Intelligent Profiling for Human Capital Utilization

Jay B.Simha[1,2]
[1]Abiba Systems
[2]RACE,
REVA University
Bengaluru, India
Jay.b.simha@abibasystems.com

Shinu Abhi
RACE,
REVA University
Bengaluru, India
shinuabhi@race.reva.edu.in

Abstract— It has been observed in recent research that Human Capital Management (HCM) is an important function that is getting much attention from analytics. One of the touch points in the HCM life cycle is the people development with effective recommendations and timely interventions. Several models cited in research use models built with both supervised and unsupervised approaches. In this study, a framework based on self-organizing neural network model and linguistic modeling using fuzzy logic is proposed for utilization and intervention. This hybrid approach enables building rules/functions for different groups of human resources separately. In the first stage, employees are segmented into clusters, that are characterized by similar features and then, in the second step, for each group, fuzzy logic is used to obtain rules that may provide profile for each segment. The main advantage of applying the integration of two techniques consists of building models that, may better profile and predict the human capital requirements better, than using each method separately. The results are compared with the results of the work available in literature. The results indicate that the proposed approach provides an alternative view of the insights.

Key words: Self-organizing Maps, Fuzzy Logic, Linguistic Modeling, Human Capital, Utilization

I. INTRODUCTION

HR as a function has moved on from being a cost function to a strategic partner contributing to the bottom line profitability of organizations. HR leaders now focus on understanding how organizations can enhance their human capital by leveraging data pertaining to their employees. This could range from talent identification, improving new hire quality, identifying and developing potential leaders, predicting turnover of employees, capturing the workforce demographics and dynamics, objective assessments and much more.

Applied properly, HR analytics can show connections, correlations and even causality between HR metrics and other business measures – all of which can be used to design effective HR strategies. In other words, HR analytics can provide a tangible link between people strategy and the organization's performance. Today, all the organizations treat employees as resources and hence aligning these resources to business goals through the right analytical techniques is essential for the long term sustenance and growth [1].

The human capital or more appropriately, human resources should be optimally utilized to create maximum economic gains, so that employee growth will result in better business performance. This requires an approach to identify, profile and deploy the right resources for the right tasks, in turn improving utilization. A better utilized human capital will result in the overall value gain and the sub-optimal utilization will lead to employee attrition and dissatisfaction. In this work, a frame work is proposed to profile the competence and the tasks assigned through "Self Organizing Maps" (SOM) and linguistic modeling using fuzzy logic, which is effective in making the neural network model to be better understood with linguistic models.

II. LITERATURE SURVEY

Human resources are always at the very core of the organization's success. Aligning the right human resources for various roles and tasks are essential for superior workforce performance [1].

The recent research studies indicate that the Human Resource/Talent analytics is to be adopted for effective utilization of the human resources and transformation of the current practices towards improved productivity [2]. It is estimated that the HR analytics function is not widely adopted due to several reasons and is likely to be a separate discipline by 2025. The author also suggests the need for value addition from HR analytics for organizational positioning [3].

One of the earliest versions of analytics is metric based analytics using Key Performance Indicators (KPI). However, use of advanced analytics in HR/HCM stands at a mere 15% compared to metric-based analytics which is at 61% [5]. Even advanced software suites do not have employee life cycle models using machine learning or other advanced analytics tools [6]. This gap is due to the non-availability of the right data, maturity of the organization in using the insights from advanced analytics and non-availability of the relevant use cases [7]. In this work a use case for utilization of the human capital has been proposed to recommend effective allocation of the tasks for better utilization and profitability.

This paper proposes a framework using Self Organizing Maps (SOM) and linguistic modeling is proposed to get insights into the utilization data and recommend effective utilization plans. The SOM is used to generate the homogeneous segments from the data and the linguistic modeling using fuzzy logic is used to build insights on the derived segments from SOM. A brief description of each of the algorithms is given in next few paragraphs.

III. SELF ORGANIZING MAPS

"Self-Organizing Maps" (SOM) imitates the function of 'grouping by categories' as done by human brain and every output processing element affecting each other. It includes a set of neurons usually arranged in a two-dimensional structure, in such a way that there is a neighborhood relationship among the neurons, which dictates the topology, or structure. The neurons are well connected to each other from input to output layers, but they are not connected to themselves [9].

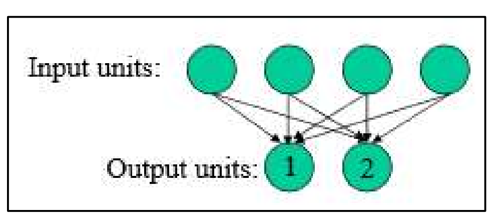

Figure 1: Self Organizing Map

SOM uses unsupervised learning which is known as self-organization to visualize topologies and hierarchical structures of high dimensional input spaces. The algorithm of SOM is initialized by assigning the values of weight vectors of each output neuron linearly or randomly. Training process of SOM starts by representing a data point randomly in the network. The distances between these data points and the weight vectors of all neurons are computed by using distance measures such as Euclidean distance. The nearest neuron wins and is thus updated to move closer to the data point [9].

