Smart Human Resource Management System to Maximize Productivity

Conference Paper · December 2020

DOI: 10.1109/ICS51289.2020.00100

4 authors:

Amila Nuwan Senarathne
Sri Lanka Institute of Information Technology

Asiri Hewage
University of Westminster

Janaka L Wijekoon
Sri Lanka Institute of Information Technology

Kushani Hettiarachchie
Sri Lanka Institute of Information Technology

Smart Human Resource Management System to Maximize Productivity

Hewage H.A.S.S
Dep. of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it17094078@my.sliit.lk

Hettiarachchi K.U
Dep. of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it17156752@my.sliit.lk

Jayarathna K.M.J.B
Dep. of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it17157124@my.sliit.lk

Hasintha K.P.C.
Dep. of Information Technolgy
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it17058292@my.sliit.lk

A. N. Senarathne
Dep. of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
amila.n@sliit.lk

J. Wijekoon
Dep. of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
janaka.w@sliit.lk

2020 International Computer Symposium (ICS) | 978-1-7281-9255-0/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICS51289.2020.00100

***Abstract*—Human resource is one of the most valuable assets in an organization. They are bounded to develop the unique and dynamic aspects that strengthen their competitive advantage to persist in an always changing market environment. In order to recruit a quality candidate for an organization, reducing human involvement and verifying details of the candidate is important in recruitment process. Furthermore, having an idea about how well or poor the employees perform, and how likely the employee attrition can occur is vital in human resource management process. This paper is an attempt to introduce smart human resource management system that can maximize the productivity of an organizational environment using machine learning and blockchain technologies. The end goal of this research is a smart human resource management system that reduces human judgment, time in the candidate selection process and predicts employee performance and attrition to motivate current employers to maximize productivity with minimal financial loss in the workplace environment. Skill assessment and resume classification have been done using unsupervised learning algorithms and natural language processing after extracting raw data from employee resumes using Object Character Recognition. Candidate details verification is done by comparing the hashes of the records which are stored in the blockchain. Employee performance and attrition are predicted using supervised machine learning classification techniques with high accuracy and the result of the final performance is generated as a score for each employee considering the multiple attributes that has been standardized and regulated by some specifically considered e-competence frameworks.**

***Index Terms*—Machine Learning, Blockchain, LDA, Human Resource Management, Recruitment, Maximizing productivity**

I. Introduction

Human resource (HR) is one of the greatest assets of any organization which enables achieving organizational goals and objectives. Human Resource Management (HRM) affects the overall productivity of the organization. It can be identified a rapid growth in Machine Learning (ML) and Blockchain technologies in HRM processes, with the development of cloud technologies in recent years [1]. Human Resource Management departments of organizations are facing a challenging task of recruiting right talent who can work in different problem domain and achieve goals it within timeline. Choosing a well-qualified candidate for a job position with accurate and verified details will help an organization to achieve desired productivity. Furthermore, having idea about the employee's suitability to the current resource pool, performance and the probability of leaving the company will help managerial positions and decision-makers to take correct decisions.

Traditional candidate selection methods are mainly focused on competency factors of individuals based on person-job fit assessment [2]. Computer software-based HR tools can support the decision making of HR Personnel. But the accuracy of decisions suggested by algorithms still cannot satisfy the HR personnel [3]. Also, there are international standards and frameworks [4] to regulate this recruitment process. Sometimes management decisions and changes in the way of work may lead to conflict of interest when contradictory interests relate to an activity by an employee. Most of the time companies hire employees who have more or less than expected skills levels and end up with a totally different resource pool compared to their business goals [2]. Verification of employee details ensures that an employee could perform the job that organization is offering. Although verification of employee records is vital task, it is not effective in time and cost for the employers. Moreover, most of the methods which are currently available for verification is manual and involve human interaction [5]. Therefore, verification process can be biased. During this process, the HR person will spend much time on reaching organization for the verification, as well as waiting them to

978-1-7281-9255-0/20/$31.00 ©2020 IEEE
DOI 10.1109/ICS51289.2020.00100

reply for the confirmation of the required records [6]. This process is extremely burden for the HR department of an organization when there are hundreds of records to be verified. Furthermore, HRM and decision-makers spend a considerable amount of time to find the best fit to execute the right job at the right time [7]. In case of an incident, this is well needed as to tackle the situation and maintain the continuous competitive advantage. The Business Dictionary of The Barron defined attrition as the Natural and uncontrollable labor force decline in age, disability, sickness, and relocation [8]. This loss is not all for money, the organization even lacks the professional employees that is the most valuable enterprise resources.

