

# Investigating the Link between Web Data Mining and Strategic Human Resources Planning

Hanadi Al-Zegaier (Ph.D)

Assistant Professor, Management Information Systems Department, Applied Science University

Amman, Jordan

Tel: 962-796-604-432 E-mail: zegaier@asu.edu.jo

Hasan Ali Al-Zu'bi (Ph.D)

Professor, Business Administration Department, Applied Science University

Amman, Jordan

Tel: 962-795-629-808 E-mail: zubi1963@yahoo.com

Samer Barakat (Ph.D)

Assistant Professor, Management Information Systems Department, Applied Science University

Amman, Jordan

Tel: 962-975-768-545 E-mail: sbarakat@asu.edu.jo

Received: January 28, 2011

Accepted: February 25, 2011

doi:10.5539/cis.v4n3p67

## Abstract

This study aims to investigate the link between web data mining and strategic human resources planning - an application of web-based data mining that integrates online data collection and data mining in developing strategies for human resource planning. The study aims to focus particularly on web mining which may benefit societal areas by extracting new knowledge, providing support for decision making and empowering the effective management of societal issues. The study population consists of top staff in management information systems and business administration departments at the housing bank. A simple random sampling technique was used to select the respondents surveyed for this study with a total of 33 questionnaires administered to respondents chosen from 180 branches. Statistical tools were used to test the hypothesis. The findings of this study indicates that there are significant relationship between web data mining usage and the importance of strategic human resources planning; the study also shows that web data mining has a significant impact and use on Strategic human resources planning.

**Keywords:** Web data mining, Human resource, Strategic human resource planning

## 1. Introduction

Data mining process aims at extracting patterns from data. Business organizations rely on Data Mining as one of the important tools to transfer data into information and consequently knowledge [[http://en.wikipedia.org/wiki/Data\\_mining](http://en.wikipedia.org/wiki/Data_mining)]. On the other hand strategic information systems are used by organizations to achieve a competitive advantage for the organization. Strategic information systems relies on data gathered at all organizational levels and from all information systems whether it was transaction processing systems, management information systems, decision support systems or executive information systems.

In data mining what is important is the extraction of patterns form data. There have been several early methods to extract and identify patterns such as the Bayes' theorem in the 1700s and the regression analysis of the 1800s. In today's computer age data storage has grown in size to unthinkable ranges that only computerized methods can be applied to find patterns among these large repositories of data available to organizations whether it was online or offline.

Neural networks, genetic algorithms, clustering, decision trees and support vector machines are an example of new developed methods used by organization to uncover hidden patterns within data.[ Kantardzic, Mehmed

(2003). In this study, we demonstrate an integral relationship between web data mining and strategic human resources systems within organizations. We conducted our study on the Housing Bank which is a leading Jordanian bank with more than 180 branches. A questionnaire has been handed out to all branch managers and their deputies. Collected data has been entered into SPSS and several valid statistical tests have been implemented.

Currently, many data mining studies are focusing on data mining techniques rather than discussing the benefits of data mining to organizations and the decision making process. The application of data mining in business is not fully explored. The major contribution of this study is the building of a solid relationship between data mining and achieving competitive advantage in today's competitive agile business environment.

## 2. Literature review

### 2.1 Data Mining Tasks

Data mining involves the extraction of helpful needed information from large repositories of data, usually it is implemented in situations and scenarios where information needed to be extracted from huge data volumes which exceeds any individual's investigative capabilities. Data mining offers business organizations enhanced intelligent functionalities to conduct data extraction from large chunks of data and providing them with vital information leading to solving several organizational problems and providing top management with the tools to achieve a competitive advantage stance within competing organizations.

Agrawal (1993) listed the following four tasks of Data mining methodologies which involve several tasks of which we shall be focusing on:

- 1) Association Rules
- 2) Clustering
- 3) Trend Detection
- 4) Sequential Pattern

### 2.2 Empirical studies on strategic human resources planning (SHRP)

Chang and Huang (2005) mentioned in their article the recent studies which examined strategic human resource management (SHRP) as a means of enhancing organizational competitive advantage. Scholars and practitioners have widely adopted this approach to organization strategic planning. The underlying assumption of SHRP is that a firm performance is influenced by a set of HRM practices. However, important questions remain, including whether SHRP guarantees positive firm performance outcome, the effect of different levels of SHRP implementation on a firm performance, and the influence of the market environment in moderating the relationship between SHRP and a firm's performance (Huselid, 1995). A critical unresolved debate involved whether a universally superior approach existed for managing human resources. Some scholars suggested that such an approach existed (Huselid and Becker, 1996; Pfeffer, 1998). Meanwhile, others noted that the effectiveness of human resource management practices depends on the specific organizational and environmental context.

