Mazandaran University of Science and Technology

Available online at www.iiec2017.com

13th International Conference on Industrial Engineering

(IIEC 2017)

Iran Institute of Industrial Engineering

Defects rule mining for continuous cold rolling mill production lines of Mobarakeh steel company

Nasrin Rezaei Abadchi[a], Javid Jouzdani[b] and Mehdi Karbasian[c]

[a] Department of Industrial Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran
Tel: +983142291111, Fax: +983142291016, E-mail: nasrin.rezaei66@yahoo.com

[b] Department of Industrial Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran
Tel: +983142291111, Fax: +983142291016, E-mail: javidjouzdani@pin.iaun.ac.ir

[c] Department of Industrial Engineering, Malek-e-Ashtar University of Technology, Esfahan, Iran
Tel: +983122945141, Fax: +983122935341, E-mail: mkarbasian@yahoo.com

Abstract

In today's competitive world, a bewildering amount of data is being generated. By utilizing appropriate techniques, the knowledge hidden within the data can be uncovered and used as a tool for quality improvements in a business. Data mining presents a collection of efficient tools for knowledge elicitation in data warehouses. Mobarakeh Steel Company, as one of major industries in Iran, has a significant impact on market and economy both in terms of extent and scale. From this point of view, it has advantages over similar companies in the region. One of the most important production lines in the company is the galvanized steel coils production line with a capacity of 20,000 tons per year. In this research, data mining techniques are applied on a database of galvanized steel coils defects in order to extract the governing rules. Having the 10 chemical analysis parameters of the products during 2012 to 2015, the associated rules are obtained. The data used in this study are obtained from the database provided from the quality control department in the Mobarakeh Steel Company. The rules are shown to be effective in modelling the knowledge within the data. In addition, it is shown that not all the defects in the products are rooted in the metallurgic properties recorded in the database. This calls for a deeper investigation for further study the effects of other factors such as human factors and equipment.

Keywords:
Continuous production line, Cold roll mill, Mobarakeh Steel Company, Associate rule mining

Introduction

The diverse preferences of customers and business competitive environment along with political and cultural changes have revolutionized production systems in today's world. The only survivors of this challenge are those companies that successfully adapt to the changes and maintain their flexibility. Knowing the fact that the core of any industrial revolution is the production system, organizations try to adapt with these changes.

Nowadays, data constitute the core of decision making in businesses and are used in solving different problems in production engineering, design, quality assurance, scheduling and sequencing, etc. Therefore, the need for a tool for storing and analyzing these data is of great importance [1]. Specifically, the data can be utilized to identify the hidden patterns governing the production processes or used in quality improvement. In many coil production lines, quality control is performed manually by experts [2]. Due to the high movement speed of the line, collecting and analyzing the data in cannot be easily completed in a small amount of time. Since the diagnosis and prediction has a significant impact on planning and management of organizations, they should be as precise and reliable as possible and based on scientific concepts and methods.

One of the major industries in many countries is steel making and its related businesses; however, quality of such products may greatly affect the market share of such industries. Therefore, quality control has been a point of attraction for decision makers in steel making plants. Identification of steel coil surface defects constitute a major part of quality control activities in cold roll mill plants. In such facilities, it is important to know the reasons behind the defects in order to take preventive measures to ensure that the quality of the products meet the needs of customers. In this paper, the rules governing the defects in the galvanized coils are obtained by applying rule mining techniques on several records provided by Mobarakeh Steel Company. This leads to proper identification of defects and their root causes which in turn results in possible corrective and preventive actions. Our main goal in the study is presenting a method to diagnose the factors that influence the surface defects in product produced in cold roll mill facility based on the data from one of the major production lines in Mobarakeh Steel Company.

Literature and Background

In steel making industry and especially in alloy steel, defects impose significant costs. One of commonly observed and important defects in low-carbon steel coils is pit/blister for which the coil should be re-worked or scrapped. The re-work is costly and time consuming. Automatic control of steel surface is one of the main steps in production of high quality steel coils in many steel making companies. Haghani [3] predicted the tearing defect in a tandem mill production line by using neural network. The data for their research were collected from the tandem mill production line of Mobarakeh Steel Company and were fed into a perceptron network with 3 hidden layers. Their results indicated that the network can be effective in predicting the tearing defect in the tandem mill production line in the company. In another research, Yazdchi, et al. [4] presented a neural network approach for detection and classification of the surface defects of cold rolling mill steel.

