

Predictive Analysis on Employee Performance

S L Tendean¹, M R Mashudi¹ and E Budiarto^{1*}

¹ Department of Information Technology, Swiss German University, Jl. Jalur Sutera Barat 15, Alam Sutera, Tangerang 15143, Indonesia

*Corresponding author, email: eka.budiarto@sgu.ac.id

Abstract. Human resources management have key roles during running some business. Some good investment in this human resources managements gives some benefit for company to gain more income. Hiring good employee and build some good organization inside company are needed to achieve good running business. The objective in this research is to determine some parameters from some performance record in X company that have key roles for human resources management to give some good hiring affecting business income. The Data that used got from the X company from 2014-2018 with around 14 parameters and 10 human performance record parameters. From this needs, this research describe some models such as multiple regression and random forest as validation focusing on analyzing the human performance record correlated to the revenue data set in PT X Company. As conclusion, this research concluded 3 parameters in human performance record that very important to determine good running business .

1. Introduction

1.1. Background

Around the world, Human Resources have several challenges that they are facing especially in small business. The big company have aware to this condition because they have spent big money to maximize their human resources management works and roles. The main role of human resources management is to build good organization so the company can run more efficient and gain more income. The main challenges that have mostly company facing in their human resources management around the world, which are (Gawke et al., 2018):

- Leadership Development: many small company or small running business doesn't mention and keep eye on their management team. This matter will cause their team lack of responsibility doing their job
- Change of Management Level : Along with the running business slightly become more bigger, the employee inside the company didn't aware about this matter and cannot adapt along the internal main progress and will make the work productivity decreased.
- Recruiting Employees: This problem still maintain in many companies that didn't invest good enough money in their HRM management. Some of company does not have any rules or standard in their human resources management to recruit new people. Some of company just recruit average people and could be become great impact for their running business

In PT X company, the problem occurring is the lack of knowledgement about the human resources management key roles. The manager sometimes didn't know current problem they are facing during running some business example like: In 1-2 years their sales income are fluctuative. The Human Resources Management also doesn't have any standard requirement for hiring new employee based on their performance record data so some of the new employee only works 1-3 months.

1.2. Human Capital and Business

Human capital is very important during running some business. The key of element in human capital were come from the human resource management. They maintain condition of organizational and employee carrer development. Firstly human capital can be described as a key component in moving forward a firm resources and workers in arrange to extend their main income as well as maintain their business. Human capital is also essentially can be described as a way that connected into instruction, preparing, and other proficient activities to build the levels of knowledge, ability, capacities and social resources of an employee

1.3. Regression Analysis

Regression analysis is sort of modelling technique that investigates correlation between a dependent and variable quantity. This technique is employed for prediction, statistic modelling and finding casual result between the variables. This research is using regression analysis for modelling for several reason (Shah and Aman, 2019):

- Regression analysis helps human resources management to ensure skills requirement for the organization and execute plan for recruitment process
- Related to revenue, regression analysis also used to predict number of employee and correlated to human performance record management according to target and current sales condition.

2. Material and Method

2.1. Research Framework

In this research, the experiments is focusing in organization area in PT X Company which is the first step to develop good individual skills. Good individual skills are develop not only from the employee itself but also got some support from human resources management to maintain their satisfaction and their skills during doing some works. The human resources management have key roles to develop good employee skills and good organization to be later can build some good operation and future strategic planning. Applying data mining method to define the selection of parameters is needed from Data Set in PT X Company. The Research framework to define the parameters is provided in Figure 1.

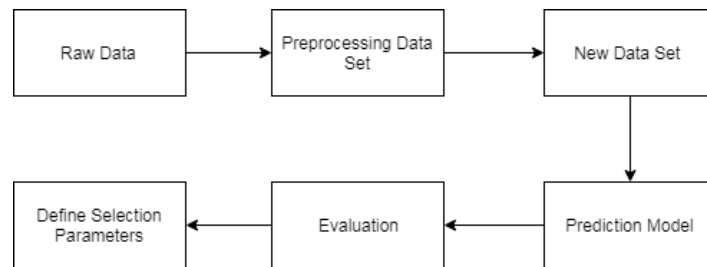


Figure 5 Research framework

2.2. System Overview for Analyze Data Set

In this subchapter, the overview of software system is given based on the research framework. Using data set of human resources management yearly data contains of employee performance and some of their personal data from X company taken around 4-5 years and conducting some propose method of experimental procedure to analyze the data is explained in Figure 2

