ARTICLE

A Hybrid WSVM-Levy Approach for Energy-Efficient Manufacturing Using Big Data and IoT

Surbhi Bhatia Khan[1,2,*], Mohammad Alojail[3], Mahesh Thyluru Ramakrishna[4] and Hemant Sharma[5]

[1]School of Science, Engineering and Environment, University of Salford, Manchester, M6 8HD, UK

[2]University Centre for Research and Development, Chandigarh University, Mohali, 140307, India

[3]Management Information System Department, College of Business Administration, King Saud University, Riyadh, 11495, Saudi Arabia

[4]Department of Computer Science and Engineering, JAIN (Deemed-to-be University), Bangalore, 560030, India

[5]Department of Computer Science and Engineering, IES College of Technology, Bhopal, 462044, India

*Corresponding Author: Surbhi Bhatia Khan. Email: s.khan138@salford.ac.uk

ABSTRACT

In Intelligent Manufacturing, Big Data and industrial information enable enterprises to closely monitor and respond to precise changes in both internal processes and external environmental factors, ensuring more informed decision-making and adaptive system management. It also promotes decision making and provides scientific analysis to enhance the efficiency of the operation, cost reduction, maximizing the process of production and so on. Various methods are employed to enhance productivity, yet achieving sustainable manufacturing remains a complex challenge that requires careful consideration. This study aims to develop a methodology for effective manufacturing sustainability by proposing a novel Hybrid Weighted Support Vector-based Lévy flight (HWS-LF) algorithm. The objective of the HWS-LF method is to improve the environmental, economic, and social aspects of manufacturing processes. In this approach, Support Vector Machines (SVM) are used to classify data points by identifying the optimal hyperplane to separate different classes, thereby supporting predictive maintenance and quality control in manufacturing. Random Forest is applied to boost efficiency, resource allocation, and production optimization. A Weighted Average Ensemble technique is employed to combine predictions from multiple models, assigning different weights to ensure an accurate system for evaluating manufacturing performance. Additionally, Lévy flight Optimization is incorporated to enhance the performance of the HWS-LF method further. The method's effectiveness is assessed using various evaluation metrics, including accuracy, precision, recall, F1-score, and specificity. Results show that the proposed HWS-LF method outperforms other state-of-the-art techniques, demonstrating superior productivity and system performance.

KEYWORDS

Sustainable manufacturing; support vector machine; random forest; levy flight optimization; weighted average ensemble

1 Introduction

The technologies, people, and procedures that make it possible to collect, store, and analyze extremely large data sets are collectively referred to as "Big Data." This can be utilized in business applications by moving the gathered information into important data using big data advancements [1]. Here are some of the applicable targets that are used for better energy management. The first and foremost objective is utilized to further develop better energy that incorporates smart energy frameworks and advancing sustainable power sources. The subsequent one is to grow the proficient versatility solutions for limiting vehicle use and greenhouse gas discharges [2]. This mostly depends on sharing data with everyone in the industry's value chain. It quickens the production arrangement cycle of an industry, employs a cutting-edge information-based strategy and reduces production complexity. These platforms will connect people, things, and systems through their applications and services [3]. By applying enormous information strategies helps to improve the strength of the examinations and helps the prescient investigation. The referenced abilities come up in numerous businesses with various factors and speeds, i.e., their requirements, in view of providers and establishment strategies [4]. In this method, the actual manufacturing process dynamically evolves over time, and during the process, a variety of unavoidable and unanticipated disturbances, such as those related to resources or jobs, may occur. There are more opportunities for production organizing to benefit from emerging technologies, such as cloud computing, big data, the Internet of Things (IoT), and others [5]. By permitting IoT solutions to find inference rules or information to incorporate into different machines or the cloud and big data permits the city to accomplish an exact understanding of an enormous arrangement of data gathered from various sources. This information generates unstructured data from observations collected through various methods [6]. Big Data Information addresses the assortment, handling, examination, and usage of immense measures of information produced inside the assembling environment. It not only involves gathering information, but also helps in removing important insights from it to drive to upgrade effectiveness and enable sustainability [7]. The investigation of big data for Intelligent and Sustainable Manufacturing presents different open doors for organizations to advance their tasks, lessen ecological effects, and make items that fulfill the needs of a consistently evolving market [8–10]. The proposed IoT-based sensor gathers temperature, dampness, accelerometer, and spinner information from the mechanical production system process while the big data handling stage handles and stores a lot of created sensor information [11–15]. This helps in recognizing the data types and information sources utilized in ML and Farming Huge Information, whether organized, semi-organized, or unstructured, as per their condition. From this perspective, this work fosters an incorporated structure of models, difficulties, and contemplations involving huge information for using big data for a smart environment. It provides such knowledge in an integrated framework, notwithstanding the way that an information-use project for brilliant urban communities requires all such information [16,17]. Big data is the main thrust behind the change of assembling into a wise and economical industry through the combination of AIoT and different learning techniques. Recently, it has developed rapidly in the assembling stage through huge volumes of information that are expected from different sensors, machines, and various cycles [18–20]. One of the foremost concerns is to expand information security protection, which handles enormous measures of assembling information and results in cyber threat issues. To address these issues, machine learning calculations can break down this information continuously to enhance creation processes, anticipate machinery failure, and further develop product quality. The inspiration to embrace these innovations and make a way towards a more shrewd, effective, and environmentally friendly assembled biological system. The above mentioned are some of the drawbacks that motivated me to do this research. In this work, a technique named Hybrid Weighted Support Vector-based Lévy flight (HWS-LF)

algorithm to enhance production sustainability is proposed. The objective of enhancing manufacturing sustainability through the integration of Big Data and IoT is to improve environmental, economic, and social aspects of manufacturing processes. This involves optimizing resource utilization, minimizing waste, and reducing environmental impact through data-driven methods. The use of Big Data and IoT aims to enhance operational efficiency, enable predictive maintenance, and promote smart decision-making. The aim is to create a more sustainable and resilient manufacturing ecosystem that balances economic growth with environmental and social responsibility.

The following are the main contributions to this work:

- In this work, we proposed a hybrid Support Vector with Lévy flight architecture to eliminate the irrelevant steps in manufacturing.
- The proposed architecture creates an effective production schedule for increased production.
- The proposed technique includes Lévy flight optimization and is used for tuning the hyperparameters in this work.

The proposed HWS-LF method introduces the practical aspects of integrating IoT devices into manufacturing environments.

The rest of the paper is organized as follows: Section 2 describes the literature survey of intelligent manufacturing using AIoT. In Section 3, the proposed methodology of this work is explained in detail. Section 4 contains the results and discussion of the experiments conducted in this work. This paper concluded in Section 5 with a few future works.

2 Related Works

The related work section is framed in such a way that certain machine learning algorithms are developed for various industrial problems and how various industrial applications are handled using big data analytics. Rosati et al. [11] evaluated a big data analytic model with the techniques of machine learning and novel IOT for Predictive Maintenance (PdM) in the industry, which is based on Decision Support System (DSS). This paper revealed how to solve PdM tasks by overcoming various challenges, such as the identification of notations in the component wears. The proposed DSS consists of a predictive model, feature extraction, cloud storage, data collection, and so on. The result of the experiment conveyed how the proposed method acts best in computation efforts as well as predictive management. In the meantime, the major characteristics, such as downtime and loss of productivity, were not clearly discussed in this research. Valaskova et al. [12] presented the deep learning technique to analyze and evaluate Big Data Analytics and Cognitive Automation in Cyber-Physical Systems. The data of five thousand six hundred respondents was examined and gathered from three major sectors: McKinsey, the World Economic Forum, and management events. The obtained data was then tested to determine the system's performance. Meanwhile, two significant aspects, production execution and robotic systems, were not analyzed in this paper.

