ISSN: 2278-3369

International Journal of Advances in Management and Economics

Available online at: www.managementjournal.info

RESEARCH ARTICLE

AI-DRIVEN HRM FRAMEWORK FOR EMPLOYEE SEGMENTATION AND PERFORMANCE OPTIMIZATION USING AUTOENCODER-BASED CLUSTERING

Mohan Reddy Sareddy[1], R. Pushpakumar[2]

[1]*Orasys LLC,Texas, USA.*

[2]*Department of Information Technology, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Tamil Nadu, Chennai, India.*

Abstract: The rapidly evolving marketing, employee performance evaluation is unarguably a data-driven approach, thus requiring employee segmentation and performance optimization. This study proposes a full employee segmentation and performance optimization process using K-means segments. This algorithm accommodates decision-making based on key employee data, such as productivity, satisfaction, and engagement. Through soundly collecting, preparing and analyzing employee data, organizations can actively design targeted interventions toward the segments they classify, ultimately impacting all interventions directed toward improving performance and retention. The findings indicate that the K-Means clustering algorithm reveals different employee profiles, which leads to much more effective HR strategies congruent with the organization's objectives.Such a study suggests an AI-based employee segmentation and performance optimization framework through K-means clustering—a machine learning technique. By integrating artificial intelligence in HR analytics the framework facilitates wiser decision-making based on principal employee data such as productivity, satisfaction and engagement.

Keywords: *HRM, human resource management, employee segmentation, performance optimization, k-means clustering, artificial intelligence, deep learning, autoencoders, machine learning, intelligent decision support, ai-driven analytics, predictive modelling.*

Article Received: 01 April 2022 **Revised: 08 April 2022** **Accepted: 18 April 2022**

INTRODUCTION

Human Resource Management is seen as vital in fostering organizational success within the current competitive business environment [1]. The rapid expansion of digital technologies and the expanding access to workforce data enabled HR departments to increasingly embed decision-making in data-driven practices [2]. Intuition and common practices that once defined traditional HRM in the past, have now, in many contexts, been replaced or augmented with data-analytic approaches [3]. One such approach includes the use of machine learning (ML) algorithms to use data from employees to provide knowledge on how better talent is managed and enhanced productivity [4].

Multiple aspects have influenced the move toward data-driven HRM. One of the factors is the sheer volume of employee data from various sources such as performance reviews, surveys on job satisfaction and engagement, attendance records and tools for managing project and team progress [5]. Organizations are becoming aware of the need to segment employees further to improve customization of engagement strategies, enhance groups and teams, and align individual objectives to the aspirations of the organization. Agile, evidence-based workforce planning needs have pressurized HR departments into hastening an embrace of data scientist methodologies in search of patterns and trends that do not otherwise become apparent through plain analysis[6].

Even in the presence of data and modern analytical tools, many a Human Resources department finds itself falling short of making quintessential use of them [7], [8]. Some commonly faced challenges include poor data cleaning, lack of technical know-how, reluctance to change, and difficulty to integrate insight into existing HR practices [9]. Also, working without proper structure may lead to employee segmentation, giving rise to improper interpretations or biased results [10]. The other outcome of this lack of structure is the underutilization of workforce insights that can be instrumental in Performance Enhancement [11].

A holistic data-driven HRM framework is proposed in this paper to untangle the issues from data collection and cleaning, through feature development and clustering, to extracting actionable performance metrics. When used on well-prepared HR data consisting of employee features through a K-Means clustering algorithm, organizations obtain relevant employee segments for interventions and strategic HR planning.

This framework helps remove technical and organizational barriers while also showcasing the potential of machine learning in making HRM more predictive and proactive. Artificial Intelligence (AI) in the form of machine learning, is transforming HRM by empowering smart segmentation, forecasting and automation of workforce decisions.

Problem Statement

In the present data-heavy corporate environment, organizations face major challenges in managing human resources due to a lack of precise analytical and scalable tools [12]. Those classical HRM practices usually do not really segment the personnel, creating generic policies often ignorant of differences in productivity, satisfaction, and risk of employees to leave [13], [14]. Although algorithms for clustering provide promising solutions, problems such as the choice of first cluster centers[115].

The presence of unrelated attributes, or the instability of the segmentation still hamper their use [16]. Hence, it is more than ever needed that companies enhance tools capable of predicting employee performance [17]. identifying inefficiencies, and forecasting attrition, especially in dynamic industries such as IT where talent optimization actually matters [18].

Data-absent HR approaches prevent decisions regarding organizational performance and employee satisfaction from being made in a timely and targeted manner [19]. Even with large HR datasets available the lack of AI-powered analytics constrains potential predictive and strategic workforce planning [20].

Objective

- Examine how various talent acquisition strategies in Human Resource Management are improved through data-driven analytics.
- Assess the effect that predictive analytics has on employee retention and the engagement initiatives within the IT sector.
- Identify loopholes in the application of traditional, intuitive approaches of HR toward contemporary workforce issues.
- Discuss the impediments to the adoption of HR analytics such as concerns with data privacy, issues in the integration with other HR functions, and skillset limitations.
- Recommend strategies for transforming HR from a reactive approach to a proactive planning model through the use of analytics.
- Assess the use of AI methods, such as unsupervised machine learning, to broaden employee segmentation, forecast attrition and refine engagement strategies.