Recent theoretical research on business strategy has indicated that a firm's competitive advantage could be created by the firm's human resources (HR). According to the resource-based view the firm could develop sustained competitive advantage through creating value in a manner that is rare and difficult for competitors to imitate. Traditional sources of competitive advantage, such as natural resources, technology and economies of scale have become increasingly easy to imitate. The concept of HR as a strategic asset has implications on this issue. HR is an invisible asset that creates value when it is embedded in the operational system in a manner that enhances the firm's ability to deal with a turbulent environment (Stuart, 2007).

The survey referred to earlier of over 1000 HR managers from organizations in 47 countries reported that HR managers believed that organizations that align their HR and business strategies are more profitable (Wright and Boswell, 2002). As a direct result of increasingly available information systems, the human resource department can use numerous strategies to contribute to the bottom line. The basic idea behind these strategies is to translate knowledge of human resources into terms that have tangible and recognizable economic benefits, especially to operating managers. Examples include analysis of the cost per hire, length of time to fill a position, and new-hire performance by recruiting strategy, with the intent of identifying the most effective strategy. (Byars & Rue, 2006)

### 3. Operational definitions

Strategic human resource planning (SHRP): Mondy and Noe (2005) define (SHRP) as the processes of systematically reviewing human resource requirements to ensure that the required numbers of employees, with the required skills, are available when and where they are needed.

-Employment security: Protection of employer's facilities and equipment from unauthorized access and protection of employees while on work premises or work assignments. (Rothwell & Dubois, 2004)

-Recruitment process: the process of developing a pool of qualified applicants who are interested in working for the organization and from which the organization might reasonably select the best individual or individuals to hire for the employment. (Denisi & Griffin, 2001)

Career planning: the process of deciding what positions the firm will have to fill, and how to fill them. (Dessler, 2003)

Human capital management: the policies and practices involved in carrying out the knowledge, education training, skills and expertise of a firm's workers. (Dessler, 2003)

Extensive training: is a plan attempted to facilitate employees' learning of job-related knowledge, skills, and behaviors. (Denisi & Griffin, 2001)

Data mining: consist the following four tasks:

1. Association Rules: The association of rules was introduced by Agrawal et al. [1993]. It involves finding relationships between variables. For example a bank might gather data on its client's withdrawal activities. Using the association rule can help the bank manager to learn which period of the month has high withdrawals rate and thereby managing the banks cash liquidity during these periods.
2. Clustering: Clustering involves discovering groups and structures in the data that are similar based on shadowy characteristics. An example is to find clusters of clients with similar levels of deposits.
3. Trend Detection: Trend detection is finding similarity among a group of objects. A good example of trend detection is: The older clients are the more savings they tend to make.
4. Sequential Pattern: This is used to detect sequences of events, such as depositing and withdrawal of funds.

### 4. The objectives of the study

The study's objectives is to build and test a suggested model that supports the relation ship between web data mining and strategic planning for human resource . We try to test the suggested model at the housing bank in Jordan. And here are some more objectives related to the statement of the research:

- 1- Determine the relationship between web data mining usage and participant opinions about the importance of strategic human resources planning (SHRP) at the housing bank.
- 2- Determine the most significant impact and use of web data mining on strategic human resources planning (SHRP) at the housing bank.

### 5. Suggested model

Based on Agrawal (1993) web data mining tasks and pfeffer (1994) strategic human resources planning Practices the researchers model is developed.

A model consists of two types of variables, the independent variable (web data mining) and the dependent variable (strategic human resources planning as shown in figure (1).

### 6. Research Hypotheses

To answer the questions posed by the authors, and based on the literature review, the researchers proposed two main null hypotheses as follows:

Ho1: There is no significant relationship between Web Data Mining usage and Participant's Opinions about the Importance of SHRP.

In order to test this first main hypothesis, there are sub-hypotheses. These sub-hypotheses are:

Ho1-a: There is no significant relationship between Association rules and Participant's Opinions about the Importance of SHRP.

Ho1-b: There is no significant relationship between Clustering and Participant's Opinions about the Importance of SHRP.