In a research by Sharafi and Esmaeli [5], a method for prediction of aforementioned faults using data mining methods is proposed. In their method, decision trees, neural networks and associated rules were utilized, compared according to efficiency and the best method was introduced. They used entropy method for the network with 24 nodes in the hidden layer and obtained the association rules for a set of four items in Iran Alloy Steel Company. In another research by Ghasemi Vinche, et al. [6] the application of data mining in predicting the mechanical properties of galvanized coils in Mobarakeh Steel Company is addressed. They used and compared artificial neural networks, regression and decision trees and concluded that neural network demonstrate a superior performance. They found out that it is possible to predict the mechanical properties of those products based on metallurgic parameters and control parameters including temperature and speed of rolling. Yazdanian [7] proposed a method for resolving the shape flaws caused by different length increases in both cold and hot coils. These flaws in turn lead to tension in the coil. In their research, they consider the amount of bending as a tool for controlling the flaws in coils in order to improve the quality of products. They presented a neural network for modelling the phenomenon in order to predict the flaws having the bending values for different stages of the rolling process. Martínez-de-Pisón, et al. [8] utilized the association rule mining in order to extract the knowledge about the industrial failures in a hot-dip galvanizing steel line.

Surface flaws in galvanized coils may be caused by several reasons as they pass through different production stages. Minor mistakes in maintenance of production machines and equipment result in several flaws and defects. Inappropriate cleaning of the surfaces causes the underlying alloy to be poorly coated by zinc leading to uncoated points in the product. Inappropriate environment or lack of temperature control may create carbon sediments or acidic films in the underlying alloy which lead to flaws in coating. The chemical properties of the bath, submerging conditions, methods of bath slag collection and air jet configurations affect the quality of coating. Obviously, effective and efficient management of these parameters reduce the flaws; however, not all defects can be eliminated this way. The metallurgic properties of the underlying alloy and its surface quality are also of great importance in creating a high quality coating. Interstitial free (IF) and similar grades of steel are easily galvanized; however, newly developed grades with complicated chemistry may have weak galvanization properties. This is mainly because these grades of steel have elements such as silicon and manganese which create active surfaces [9].

In the previous studies aimed at classification and identification of the surface defects of cold roll mill steel, a specific defect is investigated while to the best of authors' knowledge, this is the first study to consider the wide extent of factors that may cause the defects in a real world production facility.

Methodology

In this study, association rules are utilized to classify the galvanized coil surface defects in Mobarakeh Steel Company cold rolling mills. Data cleaning is an important step in data mining. In this regard, experts' judgments are considered along with other statistical tools in order to identify and elicit the major parameters for classification of coil surface defects. By using the appropriate collection of data the association rules are obtained. These rules express the relationships among parameters and the flaws.

In classification, each individual observation is specifically belongs to a class. In other words, if there is a structure in the data, classification can reveal that. Association rules can be obtained through classification algorithms. It is known that association rules usually perform better than other conventional methods such as decision trees.

Three well-known methods for association rule mining are: Classification based on Multiple Association Rules (CMAR), Classification based on Predictive Association Rules (CPAR), Classification Based on Associations (CBA). Generally, association rule mining is consisted of two main processes: first is the study of the database and search for frequent attribute-value pairs and second is the analysis of the frequent sets in order to obtain the association rules. These rules should have acceptable confidence and support values in order to be used for classification purposes. The association rules have the following general form in which P_i is a pair of attribute-value represented as (A_i, v) where A_i is an attribute that has the value v.

$$P_i \wedge P_2 \wedge \ldots \wedge P_l \Rightarrow A_{class} = C \quad (1)$$

In the above equation, the pairs are combined using "AND" operator. In this research, the attributes in the antecedent of the rule are the metallurgic and chemical properties of the products and the class in the consequence part of the rule is the defect. Therefore, in this research, we are mining the rules that relate the metallurgic and chemical properties of the product to the defects that may be observed. The obtained rules can be then utilized to classify the defects based on their root causes.