LITERATURE REVIEW

This working paper explores the change-relationship of integrating artificial intelligence (AI) within staff planning, exploring how machine learning algorithms and analytical analytics can maximize acquisition of employees, forecast workloads, maintain employees and assign resources. Using a combination of historical and real-time data, AI can project trends in the labor market, identify skill shortages, and help to build proactive recruitment and development strategies.

This working paper compares some machine learning methods, like regression, decision trees, and neural networks, as well as the accuracy of the workforce predictions into actionable analysis. This paper shows AI-enhanced planning is saving organizations time while allowing for operational efficiency and improved effectiveness, but highlights the ongoing issues with data privacy and ethical work practices. The ideas throughout the paper reaffirm there needs to be a blend of expertise and collaboration between AI experts and HR professionals to facilitate responsible and sustainable use of AI in workforce management [21].

This research article addresses the design, implementation, and authentication of a data-driven Mamdani-type fuzzy clinical decision support system based on clusters and pivot tables. The methods were applied to five public datasets (including Wisconsin and Coimbra breast cancer, wart treatment, and caesarian section) all previously addressed in the literature.

The results demonstrate acceptable levels of accuracy, with Kappa statistics approaching 1.0 and accuracies approaching 100%, which was also superior to the previous work. The proposed framework is classified as a deep learning technique which utilizes combinations of processes, or stacked or multiple processing layers, as its essential unit of learning data representations with increasing levels of abstraction [22].

This article has proposed a new model, Convolutional Neural Network-Clustering, using deep learning algorithms for diagnosis of myocarditis, a challenging cardiac condition often due to viral infections. The study took 47 subjects and nearly 100,000 pictures (98,898), achieved an accuracy of 97.41% through a 10-fold cross-validation procedure with 4 clusters. Cardiac MRI is critical to diagnose myocarditis, but the presentation of its findings could be affected by clinical factors as well as imaging factors.

This research is the first to smear deep learning algorithms for the diagnosis of myocarditis, presenting a smart and talented design which produced results that improved the accuracy of diagnosis [23].

This paper evaluates the increasing significance of Big Data Analytics (BDA) in data-driven firms. It focuses on the pressing issues and challenges linked with data generation, storage, visualizations of analytics, and processing generated data. It provides an overview of existing BDA tools and methods, their strengths and weaknesses.

It categorizes the main sources of data and applications with projected significant influence.

It also reviews the major issues and challenges related to using

BDA and posits some solutions, whilst also highlighting continuing research opportunities. Ultimately, this paper provides a basis for scholars, practitioner-industry personnel, and policymakers to understand the current complexities associated with BDA usage and move the area forward [24].

Machine learning (ML) has emerged as a powerful, game-changing automation technology that enables data-driven solutions that improve operations and accuracy across sectors. Within an implementation matrix and operationalization framework, ML applications for predictive maintenance, quality control, and real-time monitoring of supply chain optimization help minimize costs, reduce human input, and enhance decision quality.

Even more advanced models such as deep learning and neural networks further increase operational accuracy while challenges exist with respect to data quality and integration of data sources need to be accounted for. As ML continues to develop alongside edge computing and explainable AI concepts, operational automation will improve operational excellence and industrial sector innovation (i.e., manufacturing, healthcare, logistics) [25].

Automation is being transformed by machine learning with data-driven solution(s) to optimize processes and gain accuracy across numerous industries (e.g., manufacturing, healthcare, and logistics). ML potential use cases in predictive maintenance, quality control, and supply chain optimization have the potential to lower costs, improve decision making, and reduce manual labor requirements.

More advanced model types like deep learning increase accuracy, yet deployment challenges like data quality, integration, and ethics, remain to be overcome for machine learning successes to take root. The continued evolution of automation will be enhanced by edge computing (connecting closer to the data source) and explainable artificial intelligence (improved transparency and better understanding of automation outputs) [2].

This research evaluates the role of internet-inclusive finance in economic development across rural Africa, with particular focus on e-commerce. The methods were based on mainly statistical regression and econometric modelling techniques, and in some aspects machine learning.

Factors related to mobile internet admission and access, as well as financial inclusion, had positive and significant effects on income level, entrepreneurship, and business development in rural areas. These findings drew attention to the need for more internet access and financial inclusion in rural Africa, evidence being an important factor in addressing the urban-rural divide, and broader social economic development in accordance with global development goals [27].

This research describes the application of unsupervised machine learning algorithms, including K-means, agglomerative and spectral clustering methods, for event detection and flood location detection in urban stormwater drainage systems. By evaluating simulations of water depth datasets, the clustering accuracies were compared for each clustering technique using silhouette coefficient, Calinski–Harabasz index, and Davies–Bouldin index metrics.

The study found K-means and agglomerative clustering techniques to work well for short-duration events and spectral clustering to be superior for long-duration floods. Overall, this study provides avenues for improving flooding detection and ultimately early warning systems in urban stormwater systems [28].

The Earth's geosystem, including the solid Earth, oceans, and atmosphere, is a complex, interconnected system prejudiced by physical, chemical, and biological processes. As development increases, human connections with the geosystem are flattering more consequential, particularly in hazard-prone areas.

