Feature Engineering and Machine Learning Modelling for Predictive Maintenance Based on Production and Stop Events

Ariel Cedola, Rosaria Rossini, Ilaria Bosi, Davide Conzon
IoT and Robotics Research Area
LINKS Foundation
Turin, Italy
email: {ariel.cedola, rosaria.rossini, ilaria.bosi, davide.conzon}@linksfoundation.com

***Abstract*—Manufacturing systems suffer from progressive degradation due to wear, fatigue, cracking, corrosion, with respect to both age and usage. Reduced performance of system components and even catastrophic failure could be the main consequences of not being able to detect such faults at early times. Fault diagnosis and predictive maintenance aim at showing the machine working conditions, indicating current and possible future abnormal states, and allowing to take appropriate actions in advance in order to avoid damages, minimize downtime, improve the safety of the whole system and reduce manufacturing and repairing costs. In this paper, we successfully apply a data driven modelling approach designed for log data to a new scenario. The methodology proposed transforms the production and stops data of an industrial machine into a thoughtfully elaborated series of timestamped events, and applies a set of feature engineering techniques that enables to exploit the pipelines typically implemented in log-based predictive maintenance modelling. The transformed data is used to train a binary classificator that predicts with high accuracy (96.2%) the occurrence of a machine failure in the short-medium term.**

***Index Terms*—predictive maintenance, machine learning, feature engineering, manufacturing**

I. INTRODUCTION

The term 'Industry 4.0' refers to the Fourth Industrial Revolution, the recent trend of automation and data exchange in manufacturing technologies. The key fundamental principles of Industry 4.0 include data integration, flexible adaptation, cloud intranet, intelligent self-organizing, manufacturing process, optimization, interoperability, secure communication, and service orientation [1]. These innovative technologies are used to create a "smart factory" where machines, systems and humans communicate with each other to cooperate, monitor progress and connect sensors to provide data concerning the quality and the quantity of the goods, along the assembly line [2]. Apart from the evaluation of the quality of the final products, Predictive Health Management (PHM) systems use real-time and historical state information of machines, subsystems and components to provide actionable information, enabling intelligent decision-making for improved performance, safety, reliability, and maintainability in the manufacturing sector. The focus of health management is to minimize operational loss and to maximize the objectives established by the facility [3]. PHM of components or systems involves both diagnostics and prognostics: diagnostics is the process of detection and isolation of failures or faults, while prognostics is the process of prediction of the future state or Remaining Useful Life (RUL) based on current and historic conditions [4], [5]. The RUL is the prediction of the cycle-time before the performance of a component, system or process reaches an unacceptable low threshold [6]. Prognostics is based on the understanding that machines fail after a period of degradation, which if measured and thoughtfully analyzed, can be exploited to prevent system breakdown and minimize operation costs. These are in fact the goals of the predictive maintenance methodologies [7].

Data are at the core of diagnostics and prognostics operations. Manufacturers that envision the importance of data as an asset and manage to implement effective strategies to leverage data in multiple ways, gain measurable advantage over competitors. All players at a small- or large-scale produce data, but an aspect that distinguishes each other is the harnessing and management of the data that they conduct [8]. Data have different types and formats and can be collected from numerous sources using a variety of technological approaches, processed accordingly, and saved in an extensive catalogue of data stores and databases both locally and on the cloud. Data typically applied to condition monitoring and hopefully to predictive maintenance of machines are sensor and log data. In the present work, based on the lack of these kind of information, we propose the combined use of production and stops records of machines to construct artificial intelligence tools able to anticipate machine failures, and satisfactory apply this approach to predict the RUL of a machine operating in the textile sector.

The remainder of this paper is organized as follows. Section II describes the data-driven techniques commonly applied in predictive maintenance. Section III presents the scenario and the data utilized in the work. In Section IV we introduce the proposed approach and explain its three main steps: data preprocessing, feature engineering and machine learning modelling. The application of the method and the obtained results are presented in Section V. Finally, in Section VI, we draw some concluding remarks and suggest important work to address in the future.

II. Background

Model-based and data-driven approaches are two main techniques for diagnosis, monitoring and predictive maintenance of machines [9]. Model-based approaches exploit mathematical and physical models to provide insights into the failure mechanism of systems [10]. Faults are diagnosed by monitoring discrepancies between model calculations and the actual measurements. Data-driven approaches, on the other hand, are featured by building machine learning models based on large volumes of data without using the knowledge of the physical mechanisms behind the failures, and can provide excellent diagnosis results and RUL estimation [11], [12].

Data-driven predictive maintenance can be classified as sensor-based or log-based, depending on the type of data able to be generated and extracted from the machinery system and used for model training. The most frequently implemented is the sensor-based methodology, in which the streams of sensor measurements describing the working conditions of the machine are stored, aggregated, processed and applied to train a machine learning model, and subsequently to monitor and score the algorithm in order to predict future failure events [13]–[15]. Machine sensor data consist essentially in a set of time series signals, collected in general at regular time periods, and device identifiers formatted according to a hierarchical structure like machine/subsystem/component/channel. Typical machine parameters monitored by sensors are temperature, pressure, rotation speed, vibration, electrical consumption, acoustic emissions, among others. Sensor data are usually complemented with environmental, production, alerts, failures and maintenance data in this kind of approaches, enabling through a feature engineering process the generation of more solid datasets from which machine learning models with a higher prediction power can be constructed. Modern machines incorporate sensors and data processing modules from factory, but in older equipment these devices must be installed with the machine already in production. IoT devices and technologies facilitate enormously the conditioning of machines into a predictive maintenance ready status, although in some cases the high cost, the intensive use or the presence of mechanical or even regulatory constraints can turn this process unfeasible.

