Data Analytics for Transforming Towards Smart Supplier Relationship Management (A Case Study in Manufacturing Company)

Dewanti Anggrahini, Nani Kurniati, Anggi Prienda Sukma
Sepuluh Nopember Institute of Technology
Surabaya, Indonesia
dewanti@ie.its.ac.id, nanikur@ie.its.ac.id, anggiprienda@gmail.com

Abstract

Most of large-scaled manufacturing company deals with more than fifties suppliers, which are located within the country and abroad. It is also challenging to manage once each partner has different characteristics. Thus, a proper supplier relationship management (SRM) program is needed. Most of the company has gathered huge number of data; however, they rarely process it optimally. It drives to a longer time consumed to determine strategic decisions making. On the other hand, the business competition is very dynamic. As consequences new approach to utilise the big data is highly demanded. Data analytic provides a solution to manage the SRM in easier, more efficient, and effective way. This research aims to implement data analytics on managing manufacturer and strategic partners relationship, which accommodates supplier performance criteria and its dynamic behaviours. It also considers both from theoretical definition and practical implications. The analysis uses basis consideration of activities within the SRM. It follows Cross Industry Standard Process for Data Mining (CRISP-DM) framework. One data clustering tool is applied to determine supplier classification and further decision. In the end, this study finds that by using relevant data sources, data analytic successfully shows a more comprehensive analysis on SRM activities in manufacturing company.

Keywords
Clustering, CRISP-DM, Data Analytics, Manufacturing Company, and Supplier Relationship Management.

1. Research Background

Currently, modern industry is experiencing a rapid and dynamic competition. It drives manufacturing companies to provide high quality products with affordable price and shorter lead time. As mentioned in Verma (2019), it is almost desired of every company to deliver customer delight, i.e. efficiency and effectiveness and hence what customer wants over and above the products become an important aspect. It also related to how they manage the system components, as well as the stakeholders along the supply chain. The chain consists of supplier, distributor or wholesaler, and its customers. Based on the discussion with some production managers, it is known that supplier contribution takes 70% portion of the whole process. They stated that the quality of the products mainly derived by the raw materials supplied by their supply partners. Thus, the relationship between the company and its suppliers becomes very critical and should be handled properly.

Large-scaled manufacturing company mostly carries more than fifties or hundreds of suppliers out, which are placed within the country and abroad. A higher number of partners involved also increase the difficulties on managing manufacturer-supplier relations that is including supplier selection and performance evaluation. There are several steps on partner selection process, such as preparation, sample creation and complete assessment. Once passed all these steps, the candidate supplier allowed to do mass production according to the order quantity. Moreover, the suppliers will be periodically managed through set of training, inspection, and evaluation. In terms of supplier performance, the accomplishment mostly being evaluated in every month with the target of 99.5% (IQC 2017). It follows the criteria of evaluation such as incoming rejection, in-process failure, special used parts, lot received and demerit of the non-responded claim. The attainment may be varying amongst suppliers, and the corrective action will be developed by then. Some companies gathered many data of their supplier, but probably they still need longer time

to use it optimally. As consequences, companies spent longer time to process it. Since, the business competition is very challenging, a data analytics approach should be applied.

Gudivada (2017) stated that by using data analytic, the company can draw inferences based on data owned, then determine a prediction to enable innovation and help strategic decision making. The data analytic task that will be done is the research is clustering. Clustering is the process of data mining that divide the data into classes which the class form themselves from the patterns and characteristics of the datasets (Santosa and Umam 2017). Data clustering is an appropriate tool to group the suppliers based on their characteristics. The output of data clustering will be analysed to support the Supplier Relationship Management (SRM) in manufacturing company.

SRM is the process of engaging in activities of setting up, developing, stabilising and dissolving relationships with in-suppliers as well as the observation of out-suppliers to create and enhance value within relationships (Moeller *et al*. 2006) . In the presence of data analytics, it is expected that manufacturing company will consider the characteristics of their suppliers. One more important thing is how they develop the supplier performance precisely. It is aligned to Sillanpää (2015) research, which conclude that supplier development can help the company to increase the supplier capabilities gradually for further improvement.

