# PREDICTION OF HIGH CYCLE TIMES IN WHEEL RIM MOLDING WITH ARTIFICIAL NEURAL NETWORKS[1]

İnanç KABASAKAL[2], Fatma DEMİRCAN KESKİN[3]

**ABSTRACT**

*Purpose*: This study proposes a two-stage approach to determine a cycle-time threshold and predict high cycle times by examining sample molding process data obtained from a wheel-rim manufacturer.

*Methodology*: Our study firstly determines thresholds for high cycle times with two alternate approaches. Subsequently, data were labeled regarding the cycle-time threshold. Alternate models based on Artificial Neural Networks (ANNs) were developed in R to predict high cycle times.

*Findings*: Our findings include a comparison of cycle-time threshold approaches through a distance-based metric. After labeling of high cycle times, our study presents the performance of alternate predictive models. The performance of models was compared in terms of accuracy, recall and precision.

*Originality*: Process mining in wheel rim molding has been found meager in prior research, despite the abundance of process mining applications and cycle-time prediction models. Another distinctive aspect of the study is cycle-time threshold determination with multiple methods to eliminate manual labeling of processes.

*Keywords*: Cycle Time Prediction, Process Mining, Machine Learning, Artificial Neural Networks.

*JEL Codes:* M11, C38, C45.

# JANT DÖKÜMÜNDE YAPAY SİNİR AĞLARI İLE YÜKSEK ÇEVRİM SÜRELERİNİN TAHMİN EDİLMESİ

**ÖZET**

*Amaç*: Bu çalışmada, bir jant üreticisinden alınan numune kalıplama proses verileri incelenerek, çevrim süresi eşik değerleri belirleyen ve bu değere dayalı yüksek çevrim sürelerini tahmin eden iki aşamalı bir yaklaşım önerilmektedir.

*Yöntem*: Çalışmada öncelikle iki alternatif yaklaşımla çevrim süresi için eşik değer belirlenmektedir. Ardından, eşik değer uyarınca proses verileri etiketlenmektedir. Yüksek çevrim sürelerini tahmin etmek için R'da Yapay Sinir Ağları (YSA) uygulanarak alternatif sınıflandırma modelleri geliştirilmiştir.

*Bulgular*: Çalışmada uzaklık bazlı bir ölçüt aracılığıyla çevrim süresi eşiği belirleme yaklaşımları karşılaştırılmaktadır. Yüksek çevrim sürelerinin etiketlenmesini takiben alternatif tahminleme modellerinin performansları sunulmaktadır. Tahminleyici modellerin performansı doğruluk, duyarlılık ve kesinlik ölçütleri ile karşılaştırılmaktadır.

*Özgünlük*: Literatürde proses madenciliği uygulamaları ve çevrim süresi tahmin modelleri sıklıkla çalışılmış olmasına karşın, jant dökümünde proses madenciliği ile ilgili çalışmalara sık rastlanmamaktadır. Çalışmada bir diğer özgün yön ise, gecikmelerin manuel biçimde etiketlenmesi yerine, çevrim süresi için eşik değer belirleyen çoklu yaklaşım izlenmesidir.

*Anahtar Kelimeler:* Çevrim Süresi Tahmini, Proses Madenciliği, Makine Öğrenmesi, Yapay Sinir Ağları.

*JEL Kodları:* M11, C38, C45.

## 1. INTRODUCTION

Digital technologies have been increasingly used in the manufacturing industry to improve efficiency and increase profitability. Rapid progress in digital technologies has also been reshaping businesses over the last decade. Such progress is often mentioned by scholars as the Industrial Internet or Industry 4.0.

Scholars have often emphasized the term Industry 4.0 as an extensive change in the manufacturing environment. Shrouf et al. (2014) noted that the key ability for smart systems is to adopt behaviors through perceiving the information and derive findings that lead to actions in real-time. Rüßmann et al. (2015) also noticed that a significant improvement for manufacturing would be the sensing and self-learning capability in machinery. Besides, the integration of digital technologies in manufacturing systems improves the monitoring of processes by enabling real-time process data collection.

Process variations are inevitable in production systems. Process variability, which may occur due to many internal factors such as the equipment and materials used, methods applied, operational causes, or external factors, can create variability in the process cycle time (Mrugalska and Ahram, 2017). The high increase in cycle time has significant adverse effects on production performance, traceable in metrics such as productivity and on-time delivery rate. Therefore, it is crucial to predict high cycle times and take action to prevent such problems. Prediction of high cycle times might be particularly useful in industries such as the automotive and automotive supply industry, where fluctuations in productivity might lead to severe consequences. With such considerations, this study aims to present a high cycle time prediction model for molding process in wheel rim manufacturing.

Despite the availability of prior research in cycle time prediction, to our knowledge, there has been no study in this domain where sensor data was used for high cycle time prediction in wheel rim molding. Our model for the problem incorporates a two-stage approach that involves high cycle time threshold determination, and high cycle time prediction. The first stage in our approach incorporates leveraging statistical methods to determine an appropriate threshold rather than manually choosing a specific value. Particularly, a median-based approach of Leys et al. (2013) and the k-means clustering algorithm have been employed one at a time to determine cycle time thresholds. The second stage involves training and using neural networks as binary classifiers towards high cycle time prediction. The model was tested against the process data obtained from a large-scale wheel rim manufacturer in Turkey.

A secondary objective in this study is to compare the effect of threshold determination method selection on the predictive power of our model. In our case, cycle times are labeled after comparing to a threshold, and changing a threshold is a significant decision that modifies the labels within the data before analysis. Therefore, another original aspect of our study is the persistence to obtain better predictive performance at the cost of relabeling data and retraining the predictive model.

This paper is organized as follows. The following section provides a literature review on process mining and prediction of process parameters, including cycle times. Moreover, prior research on high cycle time prediction in relevant processes is summarized in this section. The next section introduces the case study. In the methodology section, the two-stage approach is presented in two stages: threshold selection and high cycle time prediction. The subsequent section involves the results where a detailed comparison of trained models is demonstrated. Finally, the high cycle time prediction model being presented is briefly summarized and reviewed in the conclusion section.