Lăzăroiu et al. [13] evolved Robotic Wireless Sensor Networks, Smart Process Planning as well as Geospatial Big Data Management Algorithms by using Deep Learning techniques. The perceptible literature reviews of Web Science, Scopus as well as ProQuest were thoroughly executed for nearly one year, and the studied research was published. Quality examining tools like Dedoose, Mixed Method Appraisal tool, Distiller SR and so on were deployed to overcome the various challenges consisting of finding robust connections through the examined topics. Meanwhile, this paper does not analyze or inspect the production execution system as well as dynamic scheduling.

Lai et al. [14] employed reinforcement learning to control the cognitive optimal settings of AIoT industrial applications. The production schedule utilizing the cognitive learning of artificial intelligence (AI) and the entire productivity was affected by mechanical control in the Artificial Intelligence of Things (AIoT). The whole control risk and accidental control were created by the issue of cognitive evaluation in deep learning (DL). To moderate the greatest value of deep reinforcement learning based on expectation in cognitive psychology of AIoT, a new technique was proposed in this article. The cognitive computing's value overfitting was reduced by the proposed technique based on exception using the Tanh softmax policy. When compared to the cognitive learning algorithm, the proposed technique attained a total score with a six-fold increase. Utilizing the IoT and surrogate models based on AI, Chiu et al. [15] established a smart management system for aquaculture farms. To solve the issues in aquaculture farming, a fish monitoring and control mechanism based on smart IoT and containing various IoT devices for collecting data was employed. This mechanism was used to monitor various parameters, such as the quality of the water in the fish ponds, and correct them automatically. A DL model was used to forecast the growth of fish. The AIoT was used to control and manage various apparatuses in the fishpond. The proposed mechanism successfully monitored and controlled various parameters in the aquaculture fish pond.

To enhance the security of AIoT in threat identification and situational awareness, Tan et al. [16] developed a HoneyNet technique. The AIoT was mostly used in Industry 4.0 and in various fields like commercial, industries, and customers. A large number of security threats were faced by the smart devices, cloud nodes, and smart gateways in AIoT. The conventional security techniques for AIoT faced various challenges. A HoneyNet technique that contains threat identification and situational awareness for increasing the security of AIoT was proposed in this paper. The proposed technique was evaluated and the results exhibited that the proposed technique proved it superiority in terms of effectiveness and feasibility. For the implicit group recommendations, Yu et al. [17] employed a secure AIoT. Social computing applications like group recommender mechanisms benefitted from the insights provided by the AIoT. There was a need for personalized services as a large number of individuals were connected across the world. Utilizing the face two issues and the aggregation, the conventional techniques used in capturing the group-level preference were developed without the feedback of implicit preference. To solve these issues, a secure AIoT for implicit group recommendations (SAIoT-GR) was proposed in this article. The SAIoT-GR was not exposed to the complex environments that contain unclear group's preferences.

Bi et al. [18] developed an enterprise architecture (EA) based on IoT, big data analytics (BDA) and digital manufacturing (DM). In this study, Shannon entropy metric was employed to measure the complexity of systems. As a result, the method enhanced system capabilities, increased flexibility, resilience and adaptability. The views, levels and organizations were highly diversified for standardization. Also, managing system complexity and achieving long-term stability. On the other hand, EAs lacked a systematic method for justification, evaluation and comparison. In the contemporary manufacturing landscape, characterized by technological advancements, both academia and industry have increasingly focused on smart manufacturing to enhance sustainability and efficiency in production processes. To illustrate the practical application of the framework, a scenario within the additive manufacturing industry is presented. Meanwhile, the method required more processing time and large datasets for training. To further describe the existing models, a comparison was made, which is depicted in Table 1.

Table 1: Descriptions of related works

Author	Techniques	Aim	Merits	Demerits
Rosati et al. [11]	Machine learning	For Predictive Maintenance (PdM) in the industry	Higher computation efforts and predictive management	Down time and loss of productivity was not clearly discussed
Valaskova et al. [12]	Deep learning	Analyze and evaluate big data analytics and cognitive automation in cyber physical system	Higher efficiency	Production execution and robotic systems was not analyzed
Lazaroiu et al. [13]	Deep learning	Big data management	Higher accuracy	Not analyzed dynamic scheduling
Lai et al. [14]	Reinforcement learning	Control the artificial intelligence of things industrial applications	Increased efficiency	Reduced overfitting
Chiu et al. [15]	Deep learning	Establish a smart management system for aquaculture farms	Efficiently monitored and controlled various parameters in the aquaculture fish pond	The computational complexity is high
Tan et al. [16]	HoneyNet	Identification and situational awareness for increasing the security of AIoT	Increased effectiveness and feasibility	Required more computational resources
Yu et al. [17]	Secure artificial intelligence of things	To capture the group-level preference without the feedback of implicit preference	Higher efficiency and accuracy	Not applicable for complex environments
Bi et al. [18]	IoT, BDA and DM	Develop enterprise architecture	Managing system complexity and achieving long-term stability	Required more computational resopurces
Majeed et al. [20]	Deep learning	Develop big data-driven framework for sustainable and smart additive manufacturing	Enhance sustainability and efficiency in production processes	Required more processing time and large datasets for training

Research Gap

The usage of machine learning algorithms in intelligent manufacturing enhances manufacturing processes. The time consumed for the production can be reduced with these algorithms, and the production rate can be increased. The existing methods DSS, SAIoT-GR, and HoneyNet used for comparing the results obtained from the proposed method contain a few limitations, such as small network size, unclear input data, lack of details about downtime and productivity loss, and dynamic scheduling. The HWS-LF is proposed in this paper to rectify these limitations. The following subsections describe the research gaps.

Effective planning: The proposed HWS-LF creates better plans for increasing production and reducing the time and cost of production by eliminating excess or unwanted steps in manufacturing. To achieve this, the HWS-LF contains the hybrid support vector and optimization algorithm of Lévy flight.

Production time: The time for production is an important parameter that decides an industry's efficiency. The proposed HWS-LF uses the Lévy flight optimization to optimize the various stages in the proposed architecture for effective processing and to achieve a reduced production time.

Accuracy in production: Production accuracy is an important parameter in deciding the quality of a product produced in an industry. The HWS-LF architecture effectively processes the input and creates the best plan for accurate production.

Integration Challenges: Traditional manufacturing systems often operate in silos with limited connectivity and data-sharing capabilities. The integration of Big Data and IoT technologies into these systems poses significant challenges. The HWS-LF method should focus on addressing interoperability issues, developing standardized communication protocols, and creating seamless interfaces to ensure the smooth integration of data generated by IoT devices with existing manufacturing processes.

Scalability Issues in Big Data Analytics: Scalability issues may arise when dealing with large volumes of data generated by IoT devices. The HWS-LF method should develop scalable and efficient algorithms, distributed computing architectures, and cloud-based solutions to handle the massive datasets generated in real-time by sensors, actuators, and other IoT devices in manufacturing environments.

Security and Privacy Concerns: The integration of Big Data and IoT in manufacturing introduces new security and privacy challenges. Traditional manufacturing systems may not be adequately equipped to handle the increased vulnerability to cyber threats. Research should focus on developing robust cybersecurity measures, encryption techniques, and privacy-preserving algorithms to safeguard sensitive manufacturing data and ensure the secure operation of interconnected systems.