There are several researches on supplier clustering using data mining techniques. Sofiani (2017) conducted a study to cluster suppliers with their dynamic behaviours in a continuous manufacturing company. Sofiani (2017) determined three clusters of supply partners with internet-based visualization. Haghighi *et al*. (2014) used Fuzzy C-Means clustering to compare each supplier's value criteria with exactly same criterion of other suppliers. This study found an enhanced quality of results compared to another research's outputs. In Rezaei and Lajimi (2018) study, the researcher combined Purchasing Portfolio Matrix (PPM) and Supplier Potential Matrix (SPM) on improving the supplier management activities. Moreover Sillanpää (2015) and Chayhan (2012) discussed about supplier development. Their main objectives are to explore the supplier development concept in empirical case studies.

According to the previous researches, it is known that supplier performance builds company competitive advantage through improving supplier factors within the development programs. Thus, this study aims to implement data analytics on managing SRM activities. It considers manufacturer performance criteria which has well defined in many previous researches and will be adjusted with the manufacturing company, especially discrete product manufacturer, needs. It also accommodates dynamic behaviour of the supply partners. This analysis uses Cross Industry Standard Process for Data Mining (CRISP-DM) model and applies one powerful data clustering, named K-Means clustering. Regarding to the output of this study, the company can use it to determine further strategic decision that more comprehensive. The output is expected to fill the gap in theoretical side and provide practical implications to manufacturing company with similar characteristics and SRM requirements.

2. Literature Review

2.1 Supplier Relationship Management

In supply chain, a company deals with many parties, such as supply partners and distributors. Suppliers refer to company partners that provide parts or product components. In the current days, people's awareness began to change regarding to the efficiency and value creation. Because this is not only about what is running on the production lines, but also in other company functions including in the procurement. Therefore, suppliers will be well considered by developing some innovation and added value creation in the future. In addition, about 40 to 70 percent of the total cost comes from the purchasing materials, in which the suppliers involved (Pujawan and Er 2010). Once suppliers can work efficiently, it is expected to increase company profits.

Since suppliers is an important part of all companies, a proportional relationship towards the suppliers is highly demanded. Every supplier has different characteristics. Some of them probably supply a few items with value of hundreds of thousands money per year, and the others provide hundreds or even thousands of items in billions money of the transaction value. This is an imbalance, and it is better to handle the relationship properly.

Moeller *et al*. (2006) defined that the SRM is set of activities on managing the setting up, developing, stabilising and dissolving relationships process with in-suppliers as well as the observation of out-suppliers. Its objective is to create and strengthen values within relationship. The research that had been conducted by Park *et al*. (2010) developed an integrated SRM framework, thus the supply chain management in any company will perform successfully. The

suppliers should be selected based on their purchasing strategy, and their performance will be periodically evaluated. Park *et al*. (2010) mentioned the result of this assessment will be the input of the development stage, and the continuous improvement will be deployed by then.

2.2 Data Analytics

Data analytics is the science of integrating heterogeneous data from diverse sources, drawing inferences, and making predictions to enable innovations, gain competitive business advantage, and help a strategic decision-making (Gudivada 2017). In data analytics, there is a data mining. Data mining is defined as a specific process of finding interesting knowledge, such as patterns, relationships, changes, peculiarities, and certain structures of big data stored in databases, data warehouse or other information storages. In Santosa and Umam (2018), data mining consists of four major tasks, i.e. clustering, classification, regression or estimation and association. Data mining may be used in many cases in our daily life. According to Santosa and Umam (2018), data mining techniques can be very useful on solving problems, with the following characteristics:

1. Requires knowledge-based decisions.
2. Has dynamic environment.
3. The current method is sub-optimal.
4. There is data that can be accessed, sufficient and relevant.
5. Provide high benefits if the decision taken is right.