Log-based approaches, conversely, use event-log data for machine diagnostics, and prevent the need of implementation or monitoring of sensor data to train machine learning models for prognostics and predictive maintenance [16]–[20]. Programmable Logic Controllers (PLC) and software applications running on machine controllers continuously produce logs containing valuable information about internal events, tasks carried out, warnings, errors, components state, dialogues between modules, etc. Logs are generated automatically at a very high rate, reaching typically volumes of hundreds of thousand records per hour. Every log is timestamped and appended into a plain text or Extensible Markup Language (XML) file. Many log files can be generated daily by machines, each one containing a limited amount of records in order to avoid complicating the reading process and the storage of very large sized files. The management of these files is in fact an important aspect to consider when preservation of log data of a machine is contemplated. Most data recorded in log files are unstructured text data, and the extraction of the small subset of data embedded into the logs, that might be useful for predictive maintenance purposes, requires a heavy preprocessing work. There is no data in the logs providing explicit information about the machine condition that can be directly applied to predict failures. This means that a careful feature extraction process must be performed in log-based approaches to obtain successful prediction results through appropriate machine learning modelling.

In both approaches depicted above, the integration of the data with machine failure records is mandatory for RUL prediction using supervised machine learning algorithms. Each instance of the dataset is composed by a vector of features and a label. Features are constructed from collected sensor or log data, whereas failure records are exploited for data labeling. Binary or multiclass classification are the most frequently applied model types in data-driven predictive maintenance scenarios, rather than regression models. In binary problems, particularly, the label usually expresses the occurrence (positive class) or not (negative class) of a failure within a prediction window ahead in time from a specific prediction point, to which the data instance is associated. These two classes define the RUL to be predicted by the model. Both sensor and log data are time-series data, and the construction of the features for every prediction point is not based only on the data at that single point in time but on the aggregation of data within a specific time window set before the prediction point. For log-based approaches, it has been also reported in the literature the use of Multiple Instance Learning (MIL), in which instances are bagged and bags are labeled positive or negative according to the labels of the instances within them [19], [21]. Typical classification models applied to predictive maintenance include XGBoost, Random Forest, Gradient Boosting Machines (GBM), Support Vector Machines (SVM), Neural Networks, among others [22]–[24].

III. Scenario and data description

A. Scenario

The equipment under study is a bleaching machine, utilized in textile industries to remove the natural yellowish-brown coloring of fabric fibers, in order to confer to the material a white appearance. The equipment owner is a big international textile manufacturer with headquarters in Turkey. The bleaching machine is a long production line that operates in continuous mode in a 24/7 regime, processing an average of 80 km of fabric per day and performing four different processes: Bleaching, Bleaching + Emulsion, Repairment and Washing. The machine operation and architecture are subdivided in four steps: pre-washing, bleaching, washing and drying of the fabric.

In the bleaching step, the fabric is subject to a chemical bath and a steamer, with controlled times, temperatures and

 ISBN: 978-1-61208-891-4

	Process	BatchNumber	FabricCode	Length	Production	StartTime	EndTime	Duration	InletTrolleyNumber	OutletTrolleyNumber
OrderID										
2000	BLEACHING	27377	IHB154	8535	13144	2020-06-01 07:55:30	2020-06-01 10:01:47	126.28	283	917
2001	BLEACHING	28431	FRF158	3666	5756	2020-06-01 10:04:57	2020-06-01 10:57:16	52.32	301	3038
2002	BLEACHING	28216	FRF208	5731	11978	2020-06-01 10:58:52	2020-06-01 12:31:49	92.95	301	3038
2003	BLEACHING	200001	FRF178	1509	2686	2020-06-01 12:36:06	2020-06-01 13:00:55	24.82	352	564
2004	BLEACHING	25409	FRF178	1652	2941	2020-06-01 13:01:05	2020-06-01 13:25:34	24.48	352	564

(a) Production data

	StartTime	EndTime	Duration	StopReason
StopID				
10000	2020-06-01 09:38:00	2020-06-01 09:41:00	153	Recipe change
10001	2020-06-01 10:01:00	2020-06-01 10:04:00	131	Unreasoned stops
10002	2020-06-01 10:04:00	2020-06-01 10:04:00	43	Unreasoned stops
10003	2020-06-01 10:04:00	2020-06-01 10:04:00	1	Unreasoned stops
10004	2020-06-01 10:04:00	2020-06-01 10:05:00	12	Unreasoned stops

