Machine Learning Algorithm with Adaptive Data Selection for Component Obsolescence Prediction

Ajay K Dass[1*], S. D. Lokhande[2]

Abstract

In the manufacturing sector, product obsolescence happens as fresh goods are developed that performs better or is more reasonably priced. A positive approach to anticipating component obsolescence can lower production costs and increase consumer contentment. In this research, a proactive method for forecasting computer components obsolescence based on machine learning algorithms is suggested. The suggested technique creates numerous models which trained using data. The proposed strategy increases the accuracy of obsolescence prediction according to the results of empirical testing. The Bayesian Regression predictive model is implanted for Optimal Inventory prediction Model with Product Obsolescence Risk. The dataset under review have the sample quantity of computer products sale. The training set is fed into the algorithm once the dataset has been divided into training and testing sets so that it can learn how to predict the values. The Bayesian Regression model compared with the linear regression and neural network regression predictive models. The Bayesian Regression model produces the best forecasting outcomes for predicting the sales of the computer items.

KeyWords: Component obsolescence; diminishing manufacturing sources and material shortages; forecasting; machine learning

DOI Number: 10.14704/nq.2022.20.5.NQ22788 **NeuroQuantology 2022; 20(5): 5109-5115**

Introduction

Businesses that sell spare parts or service components could run into situations when the majority of their Stock Keeping Units (SKUs) are gradually moving, resulting in protracted periods with little demand [1]. Forecasting is difficult for these demand series since they are irregular and intermittent. The situation gets trickier as more businesses contend for clients and are obligated to fulfill minimal service standards. Therefore, precise forecasting is essential for effective inventory management [2]. This is one of the concerning businesses when it comes to the data collection and processing of data.

A huge amount of high dimensionality about products is produced as a result of the digital realm and technological developments. Manufacturing sector facilities have evolved as a result of the utilization of different data mining technologies and machine learning algorithms which have been used to find significant patterns and identify correlations and interactions between several variables across large databases. It is a crucial tool for delivering and contrasting current facts in order to choose the best course of action. This technology enables the study of huge quantities of information by fusing several analytical approaches with cutting-edge, computational methods.

The origins of element obsolescence can really be categorized as technological innovation, accessibility, resource scarcity, commercial feasibility, and climate change, according to [3]. Early detection and management of part obsolescence issues brought on by these numerous sources is more inexpensive and effective. Treatment that works, as opposed to constantly reminding, detects and prioritizes parts that face obsolescence risk. Correctly predicting the danger of component obsolescence is crucial for proactive management to be able to propose remedies before components are terminated.

Corresponding author: Ajay K Dass
Address: [1,2]Department of E & TC, Sinhgad College of Engineering, Pune

There have been several studies on predicting the likelihood of component obsolescence, and they may be categorized into two types: those that use mathematical models based on Statistical Approaches (SAs) and those that use Machine Learning (ML).

The predictions made feasible by the models provided in the [4] allow specialists to better manage the emergence and continuous usage of field goal items in light of the anticipated life cycle of the parts incorporated into the goods. The application of the concept to integrated circuits is detailed, and estimates of the obsolescence of Dynamic Random Access Memories (DRAMs) are presented. The goals are fewer process iterations, lower inventory costs, lower sustainment costs, and lower overall life cycle product costs.

In [5], demonstrates how commercial obsolescence risk databases can be advanced and their forecasting abilities enhanced using data mining-based techniques. The method combines life cycle curve forecasts with the detection of vendor-specific windows of obsolescence for electronic parts using data mining of previous last-order or last-ship dates. The protracted technique can produce predictions for assurance levels that are defined by the user in addition to enabling more accurate obsolescence estimates. The technique has been used to demonstrate both complete modules and individual components.

The method proposed in [6] for creating methods presented in this research can be used to estimate the obsolescence dates of electronic components without obvious historical adjustable variables. The approach is based on computing buy lifetime using databases of prior obsolescence occurrences and newly introduced parts that have not yet become outdated. The technique has been tested on a variety of electronic parts, both for general part trending and for the trending of particular part properties.

