International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Machine Learning Based Prediction of Obsolescence Risk

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

Submitted: 12/11/2022 **Accepted**: 11/02/2023

Abstract*:* Rapid technological ups and downs have led to an increasingly quick jump in product overviews. Fast ups and downs enable useful life for long-life schemes, but they also provide significant issues for managing obsolescence when combined. Numerous techniques for anticipating obsolescence risk and product life cycle have been developed over time. However, gathering the data necessary for prediction is frequently difficult and independent, leading to disparities in forecasts. The goal of this paper is to develop a ML based system capable of accurately forecasting obsolescence risk and product life cycle while minimizing maintenance and upkeep of the predicting scheme in order to report these issues. Specifically, this innovative approach enables prediction of the obsolescence risk level as well as the timeframe during which a part becomes obsolete. A case study of the computer sale is presented to demonstrate the value and potency of the unique approach.

Keywords*: Stages of the life cycle, machine learning, obsolescence, maintenance for computer equipment.*

1. Introduction

Almost all industrial sectors experience some level of obsolescence, which is typically brought on by the availability of more affordable alternatives, alternatives that can achieve superior presentation and perfection, or a combination of these two. Currently, technological, practical, lawful, and sartorial obsolescence causes 3% of all electrical items in the globe each month to become obsolete. The music industry, for instance, is subject to technical obsolescence. Initial music recordings were made on vinyl, followed by eight-track cassettes, cassette tapes, and finally digital downloads. During the 1980s, cassettes were replaced by compact discs. The music business is only observing a skill transition from MP3 to music streaming facilities. Every social shift generates enormous volumes of outmoded inventory, including physical music instruments and audio players.

The difficulty of element and software obsolescence in supporting sectors has increased over the past few years as a result of the integration of electronic components and software into normally non-electronic manufactured goods. Since sensitive plans are typically more expensive than active plans, the need for practical organization grows when obsolescence occurs. Sensitive plans require higher resources, including time and materials, to handle and can result in more interruptions that lower customer satisfaction. Sensitive planning provides businesses more time to strategize and react with a practical and affordable solution [3]–[6]. An obsolescence forecasting approach is the cornerstone of a workable active obsolescence organization plan.

Many sectors can suffer greatly from obsolescence, and as a result, there is a substantial corpus of research on decision-making related to obsolescence and, more generally, on analyzing items throughout their life cycle. Cost minimization models are offered for both the product design side and the supply chain management side of obsolescence management in order to handle the economic element of obsolescence [7]–[9]. The structuring of obsolescence information has also been the subject of extensive research [10]–[12]. During the design stage of a product's life cycle, one may make more informed judgments thanks to the structuring of the information.

There are three categories of obsolescence management and decision-making techniques:

a. Reactive in the short term;

b. Reactive in the long term; and

c. Proactive.

Lifetime purchase, hobby purchase, aftermarket sources, and identification of alternative or substitute components, emulated parts, and salvaged parts are some of the most popular short-term reactive obsolescence resolution solutions [3], [13]. These tactics, however, are only short-term and may fail if the company runs out of options for obtaining the necessary components. Redesign and design refresh are better long-term choices. However, these

[1]*Department of Electronics & Telecommunication Engineering, Sinhgad College of Engineering, Pune, India.*
E-mail: ajaykdassdef@gmail.com
[2]*Department of Electronics & Telecommunication Engineering, Sinhgad College of Engineering, Pune, India.*
E-mail: sdlokhande.scoe@sinhgad.edu

solutions frequently need extensive design work and have high upfront costs.

The subcategories of obsolescence forecasting include:

a. Obsolescence risk forecasting and

b. Life cycle forecasting.

The probability that a component or other element will become obsolete is determined through forecasting for obsolescence risk. Life cycle forecasting is used to determine how long it will be before a part or component becomes obsolete. The creation date and life cycle forecast can be used by analysts to project a time frame for when a part or component will become obsolete.

Forecasting obsolescence is essential during both the product's design and manufacturing life cycles. Decisions taken during the design phase are thought to be the root of 60%–70% of the costs incurred over the course of a product's life cycle [21]. Designers can choose designs with lower risk of component obsolescence and thus lower lifetime cost effect by understanding the risk level for each component in proposed bills of materials created in the design process. Throughout a product's life cycle, obsolescence forecasting can also be used to examine anticipated component obsolescence dates and determine the best time to implement a product redesign that will get rid of the most out-of-date or at-risk-of-obsolescence parts.

