Efficacy of Internal Marketing Practice as a part of HRM in Logistic Firm" A Case Study of Leading Logistics Company.

Mukta Gupta[1], Prof.(Dr.) Pradeep K. Varshney [2]

[1]Research Scholar, [2]Professor

Department of Management and Commerce

Mewar University Chittorgarh, Rajasthan, India

[3]Dr.Lokesh Jindal,

Associate Professor, Atal bihari Vajpayee School of Management and Entrepreneurship, Jawaharlal Nehru University.

Abstract. Human Resource Management (HRM) is the function within an organization that focuses on the recruitment of, management of, and providing direction and guidance for the people who work in an organization. Employees are important clients, so don't neglect them. In this respect, foresee corporate rebate programs, special offers, and even samples which will all contribute to increasing your personnel's sense of belonging, all while allowing them to benefit from the fruit of their labour. In current scenario, every successful companies attach great importance to human resource management and internal marketing; because they are aware of the value of those activities and of strategic advantage they can bring to the organization. Enhancing the role of the employee within an organisation Internal Marketing (IM) is a new and emerging discipline that is practical in multiple companies. The purpose of this study is to see the impact of internal marketing on employee satisfaction. Data is collected through a on various parameters of employee satisfaction and internal marketing. Since Machine Learning (ML) is on the verge of seriously impacting workplace communications, therefore, a comparative analysis of various ML algorithms are done to predict the employee satisfaction with precision accuracy and recall. .Finally, the models are ranked using Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). A case of well-known logistic firm, Om Logistics is considered to validate ML models and its accuracy.

Keywords:, Internal Marketing, Employee Satisfaction, Machine Learning, Random Forest, Support Vector, Binary Logistics, K-fold Cross Validation, TOPSIS

1 Introduction

Human Resource Management (HRM) is the function within an organization that focuses on the recruitment of, management of, and providing direction and guidance for the people who work in an organization. As you can imagine, all of the processes and programs that are touched by people are part of the HR kingdom. In this case, the employee represents the client, and all the noble aspects that being a client entails. Hence, internal marketing translates into a real willingness to adapt working conditions to the employees' needs, to offer them a "product" that corresponds to their needs and to their expectations. It is a way to ensure "customer loyalty", or in other words, to engage employees and to increase your retention rate. For some, the internal marketing approach may seem like a vicious cycle, an illogical reversal of the workplace. It is quite the opposite! Remember that is a *virtuous* cycle – a set of procedures that improve a situation. No company can say no to that!

In reality, internal marketing is a skillful way to apply your business strategy in-house. Make sure to establish a constant dialogue with your employees in order to better understand them and to offer them working conditions that meet their needs. You have been warned: when marketing and HR management join forces, you get **results**!The concept of internal marketing is a tool that companies use within their workforce to communicate with their employees. Many company owners and authors of internal marketing believe this concept is as important to a company's survival as external marketing (communicating to customers). When communicating to employees this involves the communications of "corporate culture and goals, mission and vision statements, as well as personnel policies and procedures *"Internal marketing is orienting a motivating customer contact employees and supporting service people to work as a team to provide customer satisfaction.(Kotler and Armstrong 2010- add in ref section).*It has long been said that an organisation's most important asset is its people. Employees define a company, shape its culture, and often prove to be its most valuable source of inspiration, ideas, and 'shop floor' insights

In recent years, this mindset has changed. Now, in the era of digital transformation, it is believed an organisation's biggest asset is its human capital and data is the key to protect organisations' most valuable asset. People of organisation will be the driver of businesses, but this changing era of automation, robotics and artificial intelligence will make people more powerful to make future inferences. Role of the human resource team is changing as the ability to gather and analyze ever-increasing amounts of data grows, so are the opportunities for people to add more value to the organization. This is what makes human resource data such an important asset. Data is what shapes products and services, improves customer experiences, and ultimately defines brands. Over the last 12 months, machine learning and artificial intelligence has dominated much of the conversation around how technology and data will shape the way people live and work in the years to come.

