A Novel Approach for Improving Quality and Process Optimization in the Garment Industry

G.G.K.K. Wijerathne
Faculty of Computing
Sri Lanka Institute of Information Technology,
Malabe, Sri Lanka
kanchana2kw@gmail.com

U.M.M.P.K. Nawarathne
Faculty of Computing
Sri Lanka Institute of Information Technology,
Malabe, Sri Lanka
mnawarathne20@gmail.com

W.A.A.K. Weerasekara
Faculty of Computing
Sri Lanka Institute of Information Technology,
Malabe, Sri Lanka
amaliweerasekara97@gmail.com

L.H.C.J. De Silva
Faculty of Computing
Sri Lanka Institute of Information Technology,
Malabe, Sri Lanka
chanodhi20@gmail.com

Vijani Piyawardana
Faculty of Computing
Sri Lanka Institute of Information Technology,
Malabe, Sri Lanka
vijani.p@sliit.lk

Jesuthasan Alosius
Faculty of Computing
Sri Lanka Institute of Information Technology,
Malabe, Sri Lanka
jesuthasanalosius@gmail.com

Abstract - The garment industry is one of the intensive manufacturing industries operating in Sri Lanka, catering to the world's number one brands from the US and Europe. Its contribution to the economy is immense. Given the situation explained above, we need to improve the technical know-how and the product quality for meeting the deadline with regards to on-time delivery. To be more competitive in the market to improve productivity and the quality of the product along with optimizing the process workflow with using improved techniques is a must. Due to the poor workmanship and other human errors and poor human supervision, the quality of products tends to decrease. Therefore, increasing efficiency and decreasing quality cost are the main objectives. This paper presents suitable solution using machine learning techniques such as a method to identify defect co-occurrences using association rule mining, a method to optimize employee leave management using classification techniques, a method to optimize seating arrangement in sewing lines using Genetic Algorithms, and a method to implement predictive maintenance of machinery using classification techniques in data mining.

Keywords – Genetic Algorithms, Machine Learning, Predictive Maintenance, Data Science, Association Rule Mining, Classification

I. INTRODUCTION

It is a must to identify defects in the garment manufacturing process from the very beginning to provide the customers with worthy garment products at a minimal cost. Under these circumstances, the existence of an operation that can detect places where these defects occur, the defects occurring frequency, the reasons for them to occur, and the provision of solutions to such problems can aid the employees in producing quality garments. Also, machinery maintenance will play a huge role in this process.

Due to certain factors that can occur during the production process, the product quality can vary which is an unpreventable fact. For such garment defect occurances, human factors such as skill level difference, years of experience, and human prone errors can be the reasons. Therefore, manual processing of garments is a tough task as it can induce many errors during the process. Normally defects are recognized by human inspection where defect correlation is not considered. As such, these important data should be gathered as raw-data during and after the garment manufacturing process. However since the amount of collected data is huge, it is hard for the data to be analyzed which can be used to present future insights without the usage of a tool. By collecting relevant data regarding machinery and other components the risk can be minimized in advance. Also, the success of cost, quality, and services will depend on having the right people in the right position with the right skills [1]. Human Resources (HR) systems have not been improved with the current Information technology (IT) solutions. Though the HR systems had been in place, it appears that there is no provision provided in the area of the automated decision-making process using current trends in Information Technology.

This research is going to provide decision–making solutions for employee leave management, employee performance predictions, and additional salary management using Data Mining techniques and there will be common HR functions to manage the employees in the system. For the leave management predictions, the Decision Tree algorithm will be used, and the Decision Tree is a powerful algorithm that generates a tree and a set of rules. For the performance predictions, the Naïve Bayes classifier will be used, and it is an efficient and effective inductive algorithm for Machine Learning (ML) and Data Mining.

Since human skills and efficiency are different from one another, the sewing lines face the assembly line balancing problem. The Assembly line balancing problem is the inability to divide work properly within members

according to their skill levels. Within a reasonable time, the best technique is to create a plan by using a heuristic approach. To address this issue, in this research paper Genetic Algorithm (GA) has been used.

To predict a machinery failure, Predictive Maintenance (PdM) has been used to propose the solution with the use of Machine Learning due to its vast advantages [2]. Hence when a machinery component failure is predicted, this will provide the users with insights that the relevant machine is at risk due to this fact.