The study presented in [7]is a novel way for representing the degree of obsolescence as a time-dependent quantity using a probability distribution. The length of each work before obsolescence is assessed, and the degree of obsolescence is simulated as a function of time using revenue figures. The future of smartphones is evaluated in relation to the new framework. The results of this study are then analyzed, and it is advised that more research be conducted to increase the method's applicability.

How well the method may accurately estimate device obsolescence is shown, a study of the cell phone market is provided [8]. Obsolescence dates are routinely predicted within several months according to the example authors of the study obsolescence prediction results, which explicitly predict items as active or obsolete with a 98.3% accuracy rate.

The recommended algorithm in [9] works past these limitations in two different ways. An unsupervised clustering method is used to first classify the data based on how similar they are to each other in order to generate distinct machine learning models tailored for each group. A hybrid strategy made up of a number of reliable techniques is developed to boost prediction accuracy and get over the data limitation. It has been empirically demonstrated that the suggested clustering-based hybrid technique improves prediction accuracy of the obsolescence date for the electrical component data.

In [10], compares stochastic gradient boosting, stepwise logistic regression, GA-RF, and RF to ascertain this strategy's viability. With 93.3% accuracy, 90.4% sensitivity, and 95.4% specificity, test results showed that GA-RF outperformed the competition.

Deep learning (DL) algorithms, which can learn from enormous volumes of data, have recently produced encouraging results in a number of domains [11][12]. To the best of the authors' knowledge, there haven't been many researches utilizing ML to forecast the obsolescence of parts. Among them, Random Forest (RF), neural network, and support vector machine algorithms were put to the test by Jennings et al. [8] for forecasting obsolescence risk and product lifecycles. The RF approach for obsolescence forecasting was utilized by Grichi et al. and was enhanced by merging it with a genetic algorithm [13]. The lack of readily available data can be blamed for the ML or DL algorithms for obsolescence prediction's mediocre performance.

In the manufacturing sector, product obsolescence happens as fresh goods are developed that performs better or is more reasonably priced. A positive approach to anticipating component obsolescence can lower production costs and increase consumer contentment. In this research, a proactive method for forecasting computer components obsolescence based on machine learning algorithms is suggested. For the optimal inventory prediction model with product obsolescence risk, the Bayesian Regression predictive model is implanted. The dataset under consideration includes a representative sample of computer product sales. After the dataset has been split into training and testing sets, the training set is fed into the algorithm so it can learn how to predict the values. Compared to the linear regression and neural network regression predictive methods is the Bayesian Regression model. The Bayesian Regression model yields the most accurate forecasting results for estimating the sales of computer-related goods.

Methodology

Optimal Inventory Prediction Model with Product Obsolescence Risk proposed methodology is depicted in Figure 1. The analysis and forecasting of future purchases are done using information from a company's loyalty programme. Analyzing the past is what historical data does. Businesses that sell computers gather lots of information. In order to find patterns in historical and transactional data and demonstrate trends, this project combines predictive models with machine learning algorithms. Additionally, a customer-specific predictive score that is meaningful is produced. To detect dangers and possibilities, it then applies a predictive score to the most recent data. Historical data is offered in both structured and unstructured formats.

Data that has been structured is better ordered and defined and is stored in places like databases. This structure makes it simpler to access for predictive analysis. By analyzing past and present data trends, a mathematical framework is modeled in order to anticipate future behavior. A combination of historical and real-time data is typically needed for predictive analytics. Regardless of the kind of models already being used, the following procedures are used to deploy predictive models:

Treating missing data and eliminating outliers are also parts of data cleaning.

Examine whether parametric versus nonparametric predictive modeling is more effective.

Reorganize the data such that it may be used by the modeling process.

Pick a subset of the data that will be utilized in the model's training.

Fine-tune the model parameters using the training dataset.