To lower the risk for manufacturers and other businesses brought on by issues like quick technological advancements and brief technological life cycles, it is essential to precisely estimate the obsolescence cycle. Numerous statistical models have been investigated [4–7] for the precise prediction of the obsolescence risk and date. [6] describes a Weibull-based conditional probability method as a risk-based strategy for foretelling the obsolescence of microelectronic components. [8] is a summary of the references made to the issue of component obsolescence. Implementing a statistical model that can change quickly enough to anticipate the obsolescence cycle of hundreds of different component kinds is challenging, though. Additionally, it is challenging to compile the input parameters for various models.

Many techniques for forecasting future trends by learning large-capacity data and gathering relevant knowledge are being investigated in light of recent increases in computer performance. Outstanding outcomes are being shown by these teaching techniques, especially **Machine Learning (ML) and Deep Learning (DL)** techniques [9–12]. Different machine-learning techniques can be applied depending on the data type or application. Few researches have used these ML or DL techniques to forecast the DMSMS obsolescence cycle, to the best of the authors' knowledge. Two ML-based approaches for estimating the risk of obsolescence and the life cycle were proposed by Jennings et al. (2016) [13]. For data on the cell phone market, successful predictions were made utilizing random forests, artificial neural networks, and support vector machines.

Grichi et al. (2017, 2018) [14,15] proposed the use of a random forest and a random forest together with genetic algorithm searches for optimal parameter and feature selection for cell phone data, respectively. Trabelsi et al. (2021) [16] combined a feature selection and MLfor obsolescence prediction. As described above, ordinary learning methods attempted to increase the accuracy of prediction by combining the existing machine-learning methods and applying them to the component obsolescence data. Although it is necessary to present efficient methods and hybridize them, it is expected that the accuracy of prediction can be improved further if the characteristics of each part data are used for learning.

2. Learning Models

It's crucial to pick a ML or DL algorithm with strong computational and predictive capabilities for the dataset. A simple to understand tree-building method that can learn complex relationships is the decision tree (DT). An ensemble method, which outperforms each individual algorithm in terms of generalization, can be created by merging numerous techniques. The two well-liked ensemble techniques are boosting and bagging. In this study, a hybrid approach is suggested, and its benefits are contrasted with those of a number of widely used standard algorithms, including individual algorithms, bagging algorithms (random forest), boosting algorithms (gradient boosting), and DL techniques (DL network and recurrent neural network).

The following machine learning algorithms are briefly introduced, and their combinations are discussed for better outcomes.

1. Linear Regression (LR)
2. Decision Trees (DT)
3. Random Forests (RF)
4. Support Vector Machines (SVM)
5. Bayesian Regression (BR)

2.1 Linear Regression

An algorithm that corresponds to supervised machine learning is linear regression. On the basis of the data points for the independent variables, it attempts to apply relations that would forecast the outcome of an event. The relation is often a straight line that as closely as possible fits the various data points. A continuous form, or numerical value, is the output. The output could include things like revenue

or sales in money, the quantity of goods sold, etc. The independent variable in the example above can be one or many. In mathematics, linear regression can be written as Equation 1.

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (1)$$

Here, y = Dependent Variable, x = Independent Variable, β_0= intercept of the line, β1 = Linear regression coefficient, ε = random error. The final variable, random error ε is necessary because the best fit line also does not precisely encompass the data points.

2.2 Decision trees

In the shape of a tree structure, decision trees construct regression or classification models. It incrementally develops a decision tree to go along with the breakdown of a dataset into smaller and smaller subsets. A tree with decision nodes and leaf nodes is the end result. A decision node (such as Outlook) may have two or more branches, each of which represents a value for the attribute being checked (e.g., Sunny, Overcast, and Rainy). An option for the numerical aim is represented by a leaf node (for instance, Hours Played). A decision tree's root node is the topmost decision node and the best predictor. Numerical and categorical data can both be handled by decision trees. DTs that use features from training data split the data based on information gain from the root of the tree. The aim function to maximize this information gain in each division is the following:

$$f(parent, feature) = \dagger_p - \sum_{j=1}^{n} \frac{N_j}{N_p} \dagger_j \quad (2)$$