To overcome all the mentioned problems above, it needs to have a proper management tool. This paper introduces a system that helps HR personnel in the process of employee recruitment, document verification, employee performance and attrition prediction. The rest of the paper is organized and structured as; Section II discusses the existing systems in HRM domain, whilst section III introduces the objectives of this research study. The methodology carried out to develop the proposed system is described in section IV and the results of the developed system in the section V. Finally, the conclusion and the future dimensions are stated in the section VI.

II. Literature Survey

HR automation systems of the new era have introduced effective approaches for applicant resume filtering and candidate assessment via online proctored examinations. But it has been recognized as a critical need for organizations to scale and expand their human resource pool only for a specific business goal [1]. Using machine learning based rating, machine learning based predictions are black boxes to the decision makers other than using data mining for candidate selection.

Identifying the impact on Employee Performance (EP) is another challenge where ML approaches can be applied to solve major problems. A lot of researchers have tried classification techniques in solving and predicting certain aspects in the science domain [9].

Thus, prediction and analyses of employee's performance proficiency are known to be a critical problem for understanding the number of features and aspects related to the model's efficiency. Krimi JM, Motour CA (2016) focused on getting employee information from a public development firm in Kenya with and subjecting data provided to a Decision Tree model and generating the tree structure using the historical data. Then they would try to recognize the impact of feature-set on the accuracy of the Decision Tree. Furthermore, they focused on developing multiple prediction models based on different prediction techniques and choose the most suitable model for the organization [9]. Desouki M. S., Al-Daher J (2015) concentrated on a study for using Data Mining techniques in HRM. They focused on analyzing the results of Performance Appraisal. This was kept up by an organization to amplify the method of appraisal to get the measure of compatibility of practically implementing with the Performance Appraisal process objectives [10]. In order to accomplish that, they have used several techniques of data mining. The outcome of this study was that data mining techniques provide a greater significance in human resource activities to inflate employee's performance analysis [11].

III. Research Objectives

• To reduce human judgment in resume classification by introducing a topic modeling based resume scoring system and a person-skill visualization using high dimensional space.

• To motivate current employers by discovering potential talents and maximizing productivity in the workplace environment by predicting employee's performance.

• To minimize financial loss in the organization by predicting employee attrition and analyzing reasons for the retention.

• To prevent fraud and reduce time with the integrity in employee details verification process by using block chain technologies.

IV. Methodology

The resulting system is a web-based Platform as a Service which contains four main sub systems to aid the end goal. They are namely skill assessment subsystem, employee details verification subsystem, performance predicting subsystem and an attrition predicting subsystem.

• A company can sign up and start using the platform after following few steps to complete their company profile and other required information. Before using the platform, a company should have activated models for each job role. For that they can upload a set of resumes to build their own model to suit a job role or they can purchase a model from the store by providing their preferred skills. Then they can open a hiring portal to the public. Through that hiring portal, candidates will communicate with the company. Hiring managers will use the tools provided by the platform to assess the candidate. Candidate can upload their resumes through the portal, and they will be registered automatically if they have included the required information into the uploaded resume. The platform will parse contact information from the CV and contact the candidates back with their personal access token.

• Also, the end-users of the system will be able to input the details of their employees through an interactive user interface either manually or parsing a detailed document, finally to produce the result whether the employee is efficiently performing or under-performing. This result is averaged into a percentage to get a better understanding.

• System also capable of estimating service years of a particular employ, hence determine the retention factors to retain a valuable employee.

• An authorized HR person of the company can verify candidate educational certificates through the system. Furthermore, the solution allows an authorized HR person of an organization to add employees and their records to the blockchain. Once an employee is going to resign , a QR code which is linked to the employee records attached with the service letter can be issued to the relevant employee.

