

Employee Performance Prediction Using EPP Framework

M. Karthika^{*}, Dr. T. Meyyappan

Department of Computer science, Alagappa University, Karaikudi, Tamil Nadu, India

ABSTRACT

Article Info Volume 6, Issue 6 Page Number: 297-302 Publication Issue : November-December-2020 In the today's industrial world, every company's growth is depends on their employees. The company achievements are completely based on the employees in the organization. The employees' performances are measured by the targets and achievements. But some external and internal factors affect the employees' goals and achievements. Hence, the company has to find the performance of every employee and make proper solutions to improve the performance. This research work proposes a fully automated framework which can perform deep analysis of employees' performance and job fitness using data mining and prediction methods.

Article History

Accepted : 10 Dec 2020 Published : 24 Dec 2020 **Keywords :** Employee Performance, Prediction, Hybrid, Naïve Bayes, EPP, Classification

I. INTRODUCTION

Employees are the important resource of every company and the company achievement is depends on the employees' of the company. In the recruitment stage, the companies are hiring employees by based on the needs of the company. [1] On the other hand, every employee selects their job by based on various traits like working environment, salary, promotions, distance and some other factors. These factors can improve or worsen the performance of the every employee in every organization. [2]

Every company has a challenge to discover the performance of every employee and check the job fitness. It is a continuous and regular task of the human resource department. Most of the companies are conducts a survey with the internal and external factors of employees and the goals and achievements of employees. Based on the survey results, the company decides the job fitness of every employee and further activities like training, promotions, job relief and retirement. [3] The survey data is a huge collection of information and it consumes high processing time when the analytics process is manual.

The proposed model is a hybrid processing framework which includes both data mining models and predictive models. The model accepts raw data or survey data and makes the data into analytical data using data mining techniques. The model uses Naïve bayes to classify and predict the job fitness and results the performance and fitness summary to the

Copyright: © the author(s), publisher and licensee Technoscience Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

company. The hybrid framework model process the data with data mining approach and performance predictions simultaneously.

II. RELATED WORKS

Paper titled "Employee's performance analysis and prediction using KMeans clustering and Decision tree algorithm", authored by Ananya Sarkar, S.M. Shamim, Dr. Md. Shahiduz Zama and Md. Mustafizur Rahman, Published in 2018. The authors proposed a hybrid approach with K-Means Clustering and decision tree algorithms. The K-means algorithm involves in partitioning the employees' based on their performance and decision tree helps to find poor performance employee to apply certain decisions. [4]

Paper titled "A Proposed Model For Predicting Employees' Performance Using data mining techniques", authored by Mona Nasr, Essam Shaaban and Ahmed Samir, Published in 2019. The authors implemented the classification model to find the performance of employees' in MOCA. The authors used the Data mining classification model to generate a performance model and the Decision tree, Support vector Machine and Naïve Bayes models are implemented to get the important factor which affects the performance. [5]

Paper titled "Performance Predicting in Hiring Process and Performance Appraisals using Machine Learning", authored by Ali A. Mohmoud, Walid A. Salameh, Tahani AL. Shawabkeh and Ibrahim Al Amro, Published in 2019. The authors proposed an artificial intelligent model for hiring new employees based on their previous performance by analyzing the historical achievements and conditions of the new employee. [6]

Paper titled "Prediction of Employee Attrition Using Data Mining", authored by R. Shiva Sankar, J. Rajanikanth, V.V. Sivaramaraju and K VSSR Murthy. The authors proposed a multileveled employee attrition prediction approach. The model includes various classification techniques Naïve Bayes, Support Vector, KNN, Decision tree, Logistic regression and Random Forest. The model can predict the best employee by based on the environment, performance and job level. [7]

III.PROPOSED MODEL

Prediction of employees' performance is an important aspect and it is mandatory in every organization. The human resource department is responsible to measure, maintain and analyze the employees' achievements and goals. This is the primary data source and the primary data is denoted as raw data. The proposed model suggests the HR department also add the following featured data for every employee.

Table 3.1 – Featured Data List

EmpCode
Compatible With Job
Compatible With Salary
Distance Compatibility
Compatible With Co-Workers
Compatible Environment
Compatible With Transport
Satisfied in Salary Hike
Satisfied in Promotions
Satisfied in Training
Satisfied in Designation
Overall Performance Grade
E – Excellent A- Average P- Poor

The primary data and featured data are collected from HR department and Questionnaire approach. In this research work the data are collected from a Software development and IT enabled services offering company which has multiple branches in various cities.

