H2020 - INDUSTRIAL LEADERSHIP - Information and Communication Technologies (ICT)

ICT-38-2020: Artificial intelligence for manufacturing

Grant Agreement: 957362

Explainable Manufacturing Artificial Intelligence

WP4: Novel Artificial Intelligence Algorithms for Industrial Data Insights Generation

D4.1: Draft Catalogue of XMANAI AI and Graph Machine Learning Models

Deliverable Leader: A<**Deliverable**_leader>

Due Date: M15

Dissemination Level: Public

Version: F1.0

Short Abstract

This deliverable reports on the construction of the Draft catalogue of XMANAI AI and Graph ML models, as a collection of baseline algorithms and explanation methods that will be used to develop explainable solutions to address 4 generic manufacturing application scenarios, represented by the XMANAI demonstrators: i) Production Optimization, ii) Product Demand Forecasting, iii) Process/Product quality Optimization, and iv) Process Optimization and Semi-Autonomous Planning. A landscape analysis on the relevant application domains is conducted, followed by the technical description of the problems to be tackled and the identification of relevant data sources, leading to the selection of Hybrid and Graph baseline models that will populate the initial release of the XMANAI Explainable AI platform.

Document Log

Contributors	ATHENA, TYRIS, AiDEAS, Suite5, DBL, FORD, WHR, CNH, UNIMETRIK
Internal Reviewer 1	UBITECH
Internal Reviewer 2	POLIMI
Type	Report

History

Versions	Description
D0.1	Draft Table of Contents and initial inputs
D0.2	Final ToC and contributions to the manuscript
D0.3	Inclusion of references, harmonization of contents along the sections
R0.1	Revision of internal reviewer 1
R0.2	Revision of internal reviewer 2
F1.0	Final

Executive Summary

Deliverable D4.1 “Draft Catalogue of XMANAI AI and Graph Machine Learning Models” provides the initial selection of AI and Graph Machine Learning (ML) baseline algorithms that will populate the Models Catalogue of the XMANAI Explainable AI Platform. The novelty introduced by the XMANAI project mainly lies in the construction of flexible, eXplainable AI (XAI) solutions to concrete manufacturing problems by means of composite algorithms, comprised of “black-box” models equipped with additional interpretability layers.

This document was produced as a result of combined activities realizing tasks T4.1 “Graph Machine Learning Algorithm Modelling and Training” and T4.2 “Hybrid Explainable AI Modelling and Training”. The research results presented herein represent the efforts of scientific/technical partners involved in the WP4 activities, in close collaboration to the 4 industrial pilots, during months M9-M15 of the XMANAI project. Regarding the demonstrators, this document summarizes the analysis of their AI and explainability needs, technical challenges as well as their security and ethics aspects. As a result of these activities, 48 Hybrid ML models and 25 Graph ML Models are provided as part of the Draft Catalogue. These models are complemented with 15 explainability tools that help to interpret the predictions made by the models.

The conducted research is focused on the 4 manufacturing application scenarios reflected by the XMANAI demonstrators, in direct consideration of business requirements elaborated in D6.1-“Demonstrators Requirements” and the technical requirements in D1.2 “XMANAI Concept Detailing, Initial Requirements, Usage Scenarios and Draft MVP”. Building on the findings of D1.1-“State of the Art Review in XMANAI Research Domains”, a landscape analysis is conducted over AI solutions for Industrial Data Insights generation, in the application scenarios of the XMANAI demonstrators: i) Production Optimization, ii) Product Demand Forecasting, iii) Process/Product quality Optimization, and iv) Process Optimization and Semi-Autonomous Planning. With the aim to extend D1.1 results towards a more technical perspective, the analysis here explores the literature to review the adoption of various families of algorithms to solve specific problems in the context of the 4 generic application scenarios.

An in-depth analysis is further performed, delving into the technical details of the 4 application scenarios, as expected to be realized by the XMANAI demonstrators, including: i) the comprehension of each use case within the realistic environment of the demonstrator, ii) the identification of specific tasks and sub-tasks to undertake in order to address each use case, iii) the mapping of identified problems to relevant data sources that will be provided by the demonstrators to solve them, iv) the mathematical formulation and modelling of the problems at hand, by means of inputs and expected outputs, v) the comprehension of the different explainability needs that are expected to be satisfied by the XMANAI solutions, vi) the identification of technical challenges and limitations to overcome in order to reach the desired XAI solution.

Based on the results of the analysis, a set of Hybrid and Graph ML baseline models is selected with the scope to cover the needs of all the identified problems, as regards both performance and interpretability. To that end, baseline models are constructed as composites of AI and Graph ML primary algorithms, coupled with an additional explainability component. Under this approach, the XMANAI baseline models are anticipated to overcome the trade-off between AI model performance and the interpretability of the model’s behaviour by non-expert humans, since different components are dedicated to the optimization of the performance and the provision of high-quality explanations. The selected models are presented by means of algorithm cards, summarizing information on the model’s insights in four axes: i) the predictive layer, represented by the primary AI/Graph ML algorithm, ii) the explainability layer, represented by the selected explainability method, iii) the description of the general application scenario, iv) the description of the application scenario in the experimental setting of the XMANAI demonstrators.

Finally, the overview of the Draft Catalogue is presented, with respect to the selected families of AI algorithms and explainability methods. The Draft Catalogue will serve as the basis for the development of concrete XAI solutions to the XMANAI pilots, in the context of further WP4 activities.

Table of Contents

List of Figures

List of Tables

1 Introduction

This deliverable will provide the list of baseline models, that are generally appropriate for the manufacturing domain and for the different pilot scenarios to be supported by XMANAI, identifying their main parameters and structural elements, as well as their inputs, outputs and usage scenarios. As a result of the different actions carried out in the course to produce this deliverable, a draft catalogue of baseline models will be provided. This catalogue will be the basis for the work that will be performed on the rest of the WP4 deliverables jointly with the demonstrators' data.

1.1 XMANAI Project Overview

Despite the indisputable benefits that AI can bring in society as well as in any industry, humans typically have little visibility and knowledge on how AI systems make any decisions or predictions due to the "black-box effect" in which many of the machine learning/deep learning algorithms are opaque and not possible to be examined after their execution to understand how and why a decision has been made. In this context, XMANAI aims at rendering humans (especially, business experts in manufacturing) capable of fully understanding how decisions have been reached and what has influenced them in order to trust the AI systems.

Building on the latest AI advancements and technological breakthroughs, XMANAI shall focus its research activities on Explainable AI (XAI) in order to make the AI models, step-by-step understandable and actionable at multiple layers (data-model-results) in order to: (a) accelerate business adoption ("if manufacturers do not understand why/how a decision/prediction is reached, they shall not adopt/enforce it"), and (b) foster improved human/machine intelligence collaboration in manufacturing decision making without crossing lines, while ensuring regulatory compliance. In order to produce "glass box" AI models that are explainable to a "human-in-the-loop", without greatly sacrificing AI performance, XMANAI will deliver appropriate methods and techniques to address a number of AI-related challenges that currently constitute significant data scientists' pains (such as lifecycle management, security and trusted sharing of complex AI assets including data and AI models), and to effectively navigate the AI's "transparency paradox".

XMANAI aims to design, develop and deploy a **novel Explainable AI Platform** powered by explainable AI models that inspire trust, augment human cognition and solve concrete manufacturing problems with value-based explanations. Adopting the mentality that "AI systems should think like humans, act like humans, think rationally, and act rationally", a catalogue of **hybrid and graph AI models** is built, fine-tuned and validated in XMANAI at 2 levels: (a) baseline AI models that will be reusable to address any manufacturing problem, and (b) trained AI models that have been trained and validated for the different problems that the XMANAI demonstrators target, based on collaboration between humans (data scientists, data engineers and business experts). A bundle of **innovative manufacturing applications and services** are also built on top of the XMANAI Explainable AI Platform, leveraging the XMANAI catalogue of baseline and trained AI models.

XMANAI will validate its AI platform, its catalogue of hybrid and graph AI models and its manufacturing apps in **4 realistic, exemplary manufacturing demonstrators** with high impact in: (a) optimizing performance and manufacturing products' and processes' quality, (b) accurately forecasting product demand, (c) production optimization and predictive maintenance, and (d) enabling agile planning processes. Through a scalable approach towards Explainable and Trustful AI as dictated and supported in XMANAI, manufacturers will be able to develop a robust AI capability that is less artificial and more intelligent at human and corporate levels in a win-win manner.

1.2 Deliverable Purpose and Scope

Deliverable D4.1:" Draft Catalogue of XMANAI AI and Graph Machine Learning Models" reports on the results of "WP4 - Novel Artificial Intelligence Algorithms for Industrial Data Insights Generation" efforts, towards the identification of a set of baseline Hybrid and Graph ML algorithms that are suitable to address common use cases in the manufacturing domain. The selected baseline algorithms are coupled with interpretable components, in order to populate the XMANAI Explainable AI Platform and support the needs of the XMANAI manufacturing demonstrators.

The purpose of the present deliverable D4.1 is to describe in detail the construction of the ”Draft Catalogue of XMANAI AI and Graph Machine Learning Models”, as a result of consortium partners' collaborative efforts in the context of WP4 activities. This work realizes the initial stages of tasks T4.1 “Graph Machine Learning Algorithm Modelling and Training” and T4.2 “Hybrid Explainable AI Modelling and Training”, which run in parallel since M9 until the end of the project (M42) to cover the technical aspects of the selection and delivery of XMANAI AI and Graph Machine Learning Models. As a first step, a landscape analysis is conducted over the application of AI methods in industry, focused on the use cases presented by the XMANAI demonstrators. Utilizing D6.1 “Demonstrators Requirements” and the technical requirements in D1.2 “XMANAI Concept Detailing, Initial Requirements, Usage Scenarios and Draft MVP” as points of origin to define the specifics of manufacturing problems to be addressed, D1.1 “State of the Art Review in XMANAI Research Domains” is further used as a roadmap to derive novel explainable AI solutions to these problems. This initial selection of baseline algorithms is primarily driven from the comprehension of the problems put forward by the 4 XMANAI demonstrators and justified by the explainability needs and the technical challenges that arise in each case.

Although focused on the XMANAI demonstrators, the draft catalogue of explainable AI algorithms is anticipated to be reusable to address generic manufacturing problems, since the 4 demonstrators are considered representative of common AI applications in manufacturing. The draft catalogue will be implemented and made available on the XMANAI Explainable AI Platform, with the use of components defined in D5.1-“System Architecture, Bundles Placement Plan and APIs Design”, and the corresponding sub-components detailed in D3.1-“AI Bundles Methods and System Designs and D2.1-“Asset Management Bundles Methods and System Designs”. The final product is expected to be aligned with the contents and definitions in D1.2-“XMANAI Concept Detailing, Initial Requirements, Usage Scenarios and Draft MVP”.

The baseline algorithms selected in this first iteration of WP4 activities, will be subsequently leveraged to train, optimize and evaluate Hybrid and Graph ML models, fine-tuned to fit the specific needs of the XMANAI demonstrators, in terms of performance as well as interpretability. The demonstrator data used to this end, will be mapped to the XMANAI data model(s) presented in D3.1 “AI Bundles Methods and System Designs”, also processed and managed by methods and tools defined in D2.1 “Asset Management Bundles Methods and System Designs”. Model validation and optimization will take place in three sequential rounds as the activities of T4.3-“Cross-Validation and Experts Evaluation of XMANAI AI models” progress, taking security and ethics into account in each step of the process, in the framework of T4.4-“Ethics and Security in XMANAI AI Models”. Building on D4.1-”Draft Catalogue of XMANAI AI and Graph Machine Learning Models”, tasks T4.1-T4.4 will jointly result in the intermediate and final selection of trained Hybrid and Graph ML models that the XMANAI Explainable AI Platform will deliver. These trained models will be documented accordingly in the following WP4 deliverables D4.2, D4.3 and D4.4, due in months M21, M32 and M42 of the project respectively.

1.3 Impact and Target Audiences

Due to the technical content of this deliverable, the document is mainly addressed to the Data Scientists and Data engineers, while the Business user is not directly targeted at this point. This is a Draft version of the XMANAI catalogue of explainable AI models, where the foundations are set to ensure the theoretical and technical support of explainable solutions to the manufacturing use cases presented by the demonstrators. The impact of the proposed XAI solutions to all the targeted end users (Data Scientist, Data Engineer and Business User) will be assessed during the development of concrete solutions, as part of the experimental phase of the project.

1.4 Deliverable Methodology

The contents of the present document are the product of collaborative work between business users, scientific and technical partners involved in WP4 activities. The methodology adopted in order to produce the deliverable was discussed and agreed by all involved parties, and can be summarized as follows:

1. **Landscape Analysis:** The research conducted at this step is focused on the general application areas considered in the design of the XMANAI Manufacturing Apps Portfolio, including Production Optimization, Product Demand Forecasting, Process/Product Quality Optimization and Process optimization & Semi-Autonomous Planning.
2. **Demonstrator Use case analysis:** Each industrial partner, representing each of the 4 XMANAI demonstrators (namely FORD, WHIRLPOOL, CNH and UNIMETRIK), was assigned to at least one technical partner, responsible for providing the technical support in the design and implementation of each pilot. Teams then worked closely together, focusing on understanding the specifics of use cases addressed by the demonstrators, breaking down each use case in discrete sub-tasks. In parallel, each identified sub-task was related to the specific data sources that will be utilized to address it, as well as to the intended outputs.
3. **Mathematic description:** A mathematical formulation of the problems and their potential solutions is achieved, as a result of the analysis conducted in the previous step.
4. **Algorithm selection:** The selection of appropriate AI and Graph ML algorithms to address each task is made at this point. Amongst selected algorithms, those with an opaque functionality are equipped with explainability components, anticipated to provide meaningful and high-quality explanations. The methods presented here are depicted based on the insights provided from the review on explainability techniques documented in D1.1, as regards the construction of Hybrid models. In the case of Graph ML algorithms, the categorization of explainability methods is based on the taxonomy recently proposed by (Yuan, et al., 2021).
5. **Draft catalogue construction:** We further proceeded by grouping algorithms together and examining their applicability on more than one task, while the final catalogue is presented by means of algorithm cards that summarize information on the selected baseline algorithms.

1.5 Dependencies in XMANAI and Supporting Documents

The ”Draft Catalogue of XMANAI AI and Graph Machine Learning Models” presented in D4.1, was constructed in direct consideration of

- the requirements documented in D6.1-“Demonstrators Requirements”
- the results of D1.1-“State of the Art Review in XMANAI Research Domains”, extending the findings presented therein as concerns AI applications in industry towards a more technical perspective.

In addition, the draft catalogue is indirectly dependent to the contents and definitions presented in

- D1.2-“XMANAI Concept Detailing, Initial Requirements, Usage Scenarios and Draft MVP”, regarding the expected characteristics and usage of the draft and final product,
- D5.1-“System Architecture, Bundles Placement Plan and APIs Design”, in relation to the components of the XMANAI Explainable AI platform that will be used to develop and deliver the catalogue,
- D3.1-“AI Bundles Methods and System Designs”, regarding the XMANAI sub-components responsible to manage the model’s lifecycle,
- D2.1-“Asset Management Bundles Methods and System Designs”, in reference to methods and tools that will be applied to manage access to the catalogue and security issues.

1.6 Document Structure

A landscape analysis of AI methods applied to extract insights from industrial data, is outlined in the following Section 2 of this document. Common AI usage scenarios in industrial settings are described, focusing on the technical perspective and outlining emerging challenges and limitations. Section 3 provides detailed descriptions of the problems addressed by the XMANAI demonstrators, along with the data sources that will be available to solve them. Each use case is explored in depth, highlighting the needs for explainable solutions, also identifying technical challenges that are anticipated to arise. Security and ethics issues that may need to be addressed are also referred to here. Section 4 is dedicated to the selection of Hybrid and Graph ML algorithms to populate the draft catalogue of XMANAI models. The adopted methodology to collect the algorithms and summarize their use in the demonstrator manufacturing scenarios, leading to the construction of the draft catalogue, is detailed herein. Concluding arguments on the results and prospects regarding future work are presented in the final Section 5 of the deliverable.

1.7 Ethics

Ethical considerations were not encountered during the construction of the Draft catalogue of XMANAI baseline algorithms, since we are dealing here with mere mathematical objects that describe iterative algorithmic procedures, untrained models which are objective by default. AI models can only expose biases and stereotypes that are hidden in the training data, hence ethical concerns of this kind will be possible to be examined during the experimental evaluation of the selected algorithms trained on the data provided by the demonstrators. Moreover, the interference of XMANAI models with work ethics or the individuality of workers will be judged in relation to feedback provided by the business users, after the initial release and experimental use of XMANAI trained models and manufacturing applications. Consequently, ethical considerations and the fairness of XMANAI AI models will be addressed during the experimental phase of the project.

2 AI for Industrial Data Insights Generation

2.1 Introduction

This section focuses on the investigation of different manufacturing scenarios of the demonstrators, where a general analysis of the manufacturing landscape has been carried out in an attempt to describe the various possible scenarios of application. As a result, methods for 4 manufacturing scenarios have been analysed in depth. As it can be seen through the different subsections, the methods proposed so far in the state-of-the art consist of applying AI approaches leaving aside the explainability component. These manufacturing scenarios are:

- Production Optimization
- Product Demand Forecasting
- Process / Product Quality Optimization
- Process optimization & Semi-Autonomous Planning

2.2 Production Optimization

The enhancement of production planning and control, through Artificial Intelligence (AI) approaches, can lead to significant improvements in various manufacturing systems. The availability of data combined with the increased computing power have facilitated the advent of Machine Learning (ML) approaches as a powerful solution to cope with manufacturing challenges. Various implementations of ML solutions have been recently applied for production planning and control of manufacturing systems, producing robust and accurate results in different stages of the production lifecycle (Usuga Cavadid, et al., 2020).

(Gahm, et al., 2022) proposed a Neural Network (NN) approach to solve a production problem in the metal-processing industry. Specifically, the production of small parts needed to be maximized under the minimization of the utilized materials. The proposed NN was able to predict new feature vectors to describe the instances of the produced parts, and therefore they achieved the lowest expected loss (RMSE of 341.7), thus improving the production planning.

(Wang, et al., 2018) proposed a density peak-based radial basis function network (DP-RBFN) to rapidly predict the cycle time (CT) for production planning in the wafer manufacturing sector using a diverse and agglomerative CT dataset. The network method, which was based on a clustering technique, was capable of determining the density peak, while a parallel computing approach was proposed to speed up the training process with the large-scaled CT data. Finally, an experiment with respect to the semiconductor wafer fabrication system (SWFS) was presented, which demonstrated that the proposed methodology outperformed the radial basis function network, the back-propagation – network, and the multivariate regression method in terms of mean absolute deviation (MAD) and standard deviation (SD), as it accomplished a MAD and SD of 2.5 x 10-4 and 2.21 x 10-4, respectively.

