PPLICATIONS OF DATA MINING AND DATA ANALYTICS IN HR MANAGEMENT – A BIBLIOGRAPHIC REVIEW

DOI: **https://doi.org/10.5281/zenodo.10802792**

**Muhammad Farhan Bukhari*
***Heena Shabbir*
****Muhammad Abu Huraira*

ABSTRACT:

As technology spreads in Human Resource Management processes, HR analytics and information systems, which have evolved from simple transaction systems to decision support systems, have been part of this transformation and have made it possible to change the services that HR provides within organizations, as well as their role. Thus, this research aims to identify the different uses or specific applications that analytics can have in the various fields of HRM to provide tools to HR professionals as well as researchers, to decide on future investments and research that allow the use of analytics in HR to move from theory to practice and continue to expand. To accomplish the aim, a systematic literature review was reviewed by applying certain inclusion and exclusion criteria. The results obtained from the analysis revealed that data analytics and data mining have been applied in all areas of HRM, except for the area of Remuneration and Benefits, where there is potential for future research. The area that has had the greatest application over the years has been the Attraction, Recruitment, and Selection of employees, followed by the area of Performance Evaluation. The applications of analytics and data mining in all areas are very varied and respond to different problems that arise in each one.

Keywords: *Data Mining, Data Analytics, HR Analytics, Human Resource Management.*

Introduction:

The Human Resources (HR) area has been transforming for some years, going from being an administrative or maintenance area to a strategic function for the business whose purpose should be to add value to it (Ulrich & Dulebohn, 2015). As technology spreads in Human Resource Management (HRM) related processes, HR analytics and information systems, which have evolved from simple transaction systems to decision support systems, have been part of this transformation and have made it possible to change the services that HR provides within organizations, as well as their role (Dulebohn & Johnson, 2013).

……………………………

** Millennium Institute of Technology and Entrepreneurship (MiTE) Email ID: mfarhanb786@gmail.com Orcid ID: https://orcid.org/0009-0009-8121-1814*
*** Mohammad Ali Jinnah University. Email ID:Heenanaz295@gmail.com *** Millennium Institute of Technology and Entrepreneurship (MiTE) Email ID: abu.huraira@mite.edu.pk*

Although many decisions require the use of judgment, these tools allow the HR area to have greater professional credibility and rigor in decision-making. For this reason, Ulrich and Dulebohn (2015) identify HR analytics as one of the 4 areas towards which future investment in HR areas should be directed.

Despite this, the use of analytics in the areas of HR has developed more slowly than in other business functions, such as marketing, where there has been a revolution in the way companies seek to understand and provide solutions to their consumers, thanks to the fact that the ability to collect, store and analyze large volumes of data has changed the way business decisions are made, those that were previously made based on instinct or experience (Russell & Bennett, 2015). According to the research carried out by Marler and Bourdeau (2017), the term "HR Analytics" appeared for the first time in the literature in 2003, more focused on the context of business intelligence than on the current definition of Analytics and, Despite the fact that in recent years the literature around the use of analytics in HRM has been increasing, it is still considered an innovation in a state of diffusion and the percentage of companies that state that they use analytics in the definition or implementation of the HR strategy is surprisingly low.

One of the reasons analytics has been slower to evolve in the HR field than in others is that HR professionals don't have the knowledge, skills, and insight to ask the right questions about HR. The data they have at hand, added to the fact that HR professionals are usually not trained in the area of analytics, and analytics teams are usually not familiar with HR (Angrave et al., 2016). Other articles highlight the lack of skills of HR professionals around analytics as one of the main barriers to progress in their incorporation (CIPD, 2013; Bersin, O'Leonard & Wang-Audia, 2013; Jones, 2014; Ulrich & Dulebohn, 2015; Bennett & Collins, 2015). On the other hand, most of the literature that addresses the issue of analytics in HR is promotional in nature and provides little information that allows ideas to be put into practice (King, 2016); focuses on answering the question: How does analytics work in HR? and presents a theoretical framework (Marler & Boudreau, 2017). Since the implementation and dissemination of analytics in the HR areas depend to a large extent on the professionals who lead these areas, it is relevant to have references and studies on the applications or uses that analytics can have in the different sub-sectors of HR areas.

Thus, in light of the above, this research aims to identify the different uses or specific applications that analytics can have in the various fields of HRM to provide tools to HR professionals as well as researchers, to decide on future investments and research, that allow the use of analytics in HR to move from theory to practice and continue to expand. Therefore, this research seeks to answer the following questions, through a Systematic Literature Review process:

1- What applications have been given to analytics in the field of HRM and in what areas?
2- What are the most used data mining and analytics techniques or tools in HR?

3- How has the use of analytics in HR evolved over the years?
4- What are the main results obtained?

1 Literature Review

1.1 Human Resources Management

The HR area arose in organizations at the beginning of the 20th century to generate value through effective management and rationalization of relations with employees, as a result of the conditions presented by the First World War, including the labor shortages and the rapid increase in wages (Ulrich & Dulebohn, 2015). After the war, the focus of HR was internal and sought to manage the challenges presented by unions and labor relations, however, in the 80s a transformation began in HRM, from an administrative area to a more core business function that contributes to organizational effectiveness (Ulrich & Dulebohn, 2015). Changes in the environment, including globalization, changes in demographics, the search for profitability through growth, changes in technology, human capital, and the organizations themselves have led to the fact that, more and more, HRM is seen as a source of competitive advantage and becomes more important in organizations (Krishnan & Singh, 2011).

Greer (2021) explained a model for HRM made up of the following six areas or fields of work:

- **Analysis and Description of Positions:** its objective is to collect the basic information of the necessary jobs in an organization to define them correctly, including the tasks to be carried out, the context in which they are carried out, the specific requirements and the type of people whom They must occupy each position.
- **Attraction, Selection, and Incorporation:** its objective is to carry out a correct selection of people, ensuring that there is a correspondence between the expectations of the candidate and the company and seeking to generate a fruitful relationship between the two.
- **Development and Succession Plans:** seeks to increase the intellectual and behavioral capacities of employees and thus ensure their growth and career advancement, with the ultimate goal of caring for the company's intellectual capital. This field is closely related to Training and Performance Evaluation.
- **Training:** seeks to make people do their jobs better, preparing them to perform successfully in their position, in the present, and also in the future.
- **Performance Evaluation:** its objective is to assess the behavior of the employee about work and thus be able to make decisions related to development, remuneration, and training.
- **Remuneration and Benefits:** its objective is to ensure equitable payment within the company, maintaining its competitiveness in the recruitment, hiring, and retention of employees.

The Information Systems used in HRM store information about the employees in each of these processes, including demographic information, about their employment history, skills and competencies, education, performance, training, and development that the employee has had. employee, disciplinary cases, participation schemes, and attitude survey results, among others (Angrave et al., 2016), which can be used to support decision-making.

According to Ulrich and Dulebohn (2015), in the future, HR must be focused on generating value for the organization and to achieve this, it must focus its investments on:

- The HR department, that is, transforming the structure of the area
- HR practices, seeking to ensure that they are aligned with the strategy and external interest groups, that they are integrated, and that they are innovative.
- HR professionals, improving and updating the skills and competencies of professionals working in the area.
- HR Analytics, which provides greater rigor in decision-making and allows the area to have greater professional credibility.

1.2 Data Analytics and Data Mining

Analytics is a method for the technological manipulation of information that provides insight into a relevant issue and allows decision-making based on evidence (King, 2016). The data analyzed can be structured, unstructured, or a mixture of both, that is, hybrid. Therefore, analytics allows you to understand a situation from a complex mix of data and find hidden value in it. According to Edwards & Edwards (2019), there are three levels of analytics: descriptive, which describes relationships between data and historical patterns; predictive, which uses current and historical data to make predictions of the future based on probabilities and possible results; and the prescriptive, which goes beyond predictions and presents decision alternatives and their impacts. Statistics is one of the main components of analytics but analytics also involves understanding the problem, its interactions, and relationships.

Data mining is an analytical process that consists of analyzing large volumes of data that are available from different sources to transform them into useful information, that is, it is the search for new, valuable, and non-trivial information, in large volumes of data and it is the result of the joint effort between the human abilities to describe the problems and the goals and the computational abilities of search. The data mining process has 5 steps: 1) define the problem, 2) collect the data, 3) pre-process the data, 4) estimate the model and 5) interpret the model and draw conclusions. Like analytics, data mining is also classified as descriptive, when it seeks to identify patterns and describe the information so that it can be interpreted; or predictive, when it seeks to predict future or unknown values of other variables of interest (Pruengkarn, Wong & Fung, 2017; Kantardzic, 2020). Data mining techniques or methods are classified into:

Classification Techniques: These allow data to be classified into one or more previously identified categories. Some of the most used classification techniques are (Pruengkarn, Wong & Fung, 2017; Kantardzic, 2020)

- Those based on **Bayesian Decision Theory**: The objective of this theory is to minimize the probability of making a wrong decision.
- Those based on **Classification Rules**, which use IF-THEN rules to predict the class. Decision tree techniques, evolutionary algorithms, and genetic algorithms are included in this category.
- **Neural Networks**, which learn from past experience and mistakes and deliver a result based on comparison with an established limit or threshold.
- **Support Vector Machines (SVM)**, whose objective is to find an optimal hyperplane in the data space, based on the principle of structural risk minimization (SRM) and statistical learning theory.
- **K-Nearest Neighbor (KNN):** This technique seeks to identify the optimal value of K that minimizes the error in calculating the distance between the test data and the training data.
- **Reasoning based on Cases (CBR):** This technique stores the data as cases. When there is a new case to classify, it first looks to see if there is already a similar case to determine the result. If it does not exist, it looks for those that are similar to the new case.