To better understand this system, geoscientists must analyze large geo-datasets and conduct computational simulations, often using machine learning (ML) techniques. These efforts aim to recover predictive understanding of geohazards and the long-term behavior of Earth’s geosystem, which is crucial given the finite nature of resources and the growing impact of geohazards [29].

This paper investigates methods for utilizing rule mining methods to discover and represent temporal patterns from clinical sensor data streams (heart rate, respiration rate, blood pressure, etc.). This research shows a new similarity method for comparing rule sets and uses natural language generation methods for representing temporal relations between patterns. The MIMIC online database has been used to validate the method.

The extracted pieces of mined rules show that different patterns are produced predictors for various clinical conditions (respiratory failure, angina, and sepsis) and can be conveyed as textual representations. [30] Current research identifies the implementation of AI in various sectors such as HR, where algorithms like K-Means, decision trees and deep learning models are used to identify workforce trends and enhance decision-making accuracy.

PROPOSED METHODOLOGY

This is a workflow of a systematic, data-based method of streamlining HRM using machine learning. It starts by gathering data regarding productivity and satisfaction of the employees, then the cleaning and preparation of the data is done, to make the data easy to understand and correct.

Patterns and trends are then discovered with the help of exploratory analysis, resulting in the development of features that make this dataset more relevant to clustering.

The K-Means clustering is used to divide the employees into relevant clusters that allow the particular thoughts. Lastly, it is possible to accomplish performance measures to determine and improve HR strategies per the identified segments.

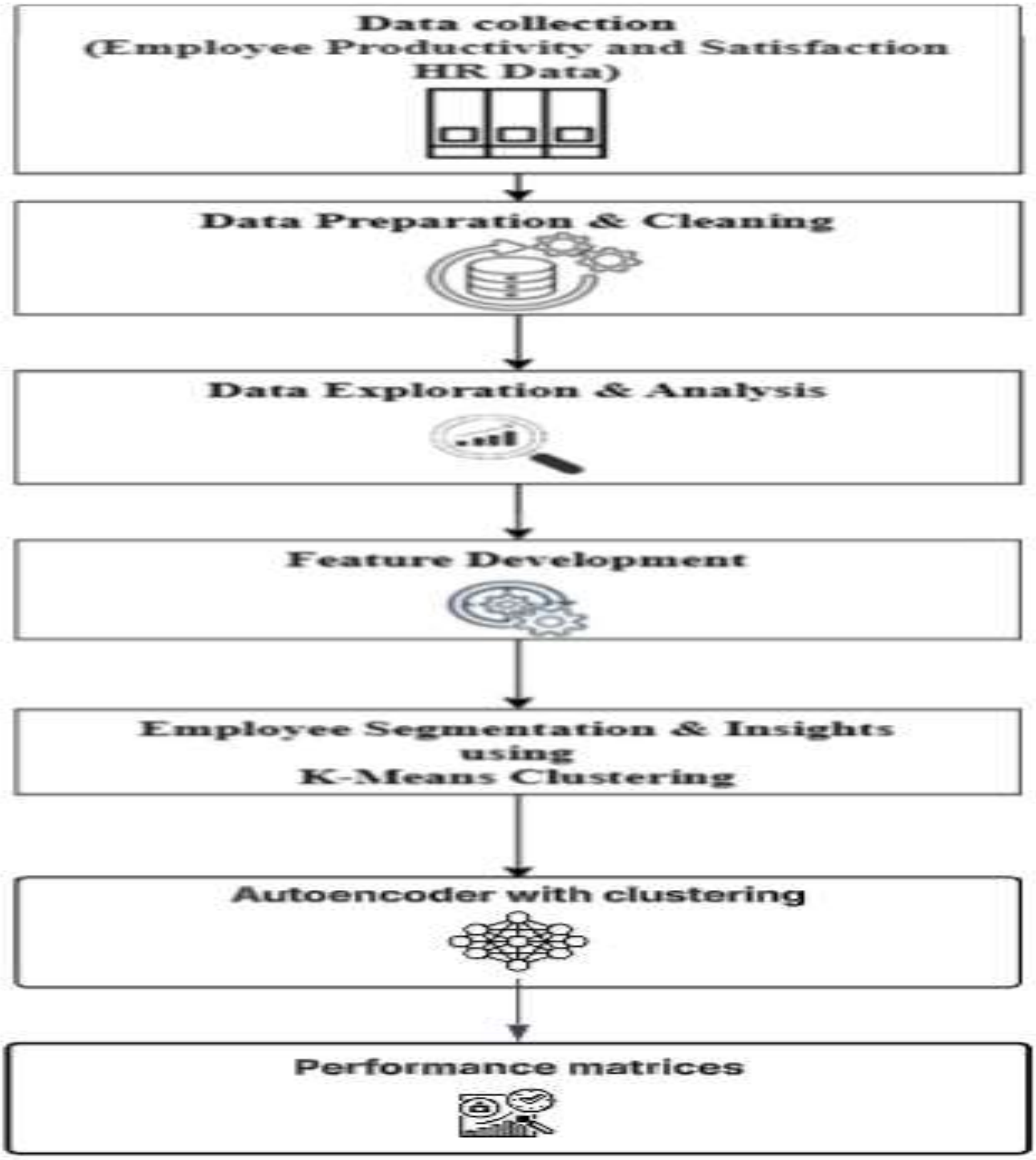

Figure 1: Data-Driven HRM workflow for employee segmentation and performance evaluation

Data Collection

The most fundamental act of this process is data collection in the establishment of any data-driven HRM framework. This type of information shown in this study entails productivity and satisfaction of the employees based on internal HR systems, performance reviews, surveys and time attendance records.