## 2. LITERATURE REVIEW

### 2.1. Process Mining

In manufacturing systems, real-time monitoring of processes, recording real-time values of process parameters, and predicting important process performance measures by applying machine learning methods is a widely followed approach to cut down product failures (Deuse et al., 2019). Monitoring of systems, capturing real-time event logs, mining data to extract potentially valuable information are among essential activities in process mining (Van der Aalst, 2011). In this regard, predicting process-related parameters, including cycle time, can be described as process mining tasks. This concept also relates to the determination of high cycle times in processes.

### 2.2. Process Parameter Prediction

In recent studies, it is noticeable that machine learning methods are often employed to predict the target variables or classify patterns obtained within data. The literature hosts numerous studies that utilize supervised learning techniques for classification and prediction in manufacturing. Yu and Xi (2009) proposed an out-of-control signal detection model based on neural networks for a manufacturing process. Chen (2015) examined the cycle time prediction problem of jobs in wafer fabrication and proposed a

methodology comprising steps of the initial prediction of jobs' cycle time based on ANN, improvements of these predictions by constructing parameters' adjustments, and post-classification using the adjusted parameters.

Quality prediction is another significant prediction task in prior process mining studies. Wang et al. (2019)'s ANN-based and Bai et al. (2019)'s SVM-based models were both proposed to predict quality-related parameters. Moreover, Köksal et al. (2010) examined process data obtained from a business in the molding industry and extracted decision trees that predict batch quality and identify the underlying factors for quality-related problems.

In a relevant study, Quintana et al. (2011) employed ANNs for the surface roughness prediction problem for the milling process. Asiltürk and Çunkaş (2011) employed Multiple Regression and ANN using the measurements taken after the turning operations to predict the steel surface roughness. Moreover, Lieber et al. (2013) used sensor data to predict the quality of products by utilizing Decision Trees, Support Vector Machines, the Nearest Neighbor Method. Table 1 lists several studies that predict process parameters with machine learning methods:

**Table 1. Machine learning models for prediction of process parameters**

| Authors | Predicted Process Parameter | Method(s) |
| --- | --- | --- |
| Yu and Xi (2009) | Out-of-control signals | ANN |
| Chen (2015) | Cycle Time in wafer fabrication | ANN |
| Wang et al. (2019) | Quality-related process parameters | ANN |
| Quintana et al. (2011) | Surface roughness prediction | ANN |
| Asiltürk and Çunkaş (2011) | Steel surface roughness | Multiple Regression and ANN |
| Lieber et al. (2013) | Product quality | Decision Trees, Nearest Neighbor Method, Support Vector Machines |
| Bai et al. (2019) | Quality-related process parameters | Support Vector Machines |

In the context of manufacturing, cycle time and throughput are essential measures that help decision-makers to assess performance (Chien et al., 2012). Cycle time is one of the most significant measures reflecting many production system performance indicators such as the production rate and quality, productivity, required inventory amounts to be held, and the on-time delivery ratio (Backus et al., 2006; Khan et al., 2014). In prior research, cycle time is often discussed along with Little's Law (Little, 1961) that relates cycle time with the WIP (work in process) and throughput. Accordingly, average throughput equals the ratio of total WIP over the average cycle time; however, the rule results in long-term results with relatively steady conditions (Rust, 2008).

A significant opportunity led by digital technologies in manufacturing is the increased availability of real-time data. In their discussion of the use of Industry 4.0 technologies in manufacturing, Kolberg and Zühlke (2015) noted that smart machines might negotiate cycle-times to contribute to a higher degree of lean automation. In this regard, the cycle-time prediction might be dealt with as a problem with a context-aware aspect. Besides, data-driven models that rely on contextual data have been proposed to predict cycle time more instantaneously.

**2.3. High Cycle Time Prediction**

Cycle time estimation or high cycle time prediction has been addressed in various domains. Due to the scarcity of prior models for this problem in wheel rim molding, this study attempts to present a classification model. On the other hand, there are processes that exhibit similar characteristics, such as the degree of variability in the process. For instance, the wafer fabrication process often hosts cycle time prediction problems. The prediction models in this domain involve techniques such as simulation, fuzzy modeling, neural networks, and hybrid techniques. Moreover, Chen et al. (2009) noticed that neural networks are often employed due to their effectiveness and accuracy for cycle time prediction in wafer fabrication.

Goodwin et al. (2004) covered the problem of cycle-time prediction for semiconductor manufacturing and underlined the necessity of statistical learning and data mining techniques due to the variety of stochastic variables in the process. Additionally, the authors presented predictions with several techniques, including ANNs.

For the prediction of cycle time and related measures, the wafer fabrication process has been occasionally revisited by scholars. Chen (2007) handled the output time prediction problem by firstly classifying the wafer lots with the k-means algorithm. Chen et al. (2008) predicted job completion times with a hybrid approach, including the use of neural networks and self-organization maps in this process.

Moreover, Chang et al. (2009) examined the flow time prediction problem about the same process and proposed a multi-step approach in which the data is clustered and flow times are individually predicted for each cluster. In prediction, the authors used a case-based reasoning approach where the feature weights are adjusted by the Genetic Algorithm. Additionally, Meidan et al. (2011) investigated the cycle time prediction problem for the wafer fabrication process by focusing on the waiting time, which is among the most significant components of cycle time. Such an approach relied on firstly determination of the key identifiers of waiting time by feature selection; then, prediction of the classes of waiting time with naive Bayesian classifier by using these identifiers. Furthermore, Wang and Zhang (2016) dealt with the cycle time prediction problem by focusing on the determination of the features that reflect the characteristics of the wafer fabrication system being analyzed. Accordingly, the authors employed a conditional mutual information-based approach in feature selection; then, a hybrid backpropagation network and fuzzy c-means approach in cycle time prediction.

