Alexandria Engineering Journal 88 (2024) 298–309

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

Original article

Machine learning and IoT – Based predictive maintenance approach for industrial applications

Sherien Elkateb [a,*,1], Ahmed Métwalli [b], Abdelrahman Shendy [c], Ahmed E.B. Abu-Elanien [d]

[a] Textile Engineering Department, Faculty of Engineering, Alexandria University, Egypt
[b] Electrical (Electronics and Communications) Engineering Department, Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt
[c] Computer Engineering Department, Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt
[d] Electrical Engineering Department, Faculty of Engineering, Alexandria University, Egypt

A R T I C L E I N F O

Keywords:
Failure
IoT
Knitting machines
Machine learning
Predictive maintenance

A B S T R A C T

Unplanned outage in industry due to machine failures can lead to significant production losses and increased maintenance costs. Predictive maintenance methods use the data collected from IoT-enabled devices installed in working machines to detect incipient faults and prevent major failures. In this study, a predictive maintenance system based on machine learning algorithms, specifically AdaBoost, is presented to classify different types of machines stops in real-time with application in knitting machines. The data collected from the machines include machine speeds and steps, which were pre-processed and fed into the machine learning model to classify six types of machines stops: gate stop, feeder stop, needle stop, completed roll stop, idle stop, and lycra stop. The model is trained and optimized using a combination of hyperparameter tuning and cross-validation techniques to achieve an accuracy of 92% on the test set. The results demonstrate the potential of the proposed system to accurately classify machine stops and enable timely maintenance actions; thereby, improving the overall efficiency and productivity of the textile industry.

1. Introduction

The manufacturing industry is always looking for ways to improve efficiency, reduce downtime, and minimize costs. One area where improvements can be made is in the maintenance of equipment, especially for machines used in the textile industry. Traditional maintenance practices have been based on reactive approaches, where equipment is repaired after a failure occurrence [1]. However, predictive maintenance, which relies on the use of data analytics and machine learning techniques, has emerged as a more proactive and cost-effective solution [2].

In this paper, we focus on the application of predictive maintenance for circular knitting machines in the textile industry. Circular knitting machines are used to produce knitted fabrics of various sizes, shapes, and textures. The smooth operation of these machines is critical to the quality of the final product, as well as to the efficiency of the production process. To achieve predictive maintenance for circular knitting machines, an Internet of Things (IoT) system is developed to capture machine speed and machine stops data. These data are then used to train a machine learning model to predict the time that the machine is likely to stop or malfunction. The machine stops' reasons are classified according to the proposed machine learning technique into maintenance, tool change, or operator error [3,4].

The proposed system is expected to offer several benefits over traditional maintenance practices, including reduced downtime, improved machine availability, and increased productivity.

Abbreviations: **AI,**, Artificial Intelligence; **CNN,**, Convolutional Neural Network; **DBN,**, Deep Belief Network; **FDP,**, Fault diagnosis and prognosis; **GP,**, Gaussian process; **GraphQL,**, Query Language; **GridSearchCV,**, Grid search cross-validations; **IIoT,**, Industrial IoT; **IoT,**, Internet of Things; **JSON,**, JavaScript Object Notation; **KNN,**, K-Nearest Neighbor; **LR,**, linear regression; **LSTM,**, Long Short-Term Memory; **ML,**, Machine Learning; **NB,**, Naïve bayes; **MQTT,**, Message Queuing Telemetry Transport; **PCB,**, Printed circuit board; **PWM,**, Pulse width modulation; **REST-API,**, Representational State Transfer Application Programming Interface; **RF,**, Random forest; **RNN,**, Recurrent neural network; **RUL,**, Remaining Useful Life; **SCADA,**, Supervisory control and data acquisition; **SPC,**, Statistical process control; **SVM,**, Support vector machine; **XGBoost,**, Extreme gradient boosting.

* Corresponding author.
E-mail address: K_sherien@alexu.edu.eg (S. Elkateb).
[1] ORCID: 0000-0002-5251-632X

https://doi.org/10.1016/j.aej.2023.12.065
Received 20 October 2023; Received in revised form 19 December 2023; Accepted 30 December 2023
Available online 20 January 2024

Furthermore, the system's capability to identify the root cause of machine stoppages would enable targeted maintenance interventions, minimizing the risk of unplanned downtime.

In this paper, the main contributions are summarized as follows:

- Novel system of data acquisition in textile industry is presented. In which, a circular knitting machine sensor data and stops are real-time monitored. Machine profiles detailing production settings like machine types and yarn configurations are introduced using improved database. The database allows easy management of multiple machine profiles and maintains detailed records of each shift activity.
- The contribution involves the introduction of a novel dataset obtained from a textile factory operating under standard conditions. This dataset serves as a repository for preserving the machine's historical records, facilitating the application of predictive maintenance methodologies through the categorization and classification of instances of machine stoppages.
- Specially designed Machine Learning (ML) - based model is used for classification of machine stops. Grid search cross validation is employed to optimize AdaBoost classifier for high classification accuracy for machine faults.
- A test rig is built in real circular knitting machine to realize the proposed condition monitoring system and prove its functionality.

The remaining of this paper is organized as follows. Section 2 presents the related work on predictive maintenance for industrial equipment. Section 3 introduces the IoT system architecture and the data collection process. Section 4 details the used machine learning methodology, including data preprocessing, feature selection, and model training. Section 5 presents the results of experimental work, including accuracy metrics and performance comparison with other maintenance strategies. Finally, Conclusion section highlights the significance of the proposed system and gives suggestions for future research.

2. Related work

In recent years, researchers have shown a growing interest in the development of predictive maintenance systems for circular knitting machines. Gao et al. [5] proposed a deep learning-based fault diagnosis method for circular knitting machines. Their system uses a Convolutional Neural Network (CNN) to automatically extract features from vibration signals, followed by a SoftMax classifier to classify the fault types. The experimental results demonstrated that their method achieved a promising accuracy in fault diagnosis for circular knitting machines. However, CNN needs a large amount of training data to achieve acceptable performance. Udo and Muhammad [6] introduced a predictive maintenance system for wind turbines using SCADA data, employing XGBoost and Long Short-Term Memory (LSTM) models for gearbox and generator monitoring. Statistical Process Control (SPC) assesses anomalies, demonstrating effectiveness in fault detection for six wind turbines, aiding in early intervention and cost-effective dynamic maintenance strategies. However, this system is not tested on knitting machines.

Lee et al. [7] explored recent advances in maintenance methods for manufacturing industries, emphasizing the shift from reliability improvement to flexible and customizable maintenance scheduling in the era of smart manufacturing. Singha et al. [8] delved into the integration of Artificial Intelligence (AI) and ML in the knitting industry, highlighting their transformative impact. The study emphasized the comprehensive application of these technologies across various stages: product sourcing, design, production, distribution, and sales. The incorporation of AI and ML facilitates advancements in fiber classification, thread prediction, fault identification, and dye recipe prediction; thereby, aiding predictive maintenance in knitting industry. A developed fuzzy decision-making system is developed [9]. It demonstrates its effectiveness in planning predictive maintenance through a sewing machine needle case study. Elkateb et al. [10,11] introduced an IoT and ML-based online monitoring system for knitting machines, contributing significantly to predictive maintenance. This system facilitates real-time tracking, statistical analysis, and issue resolution. Accordingly, it enables preventive maintenance and accurate productivity measurement. Surucu et al. [12] extensively reviewed recent literature on the efficacy of ML-based condition monitoring, emphasizing their significant contributions to predictive maintenance models. The study compared models using deep learning and Bayesian optimization, employing a Deep Belief Network (DBN) for feature extraction and a Gaussian process (GP) to optimize DBN hyper-parameters. Empirical results demonstrated precise machine failure time prediction, surpassing conventional ML methods. Therefore, cross-case performance comparisons are insufficient due to diverse complexities and unique contextual factors. Another investigation compared an intelligent Predictive Maintenance (PdM) system for industrial equipment, utilizing Industrial Internet of Things (IIoT), Message Queuing Telemetry Transport (MQTT), and machine learning (ML) algorithms [13]. Vibration, current, and temperature sensors collect real-time data from electrical motors to be analyzed by five ML models: k-nearest neighbor (KNN), support vector machine (SVM), random forest (RF), linear regression (LR), and naïve bayes (NB) for anomaly detection and failure prediction. The MQTT protocol enables efficient communication between sensors, gateways, and the cloud server. Random forest (RF) exhibits the highest accuracy in operational motors, optimizing maintenance schedules to minimize downtime and costs [13].

