



Employee Culling based on Online Work Assessment through Machine Learning Algorithm

By Khaled Redwan, Yeasin Ahammed, Masum Akram Hridoy, Fernaz Narin Nur
& A. H. M. Saiful Islam

Notre Dame University

Abstract- Job analysis, planning employee needs, recruiting the appropriate people, wages and salary management are the important theme of human resource management. Human resource management also includes evaluating performance, resolving problems, and create communication with all employees at all levels. On the other hand, Machine learning is a data analytics technique that teaches computers to do what comes naturally to humans. So through these two sectors such as computation and business administration, in this paper on employee culling based on work assessment by which machine learning algorithm such as KNN, SVM, The Decision tree can give the best result (perfect employee). We also focus on the accuracy that algorithm is performing. We marked an employee through their experience, language skills, skills, graduation, etc. we create e model by which we can get input through the companies and give them a perfect result through their requirement assessment and machine learning algorithm.

Keywords: machine learning, employee culling, human resource management, KNN. SVM.

GJCST-D Classification: F.2.1



Strictly as per the compliance and regulations of:



Employee Culling based on Online Work Assessment through Machine Learning Algorithm

Khaled Redwan ^α, Yeasin Ahammed ^σ, Masum Akram Hridoy ^ρ, Fernaz Narin Nur ^ω
& A. H. M. Saiful Islam [¥]

Abstract Job analysis, planning employee needs, recruiting the appropriate people, wages and salary management are the important theme of human resource management. Human resource management also includes evaluating performance, resolving problems, and create communication with all employees at all levels. On the other hand, Machine learning is a data analytics technique that teaches computers to do what comes naturally to humans. So through these two sectors such as computation and business administration, in this paper on employee culling based on work assessment by which machine learning algorithm such as KNN, SVM, The Decision tree can give the best result (perfect employee). We also focus on the accuracy that algorithm is performing. We marked an employee through their experience, language skills, skills, graduation, etc. we create a model by which we can get input through the companies and give them a perfect result through their requirement assessment and machine learning algorithm.

Keywords: machine learning, employee culling, human resource management, KNN, SVM.

I. INTRODUCTION

According to Human Resource Management, the process is culling employee, give them perfect training and development, appreciate their performance, maintain relationships, ensuring employee satisfaction, health and welfare management and maintain labor laws, etc. And now a day's machine learning is gaining the perfect analysis and prediction through data and they have the automated tools which can analyze the profile of business user. Through the machine learning algorithm, we provide the best employee to a company. In the busiest era, it is time-wasting matter to find the best employee from so many resumes and it is a challenging matter too. The HR department of any company has to put his valuable time to find an employee by a manual process if there is employee turnover in that company. So based on companies and online work requirement assessment, we have developed a model and through machine

learning by which we can provide them the best employees from so many employees dataset, and thus the paper determines Employee culling based on online work assessment through the machine learning algorithm. Through our Process, we can find the best employee through the skill requirement of a client and company. The process will focus on the experience education, and the skills of an employee. This process can be a future human resource management application.

II. RELATED WORK

In this paper, we create a model but, it has previous literature, which is known as the "Holy Grails" project to AMAZON. In which they put the information of the employee, and the algorithm will provide the best and talented one. But this company had some limitations in 2015. The company have realized their new system, which couldn't rating their candidates in a gender neutral way in the software developer jobs. The reason is the computer model have analyzed patterns in resumes which is submitted over a 10 year period. And in that tech industry most of the employees are male. At "Unilever" they select their employees through some process. They have to complete an application form. It is an easy process. They can apply through "Linked In" profile. On the application we can type many functions through varies and region. After that, they have to play 12 online analytical games. After completing the games they will receive personalized feedback. They have solve real- world problems through "Unilever" scenarios. A digital interview will occur. Where a candidate needs a suitable internet connection.

But "Unilever" has some limitations too, they got shortlisted employees but didn't get the perfect employee. [1][9].

Now the difference between manually employee selection process and employee selection through machine learning process is bellowed [1] [9].

Author ^{α σ ρ ω ¥}: Notre Dame University Bangladesh
Dhaka, Bangladesh. e-mails: khaled.redwan.rafsan@gmail.com,
fernazcse@ndub.edu.bd, yeasin.ahammed.sourav@gmail.com,
hridoyj@gmail.com, saiful@ndub.edu.bd

Step No.	Manually employee selection process	Employee selection through machine learning algorithm
1.	Identify your hiring needs	Load applicants list
2.	Analyze the position	Get input of client's requirement
3.	Evaluation, description and person specs	Mark an employee through his experience, skills, cgpa etc
4.	Develop a recruitment strategy	Mark an employee through his analytical skill
5.	Go through the applications	Put these marks in the algorithm
6.	Conduct interviews and tests and conduct a final interview	Get the employee according to the requirement

III. METHODOLOGY

24 *Process 1:* To find the best employee, we have to reduce the applicant's list through some basic requirements of the client.