3 Proposed Methodology

Inappropriate arrangement of workflow, inexperienced labour, inaccurate work setup, handling of non-essential products, excess stations as well as lack of resources are considered as come crucial issues in sustainable manufacturing. In order to overcome such issues, a novel method is proposed via five major stages: data collection, document systemization, operation analysis, manufacturing sustainability, and performance evaluation. The detailed description of each significant stage, as mentioned in Fig. 1, is discussed as follows, and the nomenclature and symbols used in the proposed work are depicted in Appendix A.
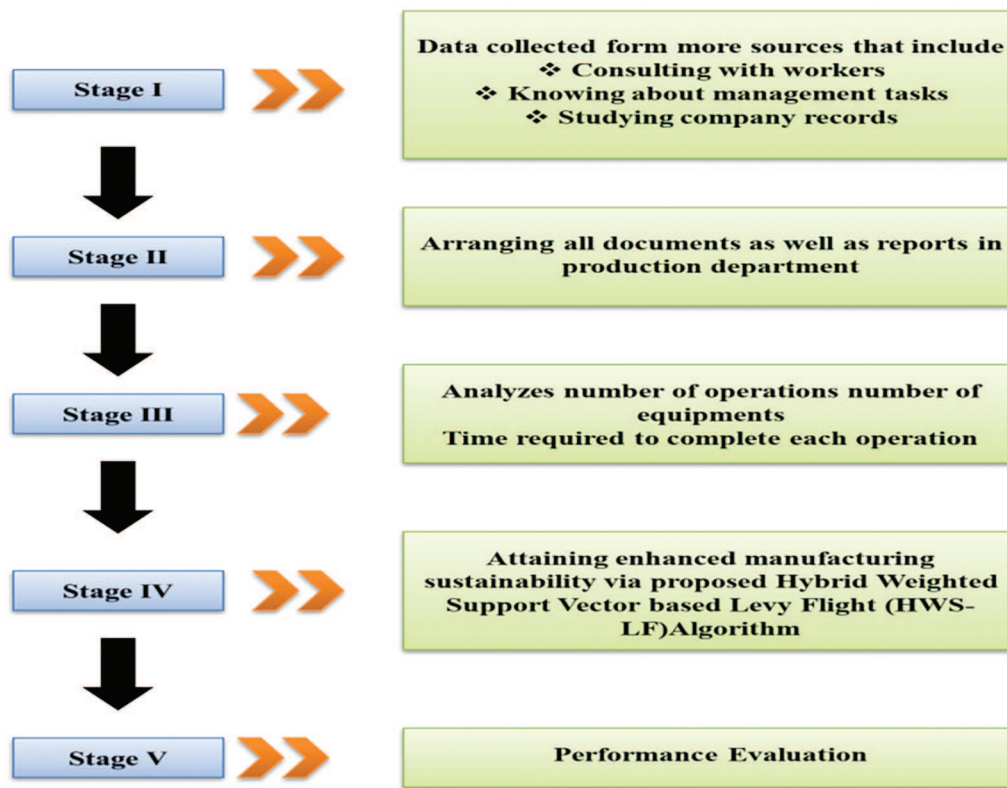
Figure 1: Overall architecture of proposed work

3.1 Collection of Data

In this phase, some of the essential qualities, which include the overall performance of all the events that take place in the manufacturing time period in the work station, are thoroughly recognized. The overall data are gathered from some reliable sources, which include consulting with the workers, knowing the history of management tasks, conducting interviews, studying company records, and so on.

3.2 Systemizing Documents

In the second stage, all the documents and reports collected in the first stage are arranged according to the production department's needs. This arrangement or classification takes place based on the machinery's requirement, the total number of technicians, the layout types, and the skills requirement.

3.3 Operation Analysis

The third stage provides the number of operators in each and every process, number of equipments used in each and every operation, the time taken to complete each operation and so on.

3.4 Enhancing Manufacturing Sustainability

In this phase, the rejection or elimination of non-productive operations takes place to improve and to achieve the maximum production level, thereby enhancing the manufacturing sustainability. This enhancement is achieved by employing a novel HWS-LF algorithm thorough examination of the work place activities. The detailed description of the novel algorithm is discussed as follows.

3.4.1 Support Vector Machine (SVM)

Different machine learning methods are used broadly by utilizing support vector machines given fundamental risk minimization standards. This helps vector machine calculation isolate the classes using the last surface and enhances the edge among the different classes using the dataset. In addition to separating the two classes, the primary objective of SVM calculations is to find an optimal hyperplane for the dataset training. Training information are presented by $\{y_j, x_j\}$. $j = 1 \ldots s$, $x_j = \{1, -1\}$, here, s demonstrates the total number of prepared samples, and the vector training, which includes two crucial classes $x_j = 1$ used in class α_1 and $x_j = -1$ in class α_2. In this if the classes are isolated persistently, it is feasible to characterize to some degree on hyper plane characterized by vector w with inclination a are been isolated by classes are displayed as follows;

$$w.y + a = 0 \tag{1}$$

In order to identify hyper plane, $x_j(w.y_j + a) \geq 1$ for $x_j = 1$ (class α_1), $x_j(w.y_j + a) \geq 1$ for $x_j = -1$ (class α_2). These two can be related in based on following:

$$x_j(w.y_j + a) - 1 \geq 0 \tag{2}$$

The ideal hyper-plane between two classes is characterized by expanding the difference between the closest classes. It helps in separating two classes by their greatest distance between help vectors and the Lagrangian multiplier can be used to define the cost function are shown as follows:

$$min \frac{1}{2} \|w\|^2 \tag{3}$$

$$K = min \frac{1}{2} \|w\|^2 - \sum_{j=1}^{s} b_j \left(x_j \left((w.y_j) + a\right) - 1\right) \tag{4}$$

Whereas, α_1 indicates the lagrangian multiplier. By introducing slack variables namely ξ_j, the constraints of non-linearly separable classes can be altered.

$$K = min \frac{1}{2} \|w\|^2 + D \sum_{j=1}^{s} \xi_j \tag{5}$$

However, D demonstrates the punishment or consistent boundaries that decide the relationship between preparing error. By utilizing a larger number of factors, feature selection issues can be taken care of by utilizing Random Forest (RF) calculation. Additionally, it is extremely effective when handling regression issues. The RF classifier effectively removes the least important features, reducing insignificant features and enhancing the proposed system's applicability.

3.4.2 Random Forest

The RF calculation is a conspicuous strategy for cooperative learning created for unsupervised learning, classification, and regression. However, this technique is most frequently used for data prediction and high-dimensional modeling. The fundamental motive of random forest is to figure out

the proper model for examining the connection between both the dependent variable and independent variable for the weight assurance of each element. This random forest algorithm contributes to the construction of numerous classification trees during the training phase, and the generation process produces the final outputs of all classification trees. In order to operate the random forest model, two principal boundaries should be taken as really important: n_{try} signifies the quantity of elements in square roots and n_{tree} indicates how many trees are needed to operate this model. The previously mentioned boundaries can be improved to limit or lessen the generation mistake. Random forest computation utilizes Gini index as an action for choosing the contamination of a given component comparable to the remaining classes. The Gini index can be figured by the probability $\sum_{i \neq j} ql = 1 - qj$ in sorting that class j. The Gini index can be communicated utilizing following condition;

$$J_L(q) = \sum_{j=1}^{i} q_j \sum_{i \neq j} ql = 1 - \sum_{j=1}^{i} q^2 j \tag{6}$$

Whereas, $j \in \{1, 2, \dots i\}$. As a result, a decision tree is designed to use a particular set of features to grow to its maximum depth. However, as with any algorithm, it should be chosen and set up carefully to solve the problem at hand. The weight capability of the support vector machine is improved by utilizing a weighted average ensemble technique which is communicated beneath.