2.3 K-Means Clustering

K-Means clustering is a tool is data mining. A group of data is classified into k clusters. In order to group the data, the number of k should be determined based on the information of the observed object. In detail, the size of dissimilarity of that can be used to group the data into clusters. It will be translated in the concept of distance, the close the distance, the similar the data, and vice versa. Santosa and Umam (2018) explained that set of data is divided into clusters by minimising its Euclidean distance between the data and the nearest cluster centre point, with the following steps:

1) Determine the number of clusters (k).
2) Determine random centre point of cluster. After that calculate the next i-cluster centroid by using this formula:

$$v = \frac{\sum_{i=1}^{n} x_i}{n} ; i = 1,2,3, \dots, n \quad (1)$$

3) Calculate the distance of the data to the centroid. One of the famous distance measurements is Euclidean distance. Below is the formula of Euclidean distance:

$$d(x,y) = \|x - y\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} ; i = 1,2,3, \dots, n \quad (2)$$

With x_i = variable on object x to-i and y_i
y_i= variable output y
n = number of objects

4) Allocate each object into nearest centroid.
5) Allocation of objects into each cluster at iteration with k-means. Where each cluster member object has been measured the proximity distance to the cluster's centre point.
6) Perform iteration until the centroid position is not change. Due to the use of MATLAB software, these formulas are included in the K-Means algorithm.

3. Methods

This research steps are classified into three important categories, i.e. supplier performance criteria, supplier cluster determination, and determining conclusion. The first step is identification the supply-partner performance criteria. It consists of three sub-steps, such as defining performance dimensions based on the literature; aligning with the most manufacturer needs; identifying indicators for each criterion; then determining supplier data which can be used as the attributes of data clustering process. The output of this steps is supplier datasets that will be proceeded in the following steps. Next, in the second step, there are five stages that follows the detail phase of CRISP-DM framework, i.e. business understanding; data understanding; data preparation; modelling; and final evaluation. Business understanding phase is aligning the data mining problem to the needs of the manufacturing company. It is followed by data

understanding, which is gathered from the previous step. In data preparation, there are two activities such as data cleaning and data transformation. One of the most important stage is modelling. This study applies K-Means clustering method, that is involving several processes such as determine number of clusters, determine clusters centroid, calculate distance of data to the centroid, allocate data to the nearest cluster, recalculate the centroid with the current cluster number, and record the final cluster member and centroid. K-Means clustering output will be evaluated. Lastly, the result from second step supports on determining conclusion. As mentioned in the previous part, this study considers both theoretical side and practical needs. Thus, the steps, terminology used, and outputs will be validated by production manager of a manufacturing company.

4. Research Findings and Discussion

Based on Dey *et al*. (2015), supplier performance evaluation is needed for organisational system for measuring supply-partner performance effectively. It is one of the mitigation strategies, especially in procurement. At the first step, this research defines supplier performance criteria. Dey *et al*. (2015) stated that supplier performance measurement is divided into two critical parts, which are performance, also capability and practices. Performance criteria consists of three indicators such as quality, delivery, and costing (Dey *et al*. 2015). While capability and practices can be used conditionally. Refers to Maestrini *et al*. (2018), the supplier performance criteria along the supply chain is considering the way the buyer company communicated the performance information to the supplier and the way suppliers react to the performance reported. It was deployed into many indicators by (Maestrini *et al*. 2018). Thus, there are five criteria that generally used based on literature.

A senior production manager in large-scaled manufacturing company with discrete products was involved on discussing the most appropriate criteria to assess its suppliers. It is assumed this company represents manufacturing company in common. The manager explained that most of the well-established companies run a monthly evaluation to their partners; however, some of them could defined specific criteria due to their needs. Point to the brainstorming, the most critical factors on assessing is quality that it is aligned to the finding in (Dey *et al*. 2018). The quality characteristics of the partners has high impacts to the whole production process, as well as the finished products. In conclusion, quality will be very fundamental and should be considered.

