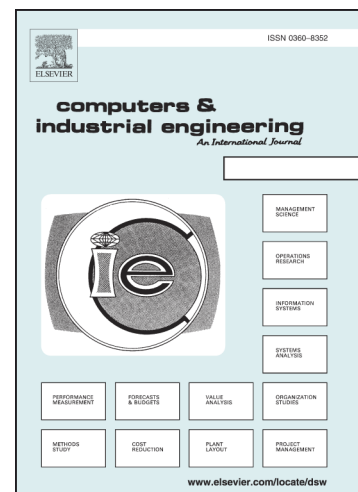# Accepted Manuscript

A data-driven algorithm to predict throughput bottlenecks in a production system based on active periods of the machines

Mukund Subramaniyan, Anders Skoogh, Hans Salomonsson, Pramod Bangalore, Jon Bokrantz

A data-driven algorithm to predict throughput bottlenecks in a production system based on active periods of the machines

Mukund Subramaniyan[a*], Anders Skoogh[a], Hans Salomonsson[b], Pramod Bangalore[b], Jon Bokrantz[a]

a. Department of Industrial and Materials Science, Chalmers University of Technology, Gothenburg – 41296, Sweden

b. Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg – 41296, Sweden

* Corresponding Author

Mukund Subramaniyan*

Department of Industrial and Materials Science
Chalmers University of Technology
Gothenburg – 41296, Sweden
Email: mukunds@chalmers.se
Phone: +46728500693

Anders Skoogh

Department of Industrial and Materials Science
Chalmers University of Technology
Gothenburg – 41296, Sweden
Email: anders.skoogh@chalmers.se
Phone: +46317724806

Hans Salomonsson

Department of Computer Science and Engineering
Chalmers University of Technology
Gothenburg – 41296, Sweden
Email: hans.salomonsson@chalmers.se
Phone: +46317728073

Pramod Bangalore

Department of Computer Science and Engineering
Chalmers University of Technology
Gothenburg – 41296, Sweden
Email: pramod.bangalore@chalmers.se
Phone: +46317721650

Jon Bokrantz

Department of Industrial and Materials Science
Chalmers University of Technology
Gothenburg – 41296, Sweden
Email: jon.bokrantz@chalmers.se
Phone: +46317723614

* Corresponding Author

Abstract

Smart manufacturing is reshaping the manufacturing industries by boosting the integration of information and communication technologies and manufacturing. As a result, manufacturing companies generate large volumes of machine data which can be potentially used to make data-driven operational decisions using informative computerized algorithms. In the manufacturing domain, it is well-known that the productivity of a production line is constrained by throughput bottlenecks. The operational dynamics of the production system causes the bottlenecks to shift among the production resources between the production runs. Therefore, prediction of the throughput bottlenecks of future production runs allows the production and maintenance engineers to proactively plan for resources to effectively manage the bottlenecks and achieve higher throughput. **This paper proposes an active period based data-driven algorithm to predict throughput bottlenecks in the production system for the future production run from the large sets of machine data. To facilitate the prediction, we employ an auto-regressive integrated moving average (ARIMA) method to predict the active periods of the machine. The novelty of the work is the integration of ARIMA methodology with data-driven active period technique to develop a bottleneck prediction algorithm.** The proposed prediction algorithm is tested on real-world production data from an automotive production line. The bottleneck prediction algorithm is evaluated by treating it as a binary classifier problem and adapted the appropriate evaluation metrics. Furthermore, an attempt is made to determine the amount of past data needed for better forecasting the active periods.

Keywords: Digitalisation; Bottleneck prediction; Big Data; Industry 4.0; Maintenance; Predictive analytics

1. Introduction

Digital Manufacturing and Industrial Internet of Things (IIoT) are new emerging technologies to increase the productivity in manufacturing (Lee, Lapira, Bagheri, & Kao, 2013). Manufacturing companies collect shop floor data in digital format using Manufacturing Execution Systems (MES), sensor technologies etc. (Hedman, Subramaniyan, & Almström, 2016). For example, one of the automotive manufacturing company in Sweden collects 100 data points of machine data per hour by MES (Subramaniyan, 2015). This means that, on 8-hour operating shift, 800 data points are collected per machine. This voluminous data at a high velocity when scaled up to a production system level or a factory level can be referred to as big data (Lee et al., 2013). With the exponential growth in the data acquired from the machines, new opportunities emerge to leverage data science to enhance the state of manufacturing and enable more data-driven decision making (Shao, Shin, & Jain, 2015). To enable such data-driven decision making, companies need high informative analytical algorithms to turn high volumes of fast-moving data into meaningful insights (Lavalle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011)(Bokrantz, Skoogh, Berlin, & Stahre, 2017). This necessitates research into data analytics which can enable efficient and effective extraction of information from the raw data to derive new knowledge and insights, which can further be applied to introducing intelligence into the control of production processes and can also improve the system-level operation of manufacturing enterprises (Wuest et al., 2016).

In the manufacturing domain, throughput is an important indicator used to evaluate the production system performance. The throughput of a production line depends on the individual throughput of the machine (Chang, Ni, Bandyopadhyay, Biller, & Xiao, 2007). It is often constrained by one or more machines in the production system, which are usually called bottlenecks (Goldrat & Cox, 1990). The prioritization of maintenance and other production improvement activities on bottleneck machines

results in substantial improvement in the overall throughput of the production system (Gopalakrishnan, Skoogh, & Christoph, 2013)(Wedel, Noessler, & Metternich, 2016). Therefore, the bottleneck machines are critical to the production system performance, especially when there are limited resources.

There are a number of research efforts focussed on developing data-driven algorithms that uses digital historical machine data collected by MES to detect the bottlenecks at a system level. For example, Li, Chang, & Ni (2009) proposed an algorithm that uses the online blockage and starvation data of the machines to detect the bottlenecks, while Subramaniyan, Skoogh, Gopalakrishnan, & Hanna (2016) proposed an MES based algorithm that uses the active states of the machine captured by MES to detect the bottlenecks. These algorithms are descriptive in nature, meaning they detect the past bottlenecks in the production system. The bottlenecks in the production line are dynamic in nature and the current bottlenecks may not be the bottlenecks for next period as the machine's behaviour changes. This is due to the system performance degradation caused by random noise and disturbances such as failure etc. (L. Li, Qing, Xiao, & Ambani, 2011). This change causes the bottlenecks to shift and in turn affects the overall throughput of the production system. To maximise the throughput from the production system, it is essential to predict the bottleneck shifts in advance of the production run so that those machines can be prioritized for maintenance (e.g. reactive and proactive maintenance strategies on bottleneck machines) and other improvement activities. Hence a system level decision support tool is required to indicate the bottleneck machines of the future production run (Jin, Weiss, Siegel, & Lee, 2016). The requirement of such a system level decision support for maintenance planning in digitalized manufacturing environment was also identified by (Bokrantz et al., 2017) in a Delphi based scenario study of the future of maintenance by 2030. Further, Bokrantz et al. (2017) indicated that the real-time data analytics will be used as a potential tool by the production and maintenance engineers to make decisions on a system level.

In the past, there have been few research efforts on predicting the future bottlenecks of the production system. Li et al (2011) proposed a time-series based predictive algorithm that predicts the blockage and the starvation of the machines and thereby the bottlenecks using a turning point method. Cao, Deng, Liu, & Wang (2012) proposed an algorithm to predict the bottlenecks in semiconductor manufacturing industries using different variables including the product types, releasing strategies, work in progress, processing times, utilization rate, and buffer length. Wedel et al. (2016) proposed an algorithm that predicts the bottlenecks based on the buffer levels. Though there are various predictive algorithms, these algorithms are developed and tested in a simulation environment using a discrete event simulation model of the production system and are limited to the real-world validation of the algorithms using the online data collected from the machines on the shop floor. Moreover, the performance of the different predictive algorithms is not assessed with the standard algorithms like naïve method (Keogh & Kasetty, 2002).

The purpose of the paper is to increase the productivity of the manufacturing systems by facilitating data-driven decision making using the real-time MES data. The aim is to propose a data-driven algorithm to predict the throughput bottlenecks in a production system using the active periods of the machines and to evaluate the prediction algorithm. The active period method of bottleneck detection, is gaining more popularity within manufacturing research, as they can detect a group of potential bottlenecks in the production system and can also be used in decoupled and coupled systems (Lima, Chwif, & Barreto, 2008; (Roser & Nakano, 2015);(Subramaniyan, Skoogh, Gopalakrishnan, Salomonsson, et al., 2016). However, there is no reported research exploring the ways by which the active period method can be used to predict the bottlenecks.

2. Literature Review

In this section, the past research efforts on bottleneck prediction are presented followed by time series analysis and the evaluation metrics used to evaluate the prediction algorithms.

2.1. Previous work on bottleneck prediction in manufacturing system

An exhaustive list on the literature of prediction of bottlenecks in manufacturing is presented in Table 1.