Arthur Samuel, a pioneer in the field of artificial intelligence and computer gaming, coined the term **"Machine Learning"**. He defined machine learning as – **"Field of study that gives computers the capability to learn without being explicitly programmed".** Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inferences. Machine learning is on the verge of seriously impacting workplace communications. Also enhancing the role of the employee within an organisation. As an example, consider the potential for a sales representative or contact centre employee. By using machine learning patterns, a sales representative will know when and where to tailor their communications approach—knowing how much they should talk versus how much they should listen. The process starts with feeding quality data and then training our machines (computers) by building machine learning models using the data and different algorithms. The choice of algorithms depends on what type of data do we have and what kind of task we are trying to automate. Today, companies are using Machine Learning to improve business decisions, increase productivity, detect disease, forecast weather, and many more things. In this work, we explore how employee job satisfaction is dependent on various parameters of internal marketing like recognition, appraisal, promotion, salary etc. The next section describes general Machine Learning approach to solve problems.

2 Case Discussion

Leading Logistics Company, India the flagship company of company is an instrument in providing innovative and value added solutions for Indian Corporate and Multinationals. It is the only leading Multi-modal logistics company with single window integrated logistics services for all the element of the supply chain management in India. Leading Logistics Company is committed to offer logistics solutions and services to the Indian Industry designed to customer's satisfaction in order to help, support and grow together in a win-win situation. To redefine customer satisfaction, herald a rationalized approach of time & cost and eventually bridge the gap between Indian requirements and global standards. Logistics Company, with its flexible set-up of transportation, offers manufacturers a number of opportunities to improve their supply chain radius across India. This can easily be regarded as one of the most significant strengths of Leading Logistics Company, as manufacturers have a real chance of saving their capital investment by way of reducing the overall cost of inventory. Logistics Company connects over a thousand destinations across the country. In addition to that, the company also owns a dedicated fleet when it comes to local distribution. According to the study, the competitors of the company have a high turnover ratio. This affects their goodwill because a high turnover comes with a number of detrimental effects for an organization. For instance, it becomes challenging to replace specific employees who are business domain experts or possess niche skill sets. The research aims to explore all the reasons behind this data. The issue of employee turnover is quite prominent in organizations in the present times due to its negative impacts on a number of problems ranging from productivity to workplace morale. It even causes disruptions in long-term growth strategies and project continuity.

It is noteworthy that employee retention has also become a significant concern for most corporates. Once employees are trained, they tend to move on to other companies for better prospects. Comfortable timings, lucrative salary, scope of growth and better ambience are some factors that prompt individuals to seek a change. When talented employees are willing to move on, the HR team and management is responsible for intervening and finding out the precise reasons that led to this decision. The management team of Leading Logistics Company comprises industry experts and savvy entrepreneurs, bringing years of experience with them to the company. Moreover, they work tirelessly and cultivate a strong culture to assist the team in growing and succeeding. The core team members also offer their vision and leadership to the other team members.

Research shows that the turnover ratio of the company is very low and employee retention is very high. Hence, this study aims to analyze how employers of the company motivate the employees to make them comfortable, leading to enhanced employee retention. Since the study is focusing on a specialised group of employees, the population is relatively small and validity and reliability issues could result in challenges further reducing the sample size. Since all employees employed with leading Logistics Company are situated in their head office, a person-administered survey was selected. This would allow the researcher to conduct the survey personally with employees to ensure that all questions and instructions were clearly understood and answered in full.

2.1   Machine Learning Approach

Machine learning algorithms are often categorized as supervised or unsupervised learning algorithms. **Supervised machine learning algorithms** talks about what has been learned in the past to new data using labelled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. In contrast, **unsupervised machine learning algorithms** are used when the information used to train is neither classified nor labelled. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabelled data.There are two more approaches used in some cases. These are s**emi-supervised machine learning algorithms and Reinforcement learning.** The former fall somewhere in between supervised and unsupervised learning, since it uses both labelled and unlabeled data for training – typically a small amount of labeled data and a large amount of unlabeled data. **Reinforcement machine learning algorithms** is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning.

Some of the popular ML models are Linear Regression, Logistic Regression, Decision **Tree**. Support Vector Machine, Naive Bayes, Neural Networks,K-Means,Random Forest etc. In this paper, we have used supervised machine learning techniques for predicting employee satisfaction given his age, salary and his rating of internal marketing parameters reward, promotion, appraisal

and recognition. The supervised learning models work by taking a set of labeled examples (training set) and discovering a function f that predicts the class for a new example (test set). The implementations of these models are available in R. R is a free and open source software licensed under GNU GPL. A brief of models is presented below subsections.