IV. LINGUISTIC MODELING

This area of fuzzy logic modelling attempts to more accurately mimic the action of human reasoning. It makes more sense to use typical linguistic expressions, for example, less, medium and more, as variables of the linguistic model, over strict quantitative numerical values. A linguistic model is a model that is described using linguistic terms in the framework of fuzzy logic, instead of mathematical equations with numerical values, or conventional logical formulae with logical symbols [10, 11, 13]. Linguistic modelling utilizes the notions of normal commonplace language to label the fuzzy sets which represent quantitative variables. Another important component of any linguistic model is the set of conditional statements, or rule base. The linguistic variables have and interdependent relationship with each other to produce the result of the decision. This is represented in an IF-THEN statement structure. For example,

"IF Experience is Medium and Complexity is Low THEN Utilization is Poor".

Typically, the Linguistic model with linguistic variables will complement the complex statistical and neural models with its white box approach, which provides. There are multiple models to derive the linguistic models from data. In this work, the method proposed by [4], will be used to derive the linguistic rule for the model from the clusters formed by the SOM from previous section.

The universe of domain is defined as the relation:
$L^i = X^j, LV^k$
Where,
L is the linguistic domain for the variable X^i
X^j is the variable to be mapped for insights
LV^k is the domain of linguistic values X^j can take

The problem is to map the discrete or continuous values of X^j to LV^k. This is achieved by converting the values of X to LV through a membership function/ In this research, S, Z, triangular and trapezoidal membership functions will be mapped. The details of the mapping is discussed in next section.

V. EXPERIMENTAL RESULTS AND DISCUSSION

Experiments were done on a benchmark data set provided in the literature [8]. The data consists of the derived variables for performance measurement. The profiles of the each of the variables are shown in fig 2-5.

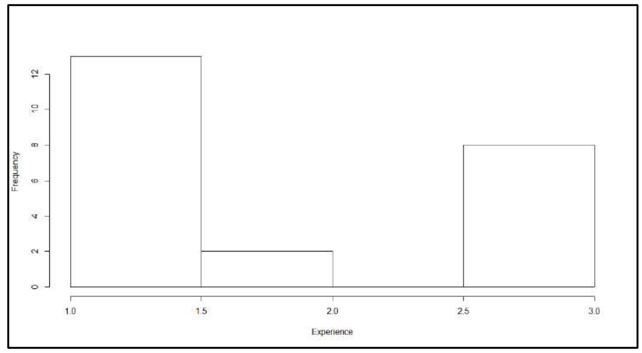

Fig 2. Experience

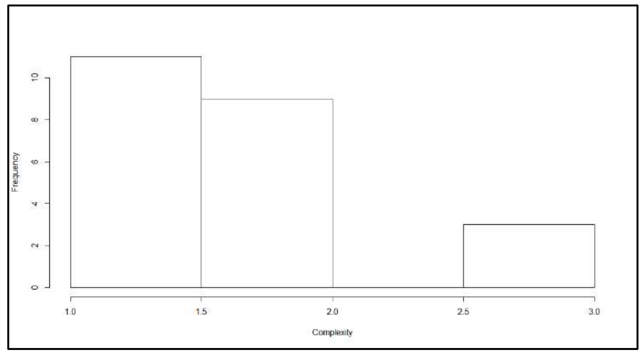

Fig 3. Complexity

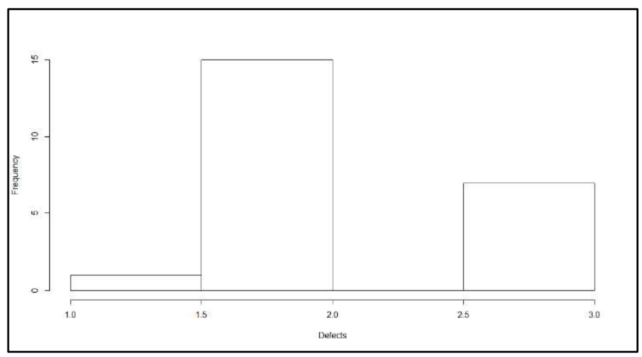

Fig 4. Defects

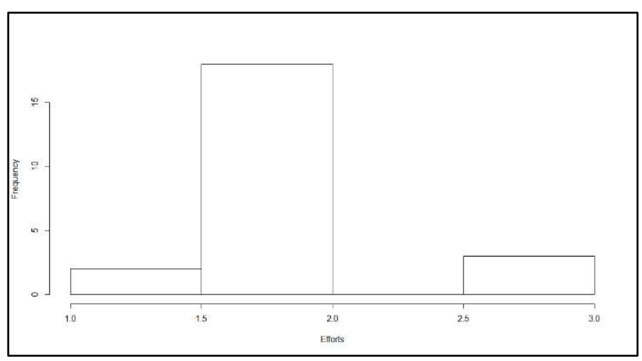

Fig 5. Efforts

The guidelines in [8] has been used in this work to map the performance variables to a linguistic domain. The following mapping were obtained after fuzzification of the variable ranges.