This paper consists of the following sections: Section 2 is a literature review related to this study. In Section 3, the methodologies used are described. In Section 4, the discussion of the system is presented.

II. Literature Review

A. *Defect Analysis*

Several previous kinds of research were carried regarding defect analysis and pattern identification in defects. The paper published by the Department of IPE, BUET, Dhaka, Bangladesh is discussed about Total Quality Management (TQM) is one of the important tools to improve quality and reduce manufacturing cost by reducing rework and scrape [3]. This paper defined the quality issues in the beginning and carried out an analysis for identifying the main causes by using a cause and effect diagram showed. A cause-and-effect diagram is a chart that identifies potential causes for quality problems that are often called the fishbone diagram [4].

Their attention in Quality Management (QM) is divided into different sub-areas such as the creation of QM environment, the introduction of workers with TQM, use of statistical process control (SPC) tools, generation of starting point, information sharing in decision making, encouraging cooperation and teamwork, customer focus as an element of design, modification of reward systems, selection of right raw materials and benchmarking [5]. But there are not any reporting services included in the solution and it is not guaranteed that the integrated detail of quality throughout the garment manufacturing process is analyzed. One of the major competitors in the Sri Lankan garment industry is the Hydramani apparel manufacturer.

Hydramani is using a very successful software solution called RES.Q[6]. As far as in the Sri Lankan garment industry this system holds a special place in analyzing real-time statistics in the garment industry.

In C.K.H. Lee's Research paper he proposed, by studying the ''IF-THEN'' relationship of defects represented in the rules, the Quality Management Team has been able to predict potential defects effectively when certain defects have occurred during production.[7]. Author Yusi Cheng proposes a GA based approach that incorporates the concept hierarchy of construction defects to discover multi-level patterns of defects from the Chinese construction industry from 2000 to 2010. [5]. Quality analysis is a business practice aiming at improving the quality of products and/or services [8].

B. *HR Process Optimization*

Himanshu Kumar Singh, Kshitij Vishnavat, R. Srinivasan [9] have used Bayesian Theory for big sales companies that consider criteria such as revenue, pay rate as the main criteria. By using the Bayesian Classifier, the score of the employee has been calculated. Then the employee leaves will be granted for a specific employee depending on the performance score.

Hamidah Jantan, Abdul Razak Hamdan, and Zulaiha Ali Othman [10] have used the C4.5 algorithm for employee talent prediction. Recommendation for promotion (yes/no) is considered as the target class in the classification process. Using the C4.5 classifier talent performance knowledge is generated from the yearly performance evaluation database.

V. Kalaivani and M. Elamparithi [11] have used Decision Tree Classification for the employee performance prediction. John M. Kirimi and Christopher A. Moturi [12] have used ID3, C4.5, and the Naïve Bayes Classifier to identify the best data mining algorithm for employee performance prediction.

C. *Machinery Failure Predictions*

The area of providing solutions for machinery failures is a widely researched area. An approach called multiple classifier ML methodology for PdM has been used to carry out the work in proving that PdM is better than other methods by [13]. The inclusion of more samples to the dataset has improved the accuracy of the result as well as reduced the skewness of the dataset giving a balanced dataset. To experiment with this, the ML classifiers; Support Vector Machines (SVM) and k-Nearest Neighbour (KNN) have been used. Finally, the outputs of these two algorithms have been compared and concluded that SVM gives the best result.

The research paper proposed by [14] revolves around sewing machines used in garment industries. To rectify the identified problems Preventive maintenance (PvM) has been proposed as an effective method. PvM is used to inspect adjustments and replacements in maintaining machines. Furthermore, according to their literature survey, it is noticeable that many researchers have used GA and linear programming in PvM research while the current authors [14] have used a method called dynamic programming. Although it is proved that PvM gives effective solutions for machinery breakdowns according to this paper, the problem of expending unnecessary measures can lead to additional costs as reviewed in the paper presented by [14]. Therefore, we believe that PdM can be

the solution over the PvM technique as it can give prior warnings with future predictions.

D. Optimized seating arrangement (Capacity plan for sewing line)

In view of the fact that the sewing process has a large impact on production, researchers have identified that the best place to focus is on the sewing lines to increase the efficiency of the garment industry.