Process 2: Remove those applicants who are not in the experience zone.

Through these two processes, we can reduce our dataset. Then the recruiter will provide some questions or games, and the workers have to answer it or should use their analytical skills. This exam will provide marks (50%)

The second stage is, the employee have to submit a video interview, where the assessor is not a human but a machine. The machine examines the candidates who has to answer questions for around 30 minutes. The natural language processing and body language analysis will determines who is fit for the job. This interview will provide the value of potentiality (50%) [4][7].

Total number = Mark (50%) + potentiality (50%) = 100%

Now a data set is given bellow

Number	Employee name	Marks	Potentiality
1	Jack	15	20
2	David	20	20
3	Jason	25	22
4	Francis	30	25
5	Chester	32	23
6	Marlin	35	32
7	Shuvashish	40	38
8	Arthur	42	39
9	Mary	43	40
10	Samuel	46	50

Dataset 1

Put the value GRAPH 1 where Here X-axis contains the value of Marks and Y-axis contains Potentiality

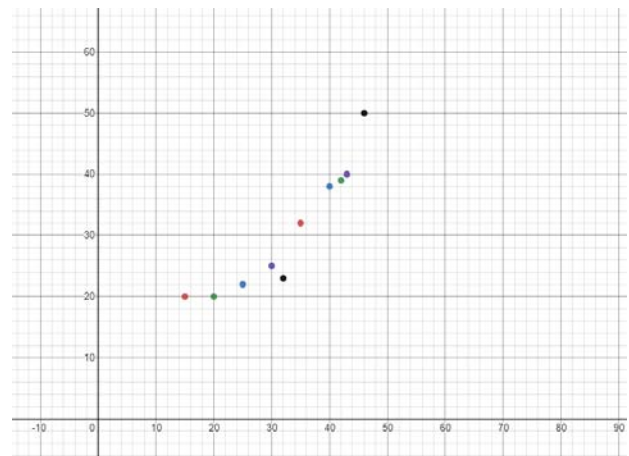


Figure 1: Graph 1

Now the client will input the value of marks and potentiality of the employee that he needs for his company.

Let the client need some of the best employees whose marks are = 40 and potentiality is= 40. And he wants to show four best employees out of 10 employees.

Here k = 4 and (x, y) = (40, 40)

Now, we have to put the point (40, 40) on the GRAPH 2

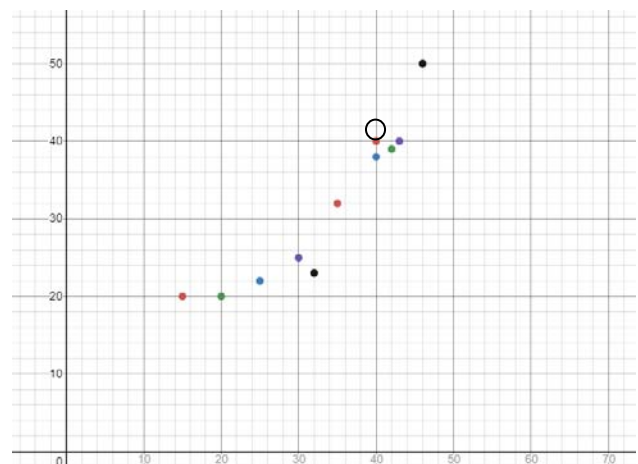


Figure 2: Graph 2

Now the K nearest neighbor algorithm will:

- Load the data
- Initialize the value = k
- For the predicted class, repeat from 1 to total number of training data points
- We have to calculate the distance between test data and training data (each row). We will use the "Euclidean distance" as our distance metric.
- We need to sort the distances in ascending order based on distance values
- We will get top k rows (from the sorted array)
- Get the most frequent class of these rows
- Return predicted class [11] [2][3]

Through this algorithm, we get Kth best employee list.

Now the process of decision tree algorithm is given bellow.

- Pick the best attribute/feature. The best attribute is one that best splits or separates the data.
- Ask the relevant question.
- Follow the answer path.
- Go to step 1 until we arrive at the answer [12]

According to the discussion, we can determine the summary of the methodology in such a process, which is defined as:

- A. Load all the application list
- B. Receive working tools (W) name from the client
- C. Receive experience range from the client. The range should be $(0 < N < L)$; N= employee working experience and L = clients experience requirement
- D. Remove applicants list
when $W \neq$ Skill of applicant
- E. Remove applicant list
if, $(0 < N < K)$ is not true
- F. Input the value of Marks and Potential from the client
Load the final list after removing
- G. Give analytical exam and a video interview by the employee
- H. Input the value of K, K= the number of employee the clients need
- I. Input the value of Marks and Potential from the client
Use K nearest neighbor, decision tree,
Find the employee list [8]