Here, † is the impurity indicator, N is the number of samples of the node, the subscript p denotes the parent node, the subscript j denotes the j-th child node, and n is the number of child nodes. As an impurity indicator, entropy $\dagger^E$ or Gini impurity $\dagger^G$ is widely used as given in Equations 3 and Equations 4 respectively.

$$\dagger_t^E = \sum_{i=1}^{m} \frac{N_i}{N_t} log_2 \left(\frac{N_i}{N_t}\right) \quad (3)$$

$$\dagger_t^G = \sum_{i=1}^{m} \left(\frac{N_i}{N_t}\right)^2 \quad (4)$$

Where m is the number of classes in the node t and the subscript i denotes the i^{th} class in node t. DTs have a few restrictions on the training data; thus they are prone to over-fitting. Therefore the maximum depth of the DT is usually controlled as a regulatory variable [10,11].

2.3 Random Forest

Ensemble learning's fundamental concept is pretty straightforward. The predictions from many ML algorithms should be combined in some way after training. This method typically produces forecasts that are more accurate than those of any one model. A model made up of numerous base models is known as an ensemble model.

So, ensemble learning is a process where multiple ML models are generated and combined to solve a particular problem. In general, ensemble learning is used to obtain better performance results and reduce the likelihood of selecting a poor model. Various types of ensemble learning exist, including:

1. Ensemble Learning in Sequence (Boosting)
2. Stacking, Bootstrap Aggregating,
3. Parallel Ensemble Learning (Bootstrap => Bagging),
4. Others (less commonly used).

A supervised learning system called Random Forest is built on several Decision Trees and the ensemble learning approach. Because Random Forest uses a bagging method, all computations are performed concurrently and there is no interaction between the Decision Trees as they are constructed. Both Classification and Regression tasks can be solved with RF.

The term "Random Forest" refers to the Bagging concept of randomizing data and creating several Decision Trees (Forest). Overall, it is an effective machine learning technique that reduces the drawbacks of a Decision Tree model (we will cover that later on). Additionally, as evidenced by the abundance of academic papers, technical articles, and Kaggle competitions, Random Forest is highly popular.

Let's look at the Random Forest's actual algorithm to help make things clear:

- Therefore, you wish to include K Decision Trees in our ensemble together with your original dataset D. You also have a number N; you will build a tree until each node has less than or equal to N samples (for the Regression, task N is usually equal to 5). Additionally, each node of the decision tree will include a random feature chosen from a pool of F characteristics. From these F features, the feature that will be used to split the node is chosen (for the Regression job, F is often equal to sqrt (number of features of the original dataset D).
- The rest is quite straightforward. K subsets of the data are created by Random Forest from the original dataset D. Samples that are "out-of-bag" are those that do not fit into any subgroup.
- K trees are constructed using just one subset. Additionally, each tree is constructed until each node

contains N samples or fewer. Additionally, F characteristics are chosen at random for each node.

- The node K trained models are divided into an ensemble using one of them, and the final outcome for the Regression task is generated by averaging the predictions of the individual trees.

2.4 Support Vector Machines

A multiobjective function is created by combining the geometrical characteristics of the tube. After that, the convex optimization—which has a single solution—is resolved uses the appropriate numerical optimization techniques. The hyperplane is represented by support vectors, which are training samples that are beyond the tube's boundaries. The most important examples that influence the shape of the tube in a supervised learning environment are the support vectors, and training and test data are considered to be independent and identically distributed (iid), coming from the same fixed but unknowable probability data distribution function.

2.5 Bayesian Regression

Each data point in the linear regression method of machine learning is represented by a pair of vectors: the input vector and the output vector. In its most basic form, linear regression makes the assumption that the k^{th} output vector was created through some linear combination of the k^{th} input vector's elements plus a constant term, to which Gaussian noise was then added. The best fitting linear relationship between the inputs and outputs can then be found using traditional linear regression and the data. With the aid of Bayesian linear regression, a somewhat natural mechanism can endure a lack of or poorly distributed data. It enables you to apply a prior to the noise and coefficients, allowing the priors to assume control in the absence of data. More crucially, you can inquire of Bayesian linear regression as to which aspects of its data fit it is confident in and which aspects are highly speculative (perhaps based entirely on the priors).