Grouping Techniques: They allow the identification of a series of categories in which the data can be grouped (Kantardzic, 2020). It is similar to the classification techniques since it groups the data into categories with the difference that in this case the categories are not previously identified. Some of the clustering techniques are (Pruengkarn, Wong & Fung, 2017):

- **K-Means:** It is one of the simplest unsupervised algorithms and is used to distinguish between similar and dissimilar data and generate meaningful clustering.
- **Fuzzy C-Means:** This technique is based on the principle that each piece of data can belong to more than one cluster and in each one it has a different value between 0 and 1, so that the sum of all is 1.

Association Rules: They are used to determine the relationship between the elements of a data set. The rules must satisfy a lower limit of support and a lower limit of confidence that are defined by the user or by experts in the area. Some association rule techniques are:

- **The Algorithm a Priori:** It is an association rule technique that is used to analyze frequent items based on prior knowledge of their properties.
- **Multidimensional Association Rules:** They are used to analyze attributes that have more than one dimension at the same time and quantitative association

rules, refer to a special type of rules in the form of X → Y in which at least one of the variables must be numeric.

- **Discovery of Sequences:** It is used to identify patterns in the events that are organized within a data set.

Regression Techniques: These allow to identification of a prediction function to predict the value of a variable (Kantardzic, 2020). There are several regression methods, including time series analysis, nonparametric regression, robust regression, ridge regression, nonlinear regression, deep learning, or machine learning (Pruengkarn, Wong & Fung, 2017).

Other data mining techniques include (Kantardzic, 2020):

- Summary (summarization), which allows to find a compact description of a series of data
- Dependency models, which allow to identification of a model that describes the dependencies between variables
- Detection of changes and deviations, which allows the identification of the most significant changes in a series of data.
- Genetic algorithms, which combine computational evolution methods with optimization algorithms.
- Rough Set Theory (RS), which is used to find structural relationships between imprecise or noisy data in the classification space.
- Those with a Fuzzy focus as Fuzzy Set Theory allows working with vague or inaccurate data

2 RESEARCH PROTOCOL/ METHODOLOGY

This section describes the research protocol or the methodological steps to answer the research questions.

The research method that was used to answer the research questions presented above is systematic literature mapping. A systematic literature mapping is a methodology to identify, evaluate, and interpret the available literature on a particular topic using already existing and published primary studies (Siddaway, Wood, & Hedges, 2019). Unlike a simple literature review, a systematic review must be rigorous and fair and is carried out following a previously defined search strategy that includes a planning stage, an execution stage, and a reporting stage (Siddaway, Wood, & Hedges, 2019). A systematic review of the literature makes it possible to synthesize the evidence on a topic, identify gaps in current research, and suggest areas for future research, as well as provide a frame of reference for it (Siddaway, Wood, & Hedges, 2019).

The research protocol is described below:

- **Bibliographic database:** Scopus.
- **Type of documents:** articles published in indexed journals.
- **Period analyzed:** all information available up to August 2023.

- **Search String:** In order to obtain the largest number of articles related to the use of analytics in HR, keywords related to these two areas of study (analytics and HR) were defined. Because the area of HR is vast and has many sub-areas and related processes, the search string used included keywords identified in articles reviewed in the initial preparation of the investigation.

((TITLE-ABS-KEY ("data analytics") OR TITLE-ABS-KEY ("analytics") OR TITLE-ABSKEY ("data mining") OR TITLE-ABS-KEY ("data science ") OR TITLE-ABS-KEY ("big data")) AND (TITLE-ABS-KEY ("talent employee*") OR TITLE-ABS-KEY (" employee recruit *") OR TITLE-ABS-KEY ("employee hiring") OR TITLE-ABS-KEY ("employee recruitment") OR TITLE-ABS-KEY ("Talent attraction*") OR TITLE-ABS-KEY ("HR") OR TITLE-ABS-KEY ("human resource management") OR TITLE-ABS-KEY ("career management*") OR TITLE-ABS-KEY ("career management*") OR TITLE-ABS-KEY ("career plan*") OR TITLE-ABS-KEY ("Employee Performance*") OR TITLE-ABS-KEY ("performance manag*") OR TITLE-ABS-KEY ("Employee Retention *") OR TITLE-ABS-KEY ("Employee Training*") OR TITLE-ABS-KEY ("Employee Turnover") OR TITLE-ABS-KEY ("employee engagement*") OR TITLE-ABS-KEY ("Talent Plan*") OR TITLE-ABS-KEY ("WORKFORCE PLAN*") OR (title AND abs KEY ("compensation" AND "employee")))) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (SRCTYPE , "j")) AND (EXCLUDE (SUBJAREA , "ENER") OR EXCLUDE (SUBJAREA , "MEDI") OR EXCLUDE (SUBJAREA , "ENVI") OR EXCLUDE (SUBJAREA , "MATE") OR EXCLUDE (SUBJAREA , "EART") OR EXCLUDE (SUBJAREA , "AGRI") OR EXCLUDE (SUBJAREA , " HEAL") OR EXCLUDE (SUBJAREA , "CENG") OR EXCLUDE (SUBJAREA , "CHEM") OR EXCLUDE (SUBJAREA , "PHYS") OR EXCLUDE (SUBJAREA , "BIOC") OR EXCLUDE (SUBJAREA , "NEUR") OR EXCLUDE (SUBJAREA , "NURS"))

- **Inclusion and exclusion criteria:** After the automatic retrieval of the search string described above, a manual review of the title and abstract of the articles was performed. Articles were included based on the following criteria:
- (1) A solution is given to a specific problem in one of the areas related to the management of Human Talent in a business.
- (2) It refers to the use of analytical tools to solve the problem.
- **Subject Areas:** The areas of Medicine, Nursing, Environmental Science, Agriculture and Biological Sciences, Health, Mathematics, Energy, Physics, Chemical Engineering, Neurology, Materials Science, Earth and Planetary Science, and Chemistry were excluded., Neuroscience, Biochemistry, Genetics, and Molecular Biology.
- **Information Collected**: The following data was extracted from each selected article to carry out the analysis of the results:
  - Authors
  - Name of the journal
  - Year of publication
  - Citations to the article.
  - H-Index of the authors.
  - Country of origin
  - Keywords
  - Description of the problem
  - Areas of Human Talent Management to which the problem addressed refers
  - Analytical and/or data mining techniques used to solve the problem

- Type of model used.

3 RESULTS

As a result of the search using the string described above, more than 400 documents were found presenting the results of studies conducted in the last 10 years (2014 to 2023). After performing the manual review and applying the inclusion and exclusion criteria, 46 articles were selected to answer the research questions, which are included in the final part of the bibliographic references. This implies that the search string only has a precision of 46/400, which is quite low and implies that a subsequent study of this work would be based fundamentally on the review by the expert rather than on the automatic obtaining of the documents, which makes the task of updating the study extremely difficult.

Figure 1 shows the number of articles written in the last 10 years. Initially, the studies were found from 2013 and subsequently it remained at very low levels in 2015. In 2016, there was a significant increase compared to previous years and sustained till 2017, while decreased again in 2018 and onwards. Subsequently, 2016 and 2017 recorded the largest number of articles, followed by an increase in 2023.

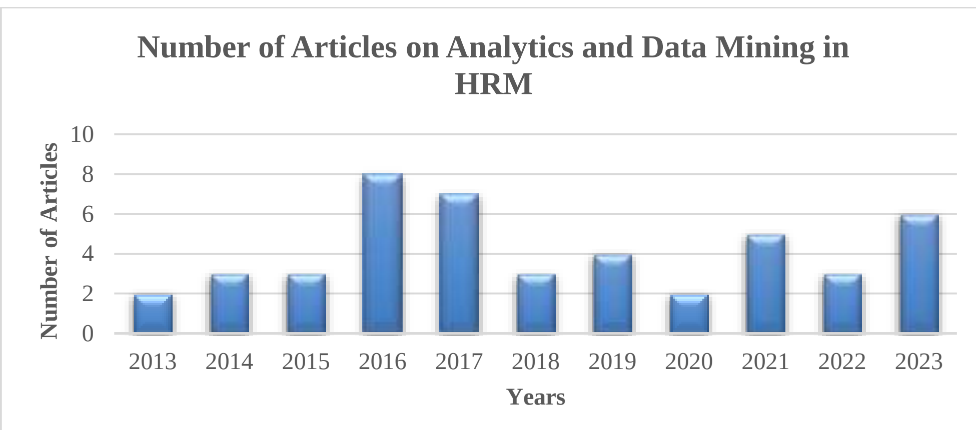

Figure 1: Number of Articles Published Per Year (Author)

These results reflect the slow evolution of the use of analytics in the fields of HRM, as mentioned in the introduction to this research, despite the great potential that exists in this area.

