Predictive Maintenance in the Production of Steel Bars: A Data-Driven Approach

(Article begins on next page)

08 November 2022

Predictive maintenance in the production of steel bars: a data-driven approach

Paolo Bethaz, Xanthi Bampoula, Tania Cerquitelli, Nikolaos Nikolakis, Kosmas Alexopoulos, Enrico Macii, and Peter van Wilgen

Abstract The ever increasing demand for shorter production times and reduced production costs require manufacturing firms to bring down their production costs while preserving a smooth and flexible production process. To this aim, manufacturers could exploit data-driven techniques to monitor and assess equipment's operational state and anticipate some future failure. Sensor data acquisition, analysis, and correlation can create the equipment's digital footprint and create awareness on it through the entire life cycle allowing the shift from time-based preventive maintenance to predictive maintenance, reducing both maintenance and production costs.

In this work, a novel data analytics workflow is proposed combining the evaluation of an asset's degradation over time with a self-assessment loop. The proposed workflow can support real-time analytics at edge devices, thus, addressing the needs of modern cyber-physical production systems for decision-making support at the edge with short response times. A prototype implementation has been evaluated in use cases related to the steel industry.

Paolo Bethaz, Tania Cerquitelli
Department of Control and Computer Engineering, Politecnico di Torino, Torino, Italy e-mail: name.surname@polito.it

Xanthi Bampoula, Nikolaos Nikolakis
Laboratory for Manufacturing Systems & Automation, University of Patras, Greece, e-mail: surname@lms.mech.upatras.gr

Kosmas Alexopoulos
Laboratory for Manufacturing Systems & Automation, University of Patras, Greece, e-mail: alexokos@lms.mech.upatras.gr

Enrico Macii
Interuniversity Department of Regional and Urban Studies and Planning, Politecnico di Torino, Torino, Italy e-mail: name.surname@polito.it

Peter van Wilgen
VDL Weweler bv, Apeldoorn, The Netherlands, e-mail: p.van.wilgen@vdlweweler.nl

1 Introduction

Predictive maintenance policies have been used throughout the years, mostly based on human knowledge and intuition as a result of experience. As technology advances, intuition is seeked to be enhanced by computer techniques [1]. The large volume of data generated on a shop floor allow for the use of artificial intelligence techniques that can analyse and create insight over production processes, and as a result complement or support the human knowledge. Nevertheless, selecting the appropriate analysis methods as well as the availability of proper datasets remain a challenge.

VDL Weweler designs, develops, and produces trailing arms, among others, to manufacture trailers, trucks, buses, and cars. The production line of VDL Weweler is fully automated, including both machinery and robots. Maintenance activities, however, are in their great majority either preventive or corrective. Knowing the equipment's working condition and causes of the production interruptions could help identify the maintenance's root cause and restore the system to an operating state. To this aim, proper data-driven predictive maintenance techniques and scheduling for the replacement of segments on the rolling milling machine are discussed and tested. As a result, adequate maintenance planning facilitates further cost reduction and better production management.

We propose and discuss several approaches regarding features extraction and data labelling, considering different subsets of features extracted from the collected data, and labeling the historical set using two different strategies. All combinations between subset of extracted features and data labelling strategy have been tested and evaluated on real data in order to estimate which one is the most performing.

The chapter is organized as follows. Section 2 describes the previous state-of-the-art works present in literature and related to the proposed scenario. Section 3 offers an overview of the Steel production bar production industry, while section 4 details the real-life setting under analysis, focusing on its manufacturing process and the resulting maintenance needs. Then, section 5 describes all the methodology and architecture implemented to provide a data analytics service to the proposed use case, including the obtained experimental results. Finally, section 6 draws conclusions, offering a general summary of what is proposed in this paper.

2 Literature Review

With the introduction of Industry 4.0, smart environments have become very popular, promoting the frequent use of the Cyber-Physical System (CPS), which promotes full integration of manufacturing IT and control systems with physical objects embedded with software and sensors. In this new type of industry, the increased communication between production components leads to a large amount of data. In addition, the integration of Cyber-Physical Systems are encouraging modern industries to transform massive data into valuable knowledge, extracting knowledge about production systems and yielding the support to optimal decision-making [2], helping managers

to improve the production processes. The new challenge of modern industries is therefore to be able to effectively collect, process and analyze large amounts of data in real time. To do these tasks, several existing works [3, 4, 5, 6] use Big Data frameworks, facing the necessity of knowledge extraction. In particular, in [3] a Big Data analytics framework is presented, capable of providing a health monitoring application for an aerospace and aviation industrial. In [4] the authors use open source technologies such as Apache Spark and Kafka to implement a scalable architecture capable of processing data both online and offline. The same open source Big Data technologies are also used in [5], in order to implement an integrated Self-Tuning Engine for Predictive maintenance in Industry 4.0. The topic of predictive maintenance in a big data environment is also addressed in [6], where, with the purpose of monitoring the operation of wind turbines, a data-driven solution deployed in the cloud for predictive model generation is presented.

The applications of predictive maintenance have had a considerable diffusion with the advent of Industry 4.0, thanks to the introduction of sensors able to constantly monitor the performance of machinery. This work [7] presents how recent trends in Industry 4.0 solutions are influencing the development of manufacturing execution systems, while in [8] authors present a framework able to implement scalable, flexible and pluggable data analysis and real-time supervision systems for manufacturing environments. In [9, 5, 10] three data-driven methodologies related to the predictive maintenance services in an Industry 4.0 context are discussed.