Due to skill differences, assigning employees to the task leads to a bottleneck problem. By identifying their skills and assigning them to the tasks by meeting the required production rate with minimum ideal time is known as assembly line balancing. This can be solved using various techniques and this is also used in different industries.

A simulation technique has been used in the respective paper [15]. The goodness of fit test has been used on the real-time data gathered, while it has used model statistics to eliminate the bottleneck problem and to perform line balancing. The paper published by Komgrit and Sukrit discussed a nursing scheduling algorithm [16] by using a mathematical model, GA, and simulation method due to the uncertainty of the number of patients who came to the hospital each day as well as to maximize the efficiency of the nursing schedule.

Khalid Alzoubi's research paper has used the Largest Candidate Rule to address the motion of the workers [17]. According to this research paper, not only workers' skill levels but also the time taken for workers' motion on the work floor also have a considerable impact on the efficiency of the production. [18,19] research papers used fuzzy logic to handle this issue.

III. METHODOLOGY

The data which was needed for all the functions were collected during several factory visits with the guidance of industry professionals and the data was preprocessed accordingly for the relevant modules using techniques such as removing null values, checking outliers, etc.

A. Association Rule Mining Module

Using this research component, the quality of the product has been analyzed to investigate the status of the existing defects, and also the defect names were discovered along with the frequencies of the defect occurrences. A root cause analysis was carried out using the data mining techniques (Association Rule Mining) which decreased the quality cost and also helped in detecting the co-occurrences of garment defects. These observations also were used to prevent the further occurrence of the garment defects by implementing suitable measures.

Following the definition of association rule mining, let I be a set of n binary attributes called items. Let D be a set of transactions. Each transaction in D has a unique transaction ID and contains a subset of the items in I. A rule is defined as an implication of the form,

$$X => Y, \text{ where } X, Y \subseteq I \qquad (1)$$

where X is called antecedent or left-hand-side (LHS) and Y consequent or right-hand-side (RHS).

B. HR Process Optimization Model

The main objective of the proposed methodology is to build a data mining model by studying certain factors that may affect and predict employee performance. Apart from that, the proposed methodology was adopted for the objectives, which is building the classification models to predict the employees' leave management and employee salary management.

C. Decision Tree algorithm

The Decision Tree algorithm recursively visits each node, selecting the optimal split, until no further splits are possible. It uses the concept of information gain to select the optimal split. The algorithm builds decision trees from a set of training data using the concept of information entropy. The training data is a set $S = S_1, S_{2, \ldots,}$ of already classified samples. Each sample S_i consists of a p-dimensional vector $(x_{1i}, x_{2i}, .., x_{pi})$, where the x_j represents attributes or features of the sample as well as the class in which S_i falls [11].

D. Naïve Bayes algorithm

The Bayesian classifier is used for classifying the new dataset based on the training data. Bayes theorem provides a way of calculating the posterior probability, P(c|x), from P(c), P(x), and P(x|c) [9].

$$P(c \mid x) = P(x \mid c) * (P(c)) / (P(x)) \qquad (2)$$

$$P(c \mid x) = P(x_1 \mid c) * P(x_2 \mid c) * P(x_n \mid c) * P(c) \qquad (3)$$

Where P(c|x) is the posterior probability of class given predictor, P(c) is the prior probability of a class, P(x|c) is the likelihood which is the probability of predictor given class and P(x) is the prior probability of predictor.

E. Machinery Failure Prediction Using PdM

The business case defined here has predicted a machinery component failure within the next 24 hours. The used dataset contains five different sets related to machinery and is been used to compute the statistics. The data of errors dataset and telemetry records have been resampled for both 6 hours and 24 hours to gain enhanced outputs. However, this number of hours used in data resampling can change according to the business incident [20]. The maintenance dataset which contains component replacement dates information has been used to calculate the component replacement frequencies in an ad-hoc manner. This type of ad-hoc method is commonly used in

feature engineering to enhance the results [20]. Also, the values of the final data frame were filled backward so that the records were having failure labels for up to 24 hours. For the evaluation process, several ML algorithms were tested on these data to optimize the solution further.