According to the scientific publications presented in Table 1, it can be understood that there are many unique predictive algorithms to predict the bottlenecks in the production system. But these algorithms are tested in a simulation environment i.e. using a discrete event simulation model of the production system. In a simulation environment, Li et al (2011) achieve a prediction accuracy 97.38% in predicting the blockage and the starvation periods of the machines while ANFIS algorithm prediction proposed by Cao et al (2012) achieves an accuracy of 92.03% in predicting the main bottlenecks. While the simulation is considered to be a vital development tool in the production analysis, the input data to the simulation follows well-defined distributions and therefore, using simulation to predict the bottlenecks can yield higher accuracy (Amar & Gupta, 1986). Compared to the simulation environment, the real-time data from the machines do not fit any standard distributions due to the random nature of production disturbances. Moreover, the simulation models might not truly represent the real-world dynamics of the production as it is difficult to update the models with the changes made in the production system and this is time-consuming (Skoogh, 2011). Therefore, in order to claim the performance of the predictive algorithms, these need to be validated on real-world datasets in order for them to be useful in the real-world operations (Amar & Gupta, 1986). **Validating the algorithms on the real-world production lines also enables the manufacturing companies to trust the algorithm and adopt the algorithm (Jin et al., 2016), thus increasing the use of scientific knowledge in the industries. The need to validate the algorithms on the real-world data sets was also emphasised by (Liao, Deschamps, Freitas, & Loures, 2017) so as to reduce the widening gap between the laboratory environments and the industrial applications.** The predictive algorithms presented in Table 1 lack the real-world validation and moreover, there are no clear explanations as to whether the data used to construct the simulation model was from a real-world production line. Also, these algorithms are not compared with the standard predictive algorithms such as naïve and therefore, it is not possible to evaluate the performance of those algorithms compared to standard algorithms (Keogh & Kasetty, 2002)(Hyndman, 2014).

It can be seen from Table 1, that there are different bottleneck detection techniques used to predict the bottlenecks. When compared to turning point method and buffer levels methods, the active period method can detect the primary and secondary bottlenecks more accurately and can be used for coupled and decoupled systems (Roser & Nakano, 2015). The turning point method is not accurate in detecting the bottlenecks in a decoupled system especially when there are minor production disturbances in the machine that doesn't necessarily cause the upstream machines to be blocked or downstream machines to be starved. The buffer method to detect the bottlenecks could be misleading especially when the buffers are shared between the machines or when the parts transfer devices fail or when the transfer time is more (J. Li et al., 2009).

Also, it can be seen from Table 1 that, there is no method found in the literature for predicting the bottlenecks using the active period based method. Therefore, this was the motivation to derive a new bottleneck prediction method based on active periods of the machines.

2.2. Theory on Active period method of bottleneck detection

In this method, the machine is in the active state when the machine is engaged in producing a product, when under breakdown, during set up, and so on, and considered to be in inactive when the machine is blocked or starved (Roser, Nakano, & Tanaka, 2001) as seen in Figure 1. **The active and inactive states are similar to equipment dependent and equipment independent states respectively as proposed by (Ron & Rooda, 2005). Equipment dependent states as the states are the machine take when the machine is in a productive state, unscheduled downstate or scheduled downstate. The equipment independent states include lack of input to the equipment, blocked state etc.** By computing the percentage of the time, the machine is active during the scheduled production time and comparing it with other machines, a group of potential bottlenecks can be identified. This method also gives an opportunity to obtain diagnostic insights as these are based on the aggregation of the different machine states to better understand the nature of bottleneck machines. Moreover, this method can be used to detect bottlenecks in different types of the production system, from job shop to flow systems with or without finite buffers (Roser & Nakano, 2015). Furthermore, the active period periods of a machine can be converted into a time series data with time resolution equal to months, weeks, days, shifts, etc. From such a time series data, an established time series based prediction methods can be applied to estimate the future active periods of the machines, and thereby predict the bottlenecks in the production line.

2.3. Time series analysis using Auto Regressive Integrated Moving Average (ARIMA)

Time series is defined as a sequence of data points collected at a constant time interval. These time series data are analyzed to determine the long-term trends to forecast the future. The aim of the time series algorithm is to explore and derive hidden patterns and insights that help in making decisions (Brockwell & Richard, 2016). The relationship between the observed sample values and the underlying stochastic process is analogous to the relationship between sample and population in hypothesis testing (Keskin, Taylan, & Terzi, 2006). Therefore, the time series is a sample from the underlying stochastic process that generated the series. **There are different time series forecasting methods available in the literature and those methods can be grouped as statistical based and artificial intelligence based (Wang, Wang, Zhang, & Guo, 2012). As the active period percentages can be expressed as a time series when collected for different production runs and they are dependent on their own historical data, an ARIMA statistical method in the form of stochastic linear difference equations can explain the dependency in the active period data of the machines. ARIMA is also proved to be a robust and an efficient method for short-term time series forecasting, especially one-step ahead forecasts (Kuvulmaz, Usanmaz, & Engin, 2005). Moreover, ARIMA has been extensively used in the field of finance for short-term prediction of stock prices etc. (Wang et al., 2012) (Adebiyi A. & Adewumi O., 2014).**

ARIMA model can be represented as ARIMA(p,d,q) where p,d,q are non-negative integers that represents the order of autoregressive (AR), integrated (I), and moving average(MA) parts of the model (Ho & Xie, 1998). In practice, most of the time series data are non-stationary. Therefore, Auto-regressive (AR), moving average (MA) or auto-regressive moving average (ARMA) cannot be applied directly. One way to convert the data into stationary data is by applying differencing. If the original data is differenced d times before fitting the ARMA model, then the model for the original time series data is called as ARIMA model, where (I) denotes the number of differencing operations (Ho & Xie, 1998). The ARIMA models not only provide the forecasted value but also other metrics related to the forecasted values including standard errors associated with the forecast, confidence and prediction intervals. To compare the performance of the ARIMA model, it can be compared to a

benchmark forecasting technique such as the naïve method, which is a simple technique in which the forecast is the same as the last observation (Hyndman, 2014).

2.4. Algorithm evaluation metrics

To evaluate the performance of the predictive algorithm with that of the naïve algorithm, different metrics are proposed in the literature by different authors. The most commonly used metrics out of the different metrics are Mean Absolute Percentage Error (MAPE) (Hyndman, 2014), accuracy, precision, recall which are metrics derived from confusion matrix (Bradley, 1997). More recently, a new informative metric, Intersection over Union (IoU) was proposed by Microsoft to evaluate the prediction algorithms (Ahmed, Tarlow, & Batra, 2015).

2.4.1. Mean Absolute Percentage Error (MAPE)

MAPE is a measure of forecast accuracy, in that it compares the forecasted value to that of the actual value and is widely used in time series forecasting (Hyndman, 2014). This MAPE is a scale-independent statistic and expresses the prediction error as a percentage. For the series of forecasted values $(\hat{F}_1, \hat{F}_2, \hat{F}_3, \ldots, \hat{F}_n)$ and the corresponding series of actual values $F_1, F_2, F_{3,\ldots}, F_n$, MAPE can be calculated according to eq.(1).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} |\left(F_i - \hat{F}_i\right) / F_i| \tag{1}$$

2.4.2. Confusion Matrix Metrics

The confusion matrix summarises the performance of the classification prediction algorithm (Bradley, 1997). This matrix is a summary of results got from a prediction algorithm on a classification problem type, e.g. a binary classification problem, and is a representation of the number of correct and incorrect predictions results. Table 2 shows the confusion matrix. The metrics that can be derived from a confusion matrix are accuracy, precision, and recall as shown in eq.(2), eq.(3) and eq.(4). These metrics give a way of assessing how the predicted responses obtained from the algorithm align with the actual responses.

The following are the metrics that can be derived from the confusion matrix:

$$Accuracy = \frac{TP+TN}{Total\ Population} \tag{2}$$

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

2.4.3. Intersection over Union (IoU)

IoU is a metric developed by Microsoft and is widely used for the evaluation of the image segmentation algorithms in the computer vision domain (Ahmed et al., 2015). This metric can be used to evaluate the multi-classifier predictive algorithms and compares the degree of overlap between the results obtained from the prediction algorithm with that of the actual objects as shown in Figure 2. In other words, IoU is the ratio between the intersection of the predicted and the actual objects to the union of the predicted and actual objects. The goal of this metric is to quantify the dissimilarities of the proposed algorithmic solution with respect to the ground truth in a meaningful way (Ahmed et al., 2015). The greater the overlapping area, the greater the algorithm performance.

3. Methodology

The methodology adopted was based on the Cross-Industry Standard for Data Mining (CRISP-DM) (Pete et al., 2000). It provides a structured methodology for a data mining project and is widely used in manufacturing (Gröger, Niedermann, & Mitschang, 2012). The CRISP-DM was adapted to design a MES based data-driven predictive algorithm. The methodology can be broadly classified into two categories: the algorithm development phase and testing phase.

3.1. Algorithm development Phase

The algorithm development phase includes literature study, the study of the MES data collected from a real-world production line and the interaction with the production domain experts to understand the operational dynamics of the production system in a detailed manner. The theory of active period based bottleneck detection method as proposed by Roser et al (2001) using a discrete event simulation model of the production system was studied in detail. In addition to that, the step by step construction of the ARIMA based time series prediction algorithm and the evaluation metrics of prediction algorithm was also studied. **ARIMA algorithm was chosen out of other time series algorithms because it is proved to work for short time series data and it can capture the trends found on the historical data and project future values (Kuvulmaz et al., 2005). Moreover, it offers the rolling window flexibility which is an enabler to estimate the length of the historical data, especially when the data is an aggregated data such as active periods (Hyndman & Yeasmin, 2008)(Vafaeipour, Rahbari, Rosen, Fazelpour, & Ansarirad, 2014). Hence ARIMA based algorithms can also be suited to data in constrained data storage environments.** A real-world MES data set from a production line as exemplified in Table 3 is also studied in detail. **This is done to determine the relevant data with respect to the active period method within the large set of MES data.** The insights gained from the MES data in combination with the literature studies and production and maintenance domain expert inputs were used to design and develop the prediction algorithm.