In summary, recent research in predictive maintenance for knitting machines has shown promising results in improving maintenance efficiency and reducing costs. These studies have utilized a variety of algorithms and different attributes to predict Remaining Useful Life (RUL) and to diagnose various faults in the machine. However, there is still a need for further research to develop more accurate, comprehensive, and efficient predictive maintenance systems for circular knitting machines. Diagnosis of different machine faults to achieve comprehensive predictive maintenance systems were not well covered in the literature. Moreover, applications of the developed methods on real working machines outside the laboratory environment were not well covered to prove the applicability in real conditions. The proposed work presents a predictive maintenance strategy that predicts machine stoppage and reason of stoppage (failure), so that long machine failures can be avoided. Multiple failure reasons are taken into consideration by using multiple sensing devices to have comprehensive maintenance strategy. ML-based classifier can access readings of these devices through IoT system. Its unique features, including real-time monitoring, a sophisticated ML model, and a comprehensive database, indicate a departure from conventional approaches to modern applicable approaches. The proposed predictive maintenance system is implemented on real circular knitting machine to prove its ability for practical applications. The system shows outstanding performance and high accuracy. Moreover, it holds substantial promise for minimizing downtime, enhancing machine availability, and optimizing productivity in the textile industry.

3. System setup

In this paper, we present an IoT-based system for predictive maintenance of circular knitting machines. The system is comprised of a custom-built Printed Circuit Board (PCB) with an ESP32 microcontroller and a variety of sensors, including speed sensors, and machine stop detection sensors. The machine stop detection sensors are designed to detect different types of machines stops, including lycra stop, gate stop, idle stop (which is triggered manually through human interaction), feeder stop, needle stop, and completed roll. The system is deployed on two different PAILONG circular knitting machines, with an ambient temperature of 26 degrees Celsius. The collected data is stored in mongoDB which is an open-source NoSQL database management

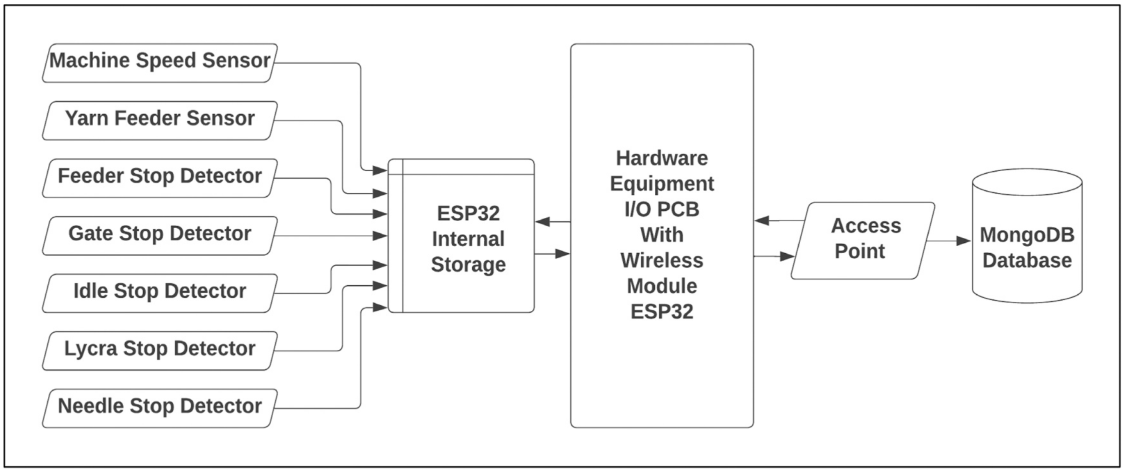

Fig. 1. System setup.

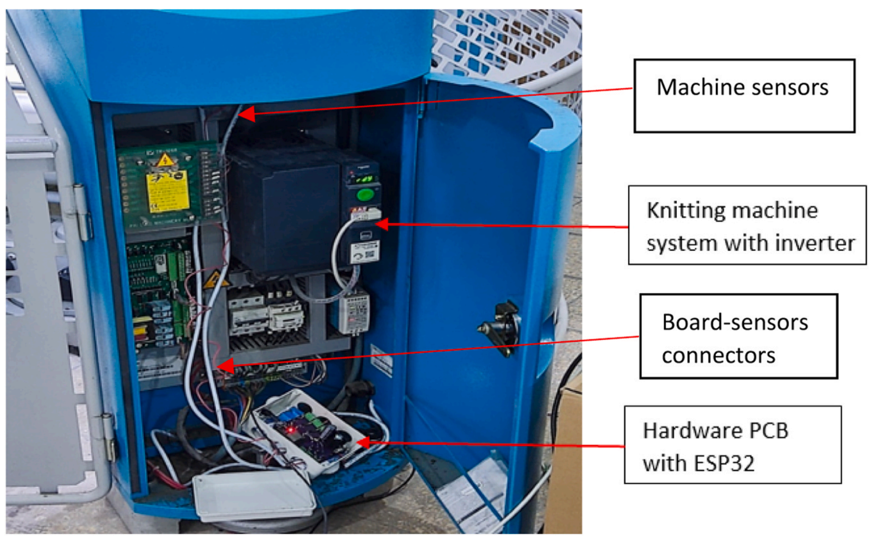

Fig. 2. System deployment.

system. The data are then used for further analysis and machine learning-based predictive maintenance. The developed system provides real-time monitoring for the circular knitting machines, enabling timely detection of potential faults and efficient maintenance scheduling. The machine built in stop detection sensors help to identify the root cause of machine stops and enable operators to take corrective actions for reducing unplanned downtime. Fig. 1 shows a block diagram of the system deployed on the knitting machine.

The machine speed and yarn feeder sensors are deployed to capture the speeds. The machine stops are used to detect any triggered event fired by feeder, gate, ideal, lycra, or needle. The pulse width modulation (PWM) signals that represent machine operation/stop are captured and delivered through the ESP32 internal storage. The ESP32 is used as a wireless module to deliver the data through the access point to the server of which the MongoDB database is used.

3.1. IoT infrastructure

The primary objective of this stage is to establish a resilient and secure environment for the monitoring system using IoT within a swiftly changing environment. This undertaking is imperative to facilitate the seamless acquisition and analysis of machine-generated logs. The experimental procedures involve the implementation of a predictive maintenance system for circular knitting machines using IoT. We built a specially designed PCB featuring ESP32 and interface with sensors to detect machine stops, records data in MongoDB through a logger system, and enables real-time monitoring for proactive maintenance. The PDB function and connection with sensors are tested and verified. Integration of GraphQL (Query Language) enhances REST-API efficiency. Moreover, the adaptable database framework improves operational efficiency across different machine types. The system ensures a secure and effective environment for utilizing machine data. The data connection to and from the server was tested for secure data traffic. The data is retrieved from the server to a personal computer via web portal or a mobile application. The retrieved data is compared with the original sensor data for assuring high data quality. Fig. 2 shows the system deployment in the main operator gate of the knitting machine. The ESP32 is mounted on a particularly designed PCB to facilitate collection of data. The output data of the different sensors are transferred to the ESP32 through the PCB available connectors. Once the system is deployed, it starts collecting the speeds and stops data.

The produced PCB has the ESP32 stack on it where the connector pins are connected through cables to the machine sensors.

In pursuit of this objective, we have judiciously decided to transfer our system architecture from its rudimentary state, as depicted in Fig. 3-a, to a more sophisticated form in order to reflect the current

