



## BENEDICTION OF DATA MINING APPLICATIONS SPOTLIGHTING ON THE INDIVIDUAL EMPLOYEE AT ORGANIZATIONAL LEVEL AND ENLACING THE PROBLEMS ACCOSTED WHILE USING DATA MINING

Dr.V.Rajesh\*      G.K.Pooranee\*\*

\*Head, Dept. of Management, Sree Amman Arts & Science College, Erode.

\*\*Assistant Professor, Dept. of Management, PKR Arts College for Women, Gobi.

### Abstract

Managing and developing Human resources are becoming center of attention nowadays in every organization. The importance of human potential for company increases proportionally with the speed of changes which appear in the business area, because human capital represents a basic qualitative parameter of fruitfulness of any changes. It has been scientifically proved that one of the supporting pillars which can contribute to the fulfillment of the personnel policy is the usage of Computer technologies in HR. Data mining in Human resource development in IT industry, hence a study is required to know the benefits of data mining applications and the problems faced while using data mining in HRD in the city due to the increase of IT companies in recent times.

**Key Words:** Data mining, Information Technology, Human Resource Development.

### Introduction

Topic of present research is related with Information technology Industry and deals with the application of data mining tool in Human resource development. The term data mining has its own eminent position in all modern industries. Data mining is the process of discovering meaningful new correlations, patterns, and trends by digging into (mining) large amounts of data stored in warehouses. Data mining, on the other hand, extracts information from a data base that the user did not know existed. Relationships between variables and behaviors that are non-intuitive are the jewels that data mining hopes to find. In the 21st Century, it is witnessing a series of changes and development all around the world and powerful IT companies established their branches all over the country. There is a turbulent change because of applications of modern scientific tools, and applications of internet etc. There are many contemporary changes that are happening in and around the countries which are responsible for changes in the IT industry. The present study deals with benefits of data mining applications on the individual employee level and organizational level, and the problems faced in the application of data mining.

### Objectives of the Study

1. To study the concept of Data mining in general and in particular to Human Resource Development.
2. To find out the benefits and problems of Data mining applications on the individual employee level and the organizational level.
3. To suggest better ways and means for maintaining, accessing the data bases of the organization and applications of Data mining for the continuous development.

### Research Methodology

Research methodology is the way in which research problems are solved systematically. It is a science of studying how research is conducted scientifically. In this study, we have adopted various steps to study a research problem, along with the underlying logic behind them.

**Research Design:** Descriptive research design is used in this study

**Sampling Design:** It is an appropriate method to find out the target population, parameters of interest of for the researcher, sampling frame, sampling method and sample size required for the study.

**Size of the sample:** The total number of respondents selected for this study is 560.

**Sampling technique:** Stratified Random sampling was used to collect data's from respondents.

**Data collection:** Both primary and secondary data's were used for this study.

**Data analysis:** The collected data's were analyzed using Factor analysis and Henry Garret Ranking method and the results are interpreted.

**Tools used in this study:** Factor analysis was used to line up the study constructs in to various factors and Henry Garret Ranking method was also used to explore the ranking of the study constructs.



### Limitations of the Study

- In this study only limited factors were conscripted to measure the benefits of data mining applications, more factors can also be included to elicit an in-depth knowledge.
- The findings of the study may not be applicable to the universe, as only 560 samples were taken.
- The time period of study can be extended to accrual more inputs which will relinquishes a broader view of this study.
- The study is purely based on the opinion of respondents, which may not confer us accurate results sometimes.

### Data Analysis and Interpretation

#### Factor Analysis

The benefits of data mining applications on the individual employee and organizational level were studied by measuring various benefits of data mining through 16 statements of cognitive components, affective component and co-native components. These 16 statements were chosen and classified in an orderly form, and factor analysis was employed and the detailed analysis and discussions are done at various stages.

#### DATA MINING MEASURES

S.No	Concepts
1	Matches the individual with job profile
2	Produces the skill gap
3	Measures the employee value
4	Helps talent forecasting to do right staffing
5	Supports the creation of desired job requirements
6	Identifies the Best Profile
7	Notices the hidden attributes of employees
8	Makes possibilities to create detailed metrics of individual, group employee activities
9	Supports and increases the effectiveness of the decision making process
10	Builds Models
11	Predicts and Optimize HRD functions
12	Contributes for corporate profitability
13	Supports the process of enhancing skills & knowledge of workforce
14	Helps to increase retention rate
15	Possibility of best possible Data storage
16	Helps the employer to establish" Employer Branding" for retaining talent

