Classification of Printing Ink Quality based on Pigment Dispersion Process: A Case Study

Muhammad Sopyan Yahya and Isti Surjandari
Department of Industrial Engineering
Faculty of Engineering, Universitas Indonesia
Kampus UI, Depok 16424, Indonesia
muhammad.sopyan@ui.ac.id
isti@ie.ui.ac.id

Abstract

Pigment dispersion process is the most crucial process in printing ink manufacturing. A multinational printing ink manufacturing company produces ink through semi-finished or intermediate products prior to finished ink production. Pigment concentrated is an intermediate with the most complex manufacturing process in the company because it involves a pigment dispersion process. During the 2019-2021 period, the Right First-Time rate of pigment concentrated was below than 70%. Along with technological developments, the use of Data Mining methods in the industrial world has been widely studied. The objective of this study is to classify printing ink quality based on the pigment dispersion process by bead mill with CART using historical production parameter and quality data from a multinational printing ink manufacturer. This research was conducted using the QM-CRISP-DM framework by Schafer et al. (2018). The data was obtained from the production logbook and the company's ERP database from 2019 to 2021 period. The model was made using 18 independent variables and particle fineness status as target variable with (0) as defected and (1) as not defected. This study shows that the fineness of printing ink is highly affected by temperature of Pass 1 during the production. The decision rule to reduce defect in pigment concentrated production process is also proposed. The use of the over-sampling method on training data and the Entropy Criteria has the highest accuracy.

Keywords
CART, Pigment Dispersion Process, Printing Ink, QM-CRISP-DM

1. Introduction

Food, beverages, household goods, electronic products from daily essentials are generally sold in a package. The packaging is made of various types of materials, such as plastic film, paper, and aluminum foil. The various types of designs consist of various colors on these packages. Singh (2006) stated that the color on the packaging is one of the important aspects that have an influence on the perception of customers when buying a product. Inconsistent colors between packages of the same brand can damage the brand's image and have an impact on product sales. The printing ink industry for packaging actually plays an important role in realizing design ideas on packaging with the desired material.

In general, color on the packaging comes from the coloring agent in the printing ink. Coloring agents that are widely used for printing ink applications on packaging are pigments. The pigments mixed in the printing ink form a colloidal system known as dispersion. The level of pigment dispersion in a printing ink depends on the particle size distribution of the pigment in the printing ink. Pigments, especially organic pigments, need to go through a grinding process to reduce their particle size so that they can be well dispersed in the polymer resin ink system and produce strong, clean color shade, with high gloss levels (Mäntynen et al., 2012). In the printing ink manufacturing process, pigment agglomerates are reduced through collisions between beads in a bead mill so that the particle size is reduced to aggregates. The grinding process is carried out continuously until finally the pigment particles reach a size of less than 5 μm. The particle size of the pigment is crushed through a cycle of collisions of the bead in a bead mill.

A multinational printing ink company produces ink through an intermediate phase. One of the intermediates produced is pigment concentrated. Pigment concentrated are semi-finished inks that only consist of resin and single pigment. It serves as a color source in the ready-to-use printing inks produced by this company. During the 2019-2021 period, the percentage value of the right-first-time (RFT) for pigment concentrated production never reached 80% (Figure 1). RFT is an important aspects parameter as it represents the effectivity of production in the supply chain directory. Failed to fulfill pigment concentrate quality at the first production would require additional processes to achieve the quality that this company has defined. The types of defects found were coarse particles, color strength, color transparency, color shade, viscosity, and total solid. The process of producing pigment concentrated is very complex and involves many factors. Improving the quality of pigment concentrated products is a very complex problem because it involves many factors. A new approach is needed to find out the root cause and solution for the appointed issue.