As stated in Montgomery (2012), there are eight quality dimensions; but the discussion found that two most important criteria is performance and responsiveness. Performance criteria defined as how the supply material perform the intended job. The company evaluate its supplier performance by using four indicators such as the lot inspected, incoming rejection (IR), in-process rejection (IPR), and specific used parts (SUP). In terms of the lot inspected, it records how much the supply partners can supply parts in the unit of lot. The part is commonly inspected based on the appearance and dimension before it is moved to the production line. Once the quality inspector finds major defects, it will be recorded as the IR. While there is a minor defect and the material still be sent to the shop floor, it will be noticed as the SUP. The second criterion is responsiveness of the supplier. It describes whether the requests and complains are promptly handled or not by company's strategic partner. Every company probably has their own time limits, which once the supplier exceeds the deadline, it will be recorded as demerit. To sum up, in this research there are five indicators deployed, and it is represented in Figure 1.

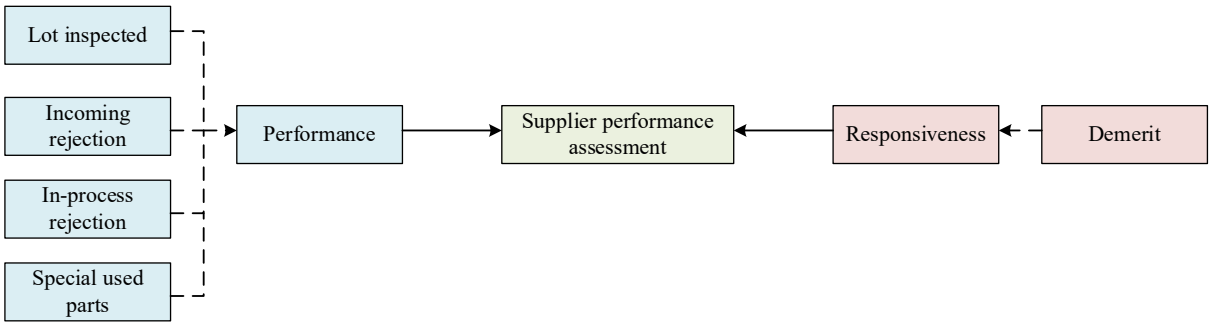

Figure 1. Performance criteria and indicators.

The data that is related to the suppliers are collected. The data used are from early 2019 to 2020. The data will be called as supplier datasets, as the input of the second step. In determining supplier cluster, this study follows five stages in CRISP-DM framework. To transform the company project, a focus discussion is conducted. Based on the

discussion, it is known that the data analytics objectives, especially clustering supplier, is aligned to the business strategy on SRM. It is followed by the sub-step of gathering supplier datasets. Data preparation consists of two steps, i.e. data cleaning and data transformation. Since there are incorrect data due to the human errors and transmission errors, the dataset is being cleaned. Data cleaning include filling the blank data, smoothing the missing data, eliminating outliers and inconsistency. Furthermore, the data transformation is done to prevent unclear results due to the large data range. To transform the data, this study used min-max normalization Thus, the data will be transformed to have the range of 0 to 1 value. Before the data is transformed, the maximum and minimum value respectively are 3440 and 0.

In the following stage, modelling involving process of selecting and applying technique to calibrate parameters to the optimal value. This study applied K-Means clustering. The datasets are transformed into coding and was run in MATLAB software. The input of coding consists of the number of clusters, datasets to be clustered and maximum iteration number. This analysis tried number of clusters (k) from 2 to 10, and the maximum iteration used varies from 5, 50 and 100. The modelling output is shown in Table 1 and Table 2.

Table 1. Silhouette index for iteration 5, 50 and 100.

k	Silhouette Index		
	Iteration 5	**Iteration 50**	**Iteration 100**
2	0.7868	0.7927	0.7927
3	0.8568	0.8634	0.8634
4	0.853	0.8724	0.872
5	0.6394	0.7898	0.6841
6	0.6561	0.787	0.7943
7	0.698	0.7723	0.7723
8	0.6114	0.7418	0.7418
9	0.6093	0.6159	0.7133
10	0.6243	0.669	0.6688

Table 2. SSE for various iterations.