The primary and featured data are merged by using a worksheet application. The proposed framework is a hybrid model which can load the worksheet and convert the raw data into noiseless data and predict the employee performance level using the naïve bayes classifier. The proposed model's algorithm is called as EPP Classifier algorithm. The EPP classifier algorithm is a hybrid algorithm which includes data mining and prediction models.

The EPP is divided into two levels. The first level is knowledge discovery using data mining and the second one is prediction using naïve bayes. The first level includes the Data cleaning stage, data preprocessing and transformation. The data cleaning stage loads the actual data from the worksheet and remove duplicate columns, records and imprecise data. The data preprocessing stage identifies the target data and creates row counts, column counts and ensures the data readability. The transformation stage process the data transformation into dataset or data table.

The second level is the classification based prediction using naïve bayes. This stage has the filters like employee designation, employee code and branch. Based on the filter value, the model identifies the exact patterns in the dataset and compares with the relevant patterns in the dataset. The naïve bayes probability is calculated between the patterns and the overall employee performance is predicted by the basis of naïve bayes results.





IV.PROPOSED ALGORITHM

The EPP framework is implemented by using the two stage algorithm called EPP algorithm. The EPP algorithm applied in this research work is given below:

Step 1: Begin

Step 2: Load EPP Framework

Step 3: Load Raw Data File (Worksheet)

Step 4: Internal Conversion into a worksheet previewStep 5: Detect improper data with data cleaning unitand eliminate

Step 6: Detect data redundancy using the data cleaning unit and eliminate

Step 7: Identify the primary and featured data from the filtered data set

Step 8: Complete the data migration and transform the results into a database table

Step 9: Load prediction filters with the values from the dataset

Step 10: Employee wise performance prediction is done by the result of naïve bayes score calculations with relevant employees which satisfy the same criteria.

P(C|X) = (P(X|C) P(C)) / P(X)

Input: Employee Attributes

Output: Prediction result of Performance

X: Employee attributes

C: Performance Grade

Step 11: Designation wise prediction results are prepared by the individual performance scores of every employee in the same designation.

P(C|X) = (P(X|C) P(C)) / P(X)

Input: Designation Wise Employee Attributes one by one

Output: Prediction result of Performance

X: Employee attributes

C: Performance Grade

Step 12: Branch wise performance prediction report is generated by the score calculations of employees' in the same branch. Performance wise chart reports are generated for the whole branch.

P(C|X) = (P(X|C) P(C)) / P(X)

Input: Branch Wise Employee Attributes one by one Output: Prediction result of Performance X: Employee attributes C: Performance Grade Step 13: End of EPP algorithm

V. EXPERIMENTAL RESULT

The experimental results of the proposed EPP framework are illustrated below:

					Data Sheet - Mphi - M	Acresoft Lacel				
Horse In	sert Page Leyo	ut Formulas Dal	ts Review View							e = 1
A CAR	Calibri	$\gamma \Pi \rangle \simeq \Lambda' \Lambda'$	= _ ₹- 5	Wrap Test	General •	- No. 😺	Normal Rad		P 🗊	I Autofum * 27 0
ifte of Format Pai	B I U	- 🖽 - 💁 - 📥 -	新新市 化化 医	Merge & Center *	5 - % + 38 43	Conditional Format	Good Neutral	= Insert	Delete Format	Q Orar Satt & Fin
Cipbeard		Font G	Alignment		Number G		Styles		Cells	Editing
A2	• 6	6 MT2MA2014045								
A		c	D	1	F		G	н		1
Instate	Working tranch	Designation	Compatible with Job	Compatible W3	h Salary Distance Con	and billy Conneth	e with Co-Workers Com	natible Instrument	Connetible	With Transport Satisfie
MT2MA2054045	Madurai	Software Developer	YES	NO	YES	YES	YES		YES	YES
MT2MA2019012	Madural	Software Developer	YES	YES	NO	YES	YES		NO	NO
MT2MA2017002	Madural	Interface Designer	YES	YES	YES	NO	YES		YES	NO
MT2MA2029003	Madural	Customer Support	NO	YES	YES	YES	YES		NO	YES
VT2MA2018005	Madural	Senior Tester	YES	NO	NO	NO	NO		YES	NO
/T2MA2017004	Madurai	Programmer	NO	YES	YES	YES	YES		NO	YES
12042034MSTN	Madural	Programmer	YES	NO	NO	NO	NO		YES	NO
AL5502012031	Bangalore	Project Manager	YES	NO	YES	YES	YES		YES	125
rT28G2019014	Bangalore	Team Leader	YES	YES	NO	YES	YES		NO	NO
172862019028	Bangalore	TeamLeader	YES	YES	YES	NO	YES		YES	NO
172862017019	Bangalore	Programmer	NO	YES	YES	YES	YES		NO	125
172862018034	Bangalore	Programmer	YES	NO	NO	NO	NO		YES	NO
172862019016	Bangalore	Programmer	NO	YES	YES	YES	YES		NO	YES
172862019057	Bangalore	Programmer	YES	NO	NO	NO	NO		YES	NO
172862018068	Bangalore	Programmer	YES	NO	NO	NO	NO		YES	NO
172862017047	Bangalore	Senior Tester	NO	YES	YES	YES	YES		NO	YES
T28G2019051	Bangalore	Senior Tester	YES	NO	NO	NO	NO		YES	NO
12062015015	Chengalpattu	Interface Designer	YES	NO	YES	YES	YES		YES	YES
172CG2018057	Chengalpattu	Software Developer	YES	YES	NO	YES	YES		NO	NO
172062019002	Chengalpatty	Interface Designer	YES	YES	YES	NO	YES		YES	NO
172052016019	Chengalpatty	Senior Tester	NO	YES	YES	YES	YES		NO	YES
172062019035	Chenzaloattu	Software Developer	YES	NO	NO	NO	NO		YES	NO
12062019074	Chengalpatty	Team Leader	NO	725	YES	YES	YE5		NO	755
172052019011	Chengalpattu	Project Manager	VES	NO	NO	NO	NO		YES	NO
T2KD2018018	Keraikudi	Software Developer	YES	NO	YES	YES	YES		YES	YES
TZKD2015002	Keraikudi	Senior Tester	YES	YES	NO	YES	YES		NO	NO
T2KD2018014	Karaikudi	Interface Designer	YES	YES	YES	NO	YES		YES	NO
T2KD2015007	Keraikudi	Software Developer	NO	YES	YES	YES	YES		NO	YES
TZKD2019004	Karaikudi	Interface Designer	YES	NO	NO	NO	NO		YES	NO
· · Sheet1	Sheet2 / Sheet	1.97		1		10	. · · ·			
	_					_		Coun	E 14 10000	1 100% (e) 0
6	📋 🖸	🦇 🔛	A 🖸	(Mi)						 4) N 6/23/26

Fig 5.1 Primary Data and Featured Data

	A	В	L	D	
1	EmpCode	WorkingBranch	Designation	Compatible With Job	Compatible
2	MTZMA2014045	Madurai	Software Developer	YES	NO
3	MTZMA2018012	Madurai	Software Developer	YES	YES
4	MTZMA2017002	Madurai	Interface Designer	YES	YES
5	MTZMA2019003	Madurai	Customer Support	NO	YES
6	MTZMA2018005	Madurai	Senior Tester	YES	NO
7	MTZMA2017004	Madurai	Programmer	NO	YES
8	MTZMA2019015	Madurai	Programmer	YES	NO
9	MTZBG2017031	Bangalore	Project Manager	YES	NO
10	MTZBG2019014	Bangalore	Team Leader	YES	YES
11	MTZBG2019028	Bangalore	Team Leader	YES	YES
12	MTZBG2017019	Bangalore	Programmer	NO	YES
13	MTZBG2018034	Bangalore	Programmer	YES	NO
14	MTZBG2019016	Bangalore	Programmer	NO	YES
15	MTZBG2019057	Bangalore	Programmer	YES	NO

Fig 5.2 Closer View

The raw data should be prepared with the primary data and featured data. The primary data is collected from the every branch of the target company and the featured data is collected by conducting the survey between the employees and administrators. The data are merged into a single worksheet using any spreadsheet application.

ata File	Proune Total Employee Records
arite	Browse Total Employee Records:

Fig 5.3 EPP Framework

ata File	Browse Total Employee Records:
¥2 Open Sola III Destey ►	- 4 Search Desitop
Organize New folder	# • 🗆 🛛
Construction Development Developm	A A Convert Data 1 EmpCole Work Convert Data 2 MT2AAA201005 Mada MT2AAA201022 Mada 4 MT2AAA2010205 Mada MT2AAA201005 Mada 6 MT2AAA201005 Mada MT2AAA201005 Mada 7 MT2AAA201005 Mada MT2AAA201005 Mada 8 MT2AAA201005 Mada MT2B04201015 Mada 9 MT2B04201015 Mada MT2B04201015 Mada 10 MT2B05201018 Mada MT2B04201015 Mada 10 MT2B05201018 Mada MT2B04201015 Mada 10 MT2B05201018 Mada MT2B04201015 Mada