3.4.3 Weighted Average Ensemble

Ensembles are made from a combination of weighting aspects from the average score of all the individual learners. High-performance ensembles can be achieved by analyzing and understanding the high-end parameters. The performance-tuning Nested method is introduced to achieve high-performance ensembles.

$$(W_a Y_a + W_b Y_b + W_c Y_c + \dots + W_n Y_n) \tag{7}$$

From the above equation, W denotes the accurate weight of the product which is based on reliable percentage whereas Y shows the count of vote.

3.4.4 Levy Flight Optimization

Levy Flight (LF) Optimization method is based upon wireless sensors network environment which is associated with the LF motions. The efficiency of the study can be increased by LF in uncertain conditions. In environment, LF has the ability to stimulate by most of the physical as well as natural inspired circumstances. Fruit Flies, Spider Monkeys and also Humans are classified and come under in natural category which all can practice the lane of LF style. In the other hand, scattering of the fluorescent particles is termed as physically stimulated phenomena. The essential qualities of the LF can be studied by the cooling and noise actions under the exact condition. In comparison with the Brownian Random Walks method, LF method acts much more effective to explore the unspecified large areas of search space. The step length and direction are the two major properties to be identified for executing the LF. The step length follows the distribution of the LF and the direction keep on moving forward and reaches the position of a target and this can be determined by the equal and stable distribution. These characteristics can be identified by many of the methods but Mantegna technique gives out the most reliable and effective results among all of other methods. In accordance with Mantegna's technique, the derivation is as follows:

$$S_L = \frac{V}{|U|^{1/\alpha}} \tag{8}$$

From the above equation, S_L defines as the Step length and α denotes the Levy distribution indices in which the boundary lies in $0 < \alpha \leq 2$, where V and U are

$$V \sim N\left(O, \sigma_V^2\right), U \sim N(O, \sigma_U^2) \tag{9}$$

The Standard deviation σ_V and σ_U are termed as

$$\sigma_U = \left\{ \frac{g(1+\alpha) \times \sin(\pi\alpha/2)}{g\left[\dfrac{1+\alpha}{2}\right] \times \alpha \times 2^{(\alpha-1)/2}} \right\}^{1/\alpha}, \sigma_U = 1 \tag{10}$$

Here, the gamma function g for an integer x is defines as

$$g(x) = \int_0^\infty h^{x-1} e^{-h} dh \tag{11}$$

Mathematically, the prescribed LFD method shows the Euclidean distance ED in the first two adjacent points which denotes as X_a and X_b using

$$ED(X_a, X_b) = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2} \tag{12}$$

where (X_a, Y_a) is the position coordinates of X_a and (X_b, Y_b) is the X_a coordinate position. The method begins its mechanism by altering the position of these points using

$$X_b(h+l) = LevyFlight(X_b(h), X_{lead}, HB, SB) \tag{13}$$

From the above equation, h denotes indices for the numbers of iteration. Where HB and SB are represented as the highest as well as the lowest values respectively in two dimensions of the search spaces. X_{lead} denotes the location of the agent which has the least quantity of neighbors and this terms as LF direction. And now the equation is written as

$$X_b(h+l) = SB + (HB - SB)\, rnum() \tag{14}$$

From the above equation, $rnum()$ denotes the function used to create the Random number R_n such that

$$rnum() = R_n, C_{sv} = 0.5 \tag{15}$$

Here, C_{sv} denotes the Comparative Scalar Values with R_n in each and every time of the position X_b. The prescribed method alters the position of X_a by using

$$X_a^{new}(h+l) = Levyflight(X_a(h+l), TP, SB, HB) \tag{16}$$

The above equation shows the last position for X_a. The total fitness target of the neighbors TFN in $X_a(h)$ is shown by

$$TFN = \sum_{s=1}^{TN} \frac{D(s) \times X_s}{TN} \tag{17}$$

where X_s is a nearer position of $X_a(h)$, S is index of the neighbor, TN denotes the total numbers of $X_a(h)$ neighbors. This will be the time complexity of a prescribed method making use of the Big O Notation is expressed as $O(n\log n)$ without considering about objective characteristics and number

of iterations *n*. The advantages of the described method are high ability of exploration, high level of exploitation, stability in between the exploitation as well as the exploration.

3.4.5 Proposed HWS-LF to Enhance Manufacturing Sustainability

Manufacturing sustainability helps in the improvement and development of the organization in a socially evenhanded and environmentally friendly way instead of limiting fossil fuel byproducts and pollution. Sustainability has become an integral part of business strategy, and the UN Sustainability 2030 agenda emphasizes this fact. The agenda includes various initiatives such as smart manufacturing, energy-efficient buildings, and low-impact industrialization, which are aimed at promoting sustainable practices in businesses. By adopting these initiatives, businesses can not only contribute to the environment but also improve their overall efficiency and profitability. This HWS-LF technique helps in pursuing a reasonable decision and ideal hyperparameters of HWS are exposed to an updated exhibition. A portion of the hyper boundaries utilized in the HWS assists with supporting systems that are changed and enhanced utilizing the demand LF algorithm. This approach is valuable in deciding the overall best arrangement with less step distance and more enhancement span. Also, it limits nearby optimum features and thus improves the execution of the calculation. The HWS algorithm's hyperparameters are adjusted by addressing complexity with the help of the LF algorithm. It assigns the best candidate solutions as a leading agent and separates the healthy and effective ones from the undesirable ones. Drawing the beneficial outcome using LF calculation, the hyperparameters are tuned. SVM, Random Forest, Weighted Average Ensemble, and Lévy flight Optimization are employed in enhancing manufacturing sustainability through the integration of Big Data and IoT.

3.4.6 Proposed HWS-LF to Validate Energy Consumption

To validate the efficiency of energy consumption using support vector machines (SVM) combined with Levy flight optimization, the process generally begins by gathering relevant data on energy consumption, such as historical usage patterns, environmental factors, and operational parameters. SVM, a machine learning technique, is used to classify or predict energy consumption patterns based on this data. However, the performance of SVM heavily depends on the proper tuning of its hyperparameters, which can be optimized using Levy flight-based optimization. Levy flight is a stochastic optimization algorithm inspired by the random movements of particles or organisms that can explore the search space efficiently by combining local exploitation and global exploration. By employing Levy flight optimization to fine-tune the hyperparameters of SVM, such as the penalty parameter (C) and kernel parameters (e.g., gamma), the model can achieve higher prediction accuracy for energy consumption efficiency. After the model is trained and optimized, its effectiveness can be validated by comparing predicted energy consumption with actual data using metrics like mean absolute error (MAE), root mean square error (RMSE), or R-squared. The workflow of the proposed HWS-LF algorithm is given in Fig. 2. The combination of SVM's classification strength and Levy flight's optimization ability ensures a robust and adaptive model that can generalize well, thus providing a valid approach to assess and enhance energy consumption efficiency in various systems.
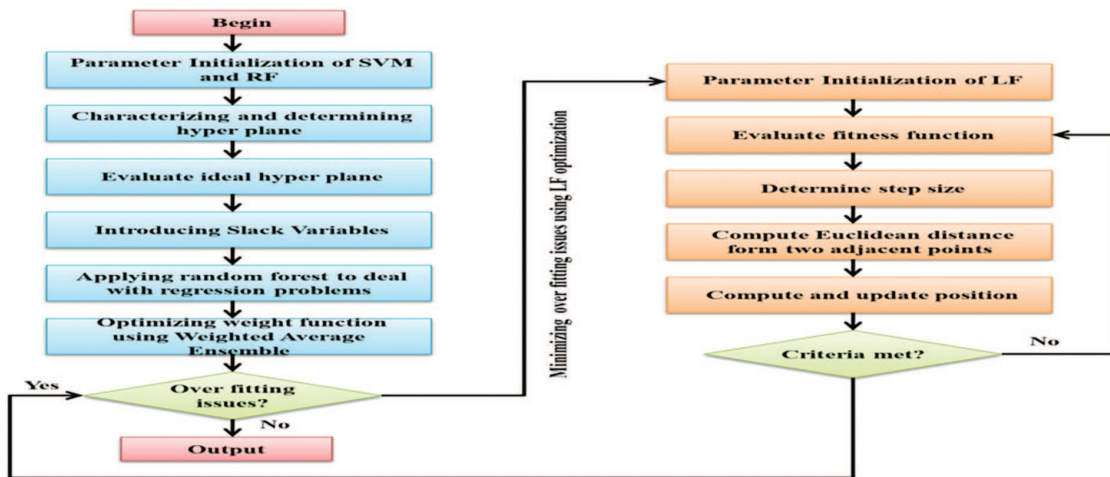
Figure 2: Flow diagram of the proposed HWS-LF algorithm

