Prediction of Human Talent using Enhanced Naive Bayes Algorithm

Rajvir Kaur[1], Ms. Poonamdeep Kaur [2], Ms. Lovepreet Kaur [3]

Student, Computer Science & Engineering, Guru Nanak Dev Engineering College, Ludhiana, India[1]

Assistant Professor, Computer Science & Engineering, Guru Nanak Dev Engineering College, Ludhiana, India[2]

Assistant Professor, Information Technology, Guru Nanak Dev Engineering College, Ludhiana, India[3]

Abstract: A modern day fast-paced, global business surroundings, coupled with the ever growing problem of a rapidly changing labor force, requires a well-planned, rigorous method of talent management. Fortunately, each HR and business frontrunners recognize that talent is an essential driver of business overall performance. It comes as no surprise after that that talent management methods are often cited as an important strategic priority. However, the actual execution of these practices is actually where companies often are unsuccessful. Thus this paper focuses on developing an efficient algorithm to overcome challenges faced by a human resource manager and segregate effective new talents from the rest.

Keywords: Naïve Bayes, Talent Acquisition, Machine Learning.

I. INTRODUCTION

Talent acquisition is the procedure of finding and acquiring skilled human being labor for organizational needs and also to meet any labor requirement [1][2]. When found in the framework of the recruiting and HR occupation, skill acquisition usually identifies the skill acquisition division or team within the recruiting section. The TA team within a company is accountable for finding, acquiring, evaluating, and hiring applicants to fill functions that must meet company goals and fill up project requirements.

Talent acquisition as a distinctive function and department is a comparatively new development. In lots of companies, recruiting itself continues to be an indistinct function of the HR generalist [1]. Within many companies, however, recruiting as a designation didn't encompass enough of the responsibilities that dropped to the organization recruiter. Another designation of skill acquisition was necessary to meet up with the advanced and unique functions. Modern skill acquisition is a tactical function of a business, encompassing skill procurement, but also labor force planning functions such as organizational skill forecasting, talent pipelining, and tactical skill evaluation and development.

Talent acquisition is quickly learning to be a unique profession, maybe even unique from the practice of general recruitment [2]. TA experts are usually skilled not only in sourcing techniques, candidate evaluation, and compliance and employing standards, but also in work branding methods and corporate and business employing initiatives. Talent acquisition as a function is becoming carefully aligned with marketing as well as recruiting.

As an art, TA is of course not new; it's the simple procedure for recruiting good skill to meet company needs. As an occupation, however, SA is quickly growing into a distinctive and important job function.

Skill Acquisition means a view of not only filling positions, but also usage of the applicants and their skills which come out of the demanding recruiting process as a way to fill similar positions in the foreseeable future also. These future positions can be recognized today by looking at the succession management plan, or by analysing the annals of attrition for several positions [3]. This helps it be easy to forecast that specific opportunities will happen at a pre-determined period with time. Taking the long-term strategic method of talent acquisition has an enormous impact about how an approach was created to a candidate.

Talent Acquisition experts recognize that each talent has something of value to provide. In addition they build associations with the best of the skill that business lead to more lucrative networking, more recommendations, more business and an incredible give and take of experience, information and knowledge. Talent acquisition takes efficient and productive processes that are simple to use and applicant centric.

Machine learning is a kind of artificial cleverness or AI that delivers computers having the ability to learn without having to be explicitly programmed Machine learning targets the introduction of computer programs that can train themselves to develop and change when subjected to new data. Machine learning programs identify patterns in data and change program actions appropriately. Essentially, it is a way of teaching computers to make and improve predictions or behaviors predicated on some data. Bayes' Theorem is an axiom of chance theory traditionally mentioned by the Reverend Thomas Bayes [5][6][7]. It could be regarded as a way of focusing on how the likelihood a theory holds true is suffering from a new little bit of evidence. It's been used in a multitude of contexts, which range from sea biology to the introduction of "Bayesian" spam blockers for email systems.

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

In the viewpoint of science, it's been used to attempt to clarify the partnership between theory and proof. Many insights in the beliefs of science including confirmation, falsification, the connection between technology and pseudoscience, and other topics can be produced more precise, and sometimes prolonged or corrected, by using Bayes' Theorem.