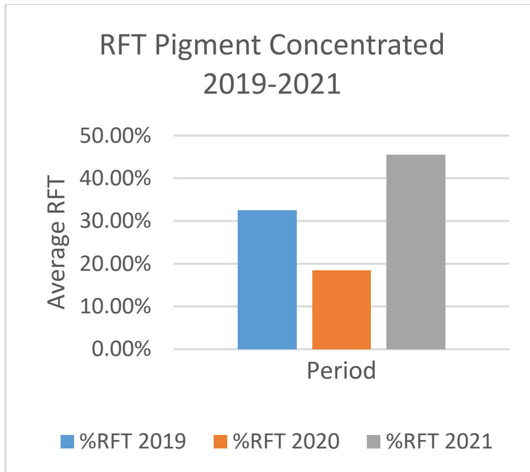

Figure 1. Right-First-Time rate of pigment concentrated product

As technology develops, the manufacturing industry has a lot of data available for use. The data is available in various sources such as environmental sensors and tool parameters (Davis et al., 2015). The use of mathematics and statistics to study large data sets is called Data Mining (Larose & Larose, 2014). Data Mining can be applied to various types of fields, one of which is manufacturing. In the last decade, the use of Machine Learning is considered more effective and efficient than statistical modelling and simple mathematics (Dogan & Birant, 2021). Quality improvement initiatives such as Six Sigma, Design for Six Sigma (DFSS), and kaizen have the disadvantages that these methods require data collection and analysis to solve quality problems (Köksal et al., 2011). Schafer et al. (2018) published their research regarding combination of Quality Management with CRISP-DM, one of well-known Data Mining methodology.

According to Abdulmalek et al. (2006), printing ink industry can be categorized as process industry as it has similar process with paint industry. Supervised Machine Learning can be applied for classification activities or quality prediction in the process industry (Ge et al., 2017). Decision Tree is a supervised machine learning algorithm that can work well on quality classification tasks in the process industry (Yeh et al., 2011). He et al. (2009) reported that CART is suitable for classification task for quality improvement initiatives.

The pigment dispersion process plays an important role in the quality of the printing ink. Many factors that affect the pigment dispersion process make it difficult to analyze the problem. In addition, the amount of data on manufacturing can be used to increase the effectiveness and efficiency of the manufacturing process so that the company's business remains competitive. Therefore, this experiment aims to analyses the causes of defects and propose parameter to decrease defect rate by constructing a model to classify quality based on the QM-CRISP-DM using the CART as Decision Tree algorithm.

1.1 Objectives

The purpose of this research is to create a model to classify printing ink quality based on the pigment dispersion process by bead mill with CART using historical production parameter and quality data from a multinational printing ink manufacturer.

2. Literature Review

2.1 Pigment Dispersion Process

There are two materials that act as color sources in printing ink, dyes and pigments. Pigment is a color source that is widely used in printing inks because it has high resistance, heat stability, solvent resistance, light resistance, and minimal migration (Gürses et al., 2016). Particle size plays an important role in the quality of the color produced by the pigment in an ink system. A good pigment dispersion process will produce an ink that is transparent, good light resistance, can be stored for a long time, gloss, and has good color strength (Leach & Pierce, 1988).

On an industrial scale, there are several types of machines to carry out the pigment dispersion process, such as three roll mill, high speed dissolver, and bead mill (Ogi et al., 2017). Figure 2 shows the stages of pigment dispersion using a bead mill which is mostly operated by ink manufacturing companies. There are 3 main stages, namely pre-mixing, milling, and let-down. Pre-mixing is done by mixing part of the ink formulation and stirring with a high-speed mixer. At this stage, the pigment particles are broken down to less than 250 μm and a wetting process is carried out between the pigment and the medium (Lin, 2003). The ready ink mixture is then further processed using a bead mill. Here the process of particle breaking and further wetting occurs with the help of friction and collision of the bead in the bead mill chamber until the particle size is less than 5 μm. Generally, there are two types of bead mill operation in printing ink manufacturing, circulation and single-pass (Miranda, 2011). Circulation mode is preferred due to shorter residence time and easy to be maintained, yet some pigments are unable to be processed by this mode. On the other hand, single-pass mode is simpler than circulation where there is no product limitation but it takes longer production time and difficult to be maintained.