(Lauer & Legner, 2019) presented a random forest (RF) ML approach for the prediction of plan instability in manufacturing master production planning. Their model, that was trained and tuned in an extended training dataset, provided an accuracy of 73%. From the results, it can be observed that the developed ML algorithm has been influenced by the relationship on-demand, hence extracting high accuracy in the second third of the planning horizon. To capture deviations from this pattern, further data and features are needed (e.g. capacity). The approach can be applied for decision-making, while the evaluation and results require further analysis.

(Gonzalez Rodriguez, et al., 2020) proposed a novel ML methodology for a Closed-Loop Supply Chain (CLSC) management problem. Their approach is a decision-making system based on fuzzy logic that is built on ML. Specifically, the real case scenario of the paper was an Industrial Hospital Laundry (IHL), with 16,000 kg of daily production integrated into a CLSC, aiming to reuse dirty hospital clothing. Therefore, there was a need for the development of a decision support system for the management of operations and the control of stocks at a tactical level. The approach was a combination of regression trees and fuzzy logic, defining a Fuzzy inference system (FIS). The performance and generalization of the regression trees were validated through a cross-validation mechanism giving an MSE of 43.12. After the learning of regression trees was finalized, a fuzzification step was applied to each one of the models. The FIS system gave an error of 5.48 ± 10 for externalization predictions and an error of 6.92 ± 9.04 after the fuzzification step for the new classification rate determination.

(Morariu, et al., 2020) presented an ML approach for predictive production planning (operation scheduling, resource allocation) and predictive maintenance. They developed a hybrid control solution that utilizes Big Data techniques and ML algorithms in order to process information streams in large-scale manufacturing. Specifically, a long short-term memory neural network (LSTM) was trained to determine possible anomalies or variations that are relative to the normal patterns of energy consumption. From the results obtained, the LSTM approach was able of predicting resource performances such as timeliness, energy consumption and precision. Large manufacturing systems with interchangeable resources could benefit from big data architectures, by extracting insights that can be adapted for triggering real-time decisions in production planning.

One of the key parameters in production planning and control is the lead time that elapses between the release of an order and its completion, as it is determined based on the observed time orders to traverse the production system. While the traditional order release models assume static lead times, they should be set dynamically to reflect the dynamics of the system. (Schneckenreither, et al., 2020) proposed a flow time estimation process to set lead times dynamically by applying an artificial neural network (ANN). Moreover, they implemented a safety lead-time to incorporate the underlying cost ratio, between the finished inventory holding and the backorder costs in the order release model. They tested the proposed methodology by utilizing a three-stage make-to-order flow-shop simulation model and compared the forecast accuracy along with the cost performance to other forecast-based order release models. The proposed ANN approach yielded high-quality flow time estimates which can help decision-makers to adapt lead times dynamically.

The data value has grown rapidly during the last years, and therefore gained a central role in modern societies. The development of sensors and new technologies to store and analyse the data are the new industrial revolution. These technologies are able, nowadays, to support smart manufacturing through the whole product lifecycle, provide new insights in each step, and also a better understanding of each step, from primary production to consumption. (Garre Perez, et al., 2020) presented an ML approach to support production planning of a food industry in the context of waste generation under uncertainty. Because food production is a complex process, uncertainty is very relevant and results in differences/anomalies between the planned production and the actual output. These anomalies lead to an economic cost for the company (e.g. waste disposal), as also in environmental impact (e.g. environmental carbon footprint). The proposed methodology consists of ML algorithms that were utilized to predict deviations in production, in order to reduce the uncertainties that were related to the amount of the waste produced in a food company that produces liquid products based on fruits and vegetables. The data have been gathered on 1,795 batches, including the characteristics of the product and the difference between input and output weights. The ML models that have been implemented, in the proposed approach, were: Linear model, Regression tree, Bagged tree, Random Forest, Gradient boosting, Lasso, Ridge regression, Elastic net and Spline. From the results it can be observed that the Gradient boosting algorithm outperforms the other models, by achieving an RMSE of 0.016, and a MAE of 0.012 for the test set, whereas it was finally verified that the proposed approach can be utilized as a tool for production anomaly detection.

Although the application of ML approaches, in production planning and control, provides great results to collaborate with such predictions, users need to demonstrate a level of confidence in them. Explainable AI (XAI) is defined as a new research area in understanding, trusting and managing the AI part. (Rehse, et al., 2019) implemented a Deep Learning (DL) methodology in a DFKI-Smart-Lego-Factory to predict processes, while they utilized State-of-the-Art XAI techniques to explain the outcomes in managers and visitors. Specifically, their approach of DL for process prediction, which is an important part of managing the business process at runtime, was based on an LSTM Recurrent Neural Network (RNN) with two hidden layers, 100 epochs, batch size equal to 20, dropout probability equal to 0.2 to avoid overfitting, a learning rate of 1.00, and LSTM forget bias equal to 0.1. From the results, the trained network was able to predict the next events, as well as the associated resources when an incomplete process instance was confronted. Moreover, through this approach, the production can be planned and optimized in real-time, while more information on the process instance leads to higher accuracy. The authors utilized XAI techniques in order to assist factory managers and production workers in better understanding, justifying and adapting the predicted outcomes by the RNN. They used textual explanations to visualize local explanations for the process outcome predictions. The user interface system introduced the area under ROC curve in the validation set, while the decision-makers could choose the probability threshold and different evaluation metrics (e.g. accuracy, precision, recall, F-score). Also, variable importance was automatically calculated and presented along with the feature importance on the global level considering the non-linear relationships presence, a process that helped decision-makers to understand the feature contribution to the extracted prediction. They utilized local explanations for particular cases and, after identifying the instances, white-box techniques such as rule induction algorithms or logistic regression were applied. Finally, they calculated saliency maps for the applied DL approach and visualization approaches (t-SNE technique) in order to make models more interpretable.

In Table 1 the aforementioned algorithms, their characteristics, and applications are presented.

Table 1 ML approaches in production planning

Task / Process	Data	Model	Output	Publication
Small parts in metal processing industry	Nesting instances data	NN	Predict new feature vectors to describe the instances of the produced parts	(Gahm, et al., 2022)
Wafer manufacturing sector	CT dataset	DP-RBFN	Predict the cycle time and determine the density peak	(Wang, et al., 2018)
Semi-conductor manufacturer	Historical data of Master Production Scheduling (MPS) process	RF	Plan instability prediction	(Lauer & Legner, 2019)
Management operations decision support in supply chain	Remaining production, differences between input and output, available time, classification rate	Regression Trees/Fuzzy logic	Classification rate Externalization	(Gonzalez Rodriguez, et al., 2020)
Operation scheduling – resource allocation in manufacturing	Energy consumption and execution time	LSTM	Predict timeliness and energy consumption	(Morariu, et al., 2020)
Flow-time estimation process	Flow time data	ANN	Lead times dynamic prediction	(Schneckenreither, et al., 2020)
Food industry	Product characteristics data	Gradient boosting	Predict deviations in production	(Garre Perez, et al., 2020)
LEGO	Production and sensors data	LSTM / t-SNE	Predict processes	(Rehse, et al., 2019)

Task / Process	Data	Model	Output	Publication
			XAI techniques for outcomes' explainability in managers and visitors	

It becomes obvious that Industry 4.0 leads to the digitization of various processes and provides opportunities and challenges. AI and ML approaches can lead to time and cost reduction in the industry, while extracting accurate and robust outcomes. The production planning and control of processes in the industry activities are of high importance because they establish smooth production processes, and as a consequence, they lead to the maximization of the outcomes. Although the ML approaches can provide benefits in the industry, they still lack understandability, while they are typically treated as black-boxes. XAI, as a new research field, can provide important insights that enhance the interpretability and explainability of the ML applications, and redefine the black-boxes approaches as grey-boxes. XAI techniques make the collaboration between humans and machine intelligence more feasible, advance the human decision-making process and define an upper level of trust that is required for autonomous AI deployment. In production planning and control, XAI is able to help managers into better decision-making through ML approaches.

2.3 Product demand forecasting

Product demand forecasting is a critical part for the supply-chain efficiency, but the limited data and the characteristics of the supply chain, alongside with the unavailability of historical data prevent forecasters pursuing advanced modelling in this field. Even when sufficient demand historical data exists, it cannot be considered as valuable because the market situation may change rapidly. The demands are affected by various hidden factors that require huge amounts of data and sophisticated models.

(Zhu, et al., 2021) presented a novel framework of supply-chain information and ML for demand forecasting in the pharmaceutical industry. Their approach was based on the following processing steps. First, they utilized cross-series training to resolve the "lack of data" issue and balance the trade-off between the sample size and sample quality. Moreover, they included two key non-demand features in the demand forecasting framework: (i) downstream inventory levels and (ii) supply chain information, designing the way to effectively include these features. They identified that the recurrent neural network (RNN) extracts better results with the cross-series learning framework and provides potential explanation of its superior performance, by the use of domain knowledge and numerical analysis. Also, they utilized two unique datasets, where they validated the performance of the proposed framework and extracted important empirical evidence. Finally, the cross-series forecasting model framework, with the grouping schemes and the non-demand features was able to demonstrate the value of ML in demand forecasting.

The product demand forecasting is also significant in the fashion industry field due to the complexity of operations, the huge product variety and the evolving retail trends. (Kharfan, et al., 2021) presented an ML technique for demand forecasting of newly launched seasonal products in a leading fashion retail company. The proposed methodology was applied to a data set obtained from a leader company of apparel and footwear in USA, in which two data types were collected from the company as sell-in (shipment) and sell-through (POS) data. Several features were included such as product, calendar, store, price and promotion, and sales units. Their proposed approach was a three-phase model, consisting of clustering, classification, and prediction. The main objective was to identify the look-alike group of products from the train set. After the identification of the product, their average sales were used as a proxy to forecast the sales with the validation and test sets. The clustering phase grouped all the styles in clusters based on similarities of features. Multiple tools were utilized in the clustering task such as feature selection, data normalization, high dimensionality reduction (t-SNE algorithm), and k-means clustering. The classification phase created links between the styles. In this phase, classification trees, random forests, and Support Vector Machines (SVM) were utilized. Finally, in the prediction phase, the future sales for the brand-new styles were forecasted in the validation and test sets. The performance of the approach was estimated through Weighted Mean Absolute Percentage Error (WMAPE), and the forecast bias by using Weighted Mean Percentage Error (WMPE), while SVM achieved the best performance with an accuracy of 93%.

In industrial enterprises active in wholesale and retail trade and especially in the modern competitive sector of air transportation, demand forecasting for new products is critical. (Smirnov & Sudakov, 2021) proposed an ML approach for forecasting new products' demand, by utilizing data from the Ozon online store. The input data consisted of features such as price, name, category, and text description of the product. To solve this regression problem, they implemented various ML algorithms such as XGBoost, LightGBM, and CatBoost. The best algorithm was the LightGBM that accomplished an RMSE of 4.00. Their proposed model can be useful for companies' analysts, to optimize sales assortment planning and logistic optimization, as well as to increase the accuracy of other prediction systems.

(Goncalves, et al., 2021) implemented a multi-variate approach for multistep demand forecasting in assembly industries, with empirical evidence from an automotive supply chain. Their approach was based on a classical SC topology that consisted of a single manufacturer that was linked with different suppliers and end-customers. The forecasted demands for the finished products were utilized to determine the component order sizes to suppliers, the supplied components were assembled by the manufacturer and after producing a set of finished products, end-customers were able to request fulfilment follows. The proposed forecasting framework is a two-stage framework for the manufacturer's demand forecasting. First, a multivariate dataset of time series, for each component, was built by utilizing the indicators that were mentioned above. A sliding window was used to create a set of training instances that were defined by the pre-specified number of time-lags, related with each input feature. An unconditional forecasting step was formulated which used only the information that is available until the forecast origin. Only lags greater or equal than the forecast horizon were included as regressors. The authors applied a multivariate expansion of Auto-Regressive Integrated Moving Average (ARIMA) model, known as ARIMAX, which allowed the inclusion of exogenous inputs apart from the autoregressive and moving average parameters. They also utilized three popular supervised ML regression models: multilayer perceptron (MLP), support vector machine (SVR), and random forest (RF), able to cope with complex nonlinear mappings. The evaluation of the forecasting framework was implemented in the logistics department of Bosch Automotive Electronics Portugal (AE/P), in order to establish a strategy for manufacturers demand forecasting improvement. The MLP models outperformed the other benchmark models in terms of forecasting performance, by achieving an NMAE of 9.48% at the 95% significance level.

Moreover, (Feizabadi, 2020) proposed a hybrid demand forecasting that combined ARIMAX and neural networks (NN) for predicting product demand for a steel manufacturer. The steel manufacturing company was operating in four different segments (retail, project, manufacturing and home appliances). For each one of the segments data was collected. The proposed methodology based on two established models; ARIMAX and two-layer feed-forward NNs with backpropagation learning were fed with both time series and explanatory factors. A correlation analysis was carried out for the task of identifying the impact of several factors to the company's sales forecast. From the reported results, the proposed ML-based forecasting approach (ARIMAX and NN) was able to improve both operational and financial metrics. The developed model was able to capture the complex and non-linear relationship among many variables, by extracting a forecasting accuracy up to 99.2% depending on the month.

In supply-chain it is crucial to control costs, in order to increase the customer's satisfaction, manage inventory and therefore improve the product. (Nguyen, et al., 2021) proposed a demand forecasting and anomaly detection framework, for fashion retailing supply-chain, which consists of AI algorithms such as LSTM neural network, autoencoder network and one-class support vector machine (OCSVM). In fashion retailing, the consumer demand is fluctuating and sensitive to fashion trends, weather and price, leading to sharp and immediate fluctuations that cannot be predicted by the POS data-based replenishment system thus generating significant profit loss. Their proposed approach aimed not only to predict the exact sales by stock-keeping unit and store, but also detect and anticipate exceptional sales in order to help the practitioners' decision making. The performance of their proposed method has been verified on benchmarking and real fashion retail datasets. The obtained results have shown that the proposed methodology was able to perform with high accuracy in both kinds of data. Specifically, for the benchmarking dataset (C-MAPSS) the proposed framework achieved an RMSE of 9.71 for the test data, while on the generated data that was used for anomaly detection achieved an accuracy of 98.36%, a precision of 98.45%, an F-score of 96.98% and a Recall of 99.55%. For the real fashion retail data, the proposed methodology achieved an accuracy of 98.45%, a precision of 98.45%, an F-score of 96.98% and a recall of 99.59%, while for the anomaly detection the LSTM-Autoencoder-OCSVM methodology was able to extract attributes from the input, and classify accurately the anomalies, extracting insights that could help the company find out the main factors that led to higher/lower product demand.

ML in the product demand field has been raised as a powerful tool for the industry, in terms of profit maximization and cost minimization. Although the AI approaches extract robust and accurate results, there is still a need for explainability. (Rozanec & Mladenic, 2021) presented a novel XAI architecture based on semantic technologies through a knowledge graph which could provide concepts that convey feature information at higher abstraction level regarding demand forecasting models in the automotive sector. Through their approach, the authors were able to link domain knowledge, forecasted values and forecast explanations in a knowledge graph. The proposed architecture novelty was based in the combination of semantic technologies and media events providing informed prediction explanations, as it could integrate predictions, gather insights/relevant features, incorporate domain knowledge and context to each prediction and provide a forecasting explanation to the end-user. The proposed methodology could support semantically enhanced explanations for demand forecasting models, as it provided demand forecast values, the associated uncertainty, high level description of features influencing the prediction, the related media events, context and means to improve the demand forecasting model. The high-level description of features provided information about the main factors that influence the forecast along with the ways to avoid exposing sensitive details and model bias.

In Table 2 the aforementioned approaches are presented in terms of algorithms, their characteristics and the associated applications.

Table 2 Production demand forecasting approaches.

Industry	Data	Model	Output	Publication
Pharmaceutical	Demand historical data	RNN	Demand forecasting	(Zhu, et al., 2021)
Fashion retail	Sell-in (shipment) and sell-through (POS) data	SVM	Demand forecasting	(Kharfan, et al., 2021)
Online store	Ozon online store data (price, name, category, and text description of the product)	LightGBM	New products demand forecasting	(Smirnov & Sudakov, 2021)
Assembly (Automotive)	Product data	ARIMAX/MLP	Multistep demand forecasting	(Goncalves, et al., 2021)

Industry	Data	Model	Output	Publication
Steel manufacturer	Multivariate data from Retail, project manufacturing and home appliances	ARIMAX/NN	Demand forecasting	(Feizabadi, 2020)
Fashion retail Supply-chain	C-MAPSS and generated dataset	LSTM-Autoencoder-OCSVM	Demand forecasting and anomaly detection	(Nguyen, et al., 2021)
Automotive	Automotive Materials, Media events, and demand data	Graph model	Information about main factors that influence the forecast, ways to avoid exposing sensitive details, and model bias.	(Rozanec & Mladenic, 2021)

Product demand forecasting is considered as a crucial part in the life cycle of almost all industries. Although the AI/ML approaches have recently levelled up the life cycle of the industries, XAI can further bring business value by establishing trust between a company and its clients. This is achieved by making the prediction outcomes of the product demands more understandable for non-experts thus upgrading the trust-levels in the production process.

2.4 Process/product quality optimization

Process quality optimization

Within the generic application domain of industrial process quality optimization, we focus on the field of **Maintenance scheduling** and explore AI methods applied for this purpose. Maintenance activities are a significant factor affecting a company's costs and availability of products. The digital transformation towards Industry 4.0 is highly connected to advances in physical systems Prognosis and Health Monitoring (PHM) via the integration of Condition Monitoring systems (CMS), made available with the use of numerous sensors along the production line. This enables the development of predictive and proactive strategies for maintenance scheduling, by exploiting the huge amounts of data produced by the monitoring process to model the health state of machine components and tools and act accordingly. The application of predictive maintenance strategy (PdM) allows to overcome limitations of earlier approaches, in particular:

- Corrective (run-to-failure R2F): Maintenance actions take place when a failure occurs. This strategy results in extended use of tools and components, but product quality may be compromised as the performance of machine components/tools degrades. Unplanned machine downtime is increased which also affects the availability of products.
- Preventive (scheduled): Maintenance actions take place in predefined time intervals, machine components and tools are replaced when half-life is reached or, according to statistical analysis of failures, regardless of their condition. As a result, maintenance costs are increased, and the potential useful operation time of components and tools is not fully exploited.
- Predictive / proactive (PdM) or Condition-Based Maintenance (CBM): The quality of operations and assets is constantly monitored, the current state of the production line is diagnosed and used to predict its future state, so that abnormal states are detected in-time and the needed maintenance actions can be undertaken at a convenient moment. Diagnosis and prognosis therefore enable informed decision making to prevent future failures or mitigate their impact.
- Prescriptive maintenance goes one step beyond, providing decision support systems based on the diagnostic and prognostic results of PdM, in order to achieve the best possible scheduling

of maintenance actions with the least effect on the overall business activities. As a result, the management of resources and inventory of spare parts is optimized, operational uptime and product availability is maximized, while safety risks are kept to a minimum.