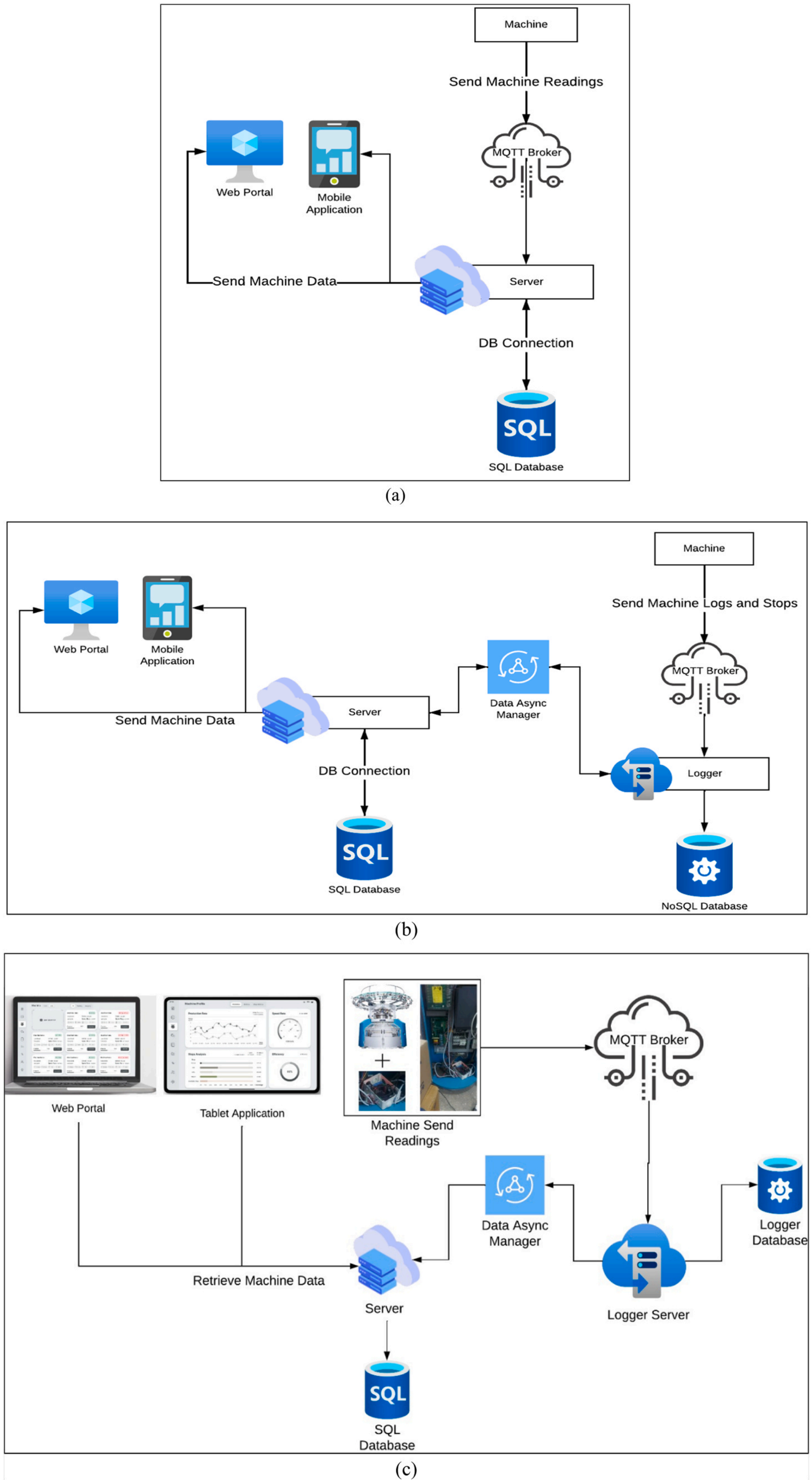

Fig. 3. The IoT system: a. IoT basic architecture, b. State of art IoT deployed system, c. system synchronization.

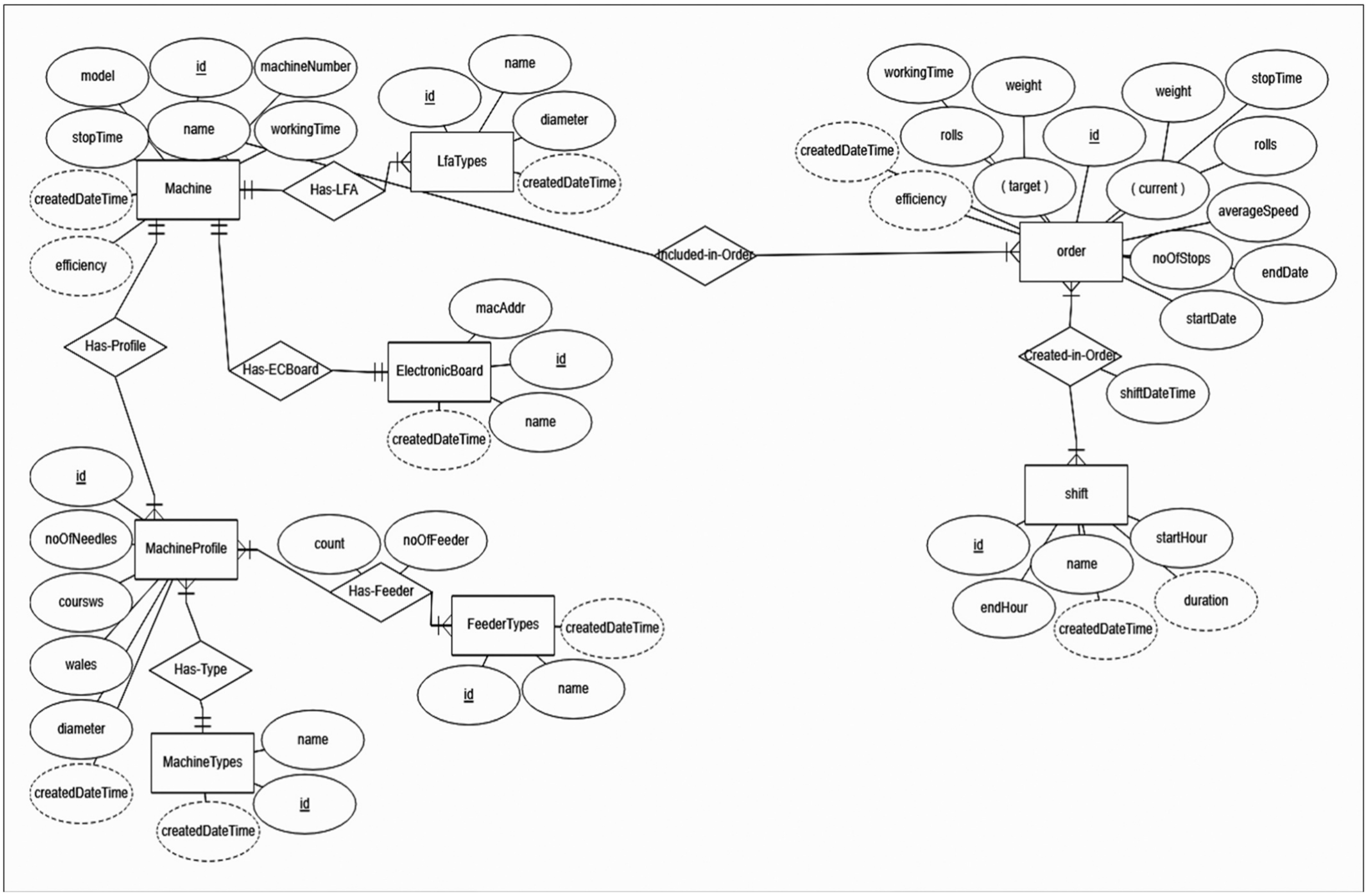

Fig. 4. Database Framework.

advancement in the field, which is reflected on quality of data collection and accuracy of incipient fault detection as shown in Fig. 3-b. In the refined Fig. 3-b architecture, we have introduced a state-of-the-art component known as the "logger." This advanced entity meticulously captures and registers every event transmitted from the machine via the Message Queuing Telemetry Transport (MQTT) broker. These critical events are promptly stored in a high-performing NoSQL database, thoughtfully chosen to optimize the speed and efficiency of our logger system; thereby, ensuring swift and accurate log processing. The logger also diligently records each event's creation date and most recent update date, enabling the construction of a precise chronological graph of events. Fig. 3-c shows the system integration and synchronization. This innovative capability empowers us to precisely reconstruct the machine's operational state during specific time intervals, facilitating comprehensive analysis and precise identification of stoppages and their underlying causes. Moreover, the logger adeptly eliminates redundant stoppages and accurately calculates and manages stoppage durations. For instance, in cases where an ongoing needle break stoppage coincides with another concurrent stoppage, the stop duration manager will diligently calculate the duration of the needle stoppage while disregarding the concurrent stoppage until the needle break is resolved. Additionally, the stop duration manager proficiently handles manual stoppages initiated by workers. It accurately computes the duration of these interruptions and integrates them into the optimal machine runtime. Furthermore, when a reason is assigned to a designated stop duration, the stoppage duration supervisor adeptly oversees and allocates it correspondingly. The logger, then, can be configured to trigger various actions, including real-time notifications or emails to managers, ensuring immediate awareness of machine stoppages and their root causes.

By segregating the Representational State Transfer Application Programming Interface (REST-API) from the machine logs and seamlessly integrating it with our user-facing platforms such as web portals or mobile applications; we enhance the speed and responsiveness of our REST-API services. This separation optimizes system resources by alleviating the burden on the machine logs, consequently laying the groundwork for real-time data transmission instead of transmission at fixed intervals.

In order to optimize REST-API performance, we have embraced cutting-edge GraphQL technology, surpassing the limitations of conventional REST-APIs. This groundbreaking transition provides us with the capabilities to overcome the inherent inefficiencies associated with server requests, particularly when faced with complex data retrieval scenarios. Traditional REST-APIs present suboptimal options when accessing database tables with numerous records and columns. One possibility requires separating the requests for each platform, resulting in redundant efforts. The alternative mandates retrieving all columns within a single request, overburdening the computer/mobile processor and causing unwarranted delays in data display. After meticulous research and evaluation, GraphQL has supplied us with a solution to this challenge. By leveraging GraphQL, each platform can precisely specify the required columns through a solitary request; accordingly, efficiency is significantly enhanced, and data retrieval processes are notably expedited.