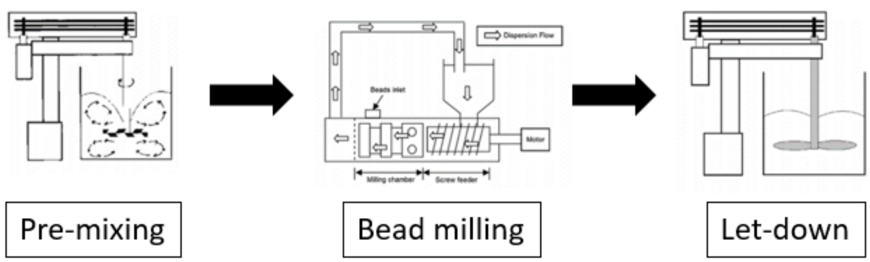

Figure 2. Pigment dispersion process in industrial scale

(Adapted from Ali & Lin (2018))

2.2 QM-CRISP-DM

Schafer et al. (2018) published his research on the combination of Quality Management with CRISP-DM, shortened to QM-CRISP-DM. This method is a fusion of DMAIC in Six Sigma with the CRISP-DM concept to facilitate industrial practitioners who are less familiar with the Data Mining method. There are 5 stages used to make the Data Mining project a success using the QM-CRISP-DM approach (Figure 2). These stages are Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment and Control.

2.3 CART

Decision Tree is a classification method consisting of a collection of decision nodes, connected by branches, extending to the root node until it breaks at the leaf node (Larose & Larose, 2014). Decision Tree is an algorithm that is widely used for classification activities. When compared with other algorithms such as Neural Network, Support Vector Machine, k-Nearest Neighbor, or Naïve Bayes, Decision Tree does not require complicated assumptions and easy to interpret. There are several algorithms in building the Decision Tree model, namely ID3, C4.5, C5.0, and CART. According to Singh & Gupta (2014), the decision tree model built with the CART algorithm is less affected by outliers compared to other algorithms. CART has a characteristic that it can only form two branches in decision making (Breiman et al., 2017). CART grows a tree by searching for all variables and all possible splitting values and selecting the optimal split based on predetermined criteria for each decision node (Figure 3-6).

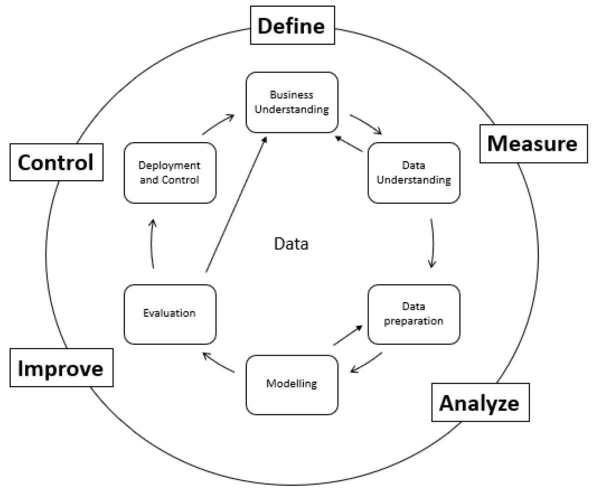

Figure 3. QM-CRISP-DM framework

(Adapted from Schafer et al. (2018))

2.4 Company Profile

The company is a multinational printing ink manufacturer based in Indonesia. The company supplied printing ink for several application such as label and flexible packaging for food, hygiene, and pharmaceutical. Their customers are printing companies located in Southeast Asia and several other countries in West Asia and East Asia. This company produces printing ink through the manufacture of intermediate or semi-finished products which are then converted into ready-to-use printing inks. The printing ink production scheme at the company can be seen in Figure 3. Before making ready-to-use printing inks, three semi-finished products are made, Intermediate A, B, and C. Intermediate C is known as concentrated pigment because it contains a high concentration of pigment. The production process of Intermediate C is the most complex because it involves a pigment dispersion process in it. This strategy is carried out by the company to produce quality printing ink and have an effective lead time.

