

Prediction of defects using machine learning techniques in order to improve quality management system – A case study

Adina Sârb^{1,*}, *Cristina Burja Udrea*¹, *Daniela Nagy – Onița*², *Liliana Itul*³, and *Maria Popa*⁴

¹Doctoral School, Faculty of Engineering, Lucian Blaga University of Sibiu, Romania

² Faculty of Mathematics and Computer Science, University of Bucharest, Romania

³Quality Department, S.C. Apulum S.A., Alba Iulia, Romania

⁴Business Administration and Marketing Department, Faculty of Economic Science, 1 Decembrie 1918 University of Alba Iulia, Romania

Abstract. According to ISO 9000, a quality management system is part of a set of related or interacting elements of an organization that sets policies and objectives, as well as the processes necessary to achieve the quality objectives. Quality is the extent to which a set of intrinsic characteristics of an object meets the requirements. Based on these definitions, the factory, considered in this paper, S.C. APULUM S.A., decided to implement a quality management system since 1998. Subsequently, the organization's attention is focus on the continuous improvement of the implemented quality management system. The purpose of this paper is to study the percent of specified defects specific to ceramic products in the future to improve the quality management system. In this regard, machine learning techniques were applied for defects forecasting for different types of products: mugs, pressed plates and jiggered plates. The experimental evaluation was performed on real data sets that contain percentages about different types of defects collected in 2018-2019. The experimental results show that for each type of product exists an algorithm that forecasts the future defects.

1 Introduction

The globalization of the world economy and the expansion of international trade have led to rapid processes of quality internationalization, which have become a crucial part of enterprises' competitiveness [1]. One of the best ways to deal with global competition is to produce a product and service of the highest quality [2]. In this context, quality management systems (QMS) are effective tools for companies to increase their competitiveness. This aspect has been studied a long time ago since Dr. E. Deming and Dr. J. Juran have started their studies and practical implementation of the quality management system and quality thinking in companies [3].

*Corresponding author: sarb.adina92@yahoo.com

In addition, the adoption of a QMS by organization represents a decision that can help improve their overall performance and provide a solid foundation for sustainable development initiatives. Also, risk-based thinking enables organizations to determine the factors which can cause deviations in their processes, implement preventive controls to maximize the opportunities that emerge and minimize negative effects. The implementation of a QMS in the functional structure of the organizations is supported by specific initiatives and models such as the Common Assessment Framework (CAF), European Foundation for Quality Management (EFQM) and International Organization for Standardization (ISO) standards, which represent important methodological support [4].

ISO 9001 is an international standard of QMS, which guarantees that the organization will provide products or services that will meet the requirements of customers and relevant stakeholders. The first ISO standard was published in 1987 by the International Organization for Standardization based in Geneva, Switzerland [5]. The last version of this standard was introduced in 2015 to improve the flexibility to respond to the internal and external changes and the complexity of business dynamics. Complying organizations can provide conformity products and services that satisfy their customers' needs and expectations while addressing relevant statutory and regulatory requirements [6]. ISO 9001:2015 is applicable to any organization, regardless of geographic location and size. To obtain an ISO 9001 certificate, it is necessary to accredit before a certifying entity that the processes are being performed according to the ISO standard for which the certification is sought. This process consists of the following phases: documentation, evaluation and rating. If the outcome of the last phase, rating, is appropriate, certification is issued. If the outcome is not appropriate, the certifying entity will list the errors detected [7]. Also, all over the world, more than one million organizations in 170 countries have an ISO 9001 certificate of compliance and the average annual growth is around 15% [8].

2 Literature review

The quality management system (QMS), a name on which there is still no unanimous agreement at the international level, using terms such as Total Quality Control (TQC), Total Quality Management (TQM), Company-Wide Quality Control (CWQC), in general, all with similar meanings, brings together in the largest proportion the global experience and concepts of top quality practitioners and organizations that have been successful in the field of quality, gaining a wide expansion lately (after 1990) [9].

QMS is designed and implemented to ensure quality in processes, products and services, while simultaneously achieving customer satisfaction. There are many definitions of QMS. For instance, L.J. Porter and A. J. Parker refers to it as a way to manage an organization to improve its general efficiency, whereas J. Juran and A. Godfrey define a QMS as an administrative system used to directly control quality matters to meet organizational objectives and goals. L. Bravi, F. Murmura and G. Santos view a QMS and its subsequent certification as a voluntary process adhered to the philosophy of quality and supported by organizational motivation, goals and policies [10], while R.B. Carey defines QMS as an integrated set of activities used to establish and control work processes, manage resources, conduct evaluations and continually improve [11].

The implementation of a QMS in an organization represents an important investment [12]. It is achieved by involving the entire organization, respectively each department existing in the company and each worker, at each level, in achieving quality requirements, related to employee and customer satisfaction [13]. The successful implementation of such a system by ISO 9001:2015 provides the company various benefits, such as alignment with business strategy, increased top management commitment, improved responsibility from the process leaders and involved employees, improved awareness regarding risk

management and its minimization, opportunities identification, cost reductions, increased product/service quality, improved processes performance, improve the use of time and resources, improve the work environment, increase customers' satisfaction and confidence, improve understanding of customers' needs and expectations [14,15].