For accuracy, we tried a new dataset [5] [6]

Academic CGPA	Experienc	Training	Professional	Cei	Publications	Skill	EngishSkill	ExtraCurricular	Availability	UserRating
1	0.96	-0.75	0.2	2	0.2	0.7	0.92	2	1	3.48
0.6	0.925	-0.5	0.2	2	0.8	0.6	2	2	-1	2.7
0.8	0.5	-0.75	2	2	2	0.3	2	2	-1	1.85
0.8	0.9125	-1	2	2	2	2	1	2	1	2.45
1	0.82	0	2	2	0.6	1	1	2	-1	3.85
0.8	0.94	-1	2	2	2	0.9	2	2	1	2.65
0.8	0.945	0.5	2	2	2	0.9	1	2	-1	2.9
0.8	0.955	-1	0.8	2	2	0.9	2	2	1	2.65
1	0.56	1	2	1	2	0.6	2	2	1	4
0.8	0.93	-0.75	2	2	2	0.4	2	2	-1	1.9
0.8	0.955	-0.5	2	2	0.6	0.9	2	2	-1	2.85
1	0.75	1	0.4	1	1	1	2	2	-1	4
1	0.75	0.5	2	1	1	2	2	2	-1	3.3
0.8	0.9775	0	0.2	2	1	0.9	1	2	-1	3.3
1	1	1	1	1	1	2	2	2	-1	3.5
1	1	1	0.4	1	1	2	2	2	-1	3.5
0.8	0.9375	-0.75	2	2	2	0.6	2	2	1	2.5
0.8	0.8275	0	0.8	2	2	2	2	2	-1	2.1
0.8	0.89	-0.75	0.2	2	2	2	2	2	1	2.2
0.5	0.86	-0.75	2	1	2	0.3	0.785	2	-1	2.55
0.6	0.8475	0	1	2	2	0.9	2	2	-1	2.35
0.5	0.925	-1	0.2	2	2	1	2	1	1	2.95
0.9	0.9	-0.75	2	2	2	0.9	0.928	2	1	3.99
0.8	0.8575	-0.75	2	2	2	0.7	2	2	-1	2.05
0.5	0.9125	0	0.2	2	2	0.9	2	2	-1	2.65
0.8	0.8625	-0.75	1	1	2	2	2	2	-1	1.7
0.6	0.6375	-0.75	0.2	1	2	0.8	0.928	2	1	2.84
0.8	0.885	-0.75	0.2	2	2	0.8	2	2	1	2.6
0.8	0.7475	0	0.4	2	2	0.8	2	2	-1	2.6
0.8	0.7875	0	2	2	2	1	2	2	-1	2.7
0.8	0.88	-1	0.2	2	2	0.4	2	2	1	2.4

Figure 3: Dataset 2

Explanation of the Dataset 2:

The attributes of the DATASET 2 given in are- Experience, Cumulative Grade Point Average, Skill, Publication and Training and Language Proficiency. the explanation of all the attributes of Dataset 2 is bellowed.

a) Experience

According to the above-given scale, a fresher candidate gets the minimum score, which is -1 according to our range of values. A candidate with experience of more than eight years gets the maximum score, which is 1. Except for the score -0.75 for less than two years of experience, in case of all other intervals, the score changes with a step of 0.5.

Table 1: [Experience]

Number of years of experience	Score
No experience (fresher)	-1
Experience < 2	-0.75
2 < experience <= 4	-0.5
4 < experience <= 6	0
6 < experience <= 8	0.5
Experience > 8	1

b) Cumulative Grade Point Average

Scoring the CGPA is much simpler. As maximum CGPA possible is 4.00, we considered that to be one under our scoring range, which is the maximum. Any other CGPA in the scale of 4 is convert into the scale of 1.

For instance, a candidate with CGPA 4.00 gets a maximum score of 1 added to his CGPA feature. If the CGPA was 3.5 then the score would be, Score = $3.5 / 4 = 0.875$

So here, the CGPA 3.5 out of the scale of 4 has got converted to 0.875, which is out of 1.

c) Skill

We have set the scoring policy depending on the types of CVs that we have come across. Candidates that have mentioned more than three skills, we fixed it to score their skills feature to be maximum that means 1. But, those who have mentioned three or less, then there we considered the degree of their skills in those areas or subjects. Which is bellowed.

