



Social Media Analytics for Human Resource Management using Association Rule

Benson Mugambi Njoni*Faculty of Computing and Information Technology,
Jomo Kenyatta University of Agriculture & Technology,
(JKUAT), Kigali, Rwanda**Ann Kibe**Faculty of Computing and Information Technology,
Jomo Kenyatta University of Agriculture & Technology,
(JKUAT), Nairobi, Kenya

Abstract— *This paper discusses and demonstrates how the Association Rule can be used in Business Analytics to mine social media data to discover hidden relationships and patterns so as to predict employee behavior, and performance. These discoveries can influence Human Resources Management to hire appropriately, identify techniques to improve employee performance, prevent crime or malpractice and therefore ensure business sustainability and success. First, a brief introduction to the main concepts is discussed, then a brief overview of current literature in Business Intelligence & Analytics (BA&I) is given in relation to Human Resource (HR) management. Finally, an analysis is given on how the Association rule can be implemented in mining employee social media data to provide insight that can be helpful in improving internal environment performance and thus leverage overall organizational performance.*

Keywords— *Business Analytics, Human Resource Management, Social Media Analytics, Association Rule, Data Mining.*

I. INTRODUCTION

All over the world, businesses are striving to outdo the competition by utilizing the information, knowledge, and experiences they have about their markets, products and competitors. This knowledge and experience, is referred to as business intelligence (BI). The more intelligence that a business has, and the better the capability to apply this intelligence appropriately, then the higher the chances of success for the business. Hsinchun et.al [12], argues that “The opportunities associated with data and analysis in different organizations have helped generate significant interest in BI&A, which is often referred to as the techniques, technologies, systems, practices, methodologies, and applications that analyze critical business data to help an enterprise better understand its business and market and make timely business decisions”.

There is no doubt that Business Intelligence and Analytics (BA&I) is now an established practice giving business users better insights, particularly from operational data stored in transactional systems [12]. But what value is there in having perfect information about the competitors and the markets (external environment) with little or no intelligence about the internal environment? A comprehensive BI&A will have to address the overall business areas including Human Resources (HR) management including other Internal environment issues.

Garrett [8] agrees: More companies today are finding the answer to regaining the value of information lies in business analytics, which enables HR executives to integrate important data from disparate sources and gain real-time visibility into metrics that can be measured against organizational goals. Perhaps HR as a function now has the greatest need to use analytics to speak more clearly to the business about value, outcomes, and risk. And business leaders, boards, shareholders and even regulators, are not only more receptive but demanding more now on the data that helps us all understand how to get more out of our people and organizations - “and to know where the bear traps might be” [3]

This article discusses how BI&A can be implemented in Human Resource Analytics to give insight into employee behavior, capabilities, talents, and performance, so as to influence optimal decision making in Human Resources management.

II. BUSINESS ANALYTICS DATA SOURCES

Using data to support decision-making is not new, and falls under the umbrella of business analytics. As a result of the ever-decreasing costs of data collection, storage, and processing it is now possible to collect humongous information about any element relevant to decision-making [18]

He, W. et.al, [10] suggests that social media analytics uses data mining platforms, tools and analytics techniques to collect, monitor and analyze massive amounts of social media data to extract useful patterns, gain insight into market requirements and enhance business intelligence.

Through advanced business analytics, IBM VOCA mines diverse data sources from routine customer service interactions - including audio recordings, call transcripts, emails, survey results and demographic data - to deliver a single, integrated view of customer sentiment to improve marketing effectiveness, enhance customer service and grow customer loyalty [12]

Data is collected using bar-code scanners in supermarkets. Often market basket databases consist of a vast number of transaction records. All items bought by a customer in a single purchase transaction are recorded in each list. Managers would be interested to find out which groups of items are systematically purchased together. They could use this data to modify store layouts (by for example placing items purchased together optimally close to each other), for cross-selling, for promotions, for catalog design, and to identify customer segments based on buying patterns [1]

Under small data analytics, you would collect historical sales data, observe any trends in the data (e.g. higher sales during the holiday season) and perform a time-series forecast of the future demand. These are examples of descriptive and predictive analytics [18]

The current problem, therefore, is hardly the availability of data or data sources, but making the meaning out of it – that is the analytics. According to Garrett [8], No matter how much information is captured, data has limited utility if it cannot be quickly distilled and analyzed to make important business decisions. Without real-time access to meaningful metrics, data loses its value as a strategic management tool.

There is a general consensus, that data is now ubiquitous and readily available. For example, an Internet-based retailers can gather a diverse range of information such as customer demographics (gender, location, age), weather, real-time inventory information from RFID (radio frequency identification) chips, and even blog post and video reviews of products. The magnitude of recorded data sets keeps growing rapidly as you recurrently record more of such relevant attributes.