#### COMMUNALITIES

Va. No.	Statements	Initial	Extraction
1	Matches the individual with job profile	1.000	.544
2	Produces the skill gap	1.000	.543
3	Measures the employee value	1.000	.574
4	Helps talent forecasting to do right staffing	1.000	.646
5	Supports the creation of desired job requirements	1.000	.654
6	Identifies the Best Profile	1.000	.617
7	Notices the hidden attributes of employees	1.000	.620
8	Makes possibilities to create detailed metrics of individual, group employee activities	1.000	.570
9	Supports and increases the effectiveness of the decision making process	1.000	.675
10	Builds Models	1.000	.824
11	Predicts and Optimize HRD functions	1.000	.613
12	Contributes for corporate profitability	1.000	.583
13	Supports the process of enhancing skills & knowledge of workforce	1.000	.639
14	Helps to increase retention rate	1.000	.552
15	Possibility of best possible Data storage	1.000	.593
16	Helps the employer to establish" Employer Branding" for retaining talent	1.000	.502



**TOTAL VARIANCE EXPLAINED**

Component	Initial Eigen values			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.407	33.791	33.791	5.407	33.791	33.791	4.643	29.020	29.020
2	3.333	20.828	54.620	3.333	20.828	54.620	3.833	23.958	52.978
3	1.013	6.332	60.951	1.013	6.332	60.951	1.276	7.973	60.951
4	.762	4.764	65.716						
5	.680	4.250	69.966						
6	.650	4.066	74.031						
7	.616	3.850	77.882						
8	.582	3.640	81.522						
9	.477	2.979	84.500						
10	.441	2.758	87.259						
11	.431	2.695	89.954						
12	.381	2.384	92.338						
13	.364	2.272	94.610						
14	.331	2.070	96.680						
15	.294	1.841	98.521						
16	.237	1.479	100.000						

**Interpretation**

It was observed that the labeled “Initial Eigen values” gives the Eigen values. The Eigen value for a factor indicates the “Total Variance” attributed to the factor. From the extraction sum of squared loadings, it was learnt that the first factor accounted for a variance 5.407 which was 33.791 % and the second factor accounted for the variance 3.333 which was 20.828%. Similarly the third factor accounted for the variance 1.013 which was 6.332%. All this five factors put together showed the total percentage of the variance with 60.951.

**Determination of factors based on Eigen values**

In this approach only factors with Eigen values greater than 1.00 are retained and the other factors are not included in this model. The six components possessing the Eigen values which were greater than 1.00 were taken as the components extracted

**COMPONENT MATRIX**

S.No	Variables	1	2	3
1	Matches the individual with job profile	.636		
2	Produces the skill gap	.586	-.443	
3	Measures the employee value	.682		
4	Helps talent forecasting to do right staffing	.638	-.489	
5	Supports the creation of desired job requirements	.710		
6	Identifies the Best Profile	.661	-.421	
7	Notices the hidden attributes of employees	.661	-.422	
8	Makes possibilities to create detailed metrics of individual, group employee activities	.670		
9	Supports and increases the effectiveness of the decision making process	.501	.546	
10	Builds Models			.796
11	Predicts and Optimize HRD functions	.538	.567	
12	Contributes for corporate profitability	.535	.505	
13	Supports the process of enhancing skills & knowledge of workforce	.602	.515	
14	Helps to increase retention rate	.468	.532	
15	Possibility of best possible Data storage	.475	.579	
16	Helps the employer to establish” Employer Branding” for retaining talent	.470	.513	

Extraction Method: Principal Component Analysis. a. 6 components extract



### ROTATED COMPONENT MATRIX

S.No	Variables	1	2	3
1	Matches the individual with job profile	.692		
2	Produces the skill gap	.734		
3	Measures the employee value	.728		
4	Helps talent forecasting to do right staffing	.804		
5	Supports the creation of desired job requirements	.791		
6	Identifies the Best Profile	.782		
7	Notices the hidden attributes of employees	.781		
8	Makes possibilities to create detailed metrics of individual, group employee activities	.733		
9	Supports and increases the effectiveness of the decision making process		.604	.553
10	Builds Models			.887
11	Predicts and Optimize HRD functions		.733	
12	Contributes for corporate profitability		.754	
13	Supports the process of enhancing skills & knowledge of workforce		.772	
14	Helps to increase retention rate		.741	
15	Possibility of best possible Data storage		.769	
16	Helps the employer to establish” Employer Branding” for retaining talent		.702	