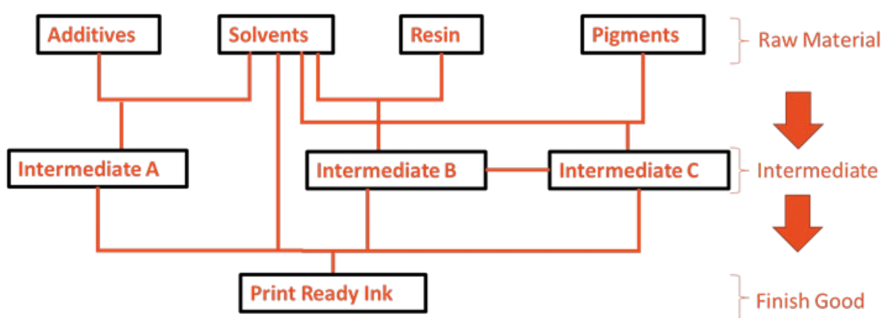

Figure 4. Production strategy in the company

3. Methods

This research identified the most commercial pigment concentrated as the priority to be improved. After that, the quality data and parameter of machine were collected. The data was explored for better understanding. Then, the independent variable and the dependent variable were determined to construct the dataset. The model was built using Jupyter Notebook with Python 3.9. Imbalanced dataset was balanced using SMOTE. The dataset was called and splitted into data training and data testing with ratio 70:30, respectively. The CART model was trained using data training and tested using data testing. Consequently, Split-validation and Cross-validation was carried out to validate model. The Decision Tree model was visualized and analyzed for further improvement proposal.

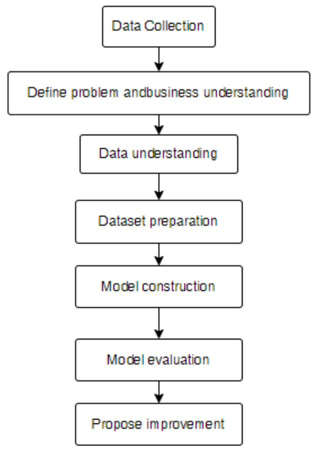

Figure 5. Research methodology

4. Data Collection

The data were collected from the quality data management on ERP and parameter logbook machine. The dataset was constructed based on these data by Microsoft Excel.

5. Case Study

This research propose model to classify quality issue during pigment concentrated production and reduce defect rate in the company that used as case study. This section shows the result of this research and implementation of the proposed methodology in this work.

5.1 Define - Business Understanding

Pigment concentrated is an intermediate product that provide color for print-ready-ink. The production of pigment concentrated is time and energy consuming due to pigment dispersion process. Pigment is dispersed in polymeric binder through mechanical shear stress and bead collision. A well-dispersed pigment concentrated provides printing ink with good color strength, clear color shade, good light fastness, and stable after storing.

Several defects can be observed during pigment concentrated production, such as particle coarse, color strength, color shade, opacity, viscosity, and total solid content. Figure 6 shows cause and effect diagram related to defect in pigment concentrated production. A defect will greatly affect the overall productivity and lead to additional process. Therefore, it decreases productivity and increases the conversion cost at the end. Analysis for improvement is often facing difficulties since many factors are involved during pigment concentrated production. The proposed model will be used to find parameter of pigment dispersion process in bead mill with lowest defect rate. The outcome of the research is expected to provide solution for a better production of pigment concentrated.

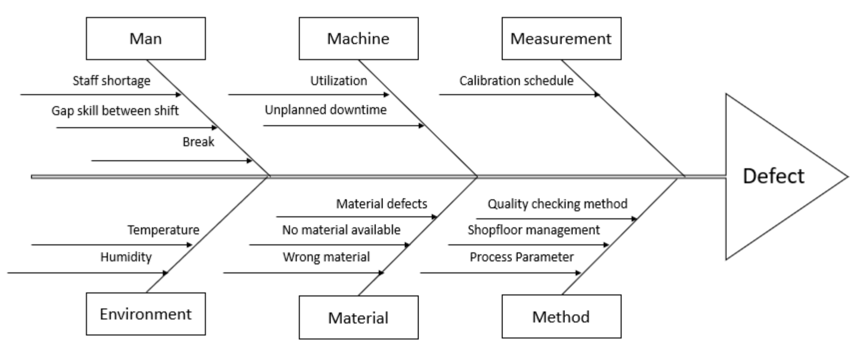

Figure 6. Cause and effect diagram defect on pigment concentrated

5.2 Measure – Data Understanding

PB (Pigment Blue) 15:4 is considered as the top three high runner pigment concentrated produced by the company. Meaning, the product is generally produced in a high quantity with a high potential of profit. PB15:4 is produced through horizontal bead mill which operated with single-pass mode. It is the only product in the top three runner that produced through single-pass mode. The research took 87 historical production data of PB15:4 from 2019 to 2021. The data were including raw material incoming quality check result, process parameter, and pigment concentrated Blue 15:4 quality record. Figure 7 is a Pareto diagram which shows that particle coarse is the most occurred defect found in this product.