F. Optimized seating arrangement by GA

Here to get the optimum seating arrangement, an approximation (heuristic) based approach: GA is used. This section presents the GA implementation for the optimum seating arrangement. Sewing workflow consists of modules. Each module is assigned with a fixed number of employees and each employee has different skill levels. Due to that employees have different efficiency levels. To represent that, employees are ranked as in Table 1.

TABLE 1. RANKING

Logic			**Rank**
Efficiency		>=75%	2
75% <	Efficiency	>=65%	1
Efficiency		<65%	0

The seating arrangement for a particular style is taken by a work breakdown chart. It contains the basic skill needed for a selected style. The assumptions for the model are as follows,

1. The seating order in the sewing line for a particular style is equal to the order of the skills given in the work breakdown chart.
2. The number of skills given in the work breakdown chart is equal to the number of employees in the module.

To find the maximum efficient seating arrangement the created fitness function is as follows.

$$Y = E1(X1) + E2(X2) + E3(X3) + E4(X4) + \dots \quad (4)$$

$$Y = \sum_{i=0}^{n} E_i X_i \quad (5)$$

Where:

E_i	i^{th} Employee
Y	Efficiency
X	The rank of i^{th} employee related
N	Employee count

One possible seating arrangement for a style is represented with one chromosome. One employee is representing one gene. Considering details of one module let, emp 1 = 0, emp 2 = 1, emp 3 = 2, emp 3 = 2, etc. Let the style number for the chromosome be, STYLE001, and the number of employees in module = 10, be represented as in Fig. 1.

2	6	7	1	0	9	3	8	4	5

Fig. 1. Chromosome

A partially mapped crossover has been used for crossover operation. By considering the seating order and from all possible employee allocation, maximum efficient employee arrangement is given as the final output.

G. Defect Analysis and Visualizations

Initially, the defect data was collected from each department. If any manual processes were available, the relevant processes were computerized. The data which was collected from the respective departments were then combined to design a data warehouse. Microsoft SQL Server Integration Services was used to cleanse the data and perform data engineering. Following the preparation of the data, the data analysis was conducted using power bi and SSAS software. Power bi was also used to create the reports, dashboards, and key performance indicators that were used for the visualizations.

IV. DISCUSSION AND RESULTS

A. HR Process Optimization

This module has been used to predict employee performance, employee leave management, and employee additional salary management and maintain common HR functions.

The attributes used in the Decision Tree model are gender, marital status, employee type, distance, week1 (weekly leave limit exceeded or not), week2, week3, week4. The week1 attribute had the maximum gain ratio, which made it the starting node and most effective attribute while week2, week3 attributes have the next most gain ratios which affects mostly the leave prediction. For the Naïve Bayes model, gender, assigned piece per day, completed pieces per day, efficiency and efficiency class have been used and analysis was performed for a data set of 3,000 records. The accuracies given by the two models are been compared in Table 2 and the Decision Tree model was selected because it had the highest accuracy. Fig. 2 shows the relevant classification matrix results for the model.

TABLE 2. MODEL ACCURACY RESULTS

Subcomponent	**Algorithm**	**Accuracy**
Employee Leave Management	Decision Tree Algorithm	84%
Employee Performance Management	Naïve Bayes Algorithm	89%

	precision	recall	f1-score	support
0	0.84	0.98	0.91	753
1	0.82	0.30	0.44	194
accuracy			0.84	947
macro avg	0.83	0.64	0.67	947
weighted avg	0.84	0.84	0.81	947

Fig. 2. Classification report for the decision tree model

B. Defect Analysis

Considering the minimum requirement is the minimum threshold value of the confidence, and the threshold value will affect the selection process of the rules regularly in the association rule mining process. Therefore, the proposed solution can be adjusted according to the requirements of

the actual production environment if necessary. Table 3 indicates the results of the proposed system.

TABLE 3. ASSOCIATION RULES

Rule	Support	Confidence	Lift
wavy seams -> skipped stitch	0.0075	0.9375	4.00641
tender fabric -> skipped stitch	0.008	1.0	3.44827
wavy seams -> Not marked by directional lines	0.0095	0.95	4.29864

It is also found that the association rules can be changed accordingly when the proposed system is operated with different threshold values for support, confidence, and lift.