III. BUSINESS ANALYTICS IN HUMAN RESOURCES MANAGEMENT

Garrett [8] opines that by utilizing the right application for the right metrics, employers can improve processes and efficiency while tying their strategic performance to organizational goals and initiatives key to their company's success. Tools that expand the use of Excel-type spreadsheets are allowing individuals who are not statisticians to see patterns in data. Dashboards are bringing more and more data to HR professionals, managers and employees. Douthitt [5] suggests that by considering training, selection, performance on key competencies, and employee attitudes, [an] organization [is] able to make fact-based decisions regarding potential investments in its people and calculate an expected ROI.

IV. BUSINESS ANALYTICS & INTELLIGENCE TECHNIQUES AND TOOLS

Whether a company develops custom analytics applications internally or purchases a ready-made solution, business analytics applications ease processes, speed results, and enable employers to get the most value out of corporate information — with positive financial results [8]. Eddy [6] argues that there are two challenges small businesses face in adopting BA&I.

“First, is just getting going, because they have little to no support from IT--likely because there are only a few IT people on staff. This challenge is usually around finding, connecting and making sense of the data, which is likely coming from spreadsheets, cloud-based applications, and some internal data.

Once they get the data prepared, they can often find initial success”. The second challenge, according to [6], it arises as a result of that success, the need for expansion and it has two aspects. "First, users will want to expand the number and types of data sources to analyze into one for more sophisticated analysis. This type of data engineering is usually beyond the skills of most business users. Second, users will want to expand the number of users. Expanding users usually means figuring out how to share information securely among many people, which again, is usually beyond the skills of most business users [6]

Garrett [8] seems to agree with this perspective: In a manual process, gathering this data and compiling it into a meaningful report can often take days to complete. Even payroll register reports become a monumental task when data cannot be modified, sorted or partitioned without a protracted, labor-intensive process. How can executives deliver the right data to the right audience at the right time when valuable metrics are so hard to obtain?

So what tools and techniques should be feasibly adopted, to the ensure success of BA& I? A business analytics application that connects directly to original data sources cuts out the intermediaries. This allows HR executives to realize the benefits of business analytics without the complexity or expense of compiling multiple data sources into a common database.

Examples of commercial BA&I tools include: Microsoft Excel, SAS, SPSS Modeler (Clementine), Statistica, Salford systems, KXEN, Angoss, and MATLAB. Popular open-source BA&I tools include: R, Weka, SQL, Tableau, Orange, KNIME and Azure ML [20]. For limited size of data, dashboards can be made using Advanced excel. But, typically Organizations use more advanced tools for creation and dissemination of tools. Business Objects, Qlikview, Hyperion are names of some such software [13]

Bryce [2], gives a breakdown of some critical qualifiers for the appropriate BA&I tools as follows:

- (I) Based on an open and integrated framework
- (ii) Able to handle data and metadata requirements
- (iii) Fully scalable to growth objectives and overall business requirements
- (iv) Low total cost of ownership
- (v) Easy to use; minimal training and disruption
- (vi) Powerful analysis and reporting power

In future, however, the next generation of BA&I tools will not be stand-alone systems, but a collection of discrete systems that can be disassembled and reassembled.

BA software will have to be integrated with a variety of other software inflections to provide additional insight and context about the business information being analyzed. New solutions will include functionality for better information access, collaboration, and business process management. Portal software (followed by master data management) is the most likely offering to be integrated with BA software among the choices provided [19].

In a nutshell, [8] and [9] opine that companies can wield business intelligence as a weapon to outsmart competitors and boost revenue, but only if data analysis techniques are used enterprise-wide and the effort is backed by senior management.

V. BA&I TECHNIQUES AND TOOLS – ASSOCIATION RULE

Association rule learning is a method for discovering intriguing relations between variables in large data sets. Association rule mining looks for interesting associations and correlations among large sets of data. Association rules portray attribute value conditions that recur in a given data set. A good example of the association rule mining is Analysis of Market Basket, which looks for items that are often purchased together. An association rule has two parts, a preceding variable or antecedent (if) and a resultant or consequent variable (then). A preceding value is an item found in the data. A resultant or consequent is an item that is found in combination with the preceding or antecedent.

Association rules are created by analyzing data for frequent if/then patterns and using the criteria support and confidence to identify the most important relationships. Support is an indication of how frequently the items appear in the database. The number of times the if/then statements are found to be true indicates confidence [17] In data mining, association rules are useful for analyzing and predicting customer behavior. Association rules help in shopping basket data analysis, merchandise clustering, catalog design and store layout.

Programmers use association rules to build programs capable of machine learning. Machine learning is a type of artificial intelligence (AI) that aims at developing programs that have the capability to become more efficient without the need for explicit programming.