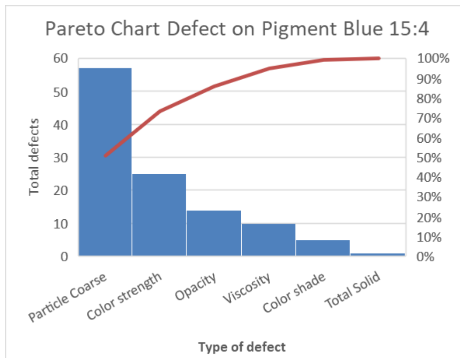

Figure 7. Pareto chart of defect on pigment concentrated Blue 15:4

5.3 Measure – Data Preparation

The dataset was prepared based on information obtained from previous steps. Table 1 shows target variable and independent variable used in this research. Particle fineness status quality or Fineness_Status which represents whether product has particle coarse or not is used as target variable. The independent variables were consisted by quality incoming data of varnish and machine parameter used.

Table 1. Independent and target variable

Description	Variable	Unit
Team_Operator	X1	-
Batch_qty	X2	kg
Temperature_Pass1	X3	°C
Agitator_Speed_Pass1	X4	rpm
Pump_Speed_Pass1	X5	rpm
Pressure_Pass1	X6	bar
Flow_Rate_Pass1	X7	kg/h
Rotor_Power_Pass1	X8	kW
Temperature_Pass2	X9	°C
Agitator_Speed_Pass2	X10	rpm
Pump_Speed_Pass2	X11	rpm
Pressure_Pass2	X12	bar
Flow_Rate_Pass2	X13	kg/h
Rotor_Power_Pass2	X14	kW
Varnish_Source	X15	-
MeToSolidContentTest_Varnish	X16	%
MeTOWaterContentTest_Varnish	X17	%
ViscosityDin6_Varnish	X18	second
Fineness_Status	Y	0/1

5.4 Analyze – Modelling

Decision Tree with CART was chosen as algorithm to classify process parameter that led to lowest defect rate during production of pigment concentrated. GridSearchCV was used to determine the best parameter with highest accuracy. SMOTE was carried out to balance the dataset and improve model accuracy. The dataset was splitted into 70% as data training and 30% as data testing. Furthermore, the model was validated using cross-validation (table 2).

Table 2. Parameter used in this research

Parameter	**Setting**
Data period	2019-2021
Algorithm	Decision Tree (CART)
Balancing method	*Over-sampling* SMOTE
Split-validation	70:30
Cross-validation fold	5
Criterion	Entropy
Splitter	Best
Maximum depth	7
Minimum sample split	2
Minimum sample leaf	2
Class weight	None
CCP alpha	0.0
Target variable	Fineness_Status
Column output	1
Fineness_Status	Defect (0) / Not defect (1)

5.4 Improve – Evaluation

Through this model, the accuracy of model to classify which process lead to defected or not defected is ca. 84%. Table 3 and Table 4 show accuracy of split-validation and cross-validation method, respectively. Based on the split-validation result, there was insignificant gap between data training and data testing. This result shows the model worked well and not over-fitted. On the other hand, the accuracy through cross-validation is quite stable which showed by low standard deviation of accuracy among each iteration.

Table 3. Split validation result

Accuracy	
Data training	**Data testing**
96%	84.84%

Table 4. Cross-validation result

Accuracy						
1st iteration	**2nd iteration**	**3rd iteration**	**4th iteration**	**5th iteration**	**Average**	**Std Dev.**
77.27%	86.36%	81.81%	85.71%	71.43%	81%	0.06

The advantage of Decision Tree algorithm is it can be easily visualized and interpretable. Figure 8 shows the decision tree based on the constructed model. Based on constructed Decision Tree, temperature during Pass 1 is the most important feature because it stands as root node.