Within the Database framework of Fig. 4, our primary focus centers on augmenting the dynamism and adaptability of the database infrastructure to accommodate all machine types without exceptions. To achieve this goal, we have introduced machine profiles that encapsulate the indispensable production settings, encompassing machine types, feeder configurations, yarn counts, and the allocation of feeders for each yarn type. Moreover, our optimized database structure enables seamless management of multiple profiles for each machine, facilitating effortless transitions between different operational configurations. Furthermore, our refined database structure empowers us to meticulously track and maintain a comprehensive historical record of each shift's activities. This pivotal functionality facilitates the seamless distribution of orders across multiple shifts or machines, facilitating automated handling and improving overall operational efficiency. By undertaking these transformative steps in our monitoring system using IoT, we are poised to establish a secure and highly efficient environment, enabling us to capture and leverage machine data effectively while enhancing operational performance and decision-making processes.

3.2. Environmental conditions

From the perspective of a knitting machine, the environmental conditions may have a significant impact on the quality of the knitted products and the performance of the machine. Some of the key environmental conditions that are important to consider for a knitting machine include:

Temperature: The temperature of the environment can affect the viscosity of the lubricants used in the machine and also impact the thermal expansion and contraction of the machine components.

Humidity: The humidity of the environment can affect the static electricity generated during the knitting process, which can cause yarn to stick together and impact the quality of the final product.

Air quality: The air quality of the environment can impact the performance of the machine and the quality of the knitted product. Poor air quality can cause dust and other contaminants to accumulate on the machine, which can impact its performance. Good ventilation and air filtration are important to maintain a clean and healthy environment for the knitting machine.

Lighting: Adequate lighting is important for the operator to be able to monitor the performance of the knitting machine and to detect any defects in the knitted products.

Noise level: Knitting machines can generate noise during operation, which can be harmful to the operator's hearing and impact their ability to detect any issues with the machine.

Vibrations: The vibrations generated by the knitting machine can impact its performance and affect the quality of the knitted products. Adequate vibration isolation and damping measures are important to reduce the impact of vibrations on the machine and its environment.

Overall, maintaining a controlled and stable environment with suitable environmental conditions is essential to ensure the optimal performance of a knitting machine and to produce high-quality knitted products.

Several studies have investigated the impact of working conditions, including temperature, humidity, lighting, and noise, on the performance and quality of knitting machines. For instance, Seitablaiev et al. [14] explored factors affecting thermal comfort and indoor air quality, providing a comprehensive understanding of environmental pollutants and their health implications. This study aligns with those findings, aiming to optimize indoor conditions for user comfort by analyzing bioclimatic parameters, thermal comfort, and air quality, with a focus on the consequential impact of pollutants on health and productivity. Rassel and Hoque [15] emphasized the significant influence of the working environment on yarn breakage, affecting both production efficiency and fabric quality. Addressing these factors is essential for enhancing competitiveness, reducing costs, and minimizing the environmental impact associated with increased power consumption. Moazzem et al. [16] conducted a comprehensive review of environmental impacts throughout the textile supply chain, emphasizing the substantial contributions of textile production and usage stages. Future research recommendations focus on evaluating fiber mixing, recycling options, and technical textiles. This study explores gaps in the environmental impact of the textile industry, emphasizing the evaluation of smart textiles, landfill effects based on fibers, and overall textile consumption for sustainable practices and informed policy development. Kohli and Dua [17] explored essential elements of Indoor Environmental Quality (IEQ), covering building parameters, lighting, noise, humidity, and acoustic quality. Emphasizing their impact on comfort, health, and productivity, the discussion explores thermal comfort, Sick Building Syndrome, and air temperature. Tailored recommendations for special needs individuals enhance health, wellbeing, comfort, and safety. IEQ significantly influences occupants' health, underscoring its importance for exposed groups. Overall goals encompass minimizing building-related health issues and ensuring high-quality indoor environments. Another study aimed to assess occupational noise exposure in the textile industry. An initial anonymous survey was conducted, followed by evaluating factory workers' noise exposure levels and selecting ear protectors based on legal criteria. Results showed that most jobs exceeded legal noise limits, with no workers using hearing protection. Symptomatic issues, like tinnitus, and reduced hearing equity correlated with years of exposure [18]. Hameed et al. [19] studied the effect of high noise levels in the knitting department pose a risk of Noise-Induced Hearing Loss (NIHL). Workers, despite experiencing symptoms like tinnitus, often neglect hearing loss complaints. Factors include age, duration of noise exposure, health conditions, and limited worker awareness. Table 1 outlines crucial environmental factors essential for the optimal operation of a knitting machine.

Table 1
The range of environmental parameters.

Parameter	Range
Temperature	20 °C to 27 °C (68°F to 81°F)
Humidity	45% to 55%
Airflow and Ventilation	6 to 10 air changes per hour
Lighting	500 to 750 lux
Noise Level	Below 85 decibels (dB)
Particulate Matter (PM)	Below 10 mg/m^3

3.3. Data collection

The data collected through MongoDB as shown in Fig. 1, are stored in JavaScript Object Notation (JSON) format. Then, the data are converted to tabular form where the preliminary data shape acquired is made of events and logs with a shape of 81000 records from 9 attributes. The events represent the time of which any machine stop is fired. On the other hand, the logs represent each record of the working machine. The records were acquired during a period of 48 days. It is formed from the machine speeds and machine stops where the attributes are the following:

Created At: A timeseries attribute which stores a record's starting time information.

Stitch Length: Stitch length is determined by the needle pitch, which is the distance between two adjacent needles in a circular knitting machine.

Yarn Feeder Speed: Yarn feeder sensors are capable of measuring yarn belt at high speeds, typically ranging from several hundred to several thousand measurements per minute. The speed is measured in revolution per minute.

Production Rate: The amount of fabric produced within a given time period, usually expressed in terms of weight or length. It is a measure of the efficiency and productivity of the machine and is influenced by various factors such as machine speed, stitch density, yarn type, and machine settings. It is measured in kilograms per hour.

Machine Progress: It refers to the movement of the knitting elements, such as needles and sinkers, as well as the fabric, around the cylinder or the needle bed. It indicates the position of the knitting elements and the fabric relative to the machine cycle and is an important parameter for controlling the knitting process and ensuring consistent

Table 2
Statistical distribution.

	Stich Lengh	Yarn Feeder	Production Rate	Progress	Machine Speed	Total Turn per Roll	Event Duration
mean	1.993838	896.599572	4.833878	50.420655	14.005796	1613.460960	84.548607
std	0.426776	224.270651	1.209121	28.861632	3.488458	923.572234	3666.011328
min	0.000000	0.000000	0.000000	0.093750	0.000000	3.000000	0.000000
25%	2.001149	865.000000	4.663514	25.656250	13.000000	821.000000	0.000000
50%	2.145279	866.000000	4.668905	50.437500	13.000000	1614.000000	0.000000
75%	2.150113	906.000000	4.884559	75.562500	16.000000	2418.000000	0.000000
max	2.283511	1320.000000	7.116576	100.000000	21.000000	3200.000000	396659.205

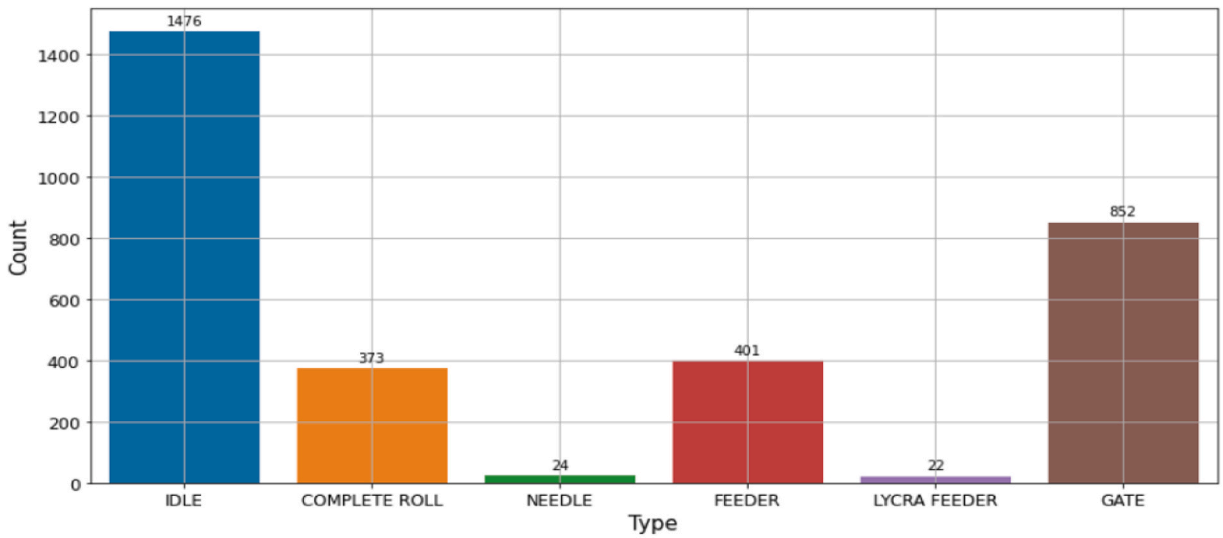

Fig. 5. Count of stop's types.