3. Methodology

The proposed method as shown in **Figure 1** first divides the training data into 'k' groups. As 'k' is the number which represent type of components. Then for each data in the test data, an appropriate model is selected for the prediction as follows. Suppose that (C_1, C_2… C_k) are the centroids of the groups in the partition of the training data. Given test data X, the distance between X and each of the centroids is measured as in Equations 5.

$$d(X, C_i) = \|X - C_i\| \quad i = 1,2,3, \dots, k \tag{5}$$

If the distance is minimized at i= $i^*, that\ is$ Equations 6.

$$d(X, C_{i^*}) = min_i\{d(X, C_i) \quad i = 1,2,3, \dots, k\} \tag{6}$$

Then the learning model obtained from the i^{th} cluster of the training data is applied for the prediction of X. The procedure is repeated for each of the test data as in Figure 1.

Varying machine-learning techniques produce different prediction outcomes, and no technique is particularly accurate. As a result, an ensemble method that modifies standard machine learning techniques is taken into consideration. Obsolescence is defined as the average of the three machine-learning algorithms' predicted obsolescence, LR, DT, RF, SVM and BR. The following Equation 7 is used.

$$Y^{Avg} = \frac{1}{5}\left(Y^{LR} + Y^{DT} + Y^{RF} + Y^{SVM} + Y^{BR}\right) \tag{7}$$

$Y^{LR}, Y^{DT} Y^{RF}$, $Y^{SVM} and\ Y^{BR}$ are the predictions from LR, DT, RF, SVM and BR respectively.

4. Algorithm

The proposed method's technique is outlined in the algorithm. It should be highlighted that the algorithm runs automatically, necessitating no human involvement from the processing of input data to the foretelling of obsolescence.

Proposed algorithm

- Convert the categorical data to numeric data;
- Define the features;
- Define the target data;
- Find the optimal size of clustering;
- Partition the training data;
- **For** each cluster **do**
- Find optimal hyper parameters;
- Find best model;
- **End**
- **For** each data in the test data **do**
- Find the closest cluster using Equation 6;
- **For** each learning method **do**
- Predict using the best model of the closest cluster;
- **End**
- Predict using the hybrid method Equation 7;