But predictive maintenance does not only refer to failures detection, in fact it can also be used to estimate the Remaining Useful Life (RUL) of a machine. In many real-life settings, the time component (sensors measure signals that evolve over time) need to be considered, as mentioned in [11]. Possible approaches to time series data mining are wavelets, recurrent neural networks and convolutional neural networks. In [11] the authors discussed the different approaches, highlighting how neural networks work better than other models, though not significantly.

Work on a context similar to this study is described in [12], where authors present an application of an image processing system in the monitoring and control of the hot-rolling of steel bars. Unlike in that work, this study does not use visual recognition to characterise the products, but appropriate features extracted from the measurements taken by the sensors during the process.

3 Maintenance needs and challenges in the Steel bar production industry

VDL Weweler systems are robust and deliver reliability and cost-effective operation for on-highway applications in the most demanding operating environments. Manufacturing facilities in Apeldoorn (The Netherlands) are highly automated with robotics to play an essential role in the cost-effective and reliable production.

The newly designed production line of VDL Weweler brings together a series of processes that were previously operated separately. The trailing arms are now

hot-formed and then tempered in a single continuous process, in which the residual heat from the forming process can be reused for the tempering process. This process cut energy requirements by 35% and significantly reduce production time.

Another significant benefit of the new production line is the ability to perform 3D forming, making it possible to build a suspension system with fewer parts that are lighter and lend themselves better to modular construction. This whole new system integrates a series of production processes, previously performed separately due to space limitations, to form a single line.

VDL Weweler provides data for testing the data-driven solution presented in this chapter as well as the SERENA's architecture. Maintenance/repairing activities within the monitored equipment are provided as output. VDL Weweler gives technical feedback to validate the project developments and suggestions to improve the overall methodology to transfer it to other industrial sectors needing similar maintenance solutions quickly.

4 Steel bar production case study: present and future

The prediction of the behavior of the segments and their maintenance is of high importance. The current cycle time of trailing arms production is estimated to forty-five seconds in a working day that no unexpected failures occur. The replacement of parts takes place approximately after the presentation of 18.000 pieces, and the visual inspection of the components usually allows increasing their lifetime at least after 25.000 repeats. In this industrial scenario, the SERENA platform aims to increase the segment's lifetime, enabling predictive maintenance techniques to replace the details on the rolling machine and provide the operators with relevant information through Augmented Reality (AR) technology for maintenance operations. In parallel, by reducing the stoppages in the production line, products' quality should be improved.

The API-pro software, currently in use, manages and schedules the maintenance activities, which experienced operators usually perform. The entire replacement/maintenance process is estimated to be around one hour. Unfortunately, it is not a fixed time, as it can vary depending on the segments' temperature, and it can reach up to four hours, including the wait time for the parts to cool.

The rolling machine supplier does maintenance activities on the hydraulics unit and the device once per year. While maintenance operations are taking place, the production line stops because of serial production and the strict relationship among all the activities. The waiting time for maintenance activities on the rolling mill can vary from 1 to 4 hours. Through API-Pro Software, instructions are provided by documents and/or smart devices to the maintenance personnel.

4.1 Maintenance needs

For the SERENA pilot case, the focus is on the forming of trailing arms through a rolling mill type machine by predicting the in-time replacement of the coating segments used by the device. The trailing components are designed to suit specific vehicle models as well as detailed operational areas. Better comfort, payload, drivability, and lifetime are assured through the high quality and long service life of the VDL Weweler production line, ensuring better comfort and road-holding needed for safe and economical transportation. The trailing arms production line starts by heating steel bars of 1000x100x50 (mm) in the responsible station while the rolling process follows as depicted in Figure 1.
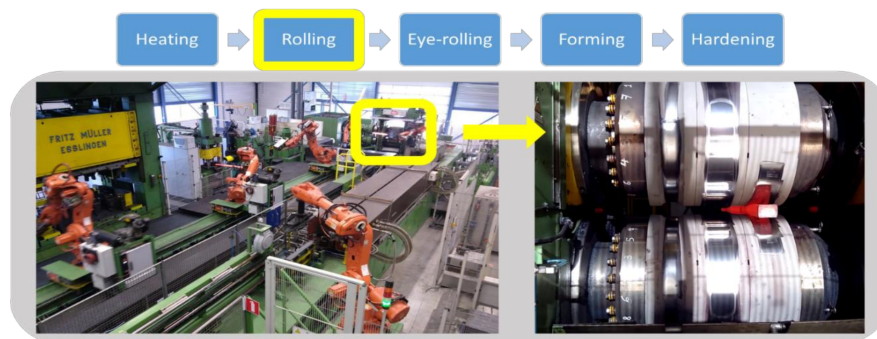

Fig. 1 VDL WEWELER Production Line

The current situation of the maintenance and repairing activities does not include a predictive maintenance approach. The SERENA platform discussed technically in Chapters 2, 3, and 4 allows exploring this possibility in various industrial sectors. The rolling mill machine used for the forming process is monitored by applying multiple sensors to the device. The acquired data is compensated with a digital twin model or a physics-based model and eventually achieves an accurate prediction of the replacement of segments. Additionally, the main benefits focus on decreasing the downtimes, reducing the exchange of components related costs, and improving the final product quality by exploiting the segments' life-cycle efficiently.