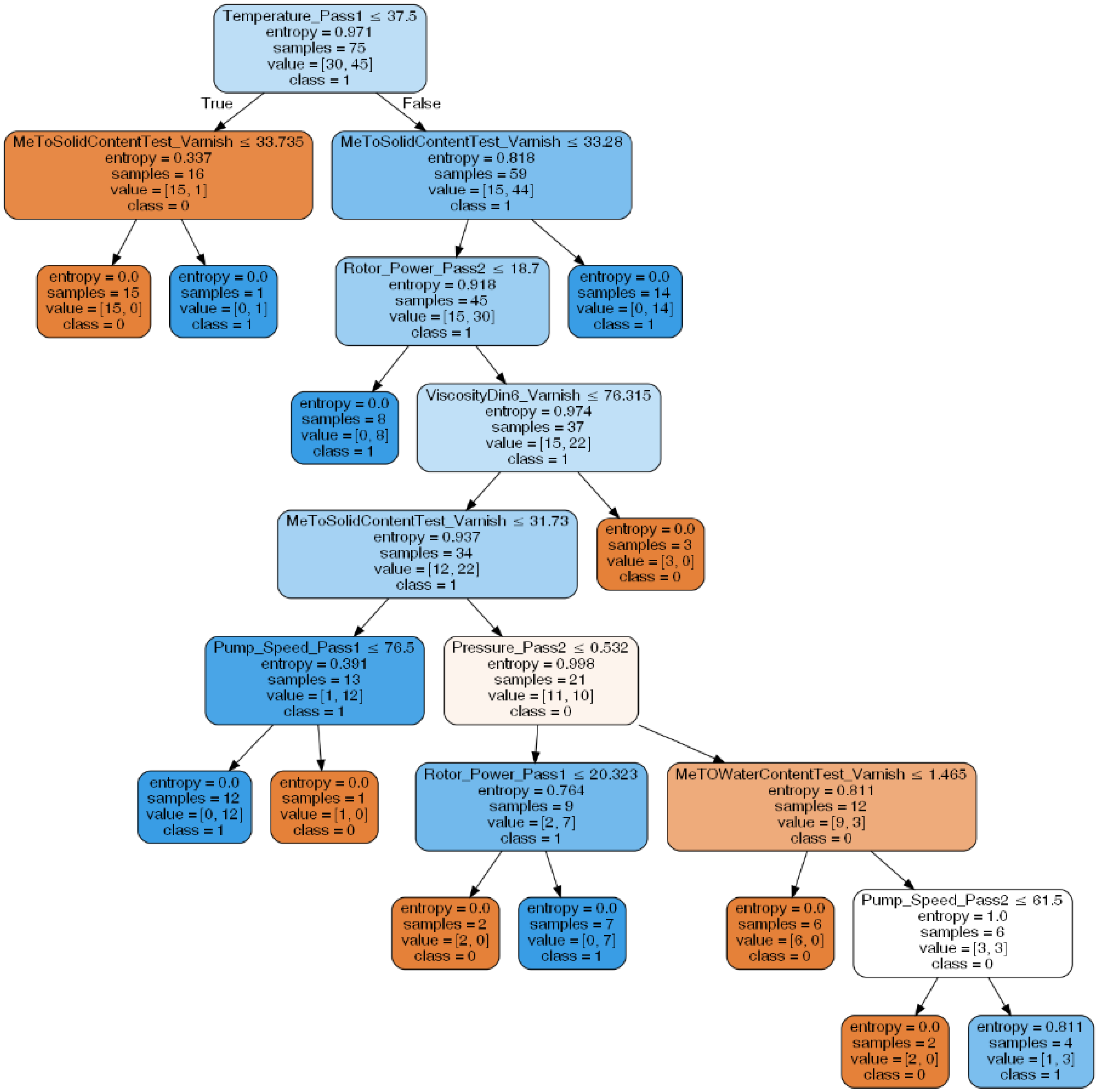

Figure 8. Decision Tree

5.5 Control-Deployment

Normally, pigment dispersion process caused heat generation. Temperature plays an important role to improve dispersion process. The company has set the standard temperature during pigment dispersion process around 40-50°C. However, several data show that some products failed to achieve the standard temperature during Pass 1. In order to achieve the purpose of Data Mining project, the Decision Rule based on constructed Decision Tree was shown on Figure 9. According to the constructed model, it was found that whenever the temperature during Pass 1 was less than or equal to 37.5°C, the particle coarse defect mostly occurred. From material quality wise, varnish with higher solid content yet lower viscosity would lead to lower defect rate. The company has been informed regarding the improvement solution.

```
|--- Temperature_Pass1 <= 37.50
|   |--- MeToSolidContentTest_Varnish <= 33.73
|   |   |--- class: 0
|   |--- MeToSolidContentTest_Varnish >  33.73
|   |   |--- class: 1
|--- Temperature_Pass1 >  37.50
|   |--- MeToSolidContentTest_Varnish <= 33.28
|   |   |--- Rotor_Power_Pass2 <= 18.70
|   |   |   |--- class: 1
|   |   |--- Rotor_Power_Pass2 >  18.70
|   |   |   |--- MeToSolidContentTest_Varnish <= 32.78
|   |   |   |   |--- Temperature_Pass2 <= 46.50
|   |   |   |   |   |--- Temperature_Pass1 <= 39.50
|   |   |   |   |   |   |--- class: 0
|   |   |   |   |   |--- Temperature_Pass1 >  39.50
|   |   |   |   |   |   |--- Agitator_Speed_Pass2 <= 892.00
|   |   |   |   |   |   |   |--- class: 0
|   |   |   |   |   |   |--- Agitator_Speed_Pass2 >  892.00
|   |   |   |   |   |   |   |--- class: 1
|   |   |   |   |--- Temperature_Pass2 >  46.50
|   |   |   |   |   |--- Pump_Speed_Pass1 <= 76.50
|   |   |   |   |   |   |--- class: 1
|   |   |   |   |   |--- Pump_Speed_Pass1 >  76.50
|   |   |   |   |   |   |--- Temperature_Pass1 <= 45.00