3 Research methodology

Weka (Waikato Environment for Knowledge Analysis) is a free data mining tool. It is developed in the Java programming language, by the University of Waikato. It contains learning algorithms. They can be applied to data sets for data mining problems. Figure 1 presents the Weka interface, which includes four other graphical interfaces: Explorer, Experimenter, Knowledge Flow and Simple CLI [16].



Fig. 1. Weka interface

For the presented research, the Explorer interface was used. This interface includes the main data mining operations that can be performed with this software: preprocessing, classification, segmentation, association, attribute selection and visualization. Since the module applied for experiments is the forecast (Forecast), which is not a default operation in Weka, it was necessary to install this plugin.

The Forecast module allows basic configuration and advanced configuration. Figure 2 presents the Basic configuration that displays the numeric type attributes from the added data set and makes it possible to select those for which you want to apply the prognosis and control the parameters. With the help of the Advanced configuration, the learning algorithm can be chosen first and its main features can be modified. For each prediction model applied to the data set, the results can be viewed both numerically and graphically.

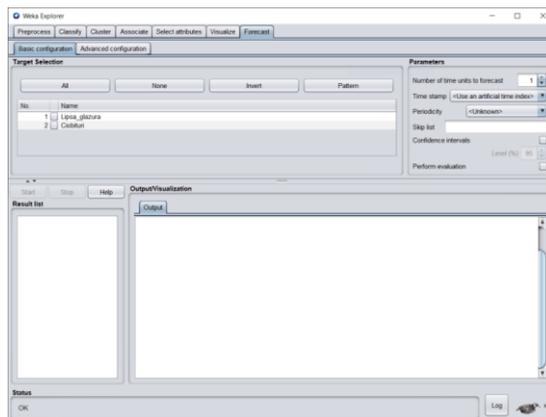


Fig.2. Forecast module – Weka interface

4 Case study

4.1 Study area

The company considered in this paper, S.C. APULUM S.A., implemented, certified and maintained a QMS since 1998, by quality standard named ISO 9001 :1994 [17, 18], thus contributing to the market position occupied by the organization – the largest porcelain manufacturer in Romania and South- East Europe [19]. Also, this QMS is reflected in the organization of the entire activity of the company, respectively: supply, production, quality control, marketing of products and provision of services.

4.2 Experimental analysis

In general, the forecast is used to establish decisions and strategies, to make assumptions based on the initial data. To obtain statistics about the QMS in the porcelain industry, machine learning techniques were applied. Weka tool was used for making predictions about the percent of defects in ceramic industry for three types of products: cups, pressed plates and jiggered plates. The scope of this research is to study the percent of a specified defect in the future in order to improve the QMS.

4.2.1 Dataset

We used a data set with text data collected from SC APULUM SA. The data set contains information about three products: mugs, pressed plates and jiggered plates. Therefore, for each type of product, in the period 2018-2019, quality surveys were conducted, thus identifying the quality defects that occurred during the considered time. Various defects specific to ceramic products have been identified in quality surveys, such as iron, cracks, deformation, bits, specks, pinholes, missing glaze, chipped. The data obtained were expressed as a percentage. From the identified defects, the first two defects that appeared most frequently for each product type were selected. Thus, the following defects were considered: for mugs – iron and bits, for pressed plates – missing glaze and chipped, for jiggered plates – iron and chipped. Therefore, machine learning techniques were applied to the selected defects to obtain predictions regarding the evolution of the defects considered for the next 12 months to improve the QMS implemented within the organization.

4.2.2 Experimental protocol

The first step for making predictions using Weka is to convert the data set from the CSV extension to the ARFF (Attribute-Relation FileFormat) extension. Such a file must contain on the first line `@relation <relation-name>`, where `<relation-name>` is of type string. The following lines are declared as follows: `@attribute <attribute-name><datatype>`. With `@attribute`, will define the attributes and the data type, in the order of the column position. `<attribute-name>` indicates the name of the attribute and must start with a letter and `<datatype>` represents the data type and can be of four types: numeric `<specified-name>` (list of possible values is entered), string or data `<data -format>`. The attributes are followed by the actual data declaration section. This section starts with a line that contains `@data` and then follows the data set values. The delimitation of the attribute values is done by commas (,). If there are missing values, they are replaced by the question mark (?) and the string values that contain space will be written in quotation marks [20].

Evaluating the quality and cleaning up the dataset is also an important step. In the case of the data set extracted from SC APULUM SA, a higher quality level of data is found, as there are no fields with missing or incomplete data. It is also not necessary to select or integrate data, because all data is relevant, being collected from a real source.