Work hours, and outcomes on the projects, rating at evaluation of job satisfaction, and feedback ratings, and demographic demographics are the simplest examples of variables contained in this data.

It is aimed at getting the big picture of employee behaviour and attitude in the organization. Valid and thorough data provides that subsequent analysis is useful and is representative of the actual dynamics of employees. A good quality of a dataset enhances a better quality of clustering and insights discovered at later stages.

Data Preparation & Cleaning

This step would entail the conversion of raw HR data to a clean and structured data format which can be analysed. It facilitates uniformity, precision and usability of the data set. The most frequent responsibilities are to correct mistakes in data entry, fix inconsistency, and data format validation. Through appropriate preparation, bias is eliminated, and downstream machine learning tasks are enhanced. It furthers the preparation of robust feature engineering and correct clustering. In the absence of such an action, even the most sophisticated algorithms will turn out to be misinformed.

Handling Missing Data

Missing data could be contributed by the non responses in surveys or errors in the system or incomplete entries. Missingness is handled in different ways according to both the type and the extent with deletion, imputation, or model-based estimation. It is also important to study the trend of missingness whether it is random or systematic. There is nothing wrong with handling the data properly, because this way, it will be representative and non-biased. Constructions that fail to take into consideration missing values can make clustering incorrect and cause bad segmentation of employees.

$$\bar{x} = \frac{1}{|N|}\sum_{i \in N} x_i \quad (1)$$

Outlier Treatment

Outliers refer to those values that differ strongly with the most of the data points normally because of mistakes or unusual actions. They may skew averages and even influence the functioning of such algorithms as K-Means that are based on distance metrics. ANOMALY Identification Outlier techniques like IQR, Z-Score or just visual plots can be used to identify anomalies. An outlier can be eliminated, ceiling, or adjust once identified based on the situation. Proper treatment improves the robustness model and accuracy in segmentation.

Standardization

Because K-Means utilizes the Euclidean distance, all variables have to be on the same scale to make a fair contribution. Standardization changes the data so that it has a sort of mean of zero as well as standard penetration of one. Normalization transforms values between 0 to 1. This prevents the dominion of variables with higher varieties over variables of smaller scales. Scaling causes proportional and significant clustering.

$$Z = \frac{X-\mu}{\sigma} \quad (2)$$

Were, Z = standardized value , X = original value, μ = mean of the feature, σ = standard deviation of the feature.

Data Exploration & Analysis

An important step is data exploration and analysis which assists one in knowing the background and quality of the data. During this step, the central variables will be summarized with the help of descriptive statistics and patterns, spreads, and possible anomalies can be observed through visualizations that may include histograms, box plot, and correlation matrices. It assists in determining correlations between variables e.g. how the level of satisfaction is correlated with the productivity or turnover. One can identify outliers and inconsistencies in data at this step as well. The knowledge obtained here will be applicable in the selection and refinement of features such that only relevant and meaningful data will be applied in the modeling process. This move will catalyze good clustering and decision-making.

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^2} \quad (3)$$

Feature Development

The development of the features entails the construction, altering or selection of limited variables that augment the predictive ability of the set of data. In HRM, this can contain by-products like tenure of employees, performance to attendance ratio or engagement index determined by responses to surveys.

To differentiate the groups of employees, effective feature engineering points out a pattern that cannot be directly observed in raw data. This procedure implies that this clustering algorithm can be used to discover meaningful features like levels of motivation, productivity rates, or risk of exit among others. Subtly designed attributes have a direct effect on precision and extensibility of the finalized division model.

Employee Segmentation & Insights using K-Means Clustering

During this step, it is assumed that the HR dataset has been prepared and feature-enhanced to be used in the K-Means clustering procedure to group employees into different collections depending on similarities with regard to the behaviors and their performance. Every cluster is linked to the distinct positions of employees, including very productive ones, ones with poor engagement, and risky ones. Such segmentation can enable HR practitioners to shift to the individualized approach to interventions, forgetting about one-size-fits-all approaches.

Through learning about regularities in the clusters, organizations can work out specific training, engagement and retention strategies. K-Means can, therefore, help in unearthing the unseen patterns and in strategically influencing the people decisions.

Autoencoder-Based Clustering Approach

Autoencoders are neural networks that learn to encode input data and then decode it. Here, we utilize the encoder's latent representation to cluster in a more semantic space. This eliminates noise and dimensionality complexity while maintaining important behavioral characteristics of employees. The segmented outputs offer richer, non-linear segmentation, which enables HR to create more customized strategies.

RESULTS AND DISCUSSIONS

In predicting the efficiency of future data-driven HRM framework in segmenting employees through K-Means clustering, the results have been shown. With PCA visualization, there was the differentiation between these three distinct groups including high performers, underperformers, and moderately engaged employees. Cluster profile analysis demonstrated the significant differences in satisfaction, productivity, feedback, and tenure, which is relevant when conducting HR-related strategies individually.

This close relation with a productivity level also highlights the importance of engagement activities among employees. The insights convey to the HR professionals the ability to create evidence-based interventions such as specific forms of training, thereby establishing recognition and retention strategies. On the whole, the clustering method gives practically usable information that would work to streamline performance and moral of the workers.