|   |   |   |   |   |   |   |--- class: 0
|   |   |   |   |   |   |--- Temperature_Pass1 >  45.00
|   |   |   |   |   |   |   |--- class: 1
|   |   |   |--- MeToSolidContentTest_Varnish >  32.78
|   |   |   |   |--- class: 0
|   |--- MeToSolidContentTest_Varnish >  33.28
|   |   |--- class: 1
```

Figure 9. Decision Rule

6. Conclusion

The pigment concentrated quality is crucial for printing ink quality because it provides color for print-ready ink. However, the pigment concentrated production process is often challenging because of many factors involved. The objective of this research is to construct model and analyzed process parameter with lowest defect rate with Decision Tree. CART was used as decision tree algorithm. SMOTE was used to balance the data and the best parameter was determined by means of GridSearchCV in order to obtain high accuracy. The model with sophisticating accuracy was constructed parameter as follow, Entropy as splitting criterion, minimum sample per leaf is 2, and maximum tree depth is 7. Temperature during Pass 1 became the most important feature in this model which stand as root node in decision tree visualization. From material quality aspect, varnish with higher solid content and lower viscosity is more desired to achieve lower defect rate. QM-CRISP-DM was helpful to solve printing ink industrial problem using Data Mining approach. For future work, the methodology in this research can be proposed to improve products that was produced using vertical bead mill and widen for the other pigment concentrated with different color index. Each pigment has its own characteristic and this model could be not applicable for the other pigments. Therefore, new model development is necessary to improve the other product quality

References

Abdulmalek, F. A., Rajgopal, J., & Needy, K. L. S, A classification scheme for the process industry to guide the implementation of lean. *EMJ - Engineering Management Journal*, *18*(2), 15–25.2006.