The first step in using Weka to obtain experimental results is to load the preprocessed dataset. Because the data set used is of the ARFF type, the upload is done using the "Open file" button.

The second step in making predictions is selecting the Forecast window. In the case of the basic configuration from the Forecast section, the attributes for which the prediction is to be made will be selected. For experiments, we decided to use both attributes for each type of product. The number of units for which the forecast is made was set to 12, so we will obtain predicted data for the next 12 months.

The advanced prediction configuration window allows the selection of algorithms and the setting of parameters that will help to evaluate it. To choose the algorithm according to the porcelain products forecast, both the numerical results, the graphical ones and the prediction evaluation parameters were taken into account.

After configuring the parameters, we tested several algorithms.

4.2.3 Results and discussions

The present case study includes captures that represent the experimental results of the algorithms. The tests were performed based on each available classifier, but we decided to present those results that have relevance, either positive or negative in choosing the most appropriate algorithm.

The first aspect of the chosen algorithm was the use of several different graphs to be able to study the large discrepancies between the prediction of the initial data and the predicted data, in the sense that we rejected the graphs that they had either constant values or values outside the initial data graph.

Mugs defects forecasting

For the mugs data set, ten algorithms were applied for forecasting. The figures below shows the graphical comparison between algorithms. Based on the graphical evaluation, it was concluded that the following algorithms performed the best: Gaussian Preprocesses, IBk and M5Rules. Results for the rest of the algorithms presents a difference from initial data.

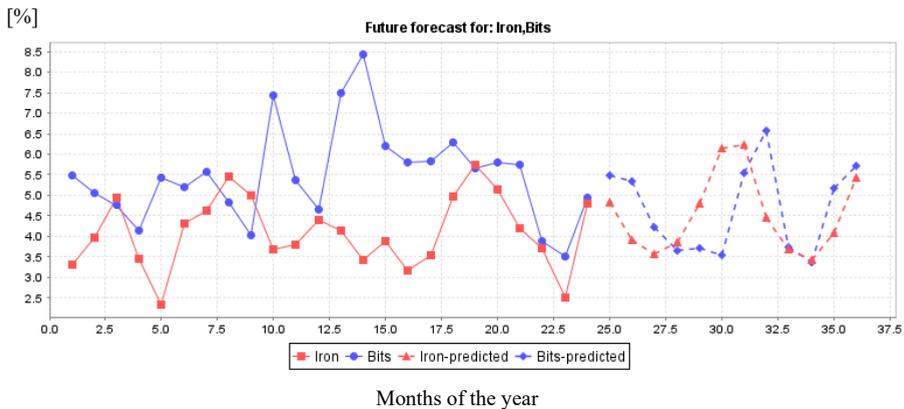


Fig. 3. Mugs defects forecasting –Gaussian Preprocesses algorithm

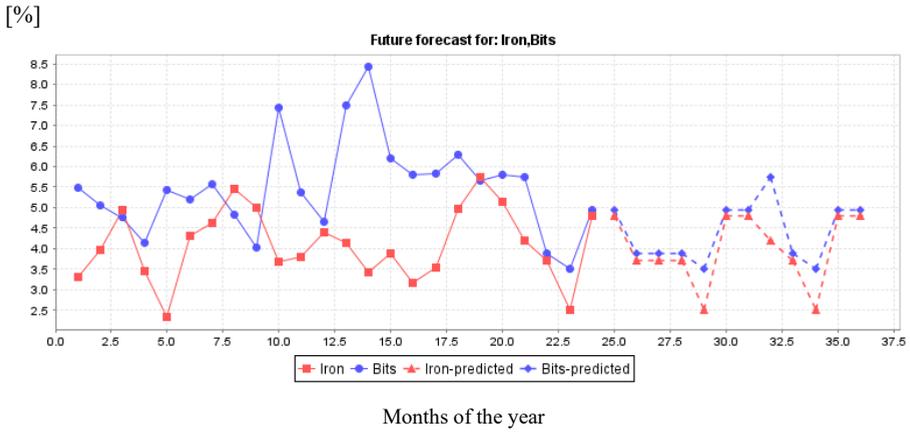


Fig. 4. Mugs defects forecasting – IBk algorithm

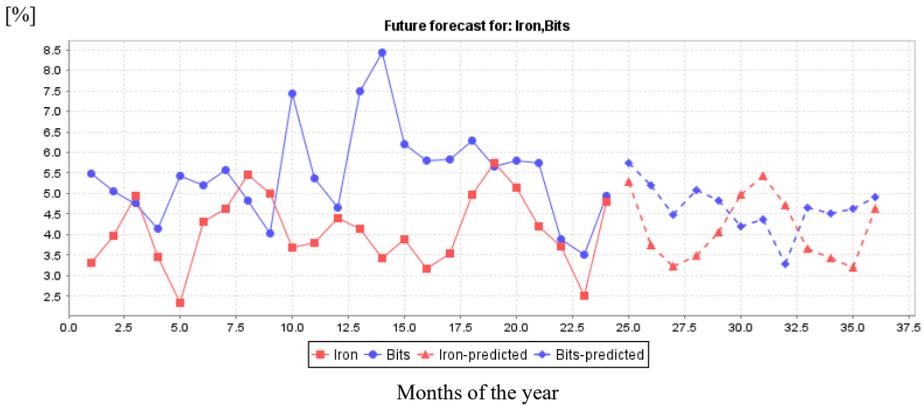


Fig.5. Mugs defects forecasting – M5Rules algorithm.