A holistic view of maintenance scheduling, based on the combined results from diagnostics and prognostics over the manufacturing system's components integrated with information from other sectors such as logistics and inventory, allows for the development of a decision support system that recommends a balanced solution, between the optimization of maintenance schedule, the minimization of costs and the maximization of product availability (Bousdekis, et al., 2019).

Given the large amount, diversity and velocity of data involved in the context of PdM, including telemetry and CMS data from various sensors such as thermometers, accelerometers, digital cameras, acoustic and vibration sensors (Pech, et al., 2021), machine component error/failure logs and tool replacement logs, AI solutions have been found to excel over physics-based (Physics-of-Failure PoF) and statistical approaches (Carvalho, et al., 2019). This is evidenced by the significant growth in publications proposing AI solutions to PdM in the last decade, including machine learning (ML), deep learning (DL), graphical and fuzzy logic approaches (Dalzochio, et al., 2020). More importantly, the benefits fostered by AI-based PdM solutions have been already justified in practice since 2018, as documented in PwC and Mainnovation Research report that involved a large number (268) of manufacturing companies from Belgium, Germany and the Netherlands (Haarman, et al., 2018).

An overview of AI techniques proposed to implement PdM in industrial settings is provided in Table 3. The list is not exhaustive, rather than indicative to the variety of AI methods and data sources that have been explored in this research and application area.

The application of AI solutions to PdM however is not a simple task, as the development and implementation of such solutions is highly dependent on the availability of data, which is related to the previously established maintenance strategy:

- Supervised methods require information on machine component failures and tool wearing in the modelling data, usually available when R2F was the previous strategy (historical data in machine maintenance cycles, i.e. between failures). If such knowledge is available, then labelled datasets can be created through data (temporal) interlinking. Regression methods can be applied to predict the remaining useful life (RUL) or the probability of failure, whereas classification methods enable to detect predefined healthy and non-healthy conditions for each tool/component in the production line.
- Unsupervised Methods come into play when machinery and process historical information is available, but failure records do not exist. This is usually the case when the previously applied strategy was preventive/scheduled maintenance. Unsupervised methods can also be applied at the data pre-processing level for dimensionality reduction and feature extraction. Methods in this category include variations of PCA and clustering (Amruthnath & Gupta, 2018), as well as embeddings. Anomaly detection algorithms are applied both at the pre-processing level, to eliminate outliers in the data and at the pattern recognition level, to detect abnormal system states.

Other data-related challenges arise from the poor quality of data (especially raw sensory data are very noisy and usually redundant), imbalance in the records between healthy (common) and unhealthy (sparse) conditions, heterogeneity of data formats, and the integration of Big Data management tools (Dalzochio, et al., 2020).

In the absence of historical data, synthetic datasets can be constructed via physical modelling of machine degradation, using digital twin technologies and Deep transfer learning (Xu, et al., 2019). In addition, there are several open-source datasets for PdM such as the Turbofan engines dataset by NASA and the FEMTO bearing dataset.

Random Forest, Artificial Neural Networks and Support Vector Machines have been depicted by earlier reviews as the most commonly used models in PdM tasks (e.g. (Carvalho, et al., 2019)). The need for time-based solutions however, instead of just alert monitoring, has recently driven researchers' interest towards Deep Learning sequence models such as Deep autoencoders, methods also able to handle unsupervised feature learning in time-series (e.g. (Ren, et al., 2018)). Deep Learning methods are currently widely explored for the development of PdM (Serradilla, et al., 2020), due to their ability to handle multidimensional data, identify latent features and uncover complex, non-linear underlying relations. Adding to that, the inclusion of Generative Adversarial Networks (GANs) in the model architecture allows for the generation of additional fault records, to surpass their lack in the data (Liu, et al., 2021). Updating and running DL models in real-time, however, requires high computational cost and may affect the system's efficiency and latency. Taking such considerations into account, (Fang, et al., 2021) developed a computationally light yet robust CNN-based solution to provide fast and accurate predictions.

Effective modelling must consider several failure mechanisms at play, in order to define the proper failure criteria/states for each component (soft: close to failure, hard: failure). This in turn allows for the recognition of degradation trends and the identification of parameters (external, such as machine workload, and internal, such as CMS sensory data) related to failure (Liu, et al., 2021). The generation of reliable quantitative indicators to achieve assets (machines, tools) and processes health monitoring ("health factors") is therefore of paramount importance for the development of accurate and robust diagnostic and predictive PdM models (Guo, et al., 2017). Domain knowledge facilitates this process, to achieve clever feature engineering and health factor identification based on physical knowledge (Zonta, et al., 2020). Further, domain expertise enables semantic knowledge representation so that ontology-based approaches can be developed (Cao, et al., 2020). Ontology-based approaches seem able to tackle the more realistic case of multi-state & multi-component optimization, instead of providing prognosis for a single machine, by taking interrelations and interdependencies among the operations of different components into account. An example is found in the work of (Ansari, et al., 2020), where the authors leverage on a Maintenance-specific data model to construct Bayesian Networks, as directed acyclic graphs (DAGs) representing causal relations in the data. Including the time dimension, these networks evolve into Dynamic Bayesian Networks able to infer the future probability distributions of nodes in the network, under the assumption that relations represented by edges in the graph remain static.

Apart from the data usage, architectural considerations of AI-based PdM systems are related to the adopted data storage system within the company (on-premises or cloud-based) and also the rate at which data is currently acquired, that is, whether data are collected at predefined intervals (batch data), or in a continuous manner (on-line or streaming data). An example driven by the work of (Nguyen & Medjaher, 2019) is provided in *Figure 1* below. Diagnosis and prognosis of the system's health is performed by LSTM network, and the results are further processed in relation to maintenance cost estimation and spare part availability by use of a heuristic algorithm, to support timely and informed decision making.

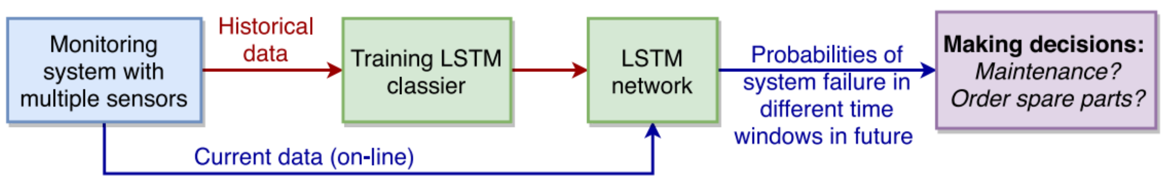


Figure 1. The dynamic predictive maintenance framework proposed by (Nguyen 2019). In this case, the AI model is trained on historical data, and designed to operate on streaming (on-line) data to support decision making regarding maintenance actions.

Additional challenges are related to the PdM system's scalability and latency. A possible solution is proposed by (Zhou & Tham, 2018), based on ensemble learning and edge computing. A base learner for health state classification is run at each component of the production line, then separate results are combined using a Graphic Ensemble, to assess the overall health state of the system.

Table 3 Overview of AI methods for Predictive maintenance in industrial settings. In case more than one models are examined, the best-performing method is depicted in bold.

Machine/Tool	Data	Model	Output	Publication
Rotating Machinery (gears)	Vibration Signals	Deep Belief Network and CNN	Feature extraction Fault Detection	(Li, et al., 2019) (Li, et al., 2019)
Gearbox	Vibration signals	Hidden Semi-Markov Model	Fault Detection, RUL	(Li, et al., 2018)
AGV, robot, milling machine, turning machine (structural parts processing)	Machinery external data (sensors) and internal data (SCADA)	**LSTM-GAN**, CNN-LSTM, WGAN, GAPCNN	State & Fault prediction	(Liu, et al., 2021)
Turbofan Engines	Multivariate sensor data (temperature, pressure, fan speed), operating conditions, fault modes	LSTM	State classification (based on probability of failure in a given time window)	(Nguyen & Medjaher, 2019)
Wind turbines generator bearings	Vibration signals	RNN	Health Index, RUL	(Guo, et al., 2017)
Motors	Vibration Signals	**CNN and Spatial Attention Mechanism**, VGG16, MobileNet, LeNet-5, MLP, SVM	Bearing Fault Detection	(Fang, et al., 2021)
Assembly Line Workstation	Temperature, humidity, acceleration, gyroscope data	DBSCAN RF	Outlier Detection Fault Detection (classification)	(Syafrudin, et al., 2018)
Turbofan Engines Aircraft Landing gear Aircraft Electrical System	Multivariate sensor data Pressure, proximity, operating conditions Voltages, currents	One-class SVM **Logistic Regression, Linear SVC, SVM, Extra Trees, RF,** AdaBoost, DT, Naïve Bayes, PCA, k-means **ARIMA**, Regression, Echo State Network, MLP, RNN-NAR, RNN-NARX	Anomaly Detection Fault classification Health levels identification RUL	(Adhikari, et al., 2018)
Industrial machines	Sensor data (voltage, rotation, pressure, vibration) Error logs, maintenance history	**RF, ANN**, DT, Naïve Bayes, k-NN	State/Fault classification	(Cardoso & Ferreira, 2021)
Exhaust fan	Vibration signal	PCA **T^2 statistic**, k-means, Fuzzy C-means, Hierarchical clustering, clustering	Feature extraction Fault Detection	(Amruthnath & Gupta, 2018)

Machine/Tool	Data	Model	Output	Publication
		based on Gaussian Mixture Model		
Turbofan engine Injection moulding	Multivariate sensor data	**Graphical Ensemble Learning** (Simple Average, Weighted Average**,** Negative correlation learning with weighted averag**e)** Base learners: **DT**, kNN, SVM, MLP	State/Fault classification Product quality classification	(Zhou & Tham, 2018)
FEMTO Bearing dataset	Vibration signals	**Autoencoder and DNN,** DNN, SVM	RUL	(Ren, et al., 2018)
Coilers parts (drums) for mill process in stainless steel production	Process parameters (steel density, coiler temperature, engine pressure etc) Logistics parameters (eg steel plate id) Configuration parameters	Discrete Bayes Filter	Probability distribution for part degradation	(Ruiz-Sarmiento, et al., 2020)

It appears that non-linear and complex ML/DL models perform better on PdM tasks, as a result of this landscape analysis, although direct comparison is not possible unless a comparative study is performed under the same setting (as for example in (Adhikari, et al., 2018), (Ren, et al., 2018), (Cardoso & Ferreira, 2021)). The development of AI-based Decision Support systems for PdM in the XMANAI context will equip such black-box models with interpretable components to provide trustful and comprehensive solutions that can be embraced by the business users at all levels, by exploiting the findings of D1.1 on state-of-the-art explainability techniques, while taking into account the explainability requirements put forth by the business users.

Product quality optimization

Industrial products quality control is an important factor that affects a company's reliability and competitiveness. This field of manufacturing operations has begun to profit from the application of AI solutions for quality assertion and defect detection, as well as for the optimization of product manufacturing process (process control). Since the usage of AI methods in the optimization of process parameters is explored in the previous section, here we focus on methods applied to quality inspection through product defect detection, in the manufacturing environment.

Product defect detection usually involves classification, localization and segmentation of geometric or surface/density defects. The most common method to achieve this is visual inspection of digital imaging data, although other signals such as acoustic, ultrasounds or Eddy currents are also explored to extract meaningful information on the quality of the product. Defect detection based on machine visual inspection outperforms human inspection, due to the ability of machine vision technologies to detect flaws invisible to the naked eye and observe across the entire electromagnetic spectrum, as opposed to the limitation of the human eye to the visible part. With the use of digital image processing and Artificial Intelligence in the manufacturing domain, machine vision inspection can offer accurate product defect detection in real-time, in an automated manner. For example, defect detection in welded joints is studied by (Launay, et al., 2021). Defects in this case are trapped gas pores observed in X-ray tomography, while their mechanical response is computed by numerical simulations based on physical laws. Classes of defects are identified according to morphological characteristics and mechanical response by k-medoids, then a MLP relates a defect's morphological representation with its mechanical response.

Computer vision has been used from manufacturing industries to assert product quality, but the selection of which features to inspect was done manually. Feature extraction also has to be done manually in case traditional ML algorithms are applied. In a dynamic manufacturing environment, manually designed features may have to be redesigned, for the inspection system to work well on different products. Deep Learning methods overcome this obstacle, being capable of learning important features from images, both in supervised and unsupervised settings, hence they prevail in recent literature focused on surface defect detection (Chen, et al., 2021).

Limitations of typical deep CNN for automated surface inspection include the difficulty of CNN to detect defects in non-flat and geometrically complex products, also the computational cost. (Wang, et al., 2020) propose a Faster R-CNN model running on embedded hardware in cloud-edge computing to overcome both issues. Faster R-CNN is built with the inclusion of a region proposal network that increases the speed in the detection of regions of interest, while exhibiting a balanced performance between the detection and classification of defects when compared to Single Shot MultiBox Detector (SSD) with Inception v3 or SSD with Mobilenet. In a different approach that aims to increase CNN classification accuracy in small-defect areas and produce a reusable model to detect defects in different automotive parts, (Qu, et al., 2018) integrated traditional image feature processing concepts for density slicing, region segmentation and area filtration into a deep CNN architecture. The proposed method exhibits high adaptability without the need for retraining.

Deep Learning models need a large amount of training data, while the validation and testing data should be drawn from the same/similar distributions. In real production environments however, this is not always possible due to the emergence of new products with different characteristics, also the fact that company records on old products may not be kept after a while or may not be accessible. To avoid retraining from scratch any time new products are manufactured and to overcome the limited availability of new data, continuous or life-long learning DNNs are built by transferring the knowledge from previously trained models on similar tasks by cloning their weights, while fine-tuning on the particular task is achieved with only a small amount of new training data (Tercan, et al., 2022). The difference from transfer learning is that continual learning methods avoid the catastrophic forgetting of the trained network when faced with a new task. (Benbarrad, et al., 2021) present a comparative study on transfer learning from trained Deep CNNs for image classification, namely VGG16, Inception v3 and EfficientNetB0. The models are fine-tuned on the cloud, because despite pre-training this is a computationally intensive task for such deep network architectures, while image classification of the current product is performed on the edge. Taking into account both classification accuracy as well as time constraints, Inception v3 is depicted as the method of choice. The authors also investigate the effect of process parameters on the quality of the final product, by means of regression techniques, where Decision Trees are found to outperform k-NN, Lasso and Linear Regression. Another solution that is aimed to perform well on variant product defect inspection without expensive re-training is proposed by (Luan, et al., 2020), combining the traditional Similarity Structure Index (SSIM) with Siamese Networks. The model learns from pairwise structural similarity differences and is found to perform well on cross-category defect detection, on unseen data.

Regarding DL methods for unsupervised defect detection that have been currently explored, Auto-Encoders (AE) and Generative Adversarial Networks (GANs) are those that appear more frequently. For example, (Mei, et al., 2018) propose a multi-scale Convolutional Denoising AE to detect and localise unseen defects at different scales while being trained only on defect-free images. The method is found to be both accurate and robust to noise by experimental results on diverse data coming from fabric textiles, LCD panels and other materials. (Liu, et al., 2019) leverage on the ability of GANs in extracting latent informative features and learning the distributions in the training data, to generate more data based on the estimated distributions. Their proposed GAN is trained on faultless steel plate images and the extracted features are used to train a One-Class SVM, which is then able to accurately classify faulty plates, based on the same feature vector, in contrast to the quality ones.

In Table 4 the aforementioned approaches are presented in terms of algorithms, their characteristics and the associated applications.

Table 4 AI methods for product quality assurance and defect detection. In case more than one models are examined, the best-performing method is depicted in bold.

Product	Data	Model	Output	Publication
Welded joints	X-ray tomography	k-mediods MLP	Defect type identification Defect classification	(Launay, et al., 2021)
Turbo blades (automotive turbo engine components)	Digital Image	**Faster R-CNN,** SSD +inception v3, SSD + Mobilenet	Defect detection and classification	(Wang, et al., 2020)
Automotive engine parts	Digital image	**Traditional image processing with Deep CNN**, SDK software, VGG (pixelwise and patchwise)	Defect detection and classification	(Qu, et al., 2018)
Plastic bricks (injection moulding)	3D CAD objects	Continual learning DNN	Defect detection	(Tercan, et al., 2022)
Casting products Iron (mining flotation process)	Digital Image Flotation process data	Transfer learning **Inception v3**, VGG16, EfficientnetB0 **Decision Tree**, k-NN, linear & lasso regression	Defect detection Predict impurities in ore concentrate	(Benbarrad, et al., 2021)
Manufactured products of different categories	Digital Image	SSIM + Siamese Networks	Defect detection and classification	(Luan, et al., 2020)
Textile surfaces (Fabric, LCD panel, various materials)	Digital Image	Multi-scale Convolutional Denoising AutoEncoder	Defect detection and localization	(Mei, et al., 2018)
Steel plates	Digital Image	GAN One-class SVM	Feature extraction Quality classification (g/b)	(Liu, et al., 2019)

2.5 Process optimization & Semi-Autonomous Planning

Another critical part in the life cycle of production is the task of process parameters' optimization. The optimization of process parameters provides advantages in the hands of companies and facilities, by extracting the best set of parameters for each individual process, extracting insights and outcomes that are vital in terms of cost/profit level and the quality of each production line. AI and ML-driven optimization approaches are more and more implemented into various industry fields, offering accurate and less time-consuming results, compared to manually optimization procedures. Process parameters can be optimized by hybrid tools for objective function enhancement.

(Deshwal, et al., 2020) deployed hybrid optimization techniques for the process parameters' optimization of tensile strength in additive manufacturing. Their proposed approach combined three optimization techniques (GA-ANN, GA-RSM, GA-ANFIS) and the best method, GA-ANN, accomplished an accuracy of 99.8%. The proposed technique facilitated the producing units to choose the optimized factor value in the input factors for Fused Deposition Modelling (FDM) parts fabrication with improved mechanical properties. The proposed hybrid models could be utilized for accurate prediction and optimization of other process parameters in similar industrial application problems.