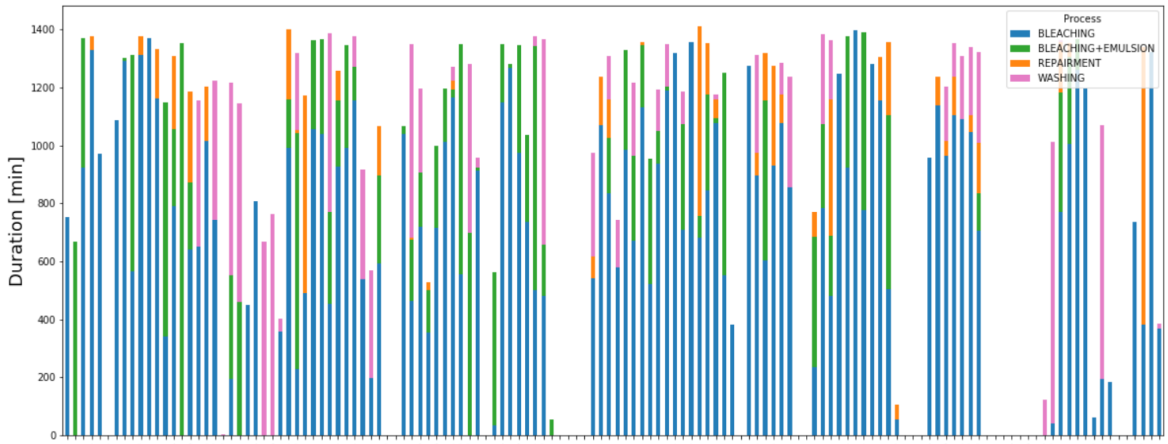
(b) Stops data

Fig. 1: Samples of the tables containing production (a) and stops (b) data.

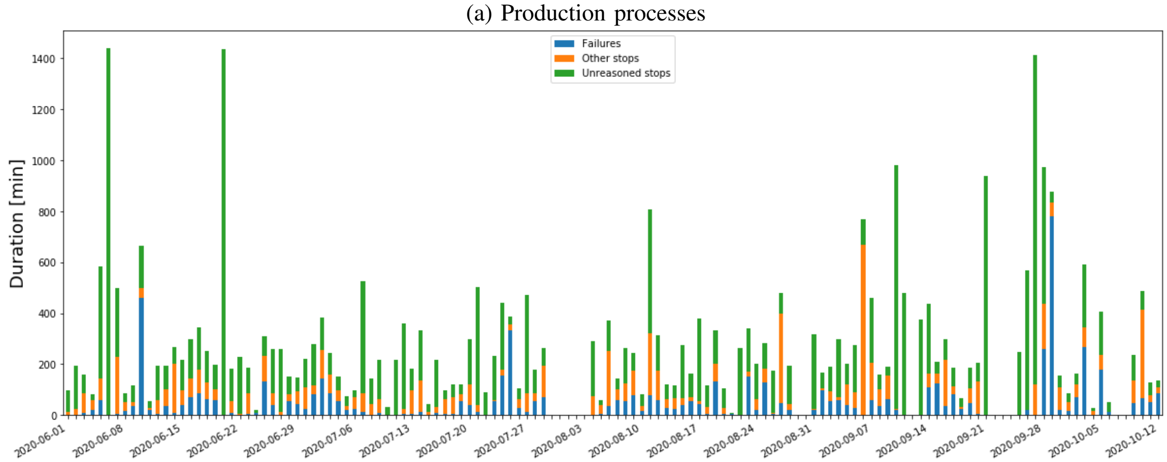
proportion of chemical bath components, which depend on the type and whiteness level of the inlet fabric. An excessive time of exposure of the material to the chemicals can produce irreversible damages, resulting in the loss of up to kilometers of textile. The long exposures to the chemicals are caused by different types of stops that the machine suffers during its operation. One of the most frequent stops are due to mechanical failures, specifically to failures of cylindrical roller bearings, which progressively deteriorate due to abrasion. The whole machine structure includes a total of 1259 roller bearings of 54 different types. Despite the intensive preventive maintenance activities, an average of 1 mechanical failure per day is reported by the machine owner. Other failure types affect the normal operation of the machine, i.e., electrical, electronic and power failures, depending on the root cause and the component or subsystem in which the issue is originated. Failures with a repairing duration exceeding 10 minutes are classified by the maintenance department as critical, based on Total Productive Maintenance (TPM) principles, and generally meet 10% of the total failures. Half of the total critical failures time is due exclusively to mechanical failures, which reach an average duration of 27 minutes. These statistics extracted from data provided by the manufacturer reveal the importance of targeting the prediction of mechanical failures to enable the predictive maintenance of the equipment. In addition to the huge quantity of roller bearings, this complex machine is composed also by 50 motors, 30 inverters, 5 chemical dosing pumps, and many other components. The machine is not equipped with a network of sensors for condition monitoring, and a collection of log files reporting a reasonable period of historical log events was not available at the moment of conducting this study. The lack of data questioned initially the possibility of executing a data-driven analysis for RUL prediction. Data provided by the manufacturer consist of four months of production and stop events in structured format, with complete information about the type of production process performed or stop undergone, and the corresponding starting and final timestamps.