Test the outcomes of the predictive model performance monitoring to ascertain the model's efficacy.

Examine the predictive modeling's precision using data that weren't utilized to modify the model. Utilize the forecasting model.

In this study, a representative selection of data points is examined, and the best prediction model can then be developed from them and recommended for data analysis. For an optimal inventory prediction model with product obsolescence risk, the Bayesian Regression (BR) predictive model is implanted.

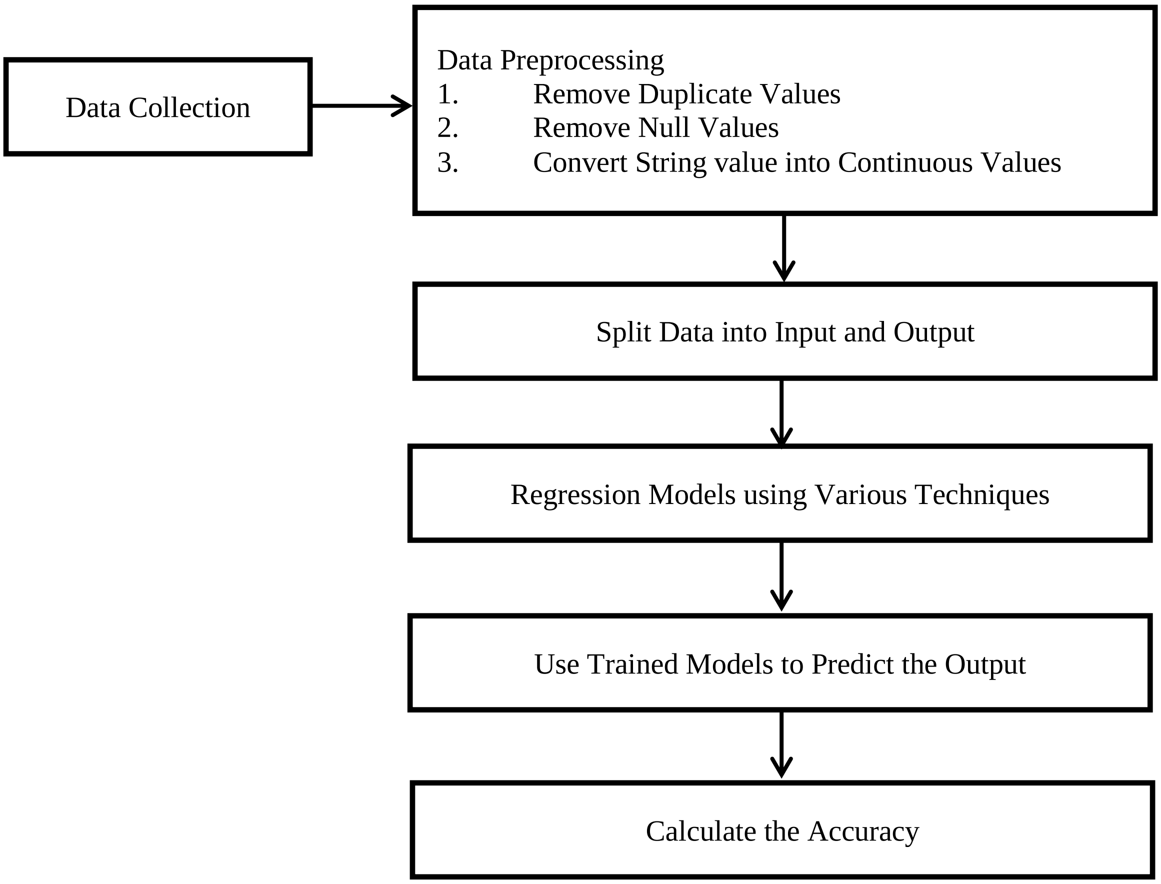

Figure 1: Proposed methodology for Optimal Inventory prediction Model with Product Obsolescence Risk