Depending on the importance of cycle time, in the literature, the prediction of cycle time has attracted considerable interest. To date, many approaches, including statistical, simulation-based, analytical, artificial intelligence, hybrids of those, have been proposed and applied to predict cycle time in different types of industries (Chen, 2007; Wang and Zhang, 2016). Siller et al. (2006) proposed an instance of such models for predicting cycle time in a high-speed milling process. Chien et al. (2005) analyzed historical semiconductor production line data for cycle time prediction using Decision Trees. Furthermore, recent studies also incorporated a combination and comparison of multiple methods for cycle-time prediction. Backus et al. (2006) noted that most studies involve statistical analysis (such as regression technique), simulation models, and neural networks to obtain cycle time; and compared the predictive performance of Classification and Regression Trees (CART), ANN, and K-Nearest Neighbors methods using mean and median absolute error values. In a more recent study, Polato et al. (2014) developed a regression model to predict the remaining time for completion, rather than the cycle time; and noted that the use of such models provides an output at real-time for operational support that help to take proper actions.

The molding process is a process where process dynamics vary depending on time; the used raw materials change state, and therefore show non-linear properties and involve multi successive steps (Zhou, 2013). The high variability of the molding process makes it harder to determine whether an increase in cycle time depends on the variability caused by the nature of the process, or if it is due to process failures that can be eliminated. To distinguish these preventable increases in cycle time is a crucial step to improve business performance. Additionally, the benefits of predicting and preventing high cycle times might be noticed through performance indicators such as production rate, product quality, on-time delivery rate, and productivity.

A prerequisite for preventing high cycle times is essentially describing a rule that signifies whether a cycle time is normal or high. In this manner, several approaches from prior research might be adopted to classify processes over their cycle times or figure out processes with high-cycle time as outliers.

For the highly variable datasets that do not follow a normal distribution, approaches based on median and median absolute deviation measures are seen as more appropriate than the mean and standard deviation-based approaches in determining outliers (Da Costa et al., 2016; Leys et al., 2013). In the literature, there have been studies adopting the median-based approach by Leys et al. (2013), which is one of the most applied median-based outlier detection approaches for determining outliers in data pre-processing (Marti-Puig et al., 2018; Herrema et al., 2019), and as a fault detection strategy (Muhammed and Shaikh, 2017).

As another approach, clustering is an unsupervised learning task that groups similar data objects into clusters. The basic idea in clustering tasks is to group similar objects into clusters that are separate from each other (Han et al., 2012:490). K-means algorithm randomly generates k clusters at the first step, then places data objects into the nearest clusters using similarity criterion while recalculating the cluster means (Han et al., 2012:453). K-means clustering method is also useful in the identification of outliers (Sumathi and Sivanandam, 2006) since outliers do not easily classify in a particular cluster.

Additionally, several studies predict measures related to cycle time in various molding processes. Yarlagadda and Khong (2001) developed an ANNs based model to predict the injection time and pressure in the injection molding process. Ramkumar et al. (2015) followed a simulation-based approach for the cycle time prediction problem for the heating process in rotational molding. Kozjek et al. (2019) handled the fault detection problem in the plastic injection molding process with data mining approaches. The authors labeled the unscheduled machine stoppage observations in their dataset as faulty and the rest observations as normal, and aimed to identify these categories using several parameters, including cycle time.

Despite the variety of process mining models for high cycle prediction in various domains, it has been noted that such studies are meager in wheel rim molding. Arguably, this originates from the recent digitalization in wheel rim manufacturing plants, and our study benefits the collection of sensor data during the molding process. Besides, prior models on high cycle time prediction often used labeled data or a predefined threshold. Our approach differentiates from prior research with the threshold selection procedure, where the predictive performances of trained models were compared to determine the best method for threshold determination.

## 3. PROBLEM DESCRIPTION

Our case study addresses the prediction of high cycle times in the wheel rim molding process. Molding is one of the initial phases in wheel rim manufacturing. Various sub-processes take place during this phase, including the heating of the metal, melting, pouring, pressurizing, and cooling consecutively. The units are formed during this complex process which even obscures the identification of individual products at this step.

The molding process data analyzed was obtained from a large-scale wheel rim manufacturer located in Turkey. The original dataset was organized as a comma-separated text file, and this study examines a dataset of 1709 recordings for wheel rims molded into a standard cast on the same day. Each data point involves sensor-generated attributes, including relevant temperature, pressure, flow data measurements recorded during the sub-processes. Moreover, the specific durations of several key steps were also involved in the dataset. The independent variables and the range of values are demonstrated in Table 2.