3.2. Testing phase

In this phase, the algorithm is tested on a real-world industrial test data of a production line. The algorithm is applied to the dataset and the active periods for future production runs are forecasted. **The forecast accuracy, MAPE as explained in section 2.4 is then calculated to evaluate the performance of the algorithm in terms of forecasting the future active periods of the machine against the naïve method. MAPE is chosen because it is a better metric to express and communicate the performance to the industrial community as it is based on percentages and percentages are easily understandable and interpretable. Moreover, MAPE being a scale-independent metric, it enables the opportunity to compare the active period time series forecasting model performance with different time scales. Thus, enabling benchmarking of different algorithms using the same metric.** From the forecasted active periods of the machine, group of bottlenecks is predicted. The bottleneck prediction algorithm is then evaluated based on the different measures such as accuracy, precision and recall and IoU, and then compared with that of the naïve approach. The results were then discussed with the production and maintenance domain experts. The modeling of the industrial data and the evaluation of the algorithm was carried out by uploading the raw MES data into R software, which is a powerful language used widely for data analytics and modeling (Hyndman & Khandakar, 2008).

3.3. Real-world industrial test study description

The algorithm was tested on the MES data of an automotive engine production line in Sweden. The layout of the production line is shown in Figure 3. The production line has 13 machines from M1, M2…., M13. Each machine has an ANDON light and the MES collects and stores the ANDON

information of the machine across a production run as shown in Table 3. There are four different ANDON lights: Red, Yellow, Green, and White. At any given time, the machine may have one light or a combination of lights. These ANDON lights can be the group to represent the machine states as shown in Table 4.

The MES system stores not more than 315 production runs data at any point in time. Each production run is of 17 hours' duration. To understand the degree of the shiftiness in the bottlenecks between the production runs, the MES data for 215 consecutive production runs were analyzed. The active period method as explained in Section 2.2 was used to identify the bottleneck machines for each production run. The bottleneck machines of each production run were checked against the bottleneck machines of the previous production run, which is the naïve method. The probability of the previous production run bottleneck machine being the true bottleneck for the next production run is 20.8%. This indicates that the bottleneck shifts between the production runs. Therefore, it is necessary to predict the bottlenecks of the future production run so that the engineers of the production system can plan and allocate the production and maintenance resources to effectively manage the bottleneck machines. In addition to that, as the MES stores the data no more than 315 production days **at any point in time**, it is desirable to find the size of the historical data that should be used to provide a better forecast for the next production run considering the dynamics of the production system.

4. Proposed Bottleneck Prediction Algorithm based on Active Period Method

The active period based bottleneck prediction algorithm consists of three steps. (1) To compute the active periods of the machine from MES data and then convert it into a time series of desired time resolution of production runs (shifts, days, weeks etc.). (2) To use time series techniques to forecast the active periods of the machines for the future production runs. (3) To detect the bottleneck machines in the production run from the corresponding forecasted active periods of the machines. Figure 4 explains these three steps. The three steps put together is called as bottleneck prediction algorithm in this paper. This algorithm takes the machine level MES data to predict the bottlenecks at a system level.

Step 1: Time series generation of the active periods of the machine

Let the elapsed time of each active state be a_{jmn}, where $j \in \{1,\ldots,J\}$ is an active state, and $n \in \{1,\ldots,N\}$ is the production run day and $m \in \{1,\ldots,M\}$ is the machine index. Let b_n be the scheduled production hours. The mean active period percentage for machine m can then be calculated as shown in eq.(5).

$$A_{\_mn} = \left(\frac{\sum_{j=1}^{J} a_{\_jmn}}{b_n}\right) \tag{5}$$

For a time, horizon T, the historical data of the active period percentages of a machine can be divided into N periods as shown in Figure 5 to form a time series. The assumption made during the construction is that the active period percentages are dependent on its own historical data i.e. A_1 depends on A_2 etc. (each machine active period percentages during a time t reflects the system dynamics). The active period percentages are calculated for all the machines in the production system

Step 2: Time series based forecasting of the active periods of the machines

This step can be further divided into two steps: the first step is to determine the size of the historical data that is required to forecast the active periods of the machine and the second step is to forecast the active periods of the machines. The first step is carried out only when designing the algorithm to fix

the size of historical data. Once the size of the historical data is determined it can be used to forecast the active periods in the second step without changing the size of the historical data.

Step 2.1: Determination of the size of historical data

An algorithm of an ARIMA process is identified that represents the past and future patterns of a time series. As the dynamics of the production in the shop floor often change due to the continuous improvement activities on the machine, it may not be valid to assume that the ARIMA model parameters (p, d, and q) are constant over time. Also, if the MES stores data only for limited production runs, it is necessary to find the amount of past data that can be used to provide better forecast the active periods. With the one-step-ahead rolling ARIMA with fixed window size, the model parameters and regression coefficients are computed for each step forecast to capture the dynamics of the production system after the improvement activities are performed on the machines. When rolling forecast has a fixed number *k* of the previous observed values as inputs, then it is called a sliding window technique (Vafaeipour et al., 2014) as presented in Figure 6.

The next step is to determine the optimum size of *k*. For this, the data set is divided into two sets: training and testing data set. Training data set is used to train the algorithm and the testing data set is used to test how developed the algorithm has been trained and to estimate the algorithm properties (Hyndman, 2014). For a given sliding window of size *k* and given the active periods of the machine A_t, A_{t-1}, A_{t-2}..., A_k at time t, the active duration needs to be predicted as the A_{t+1}. ARIMA model is applied over the historical data of active periods. Box Jenkins methodology of ARIMA application includes three iterative steps, algorithm identification, parameter estimation and diagnostic checking (Brockwell & Richard, 2016). This three-step algorithm building is typically repeated several times until a satisfactory algorithm is selected. More detailed statistics involved in the three steps is explained in (Brockwell & Richard, 2016). The three-steps are carried out automatically using the time series libraries in R (Hyndman & Khandakar, 2008). The ARIMA algorithm procedure is applied over the historical time series data of active periods of the machine to estimate the algorithm parameters and the coefficient which is represented in eq.(6).

$$A_{t+1} = \phi_0 + \phi_1 A_t + \phi_2 A_{t-1} + \phi_3 A_{t-2} + \cdots + \varepsilon_t - \theta_1 \varepsilon_t - \theta_2 \varepsilon_{t-1} - \theta_3 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \qquad (6)$$

Where, A_{t+1} represents the forecasted active duration of the next production run, at ε_t is the random variable for the disturbance and they are independent and identical distributed, $\varepsilon_t \sim N(0, \sigma^2)$ and ϕ, θ are the regression co-efficient of the auto-regressive and moving average parts of ARIMA model respectively. Eq. (6) represents the ARMA equation after the series of data is made stationary. ARIMA models with different sliding window sizes can be tested on the test data set and can be compared against each other to select the best size *k* of the sliding window. MAPE metric as explained in Section 2.4.1 can be used to compare the forecasted and actual active periods in different sliding windows sizes on the test data set. The sliding window size which has a minimum MAPE value is the optimum window size that can then be used as algorithm to forecast the future values of the active duration. The same procedure is followed for all the machines in the production line to find the optimum window size that is needed to forecast the active duration for next production run.

Step 2.2: Forecasting of the active periods

Once the sliding window size is determined, the same window size is used to forecast the future active periods of the machines and the size of the sliding window need not be determined every time a new forecast is made. The output from the ARIMA model is the forecasted value of the active period for the t+1 period and the associated standard error of the forecast.

Step 3: Bottlenecks prediction from forecasted active periods of the machines

The forecasted active duration of the machines and the standard error associated with the forecast of the machines in the production line are compared against each other to identify the group of bottlenecks. Let K be the index of the machine with highest active period percentage i.e. K=argmax $(A_{m, t+1})$, m ∈ {1,…,M}. To test the statistical significance of differences in the predicted active period for the all the machines with respect to the highest active period machine, two-tailed t-tests as shown in eq.(7) is performed for all machines m ∈ {1,..,M}\ K.

$$t_{stat} = (A_K - A_X)/\sqrt{SE_K{}^2 + SE_X{}^2} \tag{7}$$

where A_K is the mean active period percentage of machine K and A_X is the mean active period percentage of machine X. The estimated standard errors for machines K and X are represented by SE_K and SE_X, respectively.

The difference in the mean active percentage periods of the two machines is statistically significant if $t_{stat} > 1.96$ or $t_{stat} < -1.96$ at 95% confidence level (Knezevic, 2008) and hence only machine K will be the bottleneck machine. However, for $-1.96 \leq t_{stat} \leq 1.96$ at 95% confidence level the mean difference is not statistically significant and hence the machines K and X will be judged as bottlenecks in the system. The rest of the machines are classified as non-bottleneck machines.

5. Proposed Measures for Algorithm Performance Evaluation

The bottlenecks prediction algorithm can be treated as a binary classification algorithm that classifies each machine into bottleneck and non-bottleneck. Therefore, to evaluate the algorithm performance, the metrics derived from confusion matrix and IoU as explained in Section 2.4.2 and 2.4.3 respectively can be used. The adaptation of the metrics to the bottleneck prediction problem in manufacturing is explained in this section.

5.1. Metrics from Confusion Matrix

The confusion matrix for the bottleneck prediction algorithm can be constructed as shown Table 5.

The classification leads to the calculation of the metrics: accuracy, precision and recall metrics as explained in Section 2.4.2

- Accuracy is the measure how often the proposed algorithm predicts the correct bottlenecks and the non-bottlenecks.
- Precision gives the proportion of the machines that were predicted as the bottleneck machines were the actual bottlenecks.
- Recall is the proportion of machines that were the actual bottleneck machines and was also predicted by the algorithm correctly.