Ali, M., & Lin, L. , Optimisation and analysis of bead milling process for preparation of highly viscous, binder-free dispersions of carbon black pigment. Progress in Organic Coatings, 119(April 2017), 1–7.2018

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. , *Classification And Regression Trees*. Routledge.2017

Davis, J., Edgar, T., Graybill, R., Korambath, P., Schott, B., Swink, D., Wang, J., & Wetzel, J. , Smart Manufacturing. *Annual Review of Chemical and Biomolecular Engineering*, *6*(1), 141–160.2015

Dogan, A., & Birant, D. , Machine learning and data mining in manufacturing. Expert Systems with Applications, 166(February 2019), 114060.2021

Ge, Z., Song, Z., Ding, S. X., & Huang, B. , Data Mining and Analytics in the Process Industry: The Role of Machine Learning. *IEEE Access*, *5*, 20590–20616.2017

Gürses, A., Açıkyıldız, M., Güneş, K., & Gürses, M. S. , *Dyes and Pigments: Their Structure and Properties*. 13–29.2016

He, S., He, Z., Alan, G., & Li, L, Quality Improvement using Data Mining in Manufacturing Processes. *Data Mining and Knowledge Discovery in Real Life Applications*, *February*.2009

Köksal, G., Batmaz, I., & Testik, M. C, A review of data mining applications for quality improvement in manufacturing industry. *Expert Systems with Applications*, *38*(10), 13448–13467., 2011

Larose, D. T., & Larose, C. D. , Discovering Knowledge in Data: An Introduction to Data Mining Second Edition Wiley Series on Methods and Applications in Data Mining. In *IEEE Computer Society*. John Wiley & Sons.,2014

Leach, R. H., & Pierce, R. J. , The nature of printing ink. In R. H. Leach, C. Armstrong, J. F. Brown, M. J. Mackenzie, L. Randall, & H. G. Smith (Eds.), *The Printing Ink Manual* (pp. 1–13). Springer Netherlands.1988.

Lin, L. (2003). Mechanisms of pigment dispersion. Pigment and Resin Technology, 32(2), 78–88.2003.

Mäntynen, A., Zakharov, A., Jämsä-Jounela, S.-L., & Graeffe, M. , Optimization of grinding parameters in the production of colorant paste. *Powder Technology*, *217*, 216–222.2012

Miranda, S, Using an Agitator Bead Mill for Nanoparticle Dispersion and Comminution. *Nanotechnology*, *November*, 234–237.2011

Ogi, T., Zulhijah, R., Iwaki, T., & Okuyama, K, Recent progress in nanoparticle dispersion using bead mill. *KONA Powder and Particle Journal*, *2017*(34), 3–23.2017

Singh, S. (2006). Impact of color on marketing. *Management Decision*, 44(6), 783–789.2006

Singh, S, & Gupta, P. , Comparative Study Id3, Cart and C4.5 Decision Tree Algorithm: a Survey. *International Journal of Advanced Information Science and Technology (IJAIST) ISSN*, *27*(27), 97–102.2014

Schafer, F., Zeiselmair, C., Becker, J., & Otten, H. , Synthesizing CRISP-DM and Quality Management: A Data Mining Approach for Production Processes. *2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)*, 190–195.2018

Yeh, D. Y., Cheng, C. H., & Hsiao, S. C, Classification knowledge discovery in mold tooling test using decision tree algorithm. *Journal of Intelligent Manufacturing*, *22*(4), 585–595.2011

Biographies

Muhammad Sopyan Yahya started his study as Master Student at Department of Industrial Engineering with major Statistic and Quality Engineering, Faculty of Engineering, Universitas Indonesia on 2020 until now. Previously, he obtained bachelor’s degree from Department of Chemistry, Faculty of Mathematics and Science, Universitas Padjadjaran on 2018. He served as full-time employee in a leading printing ink manufacturer as R&D specialist on colorant development for paper and board application.

Isti Surjandari is a Professor and Head of Statistics and Quality Engineering Laboratory in the Department of Industrial Engineering, Faculty of Engineering, Universitas Indonesia. She holds a bachelor’s degree in industrial engineering from Universitas Indonesia and a Ph.D. degree from the Ohio State University. Her areas of interest are industrial management, quality engineering, data mining and applied statistics. She is a senior member of American Society for Quality (ASQ) and also ASQ country counselor for Indonesia. She has a vast experience in manufacturing and service industries.