**Table 2. Variables and range of values in dataset**

| No | Variable | Range | No | Variable | Range |
|----|----------|-------|----|----------|-------|
| 1 | METAL_S_SET (°C) | [705, 710] | 27 | X7_COOL_DURATION (sec) | [120, 140] |
| 2 | METAL_SAKT (°C) | [685, 712] | 28 | X8_FLOW_AVG | [931,78; 1019,04] |
| 3 | MOD | [1, 8] | 29 | X9_FLOW_AVG | [899,43; 976,58] |
| 4 | METAL_INTAKE | [0, 5] | 30 | X9_W_DURATION (sec) | [90, 100] |
| 5 | PRESSURE_MAX | [5,73; 6,11] | 31 | X10_FLOW_AVG | [908,10; 984,48] |
| 6 | COOLING_TEMP (°C) | [507, 526] | 32 | X10_W_DURATION (sec) | [90, 100] |
| 7 | TC1_SET (°C) | [450, 500] | 33 | X11_FLOW_AVG | [0; 978,22] |
| 8 | TC1_AKT | [368, 3276] | 34 | X11_W_DURATION (sec) | [80, 100] |
| 9 | TC2_AKT | [269, 422] | 35 | X11_COOL_DURATION (sec) | [100, 120] |
| 10 | PHASE_1_DURATION (sec) | [5, 30] | 36 | XX11_FLOW_AVG | [0; 19,35] |
| 11 | PHASE_1_PRESSURE | [131,22; 514,87] | 37 | XX11_W_DURATION (sec) | [45, 65] |
| 12 | PHASE_2_DURATION (sec) | [7, 35] | 38 | XX11_COOL_DURATION (sec) | [30, 60] |
| 13 | PHASE_2_PRESSURE | [301,22; 1017,06] | 39 | XX12_FLOW_AVG | [0; 13,57] |
| 14 | PHASE_3_DURATION (sec) | [7, 133] | 40 | XX12_W_DURATION (sec) | [45, 65] |
| 15 | PHASE_4_DURATION (sec) | [40, 190] | 41 | XX12_COOL_DURATION (sec) | [30, 60] |
| 16 | PHASE_4_PRESSURE | [0, 1000] | 42 | XX13_FLOW_AVG | [0; 11,11] |
| 17 | PHASE_5_DURATION (sec) | [0, 50] | 43 | XX13_W_DURATION (sec) | [45, 65] |
| 18 | X1_FLOW_AVG | [0; 1233,23] | 44 | XX13_COOL_DURATION (sec) | [30, 60] |
| 19 | X2_FLOW_AVG | [1158,35; 1372,43] | 45 | XX14_FLOW_AVG | [0; 13,79] |
| 20 | X2_W_DURATION (sec) | [80, 100] | 46 | XX14_W_DURATION (sec) | [45, 65] |
| 21 | X2_COOL_DURATION (sec) | [120, 140] | 47 | XX14_COOL_DURATION (sec) | [30, 60] |
| 22 | X3_FLOW_AVG | [0; 1352,46] | 48 | XX2_FLOW_AVG | [16,4; 17,78] |
| 23 | X3_W_DURATION (sec) | [80, 100] | 49 | XX2_W_DURATION (sec) | [100, 120] |
| 24 | X3_COOL_DURATION (sec) | [120, 140] | 50 | XX2_COOL_DURATION (sec) | [80, 100] |
| 25 | X7_FLOW_AVG | [1133; 1234,72] | 51 | CYCLE_TIME (sec) | [236, 25438] |
| 26 | X7_W_DURATION (sec) | [80, 100] | | | |

The input variables in the dataset have been adjusted into the range [0,1] with max-min normalization. Moreover, the distribution of cycle times was tested using Shapiro-Wilk's method in R. The normality test

resulted in the p-value of 2.2e-16 < 0.05. Accordingly, the distribution of the cycle time values is significantly different from the normal distribution. The density plot for cycle times is depicted below (Figure 1).
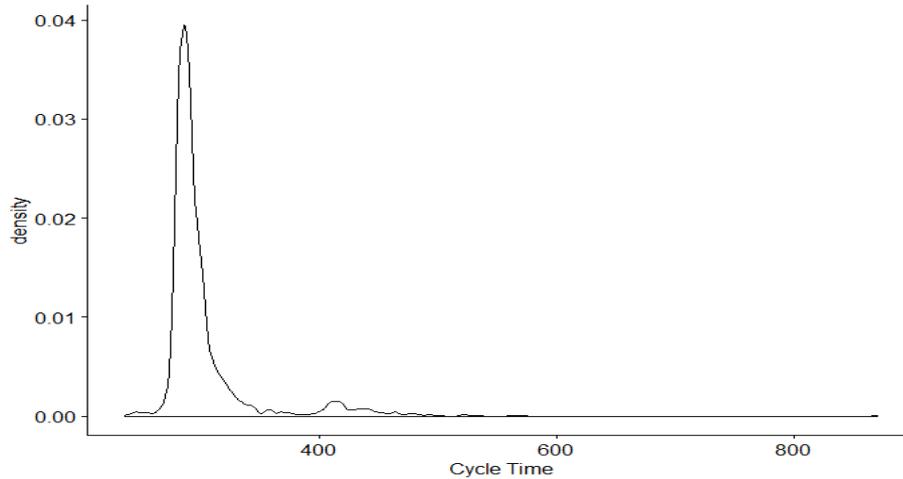


**Figure 1. Distribution of cycle time values**

## 4. METHODOLOGY

Our methodology consists of two stages for high cycle-time prediction, as summarized in Figure 2. Accordingly, our approach involves threshold determination and data labeling before training and testing prediction models.

A typical and straightforward approach would be training a prediction model in case class labels existed in the process data. However, our dataset includes cycle times, and a threshold should be determined to determine labels for data points. Choosing a high threshold would lead to a small group of data points labeled as "late", and vice versa. For this reason, our methodology involves labeling the process data with multiple thresholds one at a time and choosing the better threshold that leads to better predictive performance.
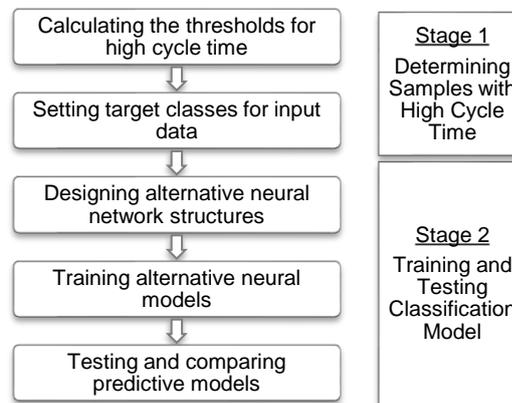


**Figure 2. Threshold determination and prediction in our methodology**

As demonstrated in Figure 2, first stage involves threshold determination to classify cycle times as high or normal, and labeling data accordingly. In the second stage, ANN models are trained and tested to predict the target values. The steps in both stages are introduced in more detail through the following subsections.

### 4.1. Determining Samples with High Cycle Time

Labeling a process based on its cycle time necessitates a threshold that differentiates normal and high cycle times. For this purpose, threshold determination was conducted as detailed in the following subsection. The next subsection briefly describes how data points are labeled right after threshold determination. Completing both steps at Stage 1 (Figure 2) prepares a labeled dataset that enables training a predictive machine learning model.