Bayesian Regression Predictive Model

Although their input variables may not always be linear, LR models all share this characteristic. LR can simulate any non-linearity from input variables to goals using non-linear basis functions of the input variables. One such example is polynomial regression, which will be illustrated later. Thus, Equation 1 can be used to create a LR model. Equation 1 can be written as Equation 2 and Equation 3.

$$Y_{(x,w)} = \omega_0 + \sum_{i=1}^{n-1} \omega_i \phi_i x \tag{1}$$

$$Y_{(x,w)} = \sum_{i=0}^{n-1} \omega_i \phi_i x \tag{2}$$

$$Y_{(x,w)} = \omega^t \phi_i x \tag{3}$$

Where n is the sum of the parametersω_i, including the bias termω_0, and i are basis functions. Here, one can apply the rule $\phi_0\ (x)$=1. The identity $\phi_0\ (x) = x$ serves as the set of basis functions in the simplest form of linear regression models, which are also linear functions of their input variables. An additive random noise (£) plus a deterministic function $Y_{(x,w)}$ determines the target variable T of an observation x as given in Equation 4.

$$T = Y_{(x,w)} + £ \tag{4}$$

It is assume that the noise is normally distributed, such that, it possesses an inverse variance Gaussian distribution with zero mean and precision (β). Equation 5 can be used to represent the relevant probabilistic model, which is the conditional distribution of T given x.

$$p(T|x,\omega,\beta) = N\left(T|Y_{(x,w)},\beta^{-1}\right) = \sqrt{\frac{\beta}{2\pi}} e^{\left(-\frac{\beta}{2}(T-Y_{(x,w)})^2\right)} \tag{5}$$

Where the regression function $Y_{(x,w)}$ is this distribution's mean. The probability function's log can be expressed as Equation 6 with training set of N independent and identically distributed observations $x_1, x_2, x_3, \dots, x_N$ and their corresponding targets$T_1, T_2, T_3, \dots, T_N$.

$$log\ logp(T|x,\beta) = \frac{N}{2} log\ log(\beta) - \frac{N}{2} log\ log(2\pi) - \beta E_{D(\omega)} \tag{6}$$

Where $E_{D(\omega)}$ is the sum-of-squares error function derived from the likelihood function's exponent. The Bayes' theorem, of course, provides the link between the variables evidence, likelihood, prior, and posterior as given in Equation 7.

$$p(\omega|T,\alpha,\beta) = \frac{p(\omega|T,\beta)p(\omega|\alpha)}{p(T|\alpha)} \tag{7}$$

As α is the precision of the prior.
Integration over model parameters also allows for direct comparison of models of varying complexity by assessing the evidence function solely on training data without the need for a validation set. We'll see an illustration of how comparing polynomial models based solely on their evidence functions allows us to directly compare polynomial models of various complexities. The models with the highest evidence are typically those with intermediate complexity, or models whose complexity is just high enough to adequately explain the facts.

Results and Discussion

Crucial theoretical information is covered in the section that follows from Linear Regression (LR) to BR. The first step in data exploration is to look over the dataset and discover more information about the suggested and existing data. The sample quantity of computer products sold, as well as the two features that have been hypothesized and are in the dataset, are included in the dataset under review, as shown in Figure 2 and Figure 3. Once the dataset has been split into training and testing sets, the algorithm is given the training set in order to teach it how to predict values. Methods for regression analysis include linear regression, Bayesian regression, and neural network regression. In addition to these boosting algorithms, accuracy has been improved by applying AdaBoost and XGBoost to the dataset. The proposed method is implemented with Sklearn 0.19.1 and Python 3.6.
A variety of models are constructed, and simulations are run. Figure 4 shows the simulation's results. In order to demonstrate the usefulness of the model, Table 3 compares and compiles the best predicting results for every regression analysis for the 0175, 0541, 16/32 bit Microcontroller, 4N33m, and 8 bit Microcontroller items. Figure 5 shows that for estimating the sales of the three computer products, the BR model yields the best results. The conclusions stated above imply that the BR model is suitable for computer forecasting of retail sales. Based on the forecasting results in Figure 5, the test is designed to determine how effectively the BR forecasting models would predict the future.