C. Machinery Failure Prediction Using PdM

ML models like Logistic Regression, SVM, and Random Forest (RF) algorithm were used to use train the model. But RF was chosen to optimize the solution further as its computation speed was higher than the rest of the classifiers. However, though the model accuracy was high in this result with proper confusion matrix and classification report, the labels had imbalanced class counts favoring only a single class.

Hence, to tackle this problem the data were resampled using the resampling techniques and Synthetic Minority Oversampling Technique (SMOTE) method from the sklearn imbalanced library. After resampling the data, the same model was used to train and to obtain the results with balanced class counts time as shown in Fig. 3.

```
Confusion Matrix:
[[19391     0   164   119    59]
 [  114 19489    46   182     0]
 [    0     0 19741     0     0]
 [    0     0    72 19873     0]
 [  596   483   151   282 18238]]
Classification Report
              precision    recall  f1-score   support

       comp1       0.96      0.98      0.97     19733
       comp2       0.98      0.98      0.98     19831
       comp3       0.98      1.00      0.99     19741
       comp4       0.97      1.00      0.98     19945
        none       1.00      0.92      0.96     19750

    accuracy                           0.98     99000
   macro avg       0.98      0.98      0.98     99000
weighted avg       0.98      0.98      0.98     99000
```

Fig. 3. Confusion matrix and classification report

Furthermore, using this tuned model, it was able to predict the failure probabilities using the predict_proba function along with the predicted class label. Finally, risk levels were allocated to present the failure warning so that if all four components were to fail, a level 1 warning is given, and the other levels will be given for the relevant component failure combinations.

D. Optimized seating arrangement by GA

When the user inputs the style, number of employees in the schedule, and types of tasks with the order, as in the methodology, steps are carried out and the output is displayed as the optimum seating arrangement. Since seating arrangement optimization is a maximization problem, the fitness values should be increased over the generation and it is considered to be an improvement in the fitness value. As shown in Fig. 4, after some point improvement of the fitness function, happens very rarely.

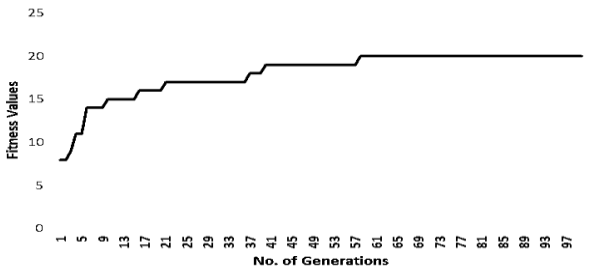

Fig. 4. Optimum Seating Arrangement (Effect of No of Generations to Fitness Value)

V. Conclusion and Future work

This research paper captures the major fields that need to be enhanced in the garment manufacturing processes which contains defect analysis, HR process optimization, machinery failure predictions, and optimized seating arrangements.

To obtain a satisfactory product, adequate amounts of data should be available for the system to process effectively. Or else, the final output might vary in accordance with the input data. Hence it is suggested that employees with appropriate knowledge of the corresponding departments to use the system be handled as per the requirements.

Other than the above-mentioned fields, many other areas should be subjected to optimization in the garment industry. These areas contain, defects analysis using computer vision, capturing employee behaviors to optimize the workflow, minimizing production costs, etc. Therefore, it is required that this system be assessed in relevant manners so that the necessary demands could be met with fine-tuned results in the future.

ACKNOWLEDGEMENT

We thank senior lecturer Mr. Jesuthasan Alosius, assistant lecturer Ms. Vijani Piyawardana and all the other lecturers at the Sri Lanka Institute of Information Technology for the assistance and guidance provided and for the comments that greatly improved the paper.

REFERENCES

[1] Phyu, T.N., (2009). Survey of classification techniques in data mining, Proceedings of the International Multi Conference of Engineers And Computer Scientists, IMECS 2009, Vol.

[2] T. Wuest, D. Weimer, C. Irgens, and K.-D. Thoben, “Machine learning in manufacturing: advantages, challenges, and applications,” Production & Manufacturing Research, vol. 4, no. 1, pp. 23–45, 2016.

[3] "Total Quality Management (TQM)", Managementhelp.org, 2020. [Online]. Available:

https://managementhelp.org/quality/total-qualitymanagement.htm. [Accessed: 22- Feb- 2020].