### 4.1.1. Calculating Thresholds for High Cycle Time

Our methodology employs two alternative approaches for high cycle-time threshold determination. The first approach calculates the median absolute deviation of cycle time recordings and determines the threshold based on this value. The second approach clusters process data regarding their cycle times. Thereby, the cluster with the higher cycle-time average involves potentially late processes. With this idea, determining a threshold separates clusters having normal and relatively high cycle time averages.

Our first approach for threshold determination is based on the method proposed by Leys et al. (2013) that relies on the use of the mean absolute deviation. The authors argued that the mean measure is sensitive to outliers, especially when the magnitude of outliers is remarkable. Due to the high variability in the cycle times and the existence of possible outliers, this approach was adopted to determine the data points with high cycle times.

Table 3 demonstrates our threshold computation procedure according to the method by Leys et al. (2013). The median and median absolute deviation measures for cycle-time were computed with *median*() and *mad*() functions in R. As detailed below, such computation resulted in a threshold of 325.58.

**Table 3. Cycle time threshold by median-based approach**

| Measure | Value (sec) |
|---|---|
| Median | 290,00 |
| Median Absolute Deviation (MAD) | 11,86 |
| Threshold = Median + 3 x MAD | 325,58 |

Our second approach to obtain a threshold was to apply the k-means algorithm to cluster data based on cycle times measured for individual parts. The technique was executed to obtain two clusters from the initial dataset, resulting in the clusters demonstrated in Table 4.

**Table 4. Initial clustering results with k-means (k=2)**

| Cluster | Members | Cluster Mean |
|---|---|---|
| 1 | 1 | 25438 |
| 2 | 1708 | 325,5 |

The clustering results in Table 4 apparently demonstrate an outlier for the record with an extremely high cycle time of 25,438. With such consideration, this record was removed as a part of the data preprocessing phase. The k-means clustering was repeated by taking k=2 and k=3 to obtain two and three clusters, respectively. The clusters obtained with both parameters are listed in Table 5:

**Table 5. Clustering results with k-means for k=2 and k=3**

| Parameter | Clusters | Mean | Range | Members | Cluster Similarity |
|---|---|---|---|---|---|
| k=2 | 1 | 4515,78 | [3121, 9761] | 9 | 77,6% |
| | 2 | 303,31 | (0, 3121) | 1699 | |
| | | | Total | 1708 | |
| k=3 | 1 | 4515,78 | [3121, 9761] | 9 | 78,6% |
| | 2 | 453,90 | [374, 3121) | 111 | |
| | 3 | 292,78 | (0, 374) | 1588 | |
| | | | Total | 1708 | |

Both sets of clusters identified a particular group of 9 processes with cycle times higher than 3121. Taking this threshold would result in 10 high cycle times, including the outlier. Moreover, clustering for k=3 discovered another group of 111 data points with relatively high cycle times. At this step, a comparison of both clustering results (k=2 and k=3) was conducted using the distance-based similarity metric, calculated as in Equation 1.

$$dist = \frac{sum\ of\ squares\ within\ cluster}{sum\ of\ squares\ to\ other\ data\ points} \tag{1}$$

The distance-based similarity for clusters obtained by k=3 was found slightly better than those obtained by k=2. (78,9% vs 77,6%). Moreover, the scarcity of members of Cluster 1 in both groups was noticed. In addition, Cluster 2 in the second clustering solution (k=3) also involved 111 members that still have relatively high cycle times. With those considerations, the second group of clusters obtained with k=3 was picked as the more appropriate option.

The minimum value of two clusters (clusters of 1 and 2, when k=3) was taken as the threshold to determine high cycle times in our k-means clustering approach.

### 4.1.2. Setting Target Classes for Input Data

The thresholds determined in the previous step were used to label the dataset in terms of cycle times. Accordingly, the target variable 'LATE' was set with binary values where 1 indicates high, and 0 indicates normal cycle times. Each record was labeled by simply comparing its cycle time with the threshold.

Since our study adopts two different approaches in threshold computation, this step was repeated for both thresholds.

### 4.2. Training and Testing Classification Model

Referring to our methodology in Figure 2, the main idea in Stage 2 is training and testing predictive models based on data labeled in Stage 1.

Jain et al. (1996) described ANNs as parallel computing systems that involve large numbers of interconnected processors, with the ability to adopt supervised and unsupervised learning abilities and perform tasks such as prediction, pattern classification, and function approximation. ANNs are often used within predictive models that analyze process data in manufacturing. One of the main advantages of ANNs is the ability to process information properties in complex non-linear and quite dynamic environments, even in datasets containing noisy information, thanks to their learning and generalization capabilities (Zhang et al., 1998; Noorzaei et al., 2007; Singh et al., 2016). Moreover, ANNs have higher fault tolerance than the supervised machine learning methods of Support Vector Machine and Random Forest. So, they are capable of handling data with missing values. ANNs are also scalable compared to Support Vector Machine and Random Forest (Abiodun et al., 2018). In this regard, our study employs ANN-based binary classifiers to predict high cycle times. Predictive model design, training, and testing steps are detailed in the following subsections.

### 4.2.1. Deriving Neural Network Structures

The selection of network structure is a crucial decision in training neural networks. However, Rafiq et al. (2001) noticed the lack of a universal method for selecting the number of hidden neurons when designing neural networks. Moreover, the authors underlined that the abundance or shortage of neurons might lead to problems such as under-fitting and over-fitting.

Due to the considerations listed, our methodology aims to minimize those risks by training multiple neural networks of different structures. Prediction models in our study are composed of ANNs with alternate structures:

- Three single-layer neural networks with hidden neuron counts of 8, 20, and 25, respectively.
- A multi-layer neural network with 19 neurons.

The variables listed in Table 2, except for the cycle-time, correspond to input nodes of ANNs. The output is a binary variable that signifies whether the cycle time is high or normal.

### 4.2.2. Training of Neural Network Models

A series of repetitions were executed with four individual settings that differ by the network structures. The *neuralnet* library (Günther and Fritsch, 2010) in R was utilized for training neural networks that predict high cycle times based on input variables.