B. Data description

Data about production processes and stop events were provided by the machine owner in two separate files, exported from the production management software and the PLC, respectively. The tables in the source files contain 1883 production records and 7663 stop events registered from June to October 2020. The structure of these tables and some example data are shown in Figure 1. The production table includes data about the performed process (Process), the ID of the order (BatchNumber), the code of the processed fabric (FabricCode), the input length (Length, in meters) and area (Production, in squared maters) of the material, the initial and final timestamps (StartTime, EndTime), and the order duration (Duration, in minutes). This duration includes the times of all eventual stops and failures undergone by the machine during the processing of the order. The stops table simply shows the type of stop experienced by the machine (StopReason), the initial and final timestamps (StartTime, EndTime), and



(a) Production processes



(b) Unplanned stops

Fig. 2: Duration of production processes (a) and unplanned stops (b) aggregated by day.

the duration (Duration, in seconds). It is important to notice that stop timestamps have minute resolution (seconds = 00), which leads to small inaccuracies when analyzing sequences of stop events during the processing of fabrics, and to the presence of records with identical StartTime and in many cases also EndTime timestamps. The indexes OrderID and StopID were added to uniquely identify each record in the tables. BatchNumber in the production table might not be unique due to one of two reasons: reprocess of bleaching when obtained whiteness degree is not acceptable (Repairment process), and reuse of BatchNumber once the processing of a fabric through all required machines in the plant is complete. The StopReason is introduced by the operators through the PLC screen of the machine, by choosing the proper option from the full list which is shown when the machine stops. The following is the list of the 17 stop reasons present in the dataset, classified as Planned and Unplanned:

- Planned stops (4): Planned maintenance, Cleaning, No production, Holiday.
- Unplanned stops (13): Unreasoned stops, Recipe change, Fabric rupture, Waiting for batch frame trolley, Fabric wrapped around a cylinder, Refreshing Stitching, Fabric construction changing, Color OK, Water changing, Mechanical Failure, Electronic failure, Electrical failure, Power failure.

The two tables integrate a total of 2050 hours of production and 1235 hours of stop events. Bleaching and Bleaching + Emulsion processes comprise the 81% of all production orders. Bleaching + Emulsion is the bleaching process executed for a specific type of fabric. More than half of the stop time corresponds to the Planned stops, mainly to the No production reason. Near 55% of the orders processed in the reported period presented unplanned stops, with a total stop duration of almost 300 hours. The large number of registered stops is due to the presence of numerous sequences of microstops of few seconds duration, usually labelled as Unreasoned stops, that should be conveniently considered as unique longer chains of these events. The duration of near 48.5% of stops does not exceed 60 seconds, whereas 81.5% last less than 5 minutes. Total stop time due particularly to the four failure types rises to 107 hours, 72 of which correspond to the mentioned critical condition. The Unreasoned stops are largely the most frequent in the dataset, reaching 67% of all stop events and 30% of the full stop time. Figure 2 shows the duration of production processes (a) and unplanned stops (b) aggregated by date. The eight stop reasons other than Unreasoned stops and Failures are combined and shown in the picture as Other stops.

IV. Event-based Modelling Approach

The approach proposed in this paper aims to transform the above mentioned production and stop data in a series of timestamped events, using a format that enables to exploit the application of the pipelines typically implemented in log-based predictive maintenance modelling. These pipelines include feature engineering steps in which the learning instances are generated by processing and aggregating data within a features window preceding each prediction instant, i.e., the instance point, subdivided into a number of time windows. Features are extracted from this time frame by applying strategies like rolling windows with different aggregation functions and temporal coverage, identification and counting of patterns, among others. A label window used to generate the instance labels is also created from each of those instants. One instance is formulated for each instance point, which is moved ahead in time along with the features and label windows by a predefined time step, in order to generate the full set of training and testing instances.

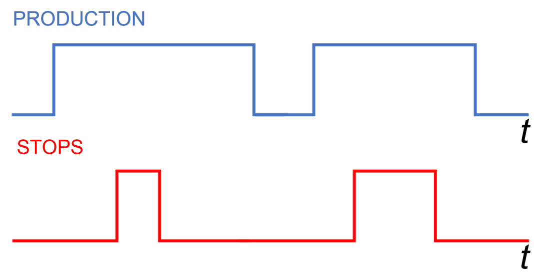
The three big steps implemented in this approach and described in the next subsections are the following:

- Data preprocessing: Events formulation and alignment
- Feature engineering: Instance and label generation
- Machine learning modelling: Binary classification model training and testing

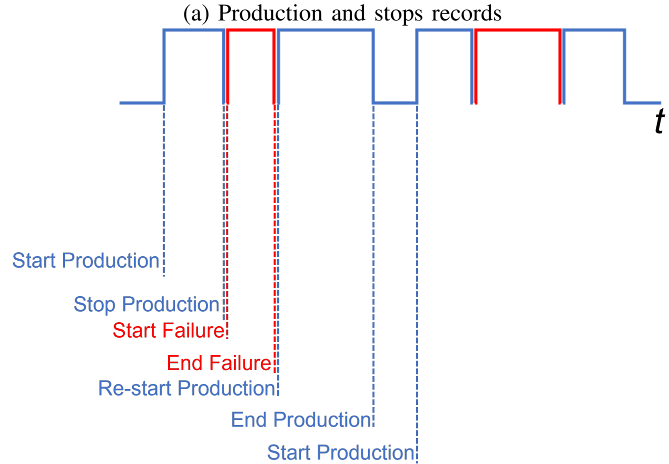
A. Data Preprocessing

The production and stop records present in the data sources have start and end timestamps, and in order to transform these records in a streaming of events, what we have done is to consider both start and end of the records as two separate events: Start Production/Stop and End Production/Stop. This sequence of events, correctly arranged by order of occurrence, is in fact the proper input data format needed for the future training and scoring of the model in production using streaming data in real-time. The collected data show that more than 70% of the registered stops occur during production processes, i.e., while the machine is in operation. This means that a failure starting after the StartTime and finishing before the EndTime of a production record splits this process into two pieces, one previous and the other subsequent to the failure. In terms of the events formulation depicted above, a situation like this one gives place to six different events, namely Start Production, Stop Production, Start Failure, End Failure, Re-start Production and End Production. Depending on the number of stops registered during the process, production records could be splitted in more than just two parts. Figure 3 shows an example of the events generated from the occurrence of two stops/failures during two production processes.