4 Results

This section evaluates the proposed HWS-LF model's sustainable manufacturing capability. It verifies sustainability with various evaluation measures and compares its outputs with existing methodologies like DSS, SAIoT-GR, and HoneyNet. The details are given below.

4.1 Experimental Setup

The method is executed on an Intel(R) core (TM) i7-8700 system with a 3.20 GHz speed processor and 32 GB memory. The OS utilized here is Ubuntu Linux 18.04.1 LTS. The simulation is performed in the MATLAB platform.

4.2 Parameter Setting

The hyperparameter is the network parameter which defines the network structure that includes, dropout, hidden units' size activation function, and weight initialization, as well as how the network parameters such as momentum, batch size, learning rate, and epochs are trained [19]. The processes of determining best settings for hyperparameters in a learning algorithm is referred to as hyperparameter tuning. Hyperparameter tuning identifies optimal values for hyperparameters to minimize the loss function, thereby enhancing results. The performance of the HWS-LF method is enhanced through applying the optimal parameters. Table 2 describes the parameters and their values. The parameter settings are given as follows:

Table 2: Parameter settings

Parameter	Values
Kernal	Radial basis function
Data range normalization	$(-1, 1)$

(Continued)

Table 2 (continued)

Parameter	Values
Cost	1
Criterion	{gini}
Maximum depth	13
Maximum features	9

4.3 Dataset Description

The data was collected using IoT devices which carry sensors like an accelerometer, gyroscopic sensor, temperature and humidity sensor, etc. The sensors are positioned on the industrial assembly line and transmit data based on a particular time (seconds). Around 20 million records were gathered within the testing time.

4.4 Evaluation Measures

Sustainable manufacturing performance of the proposed HWS-LF methodology is measured using the various metrics. The definition and details of these metrics are described below:

Manufacturing cost: The manufacturing process includes some costs which comprises all the costs consumed by the resources utilized.

Energy consumption: The manufacturing process consumes more energy. This comprises energy expended during several activities like transportation of products, operation, and maintenance.

Productivity: The productivity is the computation of economic performance. It is the difference or variation between the inputs utilized and output gained.

Resource efficiency: The effective use of the input resources will increase the sustainability such as water, raw materials, etc.

Environmental impact: The negative environmental impacts caused by the manufacturing process which causes detrimental effects to the economy. The impacts caused to the factors such as energy, water, restricted material, disposed waste, noise pollution and heat are the various kind of environmental impacts.

Operational safety: It defines the absence of risks like harm or injury to the human caused by direct and indirect ways like equipment damages and environment scenarios. For the sustainable manufacturing the safety must be high.

Accuracy: Accuracy represents the correctness of the system's predictions, reflecting the ability to correctly classify data points.

Precision: Precision measures the accuracy of positive predictions, enhancing the system's capability to avoid false positives, which is important in manufacturing to prevent unnecessary resource consumption.

Recall: Recall evaluates the system's ability to capture all relevant instances, emphasizing the importance of minimizing false negatives in manufacturing processes.

F1-score: F1-score, a harmonic mean of precision and recall, provides a balanced assessment of a system's performance, especially when there is an uneven class distribution or when both false positives and false negatives are significant concerns in the manufacturing sustainability.

Specificity: Specificity measures the system's ability to correctly identify negative instances, contributing to the reduction of false alarms and enhancing the precision of sustainability-related decisions.

4.5 Performance Analysis

In Fig. 3, we gain valuable insights into the manufacturing cost implications of the proposed HWS-LF model. Manufacturing costs encompass all the expenses involved in the production process, including both direct and indirect costs. Direct costs consist of packaging expenses, direct labor costs, and material expenditures, while indirect costs comprise fixed costs like equipment maintenance and variable costs such as repair expenses. Fig. 3a holds importance in illustrating the HWS-LF method's capacity to significantly decrease manufacturing expenses. This reduction is crucial as it directly enhances performance and efficiency. By sustaining lower manufacturing costs, the HWS-LF method facilitates the establishment of a sustainable production environment. Notably, it underscores the importance of efficient resource management, which not only enhances financial viability but also aligns with the broader objective of sustainability in manufacturing, minimizing resource utilization and fostering a more environmentally friendly and economically sound production landscape.

(a) Analysis of manufacturing cost   (b) Energy consumption
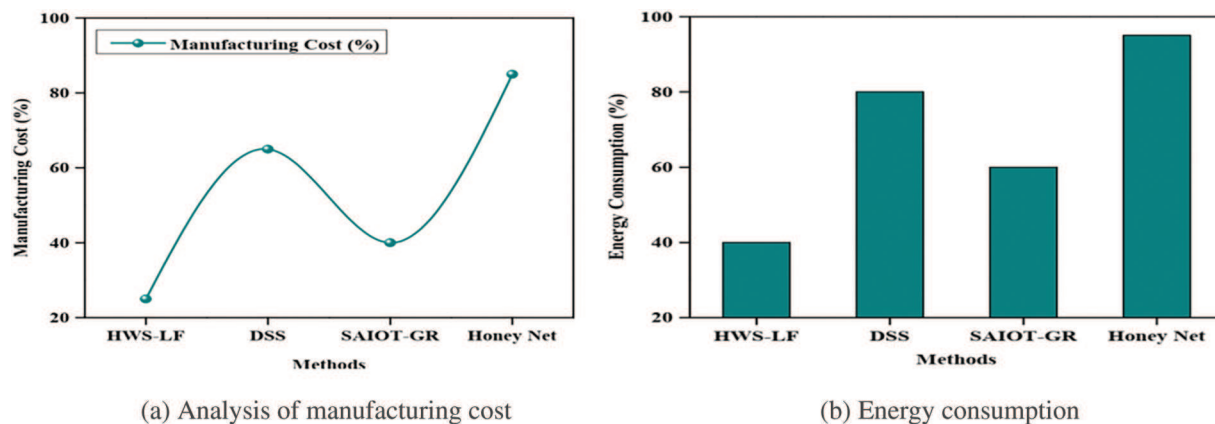
Figure 3: Manufacturing cost and energy consumption

In Fig. 3b, we delve into a critical aspect of the proposed HWS-LF model—its energy consumption, which is meticulously compared with existing methods. The energy footprint of any manufacturing process is a pivotal metric as it directly influences the sustainability and environmental impact of production. Energy consumption is distributed across various phases within the manufacturing process, making it imperative to optimize and minimize wherever possible to achieve greater sustainability goals. The figure unequivocally demonstrates that the proposed HWS-LF model excels in this regard by significantly reducing energy consumption when compared to conventional sustainable manufacturing approaches. The noteworthy finding is that the mean energy consumption for this model stands at a mere 3.78 kilowatt-hours per unit (kWh/unit). This remarkably low energy utilization is not only a testament to the model's efficiency but also underscores its potential for achieving superior performance in sustainable manufacturing.