k	Sum of Squared Error		
	Iteration 5	**Iteration 50**	**Iteration 100**
2	67.41149	67.24501	67.245008
3	41.32942	41.22319	41.22319
4	36.74817	34.02597	34.593508
5	34.18664	35.42836	34.658553
6	28.5255	26.57357	26.494523
7	27.89192	23.6306	23.630602
8	24.11921	22.42034	29.820285
9	24.1044	22.98696	21.01427
10	27.02867	19.9232	19.754952

The analysis consists of silhouette index and sum of squared error (SSE). Silhouette index is a parameter to know whether the data placement in the cluster is appropriate or not. As the silhouette index is close to 1, the data placement is appropriate. In contrast, the further value from 1 means the worse placement. Based on the programming output in Table 1 and as illustrated in Figure 2, the result shows when the number of cluster (k) is 3, the silhouette index value is 0.8568 for 5 iterations, 0.8634 for 50 iterations, and 0.8634 for 100 iterations. It has a slight difference compared to the result of k = 4, which the value of silhouette index is the highest. In short, if there are no other critical indicators to be examined, the optimal cluster number is 4.

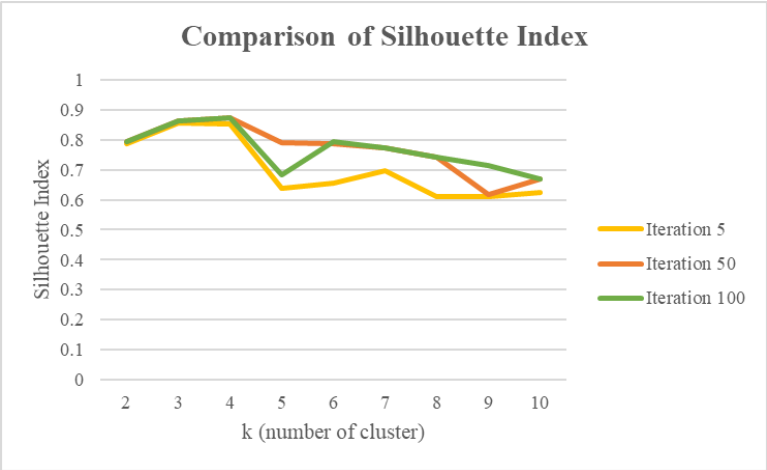

Figure 2. Comparison the silhouette index among iterations.

The Sum of Squared Error (SSE) represents the difference between data within the cluster and the mean of the cluster. Based on the reference, the smaller the SSE value means the better clustering result. According to elbow method, the optimal number of cluster (k) is the one that has significant descent when the number of clusters is changed. Refers to the output in Table 2, it is known that the optimal k equals to 3 for this case study. As illustrated in the Figure 3, the data modelling output shows there is a significant slope when the k is changed both from 2 to 3 and 3 to 4. To sum up, the number of indicators and iterations may affect the silhouette index value and sum of squared error.

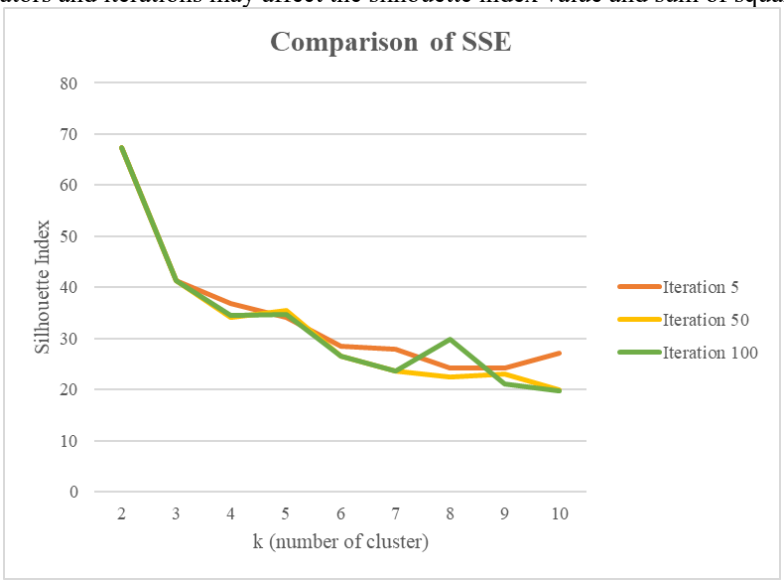

Figure 3. Comparison the SSE.