On the other hand, Figure 2 shows the number of citations of the articles, according to their year of publication.

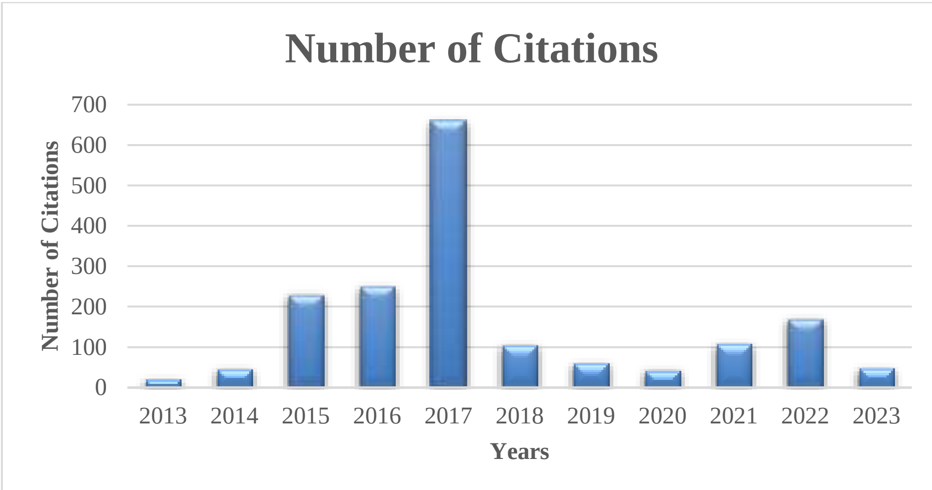

Figure 2: Number of Citations According to the Year of Publication (Author)

It can be concluded that the articles published that address the applications of analytics and data mining in the HR areas, between 2015 and 2017, are the most relevant for future studies, especially those written in the year 2017, where only 4 published articles were identified which have had the highest number of citations (above 100). As shown in Table 1, 2% of the articles have not had any citations. Besides, 25% of articles have between 1 & 10 citations while most of them have between 11 and 50 citations (54%). In contrast, only 15% have more than 50 citations, as shown in Table 1, which indicates that the researchers use the same base articles as references, which, as evidenced in Figure 2, are those articles initially written.

Citations	Number of Articles	Percentage of Total
None	1	2%
Between 1 & 10	13	28%
Between 11 and 50	25	54%
More than 50	7	15%

Table 1: Citations Per Article (Author)

Table 2 lists the articles that have had more than 50 citations, including information about their authors. The article with the highest number of citations is from 2017, followed by articles published in 2017 and 2022, which coincides with what was previously observed in Figure 2. This allows us to conclude that these articles are the most relevant for subsequent research. The year 2017 has been found with the highest citations in Figure 2 has three articles that are also included in Table 2.

Title	year	Author(H-Index)	Index-H of Authors	Citations
Artificial intelligence techniques in human resource management—a conceptual exploration	2015	Strohmeier, S. Piazza, F.	39 4	206
Knowledge discovery in software teams using evolutionary visual software analytics	2016	González-Torres, A. García-Peñalvo, F. J. Therón-Sánchez, R. Colomo-Palacios, R.	7 51 21 38	61
Competence assessment as an expert system for human resource management: A mathematical approach	2017	Bohlouli, M. Mittas, N. Kakarontzas, G. Theodosiou, T. Angelis, L. Fathi, M.	14 14 8 13 36 15	173
HR analytics and performance appraisal system: A conceptual framework for employee performance improvement	2017	Sharma, A. Sharma, T.	5 4	156
The rise (and fall?) of HR analytics: A study into the future application, value, structure, and system support	2017	Van den Heuvel, S. Bondarouk, T.	7 41	260
Human resource management in the digital age: Big data, HR analytics and artificial intelligence	2018	Lengnick-Hall, M. L. Neely, A. R. Stone, C. B.	24 7 6	86
A review of machine learning applications in human resource management.	2022	Garg, S. Sinha, S. Kar, A. K. Mani, M.	1 8 44 1	187

Table 2: Articles with More Than 50 Citations (Author)
****H-Index is as per SCOPUS**

None of the authors is repeated within Table 2 and only five of these authors are among the authors with the highest H index of above 30, as observed in Table 3: García-Peñalvo, F. J. (H index: 51), Kar, A. K. (H-index: 44), Strohmeier, S. (H-index: 39), Colomo-Palacios, R. (H-index: 38), and Angelis, L. (H-index: 36). This is also an indicator of the low dedication of the authors to the topics under study in this research, as mentioned above. That is, there is no continuity in the research in this area by the authors; this goes against the maturation of the field since the highest quality publication is expected to occur after the author has gained experience in the field. On the other hand, despite the fact that there are authors with high H indices and a large number of citations, these do not correspond to the central theme of the present investigation.

Author (H-Index)	**Index-H of Authors**	**Number of Citations in General**
Strohmeier, S.	39	3,563
García-Peñalvo, F. J.	51	10,846
Therón-Sánchez, R.	21	1,684
Colomo-Palacios, R.	38	5,077
Bohlouli, M.	14	11,430
Mittas, N.	14	818
Theodosiou, T.	13	552
Angelis, L.	36	4,068
Fathi, M.	15	988
Bondarouk, T.	23	1,940
Lengnick-Hall, M.	24	3,042
Kar, A. K.	44	7,511

Table 3: Authors with the highest H index (SCOPUS)

The 37 selected articles were written by a total of 125 authors. However, only 10 of them were found with H-indices greater than 10 and more citations. This indicates that the selected study area does not have a dedication on the part of the authors. authors and additionally these have a low academic or research activity and, on the contrary, there are very few authors who have written on the area under investigation, which have a more significant relevance at a scientific level.

Table 4 shows the names of the journals where there are articles that refer to the use of analytics in the areas of HR. It is observed that there is great dispersion in the sources that address this issue. The journals with the most articles are “Security and Communication Networks” and “Procedia Computer Science”, with a total of 2 articles respectively. The other publications only have one article.

Journal Name	**Number of Articles Published**
Journal of Social Welfare and Management	1
Indian Journal of Science & Technology	1
Journal of Leadership Studies	1
Expert Systems with Applications	1
RIMS Journal of Management	1
Turkish Journal of Electrical Engineering and Computer Sciences	1
International Conference on Decision Aid Sciences and Application	1
Tehnički glasnik	1
Journal of the Association for Information Science and Technology	1
Human Resource Management International Digest	1
Global Scientific Journal	1
INFORMS Journal on Computing	1
International Journal of Reasoning-based Intelligent Systems	1
International Journal of Productivity and Performance Management	1
Science of Computer Programming	1
Iranian Journal of Management Sciences	1
Journal of Networks	1
Asia Pacific Management Review	1
International Journal of Computing	1
Journal of Quality	1
Designing Workforce Management Systems for Industry 4.0	1
Journal of Chinese Human Resource Management	1
Management and technological challenges in the digital age	1
The International Journal of Advanced Manufacturing Technology	1
Microprocessors and Microsystems	1
Procedia Computer Science	2
Knowledge Management & E-Learning	1
Human Factors and Ergonomics in Manufacturing & Service Industries	1
International Journal of Computer Science and Information Security	1
International Journal of Scientific and Research Publications	1

International Journal of Emerging Trends in Engineering Research	1
Journal of Network and Computer Applications	1
International Journal of Sports Science & Coaching	1
The Adoption and Effect of Artificial Intelligence on HRM	1
Education and information technologies	1
Management Research Review	1
Journal of Theoretical and Applied Information Technology	1
Intelligent Techniques in Engineering Management: Theory and Applications	1
International Journal for Global Academic & Scientific Research	1
Journal of Organizational Effectiveness: People and Performance	1
Computers & Industrial Engineering	1
Security and Communication Networks	2
Human Resource Development Review	1
Plus one	1

Table 4: Number of Articles According to Publication Source (Author)

When analyzing the countries from which the articles come, it can be seen in Figure 4 that three countries lead research in the area of applied analytics in HRM. India is the country that produces the largest number of items, with a total of 11. It is followed by China and the USA, with a total of 8 and 7 respectively. Conversely, in other countries including Pakistan, there is a lack of work which could be due to the lack of experience in the use of analytics at both a business and research level.