Table 1: Employee cluster summary based on productivity and satisfaction

Cluster	% of Employees	Avg Productivity	Avg Satisfaction	Key Characteristics
1	35%	High	High	Stable top performers
2	40%	Medium	Low	At risk, needs support

The Table 1 is a summary of grouping the employees in three aggregates on the basis of productivity and satisfaction. Cluster 1 (35%) is a high performing, very satisfied employee and is perfect to implement retention and recognition programs. Cluster 2 (40%) has medium productivity and low satisfaction which implies that they need some support that can be mentoring, role changes, or engagement programs. Cluster 3 (25%) contains the low-satisfaction and low-productivity employees who are possibly new, or disengaged workforces and who needs some kind of induction and encourenagement. Such division allows the HR to deploy specific strategies depending on the need of each group to enhance the effectiveness of the workforce as a whole.

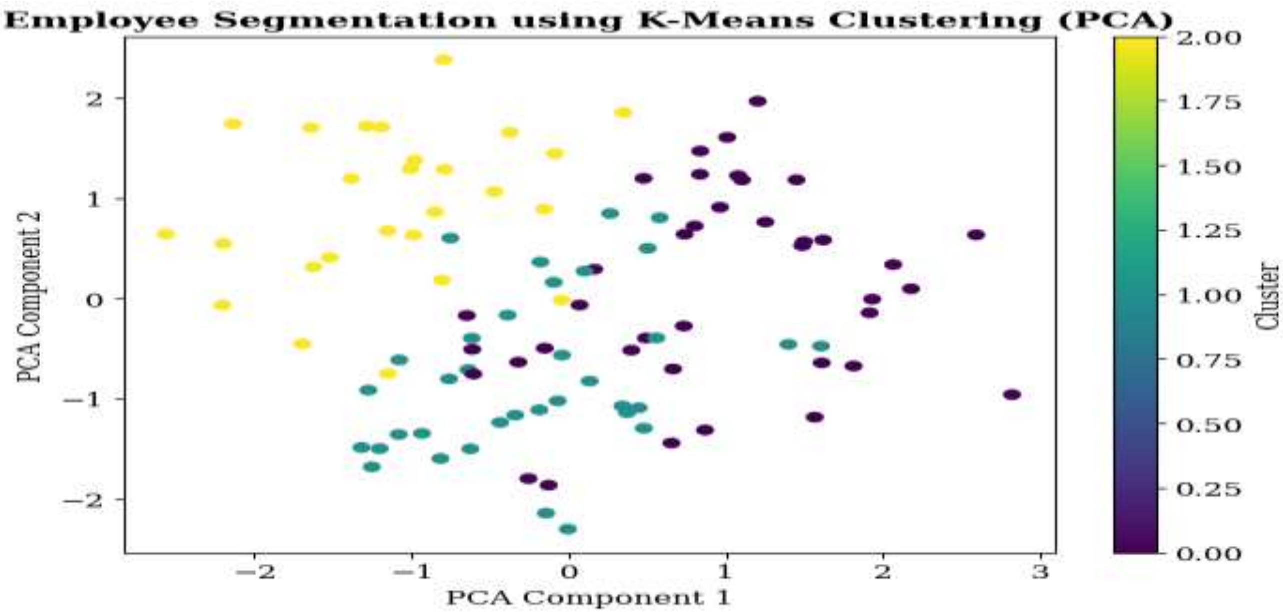

Figure 2: PCA-Based visualization of employee clusters from K-means algorithm

This Figure 2 shows the findings of employee segmentation based on the K-Means algorithm of clustering, that happens to be visualized by means of Principal Component Analysis. The gray dots denote each employee and the color is what determines the cluster that the employees are in. PCA to have an easier visualization and retain as much variance as possible.

Due to the clear grouping of points it is possible to state that the clustering algorithm has managed to identify different employee profiles based on tendencies in productivity, satisfaction, feedback and other criteria. This can be used to enhance an insight on employee grouping and it can be used to build specific intervention on each group of employees.

