatLab 2013b. The resulting analysis shows that the proposed approach is capable of detecting and classifying defects in PCB board.

Keywords: Classification, Data Mining, Feature Selection, PCB Defect

1. Introduction

Printed Circuit Boards (PCBs) are very thin a type of circuit which are made of fiberglass. These fiber glasses are connected to electronic components with copper wires and circuits. PCBs are used in various electronic equipment and products such as computers, cellular phones, televisions etc. In manufacturing industries, the product manufacturing efficiency is a crucial parameter. Due to the increasing demand of electronic devices, there is a need to speedup the manufacturing process and yield efficient output. By keeping this in mind, industries are manufacturing PCBs, electronic components, and equipment, but due to the complexity of the recent devices, it is a bottleneck for engineers to cater for it. According to the recent scenario, multi-layered circuits are required which can function faster than other conventional circuits which affect the quality of manufacturing. There are various schemes present which are used for quality measurement based on the quality test criteria but the main issue with these kinds of schemes is that it doesn't include all the parameters which affect the manufacturing quality and results in lower performance.

During manufacturing, industries store datasheets which include cost, sales, quality and parameters used for manufacturing. These datasheets can be used to improve the quality of manufacturing. To overcome the issue of manufacturing quality, data mining has been proven to be a promising technique for ensuring quality improvement by detecting the defects[1]. It is a rapid growing technique which can analyze huge amount of data with less complexity and computation time.

While conventional human examination is constrained to handle vast volume of information, the purpose of data mining is to develop new techniques and devices to dissect information and make valuable data and learning from databases where measurement, intricacy and measure of information are too huge to be processed by human perception alone. Originated from insights and machine learning, information mining is the procedure of extracting and refining information from vast databases[2]. Using data mining approach, people and associations can distinguish hidden patterns inside the information and connections of patterns in information that might be utilized to make substantial prediction.

In[3] defined data mining approach for the extraction of

* *Author for correspondence*

useful information which is not known previously. Data Mining is used for various applications such as fraud detection, market analysis etc. In[4] Introduced a hybrid approach which incorporates spatial parameters of image, Eigen decomposition and support vector based clustering which provides the estimated number of defected clusters. In[5] authors presented a multistage Extreme Learning Machine (ELM) to improve the accuracy of clustering. During the process of their ELM, the input data were divided into several stages, then, every stage was analyzed independently. Compared with individual ELM, their ELM has better function for considering the characteristics of input data. In[6] authors developed a framework for DM that consists of a Kruskal–Wallis test, k-means clustering, the variance reduction splitting criterion to investigate the huge amount of semiconductor manufacturing data and infer possible causes of faults and manufacturing process variations. In[7] introduced DM technology to fault diagnosis field and proposed a new method based on C4.5 decision tree and principal component analysis after analyzing the flaws of conventional fault diagnosis methods. Similarly, in[8,9] Hsu and Chien proposed a hybrid DM approach that integrates spatial statistics and adaptive resonance theory neural networks to quickly extract patterns from WBM and associate with manufacturing defects. a novel approach to solve the quality assurance problem in predicting the acceptance of machined components. Research presented in[10] discussed several prediction approaches for the compressive strength of cements by using soft computing approaches. A comparative study is presented in[11] and compared several data mining techniques for analyzing and prediction of liquefaction. In[12], authors reviewed applications of data mining in manufacturing engineering, in particular production processes, operations, fault detection, maintenance, decision support and production quality improvement.

PCBs are widely used in military application where these PCBs have more than 6 layer. In order to perform vibration analysis, Young's Modulus is a key component to perform PCB structure analysis[13,14]. In PCB defect detection process, data mining is also considered as a promising technique. In data mining based approaches, feature selection is considered as a challenging task. In order to address this issue, a new approach is presented in[15,16]. This technique utilizes cross validation and SVM-RFE for stable feature selection process for cancer classification. For feature class determination, Cumulative Ranking Score (CRS) parameter is introduced here. Feature selection process is known for removing irrelevant and redundant feature as pre-processing technique resulting in computational complexity reduction and accuracy improvement. In order to improve the performance of feature selection process, two algorithms, Modified Mutual Information based Feature Selection (MMIFS) and Dynamic Mutual Information based Feature Selection (DMIFS) are developed and compared[17].