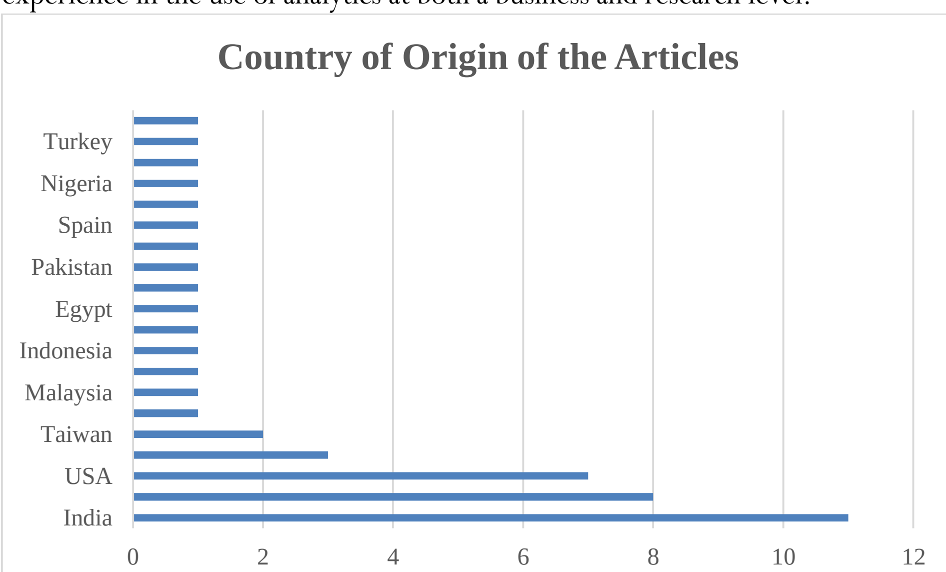

Figure 4: Number of Articles According to Country of Origin (Author)

Finally, a total of around 136 keywords were found. This widespread in the keywords used in the articles shows the wide range of areas that are considered when researching analytics in HRM. Figure 5 shows the histogram of keywords that are repeated more than once. The most used keyword is human resource management, which is found in 17 out of 46 articles, which suggests that it is one of the most considered domains in data mining and analytics research. It is followed by the data mining keyword which is used in 16 articles out of 46 representing one of the important areas of study in HRM. However, the problem is that data mining techniques from both statistics and artificial intelligence are used, so the term by itself is too broad as a keyword that helps the recovery of an article in a database of bibliographic data. This assumption will be reviewed in detail later in this investigation.

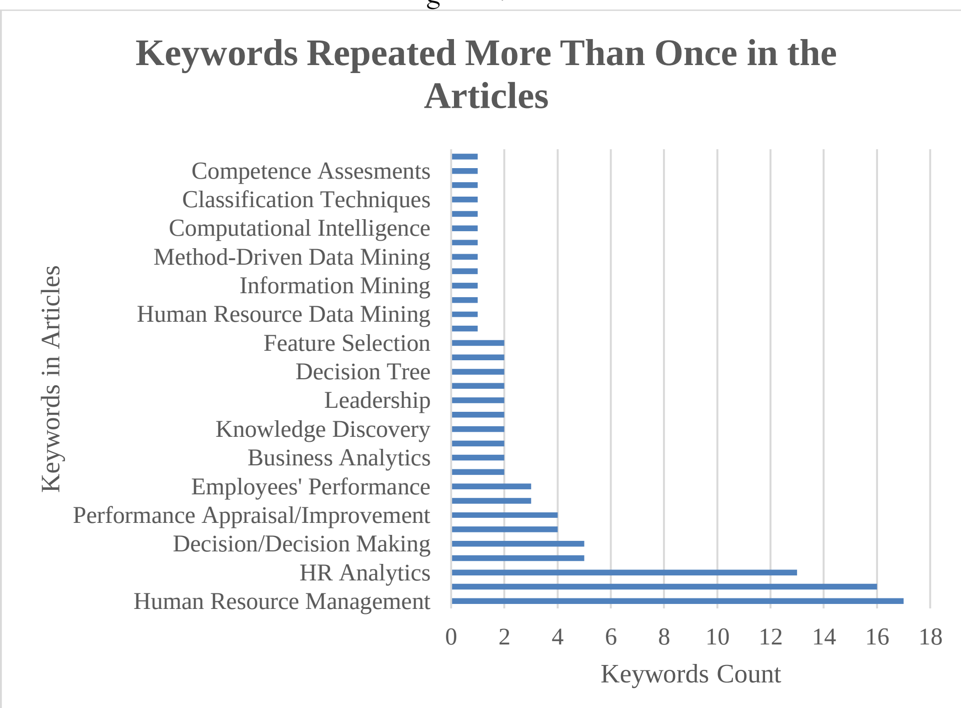

Figure 5: Repeated Keywords in the Articles (Author)

4 DISCUSSION

The research questions posed at the beginning of this document are answered below. Some of the articles describe analytics application cases that impact various processes in different fields of HR and additionally some articles describe different application cases (Bohlouli et al., 2017; Whysall, Owtram, & Brittain, 2019; Bose, 2019; Durai et al., 2019; Atchyutuni & Kumar, 2018; Shehu & Saeed, 2016; Ejo-Orusa & Okwakpam, 2018; Arora et al., 2021). In these cases, the results presented below consider each of the applications independently. On the other hand, some of the articles deal with applications of analytics and data mining in HR, however, they do not describe the model or the technique used. The above results in a total of 29 models that have 48 different applications identified in the 46 articles.

4.1 Application to Analytics and Data Mining in HRM and Its Fields

For this research, the following fields or areas within HRM, discussed by Greer (2021) were used:

1. Job Analysis and Description;
2. Attraction, Selection and Inclusion;
3. Development and Succession Plans (which includes internal procurement or employee assignment);
4. Training;
5. Performance Evaluation;
6. Remuneration and Benefits.

Additionally, an additional area called "Retention" was included, due to the relevance of this field in the results found. An additional category of "Others" was defined to include other applications that are not contained in the initially described fields. Using this classification, the results of the investigation are presented in Figure 5.

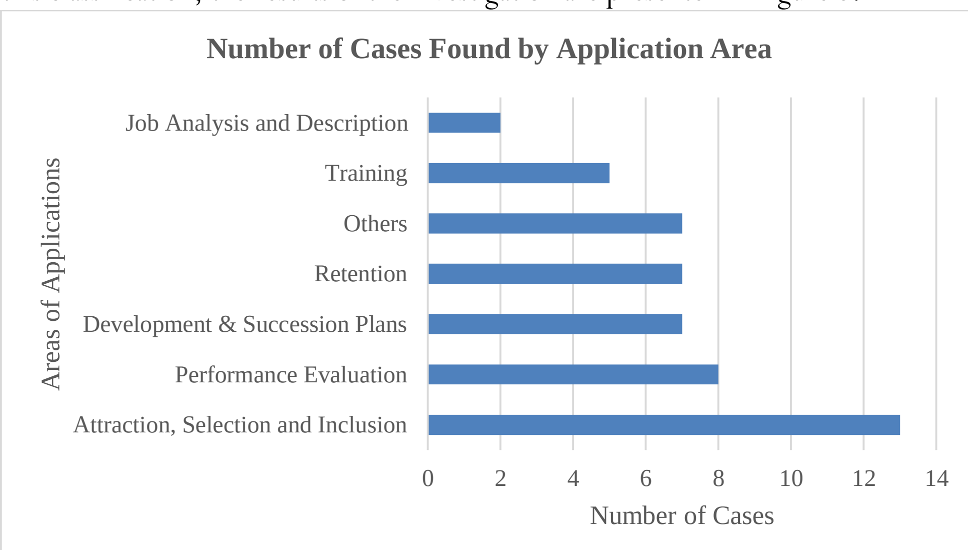

Figure 5: Number of Cases found by Application Area (Author)

As can be seen, no evidence of analytical or data mining applications in the area of Remuneration and Benefits was found in the analyzed literature. The main application that analytics has had in HRM has been in the area of Employee Attraction, Selection and Inclusion, where it has been used for the following purposes: to identify the most relevant factors or characteristics to take into account in the search for candidates that ensure their success or good performance (Whysall, Owtram, & Brittain, 2019; Pierce et al., 2017; Garg et al., 2022); develop selection models that adapt to changes in selection strategies (Shehu & Saeed, 2016); identify the candidates that best fit the requirements of a profile or position or the organizational culture, either in the initial stage of the filter or in the selection stage (Bohlouli et al., 2017; Whysall, Owtram, & Brittain, 2019; Ferneda et al., 2014; Sudha et al., 2021), facilitate the reading of resumes by extracting from them the required information (Celik, 2016), predict job behavior of candidates for a vacancy, including predictions of performance and/or retention (Sun, 2023); and evaluate the providers or sources of talent (Hajiheydari et al., 2017; Huang et al., 2023).

The field of Performance Evaluation is the second that has had the greatest use of analytics to solve the problems that arise in it. In this field, the applications that analytics have had mainly have three objectives: to improve the performance evaluation model by identifying the most relevant criteria to predict performance that have a direct impact on it (Bafna et al., 2016; Aktepe & Ersoz, 2012, Jaffar et al., 2019); improve the quality of performance evaluations, by reducing their subjectivity and/or predicting the future performance of employees (Sharma & Sharma, 2017; Li et al., 2016); and assess the competencies of employees concerning a specific role or activity (Bohlouli et al., 2017).