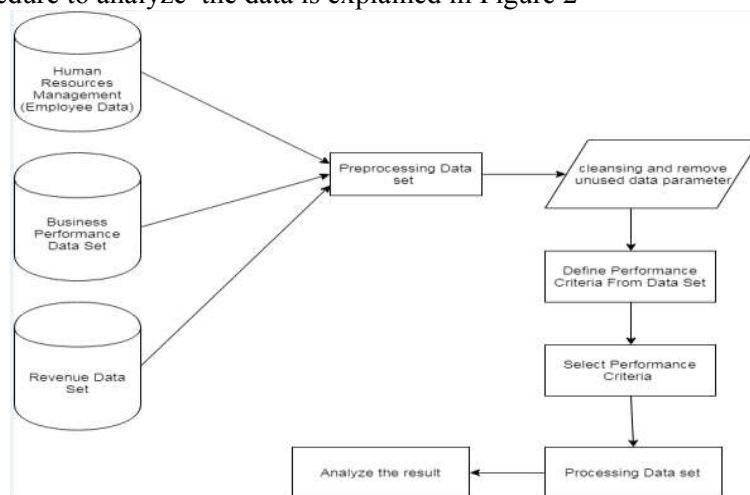


Figure 6 System Overview

2.3. Data Source Collections

In this research, the data are taken from PT. X company as data source. The data content are used by the human resource management for yearly report for employee performance along with some of their personal data. From Performance Data data set and human resources data will be later correlated to Revenue Data taken respectively from 2014-2018 to select the parameters that affecting running business. The Revenue Data set that will be used in this research is shown from Table 1 until Table 3. In this research, multiple and linear regression is best method to analyze the performance data set due to availability of data in numerical in PT X Company for selecting parameters. Revenue Data set is needed in this research to correlate between employee performance especially in Sales Department which are divided into 3 department (Disposable, Electrical and Laboratory) with the revenue in PT X Company

Table 1 Human Resources Data Parameters

| No | Features |
|----|-------------------|
| 1 | Date Joined |
| 2 | Attendance Number |
| 3 | Education |
| 4 | Place of Birth |
| 5 | Date of Birth |
| 6 | Age |
| 7 | Retirement Age |
| 8 | Position |
| 9 | Department |
| 10 | Address |
| 11 | Employee Status |
| 12 | Marital Status |
| 13 | Phone Number |
| 14 | Identity Number |

Table 2 Performance Parameters

| No | Features |
|----|-----------------------|
| 1 | Appearance |
| 2 | Knowledge |
| 3 | Enthusiastic |
| 4 | Quality |
| 5 | Responsibility |
| 6 | Innovation/Initiative |
| 7 | Behavior |
| 8 | Teamwork |
| 9 | Capability |
| 10 | Work on Time |

Table 3 Revenue Data Set

| NO | KETERANGAN | DIVISI DISPOSIBLE | DIVISI LABORATORIUM | DIVISI ELECTRIC | TOTAL |
|----|--------------------|-------------------|---------------------|-----------------|-----------------|
| 1 | PENJUALAN TH. 2014 | 198,970,761,460 | 28,089,989,853 | 7,022,497,463 | 234,083,248,776 |
| 2 | PENJUALAN TH. 2015 | 225,782,369,598 | 29,882,960,682 | 9,960,986,894 | 265,626,317,174 |
| 3 | PENJUALAN TH.2016 | 255,218,660,776 | 29,167,846,946 | 7,291,961,736 | 291,678,469,458 |
| 4 | PENJUALAN TH. 2017 | 247,192,122,409 | 37,078,818,361 | 6,543,320,887 | 290,814,261,657 |
| 5 | PENJUALAN TH. 2018 | 238,813,726,113 | 35,822,058,917 | 6,321,539,808 | 280,957,324,838 |

3. Result and Discussions

The company starts using performance data set for every supervisor and higher manager in several department to give some scoring according to their team and their colleague regarding their performance during works. The scoring of this performance record later is given to human resources management to give some credit and appraisal. However, in PT X Company, Human Resources Management only uses this performance data to arrange every employee's appraisal and benefit according to their performance. In this chapter, the result of multiple regression is applied to determine selection parameters in performance record correlated to the revenue. The Electrical Department is chosen as testing because the coefficient correlation result between parameters in performance record have highest values

