K-Means Cluster Algorithm Analysis for the Performance Categories of the Sewing Operators in the Garments Industries

S. M. Masum Alam[*]

Department of Textile Engineering, Southeast University, Dhaka - 1208, Bangladesh
*Corresponding author: E-mail: s.masum@seu.edu.bd

Abstract
A worker's career is significantly impacted by performance reviews since they offer benefits such as permitted wage raises, bonuses, termination options, and promotions. Because they only have a limited budget for hiring people, businesses can use assessments to select the finest candidates. As a result, the performance rating offered by the machine learning algorithms will be advantageous to both the employer and the employee. The research work's findings are classified as great, good, and average depending on factors including age, attendance rate, efficiency rate, experience, productivity, and daily earnings. Out of the 251 workers at the garment factory, we found that a total of 56 individuals fit the exceptional category, 87 individuals fit the decent group, and 108 individuals fit the medium category.

Keywords: K-Means clustering, Performance, Apparel, Box Plot, Productivity, Machine learning.

1. Introduction

Performance reviews have a significant impact on an employee's career because they provide advantages such allowed pay raises, bonuses, termination options, and promotions. Businesses can choose the best employees with the help of assessments because they just have modest money for providing them. As a result, both the company and the employee will benefit from the performance rating provided by the machine learning algorithms [1]. An employee's career is greatly impacted by performance reviews because they allow for benefits such as authorized pay increases, bonuses, termination choices, and promotions. As they only have a small budget for offering appraisals, it aids businesses in selecting the best personnel [2]. Therefore, the machine learning algorithms will offer a better, bias-less and more accurate performance evaluation that is advantageous to the employer and the employee.

In data science, clustering is a helpful tool. It is a technique for determining the cluster structure in a data set that is indicated by the highest degree of similarity inside a cluster and the highest degree of dissimilarity across groups [3]. The first clustering technique employed by biologists and social scientists was hierarchical clustering, whereas cluster analysis developed into a subfield of statistical multivariate analysis. Every business uses performance appraisals as part of its evaluation process to see whether its employees have been meeting its objectives and strategic plans [4]. It is beneficial to assess each and every applicant for in terms of their knowledge, abilities, and strengths weaknesses and provides feedback for upcoming development. The overall performance of the company comprises each individual contributions from each employee that lead to the outcome or actual output that is compared to the desired outcome which higher-ups in the business set out to accomplish it [5]. The evaluation process is typically erratic and frequently own opinions, which cause the organization to utilize a technique that corrects for all such abnormalities [6]. This is the implementation of an intelligent performance is presented in this study. An

evaluation model that overcomes such factors and employing machine learning methods, detect irregularities the current performance review environment includes more inaccurate manual work is done in the evaluating process [7]. It entails a number of factors where the evaluation will only be positive for a specific group of workers based on specific criteria, but We can enhance and by utilizing machine learning techniques Boost the precision to ensure that the staff will obtain the appraisal without bias and will precious a fair evaluation, which involves taking into account all necessary criteria. [8]. We develop a model for classifying the output of the employees using unsupervised machine learning K-Means systems.

1.1 Related Works

Unsupervised machine learning, which works with unlabeled data, is a component of Employee Performance evaluation. That information includes characteristics like work quality, experience, age, efficiency and other similar characteristics. The unsupervised machine learning algorithms used in multidisciplinary area in real life like cotton Fibre bale management system [9] agriculture sectors to classify seeds[10], Textile wet processing colors categories[11][12]Garments forecasting system[13] medical science to categories Diabetics [14]and data mining system[15] etc. sectors can apply this algorithm.

2. Materials and Methodology

The data of this research work was collected from the apparel industries which is located in Gazipur. A total of 2000 employees are working in that factory, among them 251 sewing operator's information we have collected. A questionnaire was prepared for the collection of the data information. The major categories which are considered for the evaluation of employee performance are based on operator's gender, age, attendance percentage, and operator's efficiency percentage, working experience, production per day and wages per day.