In the Development and Succession Plans area, analytics has been used mainly in the assignment of internal employees to roles, specific activities, or work teams, identifying those that best meet their requirements (Whysall, Owtram, & Brittain, 2019; Datta et al., 2014; Strohmeier & Piazza, 2015; Alvarado-Pérez & Peluffo-Ordóñez, 2015). It has also been used to identify gaps in skills and knowledge (development needs) of employees concerning specific roles in order to design tailored development plans and/or career plans (Khang et al., 2023); and to identify the most relevant leadership competencies that drive business results, to focus the development of leaders on these competencies (Bassi & McMurrer, 2016).

In the field of Employee Retention, analytics has been used to predict employee abandonment, identifying those with the highest risk of abandonment and/or the reasons why employees decide to leave the company, which allows strategies and actions to be implemented. to avoid it (King, 2016; Whysall, Owtram, & Brittain, 2019; Bandyopadhyay & Jadhav, 2021). It has also been used to measure the most valuable employees, towards whom retention strategies should be focused (Shehu & Saeed, 2016; Bandyopadhyay & Jadhav, 2021).

In the Training area, analytics has been used seeking to optimize the resources invested in it, through strategies such as: increasing the efficiency of Learning Management Systems (or Learning Management Systems, LMS) by providing each employee with training that is fit their learning profile (Sabitha et al., 2016); explore the quality of learning systems and identify critical indicators for their evaluation (Jou et al., 2015); to evaluate the effectiveness of the training (through the performance and behaviors of the students) (Wang et al., 2015); allow the identification and better use of informal learning networks (Khang et al., 2023) and improve the planning of training programs (Kurikala & Parvathi, 2023).

In the area of Job Analysis and Description, only one application of the analytics was identified, which aims at the automatic creation and maintenance of the competencies ontology (Malzahn et al., 2013).

Finally, in the Other category, those articles that are not related to the areas already described were grouped, including three collaboration management applications, one of which seeks to identify collaboration patterns that improve the efficiency of work teams (Fan et al., 2017); and another two aim to identify the relationships and collaborative behaviors between people, workloads and autonomy of those who are part of a workflow or a project (Park et al., 2016; González-Torres et al., 2016). Two investigations are related to the planning or assignment of personnel: on the one hand, one of these aims to generate a series of rules for the assignment of workers to certain machines or tasks that allow optimizing resources (covering production needs and at the same time provide the training required to generate flexibility in the workforce (Atchyutuni & Kumar, 2018); and the other seeks to predict future personnel needs based on possible scenarios, identifying the number of people required, the type (internal or external) and the skills required (Lengnick-Hall, et al., 2018). Another of the articles aims to identify the most relevant factors that must be considered in certain types of companies for the implementation of suggestion systems, in comparison with the successful suggestion systems in other companies (Marksberry et al., 2014; Liu et al., 2021). Finally, one of the articles analyzed refers to the application of data mining techniques in HRM in a general way to analyze the information contained in HR systems, without referring to a particular area (He & Qi, 2013; Liu, 2023).

4.2 Most Commonly Used Data Mining and Analytics Techniques or Tools in HR

Data mining is the area of analytics that has had the greatest application in HRM. In the documents analyzed, 59% of the models described use data mining techniques to solve the problem being planned and 20% use other analytical techniques. 22% of the applications found do not describe the type of technique used.

Data mining is used in the articles to solve problems related mainly to Classification (in 36% of applications), Association Rules (in 27%), Grouping (in 23%), Feature Selection (in 27%), 8%), and Regression (2%), either individually or a combination of these. Grouping techniques are those that are most frequently used in combination with others (in 24% of the applications), mainly those of Classification or Association Rules. Additionally, one of the articles combines data mining with Social Network Analytics.

Other fields different from data mining have been applied in HRM, but to a lesser extent, including Analytics or Text Mining (Whysall, Owtram, & Brittain, 2019; Sudha et al., 2021; Celik, 2016; Marksberry et al., 2014; Nasr et al., 2019), Social Network Analytics (Malzahn et al., 2013; Park et al., 2016), Visual Analytics (González-Torres et al., 2016) and Collective Intelligence (Datta et al., 2014).

Table 6 shows the number of cases in which each type of technique is applied, according to the HR area, either individually or in combination with others.

Technical \ HRM Area			Attraction, Selection and Inclusion	Performance Evaluation	Development & Succession Plans	Retention	Others	Training	Job Analysis & Description	Total
Techniques Used	Data Mining	Classification	5	3	2	4	-	3	-	**17**
		Association Rules	4	3	1	1	3	-	1	**13**
		Clustering	2	2	1	-	1	4	1	**11**
		Feature Selection	1	1	-	1	-	-	-	**3**
		Regression	-	-	-	-	1	-	-	**1**
	Other Techniques of Analytics	Text Analytics	2	-	-	1	1	-	-	**4**
		Social Network Analytics	-	-	-	-	1	1	1	**3**
		Visual Analytics	-	-	-	-	1	-	-	**1**
		Collective intelligence	-	-	-	-	-	-	-	**0**
Applications in which various techniques are combined			3	3	3	1	2	4	1	**17**
Unspecified technique			3	2	3	1	1	-	-	**10**
Total Identified Applications			**20**	**14**	**10**	**9**	**11**	**12**	**4**	**80**

Table 6: Number of Cases with Type of Analytics/Data Mining Technique as Per HRM Area (Author)

It can be observed that in all the HR areas, various data mining or analytical techniques have been applied to solve the problems inherent to them, that is, no technique is used exclusively in one of the areas or vice versa. Data mining techniques are the most widely used in all areas. Other analytics techniques have been used to a lesser extent, mainly in fields identified as "Others", in which the largest number of analytics techniques other than data mining have been applied, including Social Media, Text, and Visual Analytics.

Table 7 shows the different data mining techniques described in the articles, for those cases in which they are described in detail, which were applied either in simple or hybrid models. In the cases in which the research included a stage of evaluation of various models and selection of the best, only the technique used in the selected model is included. It is observed that there is a great variety of applied techniques. K-Means, Neural Networks, and Decision Trees are the most used, considering the applications in simple and hybrid models, followed by the a priori Algorithm and Support Vector Machines techniques. The other techniques are only used in one of the applications and are mostly used in hybrid models, except for K-nearest neighbor (in an improved version), Latent Class Analysis, Fuzzy C-means, and Gradient Boosted Machine, which were applied in simple models.

Techniques	**Type of Problem**	**Number of Times Used in Simple Models**	**Number of Times Used in Hybrid Models**	**Total Number of Times Applied**
K-means	Clustering	1	5	6
Neural Networks				
Back Propagation Neural Networks	Classification/Clustering	-	4	4
Fuzzy Neural Networks	Association Rules	1	-	1
Combined Neural Networks	Association Rules	1	-	1
Decision Tree	Classification/Association Rules	3	3	6
A priori algorithm	Association Rules	2	2	4
Support Vector Machines	Classification	1	1	2
K-Nearest Neighbors	Classification	1	-	1
Latent Class Analysis	Association Rules	1	-	1
Fuzzy C-means	Clustering	1	-	1
Gradient Boosted Machine	Classification	1	-	1
Self-Organizing Map (SOM)	Clustering	-	2	2
Regression Operating Characteristic (ROC)	Classification	-	1	1
RST (Rough Set Theory)	Feature Selection	-	1	1
Sequence Mining	Association Rules	-	1	1
Cox Regression	Regression	-	1	1
Algoritmo Scott-Knott (SK)	Clustering	-	1	1
Ward's Minimum Variance Method	Clustering	-	1	1
GRASP Algorithm	Classification	-	1	1
Simple Climbing Algorithm	Classification	-	1	1
Simulated Annealing Algorithm	Classification	-	1	1
Tabu Search	Classification	-	1	1

Table 7: Data Mining Techniques Used and Number of Applications

<u>Other Analytical Techniques</u>

In the other analytics applications, the following specific techniques were identified:

- In Social Network Analytics, discovery models, centrality measures, and prestige measures were used.
- When applying Text Analytics, Analytical Hierarchy Processes (or hierarchical analytical process), the Jaro Winkler Algorithm, Latent Semantic Analysis, Pellet Reasoner, and SMS (Semantic Matching Step) were used.
- In the case in which Collective Intelligence was applied, techniques such as Clustering Coefficient, Crawling, and Crowd Sourcing were used.

In conclusion, there are a large number of techniques used when applying analytics and data mining to HRM problems. In general, the ones that have had the greatest application have been Classification, Grouping, and Association Rules techniques, depending on the specific problem to be solved.

4.3 Evolving of the Use of Analytics in HR Over the Years

To answer this question, the evolution of analytics and data mining over the years was analyzed according to the field of application and the techniques used.

Figure 6 shows the number of applications described in the published articles, according to the HR area addressed and the year of publication.