The SERENA platform gathers data from the rolling machine within the VDL Weweler, creates models for predicting maintenance needs, correlates the rolling machine data with the digital model, and then predicts the exchange of segments. The technical feedback received from VDL Weweler during the 36-month duration improves the SERENA developments orienting the results in real industrial applications.

The most critical equipment for monitoring includes the rolling mill machine. The focus is on predicting and scheduling the replacement of the coated segments, aiming at increasing, on the one hand, their lifetime and, on the other, improving the

trailing arms quality. The main three parameters that affect the segment's lifetime are high temperature, the friction between the trailing arms and components, and finally, the forces applied.

The purpose of the SERENA platform is to predict the replacement of segments of the rolling mill machine. The target is to accurately predict when the components need to be replaced through a collection of data from the milling machine's sensors and their correlation with a digital twin model. Additionally, and as the product's quality is strictly related to the milling machine's working conditions, a measuring system is designed and developed. This measuring system precisely calculates the formed trailing arms' straightness once the milling process is completed. Based on the measured values and their correlation with the milling machine's status, accurate maintenance predictions are foreseen. Moreover, the SERENA platform schedules the maintenance activities to reduce the production stoppage time and avoid any interruptions with the production plan. The maintenance operators are equipped with AR technologies for guiding them through correct task execution and training.

4.2 Equipment description

The steel parts production pilot case is focused on the rolling milling machine in Figure 2. This machine is composed of two rolling cylinders which are rotating through the use of torque motors. The lower rolling cylinder has a fixed position, and only the upper cylinder can move vertically. Three different geometrically coated segments are attached to the rolling cylinders. The segments are used in order to form the trailing arm by applying force. Currently, the segments are replaced after 18.000 repeats as preventative maintenance and for safety reasons in order not to completely destroy the segments.

4.3 Steel bar production process

The stretching process of the rolling mill machines includes several steps. The blank entering from the heating unit is fed sequentially to the individual passes by a robot. The blank is formed analogously to the contours of the individual roller grooves. The robot arbitrarily rotates the blank by up to 180° around its longitudinal axis between the separate passes. As the rollers can be removed from the forging roll for a tool or groove change, it is possible to work with entirely circular, closed roller grooves or with grooves in the form of circular segments. The robot and the rollers work in a master-slave mode during the rolling process, with the roller rotation angle acting as the master signal. This operation mode allows shock-free working and ensures almost zero wear compared to mechanical, rigidly coupled drive systems. The robot movement is synchronized with the rollers' servo drives so that the stretched workpiece is moved according to the prevailing peripheral speed

Coated Segments

Rolling Cylinder A

Rolling Cylinder A
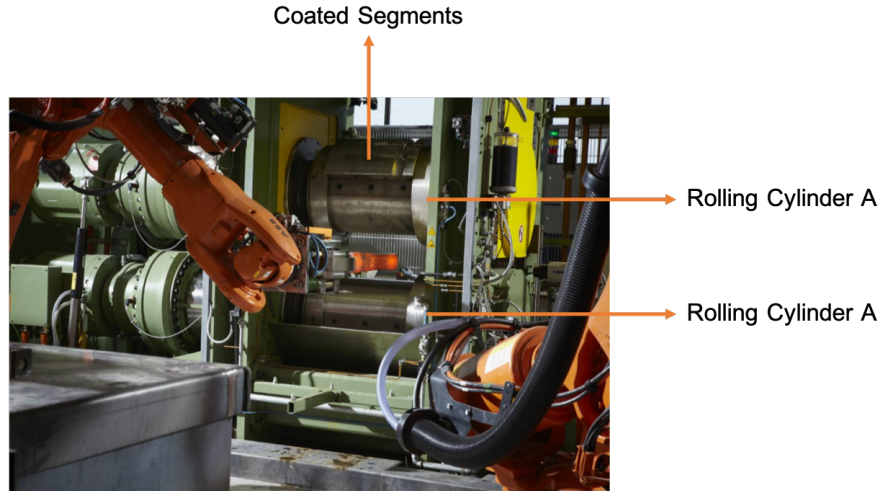
Fig. 2 Equipment Description

of the rollers. Thus, during operation, the rate of workpiece movement is matched to the speed profile.