(Rouniyar & Shandilya, 2020) presented an optimization approach for process parameters in the magnetic field assisted powder mixed electrical discharge machining (MFAPM-EDM), which is a variant EDM process, with ultimate objective to improve the surface quality, the machining rate and the stability of the process. They used Aluminium 6061 alloy due to its growing use in aviation, automotive, and naval industries. The process parameters in their study were the discharge current, spark duration, pause duration, concentration of powder and magnetic field, which were considered to analyse the effect on material erosion rate and electrode wear rate. Teacher-learning-based optimization (TLBO) was implemented for the optimal process parameters determination and maximum material erosion rate (MER) and electrode wear rate (EWR) achievement. Through the TLBO implementation and comparing the proposed optimization method with other optimization methods such as genetic algorithm (GA) and desirability function of RSM, it was observed that the TLBO approach provided minimum EWR (0.1021 mm3/min) and maximum MER (30.4687 mm3/min).

Apart from the task of process quality improvement, there are various procedures that can be also optimized in machining processes. End-users are also interested in the minimization of energy consumption during a process (e.g. cutting process). (Zhang, et al., 2017) proposed an optimization approach through a multi-objective optimization and Quantum genetic algorithm in order to reduce the energy consumption of a numerical control (NC) machine tool. The optimization model was applied to minimize the cutting specific energy consumption and the processing time, through the process parameters under actual constraint conditions in the manufacturing process. Through their methodology, they achieved a decreased processing energy consumption of the process parameters by 27.21%, while the CSEC was reduced to 32.07% and the processing time was reduced by 34.11%.

(Pfrommer, et al., 2018) presented a deep neural network (DNN) as a surrogate model for the optimization of process parameters in a composite textile draping process. Numerical experiments were linked with a Finite Element (FE) simulation model, and the surrogate-based optimization DL model was trained to predict the shear angle for more than 24,000 textile elements. The proposed methodology reduced the number of the expensive to evaluate, resource-intensive FE simulations required to find the optimized parameter configurations, while it managed to improve the best-known overall solution.

(Madhavi, et al., 2017) proposed a Taguchi-Principal Component Analysis (Taguchi-PCA) approach for the evaluation of the optimum turning process in a machining industry to achieve good surface finish and high hardness. The main objective of their study was to solve the multi-response parameter optimization problem of the turning process, though solving the correlated multiple criteria optimization problem and considering as performance characteristics the hardness and surface roughness. The input parameters were the cutting-speed, the feed, and the depth of the cut. Although the traditional Taguchi-based hybrid optimization methods rely on the assumption that quality indices are uncorrelated or independent, to overcome the Taguchi limitations the authors implemented a PCA application for converting the correlated responses into uncorrelated quality indices, which are called individual principal components. The PCA component was optimized through the Taguchi method, while an Analysis of variance (ANOVA) was applied in the PCA component to find the parameters. Tests were conducted for three different materials and the results showed that the proposed Taguchi-PCA technique could be efficient for solving multi-attribute decision-making problems, such as multi-objective product process optimization for continuous quality improvement.

(Hooda, et al., 2021) presented a random forest (RF) approach for predicting the optimum deposition angle, for any geometry, in Fused Deposition Modelling (FDM). FDM's optimum value of deposition angle varies with the product geometry, and therefore accurate predictions are of high importance. The training dataset in the proposed method was generated by utilizing different shapes and geometries, and the generated correlation-based feature selection method was applied to extract crucial product features. K-fold cross-validation was used for the effectiveness of the RF model, while the empirical evaluation achieved a prediction accuracy of 94.57%. The proposed methodology provided very accurate and robust results and could further enhance the applicability of digital manufactured products.

The ability of monitoring manufacturing processes is an important task in today's production processes, as it increases user's adaption and diffusion over technologies while users and human experts need to be provided with explanations and insights form the modules. Moreover, the unavailability of labelled historical data makes the use of ML models unfeasible. (Brito, et al., 2022) proposed an explainable artificial intelligence (XAI) approach for fault detection and diagnosis in rotating machinery. The proposed methodology consisted of three parts which were the following: feature selection, fault detection and fault diagnosis. Specifically, vibration features of the rotating machine were extracted in both time and frequency domains, unsupervised fault detection was verified through anomaly detection algorithms and fault diagnosis was achieved through Shapley Addictive explanations (SHAP) technique that interprets black-box models. In the feature extraction step, the vibrations features were extracted with respect to the type of monitored component and were then divided into training and testing groups, while the hyperparameters of the anomaly detection models were tuned. In the fault detection part, different anomaly detection algorithms were implemented, namely: clustering based local outlier factor (CBLOF), local outlier factor (LOF), isolation forest (IF), lightweight detector of anomalies (LDOA), histogram-based outlier detection (HBOD), k-nearest neighbours (kNN), fast-angle-based outlier detector (FastABOD), outlier detection with minimum covariance determinant (MCD), one-class support vector machine (OCSVM), featuring bagging (FB) and a combination of all models. In the proposed fault diagnosis step, the most relevant features were analysed through models' explainability, and the feature importance ranking was obtained through SHAP. The results were presented using the F1-score, PR-AUC (Precision-Recall Area Under the Curve) and the average confusion matrix of the iterations. In the anomaly detection task, the following performances were achieved: F1-score of 99.45% in Case 1; models: MCD, HBOS, and IF, 99.84% in Case 2; models: MCD, kNN and IF, and 99.22% for Case 3; models: HBOS, IF, and MCD. For the classification part in Case 1 a maximum accuracy of 99.57% was achieved though IF, and in Case 2 an accuracy of 96.72% was obtained by kNN and CBLOF. It was concluded that the proposed methodology could be effective and could be also utilized in many different industrial applications.

In Table 5 the aforementioned approaches for the optimization of the process parameters in terms of models, characteristics, and applications are presented.

Table 5 Optimization of process parameters approaches.

Industry	Data	Model	Output	Publication
Manufacturing	Materials	GA-ANN	Optimization of tensile strength process	(Deshwal, et al., 2020)

Industry	Data	Model	Output	Publication
			parameters in additive manufacturing	
Manufacturing	Aluminium 6061 alloy	TLBO	Process parameters optimization to improve the surface quality, the machining rate and the stability of the process	(Rouniyar & Shandilya, 2020)
Machining	Experimental machining data	Quantum GA	Minimization of the energy consumption of a numerical control (NC) machine tool	(Zhang, et al., 2017)
Manufacturing	Textile elements and draping process parameters	DNN	Optimization of process parameters in a composite textile draping process	(Pfrommer, et al., 2018)
Manufacturing	Turning process parameters	Taguchi-PCA	Evaluation of optimum turning process parameters to achieve good surface finish and high hardness	(Madhavi, et al., 2017)
Manufacturing	Cube, sphere, pyramid, rectangle, cone shaped, and wheel specimen data	RF	Optimum deposition angle prediction	(Hooda, et al., 2021)
Manufacturing	Rotating machinery data	Anomaly detection algorithms / SHAP	Fault detection and diagnosis	(Brito, et al., 2022)

Explainable AI (XAI) becomes an important part in AI research area since it leads to trustworthy, compliant, effective, and more robust systems which can increase the adoption level and the business value. Understanding which parameters each model uses in order to achieve better performance, is essential in the manufacturing processes. Through XAI it is feasible to perceive the effect of each parameter in a process, and this can lead to more persuaded optimization actions and approaches. XAI converts the confusion of AI systems into a more accessible environment, for non-experts, by enhancing stakeholders' trust through interpretability.

3 AI for the needs of XMANAI Demonstrators

3.1 Introduction

As a preliminary task to the elaboration of the Baseline Models Catalogue, the identification of the different manufacturing scenarios to be addressed in the execution of the demonstrators is mandatory in order to select the algorithms and the types of explainability that better suit their needs. Thus, this section aims to provide an overview of the different demonstrators as well as an identification of the problems to be solved and the available data sources. In addition, the explainability needs in each scenario have been taken into account in order to select not only the main model to address the particular problem, but also the explainability tool that best fits these needs. During this process, the different technical challenges, security issues and ethics that may arise in each demonstrator have been considered.

3.2 FORD - AI for production optimization

The current situation at Ford Engine Plant does not allow the power of quasi-*real-time* data to be harnessed for decision making. There are records on the status of the different operations on the production line, the quantity of engines produced and their parts, quality reports and production plans. Despite having this information, there is not a centralized database and all the information is disaggregated in different corporate databases. This lack of centralized information is the first problem that needs to be solved in order to optimize the different processes that occur on the production line. This problem implies another one, which is the lack of artificial intelligence applied to the different decision-making processes due to the impossibility of taking advantage of all the available data. The proposed application aims to mitigate these problems by means of a set of functionalities that will be explained in the following sections.

3.2.1 Holistic overview of the production with representation of the production line, unwanted scenarios alert system and workload simulations

The first use case consists of a set of actions related to the current status of the line within a shift. By means of the information provided by the different disaggregated data sources it is possible to analyse this information jointly to establish trends and to make predictions about anomalous situations in the line or the total amount of produced engines at the end of the shift. Thus, the first use case to be worked in the Ford demonstrator is divided in the following problems: **representation of the historical status of the production line** and **estimation of the production at the end of the shift, detection of unwanted scenarios** and **simulations** of new hypothetical situations.

3.2.1.1 Problems to solve

As mentioned above, the first problem is related to the **production overview and line representation**. Ford's databases have different information about the status of operations (whether an operation is cycling a new component, waiting for a new part, blocked or in another possible state), operation failures, cycle times (both actual and design time), number of parts produced in a shift and data related to the quality of the parts produced. In this use case different actions will be carried out. First, different data sources related to production data will be joined to represent the historical status of the production line and to make predictions about the number of engines produced at the end of the shift following the current trend of the line. Both information will help the business experts to understand the significant deviations that may occur between the predicted (planned) production and the actual engines produced at the end of the shift.

The second problem addressed in this use case is the development of a **warning system for undesired situations**, thanks to the information provided by the first problem. Thus, in addition to the information and predictions provided from the overview of the production line, this second application will provide an alert system of anomalous situations that will occur in case the production line behaves according to a certain trend without actions being taken to reverse this situation.

Also taking advantage of the information provided by the previous problems, a third one is included in this use case that focuses on the **simulation of different scenarios** (changes at some point of the production line) to analyze the impact of these changes on the estimated production at the end of the shift.

3.2.1.2 Specific Data Sources

In order to develop this use case, two corporate databases are going to be used: FIS and QLS-CM as also mentioned in the XMANAI Deliverable D1.2.

FIS is a computer-based data collection and reporting system that monitors Machine Performance Data, such as uptime, downtime, blocked and starved conditions, other machine states, machine faults and warnings, and other line conditions. The information is stored in a SQL database. Concretely, from FIS the following data will be taken into account:

- Machine status. The current status of each operation of the line. There are a total number of 13 states, including cycling a component or blocked by other operations, among others.
- Machine faults and warnings. It provides the information related to faults and warnings for each operation, both the number of occurrences and the duration of each one.
- Produced components. The number of components per each operation.
- Cycle Times. Design and real cycle times are available for each operation of the production line.

QLS-CM is a database used as a quality and traceability system that is used to collect birth history data for serialized assemblies and components. It is a database that can be queried to look up the birth history of a component or assembly to determine its manufacturing status, machining or assembly path, test status and quality status. The information is also stored in a SQL database. QLS-CM will provide information on:

- Traceability. Through the serial number of each component, it is possible to place each component in the production line and combine this information with the quality parts and the overview of the whole line.
- Data quality parts. Information related to the quality of the assembled components at each point of the line, such as First Time Through (FTT).

Table 6 Inputs / outputs in the 1st use case of the FORD's demonstrator.

Inputs	Outputs
Machine Status Machine Faults and Warnings Produced Components Cycle Times Traceability Data Quality Parts	Prediction of produced engines at the end of the shift
	Detection of unwanted situations
	Simulation of different scenarios based on the prediction of produced engines

3.2.2 Automated production planning

The second use case focuses its activities on the optimization of different parts of production planning performed by business experts. These production planning problems are divided into different categories, depending on the time window considered. Within a shift, it is important to divide the expected production for a day into batches and, within a batch, the sequencing of the different engines to be produced, attending to a set of constraints established by the business experts. At a higher level, there is a production planning that consists of distributing the engines demanded for a month over the available days of the month, taking into account the manual constraints. Thus, this use case is divided into three problems: the generation of automatic constraints, the sequencing and planning within a shift, and the elaboration of a month's production planning.

3.2.2.1 Problems to solve

As mentioned, the automated production planning part aims to solve different problems related to the current way of planning production, which may not be optimal due to the lack of capacity to process and take into account the information about the historical status of the production line, as well as the demand of engines in the different work shifts. Thus, the first problem of this use case will focus on the **generation of constraints for production planning** based on the current state of the production line. The second problem will focus on the **development of a system for sequencing and planning the production** of the different engines during a shift, replacing the current tool with one that takes into account both the demand and the previously generated constraints. Finally, the last problem will focus on the **development of a planning model** that proposes an automated daily planning for a whole month in addition to those proposed by the business experts, taking advantage of all available data sources.

3.2.2.2 Specific Data Sources

In this use case, in addition to the use of the information provided by the first use case, which leverages on the information provided by FIS and QLS-CM, another database will be used. The Datamart database is a mainframe computer system used to support Ford's assembly and manufacturing plants worldwide. It is used to support the following operations: shipping, receiving, inventory, scheduling, release, bar coding, warehousing and accounting. This system is also used to support electronic communication between Ford, its suppliers and customers. This system is used to view demand, production plan, available parts and components for different engines in order to plan the best production mix. Specifically, Datamart is used in this use case to analyse the following data sources:

- Availability parts. These parts provide information on the availability of the different components required to assemble an engine.
- Monthly expected production. Consists of the total number of engines of each derivative to be produced in a month.

Table 7 Inputs / outputs in the 2nd use case of the FORD's demonstrator.

Inputs	Outputs
Machine Status Machine Faults and Warnings Produced Components Cycle Times	Generation of constraints Sequencing and planning of the different engines during a shift

Traceability Data Quality Parts Availability Parts Monthly Expected Production	Automated Monthly Planning

3.2.3 Explainability needs

The explanations needed by the demonstrator will be provided in different ways. Regarding the first use case, different predictions will be made based on different data sources describing the temporal evolution of different parameters of the production line. Thus, the results of the system depend on a set of input characteristics. In order to know which of these features influence the outcome, different explainability tools will be taken into account, aiming to show the relevance of the input features in order to know which are the root causes that produces a result. The second use case focuses on the elaboration of different types of automated production schedules. In this case, business experts need to know why one plan is better than another or why a specific action is performed at a certain point in the plan rather than another. Therefore, explainability will be brought to this second use case through the employment of contrastive explanations, which aim to explain the plan by comparing different plans and establishing which one is better based on a set of reward metrics.

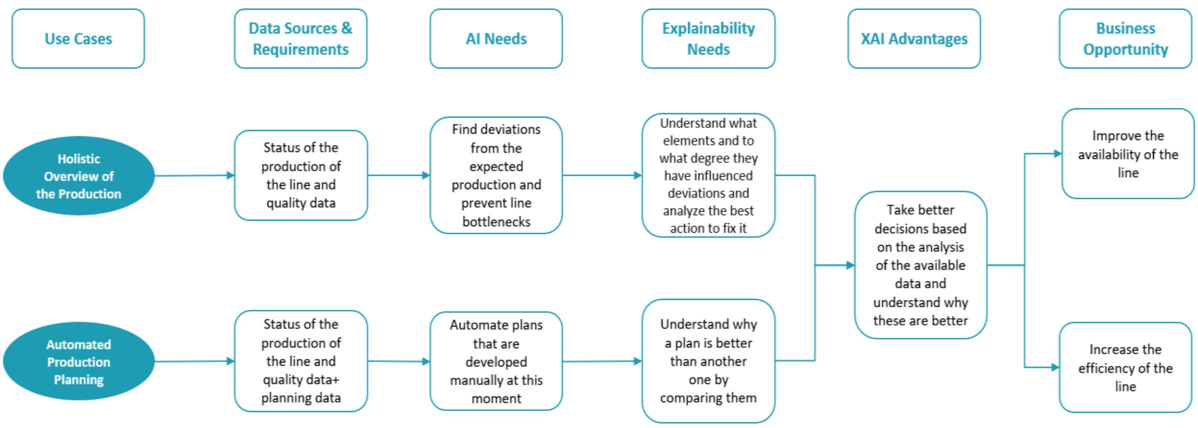


Figure 2 Explainability needs from FORD Demonstrator.

3.2.4 Technical challenges, security aspects and ethics

At the beginning, one of the main challenges to be faced is the preparation of the data. As mentioned above, different data sources will be used during the development of the demonstrator. Consequently, it is mandatory to find a way to combine all available information in order to extract relevant features from the data that can be used as input to the predictive systems. The second technical challenge is related to the elaboration of explanations. Although the draft catalogue of baseline models ensures that any model can be interpretable, these interpretations may be incomprehensible to the business user, as these explanations are provided to data scientists in a first format. Therefore, the translation of these explanations to business users is a key challenge that needs to be addressed during project implementation. In terms of security and ethics, GDPR policies will be complied with, as well as any other confidentiality needs.

3.3 WHIRLPOOL – AI for product demand planning

Whirlpool started implementing AI solutions several years ago, mainly in the manufacturing operations, through research and development projects. The main focus of the launched AI projects were around the manufacturing domain on predictive demand forecasting for white goods spare parts and on finished product demand forecasting. In 2020, though, the D2C "Open Community" was launched with the aim to deliver products within 2-5 working days, depending on the geographical area. The full activation of the D2C channel aims to effectively forecast D2C demand but also acquires a deep knowledge of the business dynamics, while protecting confidential data.

3.3.1 Demand forecasting reliability

Currently, the D2C channel uses the same algorithms as in B2B forecasting, but the goal is eventually to forecast the two channels separately and aggregate the results later on. Now statistical forecasts are generated every month for 18 months rolling based on the historical data for the specific organization, product lifecycle data, on-hand portfolio and other inputs. Added to that, every week, on Monday, there's an additional weekly review process driven by the central demand planning department, which focuses on exceptions and urgency management to support planning. The overall demand forecasting process is embedded into a SAP IBP system, which supports the generation of the Operational Demand Plan (ODP). The central planner is limited to the high-level validation of the forecast profile and to the manual data transfer and enrichment upload into the IBP system. All the information on customers' requests, choices, preferences and attitudes are only manually captured into the demand planning process during the sales enrichment phase and only for those markets which have already activated this new business channel. The forecasting capability of a reliable demand profile is currently very low and therefore there is instability in the supply flow and overproduction or stock breakages.