#### Interpretation

The rotated component is a result of VARIMAX procedure of factor rotation. Interpretation is facilitated by identifying the variables that have large loadings on the same factor. Hence, those factors with high factor loadings in each component i.e. values greater than 0.4 were selected. The selected factors were named separately and highlighted in the table below

#### Naming of statements extracted

Factor	Va. No.	Statements	Rotated Factor Loadings
Superlative	1	Matches the individual with job profile	.692
	2	Produces the skill gap	.734
	3	Measures the employee value	.728
	4	Helps talent forecasting to do right staffing	.804
	5	Supports the creation of desired job requirements	.791
	6	Identifies the Best Profile	.782
	7	Notices the hidden attributes of employees	.781
	8	Makes possibilities to create detailed metrics of individual, group employee activities	.733
Probable	9	Supports and increases the effectiveness of the decision making process	.604
	11	Predicts and Optimize HRD functions	.733
	12	Contributes for corporate profitability	.754
	13	Supports the process of enhancing skills & knowledge of workforce	.772
	14	Helps to increase retention rate	.741
	15	Possibility of best possible Data storage	.769
	16	Helps the employer to establish” Employer Branding” for retaining talent	.702
Efficacy	9	Supports and increases the effectiveness of the decision making process	.553
	10	Builds Models	.887



### Interpretation

In the above table the statements 1, 2, 3, 4, 5, 6,7 and 8 were grouped together as factor 1 and have been named as ‘Superlative’. The statements 9, 11, 12, 13, 14, 15 and 16 were grouped together as factor 2 and have been named as ‘Probable’. The statements 9 and 10 were grouped together as factor 3 have been named as ‘Efficacy’. Thus the factor analysis condensed and simplified the 16 statements and grouped them into 3 factors explaining 60.951% of the variability of all the statements.

### Henry Garrett Ranking Method Problems Faced in Data Mining

Data mining is applied effectively not only in business environment but also in other fields such as weather forecast, medicine, transportation, healthcare, insurance, government etc. Data mining has a lot of advantages when using in a specific industry. Besides those advantages, data mining also has its own limitations. For the purpose of the study, an attempt was made to identify the common problems faced by the selected sample respondents in the study area. For the purpose of this study, the problems in data mining such as data rich, information poor, lack of privacy, if used by non experts may lead to misinterpretations, universal accessibility, accuracy of the data, completeness of the data and reliability of data were chosen. Henry Garrett ranking method was employed to ascertain the magnitude of the problems. The details are furnished in the underneath table.

**PROBLEMS FACED IN DATA MINING**

S.No	Problems	Total score	Mean score	Rank
1.	Data rich, information poor	26318	46.996	5
2.	Lack of privacy	28137	50.245	4
3.	If used by non experts, may lead to misinterpretations and wrong judgments	31758	56.711	2
4.	Universal accessibility	30425	54.330	3
5.	Accuracy of the data	23223	41.470	7
6.	Completeness of the data	24295	43.384	6
7.	Reliability of data	31844	56.864	1

### Interpretation

It is witnessed from the above table that among the seven common issues faced by the respondents, the problem of “Reliability of data” was placed in the first position. It is followed by the problem “If used by non experts, may lead to misinterpretations and wrong judgments” which was ranked second and the problem of “Universal accessibility” was placed in the third rank. The issues like “Lack of privacy” and “Data rich, information poor” were placed in the fourth and fifth positions. Simultaneously, the problem of “Completeness of the data” was ranked in sixth position and the problem “Accuracy of the data” was ranked in the seventh position.

### Findings of the Study

#### Factor Analysis

- It is also found from the factor analysis that the three-factor solution accounted for 60.951% of the explained variance and it has been representing the benefits of data mining are grouped into factor 1, named as ‘Superlative’ and accounted for 33.791% of the total variance, factor 2 named as ‘Probable’ accounted for 20.828% of the total variance, and factor 3 named as ‘Efficacy’ accounted for 6.332% of the total variance. From the analysis, it is evident that out of 16 statements of merits in data mining applications, it were grouped into 3 component factors and were termed as Superlative, Probable and Efficacy.