A. Resume Data Extraction, Rating and Visualization

In order to rate each resume according to the content, a separate python module has been used which can converts all kinds of documents into a set of images and then it will extract text using Optical Character Recognition (OCR), since existing text extraction methods returns null for image type resumes. Figure 1 describes the steps of extracting resume contents.

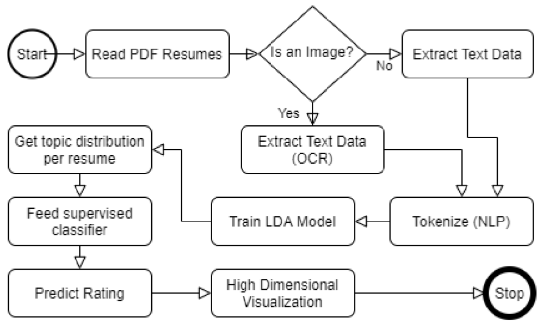

FIGURE 1. RESUME RATING WORKFLOW

After using Natural Language Processing (NLP) to generate word2vec correlation matrix of all tokens and keyword tokens, the converted tokens and chunks of words were converted into Data Frames (DF) using Pandas python library. Tokens are the instances of a sequence of characters in a document grouped together for processing. To generate a rating Latent Dirichlet Allocation (LDA) model [12] has been used.

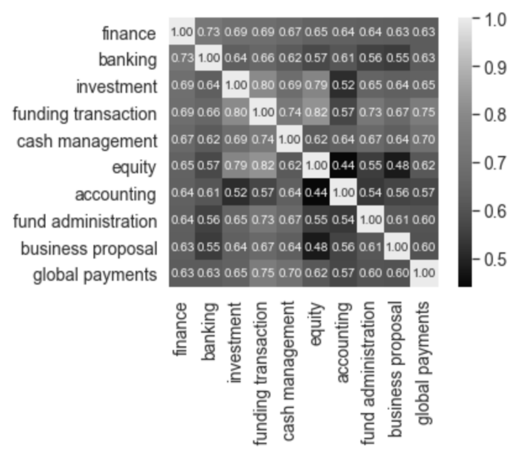

FIGURE 2. CORRELATION MATRIX- CORRELATIONS OF WORDS TO THE KEYWORD

Fixed keyword matching approach has been used and the other method is a dynamic keyword matching approach, where parameters or hyper-parameters are not needed. The parameters *number of topics* and *chunksize* were fine tuned to use with the Hoffman method [13].

After getting the list of all words from each resume, the frequency distribution of all words has been generated using python-Natural Language Tool Kit (NLTK). After getting minimum frequency for words and filtering all words whose counts are less than the minimum frequency, the number of unique words k has been retrieved.

For this research, articles with more than 50 tokens have been used, otherwise it's too short and can not be used to achieve a good accuracy. Using bag-of-words extracted from resume documents, top 5 words for each topic were selected. Since the words are only being used to compare with keywords generated, it does not matter whether words are filtered or not. During the model evaluation process could achieve a perplexity score of -8.033 using those parameters. According to the correlation matrix described in the Figure 2, the correlations are symmetric and the lower triangle can be ignored. Based on correlation, the average of top n correlations with keyword (*KM*) and the average of top n correlations without keyword (*WM*) are calculated. Then *KM* score and *WM* score for each keyword are generated using Linear Regression. Finally, the score of the resume with n number of keywords was calculated according to the equation (1). Then the top k words and the trained model have been saved in a MongoDB Database.

$$finalScore_j = \frac{1}{n^2}\left(\sum_{i=1}^{n}(KM_i) * \sum_{i=1}^{n}(KW_i)\right) \tag{1}$$

After taking the mean (μ) and standard deviation (σ) of the model, the rating if the resume j is calculated by the given equation (2) based on how much *finalScore* is away from μ.

$$rating_j = \frac{(finalScore_j - \mu_j)}{\sigma_j} \tag{2}$$

Then considering an employee as a high dimensional data point, a high dimensional visualization was generated to visualize the result of the classification. Employees were considered as nodes and kills of each employee were considered as dimensions. The rating has been used as the weight of the vertices. Then the employees were clustered according to the job role and visualized them considering a specific job role. T-Distributed Stochastic Neighbouring Entities (T-SNE) algorithm and JavaScript frameworks like D3 and React have been used to generate the visualization.