In this research, the Cross Industry Standard Process for Data Mining (CRISP-DM) approach is followed. In this approach, certain steps are followed to ensure the quality of the results. These steps are as follows [10].

- Business Understanding: understanding the objectives of the project and the business requirements and then defining the data mining problem and a plan to achieve the objectives.
- Data Understanding: starts with data collection and continues with activities to know the data and to detect interesting subsets of the data in order to investigate hypotheses for hidden information.
- Data Preparation: is performed in order to prepare the data to be fed into the model and includes attribute selection as well as transformation and cleaning of data.
- Modeling: there are several modeling techniques that can be selected and applied. Usually, more than one technique exist for the same data mining problem.
- Evaluation: At this stage a model (or models), which appears appropriate from a data analysis perspective, is built. Before proceeding to deployment step, it is important to thoroughly examine and evaluate the model. In this step, it should be investigated to determine whether some important business issues are neglected.
- Deployment: Model creation is not the end of the project even if the model is built for knowledge extraction purposes. The knowledge gained should be organized and presented in a useful format. The deployment process may be as simple as a report or as complex as an expert system that incorporates the knowledge obtained from the data mining process. In many cases the customer and not the data analyst carries out the deployment. Even if the analyst deploys the model, it is crucial for the customer to understand the activities needed for actually making use of the created models.

The sequence of the steps is not rigid and moving back and forth between the steps is often required. The process of data mining continues after a solution is found and deployed and the lessons learned can stimulate new more business questions and in turn further data mining processes will benefit from the previous ones.

Results

Mobarakeh Steel Company, as one of the largest industrial units in Iran, is located 75 kilometers southwest of Isfahan, with a capacity of 5.33 million tons of different steel products including slabs, hot and cold coils, galvanized and colored. 48% of the coils produced in hot roll mill are transported to cold roll mill in order to be transformed to several other products.

Following Cross Industry Standard Process for Data Mining (CRISP-DM) methodology, the data gathered from different sources in the company are prepared into a database covering the information regarding the products and production process for the time period between 2012 and 2015.

Table 1. Parameters and their ranges

Parameter		Allowed Range
Tank 252 Temperature		63-80
Heating Temperature	CQ	710-750
	DQ	760-800
Cooling Temperature		460-550
Dew Point		Less than 5
Hydrogen		4%-20%
Oxygen		Less than 500
Molten Temperature		450-470
Alignation in Skim Pass		0%-2%
Force in Skin		60-550
Alignation in Leveler		0%-2%
Aluminum		0.17%-0.24%
Iron		0%-0.05%
Lead		0%-0.045%
Chromium in Make Tank	Cr^{6+}	2.5%-3.5%
	Cr^{3+}	0.8%-1%
Chromium in Mix Tank	Cr^{6+}	2.5%-3.5%
	Cr^{3+}	0.8%-1%
Chromium in Work Tank	Cr^{6+}	5%-7%
	Cr^{3+}	1%-1.3%
CP in Degreasing Tank No. 1		50-80
CP in Degreasing Tank No. 2		50-80
PH in Rinsing Tank No. 1		9-11
CP in Rinsing Tank No. 2		7-9
Chromium on the Surface	Cr^{6+}	Less than 30
	Cr^{3+}	40-70
COD	Tank 1	Less than 18000
	Tank 2	Less than 18000

These data include the data regarding the surface flaws each of which is determined by a specific 3-digit code. In addition, each record of any defect is given an integer number indicating its severity and ranging from 0 to 2. The data also contain the alloy metallurgic properties and production parameters such as temperature. These parameters and their ranges are presented in Table 1. Out-of-range parameters generally result in flaws in the product.