- **End**

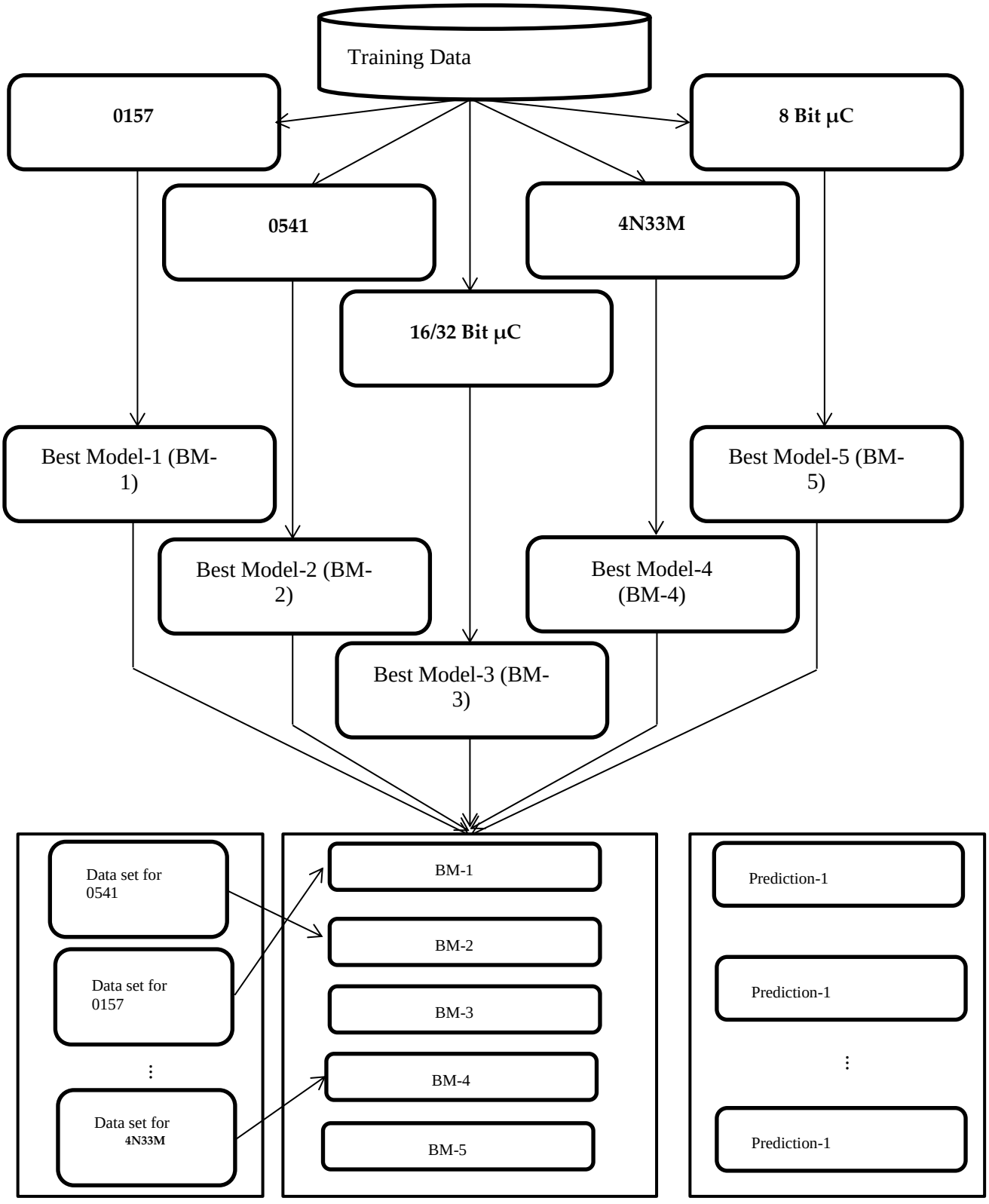

Fig 1. Proposed methodology

5. Results and Discussion

The Section 2 covers important theoretical material. Looking through the dataset and learning more about the recommended and existing data is the first step in data exploration. The dataset under consideration includes the sample number of computer products sold as well as the two attributes that have been hypothesized and are in the dataset, as illustrated in Figure 2 and Figure 3. The algorithm is given the training set once the dataset has been divided into training and testing sets in order to educate it how to predict values.

MAE and RMSE error parameters were applied in this study. **Root Mean Square Error (RMSE)** is the average error's square root. On the basis of **Equation 8**, the mean squared error is displayed.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_{obs}, i - y_{model}, i)^2}{n}} \quad (8)$$

Where, y_{obs} is actual output and y_{model} are the model output of i[th] sample. The average of all absolute errors is called the **Mean Absolute Error (MAE)**. This parameter is calculated using **Equation 9**.