#### Henry Garrett Ranking Method

- It is found from the Henry Garrett ranking technique that among the seven common problems faced by the respondents, the problem of “Reliability of data” was placed in the first position with the Garrett scores 31844 points. It is followed by the problem “If used by non experts, may lead to misinterpretations and wrong judgments” which was ranked second with the Garrett score of 31758 points.
- The problem of “Universal accessibility” was placed in the third rank by the sample respondents with the Garrett score of 30425 points. The issues like “Lack of privacy” and “Data rich, information poor” were placed in the fourth and fifth positions with the Garrett scores of 28137 and 26318 points respectively.
- Simultaneously, the problem of “Completeness of the data” was ranked in the sixth position with the Garrett scores of 24295. Lastly the problem “Accuracy of the data” was ranked in the seventh position with the Garrett score of 23223 points.



- From the analysis it is identified that the problem “Reliability of data” was the major problem in data mining highlighted by the sample respondents.

### Conclusion

Changing attitude and expectation of employees, international exposure, high-income opportunities, and improved life styles are the few key drivers influencing the application of computer and scientific tools in the present management functions. The amount of data kept in computer files and databases is growing at a phenomenal rate. At the same time, the users of these data are expecting more sophisticated information from them. Almost every field of human life has become data-intensive, which made the data mining as an essential component and produces the future trends from the past data. Data mining allows management professionals to shift their energy from managing processes to actively supporting business execution. Thus, to conclude that with the enormous growth of Information Technology companies in number and size, it forces the present Managers to innovate modern applications and management throughout the Nation.

### References

1. Alex Berson, Stephen Smith and Kurt Thearling, (2000), Building data Mining Applications for CRM, Tata McGraw Hill Publishing Company Ltd, New Delhi.
2. Arun K Pujari, (2001), Data Mining Techniques, Universities Press (India) Private Limited, Hyderabad.
3. Aswathappa k, (2005), Human Resource and Personnel Management, Fourth Edition, Tata McGraw Hill Publishing Co. Limited, New Delhi.
4. Bhambra A S, (1999), Future of Human Resource Management, First Edition, Common Wealth Publishers, New Delhi.
5. Biswajeet Pattanayak, (2002), Human Resource Management, Third Edition, Prentice Hall of India (P) Ltd, New Delhi.
6. Gary Dessler, (2002), Human Resource Management, Seventh edition, Prentice Hall of India (P) Ltd, New Delhi.
7. Jack J Philips, (1999), HRD TRENDS World wide, First edition, Gulf Publishing Company, Houston, Texas.
8. Jain J N, Ajay Bhola, (2008), Modern Human Resource Management Principles and Techniques, fourth edition, Regal Publications, New Delhi.
9. Jon M Werner and Randy L Desimone, (2006), Human Resource Development; Foundation, Framework and Application, Fifth Edition, Cengage Learning India (P) Ltd, New Delhi.
10. .Kargupta, Joshi, Siva kumar and Yesha, (2004), Data Mining Next Generation Challenges and Future Directions, PHI Learning (P) Ltd, New Delhi.
11. Margaret H. Dunham, (2008), Data Mining Introductory and advanced Topics, Pearson Education, New Delhi.
12. Nair M R R, T V Rao, (1990), Excellence through Human Resource Development, Fourth edition, Tata McGraw Hill Publishing Co. Ltd, New Delhi.
13. Pieter Adriaans, Dolf Zantinge, (2003), Data Mining, Seventh edition, Pearson Education (singapore) Pte. Ltd, Delhi.
14. Ramakrishnam, Gehrke, (2003), Data Base Management Systems, Third edition, Tata McGraw Hill Publishing Company Ltd, New Delhi.
15. Randy L Desimone, Jon M Werner, David M Harris, (2002), third Edition , Eastern Press (Bangalore) Pvt. Ltd, Bangalore.
16. .Srinivas R Kandula, (2002), Strategic Human Resource Development, Second edition, Prentice Hall of India Private Ltd, New Delhi.
17. .Subba Rao P, (1996), Essentials of Human Resource Management and Industrial Relations, First edition, Himalaya Publishing House, Mumbai.
18. Tripathi P.C, (2003), Human Resource Development, Fourth edition Sultan Chund & Sons, New Delhi.
19. Udai Pareek, T Venkateswara Rao, (1999), Designing and Managing Human Resource Systems, Second edition, Oxford & IBH Publishing Co. Pvt. Ltd, Calcutta.
20. Report of the Working Group on Information Technology Sector Twelfth Five year Plan (2012-2017), Government of India, Ministry of Communications and Information Technology, Department of Information Technology.

### Websites

1. [www.aitp.org](http://www.aitp.org)
2. [www.csi-cbe.org](http://www.csi-cbe.org)
3. [www.ijdmata.com](http://www.ijdmata.com)
4. [www.inderscience.com](http://www.inderscience.com)