A high-level illustration of the global architecture is displayed in Figure 4. As mentioned earlier, a steel metal bar is inserted in the heating station. The heating process takes approximately five minutes. Afterward, the robot R2 is responsible for picking up the heated metal bar and proceeds with the rolling mill's rolling process. For this operation, another robot, R3, is used, and the two robots are cooperating to achieve the required functions of the rolling process. At the end of the rolling process production, the robot R3 further inspects the outcome. In case of deviations in the steel bar geometry, the process parameters of the rolling mill need to be modified to reach the desirable geometrical characteristics.
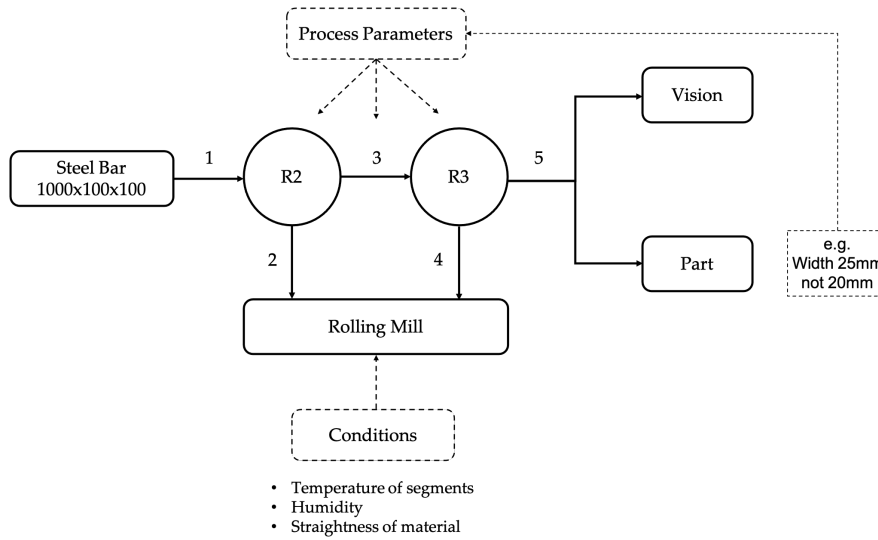
Fig. 3 Production Process Description

5 SERENA system in the steel bar production case

5.1 Architecture

The SERENA cloud platform is built on a lightweight micro-services architecture, that allows the core cloud services and the edge gateway to be managed as a single domain. The edge gateways are located close to the factory equipment that generates the data used to perform the maintenance predictions.
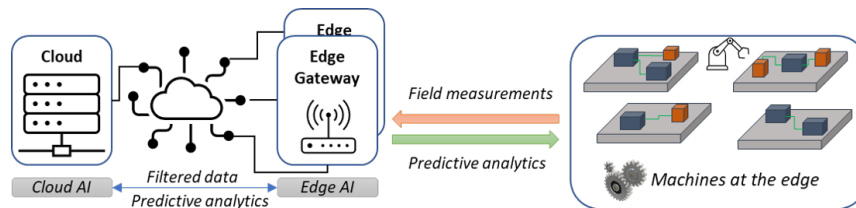
Fig. 4 *High level architecture of the global system*

All services running on the SERENA cloud platform are implemented as Docker containers and managed by Docker Swarm. The use of containers enables modularity in the SERENA cloud platform, making it simple to remove a given service and substitute it with an alternative implementation. Thus, the "plug-and-play" requirement

of the SERENA platform is achieved, giving end-users the freedom to choose the technology implementation that best suits their unique needs and corporate guidelines. It also helps to future-proof the platform. Older technologies can easily be updated and/or replaced with newer alternatives as they become available without modifying the underlying platform architecture. The plug-and-play concept is also applied to the repositories, which uses a RESTapi and canonical JSON-LD message format to facilitate communication with the repositories. In cases where tight integration is required, the APIs and the data warehouses are packaged as a group of services, and the whole group can be replaced as necessary. Services communicate via HTTP RESTapi, although other protocols, like MQTT, are also supported. The RPCA (Reverse Proxy Certification Authority) middleware acts as the interface to the edge gateways and external systems, providing security and routing services. Apache NiFi serves as the central communications broker between services.

5.2 Data connections

Figure 5 depicts the data connections for the rolling mill machine. Data from the rolling mill sensors are pre-processed to link the values collected from the force, torque, and position sensors to the rolling angle of machines' cylinders. To this aim, the data collected from the pressure, torque, and position sensors are converted into degrees in the historical data.
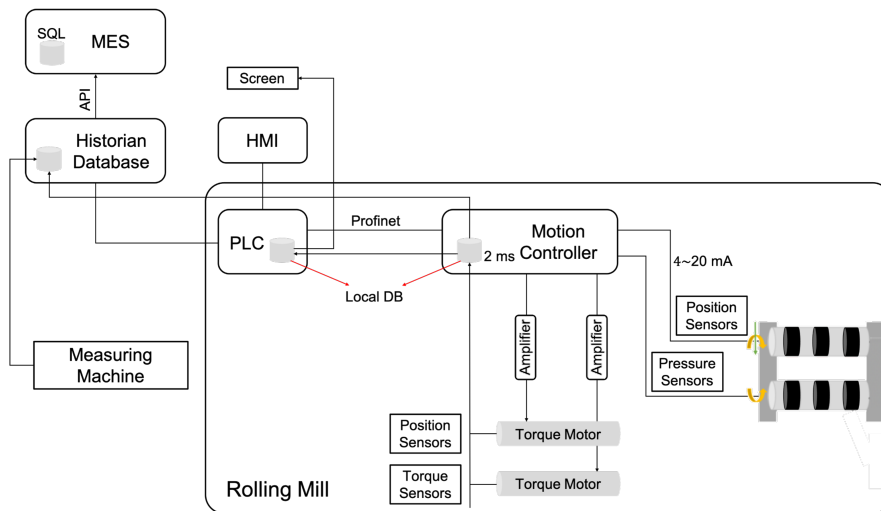

Fig. 5 Data Connections

5.3 Data acquisition

Several sensors have been integrated into the machine to acquire measurement over key parameters to support the inference of insight from machine processes data. Table 1 provides a list of the sensors applied in the rolling mill as well as to the standalone measuring system.