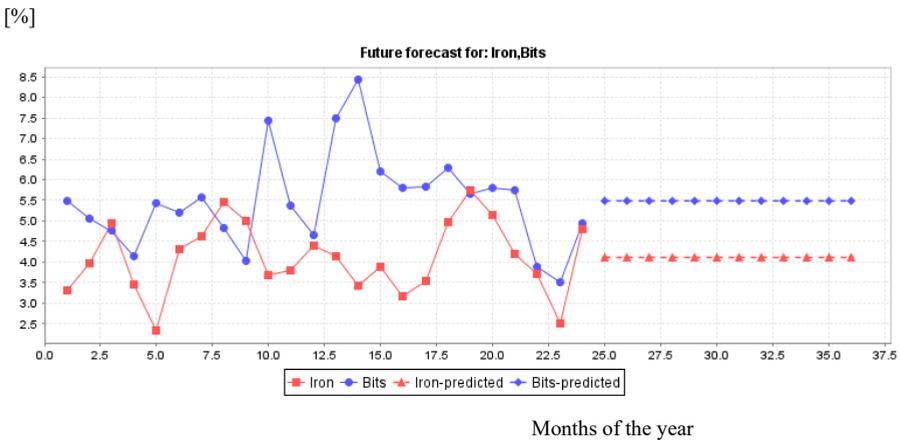


Fig.6. Mugs defects forecasting –ZeroR algorithm

Figure 6 shows the predicted data using the ZeroR algorithm. The data predicted using ZeroR algorithm presents a linear trend, so the algorithm is considered inappropriate for this forecasting problem.

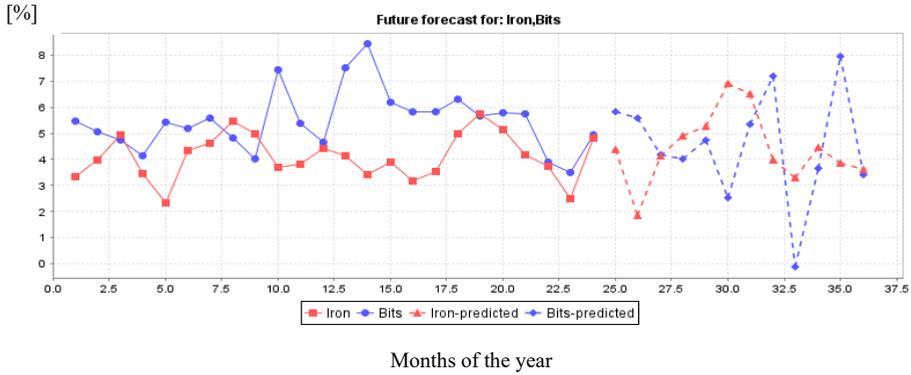


Fig.7.Mugs defects forecasting –Multilayer Perceptron algorithm

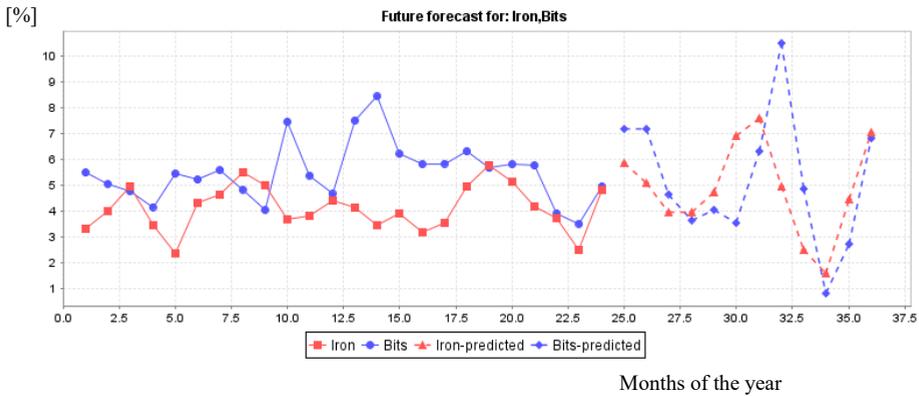


Fig. 8. Mugs defects forecasting – SMOreg algorithm

In Figure7 is presented the forecast using theMultilayer Perceptron algorithm that is inappropriate because the predicted data are well outside the range of input data. InFigure 8 is presented the forecast using the SMOreg algorithm that is inappropriate for the same reason.