The neural networks were trained using 80% of random samples from the original dataset. Consequently, the remaining 20% was used to test the models. The selection of training and test data was randomly held at each iteration, with preserving the balance of the percentage of high cycle times both in training and test data.

### 4.2.3. Testing and Comparing Predictive Models

The predictive power of machine learning models is often measured using specific measures such as accuracy, precision, recall (Géron, 2019:88-92). Accordingly, our results involved the comparison of prediction models in terms of those measures.

The predictive models are compared with respect to the following measures (Equations 2-4).

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Positives + Negatives} \tag{2}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{3}$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{4}$$

Accuracy measures the percentage of objects correctly classified, while recall corresponds to the completeness of positive classification, and precision explains what percentage of tuples labeled as positive are actually positive (Han et al., 2012:365,368).

## 5. FINDINGS

The average prediction results over 100 repetitions are cumulated in the confusion table demonstrated in Table 6. The classes 0 and 1 correspond to normal and high cycle times, respectively, regarding the cycle-time threshold.

**Table 6. Overall predictive performance of trained models (with median-based threshold=325.58)**

| Network Structure | | | | | | Performance Metrics (Average) | | |
|---|---|---|---|---|---|---|---|---|
| Layers | Hidden Neurons | Confusion Matrices | | | | Recall | Precision | Accuracy |
| 1 | 8 | | | Predictions | | | | |
| | | | | 0 (Normal) | 1 (High) | 29,25% | 38,13% | 86,87% |
| | | Actual | 0 (Normal) | 285,99 | 18,03 | | | |
| | | | 1 (High) | 26,87 | 11,11 | | | |
| 2 | (15, 4) | | | Predictions | | | | |
| | | | | 0 (Normal) | 1 (High) | 29,74% | 38,35% | 86,87% |
| | | Actual | 0 (Normal) | 285,79 | 18,18 | | | |
| | | | 1 (High) | 26,72 | 11,31 | | | |
| 1 | 20 | | | Predictions | | | | |
| | | | | 0 (Normal) | 1 (High) | 31,02% | 32,04% | 87,27% |
| | | Actual | 0 (Normal) | 288,45 | 21,25 | | | |
| | | | 1 (High) | 22,28 | 10,02 | | | |
| 1 | 25 | | | Predictions | | | | |
| | | | | 0 (Normal) | 1 (High) | 29,79% | 36,22% | 86,42% |
| | | Actual | 0 (Normal) | 284,27 | 19,86 | | | |
| | | | 1 (High) | 26,59 | 11,28 | | | |

According to Table 6, training ANNs with 20 hidden neurons led to the highest average performance. In our results, a high accuracy rate indicates the overall success of predictions; a high recall rate signifies the ability to predict high cycle times; and a high precision implies the exactness of positive predictions. The average accuracy of all trained models is 86.86%, where recall is found 29.95%, and precision is 36.19%. Low recall scores indicated that the models overlooked a significant number of occasions with high cycle times.

**Table 7. Overall predictive performance of trained models (with clustering-based threshold=374.00)**

| Network Structure | | | | | | Performance Metrics (Average) | | |
|---|---|---|---|---|---|---|---|---|
| Layers | Hidden Neurons | Confusion Matrices | | | | Recall | Precision | Accuracy |
| 1 | 8 | | | Predictions | | | | |
| | | | | 0 (Normal) | 1 (High) | 39,63% | 48,47% | 92,81% |
| | | Actual | 0 (Normal) | 307,89 | 10,11 | | | |
| | | | 1 (High) | 14,49 | 9,51 | | | |
| 2 | (15, 4) | | | Predictions | | | | |
| | | | | 0 (Normal) | 1 (High) | 38,54% | 50,11% | 92,99% |
| | | Actual | 0 (Normal) | 308,79 | 9,21 | | | |
| | | | 1 (High) | 14,75 | 9,25 | | | |
| 1 | 20 | | | Predictions | | | | |
| | | | | 0 (Normal) | 1 (High) | 40,58% | 48,22% | 92,77% |
| | | Actual | 0 (Normal) | 307,54 | 10,46 | | | |
| | | | 1 (High) | 14,26 | 9,74 | | | |
| 1 | 25 | | | Predictions | | | | |
| | | | | 0 (Normal) | 1 (High) | 39,17% | 46,33% | 92,55% |
| | | Actual | 0 (Normal) | 307,11 | 10,89 | | | |
| | | | 1 (High) | 14,6 | 9,4 | | | |

As noted, the predictive models summarized in Table 6 relied on labels (high/normal) obtained by taking the median-based threshold. Alternatively, our methodology involved another approach for threshold

determination with the use of the k-means clustering technique. Accordingly, data was relabeled by comparing each cycle time with this clustering-based threshold of 374.00. Taking this dataset, we iterate the same number of tests with identical network structures. The performance metrics obtained through 100 repetitions are presented in Table 7.

The average accuracy in Table 7 is 92.78%, where average recall and precision scores are 39.48% and 48.22%, respectively. In addition to a noticeable increase in the accuracy, the average recall was found significantly higher when the threshold was set according to the k-means clusters. Moreover, the same conclusion still holds when comparing the ANN structures one at a time. Such difference signals the impact of threshold in model performance.

## 6. CONCLUSION

Cycle times in manufacturing are among essential indicators that might signal unforeseen stoppages and failures when high cycle times occur. In this regard, predicting high cycle times might be helpful in detection of such incidents, and productivity improvements in a manufacturing plant.

This study addresses a high cycle time prediction problem and proposes an ANN-based high cycle time prediction model with two stages. The first stage involves cycle time threshold determination through statistical and clustering-based methods. The second stage involves ANN-based binary classification for predicting high cycle times. To demonstrate our model, we analyzed molding process data obtained from a wheel rim manufacturer. The dataset involved process-specific attributes collected via sensors, and cycle times recorded.