3.3.1.1 Problems to solve

Taking into consideration the Whirlpool demonstrator description as elaborated in the XMANAI Deliverable D6.1, a set of problems that are relevant to demand forecasting arises. The first problem is about **forecasting reliability** and how to improve it. The D2C demand profile per SKU/day should be generated in a time horizon of 3 months; aligned both with the weekly and monthly DFE targets (Demand Forecast Effectiveness). The second problem is regarding the **reduction of stock in the inventory for D2C**. By implementing a good demand forecasting system, the accuracy of the predictions will ensure minimization of the reserved stock, which will reduce the fixed capital and the obsolescence risk. The third problem is about **maximizing the product availability on request**, which is a key business objective of the Whirlpool supply chain. This will also be achieved with good demand predictability and it represents a crucial factor for production plan actualization in the whole supply chain.

Moving on, as a fourth problem, the XMANAI platform has to help the supply chain of Whirlpool and the D2C sales organization **understand the most important influencing factors** for demand profile and their correlations. This will ensure a higher control over the customer experience and will improve the supply process coordination. The **customers' behaviours need to be understood** (as the fifth problem to be solved); by helping the D2C marketing and sales organization classify customers behaviours and therefore optimize the web services and customer frontend. The sixth problem is about the **understanding of buying patterns**, which is another important business objective. By helping the D2C marketing and sales organization identify the buying patterns, this will support decisions on the offered product range, sales initiatives, development of new products and also website optimization.

3.3.1.2 Specific Data Sources

In order to realize this use case, historical sales transactional data from the web platform will be used to train the AI forecasting algorithms. These data are currently stored on the Google Cloud Platform and can be accessed either through direct download or using a REST-API Endpoint. They contain standard SAP sales orders that are generated when a product is purchased from the e-commerce platform and are extracted on a daily basis. The AI algorithms trained using these data will be applied to generate the forecasted sales per day and SKU. The demand profile will then be visible to the central planner within the XMANAI platform, which will then be downloaded and injected manually to the IBP Whirlpool process and then merged with the traditional B2B demand. The demand profile will also be used for D2C inventory management.

Along with the historical sales transactional data, additional google analytics data will be leveraged in order to understand the business dynamics. Google analytics data are also stored in Google Cloud Platform and can be downloaded either directly or through a REST-API endpoint. They contain clickstream data generated by Google Analytics tags embedded in the e-commerce platforms, such as visitor ID, session ID, visit start time, traffic source, device etc.

Table 8: Inputs / outputs in the use case of the Whirlpool's demonstrator.

Inputs	Outputs
Historical SAP Sales Orders Clickstream data generated by Google Analytics tags Customer and product data	D2C Demand Profile per SKU
	D2C Demand Profile per day
	Seasonal Demand Profile
	Classification of customer's behaviours
	Identification of buying patterns

3.3.2 Explainability needs

Different explainability needs emerge regarding demand forecasting and business dynamics understanding. The capability of effectively forecasting the demand for D2C and also of understanding the business dynamics are crucial for the success in this challenging market sector. So far in Whirlpool, the trust in an AI system was ensured by its capability to produce accurate results. This approach does not effectively incorporate human expertise or explain the forecasting results to humans in the process, which is crucial in certain scenarios such as if the AI system is off-line or in case of major process changes. Therefore, explainability tools can be incorporated to showcase the features (from the historical input data) which had the highest impact on the AI forecasting models decisions and, also, how much these features forced the decision in the specific direction. Similar explainability methods can be applied in order to provide feature correlations regarding the forecasted demand profiles, leading to better understanding of the key factors affecting demand evolution. Added to that, customers' behaviours can be classified into groups of similar characteristics, having statistical descriptions of each group. This will aid in the understanding of customer behaviours, by having characteristics that describe their buying preferences and patterns which are derived from the historical data of previous customers with similar behaviours. Promo initiatives can also be supported by providing impactful features derived for each recommended promotion action.

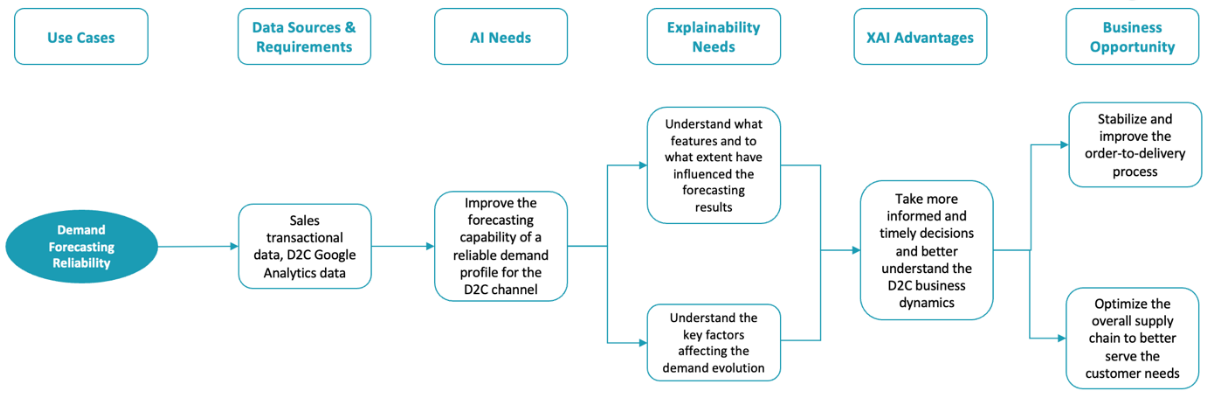


Figure 3 Explainability needs from WHIRLPOOL Demonstrator.

3.3.3 Technical challenges, security aspects and ethics

Data volume and variability are the first technical challenges that arise from this use case. The volume of data is very important as AI models generally yield better results by having more training data, therefore it is a known constraint/challenge that the models can only reach their full potential after a few years of collecting historical data. Variability of data is also very crucial, as it describes how well the data represent the problem. Low variability data might be easier to predict (as the values are more consistent), but higher variability data will explain more possible scenarios and therefore provide more information to the models. Another technical challenge is the diverse explainability needs for the different business users (e.g. D2C sales Manager, Central demand planner, D2C logistic manager, Central Inventory manager, etc.), as the ideal format and explainability information required are different for each business user.

Moving on, security and ethical challenges arise as confidential data are involved with the use cases. GDPR policies along with Whirlpool internal confidential data management policies should be fully incorporated, for data that always remains on premise; as failure to do so might be disruptive both for Whirlpool and their customers.

3.4 CNH - AI for Process quality optimization

Nowadays the use of AI for process optimization in manufacturing is gaining rapid traction with smart factories and Industry 5.0. Nevertheless, at the present time, the CNHi plant does not have any kind of AI solutions implemented on machines. For this reason, thanks to the XMANAI platform the future vision for the CNHi factory will be the implementation of XAI to improve the performance of the line production, in terms of saving costs and time for machineries maintenance and avoiding unnecessary stops of machineries.

3.4.1 Predictive Maintenance

The first problem of the XMANAI CNHi demonstrator focuses on a Heller 400 work centre, a 3 axis CNC milling with horizontal mandrel, in the Modena plant. In detail, machine learning algorithms will apply intelligent predictive maintenance on the electronic boards to prevent its failure ahead of time.

3.4.1.1 Problems to solve

The maintenance history on the Heller 400 work centre in CNHi industrial plant about electronic boards shows 4 failures in a year. Up to the present time, the failures of the electronic boards are treated by CNHi operators using an **empirical approach**; once the fault has been found, one board is replaced at a time to determine which board is the one to be replaced, identifying the broken card. The history of the Heller 400 in the CNHi factory also shows an **average temperature graph** which **corresponds badly to reality**. The main problem is that the sensor installed on the machine in the electrical panel is a single temperature sensor that detects the overall temperature of the panel in which all the electronic cards are.

3.4.1.2 Specific Data Sources

For this use case specific data will be collected from the selected Heller 400 work centre. At the moment, the machine has just one temperature sensor collecting data in all electrical panels. To develop the use case, solve the problem previously described and obtain valid data on the actual wear and/or failure of the electronic boards, CHNi intends to install two temperature and one electrical current sensor on each board. The two temperature sensors will be installed on the hottest and coldest point of the board to understand if there are differences in temperature to trace the type of fault. Once collected data for each card, the temperature and electricity data will be extrapolated to be analysed and crossed with the faults of the electronic boards. The data obtained will train the machine learning algorithm, only after being cleaned considering the surrounding conditions and after that stored on a server, to have data available for the predictive maintenance algorithm.

Table 9 Inputs / outputs in the 1st use case of the CNH's demonstrator.

Inputs	Outputs
Machine Status Machine Faults and Warnings Data from electricity sensors Data from temperature sensors Fault history of electronic boards	Predict the failure of electronic boards in advance explaining the type of fault

3.4.2 XAI – operator collaborative maintenance

Augmented Reality (AR) allows users to integrate virtual context into the physical environment in an interactive multidimensional way catching information about the surrounding environment from cameras and sensors through AR software. Implementing XAI enhances the AR experience by allowing deep neural networks to replace traditional computer vision approaches and add new features such as object detection, text analysis, and scene labelling. These features can be applied to create a collaborative industrial scenario and in detail for the implementation of the second XMANAI use case regarding the simulation of two micro-stops with guided procedure in AR for training purposes.

3.4.2.1 Problems to solve

In the CNHi production plant, the micro-stops of a particular machine are caused by an alarm and managed according to an internal procedure of the company, known only by experienced operators working in that sector, and would not require a substituting piece (which would require the intervention of maintenance technicians). This means that a young inexperienced operator, who does not know the maintenance procedure of the micro-stop, could needlessly call a maintenance technician to restore the machine's functionality. In terms of production efficiency, it means waste and loss of time, costs and a potential part of production.

The AR procedure applied in the CNHi plant will focus on the recovery of the machine micro-stop due to the warning alarm. It consists of a procedure given by a series of guidelines and/or information given by a smart device, such as a smartphone or a tablet, useful to understand and explain what the origin behind the warning is and what is the possible solution to recover the system. The XAI application in this case will provide an integration and collaboration between the machine system and the worker to apply an operator-machine collaborative maintenance.

3.4.2.2 Specific Data Sources

To train the expected XAI solution, data will be collected about faults of micro-stop history, machine warnings, machine status and tasks analysis regarding the procedures for resolving the micro-stops. The expected output is the simulation of different fault scenarios chosen and the prediction of what fault will occur based on the frequency of fault history of the machine. Task analysis, representing the activities performed during the restoration of the micro stop, will be associated also with images of the HMI and buttons involved in the specific steps. This data will be useful to implement an AR application to train the operators.

Table 10 Inputs / outputs in the 2nd use case of the CNH's demonstrator.

Inputs	Outputs
Machine status Machine warnings History fault of micro stop Task analysis of procedures	Simulation of different fault scenarios chosen and prediction of what fault will occur based on the frequency of fault history of the machine

3.4.3 Explainability needs

Two main needs were recognised as relevant for the implementation of XAI in the CNH demonstrator: the need to **reduce** i) **the time of machine downtime** and ii) **the time to detect and restore the fault** in the manufacturing machine area. The two use cases both start from those needs: the first one to reduce the time lost to understand faults on the electronic boards, and optimise maintenance, and the second one to understand how to restore full machine functionality after possible micro-stops.

In the first use case the main objective is the anticipation of electronic boards breakdown through intelligent predictive maintenance by using data coming from temperature and electricity sensors. Thus, obtaining an increase in machine production performance and reducing maintenance costs. These last needs are also valid in the second use case, the detection and recovery of micro-stops, but in this case the main objective is to train operators with an AR solution. The solution will be based on tasks analysis and images of the procedures, to avoid worthless external maintenance intervention.

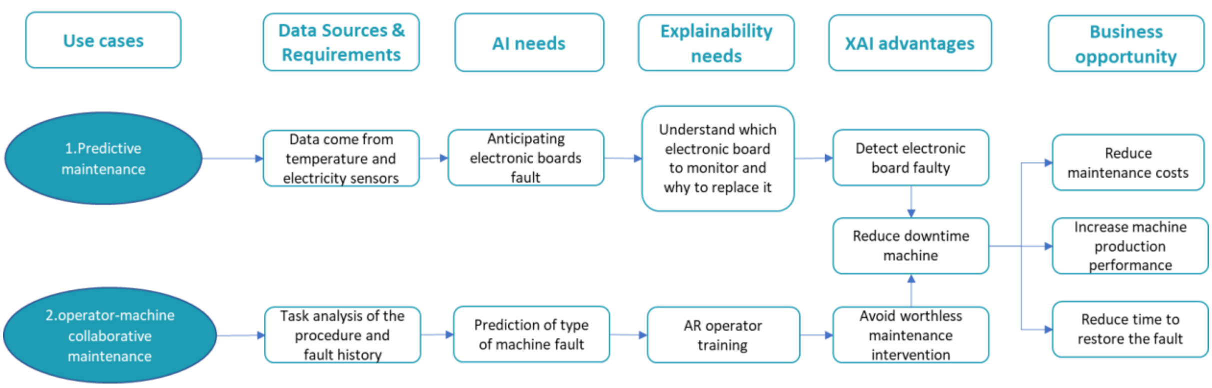


Figure 4 Explainability needs and business opportunities in CNH demonstrator.

3.4.4 Technical challenges, security aspects and ethics

The main technical challenges for those two use cases regard the availability of data. In the first use case, data are available only to a certain extent. The sensors present in the Heller 400 machine are not enough to collect rich data. So, the installation of new sensors is required to have the needed data. At the same time, for the second use case, data need to be collected through performing task analysis and collecting relevant images for the micro-stops resolution, in a process which will be human handled.

For the predictive maintenance use case, ethics aspects are very low and do not present significant risks, there could be potentially security aspects related to data which represent the status of machineries and thus proper security mechanisms should be applied. The second use case instead deals with actions to be performed directly from humans, opening up potential ethical issues related to the control of operators' activities, and actions tracking. A proper ethical assessment should be performed to identify the protective measures needed to be implemented in the design and implementation phase.

3.5 UNIMETRIK - AI for Smart semi-autonomous hybrid measurement planning (Explainable Metrology 4.0)

Unimetrik is a Metrologic service company and ENAC certified calibration laboratory that offers advanced metrology services and solutions to various industries regarding calibration, measurement of complex products and reverse engineering. Unimetrik focuses on innovative applications in metrology following the rapid pace of AI in the manufacturing domain nowadays. Given the company's innovative character and the growth of XAI, the XMANAI platform will provide solutions and improvements to well-known problems and challenges in metrology with respect to measurement and calibration, saving time and costs. UNIMETRIK's demonstrator has been divided into two use cases: (i) optimisation of the measurement plan and (ii) optimisation of point cloud. More information about the two use cases is provided below.

3.5.1 Measurement plan parameter optimization

3.5.1.1 <u>Problems to solve</u>

The first problem of the XMANAI UNIMETRIK demonstrator focuses on measurement plan parameters' optimization. Specifically, this use case copes with the optimization of a number of scanning parameters (lateral density, exposure time and direction density) with ultimate objective **to maximize the measurement accuracy of the scanning device** (thus minimizing the deviation between each data point's nominal and measured positions).

ML techniques (both interpretable/transparent and explainable) will be employed to model the interrelationships between the input parameters and the desired outputs (refer to Table 10). Point clouds, surface orientation and the associated scanning parameters will be the inputs of our data problem, whereas the measurement accuracy / tolerance along with the optimal measurement conditions (that minimize the measurement tolerance) will constitute the desired outputs.

Table 11 Inputs / outputs in the 1st use case of the UNIMETRIK's demonstrator.

Inputs	Outputs

Scanning parameters e.g. lateral density, exposure time, and direction density Point clouds Surface orientation data	Measurement accuracy and/or tolerance defined as the deviation between nominal and actual values
	Optimal measurement parameters

The role of XAI in this use case will be crucial to identify and quantify the impact of the inputs on the decision-making mechanisms of the trained ML models thus enhancing the understanding of the end user (UNIMETRIK metrologists) with respect to the effect of the scanning parameters on the measurement capacity of UNIMETRIK's scanning device.

3.5.1.2 Specific Data Sources

Calibrated artifacts (e.g. reference spheres and other calibrated objects) will be used during the data collection process. The rationale behind the selection of calibrated artifacts is depicted in Figure 5. Specifically, the actual dimensions of a calibrated artifact are almost identical to its nominal dimensions (as defined in the associated CAD file). This means that the measured tolerances (difference between the measured and the nominal dimensions) will be almost identical to the actual tolerances (difference between the measured and the actual dimensions of the object).

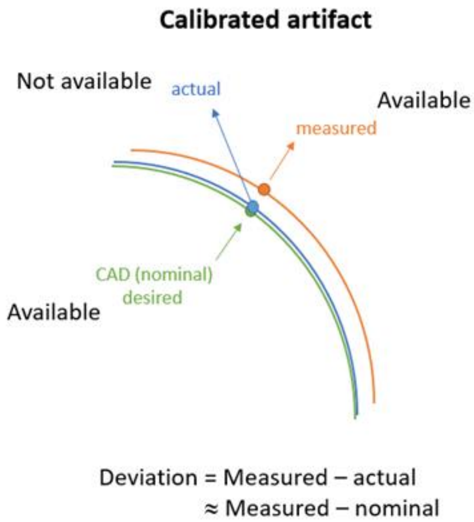


Figure 5 Graphic depicting the principle behind the measurement of tolerance using a calibrated artifact.

In this use case, an extended data collection protocol has been defined whose main points are shortly described here. Each calibrated artifact will be scanned under different scanning parameters. In total, 27 scanning conditions will be applied (3 levels in lateral density × 3 levels in direction density × 3 levels in exposure time). The acquisition of the 27 full point clouds per calibrated artifact will lead to the formulation of the final dataset that will consist of the input/output parameters that were mentioned in Table 10 above. Different validation procedures will be explored: (1) training / testing the predictive capacity of the ML models on the same calibrated artifact; (2) training using data from one calibrated artifact (e.g. sphere) and testing on another (e.g. cylinder) and (3) training / testing using all the available data (using cross-validation or leave one out mechanisms).

3.5.2 Point cloud optimization

The second case study focuses on the optimization of the generated point clouds. Point cloud optimisation is actually defined as the process of selecting or transforming the existing obtained data points towards to the generation of a new point cloud (optimal). This new 'filtered' point cloud will be optimal with respect to the following requirements:

- It will contribute to the calculation of specific object properties (e.g. the diameter in the example of the reference sphere) with higher accuracy compared to the existing techniques that utilize the initial full point clouds.
- It will consist of potentially less data points thus decreasing the computation complexity and the associated processing time that is needed.