Productivity is directly affected by human performance. The appraisal of human performance is mainly evaluated by the choice of management people and also by the conventional system which is not always effective. The factory was produced 100 per cent cotton non-denim five-pocket Longbottom for European style. The Fabrication of that product was 3/1 (Three up one down) Twill weave Fabric structure and 280 GSM (Gram per Square Meter) weight of Fabric. There were 5 Garments production lines and the total number of operators was 251 persons.
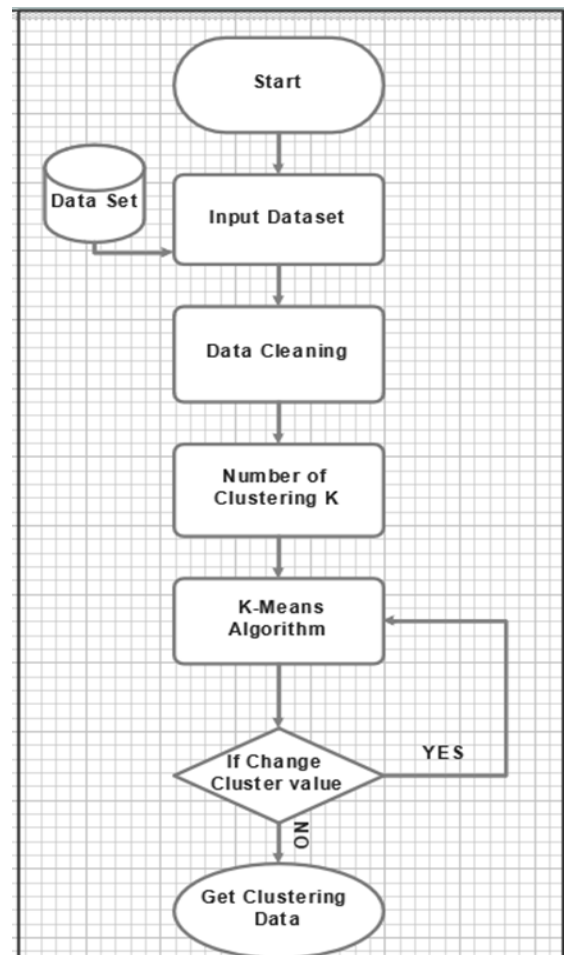

Figure 1. Data process diagram of k-means clustering.

The K-Means Clustering algorithm of Machine learning is applied here to categorize the sewing operator's performance. The manpower was divided into K=3, three categories excellent, good

and average. The excellent categorization was age=24 years, Attendance percentage = 100, Efficiency Percentage = 100, Experience = 40 month, Production per day = 900 pcs and wages = Tk.700, The good categorization was age=30 years, Attendance percentage = 95, Efficiency Percentage = 95, Experience = 60 month, Production per day = 800 pcs and wages = Tk.600 and the average categorization was age=35 years, Attendance percentage = 85, Efficiency Percentage = 85, Experience = 75 month, Production per day = 600 pcs and wages = Tk.500. But a machine learning system can categorize the human performance system automatically and effectively. The sample of the data shown below table no-1.

Table 1. Performance data of the operators.

Sl#	Operators Male/Female	Age (years)	Attendance Percentage	Efficiency (Percentage)	Experience (Month)	Productivity (Per day pcs)	Wages TK. (Per day)
1	Male	18	95	77	12	998	480
2	Female	19	78	62	12	650	430
3	Male	20	92	92	24	820	510
4	Female	21	100	76	18	860	440
5	Female	22	98	110	26	1000	560
6	Male	23	96	94	32	625	530
7	Female	24	100	100	40	900	700
-	Male	25	94	94	42	825	650
-	Male	26	90	88	45	750	600
-	Female	30	95	95	60	800	600
-	Male	32	96	90	72	725	550
251	Female	35	85	85	75	600	500

The operator's performance information is shown in table number 1. Here, the employee's gender, age, attendance percentage, efficiency percentage, productivity, experience, productivity and wages were considered for the evaluation of the human performance. The experience of the operators was considered month-wise because some workers were new and some were experienced. also, wages were calculated per day wisely because the authority reduced the salary for the number of days of absenteeism of the operators.

3. Results discussion

From the figure 2 we can analysis the performance wise categories of the sewing
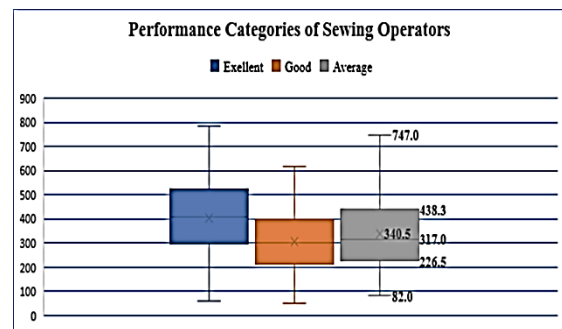
Figure 2. Performance categories of sewing operators.

The box plots used in this research work represent the various parameters graphically at a glance directly. Here, Clusters (24, 100, 100, 40, 900, 700 = K1), (30, 95, 95, 60, 800, 600 = K2) and (35, 85, 85, 75, 600, 500 = K3) shows the parameters of age, attendance percentage, efficiency percentage, experience, productivity and daily wages sequentially treat as excellent, good and average. In the graph, an excellent box plot shows the maximum value is 786 (without the outline), the lower 1st Quartile value is 295, the mean value is 406, the median value is 403, the upper 3rd quartile value is 524 and the lower minimum value is 58 (without outer line), the good box plot shows the maximum value is 618 (without the outline), the lower 1st Quartile value is 213, the mean value is 308, the median value is 306, the upper 3rd quartile value is 398.7 and the lower minimum value is

51 (without outer line) and finally, the average box plot shows the maximum value is 747 (without the outline), the lower 1st Quartile value is 226.5, the mean value is 340, the median value is 317, the upper 3rd quartile value is 438.3 and the lower minimum value is 82 (without outer line).
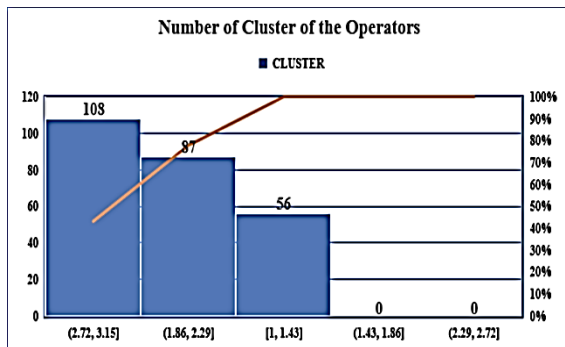

Figure 3. Number of cluster of the operators.