Pressed plates defects forecasting

In the pressed plates forecast problem more tested algorithms returned linear predicted data. Graphical analysis shows the best algorithm for pressed plates forecast among those applied is Gaussian Preprocesses (Figure 9).

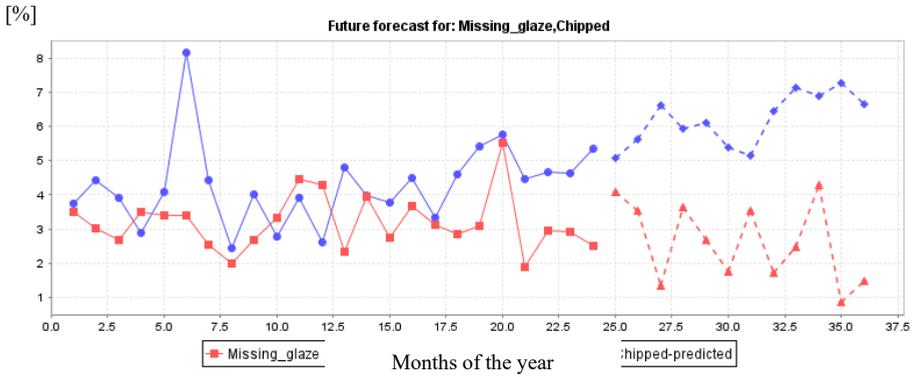


Fig. 9. Pressed plates defects forecasting – Gaussian Preprocessesalgorithm

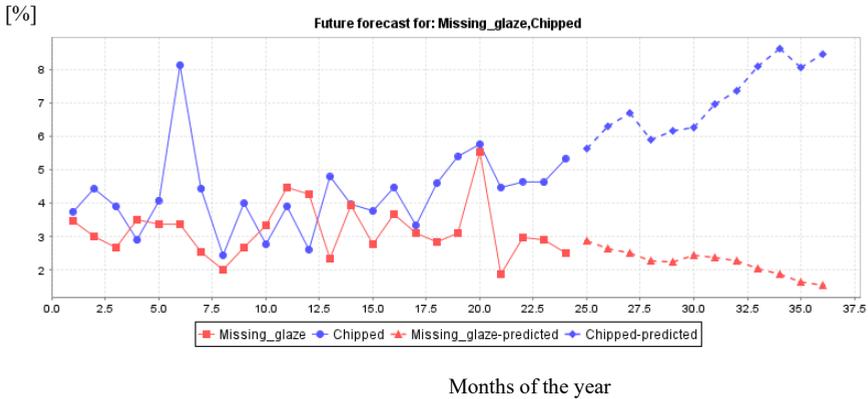


Fig. 10. Pressed plates defects forecasting – M5Rules algorithm

Graphical analysis of Figure 10 shows that the M5Rules algorithm is not appropriate for pressed plates forecasting because the graphic presents an exponential trend.

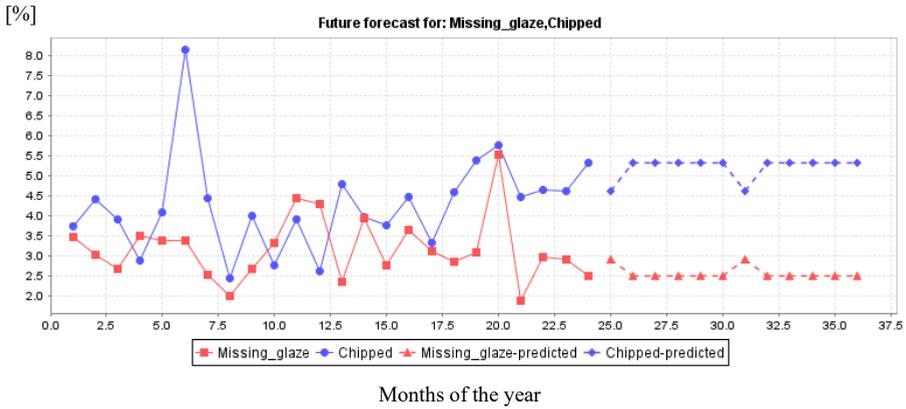


Fig. 11. Pressed plates defects forecasting – IBk algorithm

IBk algorithm (Figure 11) is not appropriate for pressed plates forecasting because the predicted data present a trend.