3.1. Regression Result

Table 4 Summary of Regression Correlation Result

| Regression Statistic Value | Disposable Department | Laboratory Department | Electrical Department |
|--------------------------------------|-----------------------|-----------------------|-----------------------|
| R-Square Value | 0.5073 | 0.7261 | 0.4443 |
| Multiple R (correlation Coefficient) | 0.7122 | 0.8521 | 0.6665 |

Based on result from Table 4, The performance parameters have been reduced from 10 into 5 by using statistical correlation. From previous study it is indicated that the correlation between performance record and revenue have good relationship with the performance record from electrical Department have high coefficient correlation result and have good R square. According to previous study, it is indicated that the value of R-squared > 0.3898 have good relationship in regression analysis and can be applied as references in financial performance of some company (Shara et al., 2019). The value of R-squared relative small because some of data in company give some similiarity of data distribution and will give some impact for prediction result. From Table 4, the selection of performance parameters cannot be determined because the result of the regression analysis indicated the total value of performance record of 5 selection parameters record. Hence another method multiple regression will be implemented

3.2. Multiple Regression Result

The result of this multiple regression result reduced from 5 parameters which are Pengetahuan Pekerjaan (Knowledge), Tanggung Jawab (Responsibility), Inovasi Pekerjaan (Innovation), Kapabilitas Dalam Berkerja (Capability), and Kerja Tepat Waktu (Work On Time) into 3 parameters which are Tanggung Jawab (Responsibility), Inovasi Pekerjaan (Innovation) and Kapabilitas Dalam Berkerja (Capability) according to P-values and significance values because it is need to be reduced until below the value below 0.05 to give more satisfied result with by reducing some error as shown on Table 5. Significance value below 0,05 means that the regression model can be accepted and each element of data have correlated into the revenue

Table 5 Regression Statistical Electrical Department Result

| Regression Statistic | Value Before Parameters Removal (5 Parameters) | Value After Parameters Removal (3 Parameters) |
|--|---|--|
| Multiple R Value (Coefficient Correlation) | 0.8945 | 0.890027 |
| R-Squared Value | 0.8008 | 0.792148 |
| Significance | 0.1403 | 0.018042 |

The statistical of each parameters of multiple regression result after parameters removal is represented in Table 6.

Table 6 Regression Value after Parameters Removal

| | Coefficient | P-values |
|------------------------------|-------------|----------|
| Intercept | 17428340686 | 0.003467 |
| Tanggung Jawab | -3007425308 | 0.012554 |
| Inovasi Pekerjaan/ Inisiatif | 2078877232 | 0.022566 |
| Kapabilitas Dalam Bekerja | -1560569854 | 0.062099 |

Based on result on Table 6, The last 3 parameters have crucial points based on the data in human resource management in PT X Company affecting running business. The multiple regression equation is represented in equation (1) below

$$y = 17428340686 - 3007425308 A + 2078877232 B - 1560569854 C \quad (1)$$

Based on equation above the following parameters indicated:

- Y=predicted Revenue
- A= Performance Record Value of (Tanggung Jawab)
- B= Performance Record Value of (Inovasi Pekerjaan/Inisiatif)
- C= Performance Record Value of (Kapabilitas Dalam Bekerja)

From Equation 1 above, from 3 important parameters, the innovation value have key important for PT X Company to increase their revenue. By providing some training to give opportunity for their employee to develop their skills and know their capability of works.

3.3. Validation Model

In this research random forest will be used to validate our selection parameters and model using regression is good or not to compare the correlation of the coefficient. This method is suitable for PT X Company data set due to small number of data size and this validation method will give more accurate prediction parameters with small number data set and this method have correlated with the multiple regression model based on it's correlation coefficient and current data availability. The result of the comparison model is given on Table 7.