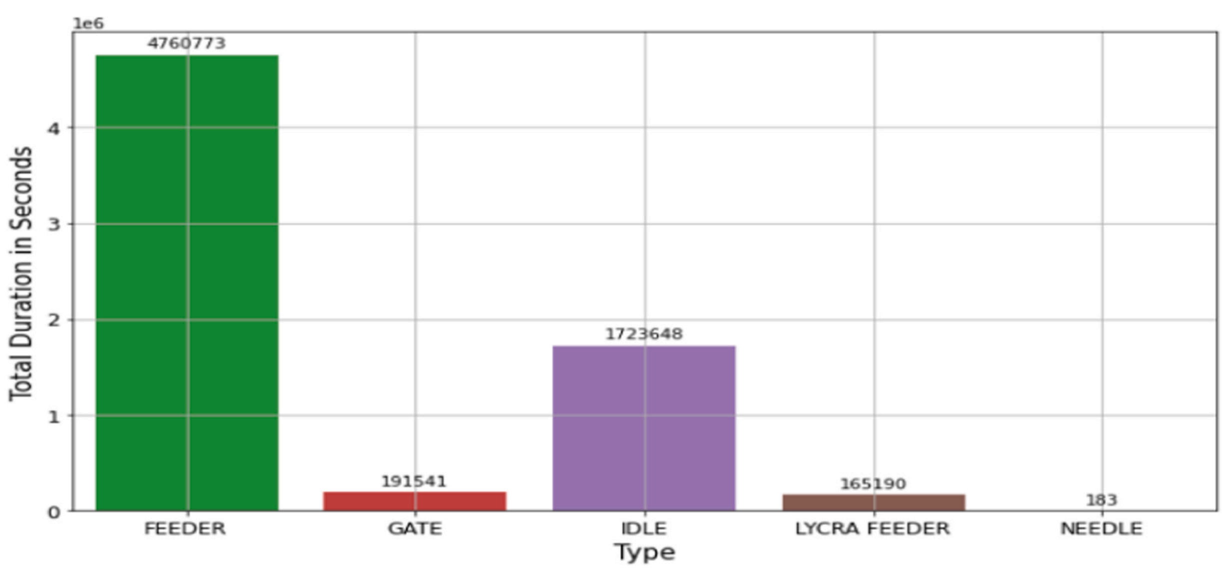

Fig. 6. Total seconds for each stop.

and uniform stitch formation. The machine progress can be measured in terms of degrees or fractions of a revolution of the cylinder or the needle bed, or in terms of machine cycle time. For example, in a single-cylinder circular knitting machine, one complete revolution of the cylinder corresponds to 360 degrees of machine progress, while in a double-cylinder machine, one complete cycle of the cylinder and dial needles corresponds to 720 degrees of machine progress.

Machine Speed: The speed of the roll turns which is captured by the proximity sensor of the machine. It is measured in revolution per minute.

Total Turns per Roll: The total turn per roll is the accumulation of the roll revolutions per minute starting from placing the roll until the machine completes 3200 turns.

Updated At: A timeseries attribute which stores a record's ending time information.

Event Duration: Event duration is the time of which the event

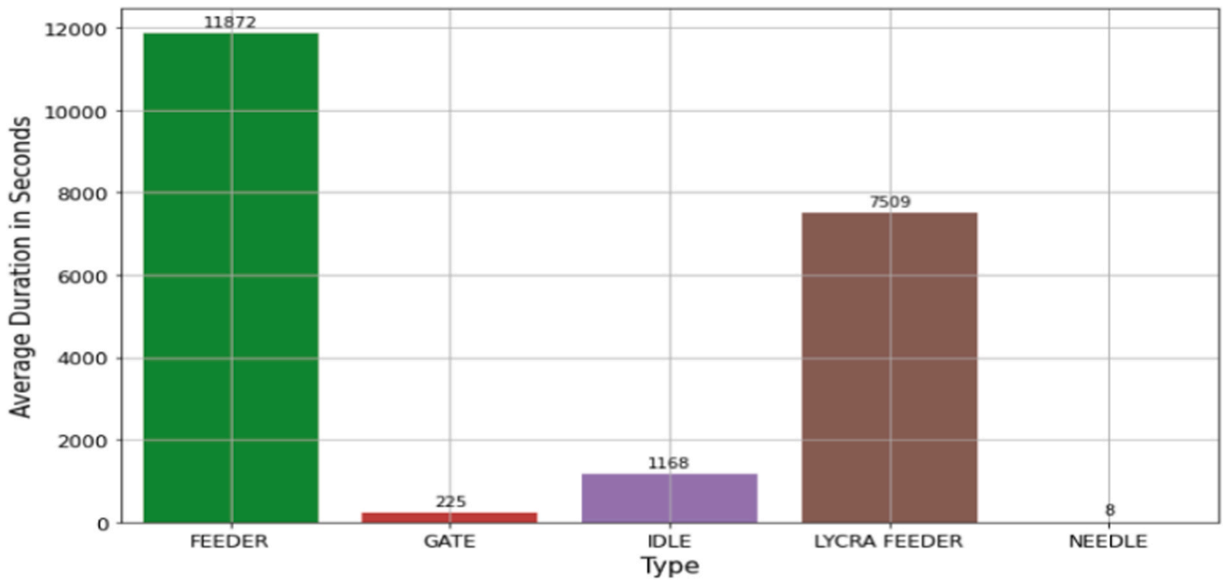

Fig. 7. Average duration for each stop.

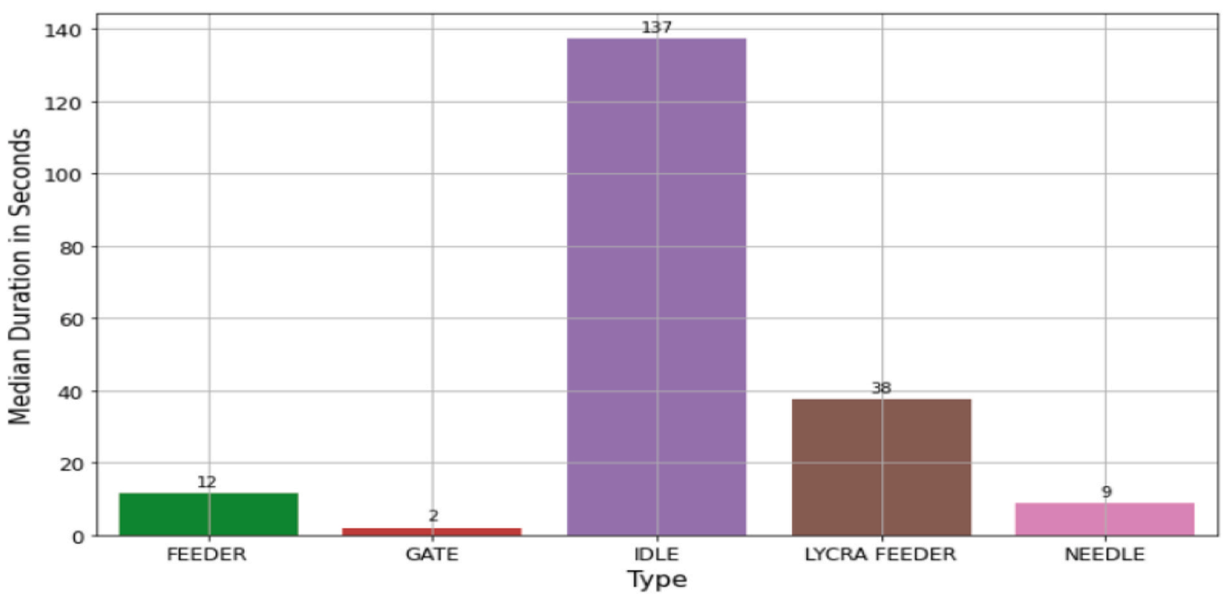

Fig. 8. Median event duration in seconds for each stop.

occurred. It is the subtraction of the ‘Updated At’ and the ‘Created At’ attributes.