Table 2: Skill

Good	Very good	Excellent
0.1	0.2	0.3

According to these criteria, suppose a candidate has mentioned about three skills MS Word, MS Excel, and PowerPoint. Consider the table below as an example,

Table 2.1 [Skill]

Skill	Good	Very good	Excellent
MS Word		√	
MS Excel		√	
PowerPoint			√

From the above skills and their degrees mentioned by a certain candidate if we calculate the score,

$$\text{Score} = \text{MS word (Very good)} + \text{MS Excel (Very Good)} + \text{PowerPoint (Excellent)} = 0.2 + 0.2 + 0.3 = 0.7$$

d) Publication and Training

In case of the features like training and publications, we have considered the number of training programs a particular candidate has attended and the number of publications they have made. We focused mostly on the numbers rather than anything else. For each of the publications that a candidate has made, 0.2 gets added to the publications feature. For example, someone with 4 publications get a score of $4 * 0.2 = 0.8$. Just the same as that for each of the training programs attended, a candidate gets same score 0.2. A candidate with five pieces of trainings attended gets a score of $5 * 0.2 = 1$.

e) Language Proficiency

For the enumeration of the feature English language proficiency, we considered the IELTS score to be the most standard scale to measure with an IELTS score of 7 or more will be taken as the maximum value 1 for this feature. If it is less than 7, then the feature value gets curbed according to the scale of 1.

For instance, if a candidate has an IELTS marks of seven then it adds 1 to the feature English Proficiency, where one is the maximum. If there is some other candidate that has a score of 6.5 then the score is, $\text{Score} = 6.5 / 7 = 0.928$; which is approximately one and the score is good.

IV. RESULT ANALYSIS

According to DATASET 1, which is applied in KNN and decision tree algorithm, we get that Shuvashish [7th employee from DATASET 1] is the perfect employee.

According to the accuracy from DATASET 2, we can deduce that SVM and Decision Tree algorithms provide the best results. Although the Decision Tree gives a higher percentage than the SVM algorithm.

Table 3: Result Analysis

Decision Tree	SVM	Multi-Linear
85%	80%	72%

But we would recommend SVM. The reason is, that a decision tree starts the process of building a tree from scratch every time the algorithm is calling but with a different root node and hence gives more volatile results as well as being more prone to over-fitting as the complexity of the dataset increases.

V. CONCLUSION

Human resources management has a vision that all the customers feel like a part of the community. Therefore a company will always be met with prepared and helpful employees. HR management considered their employees as important resources. On the contrary, today's innovation and development is machine learning. This innovative subject is much more than we think, which can develop anything as human wants. So in this paper, for employees and companies betterment, we create an internal and external bonding between human resource management and machine learning through the objective of the HR manager and machine learning algorithm.

REFERENCES RÉFÉRENCES REFERENCIAS

1. <http://www.whatishumanresource.com/human-resource-management>
2. The Implementation of K-Nearest Neighbor Algorithm in Case-Based ReasoningModel for Forming Automatic Answer Identity and Searching Answer Similarity ofAlgorithm Case Yana Aditia Gerhana1, Aldy Rialdy Atmadja2, Wildan Budiawan Zulfikar3, Nurida Ashanti4 1,2,3,4Department of Informatics UIN Sunan Gunung Djati Bandung Jl. AH Nasution No. 105, Bandung, West Java, Indonesia 1yanagerhana@uinsgd.ac.id, 2aldy@if.uinsgd.ac.id, 3wildan.b@uinsgd.ac.id, 4nurida.ashanti@student.uinsgd.ac.id
3. Comparative Analysis of K-Nearest Neighbor and Modified K-Nearest Neighbor Algorithm for Data Classification
4. The Amazing Ways How Unilever Uses Artificial Intelligence to Recruit & Train Thousands of Employees[<https://www.forbes.com/sites/bernardmarr/2018/12/14/the-amazing-ways-how-unilever->

- uses-artificial-intelligence-to-recruit-train-thousands-of-employees/#593b01ca6274]
5. Okfalisa Department of Informatic Engineering UIN Sultan Syarif Kasim Riau Pekanbaru - Indonesia okfalisa@uin-suska.ac.id
 6. An Adaptive k-Nearest Neighbor Algorithm
 7. Shiliang Sun Rongqing Huang Department of Computer Science and Technology, East China Normal University 500 Dongchuan Road, Shanghai 200241, P.R. China slsun@cs.ecnu.edu.cn, rqhuang09@gmail.com
 8. A Quick Evidential Classification Algorithm Based On K-Nearest Neighbor Rule Zhuang Wang, Wei-Dong Hu, Wen-Xian Yu Atr State Key Lab, National Univ. of Defense Technology, Cbangsha, 410073, China E-MAIL zhuangwang@sina.com
 9. <https://bizfluent.com/info-8212981-recruitment-process-hr.html>
 10. Unilever application process [<https://www.unilever.co.za/careers/graduates/application-process/>]
 11. Introduction to K nearest neighbors: A powerful machine learning algorithm(with implementation in python & R) Tavish Srivastava
 12. <https://www.hackerearth.com/practice/machine-learning/machine-learning-algorithms/ml-decision-tree/tutorial/>
 13. Amazon scraps secret AI recruiting tool that showed bias against women [<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>]