In order to obtain the association rules, the data recorded daily through 3 work shifts during 4 years are analyzed. In order to find the outlying data, the quartiles are used. More specifically, the data that lie out of the following interval are considered as outliers and are omitted.

$$\left[Q_3+1.5\times\left(Q_3-Q_1\right),Q_1-1.5\times\left(Q_3-Q_1\right)\right] \qquad (2)$$

Table 2. Available significant parameters

Defect Code	Significant Parameters	R^2
901	FE-AL-PB-TT-TM	0.863
902	N/A	0
903	Tt-Tc-Tm-H_2-O_2-Cq	0.900
904	Tt-Tc-Tm-H_2-O_2-Cq	0.763
905	Tt-Cq-Tc-H_2-Tm-AL-Fe-Pb-Cp	0.771
906	N/A	0
907	N/A	0
908	CQ-FE-O_2-H_2-TT-Tm	0.853
909	N/A	0
910	N/A	0
911	N/A	0
912	Pb	0.768
913	Pb-O_2-Cq-Tt	0.821
914	N/A	0
915	N/A	0
916	N/A	0
917	N/A	0
918	N/A	0
919	N/A	0
920	N/A	0
921	N/A	0
922	Pb-O_2-Cq-Tt-H_2	0.869
923	N/A	0
924	N/A	0
925	Pb-O_2-Cq-Tt-H_2	0.870
926	N/A	0
927	N/A	0
928	N/A	0
929	N/A	0
930	N/A	0
931	Tt-Cq-Tc-H_2-O_2-Tm	0.776
932	N/A	0
933	N/A	0
934	N/A	0
935	N/A	0
936	Pb-O_2-Cq-Tt	0.899
937	Cq-Tt-O_2-Tm-AL-Fe-Pb	0.888
938	N/A	0
939	Cq-Tt-O_2-Tm-AL-Fe-Pb	0.771
940	N/A	0
941	N/A	0
942	N/A	0

By analyzing the data, it was revealed that the data was available for only 10 parameters that play the major role in creating the flaws. These parameters are shown in Table 2 and are utilized to obtain the association rules that determine the relationship between the parameters and the different types of defects on the surface of the products. The notations used in this Table are introduced in Table 3.
In Table 4, these relationships are presented through the association rules obtained by using RapidMiner® software. In this Table, the leftmost column presents the rule and the other columns provide the confidence and the support values. In fact, these rules express the dependence of each defect to parameters.

Table 3. Parameter notations

Parameter	Notation
Tank 252 Temperature	Tt
Heating Temperature	Cq
Cooling Temperature	Tc
Hydrogen	H2
Oxygen	O2
Molten Temperature	Tm
Aluminum	Al
Iron	Fe
Lead	Pb
CP in Degreasing Tank No. 2	Cp
Chrome	Cr

Table 4. The association rules

Rule	Support	Confidence
(Tt< 63 and Tt >80) and (Tm< 450 and Tm> 470) and (Pb>045/0) and (Al< 17 and Al >24) and (Fe >05/0) → error(901)	25	50
(Tt<63 and Tt>80) and (Tc < 460 and tc <550) and (Tm<450 and Tm> 470) and (H_2 <4 and H_2 >20) and (O_2 >500) → error(903)	15	60
(Cq<710 and Cq>750) and (Tt< 63 and Tt >80) and (Tc < 460 and tc<550) and (H_2 <4 and H_2 >20) and (O_2>500) and (Tm< 450 and Tm>470) → error(904)	30	60
(Cq< 710 and Cq >750) and (Tt< 63 and Tt >80) and (Tc < 460 and tc<550) and (H_2 <4 and H_2 >20) and (Tm<450 and Tm> 470) and (Al< 17 and Al>24) and (Fe >0.05 and Pb >0.045) and (Cp <50 and Cp>80) → error(905)	40	70
(Cq< 710 and Cq >750) and (Tt< 63 and Tt >80) and (H_2 <4 and H_2 >20) and (O_2>500) and	25	60

(Tm< 450 and Tm>470) and (Fe >0.05) → error(908)		
(Pb >0.045) → error(912)	30	70
(Cq<710 and Cq >750) and (Tt< 63 and Tt >80) and (O_2 >500) and (Pb >045/0) → error(913)	40	80
(Cq< 710 and Cq >750) and (Tt< 63 and Tt >80) and (O_2 >500) and (Pb>0.045) and (H_2 <4 and H_2 >20) → error(922)	50	80
(Cq< 710 and Cq >750) and (H_2 <4 and H_2 >20) and (O_2 >500) → error(925)	30	50
(Cq< 710 and Cq >750) and (Tt< 63 and Tt >80) and (Tc<550) and (H_2<4 and H_2 >20) and (O_2>500) and (Tm< 450 and Tm>470) and (Cp <50 and Cp >80) → error(931)	30	60
(Cq< 710 and Cq >750) and (H_2 <4 and H_2 >20) and (O_2 >500) and (Pb >0.045) and (Cr3+<1 and Cr3+>8) and (Cr+6<30) → error(936)	25	60
(Cq< 710 and Cq >750) and (Tc < 460 and tc <550) and (H_2 <4 and H_2 >20) and (O_2>500) and (Tm< 450 and Tm>470) and (Al< 17 and Al >24) and (Fe >0.05) and (Pb >0.045) → error(937)	40	70