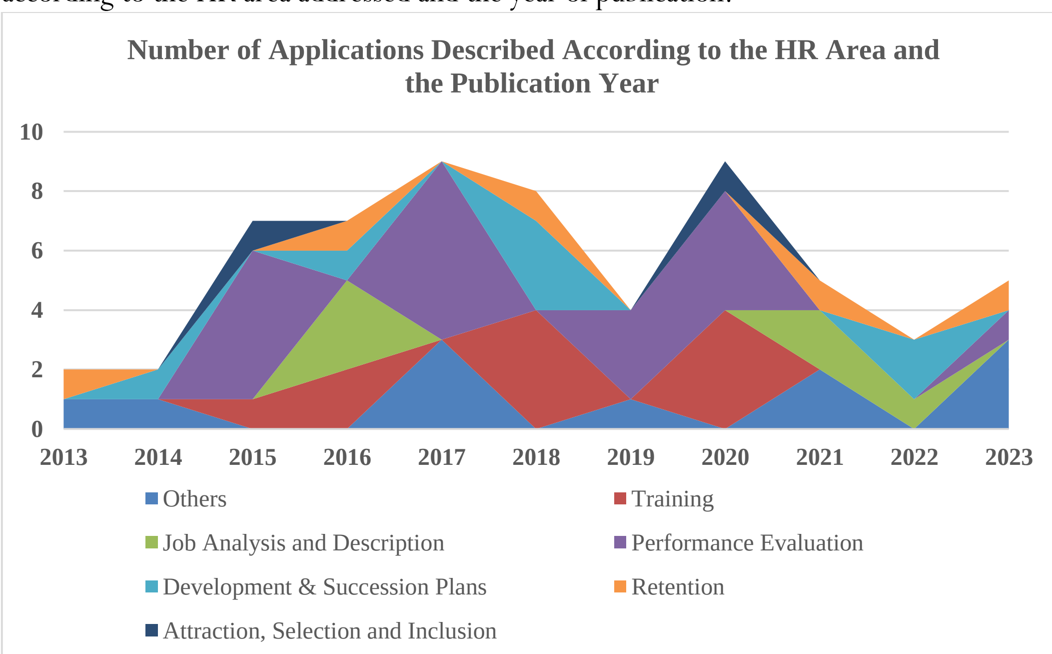

Figure 6: Number of Applications Described as Per HR Area and Publication Year

It is observed that the research published during the first 6 years that addresses the applications of analytics and data mining in the areas of HR is focused on problems of Attraction, Selection and Incorporation, Retention, as well as Development and Succession Plans. These areas are the focus of study until 2018 and continue to be studied again from 2020 to 2023. Only until 2017, in which there was an increase in the literature, were applications found in the areas of Training, Evaluation of Performance, and in Other areas, and in 2016 a new application was registered in the area of Job Analysis and Description. After this year, all areas of HR have remained current and continue to be consistently included in investigations.
Table 8 and Table 9 show, for each year, the number of models described in the articles according to their type (simple or hybrid) and according to the type of technique applied respectively. The models classified as "hybrid" are those that are built with the combination of several algorithms or techniques and those classified as "individual" are those that only include a specific technique.
As can be seen in Tables 8 and 9, in the initial years, the articles that describe analytical and data mining applications in HRM record data mining models, in which they solve Classification and Association Rules problems individually (not hybrids). Starting in 2015, hybrid models began to be registered in the applications, initially mainly combining Classification, Grouping, and Association Rules techniques and later with other techniques such as Feature Selection, Regression, and Social Network analytics. As of this year, hybrid models have been applied to a greater extent than those in which a single technique is used. As of 2016, records of applications that use analytical techniques other than data mining were found, including Social Network Analytics, Text Analytics, Visual Analytics, and Collective Intelligence. Although these techniques have been applied to a lesser extent than data mining, their applications have been increasing over the years from 2016 and onwards, also including cases in which they are applied together with other data mining techniques. Despite the investigation of new analytical techniques in the areas of HR, the investigations that apply data mining continue to occupy most of the literature, the most important being those related to Classification, Grouping, and Association Rules.

Types of Models	**Year**											**Total**
	2013	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	
Simple	1	3		4	3	2	1	4		1		**20**
Hybrid			1	3	2		1	2	2	2		**12**
Not Specified	1			2	2	1			1	1	1	**10**
Total	**2**	**3**	**2**	**9**	**7**	**3**	**2**	**6**	**3**	**4**	**1**	**42**

Table 8: Number of Models Described According to Type and Year

Types of Techniques	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**Total**
DATA MINING	**1**	**1**	**1**	**5**	**2**	**2**	**1**	**2**	**5**	**3**	**3**	**26**
Classification	1			1					3			**5**
Association Rules		1	1		1		1	1			1	**6**
Classification + Association Rules + Feature Selection					1							**1**
Classification + Clustering				2					2		1	**5**
Clustering						2				1		**3**
Clustering + Association Rules		1										**1**
Clustering + Feature Selection					1							**1**
Association Rules + Regression							1					**1**
OTHERS						**1**	**1**	**2**		**3**		**7**
Text Analytics						1		1		1		**3**

Social Network Analytics								1		1		**2**
Collective Intelligence							1					**1**
Visual Analytics										1		**1**
DATA MINING AND OTHERS				**1**	**1**							**2**
Clustering + Association Rules + Social Network Analytics					1							**1**
Clustering + Classification + Text Analytics				1								**1**
NOT SPECIFIED				**3**	**4**					**1**	**1**	**9**

Table 9: Number of Models Described as per the Type of Technique Used and the Year

When analyzing the evolution that the application of specific techniques has had over the years, it is evident that Decision Trees was the first technique used and, although it was applied again in later years, a great variety of data mining techniques has been explored. data and no one has been used consistently over the years, but on the contrary, new techniques are incorporated into the investigations. The same occurs with other analytical techniques, which were incorporated as of 2015 and as of this year, new techniques not previously used have been incorporated.

4.4 Key Results in the Investigations

Given that there is great variability in the applications described in the articles (as evidenced in Question #1), and that there is also a high variability in the analytical or data mining techniques used to solve them (as evidenced in Question #2), the results obtained in each article are specific to it and the results cannot be generalized in terms of the applications or techniques used. However, it is possible to present some conclusions about the usefulness of analytics or data mining in HRM.

First, in all the cases analyzed that include the development of a model to solve the problem being addressed, the results are satisfactory according to the parameters defined by the researchers, and in all cases valuable information is obtained to support decision-making in HRM in each of the fields described in Question number 1 (King, 2016; Shehu & Saeed, 2016; Pierce et al., 2017; Ferneda et al., 2014; Sun, 2023; Bafna et al., 2016; Li et al., 2016; Alvarado-Pérez & Peluffo-Ordóńez, 2015; Bandyopadhyay & Jadhav, 2021; Sabitha et al., 2016; Jou et al., 2015; Wang et al., 2015; Fan et al., 2017; Opatha, 2020; Atchyutuni & Kumar, 2018 & Nembhard, 2013; Marksberry et al., 2014; He & Qi, 2013). The levels of precision or confidence of the developed models vary and go from 60% to 98% (for those articles in which it is specified). However, except for one of the models that generated rules with a confidence level of 60% (Ferneda et al., 2014), all the others have results above 80% in their confidence levels. It can be concluded then that the applications of analytics and data mining in HRM allow better decisions to be made than traditional methods, which are based mainly on expert criteria. The results found in the articles show that companies can increase their efficiency in the processes, by automating or facilitating the different stages that confirm them, as well as making better decisions, by having relevant information that would not otherwise be available. It can also be concluded that the models developed can be applied in other fields or contexts only if the necessary data is available to implement the model.

Among the results obtained in the articles, the development of some software tools that can be used by professionals in the HR area to solve the problems that arise is highlighted. These tools have the following functionalities: extract information about characteristics, education and experience from unstructured resumes that are stored in databases and present it in an organized way (Celik, 2016), accurately predict the most appropriate candidate to fill a vacancy , as well as identify gaps in internal employee competencies (Bohlouli et al., 2017), identify trends in professional profiles found in the market to develop more adjusted job profiles, which facilitates recruitment and selection processes, as well as development competencies that are becoming relevant in the market (Malzahn et al., 2013; Xu & Li, 2021), automate the pre-selection processes of candidates, ranking them (Sudha et al., 2021), extract relevant information about the messages collected in job satisfaction surveys and interviews to identify the reasons for leaving and to avoid them (Whysall, Owtram, & Brittain, 2019), identify and visualize informal day-to-day learning networks in organizations and allow employees to connect with other people, according to their day-to-day learning needs and interests (Mirski et al., 2017; Pah & Utama, 2020), identify the contributions made by team members in Software development projects and identify collaboration patterns (González-Torres et al., 2016), provide recommendations for the formation of work teams, according to the area of expertise required for them and other preferences of the user (Opatha, 2020) and identify the rules for the formation of work teams, specifically for Software development (Rai & Singh, 2023).

Additionally, there is a group of articles that include the comparison between two or more models and the one that generates better results is presented (King, 2016; Shehu & Saeed, 2016; Bafna et al., 2016; Li et al., 2016; Bandyopadhyay & Jadhav, 2021; Valecha, 2022; Jou et al., 2015; Wang et al., 2015; Van den Heuvel & Bondarouk, 2017; Zhou, 2022; Zhang, 2021).