Table 1 Sensors.

Parameter	Equipment	Sensor	Measurements	Com Protocol	Location
Oil Temperature	Rolling Mill	Temperature transducer	Oil temperature	Profinet	Rolling Machine
Oil Pressure	Rolling Mill	Pressure transducer	Oil pressure	Profinet	Rolling Machine
Roughness	Standalone Measuring device	Roughness sensor	Oil pressure		Rolling Machine
Dimensions of the product	Standalone Measuring device	Thickness sensor	Thickness, width, and straightness of the product		After the rolling machine
Thickness of the coating layer	Standalone Measuring device	Thickness sensor	Thickness of the coating layer		Rolling Machine

5.4 Data Analytics

The objective of the implemented methodology is to estimate the RUL (residual useful life) value of machine cylinders, as these components are expensive and they wear out differently over time, making difficult to estimate their degradation. Having a real time prediction of the machine's RUL value every time a new product is produced is important for the company to be able to estimate the level of degradation of the machinery, thus intervening promptly with maintenance, saving time and money. The proposed methodology consists of 3 main building blocks, which are: i) *Feature engineering*, ii) *Data labeling*, iii) *Model training and validation*.

5.4.1 Feature Engineering

In this block we have extracted significant features from the various input signals collected by the machinery. Each signal is summarised through the use of several

features, able to characterise it. The extraction of these features has been done taking into account both the measuring machines and the rolling mill measurement.

From each input signal, we extracted the following 7 statistical features from the data collected by the measuring machine: minimum value, maximum value, mean value, standard deviation, variance, kurtosis (statistical index relating to the form of the distribution reflecting the concentration of data around its own average) and skewness (symmetry index of a distribution). In addition to this feature set we also include an eighth feature calculated as the distance between the signal collected in the measuring machine and the corresponding signal measured in the rolling mill. The error between the two series was calculated using the measure root-mean-square error (RMSE). In addition, the error of one measurement cycle also takes into account the errors of previous cycles, by adding them cumulatively, thus assuming the ever-increasing error over time.

5.4.2 Data Labeling

In some industrial context, it may be possible that no label is associated with the signals collected by the machinery. In these cases, in order to use a predictive methodology able to estimate the residual value of a machine, a previous data labelling step is required. This step is usually done manually by a domain expert, who is able through his knowledge to evaluate various signals and label them properly. However, this operation is very time-consuming and in an industry 4.0 context, where the various processes are robotized and automized, it would be useful to have an automatic methodology able to perform also the data labelling step.

So, in this section we offer two different methodologies able to automatically label the data collected by the machinery, using as a label the estimation of the degradation of the machinery itself. In particular, the label we want to assign is a decreasing numerical value that represents the RUL (residual useful life) of the machine. The smaller this value, the greater the risk of the machine breaking down. Knowing in advance the dates on which some components of the machinery have been replaced or in which maintenance has been done, we have assigned the same range of labels (from the highest to the lowest value) to each time interval between two of these consecutive dates. In this way, a cycle immediately after the replacement had a maximum label value, while a cycle just before the replacement had a minimum label value. In particular, we defined the RUL using two different strategies:

- the first strategy assigns the RUL a **linearly** decreasing trend over time;
- the second strategy assigns the RUL an **exponentially** decreasing trend over time.

The difference between these two strategies lies in the trend of the RUL over time. While the first strategy assumes that the degradation is constant during the various production cycles, the second strategy is based on the idea that in the first production cycles the performance of the machinery degrades much more slowly than in the final cycles.

In the first strategy, the formula for defining the RUL labels can be expressed as follows:

$$RUL = ceiling\left[(1 - X/X_{\text{tot}}) * n\right]$$

where X represents the current day or the current cycle (depending on whether we want to consider the RUL as linearly dependent with production cycles or with working days), X_{tot} is the total number of days or the total number of cycles, *n* is a parameter that we can set manually in order to define the maximum value we want to use in our labels, and ceiling is a function that can transform a floating number into its immediately greater integer.

In the second strategy, the formula defining the exponential trend of the RUL is as follows:

$$RUL = ceiling\left[-e^{\frac{ln(n+1)}{X_{\text{tot}}}*X} + (n+1)\right]$$

where the parameters take on the same meaning as in the previous formula. Here, since the generic exponential function $-e^x$ is -1 when the x value is 0, we have introduced the term *n+1* so that when x is 0 (first production cycle considered), the RUL of the machine is assigned the maximum label value. Moreover, the fraction that multiplies X to the exponent causes the RUL to be 0 (the machinery needs an intervention) when the value on the x-axis is equal to the number of total cycles.

5.4.3 Model training and validation

Finally, the purpose of this block is to build a classification model capable of predicting the correct RUL value of a new cycle, based on the values learned from the historical data. The training of the model is done using the features extracted in the features computation block, in addition to the label assigned in the data labelling block (if no original labels were present). Two state-of-the-art supervised learning algorithms have been tested in these blocks: Decision Tree and Random Forest, where the most performing parameters of each algorithm are chosen thanks to a self-tuned strategy based on a grid optimisation search over the classification algorithm parameters. The proposed methodology compares the performances obtained with the two different algorithms, highlighting the best one.