In Fig. 4, a crucial aspect of manufacturing efficiency takes center stage—the productivity levels of different methods are subjected to thorough analysis. It is widely acknowledged that higher productivity is synonymous with enhanced sustainability in manufacturing operations, as it signifies the ability to generate more output while utilizing fewer resources. The revelation in Fig. 4a is that the proposed method, represented by the HWS-LF model, excels in terms of productivity when compared to other underlying models. It is evident that the productivity percentage of the proposed model soars to an impressive 95.7%. This figure not only signifies its superior performance but also underscores its potential to significantly enhance sustainability in manufacturing. To put this into perspective, the productivity of the proposed HWS-LF model surpasses that of the SAIoT-GR method by 12.7% and outpaces the HoneyNet methodology by an even more substantial margin of 19%. This substantial increase in productivity is a testament to the model's effectiveness and efficiency in optimizing manufacturing processes.

(a) Productivity analysis

(b) analysis of resource efficiency
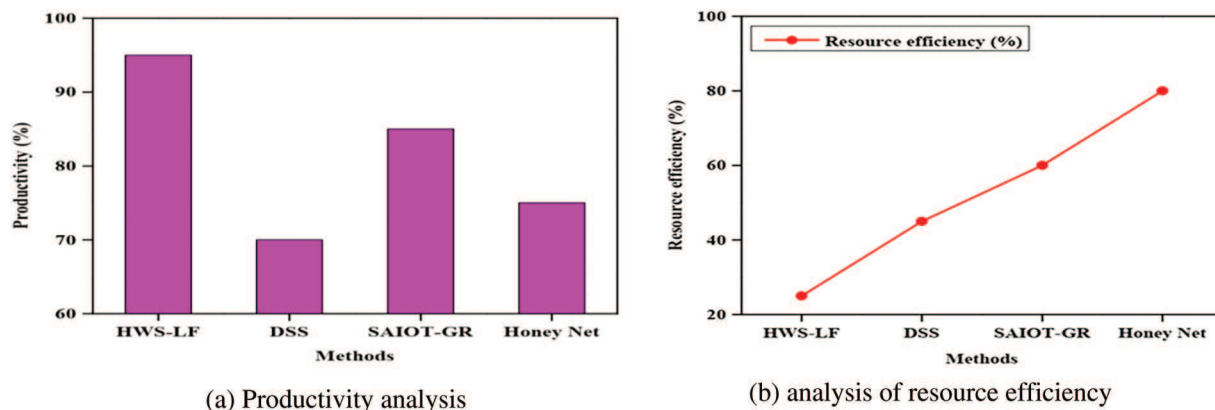
Figure 4: Productivity analysis and resource efficiency analysis

Fig. 4b provides a thorough examination of the resource effectiveness of the HWS-LF method, a pivotal aspect in the context of sustainable manufacturing. This assessment encompasses five crucial indicators aimed at quantifying resource efficiency within the manufacturing process. These indicators encompass resource usage prevention, waste generation minimization, enhanced efficiency through reduced resource utilization, process alterations, and the utilization of waste materials as valuable resources. The outcomes presented in Fig. 4b are resounding the proposed HWS-LF model consistently outperforms existing models in all five indicators. This unequivocally demonstrates its remarkable ability to optimize resource utilization, curtail waste generation, and elevate overall manufacturing efficiency. In Fig. 5a, the focus shifts to environmental impact, a paramount concern in contemporary manufacturing practices. Negative environmental impacts can lead to hazardous consequences for ecosystems and human well-being. Here, the figure serves as a stark reminder that the proposed methodology, represented by the HWS-LF model, exhibits significantly fewer adverse effects on the environment in comparison to other existing systems.

In Fig. 5b, a crucial aspect of the proposed HWS-LF methodology is meticulously evaluated: its contribution to operational safety. Operational safety is a paramount concern in any technological framework, and in this context, it signifies the method's ability to provide a secure and reliable working environment. The results presented stand out by offering a significantly higher percentage of operational safety compared to other competent models. This achievement highlights its capacity to ensure a safe and dependable operational environment, thereby extending the boundaries of

sustainability beyond what other methods can offer. Computational time refers to the average time taken by a system to complete its tasks, typically measured in seconds. This metric is critical in assessing the efficiency of a system. The analysis showcased in Fig. 6a underscores a compelling fact: the HWS-LF method boasts a distinct advantage by consuming less computational time when compared to other underlying systems. This efficiency is not just advantageous but essential for achieving optimal system performance.

(a) Environmental impacts analysis

(b) Operational safety analysis
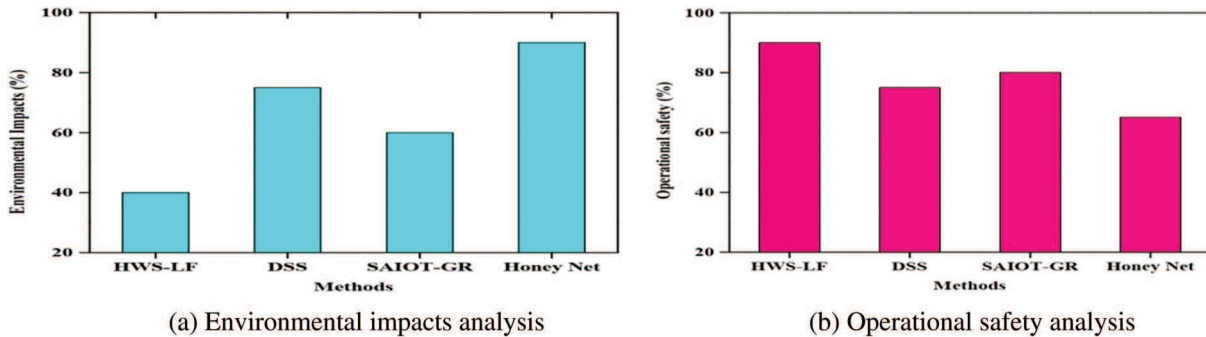
Figure 5: Environmental impacts and operational safety analysis

(a) computational time evaluation

(b) Accuracy analysis for the HWS-LF method
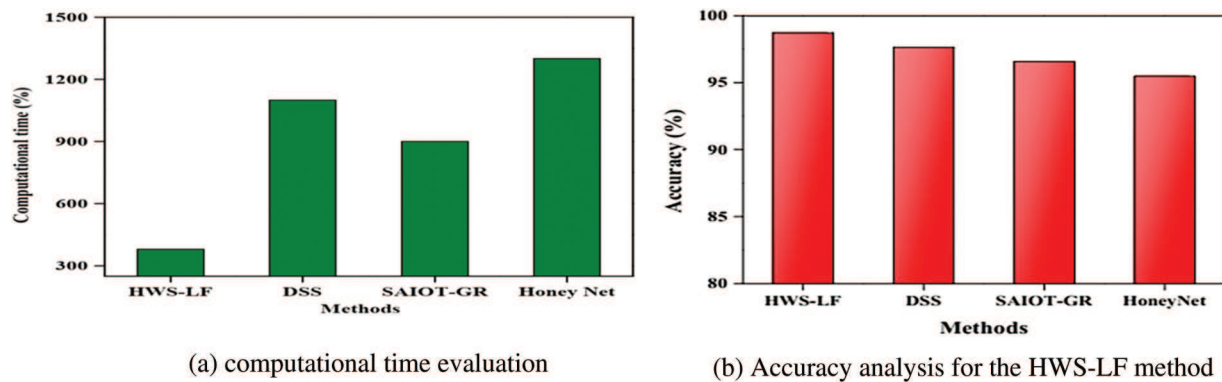
Figure 6: Computational time and accuracy analysis