3.5.2.1 Problems to solve

The main problem in the point cloud optimization use case is actually twofold: (**Obj1**) to either *filter* the full point cloud or *transform* the initial one towards the generation of the optimal point cloud that ideally includes less data points and at the same time; (**Obj2**) to *identify* the optimal measurement parameters, that are needed to measure accurately the desired feature/characteristic of a specific object (e.g. sphere diameter). The initial full point clouds along with the associated measurement conditions, under which the point clouds have been obtained, will be the main inputs in the data problem of the second use case. The outputs of the ML models of the first use case (estimated tolerances) will be also considered as potential inputs into the data problem (Table 11). ML-based optimisation methodologies and graph ML and/or CNN-based models will be explored for their suitability in implementing the task of predicting the optimal point cloud. XAI will also contribute to identify the contribution of the measurement scanning parameters on the calculation of the desired object properties.

Table 12 Inputs / outputs in the 2nd use case of the UNIMETRIK's demonstrator.

Inputs	Outputs
Full point clouds Scanning parameters Outputs of ML models in use case 1	Optimal point cloud (subset of the full point cloud that leads to the most accurate calculation of one or more desired properties)

3.5.2.2 Specific Data Sources

The same datasets, as in the case of use case 1, will be also used for the purposes of the 2nd use case. This data will be further supplemented with the outcomes of the ML models of the 1st use case (estimated tolerances per data point). The acquisition of full point clouds from several calibrated objects would significantly enlarge the dataset size allowing us to use more powerful and at the same time more complex deep learning techniques.

3.5.3 Explainability needs

Overall, in the UNIMETRIK measurement and point cloud optimization problems, explainability can provide insights on how each parameter of the exported dataset affects the data points measurement and geometrical properties' calculation. Through XAI libraries and techniques such as SHAP, LIME and gLIME the non-expert end-users will be able to clearly understand the importance of the included parameters in the specific optimization plans uncovering the decision-making mechanism of the trained ML models.

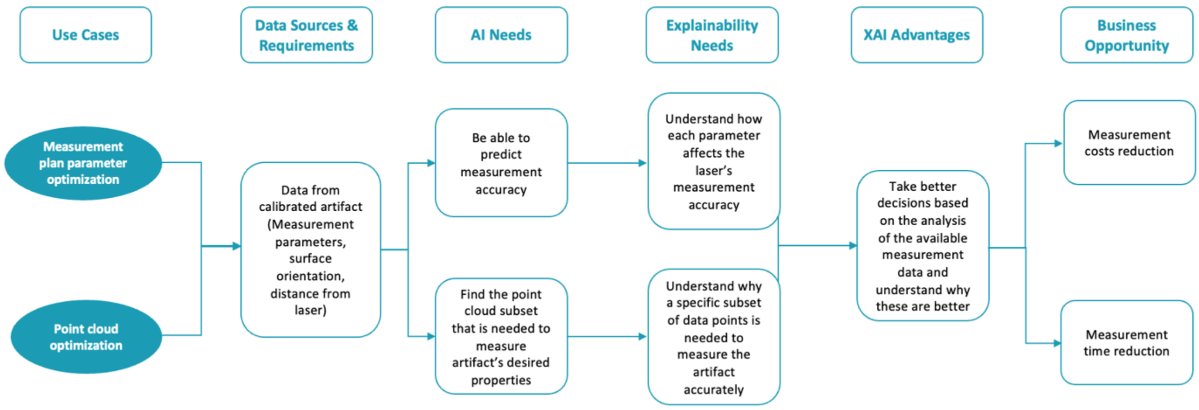


Figure 6 Explainability needs in UNIMETRIK's demonstrator.

3.5.4 Technical challenges, security aspects and ethics

Two main technical challenges rise in the UNIMETRIK optimization problems. First, the available data is limited, as it is costly to export huge amounts of data in metrology, and therefore the number of ML models that can be applied to the problem is limited. Second, there is a need for calibrated samples in order to train the models in an appropriate way. Until now, only the calibrated sphere problem has been explored, while more calibrated artefacts should be considered for validating the generalization capacity of the implemented models. Because of the high-cost data, it is very important to apply well-known SoA security mechanisms, such as back-up and encryption tools in order to preserve data integrity and confidentiality. Finally for the ethics part, we do not foresee any aspect of high importance.

4 Draft catalogue of XMANAI baseline algorithms

4.1 Overview

4.1.1 Identified tasks to be addressed by XMANAI demonstrators

The Draft catalogue of XMANAI baseline algorithms will be populated by a set of Hybrid AI and Graph ML models that cover the needs of the XMANAI demonstrators, described in detail in the previous section (Section 3) of the deliverable. Based on the technical description of the demos use cases and the identification of sub-tasks within each use case, we can summarize the set of problems to be addressed by models in the XMANAI draft catalogue:

1. Production Optimization
    - Production forecasting
    - Anomaly detection
    - Automated Production planning
2. Product demand forecasting
    - Reliable demand forecasting (daily, weekly, monthly)
    - Inventory management
    - Optimization of product availability
    - Understanding buying patterns
    - Supporting promo initiatives
3. Process/Product quality optimization
    - Intelligent diagnosis/prognosis of faults
    - Explainable decision support system for predictive maintenance
4. Process optimization and Semi-Autonomous planning
    - Optimization of process parameters
    - Optimization of point cloud

A major axis for the construction of the Draft catalogue of XMANAI baseline algorithms is to provide explainable AI solutions to the above stated problems. The selection of baseline algorithms, described in the following section, is therefore intended to satisfy the two-fold purpose of providing performant solutions that also fulfill the requirements of the different business users for explanations.

4.1.2 Algorithm selection

The selection of Hybrid and Graph ML models to populate the Draft catalogue of XMANAI baseline algorithms is a two-step procedure. Firstly, a primary model (ML/DL, Graph ML/DL) is depicted as suitable to address at least one of the above stated problems, by means of adoption by the technical partners, reported efficiency and scalability. Secondly, each primary model is coupled with an explainability component in order to provide interpretable solutions to these problems. The rationale driving the first step of primary model selection, is based on the technical partners experience regarding the suitability of algorithms to solve particular tasks. Efficiency and scalability of the selected primary models will be evaluated in practice during the experimental phase in the realistic industrial settings of the demonstrators. This evaluation will be made in order to ensure that the selected models in the final trained catalogue will be able to accomplish with the requirements in terms of efficiency and scalability. We anticipate that different sets of algorithms will fit different usage scenarios with respect to cost-effectiveness and timely constraints. The basic implementations of the primary algorithms are considered here, to facilitate the coupling of primary models with explainability techniques. Modified versions of the primary algorithms, that are meant to meet case-specific requirements (such as increasing the speed of computations) will be examined within the iterative process of experimentation, as the project progresses. As for the second step of identifying the appropriate explanatory components to equip each primary model, the selection of methods has focused on providing explanations that can provide generic interpretable knowledge using methods that are valid for data scientists. Henceforth, one of the main objectives of the remaining WP4 deliverables will be to add to these explanations complementary explanations that can be rich and valuable for business users, according to their needs.

As this is a draft catalogue, not all the algorithms will be implemented. In further steps of WP4, we will need to make a proper selection of the algorithms considering the maturity of the libraries and implementations, as well as the suitability to be adapted to the real data of the demonstrators.

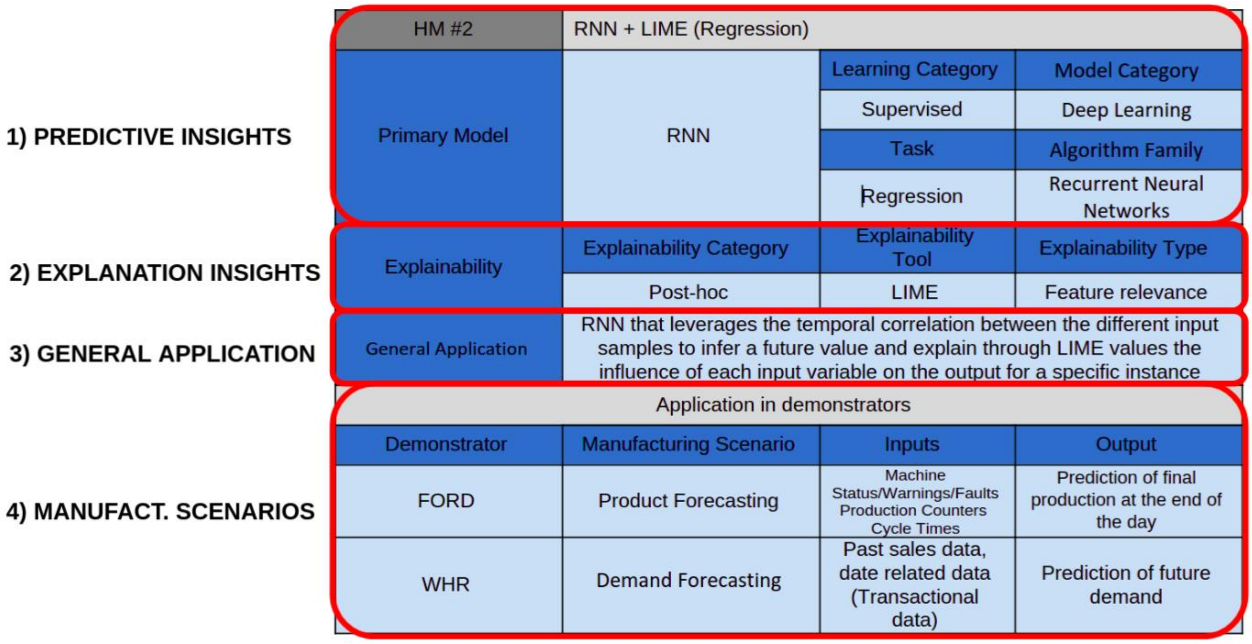
The overall process that was established to construct the draft catalogue is described as follows:

- A set of Hybrid and Graph ML models were proposed to address the problems put forth by each demonstrator. This is accomplished by collaborative effort between the consortium business and technical partners assigned to work on each pilot.
- Information on the proposed primary models and their corresponding explainability components is translated into model cards.
- The model cards were consolidated in order to:
    - Summarize the use of a single model in more than one demonstrator
    - Achieve a unified form in both Hybrid and Graph ML algorithm cards

In the following sections we describe in detail the scope and characteristics of Hybrid and Graph ML algorithms, also providing the respective model cards for each category.

4.1.3 Model Cards Description

The Model Cards aim to synthesise and group in a visual way information about the different baseline models selected from different points of view.



	HM #2	RNN + LIME (Regression)		
1) PREDICTIVE INSIGHTS	Primary Model	RNN	Learning Category	Model Category
			Supervised	Deep Learning
			Task	Algorithm Family
			Regression	Recurrent Neural Networks
2) EXPLANATION INSIGHTS	Explainability	Explainability Category	Explainability Tool	Explainability Type
		Post-hoc	LIME	Feature relevance
3) GENERAL APPLICATION	General Application	RNN that leverages the temporal correlation between the different input samples to infer a future value and explain through LIME values the influence of each input variable on the output for a specific instance		
4) MANUFACT. SCENARIOS	Application in demonstrators			
	Demonstrator	Manufacturing Scenario	Inputs	Output
	FORD	Product Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
	WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

Figure 7 Algorithm Card Structure - Example.

The information provided in the card is divided into 4 groups:

- Predictive Insights: Information about the model employed to solve a specific problem.
- Explanation Insights: Details about the kind of explainability that complements the predictive results.

- General Application: In which application the model is expected to be employed.
- Manufacturing Scenarios: Specific information of their potential implantation to solve the manufacturing scenarios from the demonstrators, considering their specific inputs and outputs.

4.2 Hybrid AI algorithms (Catalogue Model Cards)

4.2.1 Description

In this section, a number of hybrid AI algorithms are presented. First, it is important to clarify which types of algorithms fall into this group. Hybrid algorithms are those models that are composed of a traditional Machine or Deep Learning black-box algorithm and a layer of explainability using different explainability tools. Transparent models, which are interpretable by design, will also be considered in this group. The taxonomy employed to categorize the different explainability tools to be combined with the black-box models is described in D1.1 "*State of the Art Review in XMANAI Research Domains*". Although in this deliverable Transparent, Post-hoc and Hybrid approaches are grouped separately, in D4.1, all of them are considered as part of the Hybrid ML models in order to be able to include all the categories in the draft catalogue.

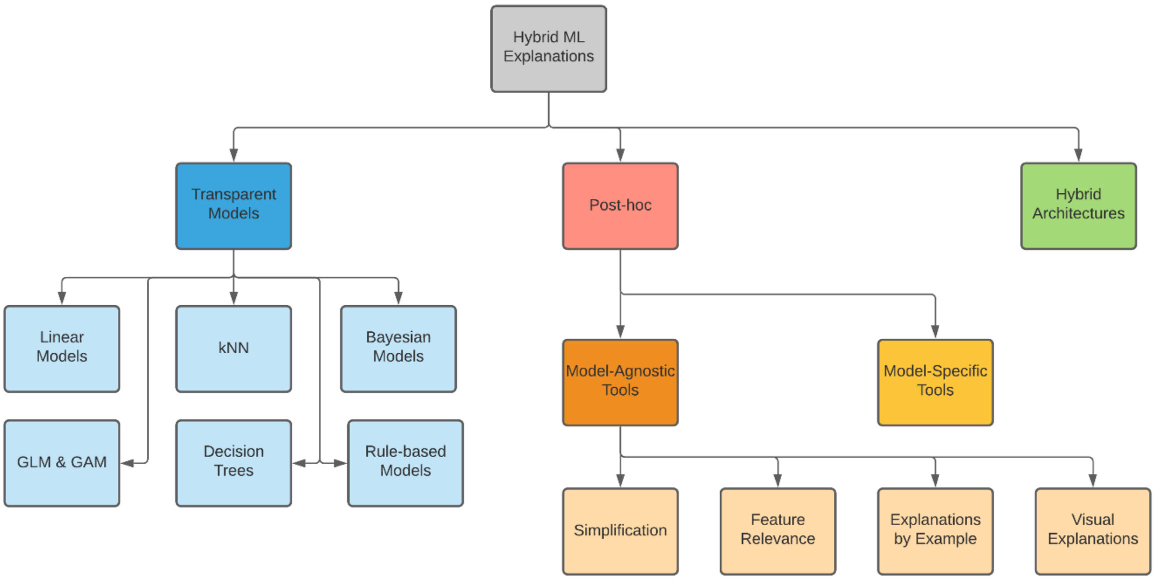


Figure 8 Hybrid ML Explainability from D1.1 "State of the Art Review in XMANAI Research Domains".

Following this approach, the next explainability categories are included:

- Transparent models: Those that are interpretable by design without the need for any explainability tool, such as decisions trees or linear models.
- Post-hoc techniques: A second method is employed to explain the decisions made by a previous black-box model. Two groups are included in this category: model-agnostic techniques and model-specific.
    - Model-agnostic: This group of explainability tools can be applied to any ML model, regardless of the family it belongs. Most of the proposed baseline models include a model-agnostic post-hoc tool to interpret their predictions. Different model-agnostic approaches can be used to provide explainability:

        - Explanations by simplification: This group includes explainability approaches that aim to approximate a black-box model with a transparent model, such as LIME (Ribeiro, et al., 2016) or G-REX (Konig, et al., 2008).
        - Feature relevance explanations: These methods highlight the influence or contribution of each input feature to the black box results, as in SHapley Additive exPlanations (Lundberg & Lee, 2017).
        - Explanations by example: This group includes the type of explanations that focus on understanding the Machine Learning model by selecting appropriate input examples, such as the counterfactual explanations.
        - Visual explanations: These techniques aim to provide to the end user a visually conceivable explanation of a black-box model's prediction, in form of heatmaps or plot dependencies among others. Often, these visualizations are combined with other explainability techniques to improve their understanding. Such explanations also allow the user to understand the internal reasoning process of transparent models with the aim of making them even more transparent.
    - Model-specific: This group includes methods that only apply to a particular family of algorithms, such as DeepRED (Zilke, et al., 2016) for multilayer neural networks.
- Hybrid architectures: Attention mechanisms will be considered in this group. These approaches can be considered as feature relevance explainability methods.

4.2.2 Algorithms (Cards)

4.2.2.1 Machine Learning

Support Vector Machines (SVM)

HM #1	SVM + SHAP (Classification)		
Primary Model	SVM	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Support Vector Machines
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature Relevance
General Application	SVM predicts when an electronic board will fail soon, finding a line/hyperplane (in multidimensional space) that separates two classes (Card faulty and Working card) and explains its predictions using SHAP.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Anomaly Detection	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

HM #2	SVM + LIME (Classification)		
Primary Model	SVM	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Support Vector Machines
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature Relevance
General Application	SVM predicts when an electronic board will fail soon, finding a line/hyperplane (in multidimensional space) that separates two classes (Card faulty and Working card) and explains its predictions using LIME.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Anomaly Detection	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

HM #3	SVM + Decision Tree (Classification)		
Primary Model	SVM	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Support Vector Machines
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Decision Tree	Simplification
General Application	SVM classification model learns a high dimensional hyper-plane based on the input data to predict whether supplier-related risks will occur. After training a less complex Decision Tree model is built based on the SVM to extract rules using the support vectors of the trained model.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Supply Base Decisions	Product/ supplier IDs, order number, quantity ordered, due date, receipt date (Transactional data, Consumer & product master)	Prediction of supplier-related risk occurrence

HM #4	SVM + SHAP (Regression)		
Primary Model	SVM	Learning Category	Model Category

		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Support Vector Machines
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature relevance / Local Explanations
General Application	Kernel SVR that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions both locally and globally using SHAP. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #5	SVM + LIME (Regression)		
Primary Model	SVM	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Support Vector Machines
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Local Explanations
General Application	Kernel SVR that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using LIME. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Predictive Maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #6	SVM + gLIME (Regression)

Primary Model	SVM	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Support Vector Machines
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	gLIME	Graphical / Local Explanations
General Application	Kernel SVR that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using gLIME. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #7	SVM + Permutation Importance (Regression)		
Primary Model	SVM	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Support Vector Machines
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Permutation Importance	Feature Relevance
General Application	Kernel SVR that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using feature permutation importance. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #8	SVM + Decision Tree (Regression)		
Primary Model	SVM	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family

		Regression	Support Vector Machines
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Decision Tree	Simplification
General Application	Measurement plan parameter optimization		
	Kernel SVR that predicts the accuracy and/or tolerance in the dimensional measurement of an object and uses a single Decision Tree (equivalently set of rules) as a global surrogate to explain its predictions. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance Optimal measurement parameters

Extreme Gradient Boosting (XGBoost)

HM #9	XGBoost + LIME (Regression)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Extreme Gradient Boosting
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature relevance
General Application	Measurement plan parameter optimization		
	XGBoost tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using LIME. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output

FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance Optimal measurement parameters

HM #10	XGBoost + LIME (Classification)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Extreme Gradient Boosting
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature relevance
General Application	XGBoost classification model is trained based on the input data to predict whether backorders will occur. After training, the LIME method assigns importance scores on the inputs based on their contribution to the prediction.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Inventory Management	Product/ supplier IDs, order number, quantity ordered, due date, receipt date (Transactional data, Consumer & product master)	Prediction of backorders occurrence

HM #11	XGBoost + gLIME (Regression)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Extreme Gradient Boosting
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	gLIME	Graphical / Local Explanations

General Application	XGBoost tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using gLIME. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #12	XGBoost + TreeSHAP (Regression)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Extreme Gradient Boosting
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	TreeSHAP	Feature relevance
General Application	XGBoost regression model is trained to predict the future demand based on the inputs. After the training, TreeSHAP is used to explain how each feature contribute to the model output.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

HM #13	XGBoost + SHAP (Regression)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Extreme Gradient Boosting
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature relevance
General Application	XGBoost tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using SHAP. The optimal measurement parameters will be defined based on model predictions.		

Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #14	XGBoost + SHAP (Classification)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Extreme Gradient Boosting
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature relevance
General Application	XGBoost classification model is trained based on the input data to predict whether backorders will occur. After training, the SHAP method assigns importance scores on the inputs based on their contribution to the prediction.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Inventory Management	Product/ supplier IDs, order number, quantity ordered, due date, receipt date (Transactional data, Consumer & product master)	Prediction of backorders occurrence

HM #15	XGBoost + Decision Tree (Regression)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Extreme Gradient Boosting
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Decision Tree	Simplification
General Application	XGBoost tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and uses a single Decision Tree (equivalently set of rules) as a global surrogate to explain its predictions. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output

UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #16	XGBoost + Permutation Importance (Regression)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Extreme Gradient Boosting
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Permutation Importance	Feature Relevance
General Application	XGBoost tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using feature permutation importance. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

Random Forest

HM #17	Random Forest + SHAP (Classification)		
Primary Model	Random Forest	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Random Forest
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature relevance
General Application	Random Forest classification model is trained based on the input data to predict whether backorders will occur. After training, SHAP assigns importance scores on the inputs based on their contribution to the prediction.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Inventory Management	Product/ supplier IDs, order number, quantity ordered, due date, receipt	Prediction of backorders occurrence

		date (Transactional data, Consumer & product master)	
CNH	Predictive Maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

HM #18	Random Forest + LIME (Classification)		
Primary Model	Random Forest	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Random Forest
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature relevance
General Application	Random Forest classification model is trained based on the input data to predict whether backorders will occur. After training, LIME assigns importance scores on the inputs based on their contribution to the prediction.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Inventory Management	Product/ supplier IDs, order number, quantity ordered, due date, receipt date (Transactional data, Consumer & product master)	Prediction of backorders occurrence

HM #19	Random Forest + LIME (Regression)		
Primary Model	Random Forest	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Random Forest
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature Relevance
General Application	Random Forest regression model is trained to predict the future demand based on the inputs. After the training, LIME is used to explain how each feature contribute to the model output.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output

FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance Optimal measurement parameters

HM #20	Random Forest + TreeSHAP (Regression)		
Primary Model	XGBoost	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Random Forest
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	TreeSHAP	Feature relevance
General Application	Random Forest regression model is trained to predict the future demand based on the inputs. After the training, TreeSHAP is used to explain how each feature contribute to the model output.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

HM #21	Random Forest + SHAP (Regression)		
Primary Model	Random Forest	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Random Forest
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature relevance

General Application	Random Forest tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions locally and globally using SHAP. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Predictive Maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #22	Random Forest + gLIME (Regression)		
Primary Model	Random Forest	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Random Forest
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	gLIME	Graphical / Local Explanations
General Application	Random Forest tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using gLIME. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #23	Random Forest + Decision Tree (Regression)		
Primary Model	Random Forest	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Random Forest
Explainability	Explainability Category	Explainability Tool	Explainability Type

	Global Surrogate	Decision Tree	Simplification
General Application	Random Forest tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and uses a single Decision Tree (equivalently ruleset) as a global surrogate to explain its predictions. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #24	Random Forest + Permutation Importance (Regression)		
Primary Model	Random Forest	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Random Forest
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Permutation Importance	Feature Relevance
General Application	Random Forest tree-based ensemble that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using feature permutation importance. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

Decision Trees

HM #25	Decision Tree (Classification)		
Primary Model	Decision Tree	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Decision Trees
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Transparent	Self-explained	Self-explained

General Application	Decision Tree classification model is trained based on the input data to predict whether supplier-related risks will occur. The model is self-explanatory through the human readable rules that allow a direct understanding of the prediction path.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Supply Base Decisions	Product/ supplier IDs, order number, quantity ordered, due date, receipt date (Transactional data, Consumer & product master)	Prediction of supplier-related risk occurrence

HM #26	Decision Tree (Regression)		
Primary Model	Decision Tree	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Decision Trees
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Transparent	Self-explained	Self-explained
General Application	Decision Tree that predicts the accuracy and/or tolerance in the dimensional measurement of an object. Model predictions are explained directly by the graphical representation of the tree and the analysis of the decision path. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Predictive Maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #27	iForest + SHAP (Classification)		
Primary Model	Decision Tree	Learning Category	Model Category
		Unsupervised	Machine Learning
		Task	Algorithm Family
		Clustering	Decision Trees

Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature Relevance
General Application	In an Isolation Forest, randomly sub-sampled data is processed in a tree structure based on randomly selected features and it explains its output through Shapley values.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Supply Base Decisions	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

Others

HM #28	k-Nearest Neighbors (Classification)		
Primary Model	kNN	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Neighbors
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Transparent	Self-explained	Self-explained
General Application	kNN Is an approach to data classification that estimates how likely a data point is to be a member of one group or the other depending on what group the data points nearest to it are in.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Supply Base Decisions	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

HM #29	Logistic Regression (Classification)		
Primary Model	Logistic Regression	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Classification	Linear Models
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Transparent	Self-explained	Self-explained

General Application	It is used to understand the relationship between the dependent variable and one or more independent variables by estimating probabilities using a logistic regression equation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Supply Base Decisions	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

HM #30	Local Outlier Factor (LOF) + SHAP (Clustering)		
Primary Model	Decision Tree	Learning Category	Model Category
		Unsupervised	Machine Learning
		Task	Algorithm Family
		Clustering	Neighbors
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature Relevance
General Application	It measures the local deviation of the density of a given sample with respect to its neighbors and explains its output through Shapley values.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Supply Base Decisions	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

HM #31	Linear Regression (Ridge, Lasso) (Regression)		
Primary Model	Linear Regression	Learning Category	Model Category
		Supervised	Machine Learning
		Task	Algorithm Family
		Regression	Linear Models
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Transparent	Self-explained	Self-explained
General Application	Linear regression model that predicts the accuracy and/or tolerance in the dimensional measurement of an object. Model predictions are explained directly by the weight coefficients of the linear transformation. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			

Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

4.2.2.2 Deep Learning

Recurrent Neural Networks (RNN)

HM #32	RNN + SHAP (Regression)		
Primary Model	RNN	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Recurrent Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature relevance
General Application	RNN that leverages the temporal correlation between the different input samples to infer a future value and explain through SHAP values the influence of each input variable on the output		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

HM #33	RNN + LIME (Regression)		
Primary Model	RNN	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Recurrent Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature relevance
General Application	RNN that leverages the temporal correlation between the different input samples to infer a future value and explain through LIME values the influence of each input variable on the output for a specific instance		
Application in demonstrators			

Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

HM #34	RNN + DeepLift (Regression)		
Primary Model	RNN	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Recurrent Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	DeepLift	Feature relevance
General Application	RNN learns the future value of product demand based on the input features. After the training, DeepLift is used to assign input features contributions to the output.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

HM #35	RNN + Decision Tree (Regression)		
Primary Model	RNN	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Recurrent Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Decision Tree	Simplification
General Application	RNN that leverages the temporal correlation between the different input samples to infer a future value. After the RNN training a decision tree is trained on the predictions of the RNN so as to add transparency.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output

FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day

HM #36	RNN + Attention (Regression)		
Primary Model	RNN	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Recurrent Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Hybrid	Feature Relevance	Attention
General Application	RNN that leverages the temporal correlation between the different input samples to infer a future value. An attention mechanism is used simultaneously to highlight the contribution of each input feature to the output.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day

__Convolutional Neural Networks (CNN)__

HM #37	CNN + SHAP (Regression)		
Primary Model	CNN	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Convolutional Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature relevance
General Application	CNN learns the future value of product demand based on the input features. After the training, SHAP is used to explain how each feature contributes to the model output.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

HM #38	CNN + LIME (Regression)

Primary Model	CNN	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Convolutional Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature relevance
General Application	CNN learns the future value of product demand based on the input features. After the training LIME is used to explain how each feature contributes to the model output.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

HM #39	CNN + DeepLift (Regression)		
Primary Model	CNN	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Convolutional Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	DeepLift	Feature relevance
General Application	CNN learns the future value of product demand based on the input features. After the training DeepLift is used to assign input features contributions to the output.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Demand Forecasting	Past sales data, date related data (Transactional data)	Prediction of future demand

Autoencoders

HM #40	Autoencoder + SHAP (Clustering)		
Primary Model	Autoencoder	Learning Category	Model Category
		Unsupervised	Deep Learning
		Task	Algorithm Family
		Clustering	Autoencoders

Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Feature Relevance	SHAP
General Application	An autoencoder that aims to reconstruct its input data and alert of anomalies (high reconstruction error). Through SHAP, the contribution of each input feature to the anomaly will be presented.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Anomaly Detection	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day

HM #41	Autoencoder + LIME (Clustering)		
Primary Model	Autoencoder	Learning Category	Model Category
		Unsupervised	Deep Learning
		Task	Algorithm Family
		Clustering	Autoencoders
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature Relevance
General Application	An autoencoder that aims to reconstruct its input data and alert of anomalies (high reconstruction error). Through LIME, the contribution of each input feature to the anomaly will be presented.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Anomaly Detection	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day

HM #42	Autoencoder + Attention (Clustering)		
Primary Model	Autoencoder	Learning Category	Model Category
		Unsupervised	Deep Learning
		Task	Algorithm Family
		Clustering	Autoencoders
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Hybrid	Attention	Feature Relevance
General Application	An autoencoder that aims to reconstruct its input data and alert of anomalies (high reconstruction error). An attention mechanism is used simultaneously to highlight the contribution of each input feature to the output		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output

FORD	Anomaly Detection	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of final production at the end of the day

Neural Networks

HM #43	Neural Network (NN) + LIME (Regression)		
Primary Model	Neural Network	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Artificial Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	LIME	Feature relevance / Local Explanations
General Application	Neural Network that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using LIME. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #44	Neural Network (NN) + SHAP (Regression)		
Primary Model	Neural Network	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Artificial Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	SHAP	Feature relevance
General Application	Neural Network that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using SHAP. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #45	Neural Network (NN) + gLIME (Regression)		
Primary Model	Neural Network	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Artificial Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	gLIME	Graphical / Local Explanations
General Application	Neural Network that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using gLIME. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #46	Neural Network (NN) + Decision Tree (Regression)		
Primary Model	Neural Network	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Artificial Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Global Surrogate	Deep Red	Simplification / Global Explanations
General Application	Neural Network that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using a global surrogate Decision Tree. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #47	Neural Network (NN) + DeepLift (Regression)		
Primary Model	Neural Network	Learning Category	Model Category
		Supervised	Deep Learning

		Task	Algorithm Family
		Regression	Artificial Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	DeepLift	Local Explanations / Feature Relevance
General Application	Neural Network that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using DeepLift. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

HM #48	Neural Network (NN) + Permutation Importance (Regression)		
Primary Model	Neural Network	Learning Category	Model Category
		Supervised	Deep Learning
		Task	Algorithm Family
		Regression	Artificial Neural Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Post-hoc	Permutation Importance	Feature Relevance
General Application	Neural Network that predicts the accuracy and/or tolerance in the dimensional measurement of an object and explains its predictions using feature permutation importance. The optimal measurement parameters will be defined based on model predictions.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

4.3 Graph ML algorithms (Catalogue Model Cards)

4.3.1 Description

The set of Graph ML algorithms that were selected for the Draft Catalogue of XMANAI baseline models is presented in this section. Similarly to Hybrid AI models, the XMANAI Graph ML baseline algorithms are composed of a primary black-box model coupled with an explainability layer. Knowledge Graph modelling is also considered here, as the interpretable by design solution to achieve explainability at the data level and examine underlying causal relations. Based on the taxonomy documented in chapter 3.4 of the deliverable D1.1-"State of the art review in XMANAI research domains", the primary black-box models to be included in the Draft Catalogue belong to the following categories and families of Graph ML/DL algorithms:

1. Geometric Deep Learning
   - Spatio-Temporal Graph Neural Networks
   - Convolutional Graph Neural Networks
   - Graph Attention Networks
2. Graph Representation Learning
   - Graph Neural Networks & Auto-Encoders
   - Random Walk based embeddings

Regarding the explainability techniques that are applied to these black-box Graph Neural Networks (GNNs) and embeddings, the D1.1 review was not extended towards that direction. We therefore adopt the taxonomy presented in (Yuan, et al., 2021) which is summarized in the below picture.

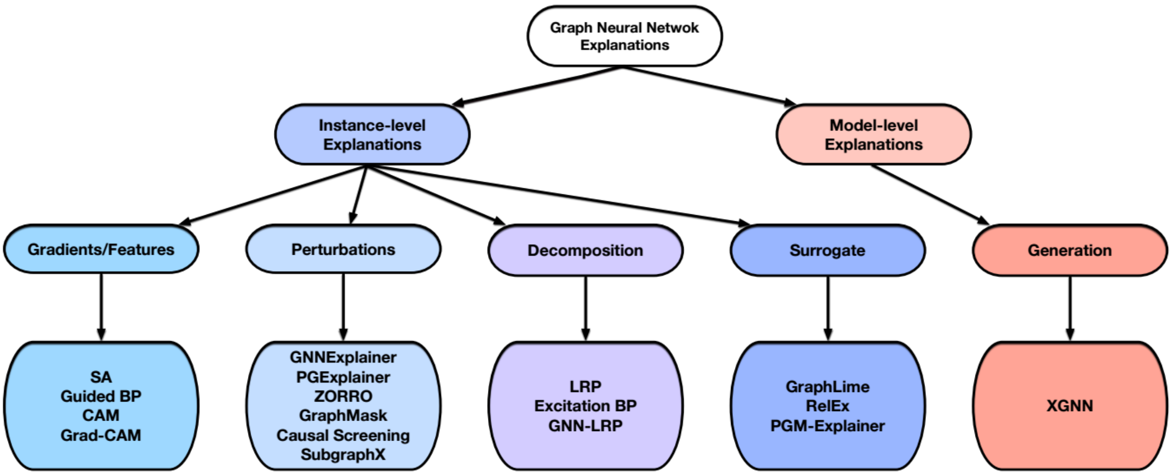


Figure 9 The taxonomy proposed by (Yuan, et al., 2021) for explaining GNNs.

The methods selected by technical partners to equip the primary Graph ML algorithms with explainability components, cover the spectrum of instance-level explanations in the proposed taxonomy, namely:

- **Gradients/Features based explanations**: Gradient-based methods for Deep Neural Networks, such as Sensitivity Analysis (SA) (Simonyan, et al., 2014) and Guided BackPropagation (Springenberg, et al., 2015), use back-propagation (BP) to explain a certain output by means of the gradients of the model's decision function with respect to the input features. The explanations are offered by both methods in the form of Saliency maps. SA and Guided BP have been applied to GNNs by (Baldassarre & Azizpour, 2019). Feature-based methods such as Class Activation Mapping (CAM) and variant gradient-weighted CAM (Grad-CAM), interpolate the hidden features extracted for a specific output to the input feature space and create hidden feature maps to explain the prediction. CAM and Grad-CAM are applied to GNNs by (Pope, et al., 2019).
- **Perturbation based explanations**: These methods study the effect of perturbing the input features to the model's outcome. Methods such as GNNExplainer (Ying, et al., 2019) and ZORRO (Funke, et al., 2021) are developed for deep GNNs, based on generating node, edge or node feature masks under perturbation of the input nodes or edges (depending on the task) and combining them with the input Graph to explain the model's decisions.

- **Decomposition based explanations**: Methods in this category decompose a model's prediction into a combination of the inputs, while all terms in the decomposition must sum up to the final prediction score. Examples of such methods applied to GNNs include Layer-wise Relevance Propagation (LRP) (Baldassarre & Azizpour, 2019) and Excitation BP (Pope, et al., 2019). To overcome limitations of these methods in explaining the importance of single nodes, (Schnake, et al., 2021) propose GNN-LRP to apply LRP for different graph walks and explain the output as a collection of walks that are relevant to the prediction.
- **Surrogate model:** These methods offer explanations by locally approximating the GNN model's predictions with an interpretable surrogate model. For example, Graph-LIME by (Huang, et al., 2020) fits a Hilbert-Schmidt Independence Criterion (HSIC) Lasso, a kernel-based method as local surrogate, on a local dataset comprised by neighbouring nodes and their corresponding predictions. The weights of HSIC Lasso are used to explain both the surrogate and the original GNN model's node prediction. A different approach based on Probabilistic Graphical Modelling is followed by (Vu & Thai, 2020). They propose PGM-Explainer to provide local explanations to a GNN model's predictions by an interpretable Bayesian Network, fit to the local dataset which is created under random perturbations of the input graph node features. PGM-Explainer can explain the outcomes of both node and graph classification tasks, whereas Graph-LIME is limited to explaining node predictions.

Finally, a recent decomposition-based approach that incorporates elements from surrogate and perturbation-based methods, is also considered as an option for explaining GNNs. (Duval & Malliaros, 2021) extend the computation of Shapley values from co-operative game theory to graphs. They propose GraphSVX explainer to decompose a local prediction into node and feature contributions, by fitting a local surrogate to a perturbed local dataset. Validated on both node and graph classification, GraphSVX is found to provide accurate, meaningful and robust explanations.

Although a single method is included in the algorithm cards for each type of instance-level explanations proposed by technical partners, it is anticipated that other methods of the same type may be also considered during the experimental phase of the project. For example, GNNExplainer may be an initial method of choice to provide Perturbation based explanations to a GNN predictions, however the use of ZORRO may also be examined in the future for the same purpose. Finally, in most cases a single Graph ML primary model is coupled with more than one type of methods (e.g., perturbation based and surrogate), with the aim to enable the provision of rich explanations to the various end users.