In these cases, the main results obtained in the comparison of the models are:

- The "Gradient Boosted Machine" model produced the best results in predicting the employees most likely to leave their employment, with an overall accuracy of 81% in its predictions, compared to other techniques such as decision trees, logistic regression, and k -means (Bohlouli et al., 2017).
- The C4.5 decision trees were the ones with the highest precision compared to the Random Tree, REP Tree, and CART methods in the generation of selection rules to support the identification of candidates that best fit a profile, with an accuracy of 88%. .7% However, the accuracy was increased above 90% by building a combined adaptive model, which allows the incorporation of the adaptive rules generated with the other three algorithms (Shehu & Saeed, 2016).
- Compared to seven traditional KNN (K-Nearest Neighbor) algorithms, the improved KNN algorithm that was developed performed better, with higher precision and lower variance. This algorithm showed a maximum prediction error of less than 2% when the weighted distance between the individual being predicted and the sample is below 0.9. When the distance is between 0.9 and 2, the error will rise to 5.31%, but when it rises above 2, the maximum error is unstable and produces unacceptable results (Li et al., 2016).

- In identifying the most relevant parameters to predict performance and that directly impact it, the continuous K-means algorithm generated better results than the K-means and "Pairwise agglomerative" algorithms, when measuring parameters such as error rate, entropy, and purity (Bafna et al., 2016).
- The C5.0 decision trees had better classification accuracy than the CHAID and CART methods, with an 89.4% accuracy rate in identifying the critical indicators for evaluating training (Jou et al., 2015).
- Backpropagation BPNN Neural Network models (traditional and adapted) obtained higher accuracy compared to decision trees and logistic regression, both in predicting student learning behaviors, as well as in predicting their performance. The adapted BPNN model was more accurate than the traditional one (98.28% vs. 95.15%) in predicting student learning behaviors, and the traditional BPNN model was more accurate than the adapted one (95, 17% vs. 91.04%) in predicting student performance (Wang et al., 2015).
- In predicting the turnover of technology professionals, a hybrid model (SOM + BPNN) obtained an accuracy of 92%, higher than that obtained with the K-Means models (63.5%) and BPNN neural networks (87, 2%) (Jou et al., 2015).
- With a confidence coefficient of 95%, it is concluded that the RST-SVMDT hybrid algorithm, which combines RST (Rough Set Theory), SVM (Support Vector Machines), and Decision Tree algorithms, has greater precision in extracting information for decision-making related to selection, performance prediction, and retention, that each of the algorithms applied independently or combined with each other. Compared to the SVM algorithm, the precision is slightly higher, although it is not statistically significant, so both models are comparable (Xu & Li, 2021).
- For the prediction of employee turnover, the SVM, Random Forest, and Naïve Bayes models were compared. Although the three models are comparable in their overall accuracy, the SVM model outperformed the other two models in recognizing true positives, which are those employees who have resigned (Bandyopadhyay & Jadhav, 2021).

Finally, there is a group of articles included in the analysis, which corresponds to 14% of the total, which mentions other potential applications of analytics and data mining in HRM; however, they do not present the development of a specific model or concrete results to support the hypotheses they propose (Sharma & Sharma, 2017; Van den Heuvel & Bondarouk, 2017; Strohmeier & Piazza, 2015; Bassi & McMurrer, 2016; Yoon et al., 2023; Zhang, 2021; Zhou, 2022). Therefore, there is an opportunity to carry out additional research to develop the most appropriate practical models to use in solving the problems identified in each case. There is also an opportunity to carry out future research focused on improving the models already developed, especially in those cases that present limitations, either through the identification of a technique that allows improving confidence levels or by incorporating additional variables or information that also generates better results. results; or in developing new models to solve problems that have not yet been addressed, including for example problems of Remuneration and Benefits, which were not included in the research analyzed.

CONCLUSION:

About each of the research questions posed, the following conclusions can be made:
As a result of the systematic analysis of the literature, it was evidenced that analytics and data mining have been applied in all areas of HRM, except for the area of Remuneration and Benefits, where there is potential for future research. The area that has had the greatest application over the years has been the Attraction, Recruitment, and Selection of employees, followed by the area of Performance Evaluation. The applications of analytics and data mining in all areas are very varied and respond to different problems that arise in each one.

Data mining is the area of analytics that has had the greatest application in HRM. Classification, Grouping, and Association Rules models are mainly used, both in hybrid and simple models. When analyzing the areas of HRM, it is possible to identify different techniques and models applied in all areas, that is, there is a great variety of techniques used in the cases analyzed and there is no tendency to use a certain type of model for every problem. K-means, Neural Networks, Decision Trees, and a priori Algorithm were the ones that were applied the most repetitively.

The areas of Attraction, Selection, and Incorporation, Employee Retention and Development, and Succession Plans have been the focus of study since the initial investigations and are still valid. Over time, problems have arisen in new areas that have been sought to be solved through analytics and data mining, including problems in the areas of Performance Evaluation, Training, and others. There have been two important moments since 2003 in which there has been a particular interest in studying the applications of analytics and data mining in HRM. The first was in 2011, the year in which there was a significant increase in the literature and in which new areas and new techniques that had not been previously studied were incorporated. Subsequently, in 2016, there was again an increase in the literature. Although no new areas were entered this year, new techniques were included, especially analytical techniques other than data mining that had not been previously studied, including Social Network Analytics, Text Analytics, Visual Analytics, and Collective Intelligence. However, these new techniques have not been subsequently applied in other areas or fields.

Despite the fact that the practical results of each of the investigations are different, it can be concluded in a general way that the applications of analytics and data mining in HRM allow for increasing the efficiency of their processes, by automating or facilitating the different stages of these, as well as making better decisions than traditional methods by having relevant information that would not otherwise be available. It can also be concluded that the models developed can be applied in other fields or contexts as long as the necessary data is available to implement the model.

REFERENCES:

Aktepe, A., & Ersoz, S. (2012). A Quantitative Performance Evaluation Model Based on a Job Satisfaction-Performance Matrix and Application in a Manufacturing Company. *International Journal of Industrial Engineering*, *19*(6).

Alvarado-Pérez, J. C., & Peluffo-Ordóńez, D. H. (2015). Artificial and natural intelligence integration. In *Distributed Computing and Artificial Intelligence, 12th International Conference* (pp. 167-173). Springer International Publishing.

Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and analytics: why HR is set to fail the big data challenge. *Human resource management journal*, *26*(1), 1-11.

Arora, M., Prakash, A., Mittal, A., & Singh, S. (2021, December). HR analytics and artificial intelligence-transforming human resource management. In *2021 International Conference on Decision Aid Sciences and Application (DASA)* (pp. 288-293). IEEE.

Atchyutuni, N., & Kumar, P. V. (2018). Applying Data Analytics for Human Resource Management: Opportunities & Challenges. *Journal of Social Welfare and Management*, *10*(2), 128.

Bafna, P., Pillai, S., & Pramod, D. (2016). Quantifying performance appraisal parameters: a forward feature selection approach. *Indian J Sci Technol*, *9*(21), 1-7.

Bandyopadhyay, N., & Jadhav, A. (2021). Churn prediction of employees using machine learning techniques. *Tehnički glasnik*, *15*(1), 51-59.

Bassi, L., & McMurrer, D. (2016). Four lessons learned in how to use human resource analytics to improve the effectiveness of leadership development. *Journal of leadership studies*, *10*(2), 39-43.

Bennett, C., & Collins, L. (2015). HR and people analytics: Stuck in neutral. *Global human capital trends 2015*, 71-77.

Bersin, J., O'Leonard, K., & Wang-Audia, W. (2013). High-impact talent analytics: Building a world-class HR measurement and analytics function. *Bersin by Deloitte*, *131*.

Bohlouli, M., Mittas, N., Kakarontzas, G., Theodosiou, T., Angelis, L., & Fathi, M. (2017). Competence assessment as an expert system for human resource management: A mathematical approach. *Expert Systems with Applications*, *70*, 83-102.

Bose, I. (2019). Metrics of Organizational Practices in Human Resource Management: Selected perspectives from HR analytics. *RIMS Journal of Management*, *4*(1), 26-35.

Celik, D. (2016). Towards a semantic-based information extraction system for matching résumés to job openings. *Turkish Journal of Electrical Engineering and Computer Sciences*, *24*(1), 141-159.

CIPD. (2013). Talent analytics and big data–the challenge for HR.

Datta, A., Yong, J. T. T., & Braghin, S. (2014). The zen of multidisciplinary team recommendation. *Journal of the Association for Information Science and Technology*, *65*(12), 2518-2533.

De Laat, M., & Schreurs, B. (2013). Visualizing informal professional development networks: Building a case for learning analytics in the workplace. *American Behavioral Scientist*, *57*(10), 1421-1438.

Dulebohn, J. H., & Johnson, R. D. (2013). Human resource metrics and decision support: A classification framework. *Human Resource Management Review*, *23*(1), 71-83.

Durai D, S., Rudhramoorthy, K., & Sarkar, S. (2019). HR metrics and workforce analytics: it is a journey, not a destination. *Human Resource Management International Digest*, *27*(1), 4-6.

Edwards, M. R., & Edwards, K. (2019). *Predictive HR analytics: Mastering the HR metric*. Kogan Page Publishers.