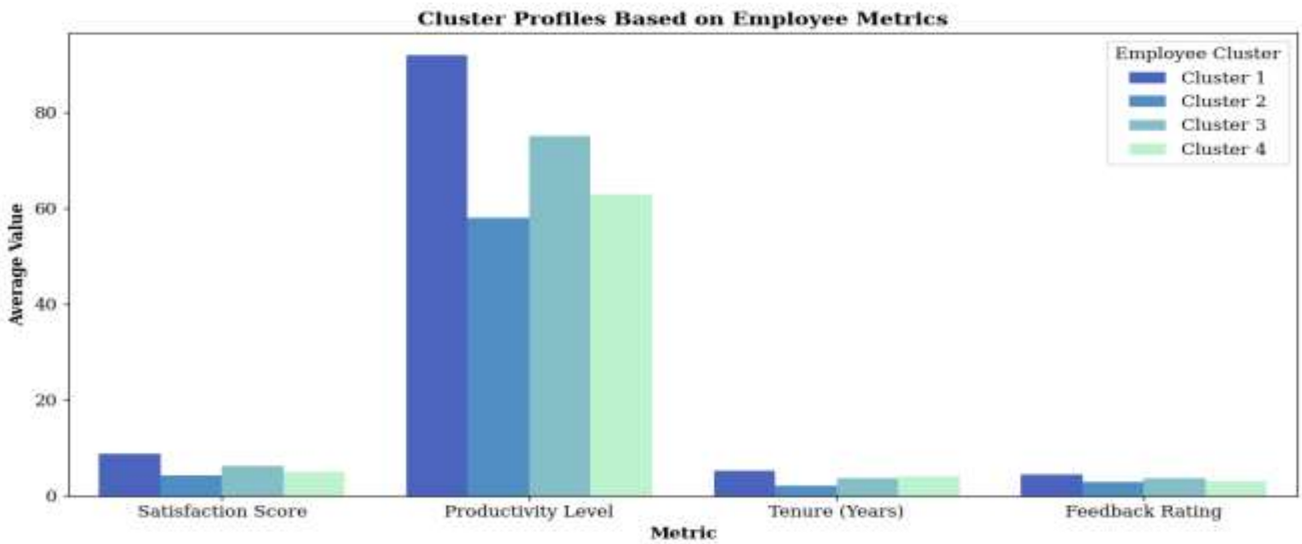

Figure 3: Cluster profiles based on employee metrics

The graph in Figure 3 is the average results of an influential HR measurement in four employee groups that are produced by the K-Means clustering procedure. The measures are Satisfaction Score, Productivity Level, Tenure and Feedback Rating. Cluster 1 demonstrates the most numbers across the board categories signifying a group of achieved, experienced and happy employees. Cluster 2 on the other hand, has the lowest figures, and this may be a group of disengaged or underperforming employees. The rest of the clusters lie somewhere between these ends as they are the mediocre groups. These can inform HR approaches on the individual level including specific training or retention programs.

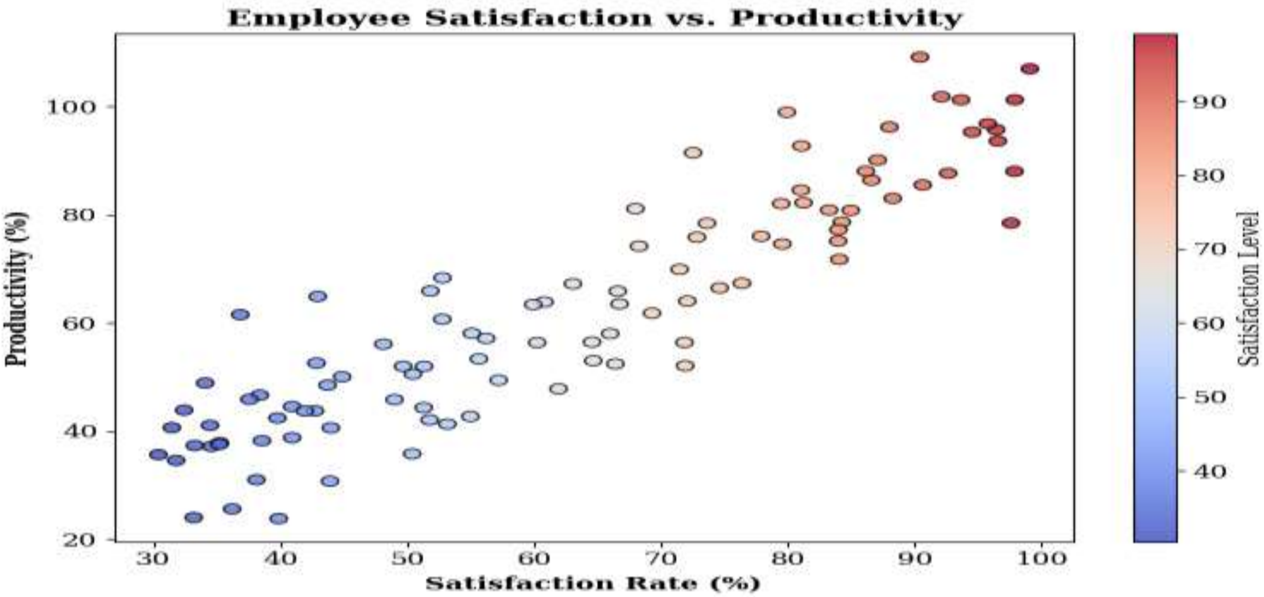

Figure 4: Employee satisfaction vs. productivity scatter plot

This scatter plot visualizes the association amid employee satisfaction rate and productivity. Each opinion signifies an individual employee, with color intensity reflecting their satisfaction level from low to high. The upward trend indicates a positive correlation, suggesting that employees who report higher satisfaction tend to be more productive. This finding supports the importance of workplace engagement and morale in driving performance. HR teams can use such insights to focus on satisfaction-improving strategies to enhance productivity organization-wide.