2. Proposed Model

This section provides the proposed solution for defect detection, prediction and classification using data mining approaches. This section involves three main categories which are used in data mining. First, data preprocessing, feature selection feature and reduction and classification. In order to perform the PCB defect detection and classification we use preprocessing steps by using missing value imputation and propose a modified approach for selection of feature and reduction using genetic algorithm and finally a neural network classifier is applied to estimate the prediction accuracy. This method is implemented by using the MATLAB tool. Training operation is performed with the help of the classifier's training approach to learn the pattern of the given dataset. In order to make the appropriate selection of the data, preprocessing is required. During this preprocessing stage unwanted attributes or features and repetitive data are removed which results in the reduction of the attributes to be processed and missing values are imputed in the dataset. Proposed scheme utilizes missing value imputation and adaptive genetic algorithm is used for feature or attributes reduction. The system flowchart is shown in Figure 1

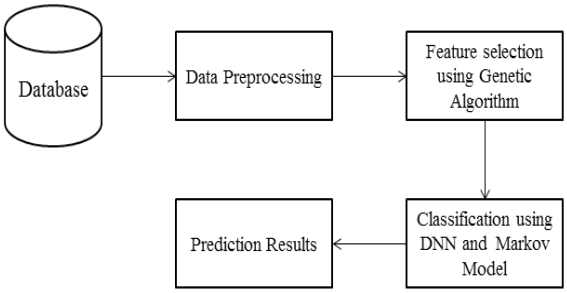

Figure 1. Overall architecture of the model.

2.1 Data Sets

This section provides a brief description of the dataset considered for the performance evaluation and pre-processing of given dataset. For this work we have considered 14 attributes as feature along with serial numbers. Database description is depicted in Table 1. During this process of data preprocessing, redundant or unwanted data is removed, attribute reduction and missing value imputation are processed to carry out the preprocessing of the data.

Table 1. Attribute list

Defect Number	Name of Defect
1	Hole missing
2	Over etching
3	Conductor missing
4	Under etching
5	Breakout
6	Open-circuit
7	Unnecessary Short
8	Shorting
9	Spur
10	Mouse-bite
11	Wrong hole size
12	Under Etching
13	Conductor closeness
14	Pin Hole

2.2 Data Preprocessing

In this section we discuss about the pre-processing of input dataset. Main aim of this process is to remove the unwanted data by reducing the low rank attributes and missing value imputation is processed to carry out the preprocessing of the data. Various approaches have been proposed to improve the data mining efficiency with the help of the data preprocessing such as cleaning of data, integration of data, transformation of data and data reduction etc.

This scheme involves two main methods: (i) imputation of missing values and (ii) feature reduction using feature selection. These steps are described in following sections.

2.3 Missing Value

Here in this section, missing value imputation is performed in a given dataset. Generally, this is the process which deals with missing data by finding the missing values in a given set and maintains resemblance with the original data.

In our proposed model, a new approach is utilized which is categorized into two sections: (i) missing value estimation and (ii) filling the values

In order to estimate missing values, a zeros matrix is initiated which is considered as a time series representation and denoted as $M = \{m_1, m_2, m_t, \ldots, m_n\}$. With the help of this time series matrix, data is modeled in a matrix of coefficients which is represented as

$$m_j = Y_j x_j + \alpha_j \tag{1}$$

After achieving the modeled data, a linear prediction approach is applied to generalized data and can be expressed as

$$\begin{bmatrix} m[c+1] \\ m[c+2] \\ \vdots \\ m[n] \\ m[1] \\ m[2] \\ \vdots \\ m[n-c] \end{bmatrix} = \Xi \times \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} + \alpha_j \tag{2}$$

α is the normal noise distribution with the zero variance σ^2

$$\Xi = \begin{bmatrix} m[c] & m[c-1] & \cdots & m[1] \\ m[c+1] & m[c] & \cdots & m[2] \\ \vdots & \vdots & \ldots & m[n-c] \\ m[2] & m[3] & \cdots & m[c+1] \\ m[n-c+1] & m[n-c+2] & \cdots & m[n] \end{bmatrix}$$

During this process, it is assumed that the coefficients are known for the input data and missing data indices are not bound to be in consecutive form.