Stoppage Type: The machine stops are categorized as follow: gate, needle, feeder, Lycra, ideal and completed roll.

Table 2 shows the statistical distribution of the data. Figs. 5–8 show some insights about the data. Fig. 5 shows the count plot of each stop, while Fig. 6 depicts the sum of all event durations for each stop. Fig. 7 portrays the average duration for each stop, while the median of event duration for each stop type is presented in Fig. 8. According to Fig. 5, during the 48 days, the most frequent stop occurred was idle stop as it counts 1476 cases. Gate, feeder, complete roll, needle, and Lycra stops are counted as 852, 401, 373, 24 and 22 cases respectively. The median is calculated as it is a robust statistical measure that is less sensitive to extreme values or outliers in the data set. The median values for feeder, gate, ideal, Lycra and needle based on collected data are 12,2137,38 and 9, respectively.

4. Classification scheme

To enable predictive maintenance of circular knitting machines, a classification scheme is developed to categorize machine stop events and identify the root cause of machine failures. The classification scheme is based on the types of machines stop detected by the sensors in the IoT system. The machine stops are classified into different categories based on the component causing the stop and the type of stop (such as whether it is a planned or unplanned stop). The proposed classification scheme allows for a more precise diagnosis of the problems encountered during machine operation, which is crucial for efficient and timely maintenance [1,2]. In the following subsections, we will describe a flow chart of the proposed classification scheme and how it is integrated into the overall IoT system for predictive maintenance of circular knitting machines. The classification scheme is shown in Fig. 9.

The data are retrieved from the database, subsequently, the data preprocessing phase is carried out, where the data are normalized, then

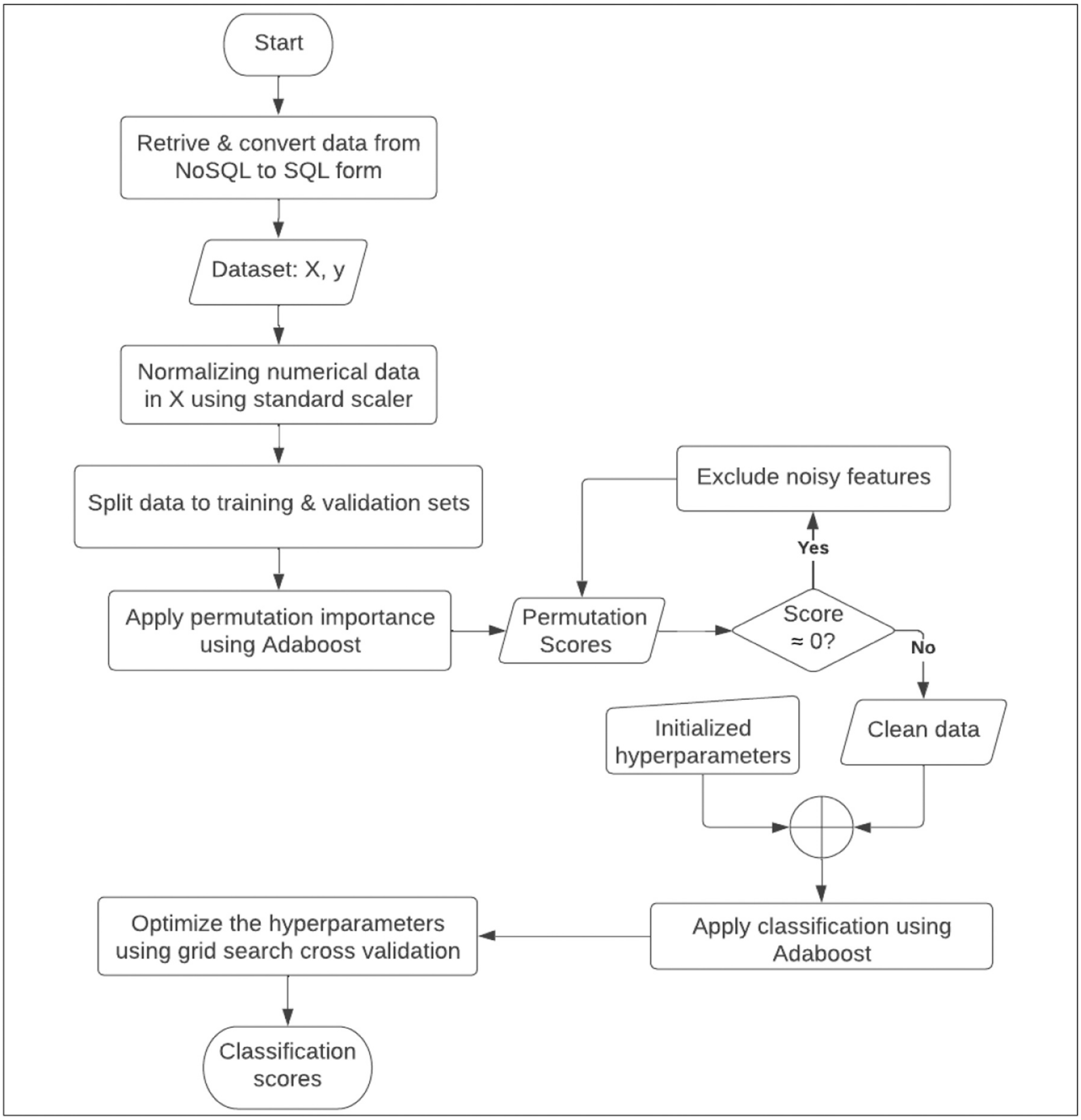

Fig. 9. Classification scheme flowchart.

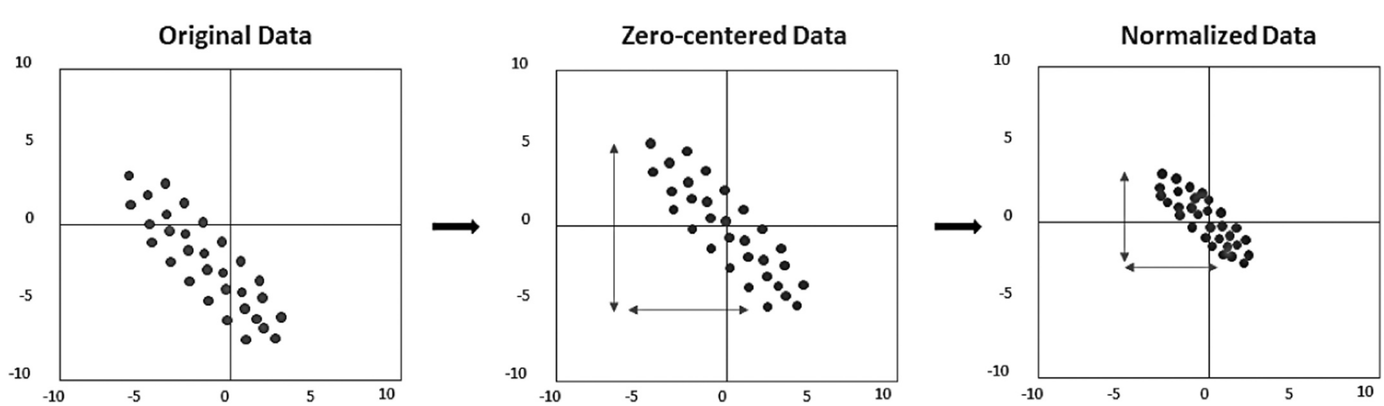

Fig. 10. Data normalization procedure.

they are divided into training set and testing and validation set. The training set is 60% and the testing and validation set is 20% and 20% respectively. The system is trained by the training set and the testing and validation set is used to validate and test the proposed method. The results are validated readily since the machine stoppage data includes the sensor reading and stoppage reason. The feature selection process is done through permutation importance technique in which the noisy features are excluded. In the end, the ML algorithm is applied then optimized using grid search cross validation to determine the best classification scores.

4.1. Data preprocessing

In this section, the data preprocessing workflow is illustrated, which includes, the data normalization and the feature selection processes.

4.1.1. Data normalization

The goal of normalization is to transform features to be on a similar scale. This scale is centralized at the origin with peak of − 1 or 1 as shown in Fig. 10. This process improves the training stability of the model. The ML models may suffer poor performance and high complexity due to input data with outliers and high variance [18–21]. Hence, z-score technique is used to normalize each data point i in data category j (D_i^j). It measures the data point distance to its mean value divided by a standard deviation. z-score is a variation of scaling that represents the number of standard deviations away from the mean. The main advantage of z-score is to ensure that the feature distributions have mean of 0 and standard deviation of 1. The advantage of normalization appears when there are a few outliers. The z-score normalization can be expressed as

$$D_j^{i'} = \frac{D_j^i - \overline{D_j}}{\sigma_{D_j}} \tag{1}$$

where.