Experience = [Low, Medium, High]
Complexity = [Low, Medium, High]

Defects = [Low, Medium, High]
Efforts = [small, Medium, Large]

Once the mapping is done, the data has been transformed and the discovery process using cross tabs has yielded the following results fig 6-7.

Experience	Complexity High	Low	Medium
High	0	7	1
Low	3	2	8
Medium	0	2	0

Fig 6. Discovery

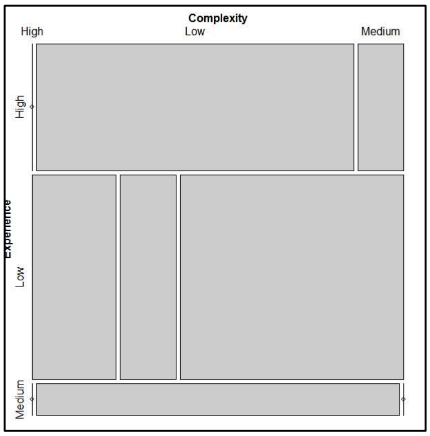

Fig 7. Experience vs Complexity of problems solved

Surprisingly, the low complexity work has been taken by the highly experienced team, which may lead to underutilization of the human resources. In addition, the complex defects resolved by junior teams demonstrate either the capability of the junior team or the latent introduction of additional bugs. In order to get insights into the full data, a segmentation is designed to be carried out on data using SOM algorithm.

The experiments consisted of creating segments using SOM. In the current work a SOM grid of 1X3, 1X5 and 3X3 are created as discussed in [8] for comparison. The results of the segmentation are given in fig 8-10.

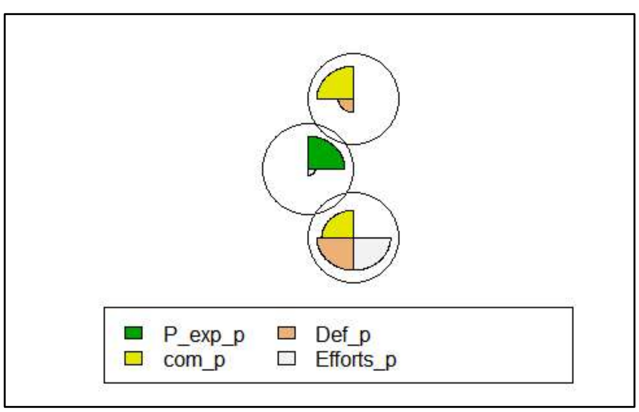

Fig 8. Three segments model

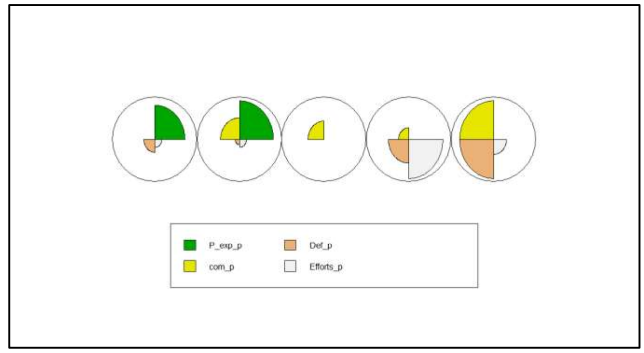

Fig 9. Five segments model

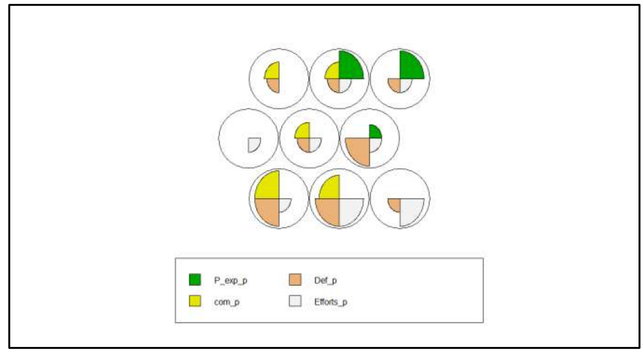

Fig 10. Nine cluster model

The deeper analysis of the critical mass of the segments has yielded a three cluster model to be better from an actionability perspective for improving the utilization.