With the aim of applying this principle to the whole available dataset, we had to perform several steps of data



(a) Production and stops records



(b) Data alignment and events creation

Timestamp	Event
16/06/2020 08:12	Start Production
16/06/2020 09:14	Stop Production
16/06/2020 09:15	Start Failure
16/06/2020 10:01	End Failure
16/06/2020 10:02	Re-start Production
16/06/2020 12:23	End Production
16/06/2020 12:28	Start Production
16/06/2020 13:56	Stop Production
16/06/2020 13:57	Start Failure
16/06/2020 16:22	End Failure
16/06/2020 16:23	Re-start Production
16/06/2020 18:41	End Production

(c) Sequence of events

Fig. 3: Example of events generation after alignment of records and splitting of production processes.

preprocessing before, to obtain a clean and consistent stream of events. In addition, we categorized the stops as those occurring during or outside the production processes. The preprocessing steps involved were:

- Chaining of stops: Trains of consecutive short stops of the same type (StopReason) were permanently joined into single and longer stop chains of identical type, and duration given by the sum of the duration of all stop components.
- Chaining of chains of stops: Consecutive and very close

Copyright (c) IARIA, 2021. ISBN: 978-1-61208-891-4

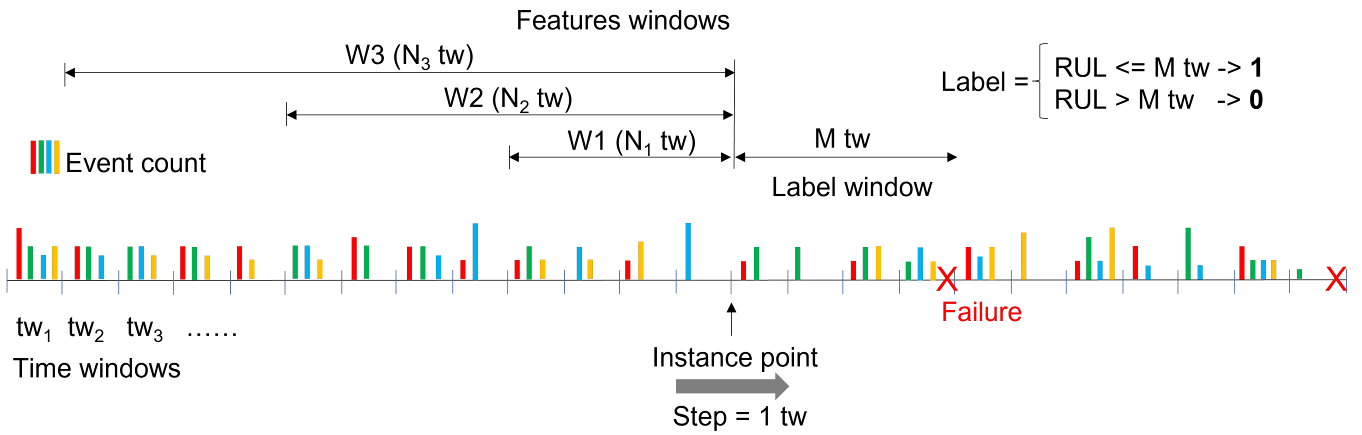


Fig. 4: Scheme applied to the generation of the training instances.

chains of stops of identical or different type were temporarily joined in order to avoid the generation of very short pieces of production records when splitting for events formulation.

- Categorization of stops: Every stop or stop chain occurring exclusively during a production process, i.e., for which StartTime (stop) $>$ StartTime (production) and EndTime (stop) $<$ EndTime (production) was denominated Stop Type I (inner stops). All remaining stops were categorized as Stop Type II (outer stops).
- Adjustment of stops EndTime: Given the minute resolution of the stop timestamps, there are many records with equal StartTime and also with identical StartTime and EndTime (around 2k), and in all cases the time gap between timestamps does not reflect the real duration of the stops. To solve this, we assumed the registered StartTime of all stops records as correct and adjusted the EndTime as StartTime + Duration. Despite this leads to slight inaccuracies regarding the exact time of occurrence of stop events, it contributes to generate a more consistent dataset.
- Partial overlapping between production and stop records: StartTime and/or EndTime timestamps of production records were slightly shifted in the cases of the frequently observed short overlapping with previous or subsequent stop records (Type II). The choice of shifting the production records was supported by the fact that registration of the start and end times of these processes is carried out manually by the operators, which means that timestamps are prone to involuntary errors and less reliable, contrarily to the stop records stored automatically by the PLC.