Figure 3 shows that the total number of 56 person operators was in the excellent categories, the good category was 87 employees and the total average quality operators were 108 persons among the total 251 person operators of that apparel factory [16].

4. Conclusion

An evaluation model that gets around these issues can spot abnormalities. We create a model utilizing K-Means methods for unsupervised machine learning. Increase the accuracy to make certain that the staff will receive the review without bias, which involves taking into account all relevant factors. The results of the research work based on the parameters of age, attendance percentage, efficiency percentage, experience, productivity and daily wages are sequentially treated as excellent, good and average. Finally, we discovered that out of the 251 operators in the garment plant, a total of 56 operators fell into the exceptional category, 87 employees fell into the good group, and 108 operators fell into the medium category.

REFERENCE

[1] S. Ilhan Omurca, C. Ozkan, and G. Aydin Keskin, "Topic extraction from user reviews View project A VARIANT PERSPECTIVE TO PERFORMANCE APPRAISAL SYSTEM: FUZZY C-MEANS ALGORITHM," 2014. [Online]. Available: https://www.researchgate.net/publication/265646832

[2] IEEE Staff, 2018 International Conference on Innovations in Information Technology (IIT). IEEE, PP 93-98, 2018.

[3] K. P. Sinaga and M. S. Yang, "Unsupervised K-means clustering algorithm," IEEE Access, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2988796.

[4] B. Kanaka Durga Prasad, "Performance Evaluation Model using Unsupervised K-Means Clustering," 2020, doi: 10.1007/s10586.

[5] E. Gothai and P. Balasubramanie, "Performance evaluation of hierarchical clustering algorithms," 2010.

[6] T. Velmurugan, "Performance based analysis between k-Means and Fuzzy C-Means clustering algorithms for connection oriented telecommunication data," Applied Soft Computing Journal, vol. 19, pp. 134–146, 2014, doi: 10.1016/j.asoc.2014.02.011.

[7] S. Hartmann, "Hartmann, Sönke A general framework for scheduling equipment and manpower on container terminals A General Framework for Scheduling Equipment and Manpower on Container Terminals," 566. [Online]. Available: http://hdl.handle.net/10419/147633

[8] H. Kong, Y. L. Zhi, A. Lim, B. Rodrigues, Y. Zhi, and B. Manpower, "Manpower Allocation with Time Windows and Job Teaming Manpower Allocation with Time Windows and Job Teaming Constraints Constraints

Xi LI Citation Citation," 2003. [Online]. Available: https://ink.library.smu.edu.sg/lkcsb_research

[9] J. I. Mwasiagi, H. Xiubao, W. Xinhou, and C. Qing-Dong, "THE USE OF K-MEANS AND KOHONEN SELF ORGANIZING MAPS TO CLASSIFY COTTON BALES."

[10] J. Immaculate, S. Terence, J. H. Immaculate, E. P. Ebenanjar, and S. J. Terence, "Applications of Machine Learning Algorithms in Agriculture", [Online]. Available: https://www.researchgate.net/publication/341179298

[11] Samane Sistani et. al "The Discovery of Major Heart Risk Factors among Young Patients with Ischemic Heart Disease Using K-Means Techniquesircrj-64781"International Cardiovascular Research Journal, vol. 13, pp. 85-90, 2019 "ircrj-64781".

[12] Z. Wang, Z. Zhuang, Y. Liu, F. Ding, and M. Tang, "Color classification and texture recognition system of solid wood panels," Forests, vol. 12, no. 9, Sep. 2021, doi: 10.3390/f12091154.

[13] P. Lai and S. Westland, "Machine learning for colour Palette extraction from fashion runway images," International Journal of Fashion Design, Technology and Education, vol. 13, no. 3, pp. 334–340, Sep. 2020, doi: 10.1080/17543266.2020.1799080.

[14] J. M. David and K. Balakrishnan, "SIGNIFICANCE OF CLASSIFICATION TECHNIQUES IN PREDICTION OF LEARNING DISABILITIES."

[15] Y. Feng, "ANALYSIS ON ALGORITHM AND APPLICATION OF CLUSTER IN DATA MINING," J Theor Appl Inf Technol, vol. 15, no. 1, 2012, [Online]. Available: www.jatit.org

[16] Dr. S. Manoharan, "Performance Analysis of Clustering Based Image Segmentation Techniques," Journal of Innovative Image Processing, vol. 2, no. 1, pp. 14–24, Mar. 2020, doi: 10.36548/jiip.2020.1.002.