2.1.1 Binary Logistic Regression

Binary Logistic Regression is a classification algorithm which is used when we want to predict a categorical dependent variable. The response variable here has binary responses or of dichotomous nature like yes or no, pass or fail, male or female, defective or non-defective etc that can be labelled as 0 or 1.The independent variables can each be either binary or continuous variables. It is a basic regression that uses logistic function or sigmoid function. If Y be a binary response variable and $X = (X_1, X_2, ..., X_k)$ be a set of explanatory variables which can be discrete, continuous, or a combination. x_i is the observed value of the explanatory variables for observation i.

$$\Pr(Y_i = 1 \mid X_i = x_i) = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}$$

or

$$\log it(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right)$$

Distribution of Y_i is $Bin(n_i, \pi_i)$, i.e., binary logistic regression model assumes binomial distribution of the response.

2.1.2 Decision Tree

Decision trees are one of the popular supervised machine learning algorithms. They have been widely used for classification and regression problems. Decision trees work very similar to the way the human brain thinks by dividing the problem into subproblems which makes them easy to understand. They work in a recursive manner by partitioning the dataset into smaller datasets on the basis of feature values until the final output is obtained. At every step, the aim is to partition the set in best possible manner i.e. that gives maximum amount of information (or reduces the level of uncertainty). A CART (Classification and Regression Trees) algorithm [D3] is used to build the decision tree model by recursively splitting the dataset around a feature that gives maximum reduction in the heterogeneity of the response variable. At every step, the heterogeneity of a node is determined using Gini Index.

$$Gini\ Index = p_{(i)}\ (1-p(i))\ for\ i = 1\ to\ K,$$

where K is the number of classes of response variable and p(i) is the probability of class i at a node.

2.1.3   Random Forest

Random forest is a popular machine learning model that works well for both classification and regression problems. It is an ensemble method proposed by Breiman [3]. During the learning phase, the algorithm works by bootstrapping multiple samples from the training data set and then creating independent decision trees for each of the bootstrapped sample. While constructing a decision tree, the algorithm randomly selects a subset of features/variables and finds out the best feature for splitting the node in that subset.

2.1.4 Support Vector Machine

Support Vector Machine takes numeric input and binary output. It is based on finding a linear plane with maximum margin to separate two classes of output. Categorical input can be turned into numeric input as before and categorical output can be modeled as multiple binary output. With a different loss function, SVM can also do regression (called SVR). I haven't used this myself so I can't talk much.The strength of SVM is it can handle large number of dimensions. With the kernel function, it can handle non-linear relationships as well.

2.2   Model Evaluation

The most important part of a classification prediction study is to evaluate the performance of the classifier. There are numerous metrics available for evaluating a classifier; some are more suitable than others depending upon the context of their use. The formula applied for all the models is given by:

Satisfaction ~ Age + Pay + Recognition + Appraisal + Reward + Promotion

Accuracy and sensitivity of the models are measured to evaluate the performance of the models used. These measures are discussed in the following subsections.

2.3 Metrics Used

Accuracy is the ratio between the number of samples correctly predicted and the total number of samples classified. The formula for calculating accuracy is:

$$Accuracy = \frac{No.\ of\ correct\ predictions}{Total\ no.\ of\ predictions\ made} \qquad (5.1.1)$$

However, accuracy may not be a good measure if there is high variance in the number of objects in the classes or imbalanced datasets.

Sensitivity measures are based on the confusion matrix. The confusion matrix shows the complete performance of the model. It is a matrix of size $n \times n$ representing the numbers of patterns of class i predicted in class j.

Table 1. The Confusion Matrix

	Predicted:YES	Predicted: NO
Actual: YES	True Positives	False Negatives
Actual: NO	False Positives	True Negatives

As shown in the Table 5.1, there are four cases:

- True Positives: Number of cases that are predicted YES and actual output is YES.
- True Negatives: Number of cases that are predicted NO and actual output is NO.
- False Positives: Number of cases that are predicted YES and actual output is NO.
- False Negatives: Number of cases that are predicted as NO and actual output is YES.