After all these operations, we finally proceeded with the generation of the events. The algorithm roughly consists of the following steps:

- Integrate Production and Stop data
- Detect Type I stops, split Production records and label events
- Label remaining events
- Unchain the temporary stop chains
- Separate instances in start and end events and relabel accordingly

B. Feature engineering

The decomposition and sorting of all production and stops processes following the actions explained above give place to a chronological sequence of discrete events. This dataset is substantially a time series of all events occurring in the bleaching machine, including the target failure event that we plan to predict by training a machine learning model. The dataset contains clean and ordered data, but it is not suitable yet for training a binary classification model for failure prediction. The next step in the pipeline is the generation of the training instances, each one containing a set of features and a label, through adequate feature engineering techniques. Figure 4 shows the scheme used for instances generation, in line with those presented in [16]–[18]. The instances are formulated at specific points in time, separated by fixed time intervals determined by the size of the time windows in which the features window and the label window are subdivided. The point in time associated to an instance is referred to as an instance point, and the features window is a time interval set before it, that finishes at the instance point and has an extension of N time windows. All the events occurring within the features window of an instance point are involved into the determination of the values of the features associated to that instance. In this work we applied the counting of the events to calculate the feature values, but other aggregation functions, statistics and strategies can be used to enrich the feature spectrum. The concept of the feature window implemented in this way is clearly the same of the rolling window commonly applied in time series analytics. All the unique events generated in the Data preprocessing step are converted in features of the learning dataset when following this method. As showed in the figure, multiple features windows can be used for each instance point, with different sizes in terms of time windows, which in turn multiplies the number of features

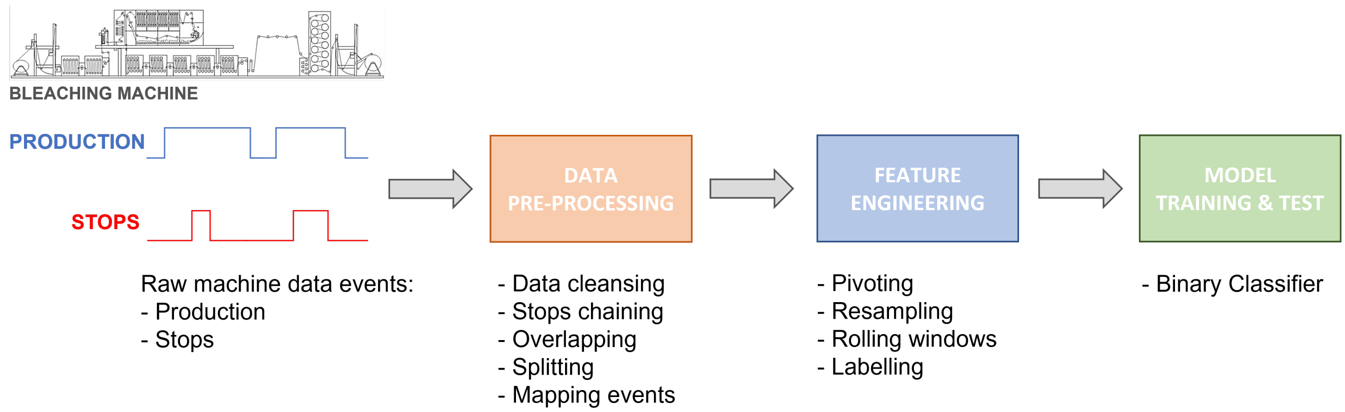


Fig. 5: Steps implemented in the present event-based approach for predictive maintenance.

created. The label window is set from the instance point and has a length of M time windows. The presence of at least one event of the targeted failure inside the label window of an instance point dictates the assignment of a positive label to that instance (class 1), otherwise of a negative one (class 0). It is worth mentioning that within the features window, the target failure events are treated identically to the rest of the stops. Once the features and label are generated and an instance is created, the instance point and the windows are moved ahead by a step of one time window in order to generate the next learning instance. By iterating through this process, we construct the entire dataset for the subsequent training of the failure prediction model.

C. *Machine Learning modelling*

The tool selected to conduct the modelling experiments is LightGBM [25], a gradient boosting decision tree framework that supports different algorithms, i.e., regression, classification and ranking. We applied this framework for binary classification, in order to predict the RUL of the machine under study considering the two classes (0 and 1) introduced in the previous subsection. A prediction of 1 raised by the model indicates the prediction of occurrence of the targeted failure event in the next period of time given by the duration of the label window, starting from the current instance point. LightGBM is a state-of-the-art framework and has demonstrated to be faster and more robust with respect to other tree-based algorithms, enabling explainability and distributed computation.

V. Results

All the steps outlined in Section IV are graphically summarized in Figure 5. We applied this approach to the production and stops data of the bleaching machine, described in Section III B. Having 4 production processes and 17 stop reasons, and considering that up to four events can be generated from every production and stop record (Start, Stop, Re-start, End for production and Start Type I, End Type I, Start Type II, End Type II for stops), we managed to formulate a total of 80 events from the integrated dataset. The size of the events table is independent of the number of the different events created, and it is around 16k records. On the other hand, each event represents one feature within each features window used to construct the training dataset. Anticipating a time window size of 1 hour and the setting of 3 features windows to proceed with the feature extraction process, a training dataset of around 3k instances with 240 features would be obtained. To prevent dimensionality issues, we decided to reduce the number of possible distinct events so as to get a larger ratio between the training dataset size and the number of dimensions. To do this we mapped the processes and stops into 9 groups, as shown in Table I, and joined the Type I and Type II stop categories, to finally reach a total of 20 features. Despite knowing their importance in log-based approaches, a thorough implementation of feature selection techniques was not in the scope of this first version of the work.