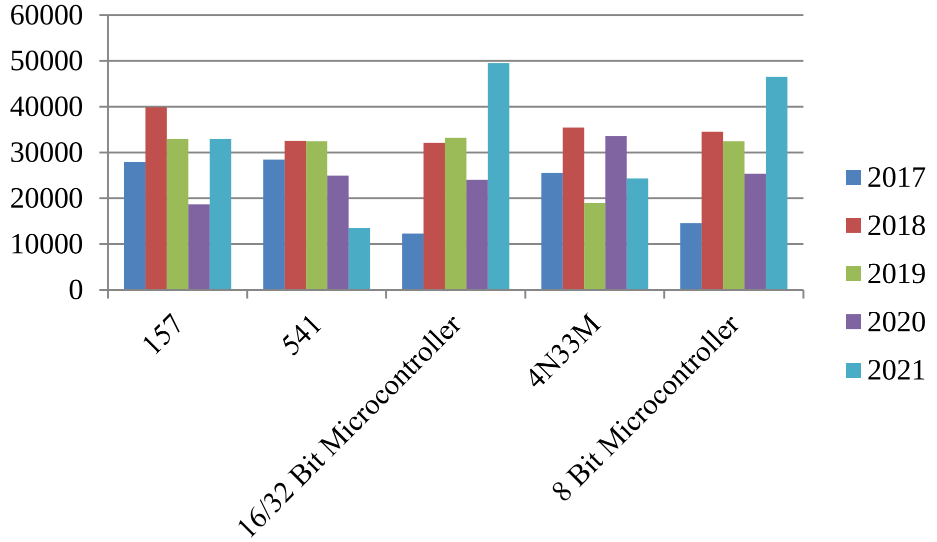

Figure 2. Sample quantity of computer products sale

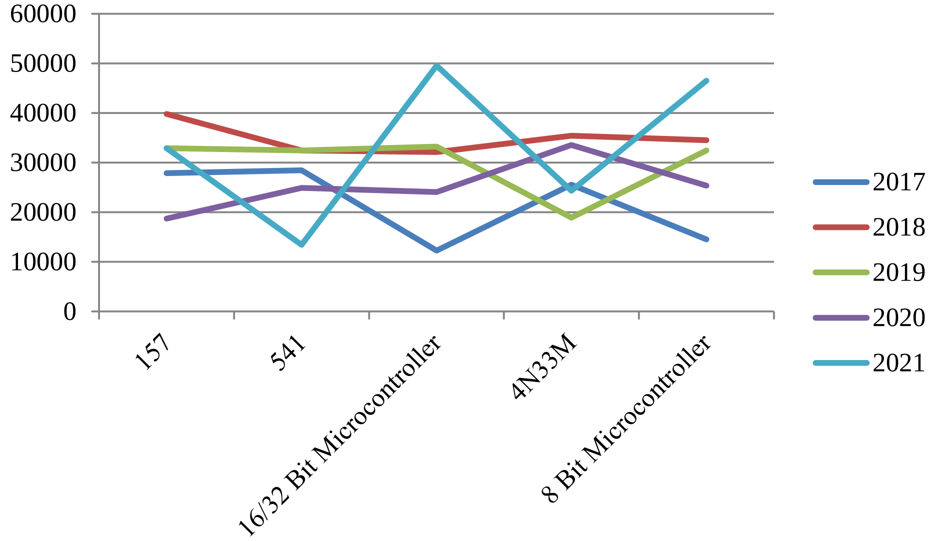

Figure 3. Graphical presentation of computer products sale

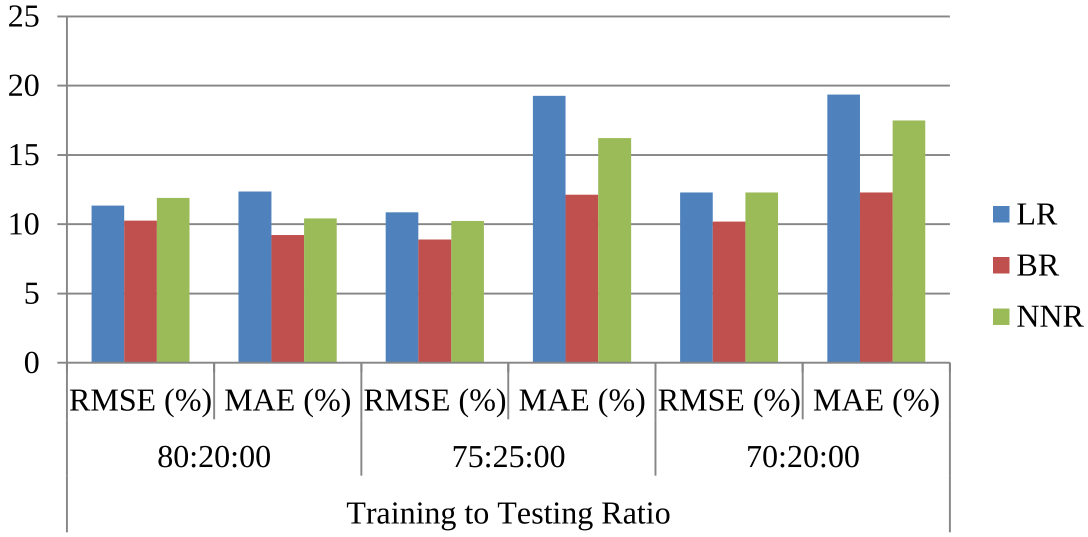

Figure 4. Simulation Results of Various Algorithms

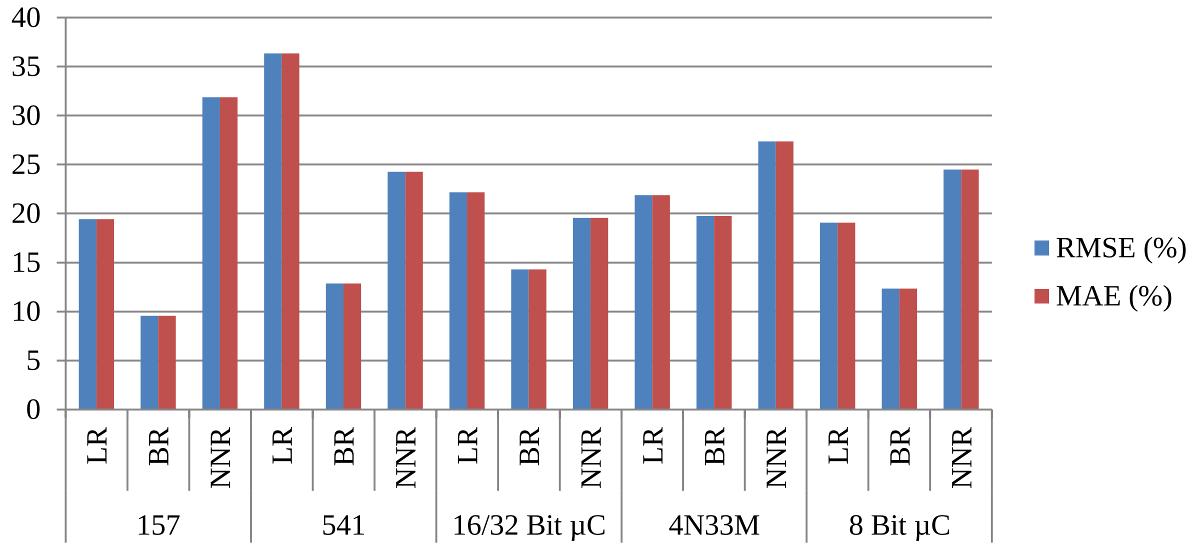

Figure 5. Simulation Results of Various Algorithms

Conclusion

The continuous technological progress necessitates the frequent replacement of computer hardware. Because of this, computer retailers must use effective sales forecasting as the basis for effective marketing and inventory management in order to compete with a wide range of rivals. This study developed machine learning algorithms for computer product sales forecasting models using linear regression, bayesian regression, and neural network regression. The empirical data were true sales figures from computer shops for the 0175, 0541, 16/32 bit Microcontroller, 4N33m, and 8 bit Microcontroller devices. A Bayesian Regression demonstrated the most promising performance when compared to other forecasting models for estimating the sales of three computer products.

References

D. Singh, "Concepts Of Inventory And Related Technical Terminologies : A Concepts Of Inventory And Related Technical Terminologies : A Literature Review," no. August, 2022.