5.2. Intersection Over Union Method (IoU Method)

Let the set of predicted groups of bottlenecks in the production system be represented as set P and the actual bottleneck set be represented as A such as P ⊂ {M1, M2, M3...} and A ⊂ {M1, M2, M3…}, where M1, M2, M3… ∈ M are the machines in the production line. Let set Z represent the intersection of set P and set A as shown in eq.(8), in that it contains the bottleneck machines that were predicted and was the actual bottlenecks. Let set X represent the union of set P and set A as shown in eq.(9) , in that it contains the total number of distinct bottleneck machines in the predicted bottleneck set and the actual bottleneck set.

$$Z = |P \cap A| \tag{8}$$

$$X = |\ P \cup A|\ \tag{9}$$

IoU determines what percentage is the predicted group of bottlenecks is the actual number of bottlenecks in the production line i.e. it is the ratio of the number of true bottlenecks that the algorithm finds to the sum of a total number of bottlenecks that are found and the number of false bottlenecks that the algorithm classified. The IoU for a production run can be calculated as shown in eq.(10).

$$IoU = \frac{Z}{X}\ X\ 100$$
(10)

This metric penalizes the algorithm performance for prediction of wrong bottlenecks in the production system. The higher the IoU metric, the better is the prediction algorithm performance. This metric can also be used for comparing different prediction algorithms including naïve and will be helpful for making decisions (e.g. trade-offs) in selecting the suitable algorithm for implementation. **Even though IoU and accuracy metric from confusion matrix looks similar, there is a distinction between the two due to the reason that IoU is not influenced by the true negatives.**

5.3. Comparison of the performance metrics of the proposed algorithm and naïve algorithm

If there are *n* production runs for which metrics accuracy, precision, recall and IoU are calculated, then the following formula can be applied to get the mean and standard error for each of the metrics for the proposed algorithm as shown in eq.(11) and eq.(12).

$$Mean\ Y = \frac{\sum_{i=1}^{n} Y}{n} \times 100; Y = accuracy, precision, recall, IoU$$
(11)

$$Standard\ error\ (SE)\ Y = \frac{Standard\ deviation\ Y}{\sqrt{n}} \tag{12}$$

The above metrics of the algorithm can be compared with that of the naïve algorithm to assess the performance of the proposed algorithm. The two-tailed t-tests as proposed by Knezevic (2008) can be used to test the statistical difference in the mean values of the proposed algorithm with that of the naïve as shown in eq.(13).

$$t_{stat} = (P_Y - N_Y)/\sqrt{SE_{PY}^2 + SE_{NY}^2} \tag{13}$$

Where P_Y and SE_{PY} is mean value and the standard error of the metric for proposed algorithm, N_Y and SE_{NY} is the mean value and the standard error of the metric for the naïve algorithm. If $t_{stat} > 1.96$ or $t_{stat} < -1.96$ at 95% confidence level then there is a statistical difference in the performance of the proposed algorithm with that of the naïve algorithm.

6. Industrial Test Case Results

The description of the Industrial test is shown Section 3.3. The MES data was collected for a period of 315 production runs from 13 machines. It is assumed that no structural change happened in the production system during the 315 production runs, meaning that there is no change in the production flow.

6.1. Time series generation of the active period percentages of the machine

The ANDON light combinations as shown in Table 3 are grouped into three states; Producing, Blocked/Starved and Down, in detailed discussions with the domain experts of the production line including the production and maintenance engineers. The Producing and the Down states of the

machine constitutes the active period of the machine. Thereafter, the active period of each machine is calculated for each of the 315 production runs to form a time series data according to the eq.(5) explained in step 1 in Section 4. The example of the active period time series data of machine M1 is shown in Table 6.

The data set is thereafter divided into train and test datasets. The training data set consists of the first 185 production run data and the test data set consists of 130 production run data of active periods for each machine. The reason to have a higher amount of test data is to evaluate the prediction performance over many production runs.

6.2. Forecasting of the active periods

The first step is to determine the sliding window size of the ARIMA model that can be used to forecast the future active periods of the machines as explained in step 2.1 in Section 4. Thereafter, the future active periods of the machines are forecasted using the same sliding window size of ARIMA model explained in Step 2.2 in Section 4.

6.2.1. Determination of the size of historical data

To estimate how much of the past data can be used for constructing an appropriate forecasting algorithm which can provide a better forecast, one-step prediction sliding ARIMA as explained in Section 4 with four different fixed window sizes was carried out. The first window consists of 50 data points (production run 1 to 50), the second consists of 100 data points (production run 1 to 100) , the third consists of 150 data points (production run 1 to 150) and the fourth set of data consists of 185 (production run 1 to 185) data points. The MAPE values from the sliding window sized ARIMA models tested on the testing dataset (so as the number of computation data points are same across periods) using different window sizes for all machines is presented in Table 7.

The most notable feature of MAPE as presented in Table 7 is that the values are almost robust among the different window sizes. This indicates there is a strong randomness associated with the active period percentages of the machines and it is difficult to forecast the randomness using only the past active period percentages of the machine data drawn from MES data with very low error. From Table 7, MAPE for one-step sliding ARIMA of 100 window size is the better out of the different window sizes except for the machines M4, M5, and M11. The MAPE for window size 50 is only slightly better than window size 100 by 0.15%, 0.14% and 0.02% for machines M4, M5, and M11 respectively. Also, the average MAPE for the production line is less for a window size of 100. A trade-off was made between the MAPE values of 50 and 100 window size for M4, M5, and M11 to have a homogenous forecasting algorithm across the machines in the production system. Therefore, sliding ARIMA algorithm with a window size of 100 is selected for all the machines in the production system.

6.2.3. Forecasting of the active periods

The sliding ARIMA model with a window size of 100 is used to forecast the active periods of the machines in the production system. An example of the forecasted active duration for all the 13 machines for 190th and 191st production run is summarised in Table 8. The MAPE values of a sliding window of size 100 ARIMA model over the test data set is compared with the naïve method as shown in Figure 7. From Figure 7, the MAPE for sliding ARIMA model with a window size of 100 historical production run data is consistently lower than that of the naïve for all machines the production system indicating that it is a better forecasting model.

6.3. Bottlenecks prediction by ARIMA algorithm

The machine which has the highest forecasted active period percentage for the next production run is the bottleneck machine. T-tests at 95% confidence level are run for other machines with respect to the bottleneck machine to test the statistical significance of differences in the forecasted active period percentage of other machines with respect to the bottleneck machine as explained in step 3 in Section 4. From this, a group of potential bottleneck machines for the next production run is estimated. The t-test results for the 190th and 191st production run are shown in Table 8. It can be observed that for 190th production run machine M1, M6 and M7 are the groups of bottlenecks and the actual bottleneck for that production run is M6. And for a 191st run, M7 and M9 are the predicted group of bottlenecks and the actual bottleneck for that run is M7.

6.4. Algorithm Evaluation

The predicted bottlenecks are compared with the actual bottlenecks to calculate the accuracy, precision, recall and IoU metrics. When identifying the actual bottlenecks for a particular production run based on the highest active periods of the machines, the bottleneck is one machine and not a group of bottleneck machines. This is due to the fact that the standard-error cannot be calculated when computing the actual active period percentage of the machines for a single production run. In other words, the sample size for calculating the actual bottleneck for a production run is one. However, the fixed window size sliding ARIMA computes the variance estimation of the forecast based on the historical data and hence group of bottlenecks can be predicted as shown in Section 4. Therefore, the metrics accuracy, precision, and recall were calculated in two different ways,

Method 1: Evaluating the actual bottleneck machine is within the group of predicted bottleneck machine

Method 2: Evaluating the actual bottleneck machine with the machine with highest predicted active period percentage (without identifying the group of bottlenecks)

T-tests at **95% confidence level** were carried out to test the statistical significance of the difference in the improvement achieved in each of these metrics as explained in Section 5.3. The results are summarised in Table 9.

6.5. Domain Experts Review of the Results

From Table 9, the proposed algorithm has higher and significant mean accuracy, precision, recall, and IoU compared to a naïve method based on the t-values for method 1. However, in method 2, the algorithm doesn't yield statistically significant improvement in recall and IoU compared to naïve. Method 1 of predicting the group of bottlenecks has several advantages for operational reasons when compared to method 2. For example, if one of the predicted bottleneck machines could be a cycle time bottleneck and less prone to production disturbances (e.g. breakdowns, setups, minor stops etc.) and in that case, the other potential bottleneck machines can be focussed for improvement activities. Out of all the metrics shown in Table 9 the domain experts, including the production and maintenance engineers of the test company were interested in the recall metric as that measures the proportion of the actual bottlenecks that was also predicted as bottlenecks by the proposed algorithm. The proposed algorithm has a statistically significant higher recall than that of the naïve method as seen in Table 9 i.e. the recall of 62.53% against the naïve method of 24.69% i.e. an improvement by 37.84% compared to naïve. On the other hand, the accuracy of the proposed metric is also statistically significantly higher than the naïve method by 9.27% for method 1. **Even though the different performance metrics proposed an algorithm in detecting the group of bottlenecks (as shown in method 1 of Table 9) is not 100%, it has surpassed the naive method of bottleneck detection for**

this production system. This indicates that the algorithm provides additional value to in predicting the bottlenecks of the future production run.

7. Discussion

The aim of this paper is to propose a data-driven algorithm to predict throughput bottlenecks of the future production run. The proposed algorithm was tested on a real-world production line. A wide range of metrics was used to evaluate the algorithm performance by comparing it with the naïve method. The proposed algorithm when tested over the real-world production line predicts a group of bottlenecks with a recall metric of 62.53% compared to the naïve method of 24.69%.