For estimation, log-likelihood data is computed which is not depending on missing data as:

$$\mathcal{L}(z) = \sum_{t=c+1}^{n} \left(m_t - \sum_{j=1}^{c} x_j Y_{t-j} \right)^2 = l^T l \tag{3}$$

where l^T is the transpose matrix which is represented as

$$l = \mathcal{A} z \tag{4}$$

and z denotes the column vector of the input data , $\mathcal{A}$ is denoted as

$$\mathcal{A} = \begin{bmatrix} -x_c & \cdots & -x_1 & 1 & 0 & \cdots & 0 \\ 0 & -x_p & \cdots & -x_1 & 1 & 0 & \vdots \\ 0 & \cdots & 0 & -a_p & \cdots & -x_1 & 1 \end{bmatrix} \tag{5}$$

By using the above expression, missing data can be represented as

$$l = \mathbb{B}m + \mathbf{C}y \quad (6)$$

$\mathbb{B} = [\mathbb{B}_1 \mathbb{B}_2 \ldots]$ and $\mathbf{C} = [\mathbf{C}_1 \mathbf{C}_2 \ldots]$ are sub matrices of the column transformed matrix which is having information about the location of missing input data.

Finally the data estimation is achieved by using least –square method such as

$$x = -\mathbb{B} * \mathbf{C}y \quad (7)$$

By using these locations of the dataset, missing values can be estimated. Next stages is to impute these values by computing the similarity score.

This similarity score is computed between two vectors to compute the measurement of data distribution in a nonparametric way. Initially dataset is standardized using score transformation based on the ranks of the each element of the vector i.e. n values of the given vector are compared to achieve ranks $\{\mathcal{R}_1, \mathcal{R}_2\}$. Then these vectors are replaced by using cumulative normal distribution. In the next stage, data is sorted and finally absolute difference is computed to achieve the raw distance between the adjacent values

$$d(v_1, v_2) = \sum_{i=2}^{n} |v_2 - v_{1_{i-1}}| \quad (8)$$

The obtained distance is equivalent to the Manhattan distance summation for the given feature vector list. This distance is normalized between zero and one.

Minimum and maximum distance given by using (9) and (10)

$$d_{min} = \mathbf{C}^{-1}\left(\frac{n}{(n+1)}\right) - \mathbf{C}^{-1}\left(\frac{1}{(n+1)}\right) \quad (9)$$

$$d_{max} = 2\sum_{i} \left|\mathbf{C}^{-1}\left(\frac{i}{n+1}\right)\right| + \mathbf{C}^{-1}\left(\frac{(n-1)}{2(n+1)}\right) \quad (10)$$

By using these minimum and maximum distance values the data is normalized to compute the similarity between two adjacent feature vector values

$$S_{norm}(v_1, v_2) = 1 - \frac{d(v_1, v_2) - d_{min}}{d_{max} - d_{\min}} \quad (11)$$

2.3.1 Imputing Missing Values

In the given datasets, non-linearity is involved which affects the response of the input data. In order to impute missing data we use kernel based scheme. Optimization techniques for missing value imputation have been discussed in recent years[7]. Computing steps of this approach are mentioned below:

- Find the minimum difference between non-missing values and input vectors

$$d = g_x - g_v$$

- Compute the weighted mean based on the Gaussian density

$$w_{ij} = e^{-\frac{(g_x - g_v)^2}{2\sigma_j^2}}$$

- Impute the missing values by using weighted mean

$$d_{i1} = \sum_{j=2}^{c} \frac{d_{ij} w_{ij}}{\sum_{j=2}^{c} w_{ij}}$$

- Measure the confidence by finding the variance of given vector

$$\beta_i^2 = \frac{\sum_{j=2}^{p} w_{ij} \left(d_{ij} - \hat{d}i_1\right)^2}{\sum_{j=2}^{c} w_{ij}}$$

- Next stage is to merge the imputing values by using weighted average by considering similarity with higher weight and prediction with low weighted variance.