To compare the performances of different models, the accuracy is given for each one of them. Moreover, for each of the predicted labels, we also report the recall and precision metrics, which are particularly useful in case the dataset is not balanced. The metrics are defined as follows:

Accuracy: Correct Predictions / Total Predictions

Recall = True Positive / (True Positive + False Positive)

Precision = True Positive / (True Positive + False Negative)

5.5 Results

This section shows the results obtained on the analyzed use case using the techniques described in Section 5.4. Here, the purpose is to verify how well the real-time prediction methodology performs on the data under analysis, using both the approaches reported in 5.4.2 to label the data.

The available data are composed of 9756 production cycles collected between 22/01/2020 and 24/02/2020. During this period, the segment was replaced on the machine at two different times: 04/02/2020 and 19/02/2020. These maintenance dates allow us to divide the production cycles into three different production groups within which the RUL of the machinery will assume all the values from the maximum (first cycle) to the minimum (cycle close to maintenance). In order to correctly assign the RUL labels, we need to know for each group the start date and end date of production. Because when we collected the data, the third production group was not yet finished, we have considered in our training model only the first two groups: a total of 5233 production cycles collected between 22/01/2020 and 19/02/2020.

Since the number of cycles is not very high and we have only two complete production groups available, we used the leave one out cross validation (loocv) technique to test the performance of the predictive model. Loocv allows us to use a single production cycle as a test set, and all other cycles as a training set. This operation is repeated iteratively, using each time a different production cycle as a test set. Then, at each iteration a new predictive model is built to predict the RUL value of the cycle used in the test set. So, at the end of this process, each cycle was used once in the test set and was labeled with a predictive RUL value.

In the following paragraphs, for each proposed configuration, the results are compared with those obtained from a baseline methodology, a more traditional approach to address feature engineering in use cases similar to the one discussed here. This baseline methodology adopts the same approach as ours with regard to data labelling; instead, in the feature engineering block, the only features extracted are the seven statistics: cumulative error between rolling mill measurements is not considered.

All the following experiments were conducted using the current production cycle as the X value in the formula introduced in section 5.4.2. This means that we consider the life of the machinery to be decreasing over the production cycles.

5.5.1 Experiment with linear approach

Since in this case we assume that the RUL has a linear trend over time, each production group will be divided into n subgroups of equal size (production cycles), where n is the number of labels we want to use in our experiments. We have manually set this parameter to 10 and 100. Table 2 and Table 3 show the accuracy results obtained using these two configurations, while the confusion matrices shown in Figure 6 and Figure 7 refer to the case with 10 labels using the Decision Tree

classifier. The columns represent the predicted labels, while the rows represent the real labels. Values in bold on the diagonal are the correct predictions.

Table 2 Accuracy using 10 labels.

	Decision Tree	Random Forest
cumulative error (c.e)	0.26	0.30
Baseline	0.68	0.72
Baseline + c.e.	0.99	0.97

Table 3 Accuracy using 100 labels.

	Decision Tree	Random Forest
cumulative error (c.e.)	0.19	0.22
Baseline	0.28	0.32
Baseline + c.e.	0.93	0.69

From these results we can see that only the statistical features (baseline) or only the cumulative error are not sufficient input variables to guarantee a good performance of the predictive model, while their combination allows to reach very good results. The low level of accuracy obtained by using only the error as an input variable can be explained by the fact that very different signals can be equally distanced from a Rolling Mill's signal measurements; so the error alone is not able to effectively characterize the signal. The same thing can be said if we use the baseline approach with only the statistical features, as signals with different trends can be characterized by very similar set of features.

The best performance obtained using both variables is confirmed in Figure 7: from this confusion matrix we can see that only 47 cycles out of a total of 5233 were not correctly labeled. In addition, most of the errors are near the diagonal of the matrix, indicating that almost all of the errors made by the classifier occurred between adjacent classes. A cycle with a high residual value has never been labeled with a low label, or vice versa.

Using a higher number of labels, obviously the possibilities of error increase and consequently the accuracy values decrease, but, as shown in Table 3, the results are still good and in line with what was obtained with 10 labels.

	1	2	3	4	5	6	7	8	9	10
1	**425**	37	10	30	12	8	0	0	0	0
2	34	**388**	61	20	20	0	0	0	1	0
3	15	60	**338**	70	35	1	0	3	1	0
4	32	19	82	**293**	76	5	1	13	2	0
5	7	17	42	82	**334**	27	0	14	0	1
6	5	0	2	4	29	**402**	78	0	1	2
7	0	0	1	2	0	90	**325**	53	41	11
8	0	0	2	15	12	3	70	**322**	68	31
9	0	0	0	1	1	2	32	57	**357**	74
10	0	0	0	0	4	1	17	37	76	**389**

RECALL	0.81	0.74	0.64	0.56	0.63	0.76	0.62	0.61	0.68	0.74
PRECISION	0.82	0.74	0.62	0.56	0.63	0.74	0.62	0.64	0.65	0.76

Fig. 6 *Baseline Confusion Matrix*

	1	2	3	4	5	6	7	8	9	10
1	**521**	1	0	0	0	0	0	0	0	0
2	1	**521**	1	0	0	1	0	0	0	0
3	0	2	**520**	1	0	0	0	0	0	0
4	0	0	1	**521**	1	0	0	0	0	0
5	0	0	0	2	**519**	3	0	0	0	0
6	0	0	0	0	2	**519**	2	0	0	0
7	0	0	0	2	0	3	**514**	2	2	0
8	0	0	0	0	0	1	4	**516**	2	0
9	0	0	0	0	0	1	3	2	**515**	3
10	0	0	0	0	0	0	0	0	4	**520**