Ho1-c: There is no significant relationship between Trend Detection and Participant's Opinions about the Importance of SHRP.

Ho1-d: There is no significant relationship between Sequential Pattern and Participant's Opinions about the Importance of SHRP.

Ho2 There is no significant relationship between Web Data Mining usage and Participant's Opinions about the Use of SHRP.

In order to test the second main hypothesis, there are sub-hypotheses. These sub-hypotheses are:

Ho2-a: There is no significant relationship between Association rules and Participant's Opinions about the Use of SHRP.

Ho2-b: There is no significant relationship between Clustering and Participant's Opinions about the Use of SHRP.

Ho2-c: There is no significant relationship between Trend Detection and Participant's Opinions about the Use of SHRP.

Ho2-d: There is no significant relationship between Sequential Pattern and Participant's Opinions about the Use of SHRP.

## 7. Research Methodology

### 7.1 Population and Sample

This section describes the population through the general characteristics of the respondents in term of gender, age, educational level, working experience, working position, as shown in table (1) below.

The data in table (1) shows that the majority of Respondents more than two third (81.1) were males and (18.2) were Females. The data also shows that respondents who are between the ages of 25-30 represented the highest percentage among respondents with a (30.3%).

Respondents less than 25 years old had a percentage of (6.1%), while respondents aged 31-35 years were (12.1%), and only (15.2%) were for ages over 45 years.

The data in table (7) show that the highest percentage (63.6%) was for the University degree holders. Only (3%) attained post graduate qualification.

The respondents had a middle and long experience in their companies, (21.2%) had (1-5) years experience and (24.2%) had (6-10) years experience and (12.1%) had more than 10 years working experience, and (42.4%) had more than 16 years work experience.

### 7.2 Data Collection

The data and information was gathered from two resources:

#### 7.2.1 Primary Resources

A study questionnaire was designed to get the primary resources and included two parts: -

First: classify data

Second: Measuring the independent and dependent variables performed together with necessary data to test the study hypothesis.

#### 7.2.2 Secondary Resources

Using the scientific references (Books, articles, etc...) concerned with the study's subject.

### 7.3 Research Instrument

Strategic human resources planning variables which include:

Variables	Source
Employment Security	<b>Chang,Huang(2005)</b>
Recruitment Process	<b>Chang,Huang(2005)</b>
<b>Career Planning</b>	<b>Chang,Huang(2005)</b>

Web data mining variables which include:

Variables	Sources
Association Rules	Agrawal, T. Imielinski, and A. Swami (1993)
Clustering	Agrawal, T. Imielinski, and A. Swami (1993)
Trend Detection	Agrawal, T. Imielinski, and A. Swami (1993)
Sequential Pattern	Agrawal, T. Imielinski, and A. Swami (1993)

#### 7.4 Instrument validity and reliability

##### 7.4.1 Instrument validity

The content validity of the questionnaire was determined by a panel of experts in the fields of, Web Data Mining, Management, and Information Systems.

##### 7.4.2 Instrument reliability

The reliability of the survey instrument was assessed through Cronbach's coefficient alpha ( $\alpha$ ). All coefficients alpha were within acceptable ranges for comparable instruments. (Sekran, 2006)

#### 7.5 Data Analysis Procedure

The analysis was conducted over two stages: instrument validation and hypothesis testing. Statistical techniques was mainly employed in order to examine the dependent variables and knowledge transfer separately

#### 7.6 Study population, and analysis unit

To be able to examine the research problem and the questions raised in that respect and in order to fulfill the objectives of this study, the researchers has chosen to conduct the research at the housing bank for trading and finance which is a one of the leading Jordanian banks with more than 180 branches.

The survey questionnaire was sent to (50) participants; (40) questionnaire were returned from the targeted population, (7) questionnaire were excluded from the analysis leaving (33) questionnaires that were included in the analysis.

#### 7.7 Unit of analysis

The unit of analysis is the Top Managers, Human Resources Managers, and Human Resources Staff at the housing bank.

## 8. Results

In order to test the first main hypothesis (Ho1), there are four minor ones. These hypotheses are:

(Ho1-a) There is no significant relationship between Association rules and Participant's Opinions about the Importance of SHRP

Pearson correlation was used to test above Hypothesis and it was found that there are significant relationship at (0.05) level between independent variable (Association) and Dependent variable (strategic human resource planning importance), we can find appositve and significant effect at function level (0.01) which supports hypothesis (Ha1-a), where ( $r=0.638^{**}$ ) which is moderate.