Our study contributes to prior studies by involving and comparing multiple threshold determination approaches used when labeling high cycle times. Before the tests, we might speculate that a high threshold would label fewer observations with higher delays. Consequently, it should be easier to distinguish those observations since they are more deviated from other data points. Alternate neural network structures were repeatedly trained and tested to predict high cycle times by taking clustering-based and median-based thresholds, which are 374 and 325, respectively. The performance of model structures was evaluated in terms of accuracy, recall, and precision in our findings. We noticed that threshold determination has fundamentally affected the average performance of classifiers in all tested configurations. The results demonstrate that taking the higher threshold obtained through k-means clustering has consistently led to better predictions in terms of accuracy, recall, and precision. Nevertheless, we underline that prediction models with different thresholds might signify delays or problems at different scales. Accordingly, it might still be beneficial to use multiple prediction models to determine the severity of alerts to be raised in real-time monitoring systems.

A limitation for the study was using a dataset that involved process data captured from a single batch. In future studies, the proposed methodology might be further tested with a broader set of attributes for the molding process by virtue of forthcoming sensors. Moreover, the model might be even extended to provide cycle time estimates rather than merely signifying high cycle times. Furthermore, the high cycle time prediction model might be modified for handling real-time data and integrated into decision support models to reinforce online monitoring systems in wheel rim manufacturing.

# REFERENCES

Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A. and Arshad, H. (2018). "State-of-the-Art in Artificial Neural Network Applications: A Survey", *Heliyon*, 4(11), e00938.

Asiltürk, I. and Çunkaş, M. (2011). "Modeling and Prediction of Surface Roughness in Turning Operations Using Artificial Neural Network and Multiple Regression Method", *Expert Systems with Applications*, 38(5), 5826-5832.

Backus, P., Janakiram, M., Mowzoon, S., Runger, C. and Bhargava, A. (2006). "Factory Cycle-Time Prediction with a Data-Mining Approach", *IEEE Transactions on Semiconductor Manufacturing*, 19(2), 252-258.

Bai, Y., Sun, Z., Zeng, B., Long, J., Li, L., De Oliveira, J.V. and Li, C. (2019). "A Comparison of Dimension Reduction Techniques for Support Vector Machine Modeling of Multi-Parameter Manufacturing Quality Prediction*", Journal of Intelligent Manufacturing*, 30(5), 2245-2256.

Chang, P.C., Fan., C.Y. and Wang, Y.W. (2009). "Evolving CBR and Data Segmentation by SOM for Flow Time Prediction in Semiconductor Manufacturing Factory", *Journal of Intelligent Manufacturing*, 20(4), 421-429.

Chen, T. (2007). "An Intelligent Hybrid System for Wafer Lot Output Time Prediction", *Advanced Engineering Informatics*, 21(1), 55-65.

Chen, T. (2015). "Combining Statistical Analysis and Artificial Neural Network for Classifying Jobs and Estimating the Cycle Times in Wafer Fabrication", *Neural Computing and Applications*, 26(1), 223-236.

Chen, T., Jeang, A. and Wang, Y.C. (2008). "A Hybrid Neural Network and Selective Allowance Approach for Internal Due Date Assignment in a Wafer Fabrication Plant", *The International Journal of Advanced Manufacturing Technology*, 36(5-6), 570-581.

Chen, T., Wu, H.C. and Wang, Y.C. (2009). "Fuzzy-Neural Approaches with Example Post-Classification for Estimating Job Cycle Time in a Wafer Fab", *Applied Soft Computing*, 9(4), 1225-1231.

Chien, C.F., Hsiao, C.W., Meng, C., Hong, K.T. and Wang, S.T. (2005). "Cycle Time Prediction and Control Based on Production Line Status and Manufacturing Data Mining", *IEEE International Symposium on Semiconductor Manufacturing (ISSM 2005),* San Jose, USA, 327-330.

Chien, C.F., Hsu, C.Y. and Hsiao, C.W. (2012). "Manufacturing Intelligence to Forecast and Reduce Semiconductor Cycle Time", *Journal of Intelligent Manufacturing*, 23(6), 2281-2294.

Da Costa, D.A., Mcintosh, S., Shang, W., Kulesza, U., Coelho, R. and Hassan, A.E. (2016). "A Framework for Evaluating the Results of the SZZ Approach for Identifying Bug-Introducing Changes**", *IEEE Transactions on Software Engineering*, 43(7), 641-657.

Deuse, J., Wiegand, M. and Weisner, K. (2019). "Continuous Process Monitoring Through Ensemble-Based Anomaly Detection", *Applications in Statistical Computing*, Editors: Bauer N., Ickstadt, K., Lübke, K., Szepannek, G., Trautmann, H., Vichi, M., Springer, Cham.

Géron, M. (2019). "Hands-On Machine Learning with Scikit-Learn, Keras & Tensorflow", O'Reilly, Canada.

Goodwin, R., Miller, R., Tuv, E., Borisov, A., Janakiram, M. and Louchheim, S. (2004), "Advancements and Applications of Statistical Learning/Data Mining in Semiconductor Manufacturing", *Intel Technology Journal*, 8(4), 325-336.

Günther, F. and Fritsch, S. (2010). "Neuralnet: Training of Neural Networks", *The R journal*, 2(1), 30-38.

Han, J., Kamber, M. and Pei, J. (2012). "Data Mining Concepts and Techniques", Morgan-Knaufmann, USA.

Herrema, F., Curran, R., Hartjes, S., Ellejmi, M., Bancroft, S. and Schultz, M. (2019). "A Machine Learning Model to Predict Runway Exit at Vienna Airport", *Transportation Research Part E: Logistics and Transportation Review*, 131, 329-342.

Jain, A.K., Mao, J. and Mohiuddin, K.M. (1996). "Artificial Neural Networks: A Tutorial", *Computer*, 29(3), 31-44.

Khan, M., Afaq, S.K., Khan, N.U. and Ahmad, S. (2014). "Cycle Time Reduction in Injection Molding Process by Selection of Robust Cooling Channel Design", *International Scholarly Research Notices*, 2014, 1-9.

Köksal, G., Batmaz, İ., Testik, M.C. and Güntürkün, F. (2010). "İmalat Sektöründe Kalite İyileştirmede Veri Madenciliği Tekniklerinin Kullanımı", *Verimlilik Dergisi*, (2), 47-65.