2.4 Proposed Feature Selection Algorithm

In order to achieve the efficient feature, we propose a new approach for feature selection based on the data relevancy information. Let us consider, the feature vector v with the feature matrix $f = (f_1, f_2, \ldots, f_N)$ where dimension of the data is represented as $\mathcal{N}$ with the class $\mathcal{C}$. Variation of the class is measured with the help the entropy which is given as

$$\mathcal{V} = entropy(\mathcal{C}) \quad (12)$$

For the given feature vector and class, the variation or uncertainty is denoted as $entropy(\mathcal{C}|f)$ and the relevance information is represented as $I_R(\mathcal{C}|f)$. Relation among these parameters is given as

$$I_R(\mathcal{C}, f) = I_R(f, \mathcal{C}) = entropy(\mathcal{C}) - entropy(\mathcal{C}|f) \quad (13)$$

$$\sum_{c \in \mathcal{C}} \int_f \frac{p_r(c,f) \log(p_r(c,f))}{\mathcal{P}_r(c) p(f)} df \quad (14)$$

Class probability is given by $\mathcal{P}_r(c)$, current feature is presented as f and $p_r(c, f)$ is the combined probability of the class and feature vector. To get the improvised

classification accuracy we perform minimization of the variation in the class vector and feature vector. According to the proposed approach classification achieves the highest accuracy with the smallest feature size.

2.4.1 Problem Formulation

Let us consider a given dataset $\mathcal{D}$ which contains f feature vectors for classification by using feature selection approach. Initial stage is to find the subsets of the dataset in a given dimension to minimize the entropy value which helps to maximize the relevance information of the dataset.

It is given as

$$D \cup f \rightarrow S \subset f \in S = \{D_1, D_2, \ldots, D_n\} \tag{15}$$

To overcome this problem we propose relevance information based feature selection to eliminate the variation in the class and feature probability. This can be written as

$$I_r\left(\mathcal{C}, {}^{\mathcal{D}_i}/_{\mathcal{D}_s}\right) = entropy\left({}^{\mathcal{D}_i}/_{\mathcal{D}_s}\right) - entropy\left({}^{\mathcal{D}_i}/_{\mathcal{C}}, \mathcal{D}_s\right) \tag{16}$$

Relevancy of the data is extracted using the chain rule which is given as:

$$I_r(\mathcal{C}, \mathcal{D}_i, \mathcal{D}_s) = entropy(\mathcal{D}_i, \mathcal{D}_s) + entropy(\mathcal{C}, \mathcal{D}_i, \mathcal{D}_s) \tag{17}$$

To maximize the relevance information of the feature, greedy approach is adapted here. By using this approach the relevance information can be written as

$$I_r\left(\mathcal{C}, \frac{\mathcal{D}_i}{\mathcal{D}_s}\right) = I_r(\mathcal{C}, \mathcal{D}_i) - \left[I_r(\mathcal{D}_i, \mathcal{D}_s) - I_r\left(\mathcal{D}_s, \frac{\mathcal{D}_i}{\mathcal{C}}\right)\right] \tag{18}$$

Ratio of the selected feature and the nearest feature gives the coefficient of relevancy given by

$$C_r = \frac{I_r(\mathcal{D}_i, \mathcal{D}_s)}{entropy(\mathcal{D}_s)} = \left[1 - \frac{entropy(\mathcal{D}_i, \mathcal{D}_s)}{entropy(\mathcal{D}_s)}\right] \tag{19}$$

It can be realized as mentioned below:

1. Initiate the parameters

Subset selection $s = "initial\ empty\ set"$

Set $D = Initial\ Feature\ vector\ values$

2. Pre-Computation of the given dataset

Find features to maximize the relevance i.e. $f \in f \rightarrow I_r(\mathcal{C}, D_i)$

3. Initiate feature selection and adapt greedy approach

Perform repetitions until desired features are selected which maximizes the relevance information

- Entropy measurement of the selected feature
- Relevancy information measurement between the features
- Next stage for feature selection if desired feature not achieved

Select feature $f \in f$ as mentioned below

$$f = \arg\left[\begin{matrix} maximize \\ f \end{matrix}\right]\left\{I_r(\mathcal{C}, D_i) - \begin{matrix} max \\ D_s \in S \end{matrix}\left[\mathcal{C}_r I_r(\mathcal{C}, \mathcal{D}_s)\right]\right\} \tag{20}$$

2.4.2 Genetic Algorithm Design

Genetic algorithm is a heuristic approach which is inspired by the natural environment and evolutions. In Nature, new living beings adjust to their surroundings through development. The executions of genetic algorithms can altogether vary in the method for developing another population. A few algorithms make a different populace of new people in each era by applying hereditary administrators (Figure2). Conventional genetic algorithm is available for optimization but due to the complexity of dataset, a new approach is required for finding optimal solution. This can be achieved by properly adapting methods of initialization of genetic algorithm and genetic operators.

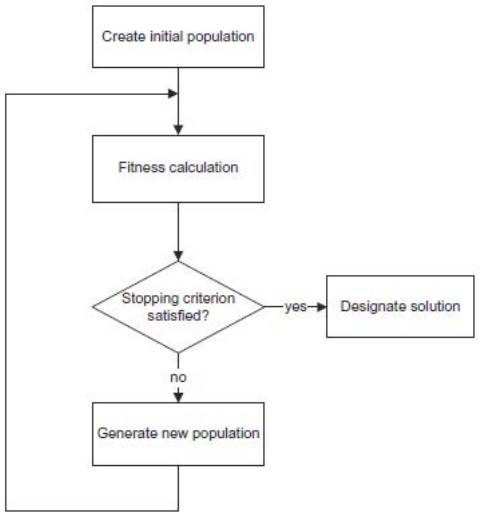

Figure 2. Genetic algorithm flowchart.

2.4.3 Proposed Model Design for Genetic Algorithm

A GA as a developmental enhancement method begins

with an underlying population moving towards a global ideal solutions and stops when the stop conditions are fulfilled. The customary methodology starts with an irregular beginning population and after that advances starting with one population then onto the next as the people experience crossover and mutation. Our methodology depends on the accompanying fundamental speculation: when from the earlier data about the conceivably appealing ranges is accessible, then the underlying population of the GA can be created in a manner that the alluring areas of the attainable location must be secured with an arrangement of focuses and the dimensionality of the issue can be diminished to those features that frame appealing zones. The consequences of prior experience and the aftereffects of the filtering systems are thought to be from the earlier data about the appealing zones.

To improvise the performance of genetic algorithm, we provide an extension by considering the existing population. This extension is depicted in Figure 3, where if stopping criteria is not satisfied then ne population is generated and added into existing population. Fitness calculation is applied for this and fittest population is selected for new population generation.

Further, a hybridized approach is implemented which is categorized into two phases which are: (1) initial solution generation, and (2) the reduced feature subset generation. Initial solutions are generated using filtering technique i.e. information gain, correlation, gain ratio and Gini index.

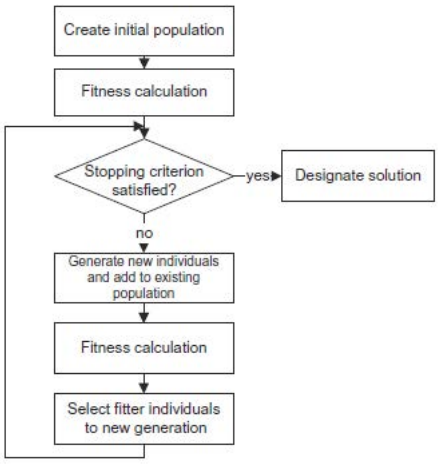

Figure 3. Algorithmic description of the genetic algorithm with the extension of the current population.