Table 7 Model Validation Result between Multiple Regression and Random Forest

| Regression Statistic | Multiple Regression | Random Forest |
|-------------------------------------|----------------------------|----------------------|
| Multiple R(Correlation Coefficient) | 0,89 | 0,8851 |
| R-Squared Value | 0,7921 | 0,7835 |

4. Conclusion

Based on purpose of this research is to determine performance record parameters that affecting running business which are revenue in PT X Company. Firstly the selection of performance record from 10 parameters is reduced to 5 parameters based on previous studies(Sutanto, 2004). Using regression analysis model can be determined the correlation between those 5 parameters performance records and revenue however the models shows that the parameters cannot be as individual parameters but as one parameters give better result. Therefore similiar Regression method called Multiple Regression applied to Electrical Department Performance Record Data due to it's coefficient correlation is very good. Further analysis shows that our selection parameters which are 5 attributes from statistical result can be reduced to 3 parameters and give good regression result and can be determined as individual parameters for human resources management in PT X Company.

To validate the regression result, another method which is random forest give nearly similiar result of coefficient correlation result. In Conclusions, our hypothesis to select several parameters can be proved by using model in this research. However based on our hypothesis, knowledge parameters doesnt give correlation to running business in PT X Company. The parameters that should be considered in human resources PT X Company are Innovation, Responsibility and Capability, this parameters also have some correlated with previous study indicated that the innovation should be considered in organization to build employee's skills and capability of their works along with their knowledge (Zarraga-Oberty et al., 2007) (Shah and Aman, 2019).

From the result of this research, it can be concluded that the 3 parameters which are Responsibility, Innovation and Responsibility from human resources performance data set parameters are important to determine good running business in PT X Company. For further research, this 3 parameters can be apply into another small or medium company with bigger data set to determine the accuracy and correlation coefficient to get better result and comparison from this small data set. More over, another method can be also applied to explore more parameters that very important in human resources department to determine good business and revenue.

References

- Shah, M., Aman, Q., 2019. The impact of human resource management practices on leadership styles: The mediating role of employee trust. *City Univ. Res. J.* 9, 58–71.
- Shankar, R.S., Rajanikanth, J., Sivaramaraju, V., Murthy, K.V., 2018. PREDICTION OF EMPLOYEE ATTRITION USING DATAMINING, in: 2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA). IEEE, pp. 1–8.
- Shara, Y., Muda, I., Rujiman, R., 2019. Role of Organizational Commitment to the Factor of Performance of Regional Financial Management, in: 1st Aceh Global Conference (AGC 2018). Atlantis Press.
- Sutanto, E.M., 2004. Forecasting: the key to successful human resource management. *J. Manaj. Dan Kewirausahaan* 2, 1–8.
- Thoman, D., Lloyd, R., 2018. A Review of the Literature on Human Resource Development: Leveraging HR as Strategic Partner in the High Performance Organization. *J. Int. Interdiscip. Bus. Res.* 5, 147–160.
- Thomas, R.J., Cheese, P., Benton, J.M., 2003. Human capital development. *Accent. Inst. High Perform. Bus.*
- Tinsley, H.E., Brown, S.D., 2000. *Handbook of applied multivariate statistics and mathematical modeling*. Academic press.
- Van Der Aalst, W., 2016. Data Mining, in: *Process Mining*. Springer, pp. 89–121.
- von Bonsdorff, M.E., Zhou, L., Wang, M., Vanhala, S., von Bonsdorff, M.B., Rantanen, T., 2018. Employee age and company performance: An integrated model of aging and human resource management practices. *J. Manag.* 44, 3124–3150.
- Xia, J., Mandal, R., Sinelnikov, I.V., Broadhurst, D., Wishart, D.S., 2012. MetaboAnalyst 2.0—a comprehensive server for metabolomic data analysis. *Nucleic Acids Res.* 40, W127–W133.
- XIONG, S., Chen, L., CHANG, L., XIE, A., 2019. Impact Analysis of Financial Early Warning Indicators Based on Random Forest. *DEStech Trans. Comput. Sci. Eng.*
- Zarraga-Oberty, C., Bonache, J., others, 2007. Human factors in the design of revenue management systems in multinational corporations. *Int. J. Revenue Manag.* 1, 141–153.
- Gawke, J.C., Gorgievski, M.J., Bakker, A.B., 2018. Personal costs and benefits of employee intrapreneurship: Disentangling the employee intrapreneurship, well-being, and job performance relationship. *J. Occup. Health Psychol.* 23, 508.
- Cheng, Z., Chen, Y., 2012. Data mining applications in human resources management system. *J. Converg. Inf. Technol.* 7.