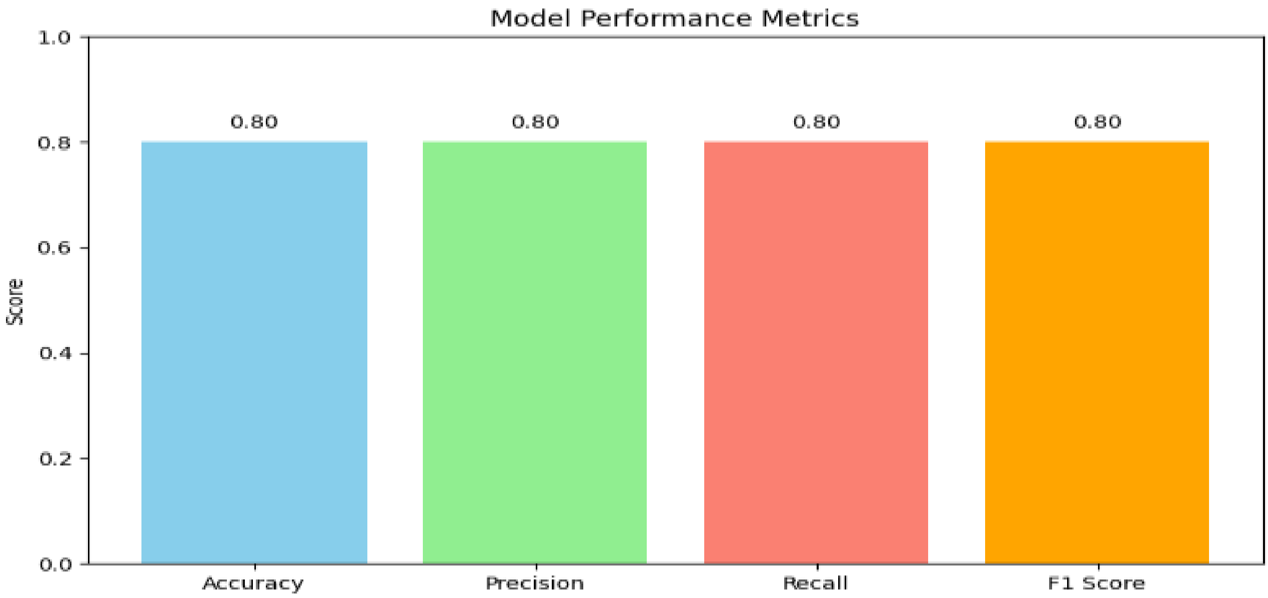

Figure 5: Bar Chart representing model performance metrics

Figure 1 shows a bar chart showing four key performance measures-Accuracy, Precision, Recall and F1 Score each measuring the performance of the classification model. All measures record a score of 0.80, which suggests a consistent and well-balanced performance. Accuracy captures the overall accuracy in the predictions. Precision shows the number of predicted positive values that actually were positive, and recall represents the ability of the model to find actual positives. F1 Score the harmonic mean of precision and recall validates the model's strength, particularly in class imbalance. The equal metric heights imply that the model has a good balance between false positives and false negatives.The findings demonstrate how AI-facilitated clustering used to analyse employee data, reveals latent patterns and facilitates evidence-based HR approaches.

CONCLUSIONS

The implementation of a data-driven HRM framework, particularly through the use of

K-Means clustering, proves to be an effective tool for enhancing employee segmentation and optimizing performance. The insights derived from clustering provide HR professionals with actionable data that can guide the development of personalized engagement strategies, training programs, and retention policies. By focusing on the specific needs of different employee segments, organizations can stand-in a more productive and satisfied staff. However, it is important to speech challenges such as data quality and the integration of insights into existing HR practices to fully leverage the benefits of machine learning in HRM. The findings highlight the growing role of analytics in transforming HR functions into more proactive and predictive systems, driving strategic workforce management. This study confirms that the incorporation of AI in HRM through clustering techniques such as improves employee segmentation and strategic decision-making. By moving from reactive to predictive HR measures AI facilitates improved workforce planning, performance improvement, and employee engagement.

REFERENCES

1. Maqsood M, Yasmin S, Mehmood I, Bukhari M & Kim M (2021), “An efficient DA-net architecture for lung nodule segmentation”, Mathematics, Vol. 9 No. 13, pp.1457.
2. Tang T, Chen S, Zhao M, Huang W & Luo J (2019), “Very large-scale data classification based on K-means clustering and multi-kernel SVM. Soft Computing”, Vol. 23, pp. 3793-3801.
3. Wang Q, Wang Y, Niu R & Peng L (2017), “Integration of information theory, K-means cluster analysis and the logistic regression model for landslide susceptibility mapping in the Three Gorges Area, China”, Remote Sensing, Vol. 9 No.9, pp. 938.
4. Lismont J, Vanthienen J, Baesens B & Lemahieu W (2017), “Defining analytics maturity indicators: A survey approach”, International Journal of Information Management, Vol. 37 No.3, pp. 114-124.
5. Azadeh A, Heydarian D, Nemati K & Yazdanparast R. (2018), “Performance optimization of unique resilient human resource management system in a coal mine industry”, International Journal of System Assurance Engineering and Management, Vol. 9, pp. 1178-1197.
6. Lopez J, Branch JW & Chen G (2019), “Line-based image segmentation method: A new approach to segment VHSR remote sensing images automatically”, European Journal of Remote Sensing, Vol. 52 No.1, pp. 613-631.
7. Torre Martínez MRDL, Calvet L, Juan AA, Hatami S & López López D (2021), “Business analytics in sport talent acquisition”, International Journal of Business Analytics, Vol. 9(1, article 6), pp. 1-20.
8. Cheng X, Fang L, Yang L & Cui S (2017), “Mobile Big Data: The fuel for data-driven wireless”, IEEE Internet of things Journal, Vol. 4 No. 5, pp.1489-1516.
9. Kaufman BE, Barry M, Wilkinson A & Gomez R (2021), “Alternative balanced scorecards built from paradigm models in strategic HRM and employment/industrial relations and used to measure the state of employment relations and HR system performance across US workplaces”, Human Resource Management Journal, Vol. 31 No.1, pp. 65-92.
10. Qiao Y, Jiao L, Tang X, Li W & Cosker D (2020), “High-quality depth up-sampling via a supervised classification guided MRF model”, Pattern Recognition Letters, Vol. 138, pp.315-320.
11. Afzalan M, Jazizadeh F & Eldardiry H (2021), “Two-stage clustering of household electricity load shapes for improved temporal pattern representation”, IEEE Access, Vol. 9, pp. 151667-151680.
12. Wang L & Cotton R (2018), “Beyond Moneyball to social capital inside and out: The value of differentiated workforce experience ties to performance”, Human Resource Management, Vol. 57 No.3, pp.761-780.
13. Wong KK, Fortino G & Abbott D (2020), “Deep learning-based cardiovascular image diagnosis: A promising challenge”, Future Generation Computer Systems, Vol. 110, pp. 802-811.
14. Richardson N (2021), “Emergency response planning: Leveraging machine learning for real-time decision-making”, Emergency, Vol. 4, pp. 14.