Algorithm

Input : Feature set
Output: Reduced feature set
Step 1: Filter techniques (Information gain, Gain ratio, Gini index and Correlation)
Step 2 : Initiate population generation
Step 3: computation matrix initialized
Step 4: Reduce feature based of step 2 matrix
Step 5: Initial population for reduced set
Step 6: compute fitness
Step 7: if stopping criteria activity
  a. Optimal solution
  b. Fix the optimal solution
Else
Repeat from step 5.
End

Figure 4 shows search space restriction stage by applying two stage procedures which involved (i) initial solution generation and (ii) generation of reduced feature set. In order to generate initial solution, we incorporate filtering scheme which provides feature ranking procedure by considering gain ratio, GINI index, correlation and earlier feature set evaluation performance. If best solution is not achieved then previously known solution is assigned as optimal solution. Novelty of this approach is presented by computing filtering based feature ranking and two stage optimal solution selection which differs from conventional genetic algorithm.

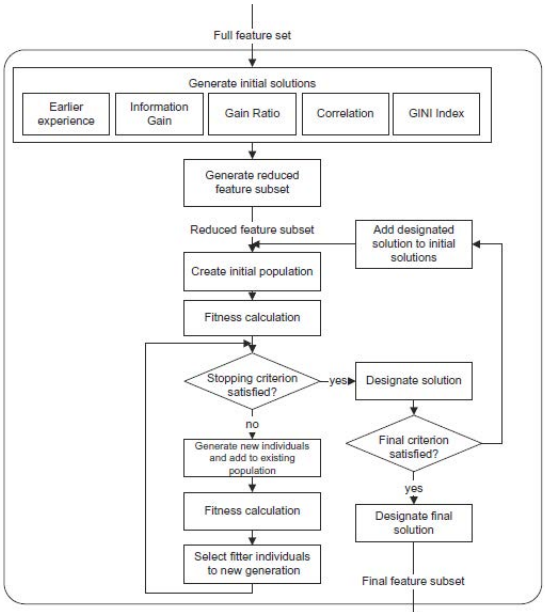

Figure 4. Flowchart of the hybrid genetic algorithm.

2.4.4 Deep Learner Neural Network

DNN is used for classification. It is feed-forward neural network which contains more hidden layer. Hidden layers are used to map the input features. A conventional mapping function is used in this work.

$$O = \frac{1}{1 + e^{-(b+fw)}}$$

Input features denoted by f , weights denoted by w , b denotes biasing and output is denoted as O

Complex relation between input and output are modeled with the help of this mapping function. This network can be trained by using back-propagation derivatives which gives the similarity between input and output for each training set. DNN pre-training can be performed by using discriminative method and supervised pre-training approach[13].

3. Results and Discussion

In this section we discuss about the experiments and results achieved under various scenarios. Experimental study for the PCB defect detection is mentioned below:

In this study we consider 5 case studies to measure the performance of the proposed approach by the following experiments: (i) Deep Learner classification (ii) hybrid classification using feature selection [11], feature reduction Deep learner.

In order to compute measure the performance of the system, we use various statistical parameters which include : (i) true positive rate ,(ii) false positive rate , (iii) precision ,(iv) False Measure , (v) ROC area , (vi) kappa measurement , (vii) mean absolute error , (viii) root mean square error , (ix) Relative Absolute Error and (x) Root Relative Squared Error.

Performance analysis of the proposed approach is mentioned in the given section by considering the various classification studies as mentioned earlier, Tables2 and 3 shows the performance based on the proposed approach by considering the various classification approaches.

True positive rate can be computed as

$$TPR = \frac{TP}{TP + FN} \tag{21}$$

where TP denotes the true positive values, FN is the representation of false negative values.