7.2. Contributions to the interdisciplinary research field of production and data sciences

The research work presented on developing a data-driven algorithm for bottleneck prediction have significant contributions to the interdisciplinary field of production and data sciences. Compared to the existing methods in the literature as shown in Table 1 in which the different predictive algorithms are tested in a simulation environment, this paper uses only MES data of the machines to develop and test the algorithm. The simulation environment is a highly-controlled environment where the variables that affect the throughput are pre-defined as inputs to a simulation model and therefore it is possible to achieve higher accuracy in predicting the bottlenecks as shown by Li et al. (2011) and Cao et al. (2012). When predicting the bottlenecks in a real-world scenario as presented in Section 6, it wasn't possible to achieve higher accuracy level than simulation. Hence, this suggests a note of caution that the total dependence on the results from the simulation is risky without those algorithms being tested on the real-world production data sets. Also, using the real-world production data sets to develop and test the bottleneck prediction algorithm will also enhance the credibility of the results among the practitioners and manufacturing companies (Amar & Gupta, 1986)**(Jin et al., 2016)(Liao et al., 2017)**. The algorithm presented in this paper is the first approach in the literature to develop and test the predictive algorithm on a real-world MES data set. Also, the existing methods in literature were limited to evaluating the performance of the bottleneck prediction algorithm whereas the proposed algorithm is evaluated based on a wide range of metrics including accuracy, precision, recall, and IoU. Thus, engineers in manufacturing companies can easily understand the algorithm performance based on these metrics. Moreover, this enables the optimisation of the algorithm based on the selected metrics that is more important to the type of decision support that the engineers need about the production line. This facilitates the integration of data sciences domain with the production for making informed decisions. The overall evaluation framework can also be used as a benchmarking tool by researchers to assess the performance of the different algorithms and to select the more appropriate metric.

The production improvements in the machines or other production resources are continuous in the production system (Roser & Nakano, 2015). Due to these changes, the data from very old production runs can no longer be representative of the current conditions and dynamics of the production system, and the usage of that data increases the chance of the predictive algorithm to cover the dynamics of more production runs than it can accurately represent. Moreover, as it is presented in the real-world industrial test study, the MES stores the data only for 315 past production runs. In such cases, the determination of the size of historical data that needs to be used in forecasting the active period is essential. In the test study presented, sliding windows of different sizes were estimated for ARIMA

model and it is found that the past 100 production run data is slightly better in predicting the performance of the machines for the next production run compared to other window sizes for the time unit of interest. Compared with the current prediction algorithms in the literature, this is the first approach to determine how much of the past active duration data of the machines is a good representative of the future.

7.2. Potential impact on management decisions of manufacturing companies

The proposed algorithm can have a significant impact on the production and maintenance management decisions. The algorithmic predictions of the potential group of bottlenecks will help engineers to understand where the production bottlenecks will be in the production line for the next production run and can thereby frame effective strategies to mitigate it. The algorithm not only gives indications on the probable future bottlenecks but also gives valuable information on the possible non-bottlenecks of the production system. The alerts given by the algorithm on bottlenecks and non-bottlenecks can be evaluated by the production and maintenance domain experts who have years of experience in their production system to decide whether they are logical and actionable to frame strategies for mitigating the bottleneck. The aim of the predictive algorithms is to complement their efforts to manage the true bottlenecks more effectively. The combined insights can then be used in the production planning and management meetings before the production run (e.g. morning meetings before the start of the production) to effectively plan proactive actions. For example, the strategies for reducing speed in the non-bottlenecks could be made which can increase the tool life etc. (Roser & Nakano, 2015) and prioritize the bottleneck machines for reactive maintenance. In this way, the engineers' knowledge of production systems can be combined with data sciences to get more transparency on the expected dynamics of the machines for the future production run and plan strategies. Hence, this type of system level decision support tool helps the engineers to focus more on bottleneck machines in the production system by which the throughput and the productivity can be increased (Jin et al., 2016)(Bokrantz et al., 2017).

7.3. Methodological discussions

The active period method used in this study detects the bottlenecks using only the states of the machine. Quality parameters, which factors the scrapped manufactured products or rework is not taken in account when detecting the bottlenecks. Hence the active period method is a method that focuses on determining the bottleneck from an availability perspective. This perspective is still important for improving the production flow and maintenance planning on bottleneck machines. (Wedel et al., 2016).

The intention of this research work is to provide an algorithm towards the bottleneck prediction in a production system using real-world MES type of data and to provides a generic framework to evaluate the algorithm based on different metrics. The strength of the proposed algorithm has been demonstrated using a real-world test study. In that test study, the MES data set is divided into training and testing data and the algorithm is tested on the testing data set. This type approach prevents the data bias (Keogh & Kasetty, 2002). Though the proposed algorithm is not tested on the MES data sets of multiple production systems, the authors are convinced that the algorithm works in an environment where MES records the different active states of the machines and their timestamps during the production runs and across different production runs. These are the only inputs to the proposed algorithm. **The same algorithm can also be used for modeling the active periods for transportation resources if the information on the states and the corresponding timestamps are available in MES. For example, when the transportation resources move to a pickup location or a drop of location, then the state of this resource is comparable to the producing state of the machine and hence it can be classified as an active state. Similarly, when the resource is waiting**

for the parts to be picked up, then it could be classified as an inactive state (Roser et al., 2001). **This when analyzed in combination with the machines active period, it can be investigated whether the transportation resources act as bottlenecks or not.** Also, though the proposed algorithm is demonstrated for a time resolution of a complete production run, the same algorithm can also be used for different time resolutions by aggregating the active period data based on the required resolution.

7.4. Future work

The proposed bottleneck detection algorithm was developed, tested and validated with the historical data of a real-world production system and evaluated its performance. On the other hand, the algorithm also needs to be evaluated in a real-time scenario and potential effects need to be studied further. This will allow the institutionalization of predictive analytics by manufacturing companies to enable data-driven decisions at the shop floor level to be more effective.

Also, as this paper is focused towards predicting the bottlenecks in the production system using MES data, the future steps will be to get diagnostic insights on the predicted bottleneck to understand the reasons of the machine disturbances**. Trying other time series algorithms such as neural networks etc. can be used to compare the prediction accuracies with that of the based ARIMA algorithm for bottleneck detection. Moreover, variety based analytics including the quality parameters of the machine etc. can be combined with the active period information to predict the bottlenecks which ensures the fact that improving the throughput from the bottlenecks also ensure high quality**. The results from the prediction algorithm can be augmented with the expert opinions and can be embedded into the system to improve prediction accuracy.