$$MAE = \frac{1}{n} \sum_{i}^{n} |e_t| \quad (9)$$

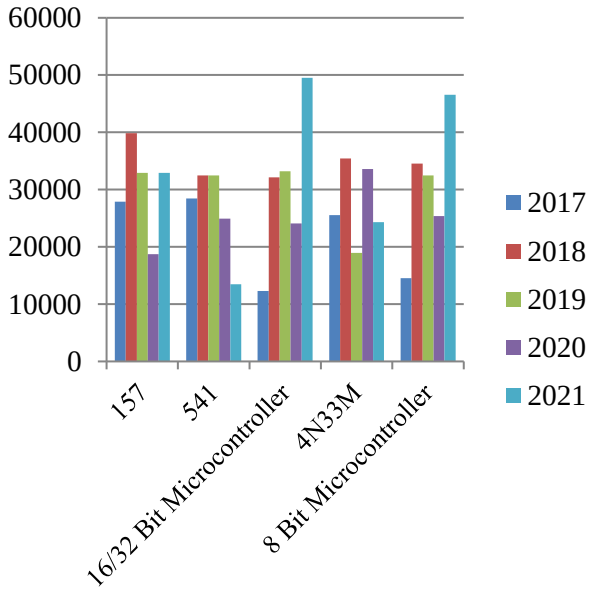

Fig. 2. Sample quantity of computer products sale

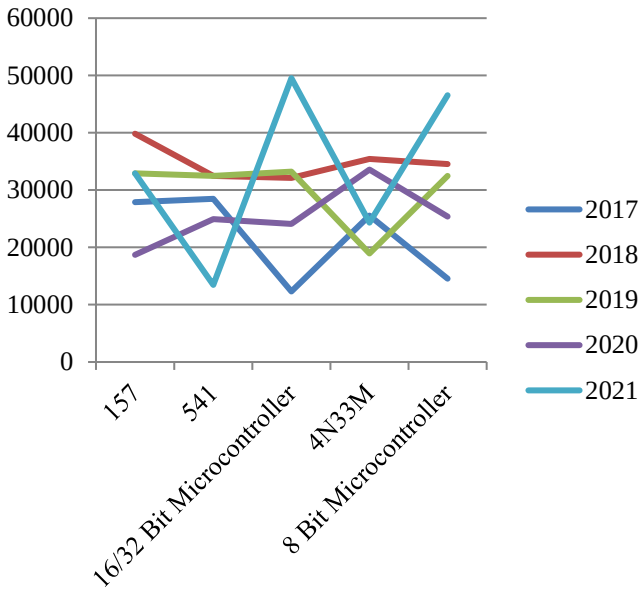

Fig. 3. Graphical presentation of computer products sale

The prediction accuracy is then compared with respect to MRE and RMSRE, two measurements. The MRE of the combine training data of all items to be predicted and separate data of items to be predicted are shown in **Table 1**. It demonstrates that training with RF over fits the provided training data and that the mistakes from the LR and DT, SVM and BR are bigger than those of the other shallow machine learning methods.

The MRE error for the test data of the combine training data of all items to be predicted and separate data of items to be predicted is shown in **Table 2**. All machine learning predictions, whether they involve clustering or not, outperform BR, and for all four categories, the four shallow machine learning methods—LR, DT, RF and SVM. In many applications, ML techniques produce accurate regression results, but in this case, the absence of data makes it impossible for them to identify the appropriate parameters.

Although the prediction of statistics with separate data of items to be predicted is superior to the test data of the combine training data of all items to be predicted, machine learning still produces the majority of the outcomes. When separate data of items to be predicted, the machine learning methods' flaws are more significant.

The LR, DT, RF, SVM and BR machine-learning approaches provide accurate forecasts for 0157, 0541, 16/32 bit µC, 4N33M and 8 bit µC, whereas the BR provide accurate predictions for 0157, 0541, 16/32 bit µC, 4N33M and 8 bit µC. Even when the same model is used, the accuracy of the models with separate data of items to be predicted is better than the accuracy of the models test data of the combine training data of all items because the data in each separate data of items to be predicted has less variation than the entire data. Therefore, the machine-

learning model trained with the separate data of items to be predicted represents the data better than a single model trained with the entire data. The hybrid technique produces good accuracy regardless of the category or the training method, which suggests that the hybrid method is dependable.

The *RMSE* errors of the training data the combine training data of all items to be predicted and separate data of items to be predicted are shown in **Table 3**. The LR, DT, SVM and BR approaches all produce greater errors than the others, and training using RF appears too overfit.

The RMSE errors of the test data the combine training data of all items to be predicted and separate data of items are shown in **Table 4**. For the given items, the predictions from all machine learning techniques combine training data of all items are superior to the naive statistical prediction. The combine training data of all items, the statistical and BR approaches produce significant inaccuracies when used to bridge rectifier diodes. In reality, for all three categories, the RMSE errors from the RF approach are substantial.

Table 1. MRE of the training for Combine training data of all items to be predicted and Separate data of items to be predicted.

Machine learning algorithms	Item	Combine training data of all items to be predicted	Separate data of items to be predicted
LR	0157	0.270	0.460
	0541	0.089	0.279
	16/32 bit µC	0.419	0.609
	4N33M	0.375	0.565
	8 bit µC	0.135	0.325
DT	0157	0.360	0.530
	0541	0.179	0.349
	16/32 bit µC	0.509	0.679
	4N33M	0.465	0.635
	8 bit µC	0.225	0.395
RF	0157	0.000	0.000
	0541	0.000	0.000
	16/32 bit µC	0.000	0.000
	4N33M	0.000	0.000
	8 bit µC	0.000	0.000
SVM	0157	0.34	0.520
	0541	0.159	0.339
	16/32 bit µC	0.489	0.669
	4N33M	0.445	0.625
	8 bit µC	0.205	0.385
BR	0157	0.36	0.160
	0541	0.179	0.179
	16/32 bit µC	0.309	0.109
	4N33M	0.265	0.165
	8 bit µC	0.125	0.125

Table 2. MRE of the training for Combine training data of all items to be predicted and Separate data of items to be predicted.