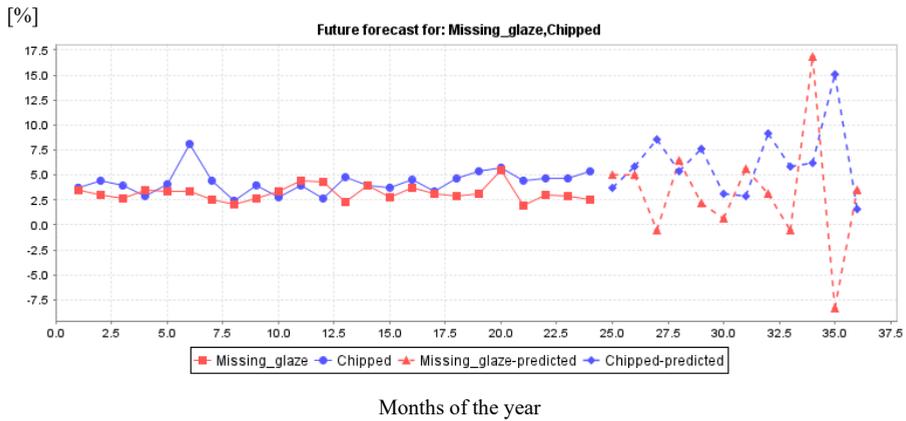


Fig. 12. Pressed plates defects forecasting – Linear Regression algorithm

Figure 12 is presented the forecast using the Linear Regression algorithm that is inappropriate because the predicted data are well outside the range of input data.

Jiggered plates defects forecasting

Several algorithms were applied for jiggered plates defects forecasting. For this product, many algorithms present a trend for predicted data. In the figures below we present a small part of the used algorithms.