After completing the stages in data modelling, the output provides a clear information about supplier classification by considering similar characteristics. The company can use this result for determining further decision on SRM. The supplier clusters should help company to define both short and long-term activities, such as monitoring future performance, providing incentives or recognition, increasing multiple investment, offering training or assessment, and enhancing competition among supply partners. Those are commonly known as supplier development program.

5. Conclusion and Future Works

According to the explained result and analysis, this study successfully applied data analytics on managing supplier relationship. By using the datasets in a manufacturing company, the modelling output provides an analysis that can be

used in strategic decision making. The advantage of this research is considering both conceptual and practical needs. However, there is still a wide opportunity to improve this research. In terms of theoretical side, future research is suggested to accommodates other performance supplier criteria, thus the analysis becomes more comprehensive. The other thing that can be improved in the future work is using various number of iterations to have a detail analysis and sensitivity in clustering the data.

References

Chavhan, R., Supplier development: Theories and practices, *IOSR Journal of Mechanical and Civil Engineering*, vol. 3, no. 3, pp. 37-51, 2012.

Dey, P. K., Bhattacharya, A., Ho, W., and Clegg, B., Strategic supplier performance evaluation: A case-based action research of a UK manufacturing organisation, *International Journal of Production Economics,* vol. 166, pp. 192-214, 2015.

Gudivada, V. N., *Data Analytics for Intelligent Transportation System,* Chapter 2, pp. 31-67, Elsevier Science Publishers B. V., Amsterdam, 2017.

Haghighi, P. S., Moradi, M., and Salahi, M., Supplier segmentation using fuzzy linguistic preference relations and fuzzy clustering, *International Journal of Intelligent Systems and Applications,* vol. 6, pp. 76-82, 2014.

IQC DO, *Supplier Performance 2019-2020*, Jakarta, 2020.

Maestrini, V., Maccarrone, P., Caniato, F., and Luzzini, D., Supplier performance measurement systems: Communication and reaction modes, *Industrial Marketing Management*, vol. 74, pp. 298-308, 2018.

Moeller, S., Fassnacht, M., and Klose, S. A., Framework for Supplier Relationship Management (SRM), *Journal of Business-to-Business Marketing*, vol. 13, no. 4, pp. 69-94, 2006.

Montgomery, D. C., *Introduction to Statistical Quality Control*, 7th ed.,Wiley, New Jersey, 2012.

Park, J., Shin, K., Chang, T., and Park J., An integrative framework for supplier relationship management, *Industrial Management and Data Systems*, vol. 110, no. 4, pp. 495-515, 2010.

Pujawan, I. N. and Er M., *Supply Chain Management*, 2nd ed., Guna Widya, Surabaya, 2010.

Rezaei, J. and Lajimi, H. F., Segmenting supplies and suppliers: Bringing together the purchasing portfolio matrix and the supplier potential matrix, *International Journal of Logistics Research and Applications*, vol. 22, no. 4, pp. 419-436, 2018.

Santosa, B. and Umam, A., *Data Mining and Big Data Analytics: Theory and Implementation, Using Phython & Apache Spark,* 2nd ed., Penebar Media Pustaka, Yogyakarta, 2018.

Shofiani, N., *Supplier Segmentation using K-Means Clustering, A Case Study in PTPN X PG Meritjan*, *Institut Teknologi*
Sepuluh Nopember, Surabaya, 2017.

Sillanpää, I., Shahzad, K., and Sillanpää, E., Supplier development and buyer-supplier relationship strategies - A literature review, *International Journal of Procurement Management*, vol. 8, no. 1/2, pp. 227-250, 2015.

Verma, A., *Supplier Matters*, Notion Press, Chennai, 2019.