8. Conclusion

In this paper, a data-driven algorithm is proposed to predict the throughput bottlenecks in a production system based on the active periods of the machines. The inputs to the algorithm are the states of the machines and the corresponding periods of those states across different production runs. **The algorithm is tested on MES data sets of a real-world production system and the performance of the algorithm is evaluated over a wide range of metrics, thus enbaling the manufacturing companies to trust the algorithm.** In the test study, the recall metric of interest for the engineers of the proposed algorithm outperformed the naïve method with an improvement of 37.84%. Compared to the current techniques in the literature, an attempt is made to determine how much of the historical data is used to predict the bottlenecks. The approach presented was developed with active involvement from the domain experts from machine learning and production field. Appropriately formulating the problem by considering the nature of the data and the real-world constraints and incorporating the appropriate metrics for evaluation are the lessons learned. In addition, this paper provides a generic framework for the evaluation of the different algorithms. Thus, this framework can be used as a benchmarking tool by other researchers and engineers to compare and evaluate the performance of different algorithms. From the indications on predicted bottlenecks combined with the knowledge of the production and maintenance engineers, fact-based decisions can be made to mitigate the bottlenecks and thus increasing the throughput and productivity of the system. This research work contributes to the cross-disciplinary field of production and data sciences.
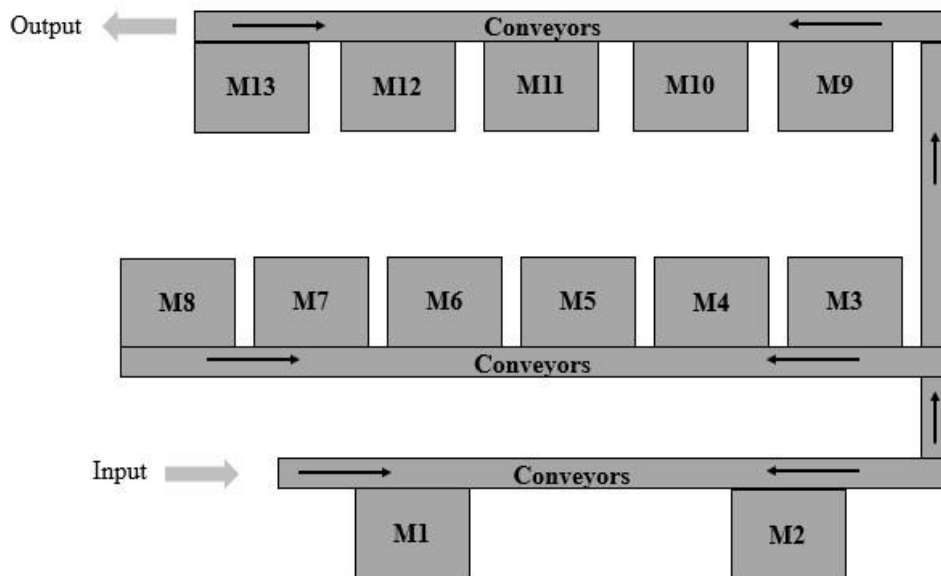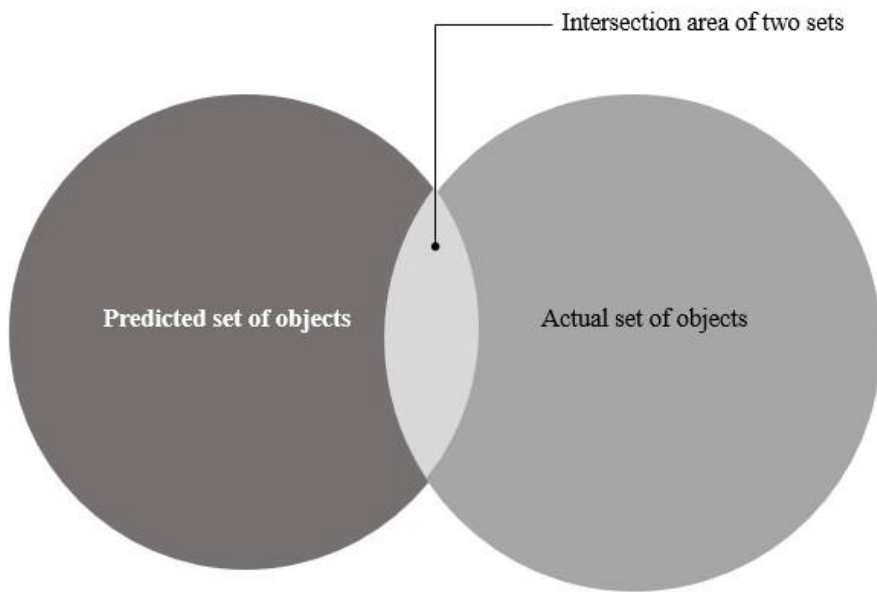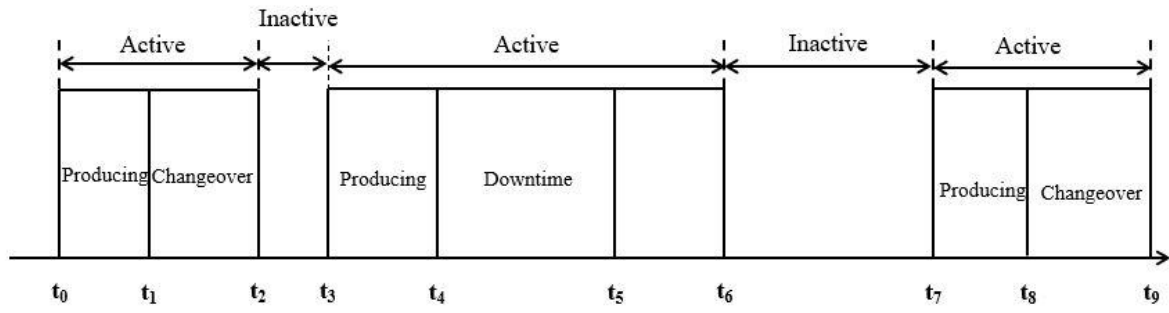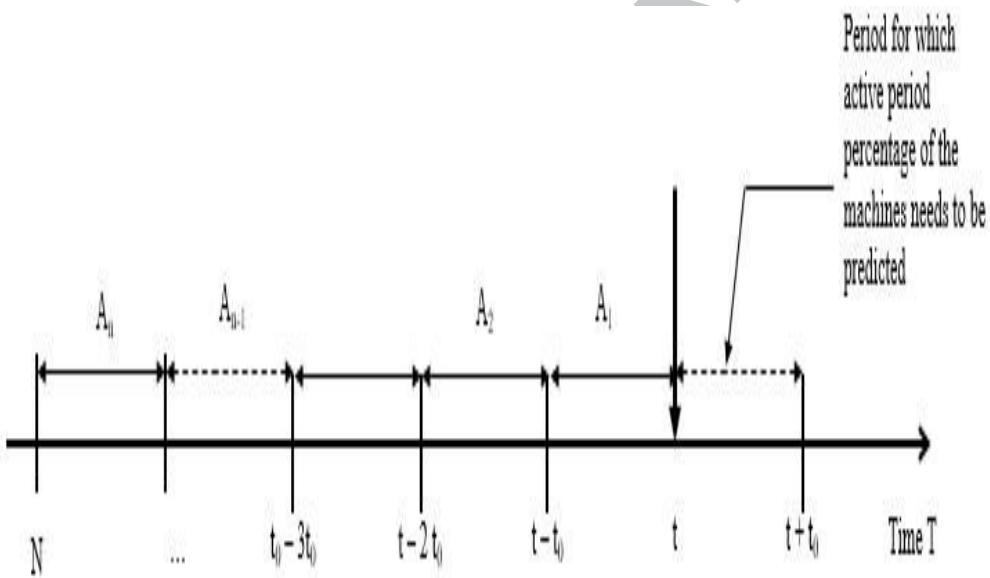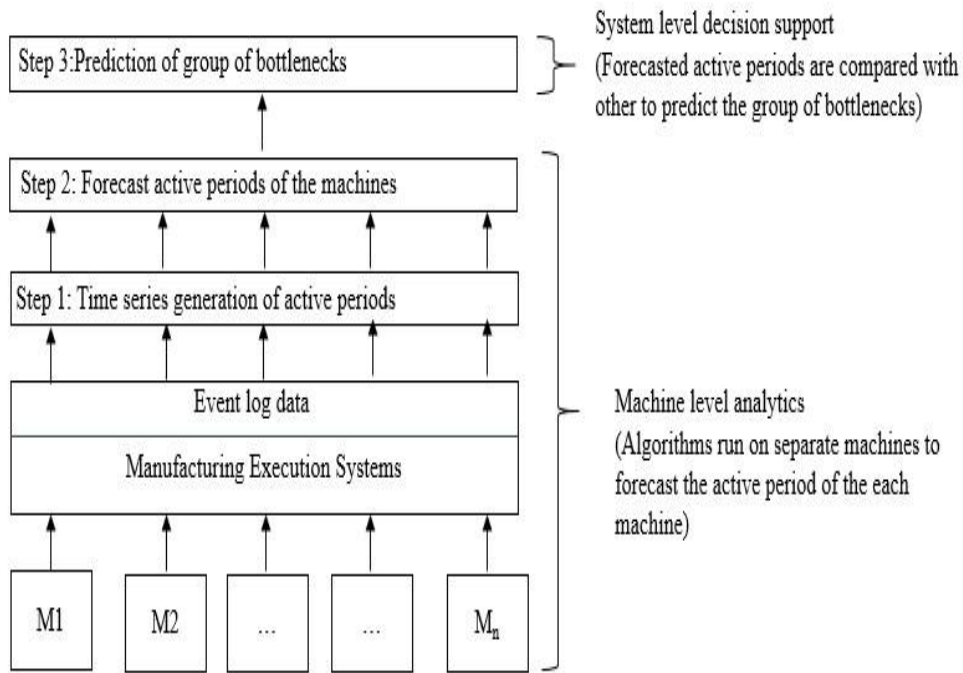
References

Adebiyi A., A., & Adewumi O., A. (2014). Stock Price Prediction Using the ARIMA Model. In *16th International Conference on Computer Modelling and Simulation Stock*. http://doi.org/10.1109/UKSim.2014.67

Ahmed, F., Tarlow, D., & Batra, D. (2015). Optimizing expected intersection-over-union with candidate-constrained CRFs. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1850–1858). http://doi.org/10.1109/ICCV.2015.215

Amar, A. D., & Gupta, J. N. D. (1986). Efficiency of Scheduling Algorithms Simulated Versus Real Life Data in Testing the Efficiency of Scheduling Algorithms. *IIE Transactions*, *18*(1), 16–25. http://doi.org/10.1080/07408178608975325

Bokrantz, J., Skoogh, A., Berlin, C., & Stahre, J. (2017). Maintenance in digitalised manufacturing : Delphi-based scenarios for 2030. *International Journal of Production Economics*.

Bradley, A. P. (1997). The Use of the Area Under the ROC Curve in the Evaluation of Machine Learning Algorithms. *Pattern Recognition*, *30*(7), 1145–1159. http://doi.org/10.1016/S0031-3203(96)00142-2

Brockwell, P. ., & Richard, D. . (2016). *Introduction to Time Series and Forecasting*. (G. (University of F. Casella, S. (Carnegie M. U. Fienberg, & I. (Stanford U. Olkin, Eds.) (Second). Springer. Retrieved from www.springer.com/series/417

Cao, Z., Deng, J., Liu, M., & Wang, Y. (2012). Bottleneck prediction method based on improved adaptive network-based fuzzy inference system (ANFIS) in semiconductor manufacturing system. *Chinese Journal of Chemical Engineering*, *20*(6), 1081–1088. http://doi.org/10.1016/S1004-9541(12)60590-4

Chang, Q., Ni, J., Bandyopadhyay, P., Biller, S., & Xiao, G. (2007). Supervisory Factory Control Based on Real-Time Production Feedback. *Journal of Manufacturing Science and Engineering*, *129*(3), 653. http://doi.org/10.1115/1.2673666

Goldrat, E. ., & Cox, J. (1990). *The Goal: A Process of Ongoing Improvement* (Third Revi). Great Barrington, MA: North River Press.

Gopalakrishnan, M., Skoogh, A., & Christoph, L. (2013). Simulation based planning of maintenance activities in the automotive industry. In R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, & M. . Kuhl (Eds.), *Proceedings of the 2013 Winter Simulation Conference* (pp. 342–353).

Gröger, C., Niedermann, F., & Mitschang, B. (2012). Data mining-driven manufacturing process optimization. In *Proceedings of the World Congress on Engineering* (Vol. III, pp. 0–6). London, UK. Retrieved from http://www.iaeng.org/publication/WCE2012/WCE2012_pp1475-1481.pdf

Hedman, R., Subramaniyan, M., & Almström, P. (2016). Analysis of critical factors for automatic measurement of OEE. *Procedia CIRP*, *57*, 128–133. http://doi.org/10.1016/j.procir.2016.11.023

Ho, S. ., & Xie, M. (1998). The use of ARIMA models for reliability forecasting and analysis. *Computers and Industrial Engineering*, *35*(1–2), 213–216.

Hyndman, R. J. (2014). *Forecasting : Principles & Practice*. University of Western Australia. Retrieved from http://robjhyndman.com/uwafiles/fpp-notes.pdf

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting : the forecast package for R Automatic time series forecasting : the forecast package for R. *Journal Of Statistical Software*, *27*(3), 1–22. http://doi.org/10.18637/jss.v027.i03

Hyndman, R. J., & Yeasmin, K. (2008). Automatic Time Series Forecasting : the forecast Package for R. *Journal of Statistcal Software*, *27*(3).