[4] "Cause and Effect Analysis: Identifying the Likely Causes of Problems", Mindtools.com, 2020. [Online]. Available: https://www.mindtools.com/pages/article/newTMC_03.htm. [Accessed: 22- Feb- 2020].

[5] Yusi Cheng "GA-based multi-level association rule mining approach for defect analysis in the construction industry - ScienceDirect".

[6] "QMS - RES.Q", RES.Q, 2020. [Online]. Available: http://resq.cloud/qms/. [Accessed: 22- Feb- 2020].

[7] C.K.H. Lee "A hybrid OLAP-association rule mining-based quality management system for extracting defect patterns in the garment industry - CityU Scholars | A Research Hub of Excellence". [28th Jan 2020]

[8] "How Total Quality Management (TQM) Works", Investopedia, 2020. [Online]. Available: https://www.investopedia.com/terms/t/total-qualitymanagement-tqm.asp. [Accessed: 23- Feb- 2020]

[9] Himanshu Kumar Singh, Kshitij Vishnavat and R. Srinivasan, "Employee performance and leave management using Data Mining technique", in International journal of pure and applied mathematics, Volume 118, No. 20, 2018, 2063 – 2069

[10] Hamidah Jantan, Abdul Razak Hamdan and Zulaiha Ali Othman, "Human talent prediction in HRM using C4.5 classification algorithm", in (IJCSE) International Journal on Computer Science and Engineering, Vol. 02, No. 08, 2010, 2526-2534

[11] V. Kalaivani and M. Elamparithi, "An efficient classification algorithms for employee performance prediction", in International Journal of Research in Advent Technology, Vol.2, No.9, September 2014 E-ISSN: 2321-9637

[12] John M. Kirimi and Christopher A. Moturi, "Application of Data Mining Classification in Employee Performance Prediction", in International Journal of Computer Applications (0975 – 8887) Volume 146, No.7, July 2016

[13] G. A. Susto, A. Schirru, S. Pampuri, S. Mcloone, and A. Beghi, "Machine Learning for Predictive Maintenance: A Multiple Classifier Approach," IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 812–820, 2015.

[14] Rama Krishna Reddy Guduru, Saddam Hussain Shaik, Satyanarayanamma Yaramala, Narayana Prakash TMS and Aurelijus Domeika, "A Dynamic Optimization Model for MultiObjective Maintenance of Sewing Machine," International Journal of Pure and Applied Mathematics, vol. 118, no. 20, pp. 33–43, Apr. 2018.

[15] S. Kursun, "Assembly Line Balancing in Garment Production by Simulation", 2020. [Online]. Available: https://www.intechopen.com/books/assembly-line-theoryand-practice/model-sequencing-andassembly-linebalancing-in-garment-production-by-simulation-workertransfer-syst#B7. [Accessed: 24- Feb- 2020].

[16] K. Leksakul and S. Phetsawat, "Nurse Scheduling Using Genetic Algorithm", Mathematical Problems in Engineering, vol. 2014, pp. 1-16, 2014. Available: https://www.hindawi.com/journals/mpe/2014/246543/.

[17] K. Alzoubi, "Facility Planning and Assembly Line Balancing in Garment Industry - IEEE Conference Publication", Ieeexplore.ieee.org, 2020. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8907686. [Accessed: 24- Feb- 2020].

[18] H. Xu, B. Xu and J. Yan, "Balancing apparel assembly lines through adaptive ant colony optimization", Textile Research Journal, vol. 89, no. 18, pp. 3677-3691, 2018. Available: 10.1177/0040517518819836 [Accessed 24 February 2020].

[19] K. chan, "Fuzzy operator allocation for balance control of assembly lines in apparel manufacturing - IEEE Journals & Magazine", Ieeexplore.ieee.org, 2020. [Online].Available:https://ieeexplore.ieee.org/abstract/document/1010885?section=abstract. [Accessed: 24- Feb- 2020].

[20] F. B. Uz, "Predictive Maintenance Modelling Guide," Azure AI Gallery, 25-Mar-2016. [Online]. Available: https://gallery.azure.ai/Collection/Predictive-MaintenanceModelling-Guide-1. [Accessed July 2, 2020].