(Ho1-b): There is no significant relationship between Clustering and Participant's Opinions about the Importance of SHRP

Pearson correlation was used to test above Hypothesis and it was found that there are significant relationship at (0.05) level between independent variable (Clustering) and Dependent variable (strategic human resource planning importance), we can find appositve and significant effect at function level (0.01) which supports hypothesis (Ha1-b), where ( $r=0.660^{**}$ ) which is moderate.

(Ho1-c) There is no significant relationship between Trend Detection and Participant's Opinions about the Importance of SHRP.

(Strategic Pearson correlation was used to test above Hypothesis and it was found that there are significant relationship at (0.05) level between independent variable (Trend detection) and Dependent variable (human

resource planning importance), we can find appositve and significant effect at function level (0.01) which supports hypothesis (Ha1-c), where ( $r=0.479^{**}$ ) which is moderate.

(Ho1-d) There is no significant relationship between Sequential Pattern and Participant's Opinions about the Importance of SHRP.

Pearson correlation was used to test above Hypothesis and it was found that there are no significant relationship at (0.05) level between independent variable (Sequential pattern) and Dependent variable (strategic human resource planning importance), we can find negative and no significant effect, where ( $r=0.316$ ) which is moderate.

In order to test the second main hypothesis (Ho2), there are four minor ones. These hypotheses are:

(Ho2-a) There is no significant relationship between Association rules and Participant's Opinions about the Use of SHRP.

Pearson correlation was used to test above Hypothesis and it was found that there are no significant relationship at (0.05) level between independent variable (Association rules) and Dependent variable (strategic human resource planning use), we can find negative and no significant effect, where ( $r=0.274$ ) which is moderate.

(Ho2-b) There is no significant relationship between Clustering and Participant's Opinions about the Use of SHRP.

Pearson correlation was used to test above Hypothesis and it was found that there are no significant relationship at (0.05) level between independent variable (Clustering) and Dependent variable (strategic human resource planning use), we can find negative and no significant effect, where ( $r=0.330$ ) which is moderate.

(Ho2-c) There is no significant relationship between Trend Detection and Participant's Opinions about the Use of SHRP.

Pearson correlation was used to test above Hypothesis and it was found that there are significant relationship at (0.05) level between independent variable (Trend detection) and Dependent variable (strategic human resource planning use), we can find appositve and significant effect at function level (0.05) which supports hypothesis (Ha1-c), where ( $r=0.395^{*}$ ) which is moderate.

(Ho2-d) There is no significant relationship between Sequential Pattern and Participant's Opinions about the Use of SHRP.

Pearson correlation was used to test above Hypothesis and it was found that there are significant relationship at (0.05) level between independent variable (Sequential Pattern) and Dependent variable (strategic human resource planning use), we can find appositve and significant effect at function level (0.01) which supports hypothesis (Ha1-d), where ( $r=0.486^{**}$ ) which is moderate.

## 9. Conclusion

This study has investigated the relationship between web data mining and strategic human resources planning through developing a practical model. The tools of analysis for the web data mining performed were four tasks while we used three variables to test the strategic human resources planning. According to statistical results the first and second null hypotheses were rejected which means that there is a significant relationship between web data mining usage and the importance of strategic human resources planning. Furthermore, the results showed a good relationship between the usage of web data mining and strategic human resources planning, especially at the recruitment process at the housing bank. Also this paper has intended to draw the attention of the web data mining community to research opportunities in more practical and benefit areas. On the other hand, this work also intended to draw the attention of researchers working on the web in various disciplines such as strategic planning in order to consider the web data mining for improving their applications.

## 10. Study Limitations

The limitations that confronted the researchers while conducting the study can be summarized as follows:

Personal bias: some of the information in this study reflects the judgment of individuals who provided the information and may have biases favoring the success of their companies.

The other limitations is the lack of sufficient levels of data, the reason behind that is that many companies consider data related to human resources confidential, or at least should not be publicized, however, the appropriate analysis unit and ensuring the anonymity of responses should solve this problem.

## 11. Suggestions for Future Research

The future directions of the research is to investigate other web mining tasks and their available web mining technologies for analyzing various types of data, and making comparisons of the capabilities of these tasks between and among each other as well as the four tasks used in this paper. Future research may also include the acquisition of other data sets to perform new analyses and comparisons. Future directions would also include applications by mining of web usage, web content, and web structure and the availability of technologies for each of these.