$D_j^{i'}$: normalized data point I in data category j,

D_j^i: data point I in data category j,

$\overline{D}_j$:mean of data category j,

σ_{D_j}: standard deviation of data category j.

4.1.2. Feature selection

Permutation importance is a technique that can be used to determine the significance of features in a machine learning model. This method involves randomly shuffling the values of a feature in the test dataset and observing the resulting reduction in the model's accuracy. A large drop in accuracy indicates that the feature is important to the model's performance. The use of permutation importance in classification has several benefits, such as aiding in feature selection and model evaluation [15,16]. When conducting permutation importance using Adaboost, the feature importance is measured for the specific classifier used in the classification process. This approach can provide more accurate and relevant information since different classifiers may assign varying levels of importance to features. Adaboost is a type of ensemble method that combines multiple weak classifiers to create a strong classifier. Measuring feature importance using Adaboost can provide insight into how the various weak classifiers in the ensemble contribute to the model's overall performance [21–23]. The steps of measuring feature importance using Adaboost are listed below:

- Input: fitted predictive model m, tabular dataset (training or validation) D.
- Calculate the reference score s of the model m on dataset D (for instance classifier accuracy).
- For each feature j:
  o For each iteration k in $1,..,\mathrm{k},\ldots,K$:
    - Shuffle column j of dataset D to generate a distorted version of the data named $\widetilde{D}_{k,j}$
    - Calculate the score $s_{k,j}$ of model m on data $\widetilde{D}_{k,j}$
    - Calculate importance i_j for feature f_j as follows.

$$i_j = s - \frac{1}{K}\sum_{k=1}^{K} s_{k,j} \tag{2}$$

Where K: the total number of points in a data set, and.

k: the k$^{\mathrm{th}}$ data point.

4.2. Machine Stops Classification

Adaboost is a popular ensemble learning technique that aggregates weak learners into a strong classifier. Moreover, it is short for Adaptive Boosting. Therefore, it is used in this study for classification purposes. In the present work, multiple weak classifiers are trained on different subsets of the training data, and their predictions are combined to make a final stronger prediction. During each iteration, Adaboost assigns higher weights to misclassified data points to force the weak classifiers to focus on the most challenging examples. The final classifier is a weighted sum of the weak classifiers, where the weights are determined based on their individual performance. Adaboost has been widely used in various purposes such as text classification, face recognition, and bioinformatics [24]. One of the main advantages of Adaboost is its ability to handle complex data distributions and achieve high accuracy with simple classifiers. Adaboost has also been shown to be robust to noise and outliers, making it suitable for real-world scenarios [25]. Moreover, Adaboost is a flexible algorithm that can be combined with different weak learners, such as decision trees, SVMs, and neural networks, to improve their performance [26]. Classification algorithm using Adaboost is given below.

- Given: $(x_1, y_1), \ldots, (x_m, y_m)$ where $x_i \in \mathscr{X}, y_i \in \{-1, +1\}$.
- Initialize $D_1(i) = 1/m$ for $i = 1, \ldots, m$.
- For $= 1, \ldots, T$:

- Perform training for weak learner using distribution D_t.
- Obtain weak hypothesis $h_t : \mathscr{X} \rightarrow \{-1, +1\}$.
- Select h_t with low weighted error. The error is calculated as follows:

$$\varepsilon_t = \Pr_{i \sim D_t}\left[h_t(x_i) \neq y_i\right] \tag{3}$$

- Choose $\alpha_t = \frac{1}{2}\ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$
- For $i = 1, \ldots, m$, update

$$D_{t+1}(i) = \frac{D_t(i)\exp(-\alpha_t y_i h_t(x_i))}{Z_t} \tag{4}$$

where Z_t is a normalization factor (It is selected such that D_{t+1} is a distribution).

- Find the final hypothesis:

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right) \tag{5}$$

The grid search cross-validations (GridSearchCV) are then applied. GridSearchCV's a commonly used hyperparameter tuning technique in machine learning. It is a systematic approach to search for the best combination of hyperparameters by exhaustively testing all possible parameter combinations within a predefined range. The technique evaluates the performance of each parameter combination by cross validating the model on different subsets of the training data. GridSearchCV is widely used in various applications such as image classification, natural language processing, and predictive modeling [27].

Table 3
Permutation importance parameters.

Base estimator	Number of estimators	Learning rate	Number of repeats
Decision tree classifier maximum depth = 8	50	1	10

One of the advantages of GridSearchCV is its ability to optimize the model hyperparameters in a systematic and efficient manner. The technique provides a comprehensive search over a predefined parameter grid, which enables the user to find the optimal hyperparameters without relying on intuition or trial-and-error. Moreover, GridSearchCV can help prevent overfitting by using cross-validation to evaluate the performance of each parameter combination on independent subsets of the training data [28].

5. Results

The AdaBoost classifier permutation importance parameters are listed in Table 3. Permutation importance results in feature selection is shown in Fig. 11.

The permutation importance test showed that the stich length, yarn feeder speed and machine speed are considered as noisy data and does not affect the output since there score is zero. The output here refers to the stop type during the classification process. Table 4 shows the actual score for each feature that is not dropped.

The event duration showed the best promising results in feature selection process with score of 0.345. The parameters used by grid search to determine the optimum (best parameter), are presented in Table 5.

The AdaBoost model is applied after eliminating the unwanted features and hyper-tuning the model. The model achieved 92% accuracy of classifying the stop only using the event duration. The classification results are reported in Table 6.

In the classification report for knitting machine stops, the "Complete" category has a perfect score of unity with respect to precision, recall, and F1-score, based on 144 samples. The "Feeder" stops have a precision of 0.91, recall of 0.66, and F1-score of 0.76, with 163 samples. The "Gate" stops exhibit high precision and recall at 0.92 and 0.95, respectively, with an F1-score of 0.93 from 332 samples. The "Idle" stops, the largest group with 602 samples, have a precision of 0.91, a recall of 0.96, and an F1-score of 0.93. The "Lycra" stops, though limited in number (8 samples), showcases a balanced precision, recall, and F1-score, all at 0.75. The "Needle" stops, based on 11 samples, have a lower precision of 0.44, a recall of 0.36, and an F1-score of 0.4. The overall accuracy of the model across all categories is 0.92 based on 1260 samples. The macro average, which gives equal weight to each category, has a precision of 0.82, recall of 0.78, and F1-score of 0.8. The weighted average, considering the number of samples in each category, is 0.92 for precision, 0.92 for recall, and 0.91 for the F1-score.

Table 4
Permutation importance score for each feature.

Parameter	Created At	Progress	Total Turn per Roll	Updated At	Event Duration
Score	0.175	0.165	0.175	0.145	0.345

Table 5
Grid search hyperparameter optimization.

Parameter	Train/test split ratio	Number of estimators	Learning rate
Range	0.1 to 1	10 to 150	0.1 to 1
Optimum	0.4	100	0.1

Table 6
Classification report.

	Precision	Recall	F1-Score	Support (Samples)
Complete	1	1	1	144
Feeder	0.91	0.66	0.76	163
Gate	0.92	0.95	0.93	332
Idle	0.91	0.96	0.93	602
Lycra	0.75	0.75	0.75	8
Needle	0.44	0.36	0.4	11
accuracy			0.92	1260
macro avg	0.82	0.78	0.8	1260
weighted avg	0.92	0.92	0.91	1260

6. Conclusion

In conclusion, this study focuses on addressing the challenges posed by unplanned outages in the textile industry through the