B. Employee Performance Predication

In the process of making a predictive ML model for the employee performance, a sample data set was selected. This data set was taken from the IBM HR analytical data published in the Kaggle [15]. The sample data set was pre-processed to grasp the most significant attributes. At the end, 17 attributes were selected to predict the prediction model. Later, the processed data set was subjected to three different ML algorithms namely, Decision Tree, Random Forest and Naive Bayes. Accuracy of each the algorithm was generated by changing their hyper-parameters. According to the results generated from each algorithm, it can be concluded that Naive Bayes Gaussian algorithm generated the most accurate result. In the employee performance analysis, a strong assumption has been taken that the attributes are not co-related, I.e. they are independent of each other and continuous values associated with the target class are behaving according to a normal distribution. The Naive Bayes theorem (3) can be applied to an example scenario where the employee performance is measured against salary range of the employee as below

$$P(a/b) = P(b/a) * P(a) / P(b) \quad (3)$$

According to the equation (3), P(a/b) is the probability of employee performance being ”true” given that salary is ”true”. *P(b/a)* is the probability of salary being ”true” given that employee performance is ”true”. The probability of employee performance being” true” is denoted by *P(a)* and the probability of salary being ”true” is denoted by *P(b)*. Likewise, the other significant attributes are too measured against the target class (Employee Performance) with the Naive Bayes formulae to find its probability.

According to the Figure 3, the performance of the Naive Bayes algorithm can be described with a help of a confusion matrix. The number of correctly classified and unclassified values are broken down into each class ”0” and ”1”. According to the Figure 3, the correlations shows a symmetric distribution and there are few instances where the model is confusing the two classes. But this is negligible compared to the correctly classified instances.

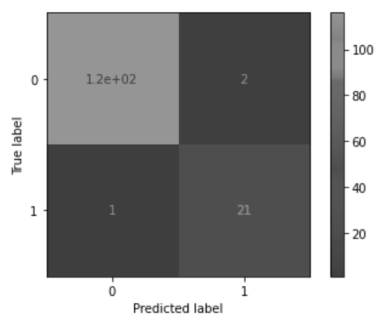

FIGURE 3. CONFUSION MATRIX TO UNDERSTAND THE PERFORMANCE ACCURACY

C. Employee Attrition Predication

As in the employee attrition part separate machine learning module developed by using supervised learning methods. A sample data set was chosen in the process of developing a predictive ML model for the employee Attrition. The IBM HR analytical dataset [6] included various important features such as average of monthly hours, number of projects, years spent in the company, work accident whether employee received a promotion. There were total of nine features. Then the module accuracy has been tested using few algorithms and the most exact result of 88.88% is generated by random forest algorithm. Additional four algorithms; support vector machine, k-neighbors classifier, bagging classifier and extra-trees classifier. The Random Forest algorithm is chosen as base algorithm for employee attrition prediction with its high accuracy level. To build this module, The program will need an old dataset with necessary features and then train and construct the module. Then the end user will need to enter all the values for the necessary fields afterwards outcome is presentable in humans' Readable interface of GUI's with descriptions of “estimated service years of a particular employee”. Then determine the retention factors to retain a valuable employee. Predicting employee attrition provides a visual way to compare turnover and engagement of employees across the entire company.

D. Employee Details Verification

Roles associated with the employee details verification can be divided as below,
1. Employer/Manager- Registers the employee to the blockchain and adds relevant employee records and verifying certificates validity.
2. Employee/Applicant- Shares the QR code attached service letter and hashes of the certificates with the recruiting organizations.
3. Verifier/Recruiter/Contributor- Contributes and verifies the relevant employee records by scanning the QR code.

To develop the employee details verification using blockchain, this component makes the use of Ethereum, MetaMask, Web3.js, Infura and smart contracts written in Solidity.This solution uses four main smart contracts named EmployeeFactory, Employee, CertFactory and CertAdd contracts.Those are written according to the requirements of the system. EmployeeFactory Contract is used to add employees to the blockchain by calling the Employee Contract and Employee Contract is used to add records of the prospective employee to the blockchain.