False positive rate computation is carried out using below given equation

$$FPR = \frac{FP}{FP + TN} \tag{22}$$

Precision is computed using

$$precision = \frac{TP}{TP + FP} \tag{23}$$

False score is defined as

$$fscore = \frac{2TP}{2TP + FP + FN} \tag{24}$$

Kappa measurement is given as

$$Kappa = \frac{observed\ Class - Expected\ Class}{1 - Expected\ Class} \tag{25}$$

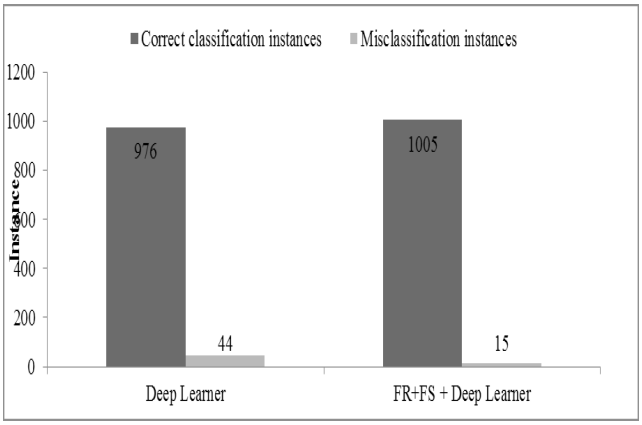

Figure 5. Correct classification and misclassification performance.

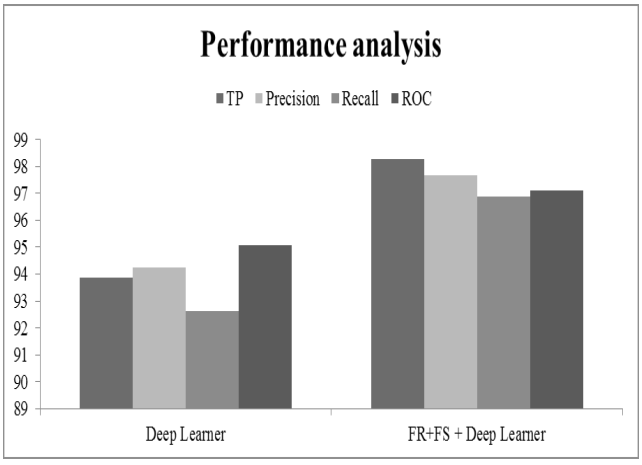

Figure 6. Performance analysis of proposed mode.

Table 2. Conventional Neural Network Performance

True Positive Rate	False Positive Rate	Precision	Recall	F-Measure	ROC Area	Class	Correctly Classified	976
95.67	0.62	95.27	91.86	92.64	95.43	TRUE	Incorrect Classified	44
93.87	0.691	94.24	92.63	91.07	95.09	FALSE	Kappa Measurement	86.24
							Mean Absolute Error	19.83
							Root Mean Square Error	27.49
							Relative Absolute Error	18.02
							Root Relative Squared Error	29.49
							Total Number of Instance	1020

Table 3. Feature selection and Feature reduction with neural network

True Positive Rate	False Positive Rate	Precision	Recall	F-Measure	ROC Area	Class	Correctly Classified	1005
98.56	0.0029	98.31	97.68	96.87	97.81	TRUE	Incorrect Classified	15
98.26	0.0038	97.67	96.89	96.79	97.09	FALSE	Kappa Measurement	97.96
							Mean Absolute Error	10.09
							Root Mean Square Error	15.81
							Relative Absolute Error	14.93
							Root Relative Squared Error	16.22
							Total Number of Instance	1020

In Figure 5 and 6, we show the performance analysis of Deep learner and hybrid approach in terms of classification and misclassified instances whereas Figure 6, shows the performance in terms of true positive rate, precision, recall and ROC.