D. Kumar, P. K. Sarangi, and R. Verma, "A systematic review of stock market prediction using machine learning and statistical techniques," Mater. Today Proc., vol. 49, no. January, pp. 3187–3191, 2020, doi: 10.1016/j.matpr.2020.11.399.

Colbert Edd, "The Role Of Demonstration Farms In Agroecological Transitions-Nexus For Scaling Up And Scaling Out Agroecology Edd Colbert MSc Agroecology Double Degree," pp. 1–90, 2020.

C. Ramsey, "Forecasting risk. Risk management will become increasingly critical in accountable care models.," Health Manag. Technol., vol. 32, no. 5, p. 22, 2011.

P. A. Sandborn, F. Mauro, and R. Knox, "A data mining based approach to electronic part obsolescence forecasting," IEEE Trans. Components Packag. Technol., vol. 30, no. 3, pp. 397–401, 2007, doi: 10.1109/TCAPT.2007.900058.

P. Sandborn, V. Prabhakar, and O. Ahmad, "Forecasting electronic part procurement lifetimes to enable the management of DMSMS obsolescence," Microelectron. Reliab., vol. 51, no. 2, pp. 392–399, 2011, doi: 10.1016/j.microrel.2010.08.005.

I. Trabelsi, M. Zolghadri, B. Zeddini, M. Barkallah, and M. Haddar, "Prediction of obsolescence degree as a function of time: A mathematical formulation," Comput. Ind., vol. 129, 2021, doi: 10.1016/j.compind.2021.103470.