The authorized employers (managers) can verify the candidate certificates by inputting the hash of the relevant certificate. The system can identify the hashes of the certificates which are stored in the blockchain. Also, some other additional details regarding the certificate verification can be obtained through the system. Furthermore, the authorized employers can register the employees into the block-chain. Once the registration of an employee is done specific blockchain address will be generated. Next the authorized employers can add prospective employee records to the blockchain. This record included with details of the employee such as full name, position, department, company, and the description of the working period. Once the record is added to the block-chain the hash of the record will be generated. The QR code for the added employee records can be generated once an employee going leave the company. Once the resigned employee (applicant) applying for a job, her or she can share the QR code attached service letter with the recruiting company. Once the recruiter received the service letter, he or she can verify the service letter by scanning the QR code using online QR scanner.

V. RESULTS AND DISCUSSION

The implemented system was tested for different kinds of scenarios to identify strengths and weaknesses. Employee performance evaluation and attrition prediction were tested against a dataset that was taken from the published IBM HR analytical dataset. Smart contracts that are used for employee details verification were tested with mocha testing framework. Results are further discussed comprehensively in the below subsections respectively.

A. *Data extraction and visual representations*

Data extraction from resumes was somewhat challenging since it takes more time and it consumes more computing power to process many documents. When considering a single document, it can be either in plain text format or in image format. The dataset contained both types of resumes and normal text extraction methods were not practically applicable to extract data from those resumes. So each resume was converted into an image and then the data was extracted using OCR and Spacy NLP to avoid that problem. Then Genism LDA unsupervised machine learning model has been used according to Hoffman method [13] to rate resumes. According to the Figure 4, the rating was acceptable since the model rates resumes well by considering relevant skills of a particular job.

	job_role	rating	skills	totalWorkExperience
940	Financial Assistant	7.05	[Due diligence, Legal, Banking, Documentation,...	8 years 11 months
237	Financial Assistant	5.81	[Programming, Math, Communication, Technical, ...	0 months
450	Financial Assistant	5.28	[Teaching, English, Audit, Documentation, Fina...	5 years 0 months
951	Financial Assistant	8.25	[Budget, Internal controls, Controls, Banking,...	15 years 4 months
1294	Financial Assistant	7.73	[Strategy, Research, Financing, Healthcare, Fi...	0 months
1141	Financial Assistant	7.30	[Budget, Controls, Banking, P, Staffing, R, Op...	2 years 2 months
1050	Financial Assistant	9.73	[Media relations, Due diligence, Rfps, Pr, Ana...	-16 months
787	Financial Assistant	5.17	[C++, Jsp, Php, Android, Website, Email, Softw...	0 months

FIGURE 4. RATINGS PROVIDED BY THE LDA MODEL

The fixed-keywords method was more in-line with how humans select resumes. But the dynamic-keywords method learns the underlying keywords from the entire set of resumes, with the assumption that the resume dataset is generated by some hidden keywords. There was a high variance in dynamic-keyword method because it uses the Gibb's Sampling. But the variance can be minimized by ensembling. Two passes of the data has been passed since the distribution should be be stabilized. For this research, paralleled computing has not been used with multi-core machines to speed up LDA model training. Visualizing the relation between employees and skills in a high dimensional space was challenging since it consumes more memory and processing power in run-time. The resulting visualizations can be used in HR decision making process since it visualizes the relations in the whole workforce and related skills in a human understandable manner.

B. *Candidate performance prediction*

The objective of this system component is to predict and analyse the employee performance in a working environment to find the most appropriate employees for the right jobs at the right time and help decision-makers and HRM to decisions to keep up with the competitive advantage. This has been achieved with different ML techniques to produce 3 different predictive models. The accuracy of these ML model was evaluated and measured using 4-fold cross validation method. By the results shown in the Table I, each Machine Learning model generated was able to produce an accuracy greater than the moderate accuracy of 70%. In all three experiments using three models, they produced a satisfactory accuracy.Out of three experiments, the Na¨ıve Bayes classification algorithms produced the highest accuracy of all. These models have been fine-tuned with their respective hyper parameters to generate the optimal output with highest accuracy that can generate.