The three cluster solution has provided additional insights to the rules derived from contingency table. The average values of different parameters for each of the segment are shown in table 1. It can be observed that the low experience resources have been working on higher complexity defects and high experience resources are doing low complexity defects.

Table 1. Profiles of the segments from best model

Segment	Experience	Complexity	Defects	Efforts
1	1.000000	2.000000	2.666667	3.000000
2	2.636364	1.090909	2.090909	2.000000
3	1.000000	2.222222	2.333333	1.777778

Surprisingly, less experienced resources are resolving high complexity defects with lesser effects than the highly experienced resources. This is an issue for further investigation by the concerned managers.

Once the segments are derived from the data, the profiles are extracted with linguistic variables defined over the universe of discourse. The profiles resulted in SOM/cross tab, are shown in table 2.

Table 2. Discovered profiles

Experience	Complexity	Effort	Defects
Low	Medium	Any	Any
Low	High	Any	Any
High	Low	Any	Any

The Fuzzy Knowledge Base using Fuzzy Associative Memory (FAM) or fuzzy rules are used as the base for recommendation. This is derived as intransigent domain expectation and is derived from domain experts. A sample of the fuzzy rules used for recommendation in this work are shown in table 3.

Table 2. Fuzzy Rules (FAM) table

Experience	Complexity	Effort	Utilization
High	High	Low	High
Medium	High	Medium	High
Low	High	High	Medium
High	Low	Low	Low

The discovered profiles are compared with expected performance given by a FAM table and recommendations are derived.

The segments discovered from SOM are analyzed using FAM table for further insights. It has been observed that the discrepancies in the expected performance are clear in the segments. The visual clues are further profiled using the linguistic variables as discussed in the previous section.

The linguistic profiles were derived for the utilization. It has been observed that the less experienced resources are solving more complex problems and experienced resources are solving less complex problems, which shows the inefficiency in the system. This information can be used to recommend suitable utilization of the resources, as observed in the literature [12].

Conclusion

In the paper, a framework for analyzing the resource utilization using SOM and linguistic modeling is proposed. The presented framework allows for building different profiles for different segments of resources, which provide the best results for that segment. In the proposed approach, each resource is assigned to the most similar group of resources from the training data set and utilization is evaluated by applying the linguistic model proper for this group. The results obtained on the real data set show simplicity of models obtained for each cluster than for model developed with the whole data set.

References

[1] Momin, Weena Yancey M. "HR analytics transforming human resource management." IJAR 1, no. 9 (2015): 688-692.
[2] Angrave, David, Andy Charlwood, Ian Kirkpatrick, Mark Lawrence, and Mark Stuart. "HR and analytics: why HR is set to fail the big data challenge." Human Resource Management Journal 26, no. 1 (2016): 1-11.
[3] Van den Heuvel, Bondarouk. "HR analytics: A study into the current state of HR analytics and predictions for its future", Human Resource Management International Digest, Vol. 25 Issue: 7, pp.9-11, (2017)
[4] Dua S, Iyengar S.S., Bharatheesh T.L., “A new method for information discovery in plant databases”, International journal of information science, August 2000
[5] Martin A, “What is happening with HCM Business Intelligence and Analytics Technologies Adoption?”, Workforce Solutions Review, April/May.2013, 2013
[6] Prebuilt analytics for HCM, https://www.doag.org/formes/pubfiles/7055339/2015-BSK-BIG-Eva_Michel-OTBI-E_-_fertige_Analytics_fuer_HCM___CX_aus_der_Oracle_Cloud-Praesentation.pdf, Jun.2015, 2015.
[7] Roberts G, “Using Predictive Workforce Analytics to Drive Business Performance”, PCI HR Conference, April 23-25, 2017
[8] https://www.shrm.org/foundation/ourwork/initiatives/preparing-for-future-hr-trends/Documents/Workforce%20Analytics%20Report.pdf
[9] Kohonen, Teuvo. "The self-organizing map." Proceedings of the IEEE 78, no. 9 (1990): 1464-1480.
[10] Sugeno, Michio, and Takahiro Yasukawa. "A fuzzy-logic-based approach to qualitative modeling." IEEE Transactions on fuzzy systems 1, no. 1 (1993): 7.
[11] Zadeh, Lotfi A. "Outline of a new approach to the analysis of complex systems and decision processes." IEEE Transactions on systems, Man, and Cybernetics 1 (1973): 28-44.
[12] Ajmani S, Prakash K.S.S, and Prasad D.S.U.M, “Human Resource Utilization Analysis and Recommendations Using Software Analytics” , Lecture Notes on Software Engineering, Vol. 4, No. 3, August 2016
[13] Zadeh, L.A. “Outline of a New Approach to the Analysis of Complex Systems and the Decision Process.”, IEEE Trans. Sys Man & Cyber., Vol.SMC-3, No.1 , Jan.1973, 1973