Analytics applications also utilize queries and time-lines to provide better visibility on operational metrics, and create an early warning system through threshold monitoring and notifications [8]

VI. SOCIAL MEDIA BA&I

There is no doubt that social media today has more than adequate data (Big Data) that can be relied by businesses for BA&I ranging from descriptive analytics, reporting analytics to predictive analytics. Winterberg [23], suggests that popular web sites such as Facebook, LinkedIn, and Twitter provide compelling and previously unimaginable ways to communicate with family, friends, allied professionals, clients, and prospects.

But with this rapidly growing medium comes significant uncertainty about how planners can leverage social media to meet business objectives and do so without violating regulatory or compliance requirements. For example [16] proposes “a new method for recommending tourist locations that are relevant to users (i.e., personalization) in the given context (i.e., context awareness)” based on users' traveling preferences elicited from their travel experiences exposed on social media sites by sharing geotagged photos on sites like Flickr, Facebook, Twitter and others. Konopnicki et.al [15], presents a way of analyzing social media conversational data in order to better understand customers. In addition, the last couple of years have seen the emergence of advanced BA&I tools modeled on social media.

For example, [12] reports that “Beyond the Arc can now deliver tangible customer, product and market insights based on daily access to millions of customer comments from Facebook, Twitter and YouTube”. Another example InfoExtractor, a Web-based tool for collecting data and metadata from focused social media content. According to [7], InfoExtractor then provides these data in various structured and unstructured formats for manipulation and analysis.

The tool allows social science researchers to collect data for quantitative analysis, and is designed to deliver data from popular and influential social media sites in a useful and easy-to-access format. It is clear therefore that social media data mining is the next frontier BI&A and we need to embrace it and develop adequate methods and techniques for this activity. Walter [21], advises: Get used to it.

The gathering of such data, whether by private commercial enterprises, hackers or governments - ours or foreign ones - is part of 21st- century life.

The significance of social media supersedes BI&A as noted by [4] : As much as it can be a means of marketing and communication with the public, social media also is a minefield for attorneys. One wrong move can mean anything from a self-sullyng reputation to possibly an ethics violation. Social media also exposes the reputation of clients, employers and potential employees to a higher level of scrutiny that was not there a few years ago.

For example, [4] notes that there wasn't much public perception of individual lawyers. But now that's changed. So much of the practice of law involves the lawyer's reputation. You as a prospective lawyer need to start thinking about your reputation now. As part of reputation control, many social users already regularly monitor their personal social media accounts to make sure anyone they have friended don't post photos of them that could present them in a bad light.

This emphasizes the value of social media data in BA&I that is hardly utilized by the HR in predictive Analytics but that could be the key to unlock HR issues related to employee competence. Social media analytics uses data mining platforms, tools and analytics techniques to collect, monitor and analyze massive amounts of social media data to extract useful patterns, gain insight into market requirements and enhance business intelligence. But before any analytics on social media can commence, it is important to have the goals of the BA&I well stated and stream-lined with the overall goals of the organization which the analytics seeks to achieve. Clarity and transparency is of essence.

VII. HYPOTHETICAL CASE USING SOCIAL MEDIA AND ASSOCIATION RULE

In this hypothetical case, a methodology for mining social media data using Association Rule is proposed. The purpose is to track the non-obvious relationships among persons (for example potential employees of a company) to gain an insight into the behavior, character, and other properties that can affect the performance of the employee.

In this hypothetical case, Facebook is chosen. In this basic example, the goal is to mine and look for relationships among a Facebook user and their friends using the fuzzy association rule.

Table 1. Hypothetical Social Media Analytics model showing person and corresponding friends

Person	Friends
A	{B, D, E, F}
B	{A, B, C, F, G}
C	{B, E, F, G}
D	{B, F, G}
E	{A, C, F, G}
F	{A, B, C, D, E}
G	{B, C, D, E }

Looking directly at the hypothetical model above, it is clear that B and F have more friends than the rest in the dataset. This is an obvious relationship. Advanced analysis of the dataset should reveal more information. For example, assuming D is flagged as a convicted criminal, then it is possible to find out the unobvious relationships between the flagged person D and the rest of the datasets. E.g D has B, F, G friends and one of those friends B has a friend A who is also a convicted criminal. This squarely puts B on the spot and therefore may need to be scrutinized further before being recruited.

This is a very simple illustration of how the Association Rule can be used to mine social media data for HR BI&A. The hypothetical case, definitely requires further expansion and modeling using a recognized algorithm to make it applicable in developing tools that can be applied on huge data sets.