Jin, X., Weiss, B. A., Siegel, D., & Lee, J. (2016). Present Status and Future Growth of Advanced Maintenance Technology and Strategy in US Manufacturing. *International Journal of Prognostics and Health Management*, *7*.

Keogh, E., & Kasetty, S. (2002). On the need for time series data mining benchmarks. In N. Hand & Keim (Eds.), *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '02* (p. 102). Kluwer Academic Publishers. http://doi.org/10.1145/775047.775062

Keskin, M. E., Taylan, D., & Terzi, Ö. (2006). Adaptive neural-based fuzzy inference system (ANFIS) approach for modelling hydrological time series. *Hydrological Sciences Journal*, *51*(4), 588–598. http://doi.org/10.1623/hysj.51.4.588

Knezevic, A. (2008). Overlapping Confidence Intervals and Statistical Significance. *Cornell Statistical Consulting Unit StatNews # 73*.

Kuvulmaz, J., Usanmaz, S., & Engin, S. N. (2005). Time-Series Forecasting by Means of Linear and Nonlinear Models. In A. Gelbukh, Á. de Albornoz, & H. Terashima-Marín (Eds.), *MICAI 2005: Advances in Artificial Intelligence: 4th Mexican International Conference on Artificial Intelligence, Monterrey, Mexico, November 14-18, 2005. Proceedings* (pp. 504–513). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/11579427_51

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big Data , Analytics and the Path From Insights to Value Big Data , Analytics and the Path From Insights to Value. *MIT Sloan Management Review*, *52*(2), 21–32.

Lee, J., Lapira, E., Bagheri, B., & Kao, H. (2013). Recent advances and trends in predictive manufacturing systems in big data environment. *Manuf Lett*, *1*. http://doi.org/10.1016/j.mfglet.2013.09.005

Li, J., Blumenfeld, D. E., Huang, N., Alden, J. M., Li, J., Blumenfeld, D. E., … Alden, J. M. (2009). Throughput analysis of production systems : recent advances and future topics. *International Journal of Production Research*, *7543*(November). http://doi.org/10.1080/00207540701829752

Li, L., Chang, Q., & Ni, J. (2009). Data driven bottleneck detection of manufacturing systems. *International Journal of Production Research*, *47*(18), 5019–5036. http://doi.org/10.1080/00207540701881860

Li, L., Qing, C., Xiao, G., & Ambani, S. (2011). Throughput Bottleneck Prediction of Manufacturing Systems Using Time Series Analysis. *Journal of Manufacturing Science and Engineering*, *133*, 1–8. http://doi.org/10.1115/1.4003786

Liao, Y., Deschamps, F., Freitas, E. De, & Loures, R. (2017). Past , present and future of Industry 4 . 0 - a systematic literature review and research agenda proposal. *International Journal of Production Research*, *7543*(December), 0. http://doi.org/10.1080/00207543.2017.1308576

Lima, E., Chwif, L., & Barreto, M. R. . (2008). Metodology for selecting the best suitable bottleneck detection method. In S. . Mason, R. Hill, L. Mönch, O. Rose, T. Jefferson, & J. . Fowler (Eds.), *Proceedings of the 2008 Winter Simulation Conference* (pp. 1746–1751).

Pete, C., Julian, C., Randy, K., Thomas, K., Thomas, R., Colin, S., & Wirth, R. (2000). *CRISP-DM 1.0*. *CRISP-DM Consortium*.

Ron, A. J. De, & Rooda, J. E. (2005). Equipment Effectiveness : OEE Revisited. *IEEE Transactions on Semiconductor Manufacturing*, *18*(1), 190–196.

Roser, C., & Nakano, M. (2015). A Quantitative Comparison of Bottleneck Detection Methods in Manufacturing Systems with Particular Consideration for Shifting Bottlenecks. *IFIP International Federation for Information Processing 2015*, *1*, 273–281. http://doi.org/10.1007/978-3-319-22759-7

Roser, C., Nakano, M., & Tanaka, M. (2001). A Practical Bottleneck Detection Method. In B. . Peters,

J. . Smith, D. . Medeiros, & M. . Rohrer (Eds.), *Proceedings of the 2001 Winter Simulation Conference*.

Shao, G., Shin, S. J., & Jain, S. (2015). Data analytics using simulation for smart manufacturing. In A. Tolk, S. . Diallo, I. . Ryzhov, L. Yilmaz, S. Buckley, & J. . Miller (Eds.), *Proceedings - Winter Simulation Conference* (pp. 2192–2203). http://doi.org/10.1109/WSC.2014.7020063

Skoogh, A. (2011). *Automation of Input Data Management. Chalmers University of Technology*. Chalmers University of Technology.

Subramaniyan, M. (2015). *Production Data Analytics – To identify productivity potentials. Chalmers University of Technology*. Chalmers University of Technology.

Subramaniyan, M., Skoogh, A., Gopalakrishnan, M., & Hanna, A. (2016). Real-time data-driven average active period method for bottleneck detection. *International Journal of Design & Nature and Ecodynamics*, *11*(3), 428–437. http://doi.org/10.2495/DNE-V11-N3-428-437

Subramaniyan, M., Skoogh, A., Gopalakrishnan, M., Salomonsson, H., Hanna, A., & Lämkull, D. (2016). An algorithm for data-driven shifting bottleneck detection. *Cogent Engineering*, *54*, 1–19. http://doi.org/10.1080/23311916.2016.1239516

Vafaeipour, M., Rahbari, O., Rosen, M. A., Fazelpour, F., & Ansarirad, P. (2014). Application of sliding window technique for prediction of wind velocity time series. *International Journal of Energy and Environmental Engineering*, *5*(2–3), 1–7. http://doi.org/10.1007/s40095-014-0105-5

Wang, J., Wang, J., Zhang, Z., & Guo, S. (2012). Stock index forecasting based on a hybrid model. *Omega*, *40*(6), 758–766. http://doi.org/10.1016/j.omega.2011.07.008

Wedel, M., Noessler, P., & Metternich, J. (2016). Development of bottleneck detection methods allowing for an effective fault repair prioritization in machining lines of the automobile industry. *Production Engineering*, *10*(3), 329–336. http://doi.org/10.1007/s11740-016-0672-9

Wuest, T., Weimer, D., Irgens, C., Thoben, K., Wuest, T., Weimer, D., … Thoben, K. (2016). Machine learning in manufacturing : advantages , challenges , and applications. *Production & Manufacturing Research*, *3277*(February 2017), 1–23. http://doi.org/10.1080/21693277.2016.1192517

Step 3:Prediction of group of bottlenecks

System level decision support
(Forecasted active periods are compared with
other to predict the group of bottlenecks)

Step 2: Forecast active periods of the machines

Step 1: Time series generation of active periods

Event log data

Manufacturing Execution Systems

Machine level analytics
(Algorithms run on separate machines to
forecast the active period of the each
machine)

M1   M2   ...   ... M_n

Period for which
active period
percentage of the
machines needs to be
predicted

A_n A_{n-1} A_2 A_1

N ... $t_0 - 3t_0$ $t - 2t_0$ $t - t_0$ t $t + t_0$ Time T
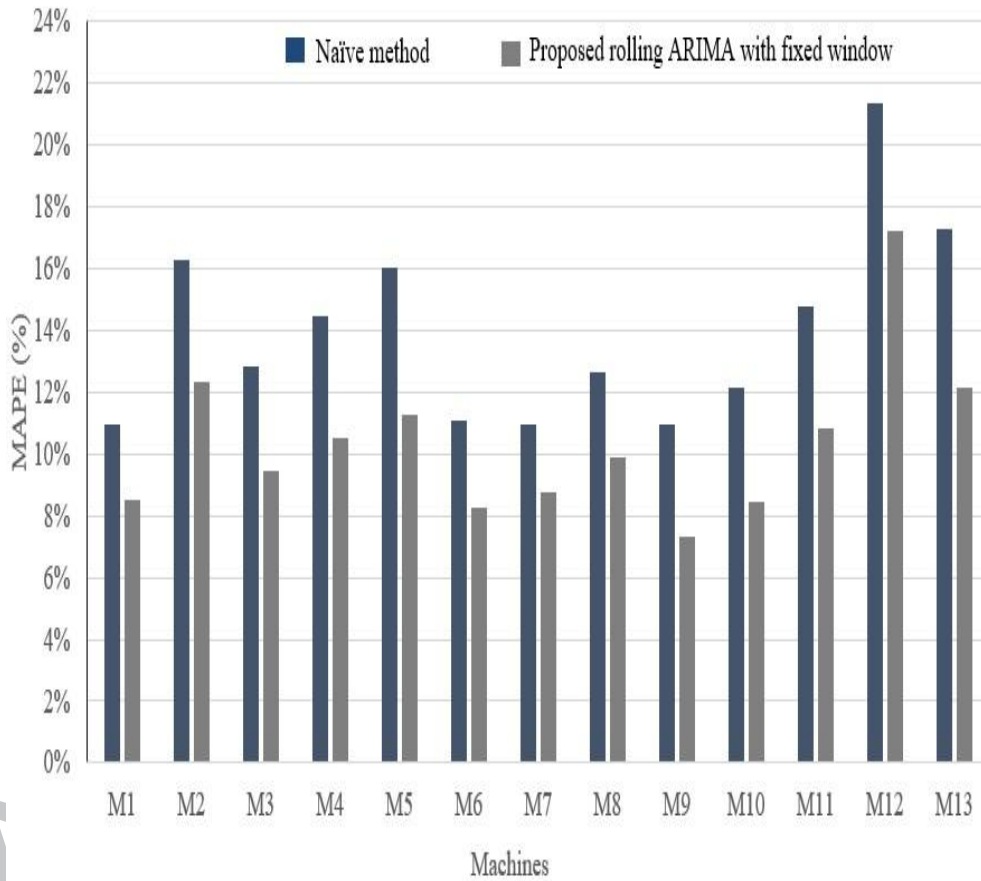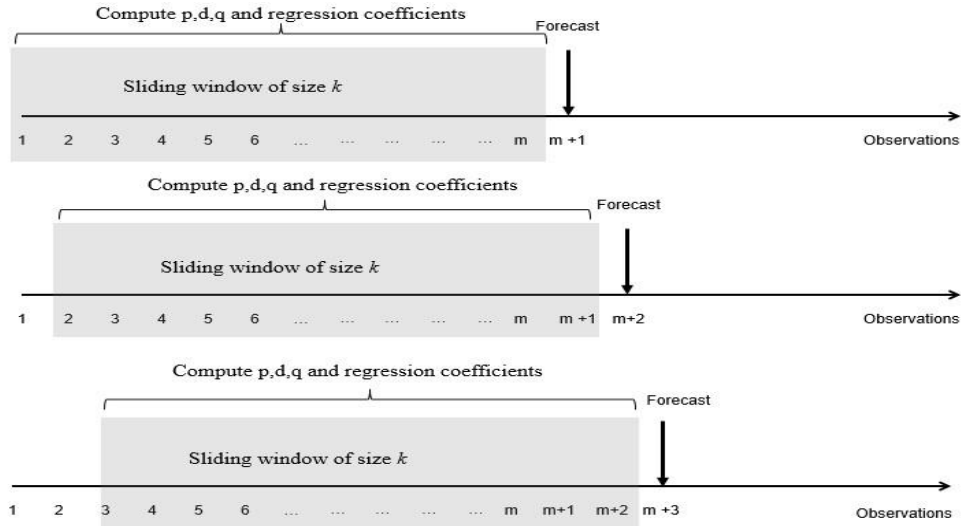
Table 1: Literature collection on bottlenecks prediction in production system