Begin by taking a look at the theorem, displayed below.

$$P(A_i/B) = \frac{P(A_i) * P(B/A_i)}{\sum_{i=1}^{k} P(A_i) * P(B/A_i)} \ .$$

II. METHODOLOGY

This paper formulates an algorithm based on naive bayes theorem to precisely distinguish talent from the rest of the applicants by using their previous employment records.

This novel calculation could turn out to be useful for the HR on filling positions quickly and focus more on aligning skill acquisition with the business enterprise. The flowchart is shown below:
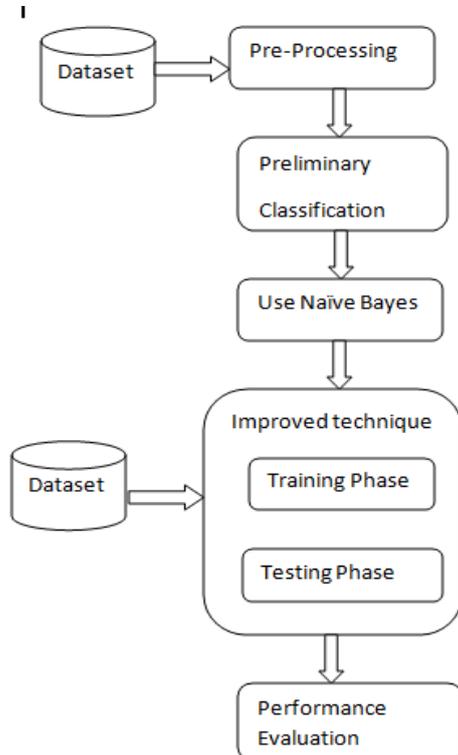

Fig.1. Flowchart of proposed Algorithm

A. Inputs

For this work 9 attributes taken from 474 different observations is the input to the proposed system. The selected parameters are described in table 1:

TABLE I: Description of attributes

Attribute	Description
Gender	Male=1 Female=0

Education	Masters=16 Graduate=15 Other=8
Job Category	Professional=1 Supportive=2
Salary	Quantitative
Salary Beginning	Quantitative
Technical Skill	Good=1 Average=2
Previous Experience	Less than or equal to 5 years=1 Less than or equal to 10 years=2 More than 10 years=3
Minority	General=1 Other=0
Date of Birth	Only the year of birth
Result	Selected=1 Not selected=0

B. Algorithm

The algorithm of the proposed methodology is given below:

Let Selected Attribute = S and Non Selected Attribute = S'

Step 1: Read Attributes from dataset.

Step 2: Compute X1 = $\mu(S)$

Step 3: Compute X1Prime = $\mu(S')$

Step 4: Compute Y1= $\sigma(S)$

Step 5: Y1Prime = $\sigma(S')$

Step 6: Repeat step 2 to 5 until μ and σ of each attributes are computed.

Step 7: Compute Normal Distribution of each Attributes.

Step 8: k = $\prod_1^n ND(S)$

Step 9: kprime = $\prod_1^n ND(S')$

Step 10: if K >kprime: Candidate Selected else Not Selected

Step 11: Stop.

C. Explanation

The proposed algorithm works on the principle of maximization of posteriori, for this a table T consisting of previous records of applicants are generated. For the notation convenience each attribute of the applicant is denoted by X ∈ {$X_1, X_2, X_3 \ldots X_n$} and the posteriori class be denoted by S and S', where S denotes "The applicant has been selected given X" and S' denotes "The applicant has been rejected given X". And mean and standard deviation is denoted by μ and σ respectively. Also, Normal Distribution is denoted by ND.

III. RESULTS

To test the accuracy of the system, 500 different test cases were randomly chosen, out 500 test samples 50 test samples were taken from the original test cases themselves.

To contrast the efficiency of the proposed system, 3 best algorithms were taken namely, ID3, Naïve Bayes and C4.5. And the proposed methodology turns out to be more efficient than others. The analysis has discovered that several factors may have a great influence on employee performance. One of the very most effective factors is the work name.