This research component has discovered the most significant set of attributes to determine and analysis the employee performance in a firm. This can be used by decision-makers and HR management personnel to take necessary actions where employee performance is used, hence increasing the productivity and to keep up with the constantly changing competitive advantage. The process of classification is carried out with three algorithms namely, Decision Tree (DT), Random Forest (RF) and Naive Bayes (NB) [14]. In the staging of training the models, the selected set of variables were subjected to the three ML models namely, DT, RF and NB. Table I shows the percentage of accuracy and the value of hyper-parameters used to tune each model to yield the maximum accuracy that model can produce. With reference to the Table I, the DT algorithm produces an optimal accuracy of 85.03% with the hyper-parameter value *ccp alpha* as 10, and RF produces an optimal accuracy of 84.69% with the hyper-parameter value *ccp alpha* as 0.1% and finally NB – Gaussian produces an optimal accuracy of 98.57% . Here the hyper-parameters are not set for the Gaussian as it produces a higher accuracy and if the hyper parameters were specified, the previous probabilities of the classes will not be adjusted automatically according the data samples.

TABLE I: PERCENTAGE OF MODEL ACCURACY

No.	Technique	Hyper-parameter value	Accuracy
1	Decision Tree	ccp alpha (10)	85.03%
2	Random Forest	ccp alpha (0.1)	84.69%
3	Naïve Bayes (Gaussian)	-	98.57%

To know if the model is working and producing the results as it is intended to, 20% of the dataset was taken as test data sample and validated the predictions to find the extent of accuracy. This way it was possible to find that each tuned model was able to output the higher accurate results as as shown in Table I. Furthermore, set of real values were extracted from a different dataset to test the model's accuracy. Most of the obvious attributes were there and some attributes which are not making a significance sense (such as name, employee id) were manually added to do the prediction and it turned out the model is performing well and producing the exact result as mentioned in that dataset.

C. *Candidate attrition prediction*

Employee attrition prediction process was carried out by using few algorithms such as RF, Support Vector Machine (SVM), k-Neighbors classifier (KNN), Bagging Classifier (BAG) and Extra-Trees (ET) classifier and they produced different accuracy levels as shown in Table II. The dataset included nine features, 2800 raw records and two categorical features [15]. Algorithms were tested using more than 2500 unique employee records.

The results of this section demonstrates the RF classifier's dominance in terms of precision and predictive performance.

It is successful when used with its best configuration approach that gives precise results given the noise in the data set, which is a significant machine learning problem and the algorithms. The authors also encourage the use of the RF classifier to reliably forecast employee turnover in a company, which helps HRM to reduce the business risk and take appropriate steps to keep workers who are expected to be at risk of leaving the company. Hence this will achieve the objective of minimizing the financial loss of the organization.

TABLE II: PREDICTION MODEL PERFORMANCE MEASUREMENTS

Algorithm tested	SVM	KNN	BAG	RF	ET
Accuracy	80.50%	16.7%	87.2%	88.88%	88.02%

D. Employee details verification

The objective of this component is to prevent fraud in candidate employee records including their certificates and to verify them in real time rather than using manual verification methods. The resulting system can verify the certificates of the candidate by providing the hash of the relevant certificate. If the certificate hash is valid which means the certificate is already in the blockchain and system will identify it as a valid hash. If the hash is incorrect, system will identify it as invalid.

TABLE III: MOCHA TEST CASES AND THE RESULTS

Test Case	Result
Deploys the factory and employee contract	Passed
Marks the caller as manager of the employee	Passed
Adds manager to the manager mapping.	Passed
Allows people to contribute and mark them as contributor	Passed
Requires the minimum contribution when contributing	Passed
Allows manager to create a record	Passed

In order to rapidly test the accuracy of smart contracts without doing manual testing, custom Mocha test runner is used with Ganache local test network with the use of Web3. Ganache local node is used to do the testing. Table III shows the test cases and the results of the written six test cases. Since the written all test cases were passing, it can be ensured that the smart contracts are working according to the requirements. This method is useful to reduce the time spent for official document verification as well as to make easiness of sharing official records.