C. Jennings, D. Wu, and J. Terpenny, "Forecasting obsolescence risk and product life cycle with machine learning," IEEE Trans. Components, Packag. Manuf. Technol., vol. 6, no. 9, pp. 1428–1439, 2016, doi: 10.1109/TCPMT.2016.2589206.

K. S. Moon, H. W. Lee, H. J. Kim, H. Kim, J. Kang, and W. C. Paik, "Forecasting Obsolescence of Components by Using a Clustering-Based Hybrid Machine-Learning Algorithm," Sensors, vol. 22, no. 9, pp. 1–21, 2022, doi: 10.3390/s22093244.

Y. Grichi, T. M. Dao, and Y. Beauregard, "A new approach for optimal obsolescence forecasting based on the random forest (RF) technique and meta-heuristic particle swarm optimization (PSO)," Proc. Int. Conf. Ind. Eng. Oper. Manag., vol. 2018, no. JUL, pp. 1680–1688, 2018.

M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," Science (80-.)., vol. 349, no. 6245, pp. 255–260, 2015, doi: 10.1126/science.aaa8415.

O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," Appl. Soft Comput. J., vol. 90, pp. 2005–2019, 2020, doi: 10.1016/j.asoc.2020.106181.

K. S. Moon, H. W. Lee, and H. Kim, "Adaptive Data Selection-Based Machine Learning Algorithm for Prediction of Component Obsolescence," Sensors, vol. 22, no. 20, pp. 1–21, 2022, doi: 10.3390/s22207982.