The parameters of the feature extraction scheme selected to conduct the experiments are the following:

- Time windows size = 1 hour
- Number of features windows = 3
- Sizes of features windows (N) = 12, 24, 36 time windows
- Size of label window (M) = 12 time windows

The size of the time windows is the parameter that defines the quantity of instances composing the ultimate dataset that is used to train and test the model. A size of 1 hour is reasonable given the typical duration of the production processes and the frequency of occurrence of the unplanned stops and failures. Considering the parameters detailed above, the final dataset consists of 3161 records, 60 features and the label. The event selected as the target failure to construct the label is the Start Mechanical Failure, whose relevance was highlighted in Section III A. The resulting ratio between classes is approximately 1:5, being the minority class the positive one, i.e., the class indicating the occurrence of at least one Start Mechanical Failure event in the next 12 hours from an instance point. Although a slight imbalance between classes is observed, it is not large enough as to consider the problem a severely imbalanced classification.

The LightGBM binary classifier was trained and evaluated

 ISBN: 978-1-61208-891-4

TABLE I: Mapping of production and stop records into groups.

Process or stop reason	***Group***
Bleaching, Bleaching+Emulsion, Repairment, Washing	Production
Planned maintenance, Cleaning, No production, Holiday	Planned
Electrical failure, Power failure	Electrical/Power Failure
Fabric construction changing, Water changing, Recipe change, Color OK	Change
Fabric wrapped around a cylinder, Fabric rupture, Refreshing Stitching	Fabric issue
Unreasoned stops	Unreasoned stops
Waiting for batch frame trolley	Waiting for batch frame trolley
Mechanical Failure	Mechanical Failure
Electronic failure	Electronic failure

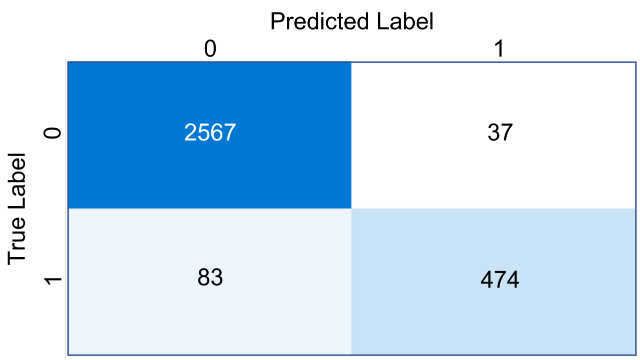


Fig. 6: Confusion matrix.

by applying cross-validation with 5 folds. Hyperparameter tuning was conducted using a grid search to find the model with better performance. Due to the observed class imbalance, the macro average of Area Under Receiver Operating Characteristic Curve (AUC), Precision and Recall metrics were calculated in order to weight equally both the minority and majority classes. The following are the obtained values:

- Accuracy = 0.962
- AUC(macro) = 0.987
- Precision(macro) = 0.948
- Recall(macro) = 0.918

As can be corroborated the metrics are very encouraging. When considering only the minority class, the Recall metric falls to 0.85. This is due to the non-negligible number of false negatives predicted by the model, as can be noticed from the confusion matrix shown in Figure 6. There are 474 and 83 positive samples predicted correctly and wrongly, respectively. There is room to improve even more the classification quality, for example, by deepening the search of the optimal hyper-parameters of the model and conducting experiments with different classificators.

VI. Conclusion and Future Work

Knowing the principles of the log-based approaches for predictive maintenance, we decided to transform the available production and stops data from a textile machine in a large sequence of events having a similar format to the events typically extracted from log files, then apply the proper preprocessing, feature engineering and labelling steps, and finally train a supervised machine learning classificator to predict the time to mechanical failure of the equipment.

The preliminary results evidence the potentiality of the proposed method for the predictive maintenance of industrial machines, and its extreme utility in scenarios in which the lack of collections of sensor measurements and log files prevents the application of more traditional data-driven schemes. Despite the approach has been proven to be effective in a particular use case, we consider that the application can be extended to a wide variety of manufacturing sectors in which the predictive maintenance of the equipment is known to deliver a high business impact.

The work also shows the high extra value that the production and failure data of a machine might provide to a manufacturer. These data are commonly available in industrial plants at all scales, and the study contributes to highlight the reasons because data are considered so valuable assets in the era of Industry 4.0.

It is worth mentioning that the unavailability of detailed and quality maintenance data regarding repairing, reconditioning or replacement of specific mechanical components of the bleaching machine, addressed the study to the RUL prediction of the roller bearings as a whole subsystem. Based on their experience, the domain experts and senior operators can exploit these predictions to elucidate what are the most probable roller bearings to fail when an alert is raised by the model.

The future work will be focused not only on the improvement of the prediction performance of the model by hyperparameters tuning and feature selection, but also on the exploration of the incidence of different combinations of parameters in the feature extraction method, namely the size of the time windows, the number and size of the features windows and the size of the label window. In addition, efforts will be addressed to investigate other aggregations and featuring strategies to apply to the events enclosed into the features windows.

Acknowledgement

This work was developed in the framework of the project ”RECLAIM- RE-manufaCturing and Refurbishment LArge Industrial equipMent” and received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 869884. The authors would like to acknowledge the valuable collaboration of Zorluteks, for providing data and domain expert support.