## References

- Agrawal, R, Imielinski, T, and Swami, A. (1993). Mining Association Rules Between Sets of Items in Large Databases. In Proc. 1993 ACM-SIGMOD Int. Conf. Management of Data.
- Aswathappa, K. (2002). *Human resource management*. 3rd edition. New York U.S.A.: Tata McGraw-Hill Publishing Company.
- Boris, Grigor & Evich, Mirkin. (2005). *Clustering for data mining: a data recovery approach*. Illustrated ed., Chapman & Hall/CRC.
- Boxall, P., Purcell, J. (2000). Strategic human resource management: where have we come from and where should we be going?. *International Journal of Management Reviews*, 2, (2), 183-203.
- Byars, L & Rue L. (2006). *Human resource management*. 8th edition. New York, U.S.A, McGraw-Hill.
- Byham, William. (2003). developing dimension/competency-based human resource systems. Development dimensions international press, 1-33.
- Chang, w., Huang, T. (2005). Relationship between strategic human resource management and firm performance, *international journal of manpower*, 26, (5), 434-449.
- Conner,J & Ulrich,D. (1996 ). human resource roles: creating value, not rhetoric. *human resource planning*, 19,(3),1-38.
- Currie, G., Kerrin, M. (2003). Human resource management and knowledge management: enhancing knowledge sharing in a pharmaceutical company. *International Journal of Human Resource Management*, 14, (6), 1027-45.
- David, Fred. (2001). *strategic management, concepts and cases*. 8th ed, New Jersey, prentice hall.
- DeNisi & Griffin. (2005). *Human resource management*, New York U.S.A. Houghton Mifflin Company.
- Dessler, Gary. (2003). *Human resource management*. 9th ed, prentice hall, new jersey, U.S.A.
- Dubois,D.&Rothwell,W. (2004). *Competency-based human resource management*. Davies-black publishing, U.S.A.
- Huselid, M.A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, 38, 635-72.
- Huselid, M.A., Becker, B.E. (1996). Methodological issues in cross-sectional and panel estimates of the human resource-firm. *Industrial Relations*, 35, 400-22.
- Huselid, M.A., Jackson, S.E., Schuler, R.S. (1997). Technical and strategic human resource management effectiveness as determinants of firm performance. *Academy of Management Journal*, 40,171-88.
- Kantardzic, Mehmed. (2003). *Data Mining: Concepts, Models, Methods, and Algorithms*. John Wiley & Sons.
- Mcnurlin, B&Sprague, R. (2004). *Information systems management in practice*. 6th ed, Jersey U.S.A., Prentice- Hall.
- Mondy, R & Noe, R. (2005). *Human resource management*. 9th ed, New Jersey U.S.A., Prentice- Hall.
- Narasimha, S. (2000). Organizational knowledge, human resource management and sustained competitive advantage: toward a framework. *Competitiveness Review*, 10, (1), 123-35.
- Neg, E.W.T, Wat, F.K.T. (2006). Human resources information systems: a review and empirical analysis. *Personal Review*, 35, (3), 1-13.
- Ng, S.T., Skitmore, R.M., Sharma, T. (2001). Towards a human resource information system for Australian construction companies. *Engineering, Construction and Architectural Management*, 8, (4), 238-49.
- Noe, R., Hollenbeck, J., Gerhart, B. & Wright, P. (2000). *Human resource management: gaining a competitive advantage*. 3rd ed, USA, Irwin mc Graw – hill.