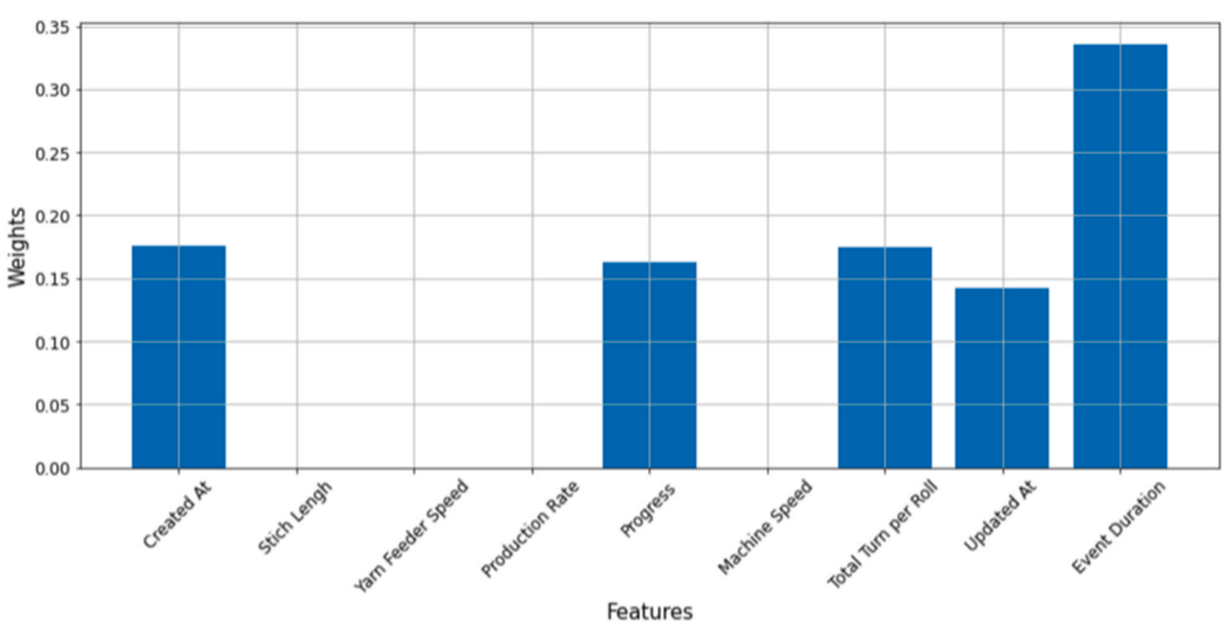

Fig. 11. Permutation importance results.

implementation of a predictive maintenance system. By leveraging IoT-enabled devices and machine learning algorithms, specifically AdaBoost, the study successfully classified different types of machines stops in real-time. The data collected from the machines, including machine speeds and steps, underwent pre-processing were used as inputs to the machine learning model. The model was trained and optimized using a combination of hyperparameter tuning and cross-validation techniques. As a result, it achieved an impressive accuracy of 92% on the test set. The outcomes of this research demonstrate the potential of the proposed system to accurately classify various machine stops, encompassing gate stops, feeder stops, needle stops, completed roll stops, idle stops, and Lycra stops. By enabling timely maintenance actions, the system has the capability to enhance overall efficiency and productivity within the textile industry. Implementing the proposed predictive maintenance system holds significant advantages for the textile industry such as mitigating machine failures, reducing unplanned downtime, minimizing production losses, and reducing maintenance costs. Moreover, it empowers businesses to adopt a proactive maintenance approach, resulting in improved operational efficiency and resource optimization. Accordingly, this study has significant potential impact in the textile industry. It enhances machine lifetime, improves product quality, and improves manufacturer income. Future work includes expanding the data inputs by more stop types to cover all incipient faults. Moreover, the presented technique will be applied on different types of machines.

Funding

STDF grant number 43467.

Declaration of Competing Interest

I am the corresponding author and I declare that I have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This paper is based upon work supported by Science, Technology & Innovation Funding Authority (**STDF**) under grant number **43467**.

References

[1] W. Lee, H. Wu, H. Yun, H. Kim, M. Jun, J. Sutheralnd, Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data, Procedia CIRP 80 (2019) 506–511, https://doi.org/10.1016/j.procir.2018.12.019.
[2] T. Zonta, C.André da Costa, R. da Rosa Righi, M. de Lima, E. da Trindade, G. Li, Predictive maintenance in the Industry 4.0: a systematic literature review, Comput. Ind. Eng. 150 (2020) 106889, https://doi.org/10.1016/j.cie.2020.106889.
[3] W. Lee, H. Wu, H. Yun, H. Kim, M. Jun, J. Sutherland, Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data, Procedia CIRP 80 (2019) 506–551, https://doi.org/10.1016/j.procir.2018.12.019.
[4] C. Mgbemena, F. Okeagu, Development of an IoT-based real-time remote monitoring device for the maintenance of injection moulding machines in plastic industries, UNIZIK J. Eng. Appl. Sci. 2 (1) (2023) 260–278.
[5] Y. Gao, C. Chai, H. Li, W. Fu, A deep learning framework for intelligent fault diagnosis using automl-cnn and image-like data fusion, Machines 11 (10) (2023) 932, https://doi.org/10.3390/machines11100932.
[6] W. Udo, Y. Muhammad, Data-driven predictive maintenance of wind turbine based on SCADA data, IEEE Access 9 (2021) 162370–162388, https://doi.org/10.1109/ACCESS.2021.3132684.
[7] J. Lee, J. Ni, J. Singh, B. Jiang, M. Azamfar, J. Feng, Intelligent maintenance systems and predictive manufacturing, J. Manuf. Sci. Eng. 142 (2020) 1–40, https://doi.org/10.1115/1.4047856.
[8] K. Singha, S. Maity, P. Pandit, Use of AI and machine learning techniques in knitting, Ch.6, Text. Inst. Book Ser. (2022) 161–180, https://doi.org/10.1016/B978-0-323-85534-1.00021-0.
[9] C. Baban, M. Baban, S. Darius, Using a fuzzy logic approach for the predictive maintenance of textile machines, J. Intell. Fuzzy Syst. 30 (2) (2016) 999–1006, https://doi.org/10.3233/IFS-151822.
[10] S. Elkateb, A. Métwalli, A. Shendy, An Innovative Online Monitoring System in Knitting Industry, The 16th Textile Bioengineering and Informatics Symposium; Blended Conference. 412–419, August 22–25(2023). DOI: TBIS 10.3993/tbis (2023).
[11] S. Elkateb, A. Métwalli, A. Shendy, K. Moussa, A. Abu-Elanien, Online monitoring-based prediction model of knitting machine productivity, Fibres Text. East. Eur. 31 (4) (2023) 46–52, https://doi.org/10.2478/ftee-2023-0035.
[12] O. Surucu, S.Andrew Gadsden, J. Yawney, Condition monitoring using machine learning: a review of theory, applications, and recent advances, Expert Syst. Appl. 221 (2023) 119738, https://doi.org/10.1016/j.eswa.2023.119738.
[13] N. Mohammed, O. Abdulateef, A. Hamad, An IoT and machine learning-based predictive maintenance system for electrical motors, J. Eur. Des. Systèmes Autom. 56 (4) (2023) 651–656, https://doi.org/10.18280/jesa.560414.
[14] M. Özdamar Seitablaiev, F. Umaroğulları, Thermal comfort and indoor air quality, Int. J. Sci. Res. Innov. Technol. 5 (2018) 90–109.
[15] M. Rassel, M. Hoque, Re-evaluation on causes of circular knitting machine production efficiency and their impact on fabric quality, Eur. Sci. J. ESJ 15 (2019), https://doi.org/10.19044/esj.2019.v15n21p448.
[16] S. Moazzem, E. Crossin, F. Daver, L. Wang, Environmental impact of apparel supply chain and textile products, Environ., Dev. Sustain. 24 (2022), https://doi.org/10.1007/s10668-021-01873-4.
[17] N. Kohli, K. Dua, Indoor Environment: Light, Noise, Humidity and Temperature. Ch.5:57-75, Integrated Publications, Rohini, Delhi-110085, India, 2022.
[18] M. Alves, H. Simões, J. Pereira, J. de Figueiredo, A. Ferreira, Evaluation of workers' exposure to occupational noise in the textile industry (Case Study), SSDC 492 (2023).
[19] H. Hameed, A. Eleue, F. Hussein, Noise induced hearing loss (NIHL) in wasit corporation textile industries, J. Otolaryngol. -ENT Res. 11 (2) (2019).
[20] A. Fisher, C. Rudin, All models are wrong but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously, J. Mach. Learn. Res. 18 (1) (2018) 3754–3786.
[21] A. Niculescu-Mizil, R. Caruana, Predicting good probabilities with supervised learning, Proc. 22nd Int. Conf. Mach. Learn. (2005) 625–632.
[22] T. Hastie, R. Tibshirani, J. Friedman, The elements of statistical learning: data mining. Inference, and Prediction, Springer,, 2009.
[23] P. Braga, A. Oliveira, G. Ribeiro, S. Meira, Bagging predictors for estimation of software project effort, Int. Jt. Conf. Neural Netw., Orlando, FL, USA (2007) 1595–1600, https://doi.org/10.1109/IJCNN.2007.4371196.
[24] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, J. Comput. Syst. Sci. 55 (1) (1997) 119–139.
[25] C. Hsu, C. Lin, C. Chang, A practical guide to support vector classification. Department of Computer Science, National Taiwan University, Taipei, Taiwan. (2003).
[26] R.E. Schapire, A brief introduction to boosting, Proc. Sixt. Int. Jt. Conf. Artif. Intell. (1999) 1401–1406.
[27] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, J. Mach. Learn. Res. 13 (2012) 281–305.
[28] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, E. Duchesnay, Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830.