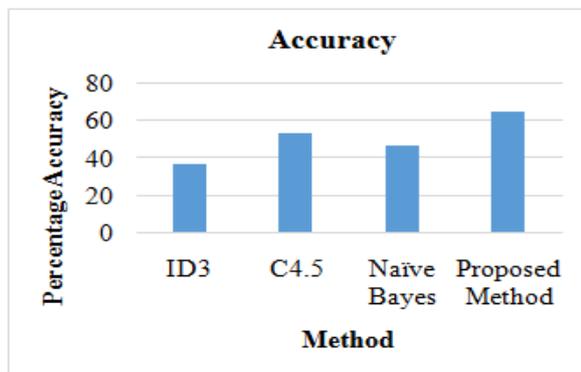The comparison table is shown in the picture.

Fig.2. Comparison of Accuracy with other algorithms

The comparison of the proposed system is based on their accuracy for 500 data items taken from different data sources. The formula used for the evaluation for the accuracy is given by:

Accuracy= Total number of correct Prediction/ (Number of correct Prediction + Number of incorrect Prediction

---

IV. CONCLUSION

This paper attempts to classify HR Talent Acquisition data in to the sets of Promotion Recommendation using Enhanced Naïve Bayes algorithm. For this work, an exercise group of an arbitrary test of 474 information and then applied the classification guideline to obtain the Employee test dataset. The obtained dataset here's 98% in working out phase. This paper has compared the performance of Enhanced Naïve Bayes algorithm with other classification techniques. The probability of creating a classification model for predicting the talent's recruitment by focusing on performance, many attributes have been tested, plus some of them are located effective on the performance prediction. The work category was the most powerful feature, the salary then, with minor after effect of time and experience. This attribute didn't show any clear effect as the gender show some effect in a few of the experiments for predicting the performance. Salary, quantity of pervious companies, test years and job satisfaction, each experienced an amount of influence on predicting the performance. For companies managements and recruiting departments, this model, or an improved one, can be utilized in predicting the newly candidate staff performance. Several activities can be studied in cases like this to avoid any risk related to employing low performance worker.

V. FUTURE SCOPE

As future work, it is strongly recommended to gather more proper data from several companies. Directories for current employees and even earlier ones can be utilized, to truly have a right performance rate for every one of these. When the correct model is generated, software could be developed to be utilized by the HR like the rules produced for predicting performance of employees. For data cleaning and transformation logical methods are used, but for large amount of dataset this methods can turn out to ineffective, hence standard data mining tools should be used for storing, cleaning and transforming the data. Use of Some standard SQL server is recommended as opposed to file based system for ease of retrieval and updation. The run time of the proposed method is $O(n)$ but space complexity of data is $O(n^2)$, hence some compression technique could be employed to reduce space complexity.

ACKNOWLEDGMENT

REFERENCES

[1] Beechler, S. and Woodward, I. C. (2009) *"The global "war for talent""*, Journal of International Management, vol. 15, pp. 273-285.
[2] Boudreau, J. W. and Ramstad, P. M. (2005) *"Talentship and the new paradigm for human resource management: From professional practices to strategic talent decision science",* Human Resource Planning, vol. 44, pp. 129-136.
[3] Cappelli, P. (2008) *"Talent on Demand",* Challenges, vol. 39, pp.414-431
[4] Jantan, O.H. (2011) *"Talent Knowledge Acquisition using Data Mining Classification Techniques* ", 3rd Conference on Data Mining and Optimization, pp. 32-37.
[5] Radaideh, Al. and Qasem, E.A.N.(2012)" *Using Data Mining Techniques to Build Classification Model for Predicting Employees Performance*", International Journal of Advanced Computer Science and Applications(IJACSA), vol. 3..
[6] Rennie, J.D.M., Shih,L., Teevan, J. and Karger, D.R.(2003)" *Tackling the Poor Assumptions of Naive Bayes Text Classifiers",* Proceedings of the Twentieth International Conference on Machine Learning (ICML), vol. 20, pp. 616-623..
[7] Shan, H. and Banerjee, A. (2011) *"Mixed-membership naive bayes models",* Data Mining and Knowledge Discovery, vol. 23, pp. 1-62.