- Orlando, C.R., Johnson, N.B. (2004). High performance work practices and human resource management effectiveness: substitutes or complements?. *Journal of Business Strategies*, 21, (2), 133-48.
- Patterson, S. (2002). Lack of money, staff and time: toughest part of managing HRIS. *Human Resource Department Management Report*, 2, (12), 3-5.
- Patton, Witt, Iovrich & Fredericksen. (2002). *Human resource management*, Houghton Mifflin Company, Boston, U.S.A.
- Pfeffer, J. (1994). Competitive advantage through people. *California Management Review*, 36, (2), 9-28.
- Pfeffer, J. (1998). Seven practices of successful organizations. *California Management Review*, 40, (2), 96-124.
- Sekran, uma. (2006). *Research methods for business*. 3rd ed, John Wiley & sons, USA.
- Stuart, M. (2007). The role human resource management in information development. *Management decision*, 45, (2), 252-264.
- Wright, P.M., Boswell, W.R. (2002). Desegregating HRM: a review and synthesis of micro and macro human resource management research. *Journal of Management*, 28, 247-76.
- Wright, P.M., McMahan, G.C. (1992). Theoretical perspectives for strategic human resource management. *Journal of Management*, 18, 295-321.
- Wright, P.M., Sherman, A.S. (1999). Introduction: strategic human resource management research in the 21st century. *Human Resource Management Review*, 8, (3), 187-91.
- Wright, P.M., Snell, S.A. (1991). Toward an integrative view of strategic human resource management. *Human Resource Management Review*, 1, 203-25.
- [http://en.wikipedia.org/wiki/Data\\_mining](http://en.wikipedia.org/wiki/Data_mining)

Table 1. Demographic Characteristics of respondents (n=33)

Characteristics	Frequency	Percentage
Gender:		
Male	27	81.8
Female	6	18.2
Age:		
Less than 25 years	2	6.1
25 – 30	10	30.3
31 – 35	4	12.1
36 – 40	5	15.2
41 – 45	7	21.2
46 years and more	5	15.2
Educational Level:		
Bachelors	21	63.6
Master	11	33.3
PhD	1	3.0
Experience Years:		
1 – 5	7	21.2
6 – 10	8	24.2
11 – 15	4	12.1
16 years and more	14	42.4

Table 2. Means, Standard Deviations and Pearson Correlation among the research variables

variables	M	S.D	1	2	3	4	5	6
Association rules	3.96	0.420	-					
Clustering	3.79	0.506	0.550**	-				
Trend Detection	3.62	0.635	0.666**	0.495**	-			
Sequential Pattern	3.81	0.493	0.278	0.556**	0.489**	-		
Importance of SHRP	4.11	0.552	0.638**	0.660**	0.479**	0.316	-	
Use of SHRP	3.99	0.524	0.274	0.330	0.395*	0.486**	0.324	-

\*\* Correlation is significant at the 0.01 level.

\* Correlation is significant at the 0.05 level.

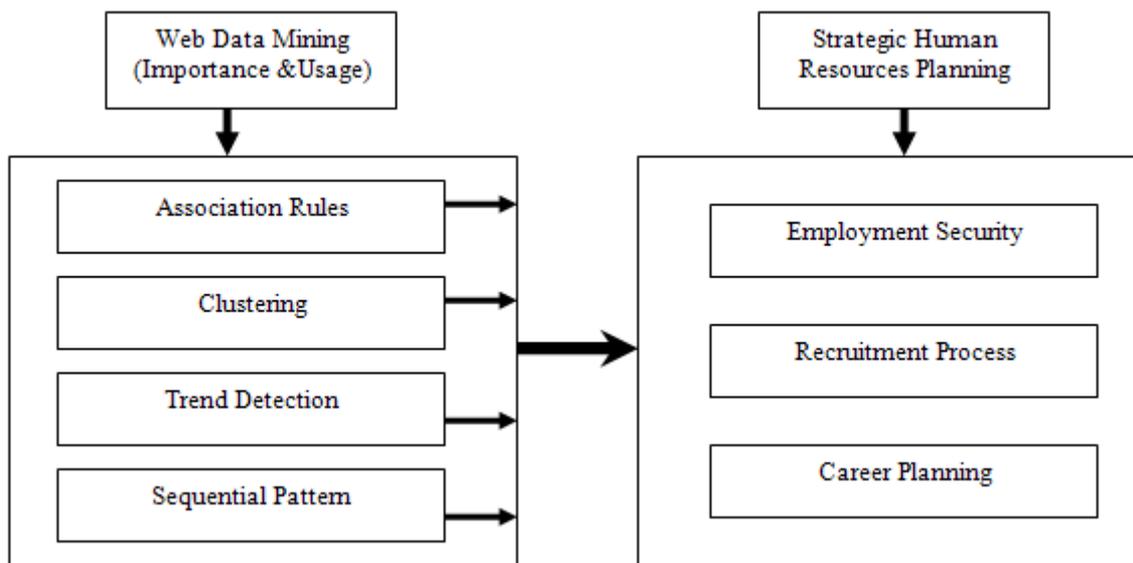


Figure 1. Research Model