As an example, the first rule expresses that if the temperature in Tank 252 is lower than 63 or higher than 80 and the heating degree below 450 or larger than 470 and lead percentage is more than 0.045 and aluminum percentage is below 17 or higher than 24 and Iron percentage is more than 0.05, then defects with code 901 may occur. The likelihood of occurrence of the flaws are determined by the other two columns. Specifically, the support of a rule shows the percentage of the data records that have the conditions in the "if part" of the rule. The other column, i.e., confidence shows the percentage of data records that for which the rule is found to be true. For example, for the first rule, support is 25 showing that 25% of the data match the "if part" of the rule and confidence is 50 indicating that 50% of the data are in agreement with the rule.

Conclusions and Future Works

The main objective of this research was to classify the possible products defects in the cold roll mill of Mobarakeh Steel Company using association rule mining. The outcomes of this research can shed light on the mechanisms of the emergence of the 13 identified defects and their root causes through the obtained rules. The results also indicate that for 29 defects out of 42 identified defects in the galvanized coils production line, no parameters are identified. This is an alert for decision makers and calls for extensive studies to identify other parameters such as human factors and equipment errors that result in the aforementioned flaws.

Predicting the defects using other methods in data mining such as neural networks in combination with other methods is an interesting subject of future research that may further provide insight to the problem. In addition, similar applications of associate rule mining in other production lines can greatly facilitate the decision making for quality managers.

References

[1] M. Jelali, "Performance assessment of control systems in rolling mills–application to strip thickness and flatness control," *Journal of Process Control*, vol. 17, pp. 805-816, 2007.

[2] P. K. Rai, A. Mohammad, and H. Z. Jafri, "Causes & Preventation of Defects (Burr) In Sheet Metal Component," *International Journal of Engineering Research and Applications (IJERA) ISSN: 2248*, vol. 9622, pp. 511-515, 2013.

[3] A. Haghani, "Predicting Strip Tearing in Cold Rolling Tandem Mill using Neural Network," *International Journal of Advanced Design and Manufacturing Technology*, vol. 8, pp. 67-75, 2015.

[4] M. R. Yazdchi, A. G. Mahyari, and A. Nazeri, "Detection and Classification of Surface Defects of Cold Rolling Mill Steel Using Morphology and Neural Network," presented at the Proceedings of the 2008 International Conference on Computational Intelligence for Modelling Control & Automation, 2008.

[5] S. M. Sharafi and H. Esmaeli, "Application of data

mining in prediction of steel surface flaws," presented at the Fourth Cinference on Data Mining, 1389.

[6] Z. Ghasemi Vinche, A. Zeynal Hamedani, S. Alimoradi, and A. Nazeri, "Application of data mining in prediction of mechanical properties of the products in galvanized line in Mobarakeh Steel Compoany," presented at the 8th International Conference on Industrial Engineering, 1389.

[7] M. Yazdanian, "Investigating the wave flaws in steel coils and its resolution," presented at the 3rd National Conference on Production and Manufacturing Engineering, 1390.

[8] F. J. Martínez-de-Pisón, A. Sanz, E. Martínez-de-Pisón, E. Jiménez, and D. Conti, "Mining association rules from time series to explain failures in a hot-dip galvanizing steel line," *Computers & Industrial Engineering,* vol. 63, pp. 22-36, 8// 2012.

[9] P. Rahrig, "Hot Dip Galvanizing," *Modern Steel Construction,* vol. 35, 1995.

[10] G. Harper and S. D. Pickett, "Methods for mining HTS data," *Drug Discovery Today,* vol. 11, pp. 694-699, 8// 2006.