Kolberg, D. and Zühlke, D. (2015). "Lean Automation Enabled by Industry 4.0 Technologies", *IFAC-PapersOnLine*, 48(3), 1870-1875.

Kozjek, D., Kralj, D., Butala, P. and Lavrač, N. (2019). "Data Mining for Fault Diagnostics: A Case for Plastic Injection Molding", *Procedia CIRP*, 81, 809-814.

Leys, C., Ley, C., Klein, O., Bernard, P. and Licata, L. (2013). "Detecting Outliers: Do Not Use Standard Deviation Around the Mean, Use Absolute Deviation Around the Median", *Journal of Experimental Social Psychology*, 49(4), 764-766.

Lieber, D., Stolpe, M., Konrad, B., Deuse, J. and Morik, K. (2013). "Quality Prediction in Interlinked Manufacturing Processes Based on Supervised & Unsupervised Machine Learning", *Procedia CIRP*, 7, 193-198.

Little, J.D.C. (1961). "A Proof of the Queueing Formula: L=λW", *Operations Research*, 9, 383-387.

Marti-Puig, P., Blanco-M, A., Cárdenas, J.J., Cusidó, J. and Solé-Casals, J. (2018). "Effects of the Pre-Processing Algorithms in Fault Diagnosis of Wind Turbines", *Environmental Modelling & Software*, 110, 119-128.

Meidan, Y., Lerner, B., Rabinowitz, G. and Hassoun, M. (2011). "Cycle-Time Key Factor Identification and Prediction in Semiconductor Manufacturing Using Machine Learning and Data Mining", *IEEE transactions on semiconductor manufacturing*, 24(2), 237-248.

Mrugalska, B. and Ahram, T. (2017). "Managing Variations in Process Control: An Overview of Sources and Degradation Methods", *Advances in Ergonomics Modeling, Usability & Special Populations*, 377-387.

Muhammed, T. and Shaikh, R.A. (2017). "An Analysis of Fault Detection Strategies in Wireless Sensor Networks", *Journal of Network and Computer Applications*, 78, 267-287.

Noorzaei, J., Hakim, S.J.S., Jaafar, M.S. and Thanoon, W.A.M. (2007). "Development of Artificial Neural Networks for Predicting Concrete Compressive Strength", *International Journal of Engineering and Technology*, 4(2), 141-153.

Polato, M., Sperduti, A., Burattin, A. and De Leoni, M. (2014). "Data-Aware Remaining Time Prediction of Business Process Instances", *2014 International Joint Conference on Neural Networks (IJCNN 2014)*, Beijing, China, 816-823.

Quintana, G., Garcia-Romeu, M.L. and Ciurana, J. (2011). "Surface Roughness Monitoring Application Based on Artificial Neural Networks for Ball-End Milling Operations", *Journal of Intelligent Manufacturing*, 22(4), 607-617.

Rafiq, M.Y., Bugmann, G. and Easterbrook, D.J. (2001). "Neural Network Design for Engineering Applications", *Computers & Structures*, 79(17), 1541-1552.

Ramkumar, P.L., Ramesh, A., Alvenkar, P.P. and Patel, N. (2015). "Prediction of Heating Cycle Time in Rotational Moulding", *Materials Today: Proceedings*, 2(4-5), 3212-3219.

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P. and Harnisch, M. (2015). "Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries", Boston Consulting Group, 9(1), 54-89.

Rust, K. (2008). "Using Little's Law to Estimate Cycle Time and Cost", *IEEE Winter Simulation Conference (IEEE WSC 2008)*, Florida, USA, December 7–10: 2223-2228.

Shrouf, F., Ordieres, J. and Miragliotta, G. (2014). "Smart Factories in Industry 4.0: A Review of the Concept and of Energy Management Approached in Production Based on the Internet of Things Paradigm", *21st IEEE International Conference on Industrial Engineering & Engineering Management (IEEM 2014)*, China, 697-701.

Siller, H., Rodriguez, C.A. and Ahuett, H. (2006). "Cycle Time Prediction in High-Speed Milling Operations for Sculptured Surface Finishing", *Journal of Materials Processing Technology*, 174(1-3), 355–362.

Singh, A., Thakur, N. and Sharma, A. (2016). "A Review of Supervised Machine Learning Algorithms", *3rd International Conference on Computing for Sustainable Global Development*, March 16-18, 1310-1315.

Sumathi, S. and Sivanandam, S.N. (2006). "Introduction to Data Mining and Its Applications", Springer, Berlin.

Van Der Aalst, W. (2011). "Process Mining: Discovery, Conformance and Enhancement of Business Processes", Springer, Heidelberg.

Wang, G., Ledwoch, A., Hasani, R. M., Grosu, R. and Brintrup, A. (2019). "A Generative Neural Network Model for the Quality Prediction of Work in Progress Products", *Applied Soft Computing*, 85, 105683.

Wang, J. and Zhang, J. (2016). "Big Data Analytics for Forecasting Cycle Time in Semiconductor Wafer Fabrication System", *International Journal of Production Research*, 54(23), 7231-7244.

Yarlagadda, P.K. and Khong, C.A.T. (2001). "Development of a Hybrid Neural Network System for Prediction of Process Parameters in Injection Moulding", *Journal of Materials Processing Technology*, 118(1-3), 109-115.

Yu, J.B. and Xi, L.F. (2009). "A Neural Network Ensemble-Based Model for On-Line Monitoring and Diagnosis of Out-of-Control Signals in Multivariate Manufacturing Processes", *Expert systems with Applications*, 36(1), 909-921.

Zhang, G., Patuwo, B.E. and Hu, M.Y. (1998). "Forecasting with Artificial Neural Networks: The State of the Art", *International Journal of Forecasting*, 14(1), 35-62.

Zhou, H. (2013). "Computer Modeling for Injection Molding: Simulation, Optimization, and Control", John Wiley & Sons, Singapore.