Fig. 6b illustrates the accuracy performance of the HWS-LF and the existing methods for enhancing manufacturing sustainability using Bigdata and IoT. The HWS-LF method outperforms alternative approaches, achieving an accuracy of 98.72%. The DSS, SAIOT-GR and HoneyNet methods follow closely with an accuracy of 97.64%, 96.58% and 95.47%, respectively, demonstrating their effectiveness. The graphical representation indicates the better accuracy of the HWS-LF method, emphasizing its efficiency. In Fig. 7a, a precision analysis is presented to validate the performance of various methods. The HWS-LF method demonstrates better precision, achieving 97.83%. The DSS method closely follows with a precision of 96.71%, displaying its performance in positively identifying true positive instances. The SAIOT-GR and HoneyNet methods exhibit precision values of 96.57% and 95.42%, respectively, demonstrating their effectiveness in minimizing false positives.

Fig. 7b depicts the performance evaluation of HWS-LF and existing methods based on recall. The HWS-LF method exhibits better recall, achieving a value of 97.81%. The DSS method closely follows with a recall of 96.78%, emphasizing its ability to effectively identify true positive instances. The SAIOT-GR and HoneyNet methods obtain recall values of 95.62% and 94.53%, respectively,

highlighting their efficiency in capturing actual positive instances. The performance validation highlights the better recall of the HWS-LF method, emphasizing its efficiency. In Fig. 8a, a graphical representation based on the F1-score is presented to assess the performance of HWS-LF and existing methods. The HWS-LF method achieves an F1-score of 97.79%, displaying a balanced trade-off between precision and recall. The DSS method closely follows with an F1-score of 96.65%, indicating a balance between accurate positive predictions and high recall of actual positive instances. The SAIOT-GR and HoneyNet methods exhibit F1-score values of 95.56% and 94.27%, respectively. In Fig. 8b, a performance evaluation based on specificity is illustrated for HWS-LF and existing methods. The HWS-LF method displays a better specificity value of 97.75%, demonstrating its efficiency. The DSS, SAIOT-GR and HoneyNet methods exhibit specificity values of 96.67%, 95.52% and 94.48%, respectively. The performance evaluation illustrates the HWS-LF method's efficiency in accuracy, precision, recall, F1-score and specificity for enhancing manufacturing sustainability using Bigdata and IoT. Manufacturing and production are very generalized areas which are there in almost all industries. The proposed models were executed through a dataset that has three major production units, which means that the models are not too specific to analyze the performance in a single productivity area. By considering the nature of the dataset and also the computational methods that are used for the analysis, it is evident that the proposed model is a generalized approach for all the manufacturing and production industries to analyze performance and to improve it efficiently.
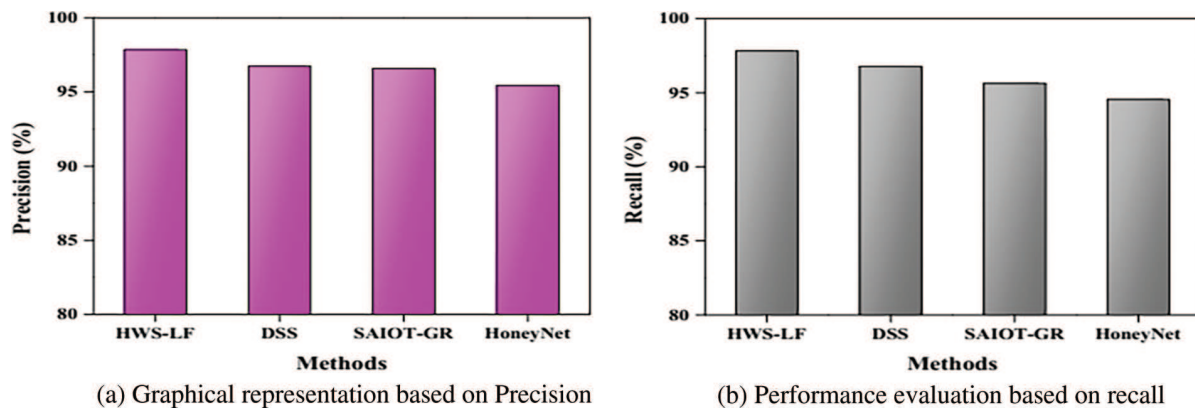

(a) Graphical representation based on Precision   (b) Performance evaluation based on recall

Figure 7: Performance evaluation based on precision and recall

(a) F1-score analysis for HWS-LF method   (b) Graphical representation based on specificity
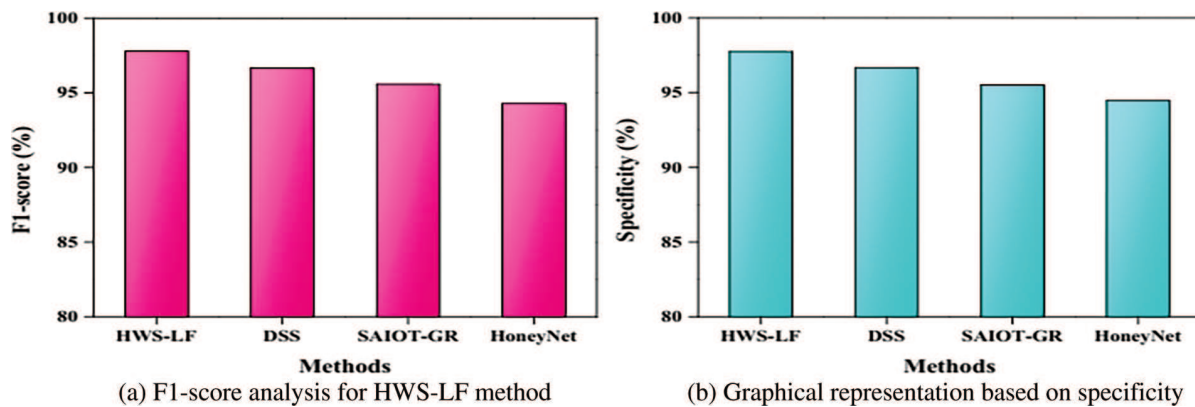
Figure 8: F1-score and specificity analysis

5 Conclusion

To address challenges that arise in the manufacturing process and enhance sustainability, this study introduces an HWS-LF technique aimed at improving efficiency by analyzing workplace activities. The analysis considers factors such as operation times and required equipment. Increased manufacturing capacity is achieved through the implementation of the HWS-LF technique, which incorporates SVM and RF algorithms, with parameters optimized using the LF algorithm. The proposed HWS-LF technique achieved superior performance metrics, including an accuracy of 98.72%, precision of 97.83%, recall of 97.81%, F1-score of 97.79%, and specificity of 97.75%, surpassing all other existing models. The goal of enhancing manufacturing sustainability through the integration of Big Data and IoT is to improve the environmental, economic, and social aspects of manufacturing processes. The proposed method was evaluated using various metrics, including manufacturing costs, energy consumption, productivity, resource efficiency, environmental impact, operational safety, and computation time. The evaluation results demonstrate that the model achieves enhanced sustainability, with increased productivity and reduced computational and energy demands, consuming an average of 3.78 kWh per unit. In the future, this method will be extended to assess its effectiveness across multiple manufacturing industries, each with unique operational contexts, challenges, environmental conditions, and requirements.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Surbhi Bhatia Khan and Mohammad Alojail; data collection: Mahesh Thyluru Ramakrishna; analysis and interpretation of results; Surbhi Bhatia Khan and Mahesh Thyluru Ramakrishna; draft manuscript preparation: Hemant Sharma and Mohammad Alojail. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

[1] M. Johnson *et al.*, "Impact of big data and artificial intelligence on industry: Developing a workforce roadmap for a data driven economy," *Glob. J. Flexible Syst. Manag.*, vol. 22, no. 3, pp. 197–217, 2021. doi: 10.1007/s40171-021-00272-y.