Fig 5.4 Load Raw Data

The EPP framework has the utility to load the raw data from the spreadsheet application. After the selection of raw data file it converts into a worksheet preview of actual data. The convert feature begins it process after the successful load. The convert feature does the data cleaning, data preprocessing and data transformation. After the successful completion of data transform the prediction filters are loaded with values from the data set.

ata F	ile C:\Users	\Ganesh\Deskto	op\Data Sheet - Mpi	hilxlsx			Browse Total E	mployee Records C	ount: 29
	EmpCode	WorkingBranch	Designation	Compatible With Job	Compatible With Salary	Distance Compatibility	Compatible with Co-Workers	Compatible Environment	Compatible With Transp
1		Madurai	Software Developer	YES	NO	YES	YES	YES	YES
1	ATZMA2018012	Madurai	Software Developer	YES	YES	NO	YES	YES	NO
1	4TZMA2017002	Madurai	Interface Designer	YES	PredictPerformV1	× 1	NO	YES	YES
1	ATZMA2019003	Madurai	Customer Support	NO	-		YES	YES	NO
1	ATZMA2018005	Madural	Senior Tester	YES	and the second	100	NO	NO	YES
1	ATZMA2017004	Madural	Programmer	NO	Successfully Saved	ed	YES	YES	NO
1	ATZMA2019015	Madurai	Programmer	YES			NO	NO	YES
1	ATZBG2017031	Bangalore	Project Manager	YES			YES	YES	YES
1	4TZBG2019014	Bangalore	Team Leader	YES			YES	YES	NO
1	AT28G2019028	Bangalore	Team Leader	YES	-	100	NO	YES	YES

Fig 5.5 Completion of Level1 in EPP





The various stages of prediction summary and results of EPP algorithm framework are shown below:



Fig 5.7 Employee Wise Prediction









VI. CONCLUSION

The EPP (Employee Performance Predictor) framework algorithm is proposed to offer a quick and accurate prediction of employees' performance. The naïve bayes prediction results are accurate and very faster than other prediction approach. This research work uses the data from a software development company having various branches in familiar cities. This proposed work is tested with the live data of employees' in the company. The EPP framework lets the human resource department of any company to predict the performance of their employees' by preparing a data sheet with the prescribed format.

VII. REFERENCES

- Nitin Kumar, Sumika Jain and Kuldeep Chauhan, "Knowledge Discovery From Data Mining Techniques", International Journal of Engineering and Technology, ISSN: 2278 -0181.
- [2]. S.E. Viswapriya, "Survey on Predicting Performance of An Employee Using Data Mining Techniques", International Journal of Engineering research and technology, ISSN: 2278 - 0181.
- [3]. Pooja Akulwar, Sujata Pardeshi and Arvind Kamble, "Survey On Different Data Mining Techniques For Prediction" International Conference on I-SMAC, ISBN: 978-1-5386-1442-6.
- [4]. Ananya Sarker, S.M. Shamim, Dr. Md. Shahiduz Zama and Md. Mustafizur Rahman, "Employee's performance Analysis and prediction Using K-Means Clustering and Decision Tree Algorithm", Global Journal of Computer Science and Technology, ISSN: 0975 - 4172.
- [5]. Mona Nasr, Essam Shaaban and Ahmed Samir, "A proposed model for Predicting Employees' performance Using Data Mining techniques", International Journal of Computer Science and Information Security, ISSN: 1947 - 5500.
- [6]. Ali A. Mohmoud, Walid A. Salameh, Tahani AL. Shawabkeh and Ibrahim Al Amro, "Performance Predicting In Hiring Process And Performance Appraisals Using Machine Learning", International Conference On Information and Communication System, ISBN: 978-1-7281-0045-6.
- [7]. R. Shiva Sankar, J. Rajanikanth, V.V. Sivaramaraju and K VSSR Murthy, "Prediction of Employee Attrition Using Data Mining".

Cite this article as :

M. Karthika, Dr. T. Meyyappan, "Employee Performance Prediction Using EPP Framework", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 6, pp. 297-302, November-December 2020. Available at

doi : https://doi.org/10.32628/CSEIT206654 Journal URL : http://ijsrcseit.com/CSEIT206654