REFERENCES

[1] C. Ji, Q. Shao, J. Sun, S. Liu, L. Pan, L. Wu, and C. Yang, "Device data ingestion for industrial big data platforms with a case study," *Sensors*, vol. 16, no. 3, 2016.

[2] G. Peralta, M. Iglesias-Urkia, M. Barcelo, R. Gomez, A. Moran, and J. Bilbao, "Fog computing based efficient iot scheme for the industry 4.0," in *2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM)*, 2017, pp. 1–6.

[3] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, "Prognostics and health management design for rotary machinery systems—reviews, methodology and applications," *Mechanical Systems and Signal Processing*, vol. 42, pp. 314–334, 01 2014.

[4] A. Mathur, K. F. Cavanaugh, K. R. Pattipati, P. K. Willett, and T. R. Galie, "Reasoning and modeling systems in diagnosis and prognosis," in *Component and Systems Diagnostics, Prognosis, and Health Management*, P. K. Willett and T. Kirubarajan, Eds., vol. 4389, International Society for Optics and Photonics. SPIE, 2001, pp. 194 – 203.

[5] L. Barajas and N. Srinivasa, "Real-time diagnostics, prognostics and health management for large-scale manufacturing maintenance systems," *Proceedings of the ASME International Manufacturing Science and Engineering Conference, MSEC2008*, vol. 2, 01 2008.

[6] J. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1803–1836, 2011.

[7] M. Ben-Daya, S. Duffuaa, A. Raouf, J. Knezevic, and D. Ait-Kadi, *Handbook of Maintenance Management and Engineering*, 01 2009.

[8] R. L. de Moura, L. B. Werner, and A. Gonzalez, "Management and ownership: A data strategy in the industry 4.0 context," in *Proceedings of the 3rd International Conference on Big Data and Internet of Things*, ser. BDIOT 2019. New York, NY, USA: Association for Computing Machinery, 2019, p. 23–28.

[9] Y. Peng, M. Dong, and M. Zuo, "Current status of machine prognostics in condition-based maintenance: a review," *Int J Adv Manuf Technol*, vol. 50, p. 297–313, 2010.

[10] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—part i: Fault diagnosis with model-based and signal-based approaches," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3757–3767, 2015.

[11] Y. Yuan, X. Tang, W. Zhou, W. Pan, X. Li, H.-T. Zhang, H. Ding, and J. Goncalves, "Data driven discovery of cyber physical systems," *Nature Communications*, vol. 10, 10 2019.

[12] Y. Yuan, H.-T. Zhang, Y. Wu, T. Zhu, and H. Ding, "Bayesian learning-based model-predictive vibration control for thin-walled workpiece machining processes," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 509–520, 2017.

[13] H. M. Hashemian, "State-of-the-art predictive maintenance techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 1, pp. 226–236, 2011.

[14] W. Zhang, D. Yang, and H. Wang, "Data-driven methods for predictive maintenance of industrial equipment: A survey," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2213–2227, 2019.

[15] M. Canizo, E. Onieva, A. Conde, S. Charramendieta, and S. Trujillo, "Real-time predictive maintenance for wind turbines using big data frameworks," in *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, 2017, pp. 70–77.

[16] C. Gutschi, N. Furian, J. Suschnigg, D. Neubacher, and S. Voessner, "Log-based predictive maintenance in discrete parts manufacturing," *Procedia CIRP*, vol. 79, pp. 528–533, 2019, 12th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 18-20 July 2018, Gulf of Naples, Italy.

[17] J. Wang, C. Li, S. Han, S. Sarkar, and X. Zhou, "Predictive maintenance based on event-log analysis: A case study," *IBM Journal of Research and Development*, vol. 61, no. 1, pp. 11:121–11:132, 2017.

[18] M. Calabrese, M. Cimmino, F. Fiume, M. Manfrin, L. Romeo, S. Ceccacci, M. Paolanti, G. Toscano, G. Ciandrini, A. Carrotta, M. Mengoni, E. Frontoni, and D. Kapetis, "Sophia: An event-based iot and machine learning architecture for predictive maintenance in industry 4.0," *Information*, vol. 11, no. 4, 2020.

[19] R. Sipos, D. Fradkin, F. Moerchen, and Z. Wang, "Log-based predictive maintenance," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 1867–1876.

[20] S. Xiang, D. Huang, and X. Li, "A generalized predictive framework for data driven prognostics and diagnostics using machine logs," in *TENCON 2018 - 2018 IEEE Region 10 Conference*, 2018, pp. 0695–0700.

[21] O. Maron and T. Lozano-Pérez, "A framework for multiple-instance learning," in *Advances in Neural Information Processing Systems*, M. Jordan, M. Kearns, and S. Solla, Eds., vol. 10. MIT Press, 1998.

[22] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to rul prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 799–834, 2018.

[23] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical Systems and Signal Processing*, vol. 107, pp. 241–265, 2018.

[24] T. Xia, Y. Dong, L. Xiao, S. Du, E. Pan, and L. Xi, "Recent advances in prognostics and health management for advanced manufacturing paradigms," *Reliability Engineering & System Safety*, vol. 178, pp. 255–268, 2018.

[25] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30. Curran Associates, Inc., 2017.