[2] A. Chehri, I. Fofana, and X. Yang, "Security risk modeling in smart grid critical infrastructures in the era of big data and artificial intelligence," *Sustainability*, vol. 13, no. 6, 2021, Art. no. 3196. doi: 10.3390/su13063196.

[3] H. Singh, "Big data, Industry 4.0 and cyber-physical systems integration: A smart industry context," *Mater. Today: Proc.*, vol. 46, pp. 157–162, 2021. doi: 10.1016/j.matpr.2020.07.170.

[4] Z. Shahbazi and Y. -C. Byun, "Integration of blockchain, IoT and machine learning for multistage quality control and enhancing security in smart manufacturing," *Sensors*, vol. 21, no. 4, 2021, Art. no. 1467. doi: 10.3390/s21041467.

[5] F. Qiao, J. Liu, and Y. Ma, "Industrial big-data-driven and CPS-based adaptive production scheduling for smart manufacturing," *Int. J. Prod. Res.*, vol. 59, no. 23, pp. 7139–7159, 2021. doi: 10.1080/00207543.2020.1836417.

[6] G. Sun, D. Liao, D. Zhao, Z. Xu, and H. Yu, "Live migration for multiple correlated virtual machines in cloud-based data centers," *IEEE Trans. Serv. Comput.*, vol. 11, no. 2, pp. 279–291, 2018. doi: 10.1109/TSC.2015.2477825.

[7] E. Wang, M. Zhang, B. Yang, Y. Yang, and J. Wu, "Large-scale spatiotemporal fracture data completion in sparse crowdsensing," *IEEE Trans. Mob. Comput.*, vol. 23, no. 7, pp. 7585–7601, 2024. doi: 10.1109/TMC.2023.3339089.

[8] C. Lim, K. -J. Kim, and P. P. Maglio, "Smart cities with big data: Reference models, challenges, and considerations," *Cities*, vol. 82, no. 1, pp. 86–99, 2018. doi: 10.1016/j.cities.2018.04.011.

[9] E. Wang, Y. Yang, J. Wu, W. Liu, and X. Wang, "An efficient prediction-based user recruitment for mobile crowdsensing," *IEEE Trans. Mob. Comput.*, vol. 17, no. 1, pp. 16–28, 2018. doi: 10.1109/TMC.2017.2702613.

[10] A. Cravero, S. Pardo, P. Galeas, J. López Fenner, and M. Caniupán, "Data type and data sources for agricultural big data and machine learning," *Sustainability*, vol. 14, no. 23, 2022, Art. no. 16131. doi: 10.3390/su142316131.

[11] R. Rosati *et al.*, "From knowledge-based to big data analytic model: A novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0," *J. Intell. Manuf.*, vol. 34, no. 1, 2022, Art. no. 107121. doi: 10.1007/s10845-022-01960-x.

[12] K. Valaskova, P. Ward, and L. Svabova, "Deep learning-assisted smart process planning, cognitive automation, and industrial big data analytics in sustainable cyber-physical production systems," *J. Self-Gov. Manag. Econom.*, vol. 9, no. 2, 2021, Art. no. 9. doi: 10.22381/jsme9220211.

[13] G. Lăzăroiu, M. Andronie, M. Iatagan, G. Geamănu, R. Ştefănescu and I. Dijmărescu, "Deep learning-assisted smart process planning, robotic wireless sensor networks, and geospatial big data management algorithms in the internet of manufacturing things," *ISPRS Int. J. Geo-Inf.*, vol. 11, no. 5, 2022, Art. no. 277. doi: 10.3390/ijgi11050277.

[14] Y. -H. Lai, T. -C. Wu, C. -F. Lai, L. T. Yang, and X. Zhou, "Cognitive optimal-setting control of AIoT industrial applications with deep reinforcement learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, 2021, Art. no. 21162123. doi: 10.1109/tii.2020.2986501.

[15] M. -C. Chiu, W. -M. Yan, S. A. Bhat, and N. -F. Huang, "Development of smart aquaculture farm management system using IoT and AI-based surrogate models," *J. Agric. Food Res.*, vol. 9, 2022, Art. no. 100357. doi: 10.1016/j.jafr.2022.100357.

[16] L. Tan, K. Yu, F. Ming, X. Cheng, and G. Srivastava, "Secure and resilient artificial intelligence of things: A HoneyNet approach for threat detection and situational awareness," *IEEE Consum. Electron. Mag.*, vol. 11, no. 3, 2022, Art. no. 6978. doi: 10.1109/mce.2021.3081874.

[17] K. Yu, Z. Guo, Y. Shen, W. Wang, J. C. -W. Lin and T. Sato, "Secure artificial intelligence of things for implicit group recommendations," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2698–2707, 2022. doi: 10.1109/jiot.2021.3079574.

[18] Z. Bi, G. Wang, and L. D. Xu, "A visualization platform for internet of things in manufacturing applications," *Internet Res.*, vol. 26, no. 2, 2016, Art. no. 377401. doi: 10.1108/intr-02-2014-0043.

[19] D. A. Otchere, T. O. Arbi Ganat, R. Gholami, and S. Ridha, "Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models," *J. Pet. Sci. Eng.*, vol. 200, 2021, Art. no. 108182. doi: 10.1016/j.petrol.2020.108182.

[20] A. Majeed *et al.*, "A big data-driven framework for sustainable and smart additive manufacturing," *Robotics Comput. Integr. Manuf.*, vol. 67, 2021, Art. no. 102026. doi: 10.1016/j.rcim.2020.102026.

Appendix A.

Table A1

Symbols	Description
s	Total number of prepared samples
w	Degree on hyper plane characterized by vector
a	Inclination
α_1	Lagrangian multiplier
ξ_j	Slack variables
D	Punishment or consistent boundaries that decide the relationship between preparing error
n_{try}	Quantity of elements in square roots
n_{tree}	How many trees are needed to operate this model
W	Accurate weight of the product
Y	Count of vote
S_L	Step length
α	Levy distribution
σ_V, σ_U	Standard deviation
g	Gamma function
(x_a, y_a)	Position coordinates of x_a
(x_b, y_b)	x_b coordinate position
h	Indices for the numbers of iteration
HB and SB	Highest as well as the lowest values of the search spaces
X_{lead}	Location of the agent
$rnum\,()$	Function used to create the Random number R_n
C_{sv}	Comparative scalar values
TFN	Total fitness target of the neighbors
X_S	Nearer position of $X_a\,(h)$
S	Index of the neighbor
TN	Total numbers of $X_a\,(h)$ neighbors