RECALL	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.98	0.98	0.99
PRECISION	0.99	0.99	0.99	0.99	0.99	0.98	0.98	0.99	0.98	0.99

Fig. 7 *New Methodology Confusion Matrix*

5.5.2 Experiment with exponential approach

The purpose of this section is to evaluate the performance of the predictive model, assuming that the degradation of the machinery has an exponentially decreasing trend. Labeling the production cycles based on this hypothesis, the dataset will no longer be balanced; in fact, since the degradation will increase much faster at the end of the production group, the cycles that will be assigned a high RUL value will be more numerous than those with a low RUL value. In particular, each class will contain always fewer elements than the previous class. Table 4 contains the percentages of elements that have been assigned to each class after the data labelling step.

Table 4 Labels percentage distribution.

Label	10	9	8	7	6	5	4	3	2	1
Occurrences(%)	28.9	16.8	12.0	9.3	7.6	6.5	5.6	4.9	4.5	3.9

Also in this case the results obtained are very good and are on average even better than those obtained in the linear case. Table 5 contains the accuracy values with the different variables used as inputs and we can see that with our methodology the performance of the predictive model is improved compared with other conditions, as also demonstrated by the two confusion matrices shown in Figure 8 and Figure 9.

Table 5 Accuracy value using exponential approach.

	Decision Tree	Random Forest
cumulative error (c.e.)	0.39	0.42
Baseline	0.70	0.75
Baseline + c.e.	0.99	0.97

A comparison between the matrices shows that the baseline methodology performs well when it has to recognise the optimal condition of the machinery (label 9 or 10), as, due to the exponential trend in degradation, these labels are the most numerous. However, when degradation increases and the labels defining the RUL of the machinery rapidly change, this methodology is no longer able to assign each production run to the correct label, as demonstrated by the recall and accuracy values in Figure 8. Instead, using our new approach, the predictive model is able to correctly recognise the labels not only of the initial classes, but also of those in which degradation is greater, like shown in Figure 9.

	1	2	3	4	5	6	7	8	9	10
1	**122**	30	28	8	1	2	4	12	0	0
2	32	**144**	20	1	8	5	15	5	0	0
3	19	17	**141**	53	6	6	11	4	0	0
4	9	2	45	**167**	36	11	17	5	0	0
5	3	8	0	36	**194**	51	32	12	0	0
6	1	2	7	17	60	**203**	63	40	1	4
7	1	11	11	10	25	78	**247**	81	0	22
8	13	4	4	6	16	35	90	**412**	32	17
9	0	0	0	0	0	2	1	40	**718**	123
10	0	0	0	0	1	5	19	20	133	**1336**

RECALL	0.59	0.62	0.54	0.57	0.57	0.51	0.50	0.65	0.81	0.88
PRECISION	0.61	0.66	0.55	0.56	0.55	0.51	0.49	0.65	0.81	0.88

Fig. 8 *Baseline Confusion Matrix*

	1	2	3	4	5	6	7	8	9	10
1	**207**	0	0	0	0	0	0	0	0	0
2	1	**227**	1	1	0	0	0	0	0	0
3	0	1	**254**	2	0	0	0	0	0	0
4	0	0	2	**289**	1	0	0	0	0	0
5	0	0	0	1	**335**	0	0	0	0	0
6	0	0	0	0	1	**396**	1	0	0	0
7	0	0	0	2	0	1	**484**	1	0	0
8	0	0	0	1	0	0	2	**624**	2	0
9	0	0	0	0	0	2	0	3	**878**	1
10	0	0	0	0	0	0	0	0	4	**1510**

RECALL	1	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99
PRECISION	0.99	0.99	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99

Fig. 9 *New Methodology Confusion Matrix*

5.6 System deployment

The proposed data-driven methodology has been implemented in a python package and integrated as a service in the SERENA cloud platform described in Chapters 2 and 4. The cloud platform is built upon lightweight micro-services bridging the gap between the edge and the cloud. The services enabling its functionalities are implemented as Docker containers enabling scalability and modularity. Whilst edge devices usually consist of a separate network from the cloud infrastructure, cloud services and edge gateways along with edge devices/sensors can be managed as a single domain.

Edge components are deployed at the edge gateway as docker containers. All the services are centrally managed via the Docker Swarm. The use of containers allows for a single service to be easy to replace in terms of functionality and/or technology. Therefore, the cloud platform with its functionalities is implemented onto a stack of indicative technologies that can be replaced with alternative ones, granting its user with the freedom to select the technologies that suit his/her specific needs.

The communication throughout the system is enabled using HTTP REST APIs and canonical JSON-LD messages without excluding other types of communication such as MQTT. The JSON data format is partially built upon the MIMOSA open standard for asset maintenance and, in particular, the CRIS 3.2 schema. The JSON messages are forwarded from the edge gateways to the cloud via RPCA, creating a security middleware between edge gateways and external systems. A central communication broker is implemented using Apache NiFi.