VI. CONCLUSION AND FUTURE DIMENSIONS

The results of studying four main subsystems mentioned above can be used to define the accuracy of the overall system with respect to the data set that are being used. Implementing Smart Human Resource Management System as well as achieving above mentioned objectives through the system are met successfully. As future dimensions, to extract data from employee resumes, a NER approach would work much better than a rule-based approach if can get a labelled resume data set suitable for the company. Big companies with many employees would benefit very well with such an approach.

However, rule-based approach still works best for smaller companies with a smaller set of resumes if resumes contain relevant information. The proposed model on Employee Performance aids decision-makers and HRM to take instantaneous decisions about the employees for the growth of the organization as it focuses on the capabilities of employees and generating a rating-based analysis on their performance. Furthermore, this approach will be of service to determine the right employee to carry out the right job at the right time making the organization to keep up with the competitive advantage. Employee performance analysis is, as of now, only confined to the specified set of domains or job roles. Mostly the people related to Information Technology domain. This study can be improved to extend for other professional domains to cover another set of job role in the industry. The product features can be extended with new visualizations and predictions.

REFERENCES

[1] Chen Zhu, Hengshu Zhu, Hui Xiong, Peng liang Ding, and Fang Xie,“Recruitment market trend analysis with sequential latent variable models”,Proceedings of the 22nd ACM SIGKDD International Con-ference on Knowledge Discovery and Data Mining, Aug. 2016, doi:10.1145/2939672.2939689.

[2] J. Chen, K. Li, Z. Liu, T. Zhang, W. Wen, Z. Song, Y. Wang, Y.Jin, and T. Huang, “Data analysis and knowledge discovery in web recruitment—based on big data related jobs”, 2019, pp. 142–146.

[3] V. Mohanraj, R. Lakshmipathi, M. Chandrasekaran, J. Senthilkumar, and Y. Suresh, “Intelligent agent based talent evaluation engine using a knowledge base”, 2009, pp. 257–259.

[4] International Human Rights Instruments,International Recruitment Integrity System, 2020. [Online]. Available: https ://iris.iom.int/iris-standard. [Accessed 2020-02-21]

[5] Singhal A. and S. Pavithr R., “Degree certificate authentication using qr code and smartphone”,International Journal of Computer Applications, vol. 120, 2015, pp. 38–43.

[6] Ghazali O. and Saleh O.S, “A graduation certificate verification model via utilization of the blockchain technology”,Telecommun. Electron.Comput. Eng. JTEC, vol. 10, 2018, pp. 29–34.

[7] M. Nickel, K. Murphy, V. Tresp, and E. Gabrilovich, “A review of relational machine learning for knowledge graphs”,Proceedings of the IEEE, vol. 104, no. 1, 2016, pp. 11–33.

[8] Alao D. Adeyemo A. B, “Analyzing employee attrition using decision tree algorithm, computing”,Information Systems and Development Informatics, vol. 1, 2013, pp. 17–28.

[9] Jantan H. Hamdan A. R. Othman Z. A., “Human talent prediction in HRM using c4.5 classification algorithm”,International Journal on Computer Science and Engineering, vol. 2, 2010, pp. 2526–2534.

[10] Al-Radaideh Q.A. Al-Nagi E., “Using data mining techniques to build a classification model for predicting employees performance”,International Journal of Advanced Computer Science and Applications,vol. 2, 2012, pp. 144–151.

[11] Desouki M. S. Al-Daher J, “Using data mining tools to improve the performance appraisal procedure, hiast case”,International Journal of Advanced Information in Arts, Science and Management, vol. 2, 2015.

[12] Chris Dyer Rajarshi Das Manzil Zaheer, “Gaussian LDA for topic models with word embeddings”,d Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, vol. 7, 2015, pp. 795–804.

[13] Thomas Hofmann, “Unsupervised learning by probabilistic latent semantic analysis”,Machine Learning, vol. 42, pp. 177–196, Jan. 2001, doi: 10.1023/A:1007617005950.

[14] Surjeet K. Y. Brijesh B. Saurabh P., “Data mining applications: A comparative study for predicting student’s performance”, International Journal of Innovative Technology and Creative Engineering, vol. 1, 2011, pp. 13–19.

[15] Keggle, IBM HR Analytics Employee Attrition and Performance, 2020, [online] Available: https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset, [Accessed 2020-02-21]