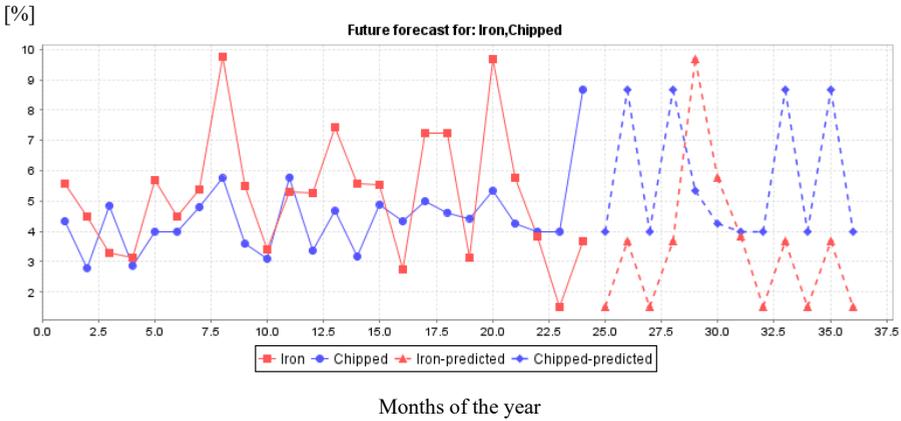


Fig. 13. Jiggered plates defects forecasting – IBk algorithm

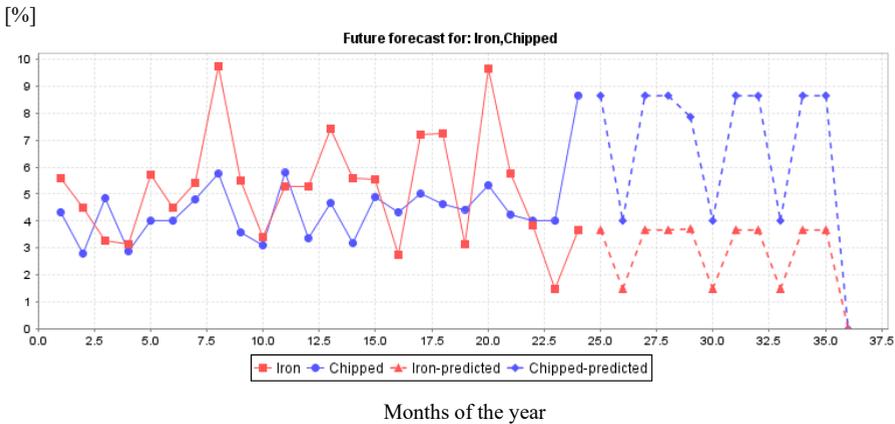


Fig. 14. Jiggered plates defects forecasting – KStar algorithm

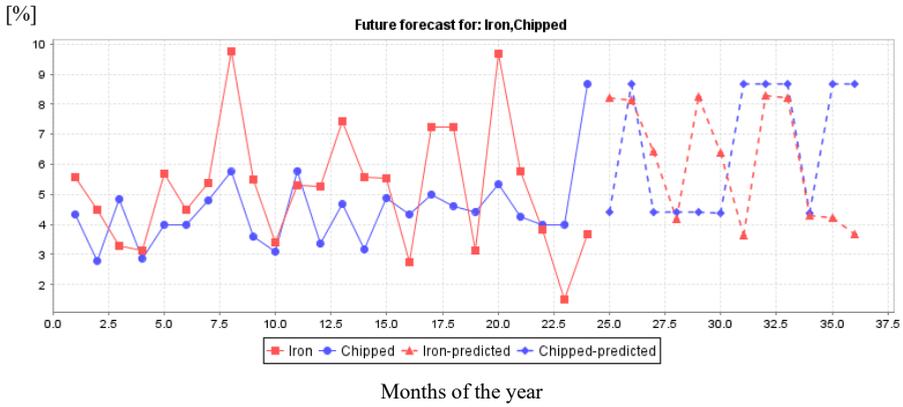


Fig. 15.Jiggered plates defects forecasting – LWL algorithm

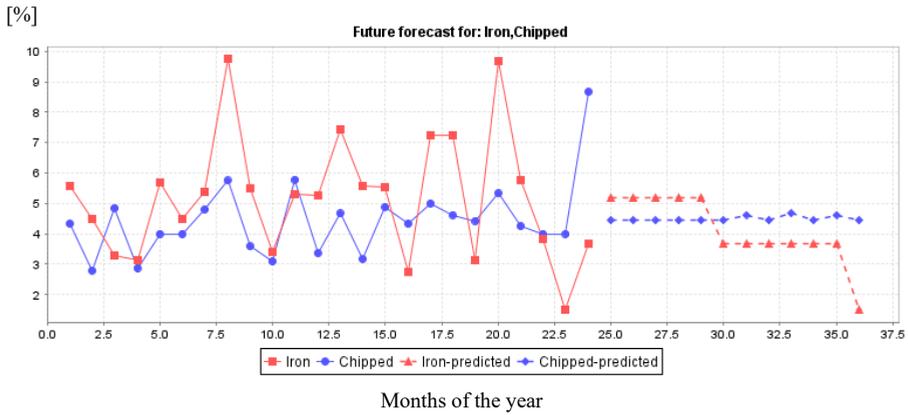


Fig. 16.Jiggered plates defects forecasting – Decision Table algorithm

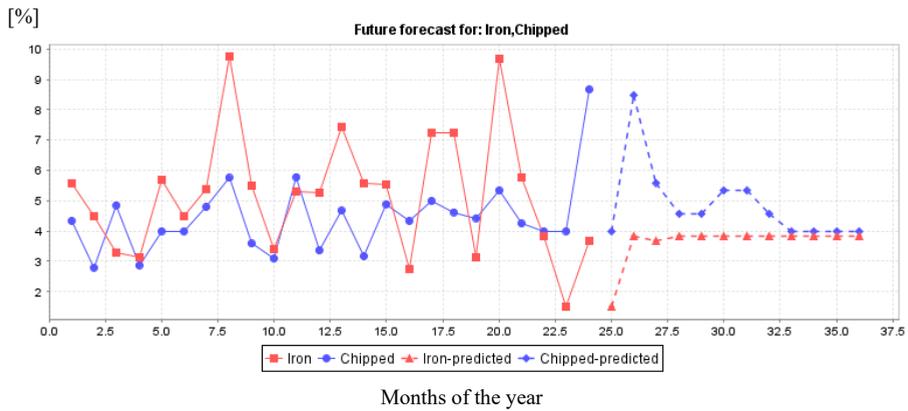


Fig. 17. Jiggered plates defects forecasting – Random Tree algorithm

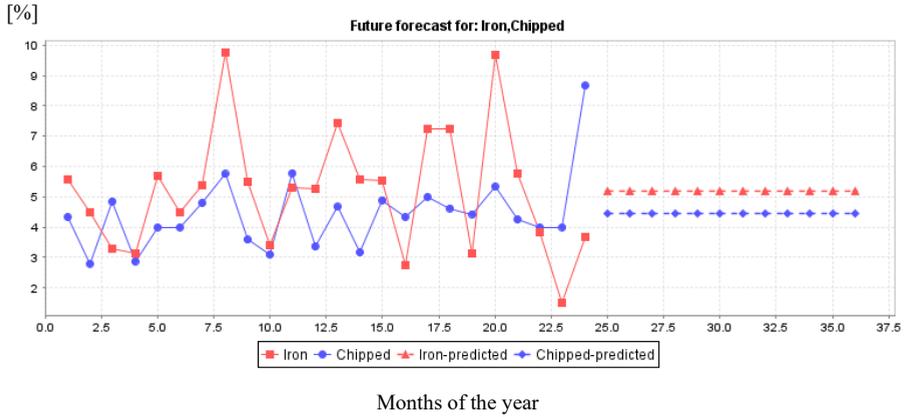


Fig. 18.Jiggered plates defects forecasting – ZeroR algorithm

Graphical results presented in the figures above are in appropriate for jiggered plate defects forecasting because the graphs show a trend or the predicted data are linear.

Figure 19 presents the forecasting of jiggered plate defects using the Gaussian Preprocesses algorithm. Graphical analysis shows that the Gaussian Preprocess algorithm is the most representative of the used algorithms.