Predictive Method	Inputs	Output	Bottleneck Detection Technique
Auto – Regressive Moving Average (ARMA) (Li et al., 2011)	Time series data blockage and starvation of machines	Forecasted value of blockage and starvation for next production run	Turning point method

Adaptive Network Based Fuzzy Inference System (ANFIS) (Cao et al., 2012)	Processing time, utilisation rate, buffer length, mean time between failure (MTBF), mean time to Repair (MTTR), work in progress inventory (WIP), job types	Main bottleneck, sub bottlenecks	Work load on the machines calculated from the input variables
Varying buffer levels (Wedel et al., 2016)	Machine cycle time, buffer capacity between the machines, current buffer levels, MTTR, MTBF	Downtime bottlenecks for maintenance purposes	Buffer level before the machines

Table 2: Confusion Matrix

		Predicted Responses	
		True	**False**
Actual response	**True**	True Positive (TP)	False Negative (FN)
	False	False Positive (FP)	True Negative (TN)

Table 3: Example MES record of a machine

Machine	Red	Yellow	Green	White	Duration (sec)	Date and Time
M1	0	0	1	1	80	01-07-2016 06:02:18
M1	0	0	0	1	997	01-07-2016 06:03:38
M1	0	0	1	1	39	01-07-2016 06:20:15
M1	0	0	0	1	997	01-07-2016 06:20:54
M1	0	0	1	1	39	01-07-2016 06:37:31

Table 4: Grouping of different ANDON lights

ANDON Light	States
Yellow	Producing
Green	
White	
Yellow+White	
Yellow+Green	Blocked/Starved/ Idle
Yellow+Green+White	
Green+White	
Red	Down
Red	
No light	

Table 5: Adapted Confusion matrix

	Predicted bottleneck machines	Predicted non-bottleneck machines
Actual bottleneck machines	True Positive (TP)	False Negative (FN)
	How many predicted bottleneck machines were actual bottleneck machines?	How many predicted non-bottleneck machines were actual bottleneck machines?
Actual non-bottleneck machines	False Positive (FP)	True Negative (TN)
	How many predicted bottleneck machines were actual non-bottleneck machines?	How many predicted non-bottleneck machines were actual non-bottleneck machines?

Table 6: Example of time series representation of the active period data for each machine M1

Production run	Active period (%)
1	86.97
2	84.96
3	74.12
…	…
…	…
…	…
315	80.33

Table 7: MAPE values of different window sizes

	Different window sizes and their MAPE values from ARIMA			
Machines	50	100	150	185
M1	8.64%	8.50%	8.66%	8.56%
M2	12.45%	12.36%	12.91%	12.98%
M3	9.88%	9.42%	9.87%	9.69%
M4	10.35%	10.50%	10.70%	10.60%
M5	11.13%	11.27%	11.75%	11.69%
M6	8.44%	8.25%	8.83%	8.74%
M7	8.75%	8.75%	8.92%	8.89%
M8	10.01%	9.89%	10.62%	10.07%
M9	7.34%	7.35%	7.38%	7.50%
M10	8.54%	8.43%	8.55%	8.53%
M11	10.78%	10.80%	11.32%	11.60%
M12	17.30%	17.21%	17.28%	17.43%
M13	12.24%	12.13%	12.48%	12.35%
Average for the production line	10.45%	10.37%	10.71%	10.66%

Table 8: Forecasted active duration results for 13 machines

Run	Machine	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
190	ARIMA (p,d,q)	4,1,3	0,1,1	0,1,1	4,1,1	3,1,1	5,1,2	3,1,2	0,0,1	1,0,1	1,1,0	0,0,3	0,1,0	2,0,3
	Forecast (%)	84.08	76.43	77.22	75.93	74.50	83.08	82.08	79.41	79.72	75.71	72.63	63.94	60.87
	Standard error (%)	0.84	1.03	0.91	0.94	0.95	0.76	0.88	0.99	0.85	0.86	1.05	1.20	0.91
	t_{stat}	-	5.74	5.54	6.48	7.58	0.89	1.65	3.59	3.65	6.98	8.53	13.75	18.73
	Actual (%)	62.11	61.23	53.41	54.94	53.36	67.13	64.46	63.37	62.25	65.24	54.11	43.19	47.99
191	ARIMA (p,d,q)	1,1,1	2,1,1	0,1,1	4,1,2	4,1,1	3,1,1	3,1,1	3,1,1	4,1,4	0,0,0	2,0,2	0,1,0	1,0,3
	Forecast (%)	75.55	73.49	75.41	72.99	72.04	73.35	78.64	74.58	79.55	75.59	66.03	63.72	60.70
	Standard error (%)	0.87	1.04	0.94	0.92	0.97	0.78	0.84	0.94	0.87	0.86	1.02	1.22	0.92
	t_{stat}	3.26	4.47	3.25	5.19	5.79	5.32	0.75	3.88	-	3.24	10.09	10.60	14.93
	Actual (%)	83.04	82.98	82.41	78.07	79.82	83.65	83.77	83.75	80.03	76.52	71.12	66.12	53.30

Table 9: Evaluation metrics results

Method		Proposed Algorithm				Naïve method			
		Accuracy (%)	Precision (%)	Recall (%)	IOU Metric (%)	Accuracy (%)	Precision (%)	Recall (%)	IOU Method (%)
1	Mean	86.13	36.34	62.53	29.89	78.82	29.40	24.69	20.81
	Standard Error	0.72	2.14	3.08	1.95	0.59	2.81	2.73	2.98
	t_{stat}	-	-	-	-	7.85	1.97	9.19	2.54
2	Mean	89.21	39.06	28.54	26.52	78.82	29.40	24.69	20.81
	Standard Error	0.57	3.22	2.69	1.63	0.59	3.15	2.81	2.98
	t_{stat}	-	-	-	-	12.66	2.144	0.98	1.68
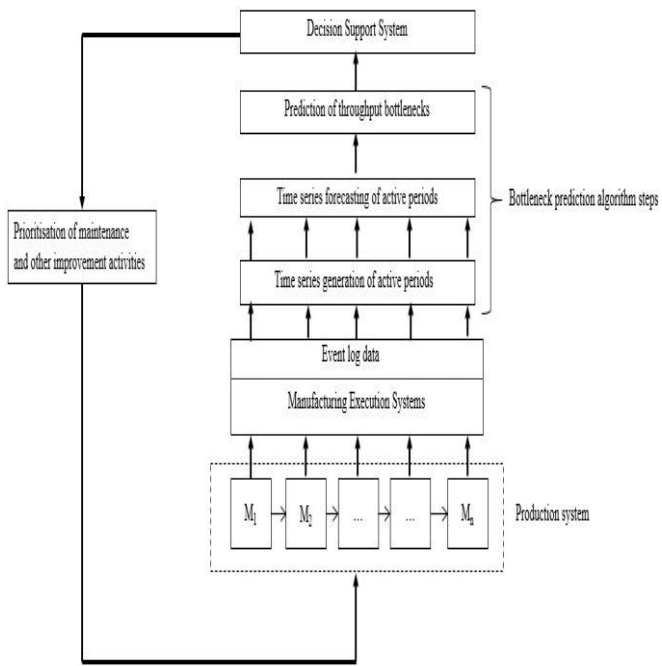
Highlights

- Proposes an algorithm to predict the throughput bottlenecks from real-time data of the machines for engineers to take proactive decisions
- Proposes an approach to determine the size of historical data that is a good representative of the future
- Proposes a novel evaluation framework to evaluate the bottleneck prediction algorithm by treating it as a binary classifier problem
- Tests the algorithm and evaluated its performance on a real-world production system and compared the results with naïve method
- Proposed algorithm outperformed the naïve method of bottleneck prediction