4. Conclusion

In this work, a hybrid data mining approach is presented for PCB defect detection and classification. Proposed approach involves three main stages which are (i) dataset pre-processing, (ii) reduced feature selection and (iii) classification. Initially, missing values are estimated and imputed in the dataset by using confidence measurement of similarity score. Later, an optimization problem is formulated for feature reduction and selection of the efficient features. In order to solve this optimization problem for PCB defect detection, an improved genetic algorithm is applied for feature selection. Selected features or attributes are filtered using gain ratio, GINI index correlation and earlier feature set evaluation performance and the final set of selected features are obtained. Finally, the deep neural network is applied for classification. An experimental study is carried out under considering various scenarios i.e. feature selection and reduction classification. The result of classification accuracy using proposed model shows the efficiency and robustness of the system. This work is first of its kind in the available literature.

5. Acknowledgement

One of the author Vaddin Prathiba is grateful for the Rayalaseema University for providing the facilities to carry out the research work.

6. References

1. De La Calle FJ, Bulnes FG, Garcia DF, Usamentiaga R, Molleda JA. Parallel genetic algorithm for configuring defect detection methods. In IEEE Latin America Transactions. 2015 May; 13(5):1462–8.
2. Cheung DW, Lee SD, Xiao Y. Effect of data skewness and workload balance in parallel data mining. In IEEE Trans-

actions on Knowledge and Data Engineering. 2002 May; 14(3):498–514.
3. Frawley W, Piatetsky-Shapiro G, Maktheus CW. Knowledge discovery in databases: an overview. AI Magazine. 1992; 13(3):213–38.
4. Wang C. Separation of composite defect patterns on wafer bin map using support vector clustering. Expert System Application. 2009 Mar; 36(2):2554–61.
5. Hu X-F, Zhao Z, Wang S, Wang F-L, He D-K, Wu S-K. Multi-stage extreme learning machine for fault diagnosis on hydraulic tube tester. 2008 Aug; 17(4):399–403.
6. Chien C, Wang W, Cheng J. Data mining for yield enhancement in semiconductor manufacturing and an empirical study. Expert Syst Appl. 2007 Jul; 33(1):192–8.
7. Sun W, Chen J, Li J. Decision tree and PCA-based fault diagnosis of rotating machinery. Mechanical Systems and Signal Processing. 2007 Apr; 21(3):1300–17.
8. Hsu SC, Chien CF. Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing. International Journal of Production Economics. 2007 May; 107(1):88–103.
9. Tseng TLB, Kwon Y, Ertekin YM. Feature-based rule induction in machining operation using rough set theory for quality assurance. Robotics and Computer-Integrated Manufacturing. 2005 Dec; 21(6):559–67.
10. Baykasoglu A, Dereli T, Tanis S. Prediction of cement strength using soft computing techniques. Cement and Concrete Research. 2004 Nov; 34(11):2083–90.
11. Baykasoglu A, Cevik A, Ozbakir L, Kulluk S. Generating prediction rules for liquefaction through data mining. Expert Systems with Applications. 2009 Dec; 36(10):12491–9.
12. Harding JA, Shahbaz MM, Srinivas, Kusiak AA. Data Mining in Manufacturing: A Review. ASME J Manuf Sci Eng. 2005; 128(4):969–76.
13. Turian J, Ratinov L, Bengio Y. Word representations: a simple and general method for semi-supervised learning. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics. 2010 Jul; 384–94.
14. Venkat, Siva Rao PV, Singh AK, Srikrishna, Ranganath. Vibration Analysis of Printed Circuit Board Plate with Varying Boundary Conditions. Indian Journal of Science and Technology. 2016 Aug; 9(34):1–3.
15. George GVS, Raj VC. Accurate and Stable Feature Selection Powered by Iterative Backward Selection and Cumulative Ranking Score of Features. Indian Journal of Science and Technology. 2015 Jun; 8(11):1–6.
16. Devi KS, Ravi R. A New Feature Selection Algorithm for Efficient Spam Filtering using Adaboost and Hashing Techniques. Indian Journal of Science and Technology. 2015 Jul; 8(13):1–8.
17. Priyadharshini G, Sarojini B. Comparative Study on the Performance of MMIFS and DMIFS Feature Selection Algorithms on Medical Datasets. Indian Journal of Science and Technology. 2015 Dec; 8(35):1–5.