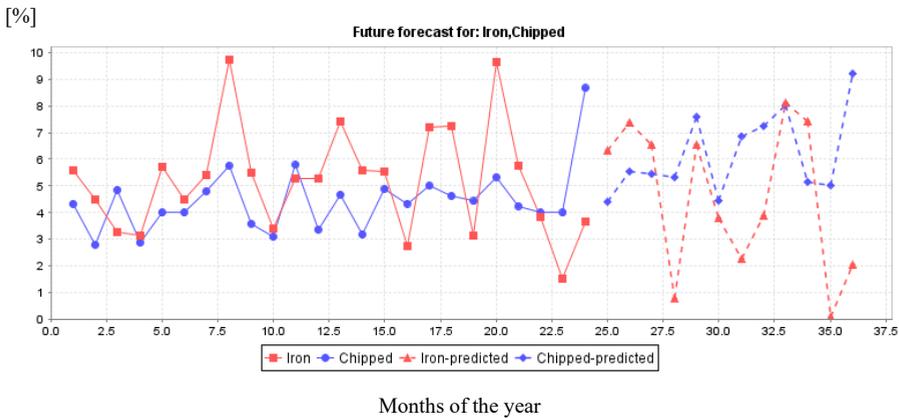


Fig. 19. Jiggered plates defects forecasting – Gaussian Preprocess algorithm

5 Conclusions

Our experimental results show that using the Weka tool we can predict the percentage of defects for different types of products.

The initial values of the considered defects obtained from quality surveys carried out for each type of product highlight the instability of the process regarding the occurrence of defects, so it is quite difficult to establish the prediction that is closest to reality. However, we can affirm that the Gaussian Preprocesses algorithm has the best results for each type of product compared with the others applied algorithms. We used a graphical analysis for evaluating the results of algorithms.

To determine concretely whether machine learning techniques are suitable for defects forecasting to improve the quality management system, we plan to further extend our approach by comparing the results with real data collected in 2020. Also, this evaluation step can improve the final decision about the chosen algorithm in forecasting for each type of product.

Also, if the results obtained from the comparison between the data obtained by applying machine learning techniques and the real data recorded by quality surveys confirm that these techniques can be used successfully in this field, in the future, we will apply machine learning techniques to all quality defects that occur on ceramic products within the organization considered.

References

1. C. ParraLópez, A. HinojosaRodríguez, C. CarmonaTorres, S. Sayadi, *Food Control*, **62**, 23, (2016)
2. L.D.Yanti, A.Firdaus, Millah:*JurnalStudi Agama*, **17**, 31, (2017)
3. J. Priede, *ProcediaSoc.Behav. Sci.*, **58**, 1468, (2012)
4. N. Africano, A.S. Rodrigues, G. Santos, *Qual. Innov. Prosper.*,**23**, 123, (2019)
5. R.Nurchahyo, Zulfadlillah, M. Habiburrahman, *Heliyon*, **7**, 1, (2021)
6. M.A.T.Betegon, N.B.Ramírez, A.M. Arce, J.R. Guzmán, F.J.R. Calvo, S. Avanzada, *Clin.Simul. Nurs.*,**51**, 29, (2021)
7. I. BetllochMas, R. RamónSapena, C. AbellánGarcía, J.C. PascualRamírez,*Actas Dermosifiliogr.*, **110**, 94, (2019)
8. A. Pacana, R. Ulewicz, *Pol. J.Manag.Stud.*, **21**, 284, (2020)
9. C. Oprean, C.V. Kifor, O. Suciu, C. Alexe, *Managementulintegrat al calității* (Editura Academiei Române, București, 2012)
10. M.A. SanchezLizarraga, J. LimonRomero, D.Tlapa, Y. BaezLopez, C.Puente, L. PurtaSierra, S.Ontiveros, *IEEE Access*,**8**, 190677, (2020)
11. R.B. Carey, *Clin. Microbiol. Newsl.*, **40**, 184, (2018)
12. J. AndresJimenez, J.A. MedinaMerodio, L. FernandezSanz, J.J. MartinezHerraiz, E. RuizPardo, *Sustainability*, **12**, 2, (2020)
13. I.N. Hădărean Șomlea, L. Marian, I.S. Ferencz, *Procedia Econ. Finan.*, **15**, 1072, (2014)
14. L.M. Fonseca, J.P. Domingues, *QIP*, **22**, 32, (2018),
15. A. Sârb, L. Itul, M. Popa,*Emerging Markets Economics and Business. Contributions of Young Researchers. Proceedings of the 9th International Conference of Doctoral Students and Young Researchers* (EdituraUniversității din Oradea, Oradea,2018)
16. <http://www.cs.waikato.ac.nz/ml/weka/>, [Accessed 10.03.2021]
17. A. Sârb, I. Glevitzky, L. Itul, M. Popa, *MATEC Web Conf.*, **290**, 5, (2019)
18. A. Sârb, L. Itul, M. Popa, *OJBE*,**4**, 50,(2019)
19. A. Sârb, L. Itul, A. Goleanu, M. Popa, *J. Environ. Prot. Ecol.*, **20**, 1988, 1989, (2019)
20. <http://www.cs.waikato.ac.nz/ml/weka/arff.html>, [Accessed 10.03.2021]