VIII. CONCLUSION

BA&I relies on huge data sets, often referred to as big data. Today, this data is available and surpasses the capabilities the quantities that existing tools can process and analyze efficiently. Such data is available ubiquitously, freely and most often than not its public on the web and precisely on social media. This data can be of significant help in BA&I if well utilized. To utilize this data in BA&I however, requires more advanced tools and techniques that can handle unstructured, and often stream data in a manner that creates value for businesses. The methods and techniques like Fuzzy association rules can be combined with appropriate algorithms to build tools that can address this challenge

REFERENCES

- [1] Association Rules | Solver. (n.d.). Retrieved August 20, 2016, from <http://www.solver.com/xlminer/help/association-rules>
- [2] Bryce, B. (2004). Assessing Business Intelligence (BI) and Analytics Tools for Your Organization. *Treasury & Risk Management*, 15(1), 18-19.
- [3] Cheese, P. (2013). An information revolution. *People Management*, 5.
- [4] Douglas, L. (2016). Experts: Used wrongly, social media can sting. *Michigan Lawyers Weekly*,
- [5] Douthitt, S., & Mondore, S. (2014). Creating a Business-Focused HR Function with Analytics and Integrated Talent Management. *People & Strategy*, 36(4), 16-21.
- [6] Eddy, N. (2014). Business Intelligence Software Adoption Still Lagging. *Eweek*, 1.
- [7] File, C., & Shah, C. (2012). InfoExtractor—A Tool for Social Media Data Mining. *Journal Of Information Technology & Politics*, 9(3), 269-278. doi:10.1080/19331681.2012.665259
- [8] Garrett, S. (2007). Winning the numbers game: Finding metrics that matter. *Employee Benefit News*, 21(3), 13-14.
- [9] Havenstein, H. (2006). Data Analysis Tools Can Deliver Competitive Advantage. *Computerworld*, 40(4), 7.
- [10] He, W., Shen, J., Tian, X., Li, Y., Akula, V., Yan, G., & Tao, R. (2015). Gaining competitive intelligence from social media data. *Industrial Management & Data Systems*, 115(9), 1622-1636. doi:10.1108/IMDS-03-2015-0098
- [11] Hsinchun, C., Chiang, R. L., & Storey, V. C. (2012). BUSINESS INTELLIGENCE AND ANALYTICS: FROM BIG DATA TO BIG IMPACT. *MIS Quarterly*, 36(4), 1165-1188.
- [12] IBM to Apply Analytics to Improve Customer Service and Loyalty; New 'Voice of Customer' Mines Diverse Data Sources to Deliver Consolidated View Into Customer Satisfaction Issues. (2009). M2PressWIRE,
- [13] JAIN, K. (2013). What is Business Analytics and which tools are used for analysis? Retrieved August 22, 2016, from <https://www.analyticsvidhya.com/blog/2013/06/business-analytics-spectrum/>

- [14] Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). EMERGING TRENDS IN BUSINESS ANALYTICS. *Communications Of The ACM*, 45(8), 45-48.
- [15] Konopnicki, D., Shmueli-Scheuer, M., Cohen, D., Sznajder, B., Herzig, J., Raviv, A., & ... Mass, Y. (2013). A statistical approach to mining customers' conversational data from social media. *IBM Journal Of Research & Development*, 57(3/4), 1-13. doi:10.1147/JRD.2013.2251833
- [16] Majid, A., Chen, L., Chen, G., Mirza, H. T., Hussain, I., & Woodward, J. (2013). A context-aware personalized travel recommendation system based on geotagged social media data mining. *International Journal Of Geographical Information Science*, 27(4), 662-684. doi:10.1080/13658816.2012.696649
- [17] Rouse, M. (2011, February). What is association rules (in data mining)? - Definition from WhatIs.com. Retrieved August 22, 2016, from <http://searchbusinessanalytics.techtarget.com/definition/association-rules-in-data-mining>
- [18] Vahn, G. (2014). Business Analytics In The Age Of Big Data. *Business Strategy Review*, 25(3), 8-9. doi:10.1111/j.1467-8616.2014.01083.x
- [19] Veset, D., & McDonough, B. (2007). The Next Wave of Business Analytics. *DM Review*, 17(3), 20-22.
- [20] Vohra, G., & Singh, G. (2011). 10 most popular analytic tools in business | Analytics Training Blog. Retrieved August 22, 2016, from <http://analyticstraining.com/2011/10-most-popular-analytic-tools-in-business/>
- [21] Walter, P. (2013, August). The data-mining of social media: Get used to it. *Washington Post*, The.
- [22] Welbourne, T. M. (2014). Does Size Really Matter? Getting to Results with Small, Medium and Big Data. *Workforce Solutions Review*, 5(2), 9-10.
- [23] Winterberg, B. (2010). Tools and Techniques to Leverage Social Media. *Journal Of Financial Planning*, 23(5), 40-43.