Machine learning algorithms	Item	Combine training data of all items to be predicted	Separate data of items to be predicted
LR	0157	0.170	0.300
	0541	0.110	0.119
	16/32 bit µC	0.319	0.449
	4N33M	0.275	0.405
	8 bit µC	0.195	0.165
DT	0157	0.260	0.370
	0541	0.079	0.189
	16/32 bit µC	0.409	0.519
	4N33M	0.365	0.475
	8 bit µC	0.125	0.235
RF	0157	0.000	0.000
	0541	0.000	0.000
	16/32 bit µC	0.000	0.000
	4N33M	0.000	0.000
	8 bit µC	0.000	0.000
SVM	0157	0.240	0.360

	0541	0.059	0.179
	16/32 bit µC	0.389	0.509
	4N33M	0.345	0.465
	8 bit µC	0.105	0.225
BR	0157	0.090	0.100
	0541	0.079	0.119
	16/32 bit µC	0.109	0.149
	4N33M	0.165	0.105
	8 bit µC	0.185	0.165
	16/32 bit µC	0.519	0.679
	4N33M	0.475	0.635
	8 bit µC	0.235	0.395
BR	0157	0.227	0.270
	0541	0.209	0.289
	16/32 bit µC	0.239	0.319
	4N33M	0.295	0.275
	8 bit µC	0.315	0.335

Table 3. RMSE of the training for Combine training data of all items to be predicted and Separate data of items to be predicted.

Machine learning algorithms	Item	Combine training data of all items to be predicted	Separate data of items to be predicted
LR	0157	0.317	0.470
	0541	0.240	0.289
	16/32 bit µC	0.449	0.619
	4N33M	0.405	0.575
	8 bit µC	0.325	0.335
DT	0157	0.390	0.540
	0541	0.209	0.359
	16/32 bit µC	0.539	0.689
	4N33M	0.495	0.645
	8 bit µC	0.255	0.405
RF	0157	0.240	0.370
	0541	0.220	0.389
	16/32 bit µC	0.149	0.419
	4N33M	0.231	0.375
	8 bit µC	0.276	0.435
SVM	0157	0.370	0.530
	0541	0.189	0.349

Table 4. RMSE of the testing for Combine training data of all items to be predicted and Separate data of items to be predicted.

Machine learning algorithms	Item	Combine training data of all items to be predicted	Separate data of items to be predicted
LR	0157	0.32	0.56
	0541	0.26	0.379
	16/32 bit µC	0.469	0.709
	4N33M	0.425	0.665
	8 bit µC	0.345	0.425
DT	0157	0.41	0.63
	0541	0.229	0.449
	16/32 bit µC	0.559	0.779
	4N33M	0.515	0.735
	8 bit µC	0.275	0.495
RF	0157	0.26	0.46
	0541	0.24	0.479
	16/32 bit µC	0.169	0.509
	4N33M	0.251	0.465
	8 bit µC	0.296	0.525
SVM	0157	0.39	0.62
	0541	0.209	0.439
	16/32 bit µC	0.539	0.769

	4N33M	0.495	0.725
	8 bit µC	0.255	0.485
BR	0157	0.24	0.36
	0541	0.229	0.379
	16/32 bit µC	0.259	0.409
	4N33M	0.315	0.365
	8 bit µC	0.335	0.425

6. Conclusion

Rapid technological ups and downs have caused product overviews to move up and down with increasing frequency. Fast ups and downs extend the useful life of long-term plans, but when coupled, they also provide considerable challenges for managing obsolescence. Over time, many methods have been developed for predicting the risk of obsolescence and the product life cycle. The difficulty and independence of obtaining the data needed for prediction, however, usually results in discrepancies in forecasts. The purpose of this study is to create a machine learning (ML) based system capable of projecting product life cycle and obsolescence risk accurately while minimizing maintenance and upkeep of the prediction scheme in order to report these issues. The ability to estimate the level of obsolescence risk and the period in which a part will become obsolete is specifically provided by this novel approach. In particular, this novel approach makes it possible to forecast the level of obsolescence risk and the period of time during which a part would become obsolete. A case study of the computer sale is presented to demonstrate the value and potency of the unique approach.

References

[1] D. Singh, "CONCEPTS OF INVENTORY AND RELATED TECHNICAL TERMINOLOGIES : A CONCEPTS OF INVENTORY AND RELATED TECHNICAL TERMINOLOGIES : A LITERATURE REVIEW," no. August, 2022.

[2] 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.

[3] 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.

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

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.