15. Parimi SSR (2019), “Investigating how sap solutions assist in workforce management, scheduling and human resources in healthcare institutions”, IEJRD- International Multidisciplinary Journal, Vol. 4 No.6, pp.10.
16. Lincaru C, Pîrciog S, Atanasiu D, Stroe C, Ciucă V & Grigorescu A (2020), “Patterns of mainly tourism sectors at local level by employee’s characteristics using gis multivariate clustering analysis-Romania case study”, Regional Science Inquiry, Vol. 12 No.1, pp.261-290.
17. Ryait H, Bermudez-Contreras E, Harvey M, Faraji J, Mirza Agha B, Gomez-Palacio Schjetnan A, Luczak A (2019), “Data-driven analyses of motor impairments in animal models of neurological disorders”, PLoS biology, Vol. 17 No.11, pp. e3000516.
18. Moradi S & Mokhatab Rafiei F (2019), “A dynamic credit risk assessment model with data mining techniques: evidence from Iranian banks”, Financial Innovation, Vol. 5 No.1, pp.1-27.
19. Kuo YH & Kusiak A (2019), “From data to big data in production research: The past and future trends”, International Journal of Production Research, Vol. 57 No.15-16, pp.4828-4853.
20. Kanda E, Epureanu BI, Adachi T, Tsuruta Y, Kikuchi K, Kashihara N, & Nitta K (2020), “Application of explainable ensemble artificial intelligence model to categorization of hemodialysis-patient and treatment using nationwide-real-world data in Japan”, Plos one, Vol. 15 No.5, pp. e0233491.
21. Kalusivalingam AK, Sharma A, Patel N & Singh V (2020), “Optimizing workforce planning with AI: leveraging machine learning algorithms and predictive analytics for enhanced decision-making”, International Journal of AI and ML, 1(3).
22. Hernández-Julio YF, Prieto-Guevara MJ, Nieto-Bernal W, Meriño-Fuentes I, Guerrero-Avendaño A (2019), “Framework for the development of data-driven Mamdani-type fuzzy clinical decision support systems”, Diagnostics, Vol. 9 No.2, pp.52.
23. Kaufman BE (2015), “Market competition, HRM, and firm performance: The conventional paradigm critiqued and reformulated”, Human Resource Management Review, Vol. 25 No.1, pp. 107-125.
24. Simón C & Ferreiro E (2018), “Workforce analytics: A case study of scholar–practitioner collaboration”, Human Resource Management, Vol. 57 No.3, pp. 781-793.
25. Adekunle BI, Chukwuma-Eke EC, Balogun ED & Ogunsola KO (2021), “Machine learning for automation: Developing data-driven solutions for process optimization and accuracy improvement”, Machine Learning, 2(1).
26. Tirkolaee EB, Sadeghi S, Mooseloo FM, Vandchali HR & Aeini S (2021), “Application of machine learning in supply chain management: A comprehensive overview of the main areas”, Mathematical problems in engineering, Vol. 1, pp. 1476043.
27. Boyapati S. (2021), “Bridging the urban-rural divide: A data-driven analysis of internet inclusive finance in the e-commerce era”, Int. J. Eng, Vol.11 No.1.
28. Li J, Hassan D, Brewer S & Sitzenfrei R (2020), “Is clustering time-series water depth useful? An exploratory study for flooding detection in urban drainage systems”, Water, Vol. 12 No. 9, pp. 2433.
29. Bergen KJ, Johnson PA, de Hoop MV & Beroza GC (2019), “Machine learning for data-driven discovery in solid Earth geoscience”, Science, Vol. 363 No. 6433, pp.eaau0323.
30. Banaee H, & Loutfi A (2015), “Data-driven rule mining and representation of temporal patterns in physiological sensor data”, IEEE Journal of Biomedical and Health Informatics, Vol. 19 No. 5, pp. 1557-1566.
31. Dataset Link: https://www.kaggle.com/datasets/adityaab1407/employee-productivity-and-satisfaction-hr-data/data