Ejo-Orusa, H., & Okwakpam, J. A. A. (2018). Predictive HR analytics and human resource management amongst human resource management practitioners in Port Harcourt, Nigeria. *Global Scientific Journal*, *6*(7), 254.

Fan, S., Li, X., & Zhao, J. L. (2017). Collaboration process pattern approach to improving teamwork performance: A data mining-based methodology. *INFORMS Journal on Computing*, *29*(3), 438-456.

Ferneda, E., Do Prado, H. A., Sobrinho, A. G. C., & Balaniuk, R. (2014). Using data mining in the selection process of high performance managers. *International Journal of Reasoning-based Intelligent Systems 16*, *6*(1-2), 40-48.

Garg, S., Sinha, S., Kar, A. K., & Mani, M. (2022). A review of machine learning applications in human resource management. *International Journal of Productivity and Performance Management*, *71*(5), 1590-1610.

González-Torres, A., García-Peñalvo, F. J., Therón-Sánchez, R., & Colomo-Palacios, R. (2016). Knowledge discovery in software teams by means of evolutionary visual software analytics. *Science of Computer Programming*, *121*, 55-74.

Greer, C. R. (2021). *Strategic human resource management*. Pearson Custom Publishing.

Hajiheydari, N., Khabiri, S. H., & Talafi Daryani, M. (2017). A framework for data mining approach applications in human resource management. *Iranian journal of management sciences*, *12*(47), 21-50.

He, L., & Qi, J. (2013). Enterprise human resources information mining based on improved apriori algorithm. *Journal of Networks*, *8*(5), 1138.

Huang, X., Yang, F., Zheng, J., Feng, C., & Zhang, L. (2023). Personalized human resource management via HR analytics and artificial intelligence: Theory and implications. *Asia Pacific Management Review*.

Jaffar, Z., Noor, W., & Kanwal, Z. (2019). Predictive human resource analytics using data mining classification techniques. *Int. J. Comput*, *32*(1), 9-20.

Jones, K. (2014). Conquering HR Analytics: Do you need a rocket scientist or a crystal ball. *Workforce Solutions Review*, *5*(1), 43-44.

Jou, Y. T., Wu, Y. C., & Lin, W. T. (2015). Applying decision trees and neural network to raise the performance of human training quality. *品質學報*, *22*(5), 383-403.

Kantardzic, M. (2020). *Data mining: concepts, models, methods, and algorithms*. John Wiley.

Khang, A., Gupta, S. K., Dixit, C. K., & Somani, P. (2023). Data-driven application of human capital management databases, big data, and data mining. In *Designing Workforce Management Systems for Industry 4.0* (pp. 105-120). CRC Press.

King, K. G. (2016). Data analytics in human resources: A case study and critical review. *Human Resource Development Review*, *15*(4), 487-495.

Krishnan, S. K., & Singh, M. (2011). Strategic human resource management: A three-stage process model and its influencing factors. *South Asian Journal of Management*, *18*(1), 60-82.

Kurikala, V., & Parvathi, V. (2023). Identifying the characteristics that lead to effective data analytics deployment in human resource management—a paradigm for consequences. *J Chin Hum Resour Manag*, *14*(1), 60-89.

Lengnick-Hall, M. L., Neely, A. R., & Stone, C. B. (2018). Human resource management in the digital age: Big data, HR analytics, and artificial intelligence. In *Management and technological challenges in the digital age* (pp. 1-30). CRC Press.

Li, N., Kong, H., Ma, Y., Gong, G., & Huai, W. (2016). Human performance modeling for manufacturing based on an improved KNN algorithm. *The International Journal of Advanced Manufacturing Technology*, *84*, 473-483.

Liu, J. (2023). Design and Application of Human resource management system Based on Data Mining Technology. *Procedia Computer Science*, *228*, 241-252.

Liu, P., Qingqing, W., & Liu, W. (2021). Enterprise human resource management platform based on FPGA and data mining. *Microprocessors and Microsystems*, *80*, 103330.

Malzahn, N., Ziebarth, S., & Hoppe, H. U. (2013). Semi-automatic creation and exploitation of competence ontologies for trend-aware profiling, matching, and planning. *Knowledge Management & E-Learning*, *5*(1), 84.

Marksberry, P., Church, J., & Schmidt, M. (2014). The employee suggestion system: A new approach using latent semantic analysis. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *24*(1), 29-39.

Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR Analytics. *The International Journal of Human Resource Management*, *28*(1), 3-26.

Mirski, P., Bernsteiner, R., & Radi, D. (2017). Analytics in human resource management the OpenSKIMR approach. *Procedia computer science*, *122*, 727-734.

Nasr, M., Shaaban, E., Samir, A., & Nasr, M. (2019). A proposed model for predicting employees' performance using data mining techniques: Egyptian case study, *International Journal of Computer Science and Information Security, 17(1),* 31-40.

Opatha, H. H. D. P. J. (2020). HR analytics: a literature review and new conceptual model. *International Journal of Scientific and Research Publications*, *10*(6), 130-141.
Pah, C. E. A., & Utama, D. N. (2020). Decision support model for employee recruitment using data mining classification. *International Journal of Emerging Trends in Engineering Research, 8*(5), 1511-1516.
Park, M., Ahn, H., & Kim, K. P. (2016). Workflow-supported social networks: Discovery, analyses, and system. *Journal of Network and Computer Applications*, *75*, 355-373.
Pierce, D. A., Johnson, J. E., Krohn, B. D., & Judge, L. W. (2017). Who should we hire?: Examining coaching succession in NCAA Division I women's basketball. *International Journal of Sports Science & Coaching*, *12*(2), 151-161.
Pruengkarn, R., Wong, K. W., & Fung, C. C. (2017). A review of data mining techniques and applications. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *21*(1), 31-48.
Rai, A., & Singh, L. B. (2023). Artificial Intelligence-based People Analytics Transforming Human Resource Management Practices. In *The Adoption and Effect of Artificial Intelligence on Human Resources Management, Part A* (pp. 229-244). Emerald Publishing Limited.
Russell, C., & Bennett, N. (2015). Big data and talent management: Using hard data to make the soft stuff easy. *Business Horizons*, *58*(3), 237-242.
Sabitha, A. S., Mehrotra, D., & Bansal, A. (2016). Delivery of learning knowledge objects using fuzzy clustering. *Education and information technologies*, *21*, 1329-1349.
Sharma, A., & Sharma, T. (2017). HR analytics and performance appraisal system: A conceptual framework for employee performance improvement. *Management Research Review*, *40*(6), 684-697.
Shehu, M. A., & Saeed, F. (2016). An adaptive personnel selection model for recruitment using domain-driven data mining. *Journal of Theoretical and Applied Information Technology*, *91*(1), 117.
Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual review of psychology*, *70*, 747-770.
Strohmeier, S., & Piazza, F. (2015). Artificial intelligence techniques in human resource management—a conceptual exploration. *Intelligent Techniques in Engineering Management: Theory and Applications*, 149-172.
Sudha, G., Sasipriya, K. K., Nivethitha, D., & Saranya, S. (2021, December). Personality prediction through CV analysis using machine learning algorithms for automated e-recruitment process. In *2021 4th international conference on Computing and Communications Technologies (ICCCT)* (pp. 617-622). IEEE.
Sun, Z. (2023). A Method for Evaluating the Competitiveness of Human Resources in High-tech Enterprises Based on Self-organized Data Mining Algorithms. *International Journal of Advanced Computer Science and Applications*, *14*(7).

Ulrich, D., & Dulebohn, J. H. (2015). Are we there yet? What's next for HR?. *Human Resource Management Review*, *25*(2), 188-204.

Valecha, N. (2022). Transforming human resource management with HR analytics: A Critical Analysis of Benefits and Challenges. *International Journal for Global Academic & Scientific Research*, *1*(2), 56-66.

Van den Heuvel, S., & Bondarouk, T. (2017). The rise (and fall?) of HR analytics: A study into the future application, value, structure, and system support. *Journal of Organizational Effectiveness: People and Performance*, *4*(2), 157-178.

Wang, J., Lin, Y. I., & Hou, S. Y. (2015). A data mining approach for training evaluation in simulation-based training. *Computers & Industrial Engineering*, *80*, 171-180.

Whysall, Z., Owtram, M., & Brittain, S. (2019). The new talent management challenges of Industry 4.0. *Journal of management development*, *38*(2), 118-129.

Xu, M., & Li, C. (2021). Data mining method of enterprise human resource management based on simulated annealing algorithm. *Security and Communication Networks*, *2021*, 1-9.

Yoon, S. W., Han, S. H., & Chae, C. (2023). People Analytics and Human Resource Development–Research Landscape and Future Needs Based on Bibliometrics and Scoping Review. *Human Resource Development Review*, 15344843231209362.

Zhang, A. (2021). Influence of data mining technology in information analysis of human resource management on macroscopic economic management. *Plos one*, *16*(5), e0251483.

Zhou, D. (2022). Application of data mining technology in enterprise digital human resource management. *Security and Communication Networks*, *2022*.