The platform is designed and implemented to store and process vast volumes of semi-structured data. To this end, the Apache Hadoop framework has been adopted, allowing for distributed processing across computer clusters. Data are consumed by integrated applications for analytic, scheduling, visualization purposes that can be extended to include additional functionalities. In this work, the focus is placed on the analytics service.

6 Discussion and Conclusion

The steel products manufacturing industry represents one of the SERENA use cases where the platform's innovative solutions were deployed and tested. The main cloud platform was instantiated to best address the specific requirements of the use case, while the focus was paid upon data driven analytics.

The aforementioned use case provides a testbed for testing and validating the proposed analytics methodology towards detecting abnormal events and imminent failures. It should be noted that while the SERENA solutions were fitted in the existing production facilities relatively seamlessly, the proper training of the AI algorithms was challenging. However, the results obtained on a collection of real-data are promising. There is a clear improvement over the performance obtained with the traditional approach. This indicates that considering both rolling mill and

measuring machine measurements can bring significant benefit to the predictive model. Moreover, the proposed methodology is general purpose and takes into account different degradation functions over time. A methodology like the one proposed in this chapter is able to adapt to the RUL prediction of different components with different degradation trends.

The proposed approach can therefore be decoupled from the specific use-case on which it was tested in this chapter, generalising it to other contexts of industrial steel production.

References

1. T. Cerquitelli, D.J. Pagliari, A. Calimera, L. Bottaccioli, E. Patti, A. Acquaviva, M. Poncino, Manufacturing as a data-driven practice: Methodologies, technologies, and tools, Proceedings of the IEEE (2021)

2. J. Lee, H.D. Ardakani, S. Yang, B. Bagheri, Industrial big data analytics and cyber-physical systems for future maintenance & service innovation, Procedia CIRP **38**, 3 (2015)

3. B. Xu, S.A. Kumar, Big data analytics framework for system health monitoring, in *2015 IEEE International Congress on Big Data* (2015), pp. 401–408. DOI 10.1109/BigDataCongress.2015.66

4. G.M. D'silva, A. Khan, Gaurav, S. Bari, Real-time processing of iot events with historic data using apache kafka and apache spark with dashing framework, in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)* (2017), pp. 1804–1809. DOI 10.1109/RTEICT.2017.8256910

5. D. Apiletti, C. Barberis, T. Cerquitelli, A. Macii, E. Macii, M. Poncino, F. Ventura, istep, an integrated self-tuning engine for predictive maintenance in industry 4.0, in *IEEE International Conference on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications, ISPA/IUCC/BDCloud/SocialCom/SustainCom 2018, Melbourne, Australia, December 11-13, 2018*, ed. by J. Chen, L.T. Yang (IEEE, 2018), pp. 924–931

6. M. Canizo, E. Onieva, A. Conde, S. Charramendieta, 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. DOI 10.1109/ICPHM.2017.7998308

7. S. Jaskó, A. Skrop, T. Holczinger, T. Chován, J. Abonyi, Development of manufacturing execution systems in accordance with industry 4.0 requirements: A review of standard- and ontology-based methodologies and tools, Computers in Industry **123**, 103300 (2020). DOI https://doi.org/10.1016/j.compind.2020.103300. URL http://www.sciencedirect.com/science/article/pii/S0166361520305340

8. R.S. Peres, A. Dionisio Rocha, P. Leitao, J. Barata, Idarts – towards intelligent data analysis and real-time supervision for industry 4.0, Computers in Industry **101**, 138 (2018). DOI https://doi.org/10.1016/j.compind.2018.07.004. URL http://www.sciencedirect.com/science/article/pii/S0166361517306759

9. S. Proto, F. Ventura, D. Apiletti, T. Cerquitelli, E. Baralis, E. Macii, A. Macii, Premises, a scalable data-driven service to predict alarms in slowly-degrading multi-cycle industrial

processes, in *2019 IEEE International Congress on Big Data, BigData Congress 2019, Milan, Italy, July 8-13, 2019*, ed. by E. Bertino, C.K. Chang, P. Chen, E. Damiani, M. Goul, K. Oyama (IEEE, 2019), pp. 139–143

10. F. Ventura, S. Proto, D. Apiletti, T. Cerquitelli, S. Panicucci, E. Baralis, E. Macii, A. Macii, A new unsupervised predictive-model self-assessment approach that scales, in *2019 IEEE International Congress on Big Data, BigData Congress 2019, Milan, Italy, July 8-13, 2019*, ed. by E. Bertino, C.K. Chang, P. Chen, E. Damiani, M. Goul, K. Oyama (IEEE, 2019), pp. 144–148

11. S. Munikoti, L. Das, B. Natarajan, B. Srinivasan, Data-driven approaches for diagnosis of incipient faults in dc motors, IEEE Transactions on Industrial Informatics **15**(9), 5299 (2019). DOI 10.1109/TII.2019.2895132

12. P. O'Leary, Machine vision for feedback control in a steel rolling mill, Computers in Industry **56**(8), 997 (2005). DOI https://doi.org/10.1016/j.compind.2005.05.023. URL http://www.sciencedirect.com/science/article/pii/S0166361505001351. Machine Vision Special Issue