Sensitivity is also known as the true positive rate (TPR). This measure tells us the percentage of employees that are correctly identified as satisfied employees.

$$\text{Sensitivity} = \text{No. of true positives} \quad \text{No. of true positives} + \text{No. of false negatives} \qquad (5.1.2)$$

In this paper, we have used K-fold cross validation for measuring the performance of the predictive models over new data. In K-fold CV, the original sample data set is divided into k equal size subsets. Out of the k subsets, one subset is retained as validation set and the remaining (k-1) subsets are used for training the model. The process is repeated k times (the folds) while making sure that each subset is used as validation set exactly once. The evaluation metrics accuracy and sensitivity are computed for each run and then averaged to produce a single estimation.

3. Model Selection

In this paper, we have used TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) method for selecting the best machine learning model based on their performance in terms of accuracy and sensitivity. TOPSIS is a Multiple criteria decision-making (MCDM) technique that was first developed in 1981 by Yoon and Hwang [D4]. TOPSIS is based on the concept that the chosen alternative should have the shortest geometric distance from the positive ideal solution (PIS) and the longest geometric distance from the negative ideal solution (NIS) [4,5].

4.Results

In this section, the results of all four machine learning models on the given dataset have been analysed. The accuracy has been computed using the equation 5.1.1 and sensitivity has been calculated using equation 5.1.2 for every model used. As discussed in section 5.2, performance of the models used was further analysed using TOPSIS to select the optimal model for our case. The table 5.3 shows the results.

Table 5.3: Performance comparison of all models.

Models	Accuracy (%)	Sensitivity	Rank given by TOPSIS
Random Forest	87.3	0.89	1
Decision Tree	81.1	0.85	2
Support Vector Machine	78.5	0.84	3
Binary Logistic Regression	77.0	0.84	4

As shown in Table 5.3, random forest performs best among all models while Binary Logistic Regression performing the worst. Random forest gives good sensitivity score of 0.89.

5. Conclusion

We have successfully demonstrated the application of four popular machine learning models for predicting the employee satisfaction for Leading Logistics Company.

6. References

[D1] https://www.rdocumentation.org/packages/mice/versions/3.6.0/topics/mice

[D2] Nowok, B., Raab, G. M., & Dibben, C. (2016). synthpop: Bespoke creation of synthetic data in r. Journal of Statistical Software, 74, 1–26. doi:10.18637/jss.v074.i11

[D3] L. Breiman, J.H. Friedman, R.A. Olshen, , and C.J Stone. Classification and Regression Trees. Wadsworth, Belmont, Ca, 1983.

[D4] Yoon, K.; Hwang, C.L. Multiple Attribute Decision Making: Methods and Applications; Springer: Berlin/Heidelberg, Germany, 1981.

[D5] Yoon, K. System Selection by Multiple Attribute Decision Making. Ph.D. Dissertation, Kansas State University, Manhattan, KS, USA, 1980.

[D5] Behzadian, M.; Otaghsara, S. K.; Yazdani, M.; Ignatius, J. 2012. A state-of the-art survey of TOPSIS applications, Expert Systems with Applications 39(17): 13051–13069. http://dx.doi.org/10.1016/j.eswa.2012.05.056

[D6] Heyburn, R., Bond, R. R., Black, M., Mulvenna, M., Wallace, J., Rankin, D., & Cleland, B. (2018). Machine learning using synthetic and real data: Similarity of evaluation metrics for different healthcare datasets and for different algorithms. In Data Science and Knowledge Engineering for Sensing Decision Support (pp. 1281-1291).

[D7] Sabay, A., Harris, L., Bejugama, V., & Jaceldo-Siegl, K. (2018). Overcoming Small Data Limitations in Heart Disease Prediction by Using Surrogate Data. SMU Data Science Review, 1(3), 12.

[D8] Bouhlila, D. S., & Sellaouti, F. (2013). Multiple imputation using chained equations for missing data in TIMSS: a case study. Large-scale Assessments in Education, 1(1), 4.

[D9] https://www.thebalancecareers.com/what-is-human-resource-management-1918143

[D10] https://atmanco.com/blog/working-environment/internal-marketing-when-marketing-human-resources-collide/