The composite Graph ML baseline algorithm cards are presented in the next section, categorized with respect to the primary models.

4.3.2 Algorithms (Cards)

4.3.2.1 Geometric Deep Learning

Spatio-Temporal Graph Neural Networks

GM #1	ATGN + GNNExplainer (Clustering)		
Primary Model	Attentive Temporal Graph Network	Learning Category	Model Category
		Unsupervised	Geometric Deep Learning
		Task	Algorithm Family
		Clustering	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type

	Instance-level	GNNExplainer	Perturbation based
General Application	Temporal Graph Neural Network with attention mechanism for anomaly detection. Local explanations will be produced by GNNExplainer via perturbations of the input graph nodes features.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Anomaly Detection	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of unwanted situations

GM #2	ATGN + PGM-Explainer (Clustering)		
Primary Model	Attentive Temporal Graph Network	Learning Category	Model Category
		Unsupervised	Geometric Deep Learning
		Task	Algorithm Family
		Clustering	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	PGM-Explainer	Surrogate
General Application	Temporal Graph Neural Network with attention mechanism for anomaly detection. Local explanations will be provided by PGM-Explainer via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Anomaly Detection	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of unwanted situations

GM #3	STGCN + GNNExplainer (Regression)		
Primary Model	Spatio-Temporal Graph Convolutional Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Regression	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	GNNExplainer	Perturbation based

General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting (CNN-based). Local explanations will be produced by GNNExplainer via perturbations of the input graph nodes features.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Product Demand Forecasting	Previous and current product demand, web visits sequential data	Future product demand

GM #4	STGCN + Guided BP (Regression)		
Primary Model	Spatio-Temporal Graph Convolutional Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Regression	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	Guided BP	Gradient/Feature based
General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting (CNN-based). Local explanations for node predictions will be produced by Guided BP.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Product Demand Forecasting	Previous and current product demand, web visits sequential data	Future product demand

GM #5	STGCN + PGM-Explainer (Regression)		
Primary Model	Spatio-Temporal Graph Convolutional Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Regression	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	PGM-Explainer	Surrogate

General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting (CNN-based). Local explanations will be provided by PGM-Explainer via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Product Demand Forecasting	Previous and current product demand, web visits sequential data	Future product demand

GM #6	GCRN + GNNExplainer (Regression)		
Primary Model	Graph Convolutional Recurrent Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Regression	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	GNNExplainer	Perturbation based
General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting (RNN-based). Local explanations will be produced by GNNExplainer via perturbations of the input graph nodes features.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Product Demand Forecasting	Previous and current product demand, web visits sequential data	Future product demand

GM #7	GCRN + Guided BP (Regression)		
Primary Model	Graph Convolutional Recurrent Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Regression	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	Guided BP	Gradient/Feature based

General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting (RNN-based). Local explanations for node predictions will be produced by Guided BP.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Demand Forecasting	Previous and current product demand, web visits (sequential, multivariate)	Future product demand

GM #8	GCRN + PGM-Explainer (Regression)		
Primary Model	Graph Convolutional Recurrent Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Regression	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	PGM-Explainer	Surrogate
General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting (RNN-based). Local explanations will be provided by PGM-Explainer via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Demand Forecasting	Previous and current product demand, web visits (sequential, multivariate)	Future product demand

GM #9	TGN + GNNExplainer (Regression)		
Primary Model	Temporal Graph Networks	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	GNNExplainer	Perturbation based

General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting. Local explanations will be produced by GNNExplainer via perturbations of the input graph nodes features.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times Data quality parts	Prediction of final production at the end of the day
WHR	Demand Forecasting	Previous and current product demand, web visits sequential data	Future product demand

GM #10	TGN + PGM-Explainer (Regression)		
Primary Model	Temporal Graph Networks	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Regression	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	PGM-Explainer	Surrogate
General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting. Local explanations will be provided by PGM-Explainer via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times Data quality parts	Prediction of final production at the end of the day
WHR	Demand Forecasting	Previous and current product demand, web visits sequential data	Future product demand

GM #11	TGN + Guided BP (Regression)		
Primary Model	Temporal Graph Networks	Learning Category	Model Category
		Supervised	Geometric Deep Learning

		Task	Algorithm Family
		Regression	Spatio-temporal GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	Guided BP	Gradients/Features
General Application	A group of GNNs that capture the spatial and temporal dependencies of a graph nodes and can be used for future node values forecasting. Local explanations for node predictions will be produced by Guided BP.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Forecasting	Machine Status/Warnings/Faults Production Counters Cycle Times Data quality parts	Prediction of final production at the end of the day
WHR	Demand Forecasting	Previous and current product demand, web visits sequential data	Future product demand

Convolutional Graph Neural Networks

GM #12	GCN + ZORRO (Classification)		
Primary Model	Graph Convolutional Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Convolutional GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	ZORRO	Perturbation based
General Application	Graph Convolutional Network is based on an efficient variant of convolutional neural networks which operate directly on graphs. Local explanations will be produced by ZORRO via perturbations of the input graph nodes features.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Predictive maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Prediction on which electronic board will fail soon

GM #13	GCN + Graph-LIME (Classification)

Primary Model	Graph Convolutional Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Convolutional GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	Graph-LIME	Surrogate
General Application	Graph Convolutional Network is based on an efficient variant of convolutional neural networks which operate directly on graphs. Local explanations will be provided by Graph-LIME via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Predictive maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Prediction on which electronic board will fail soon

GM #14	DCNN + GNNExplainer (Classification)		
Primary Model	Diffusion- Convolutional Neural Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Convolutional GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	GNNExplainer	Perturbation based
General Application	Diffusion-based representations can be learned from graph structured data and used as an effective basis for node classification. Local explanations will be produced by GNNExplainer via perturbations of the input graph nodes features.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Predictive maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Prediction on which electronic board will fail soon

GM #15	DCNN + Graph-LIME (Classification)

Primary Model	Diffusion- Convolutional Neural Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Convolutional GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	Graph-LIME	Surrogate
General Application	Diffusion-based representations can be learned from graph structured data and used as an effective basis for node classification. Local explanations will be provided by Graph-LIME via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
CNH	Predictive maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Prediction on which electronic board will fail soon

GM #16	GCNN + Graph-LIME (Classification)		
Primary Model	Graph Convolutional Neural Network	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Convolutional GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	Graph-LIME	Surrogate
General Application	A group of GNNs that perform convolutions by information propagation considering node neighborhoods. Local explanations will be provided by Graph-LIME via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Point Cloud optimization	Point Cloud, Lateral density, Direction density, Exposure time	Optimal Point Cloud

GM #17	GCNN + GraphSVX (Classification)		
Primary Model		Learning Category	Model Category

	Graph Convolutional Neural Network	Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Convolutional GNNs
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	GraphSVX	Decomposition based
General Application	A group of GNNs that perform convolutions by information propagation considering node neighborhoods. Local explanations will be produced by GraphSVX, by decomposing the model's prediction into node and feature contributions using Shapley values.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Point Cloud optimization	Point cloud, Lateral density, Direction density, Exposure time	Optimal Point Cloud

Graph Attention Networks

GM #18	GAT + ZORRO (Classification)		
Primary Model	Graph Attention Networks	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Graph Attention Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	ZORRO	Perturbation based
General Application	GAT is a representative of spatial convolutional GNNs. Local explanations will be produced by ZORRO via perturbations of the input graph nodes features.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Anomaly Detection	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of unwanted situations
WHR	Understanding demand evolution / buying patterns	Past sales, promotions and web visits sequential data	Node importance, extraction of paths in the Graph
CNH	Anomaly Detection	Historical data of electronic boards faults correlated to the	Detection on which electronic board failed

		average panel temperature	

GM #19	GAT + PGM-Explainer (Classification)		
Primary Model	Graph Attention Networks	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Graph Attention Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	PGM-Explainer	Surrogate
General Application	GAT is a representative of spatial convolutional GNNs. Local explanations will be provided by PGM-Explainer via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Anomaly Detection	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of unwanted situations
WHR	Understanding demand evolution / buying patterns	Past sales, promotions and web visits (sequential data)	Node importance, extraction of paths in the Graph
CNH	Anomaly Detection	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

GM #20	GAT + Guided BP (Classification)		
Primary Model	Graph Attention Networks	Learning Category	Model Category
		Supervised	Geometric Deep Learning
		Task	Algorithm Family
		Classification	Graph Attention Networks
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	Guided BP	Gradients/Features

General Application	GAT is a representative of spatial convolutional GNNs. Local explanations for node predictions will be produced by Guided BP.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Anomaly Detection	Machine Status/Warnings/Faults Production Counters Cycle Times	Prediction of unwanted situations
WHR	Understanding demand evolution / buying patterns	Past sales, promotions and web visits (sequential data)	Node importance, extraction of paths in the Graph
CNH	Anomaly Detection	Historical data of electronic boards faults correlated to the average panel temperature	Detection on which electronic board failed

4.3.2.2 Graph Representation Learning

Graph Neural Networks & Autoencoders

GM #21	VGAE + GNNExplainer (Clustering)		
Primary Model	Variational Graph Autoencoder	Learning Category	Model Category
		Unsupervised	Graph Representation Learning
		Task	Algorithm Family
		Clustering	GNN & Auto-Encoders
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	GNNExplainer	Perturbation based
General Application	A framework for unsupervised learning on graph-structured data based on the Variational Auto-Encoder. Local explanations will be produced by GNNExplainer via perturbations of the input graph nodes features.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Planning	Generated constraints Demand - Datamart Availability parts Monthly expected production	Day / Month Production Planning
WHR	Demand Forecasting	Previous and current product demand,	Future product demand

		web visits sequential data	
CNH	Predictive Maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Prediction on which electronic board will fail soon

GM #22	VGAE + PGM-Explainer (Clustering)		
Primary Model	Variational Graph Autoencoder	Learning Category	Model Category
		Unsupervised	Graph Representation Learning
		Task	Algorithm Family
		Clustering	GNN & Auto-Encoders
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	PGM-Explainer	Surrogate
General Application	A framework for unsupervised learning on graph-structured data based on the Variational Auto-Encoder. Local explanations will be provided by PGM-Explainer via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
FORD	Production Planning	Generated constraints Demand - Datamart Availability parts Monthly expected production	Day / Month Production Planning
WHR	Demand Forecasting	Previous and current product demand, web visits sequential data	Future product demand
CNH	Predictive Maintenance	Historical data of electronic boards faults correlated to the average panel temperature	Prediction on which electronic board will fail soon

GM #23	DGI + Graph-LIME (Clustering)		
Primary Model	Deep Graph Infomax	Learning Category	Model Category
		Unsupervised	Graph Representation Learning
		Task	Algorithm Family
		Clustering	GNN & Auto-Encoders

Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	Graph-LIME	Surrogate
General Application	A general approach for learning node representations within graph-structured data in an unsupervised manner. Local explanations will be provided by Graph-LIME via surrogate model approximation.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Supporting promo initiatives	Past sales, promotions and web visits sequential data	Recommendations for promotional actions

Random Walk based embeddings

GM #24	Node2Vec + GNN-LRP (Clustering)		
Primary Model	Node2Vec	Learning Category	Model Category
		Unsupervised	Graph Representation Learning
		Task	Algorithm Family
		Clustering	Random Walk based embeddings
Explainability	Explainability Category	Explainability Tool	Explainability Type
	Instance-level	GNN-LRP	Decomposition
General Application	Random walk based Embedding methods are used to approximate some of the properties of a graph, including node centrality and similarity. Local explanations for node predictions will be produced by GNN-LRP.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
WHR	Understanding customer behaviors	Past sales, promotions and web visits sequential data	Clustering of customer behaviors

4.3.2.3 Graph Modelling

GM #25	Graphical Modelling / FCM (Classification)		
Primary Model	Graphical Modelling / FCM	Learning Category	Model Category
		Supervised	Graph Representations
		Task	Algorithm Family
		Classification	Graph modelling

Explainability	Explainability Category	Explainability Tool	Explainability Type
	Transparent	Self-explained	Self-explained
General Application	Fuzzy-graph structures for representing causal reasoning.		
Application in demonstrators			
Demonstrator	Manufacturing Scenario	Inputs	Output
UNIMETRIK	Measurement plan parameter optimization	Position, Direction, Lateral density, Direction density, Exposure time	Measurement accuracy and/or tolerance
			Optimal measurement parameters

4.4 Catalogue Overview

As Section 4.2 and 4.3 conform the consolidated Baseline Models Catalogue, the following charts analyse the proposed methods. In this catalogue it has been described a total of 73 models, that can be grouped according to three different perspectives. They can be classified according to their category, their algorithm family and their explainability type. Regarding the learning category and the target tasks, the whole catalogue is split into learning category, algorithm family and explainability category. We have grouped all models to have an overview of the whole catalogue including both Hybrid ML models and Graph ML models.

Although the explainability tools for hybrid ML proposed in the draft catalogue include mostly simplification and feature relevance tools, the use of other tools, such as explanations by example, has not been ignored. These types of tools will be considered during the experimentation in the different demonstrators in order to adapt to the data provided by the demonstrators, but at this point in the project, the tools that have been selected can a priori complement the proposed black box models without any problems.

Regarding the **"Learning Category"** we have grouped the models in **48 hybrid algorithms** and **25 graph algorithms.** The 73 algorithms can be further classified as follows:

- 63 supervised models. These supervised models are composed by 21 classification algorithms and 42 other algorithms for regressions tasks.
- 10 unsupervised models. All of them are employed for clustering purposes.

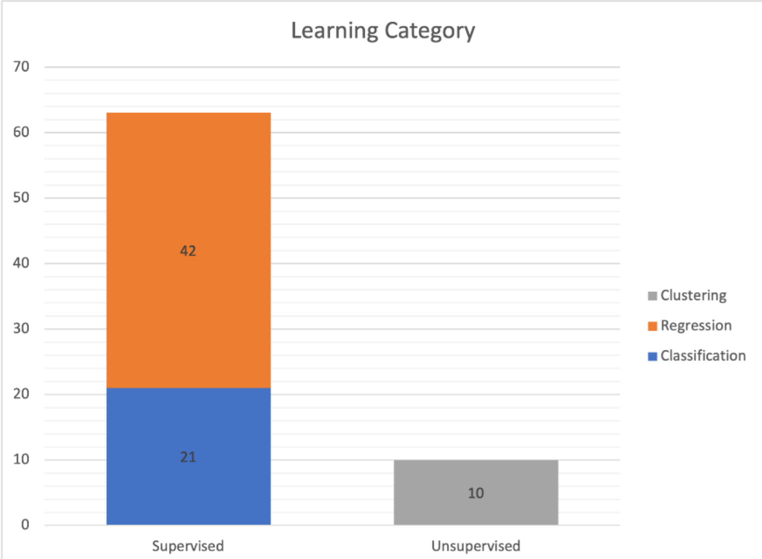


Figure 10 Overview of XAI Baseline Models from a learning category perspective.

Regarding the **"Algorithm Family"**, the cards are divided into:

- **31 Machine Learning models**. Several algorithms families constitute this category as follows: 3 Decision Trees, 8 SVMs, 8 XGBoosts, 8 Random Forests, 2 Linear Models and 2 Neighbouring algorithms.
- **17 Deep Learning models.** This group is composed by 6 DNNs, 3 CNNs, 3 Autoencoders and 5 RNNs.
- **20 Geometric Deep Learning**. This group is composed by 11 spatio-temporal GNNs, 3 Graph Attention Networks and 6 Convolutional GNNs.
- **4 Graph Representation Learning**. Composed by 3 Autoencoders and 1 Random Walk based embeddings.
- **1 Graph Representation**. Composed by 1 Graph Modelling.

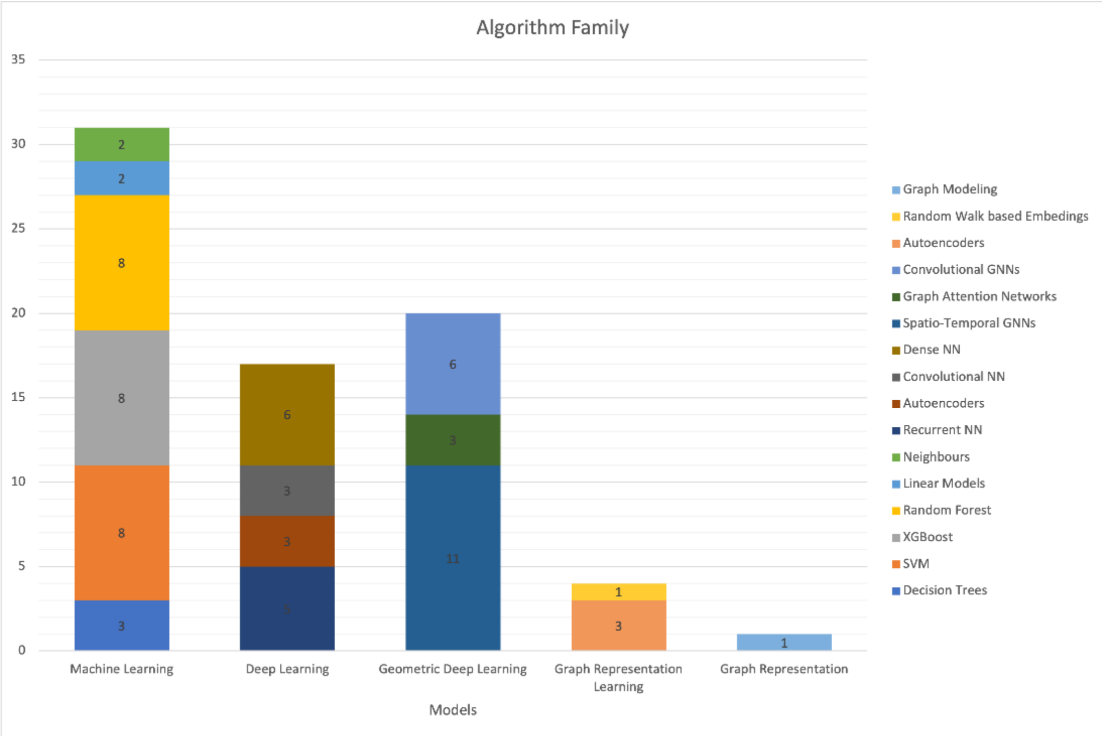


Figure 11 Overview of XAI Baseline Models from an algorithm family point of view.

Besides these two categories, the catalogue is analysed from an **"Explainability Category"** perspective:

- 41 black-box models are explained by via post-hoc techniques.
- 6 of the 73 models are transparent and do not need any explainability tools, as they are interpretable by design.
- 2 models will provide explainability through hybrid architectures.
- **24 instance-level** as graph explainability models.

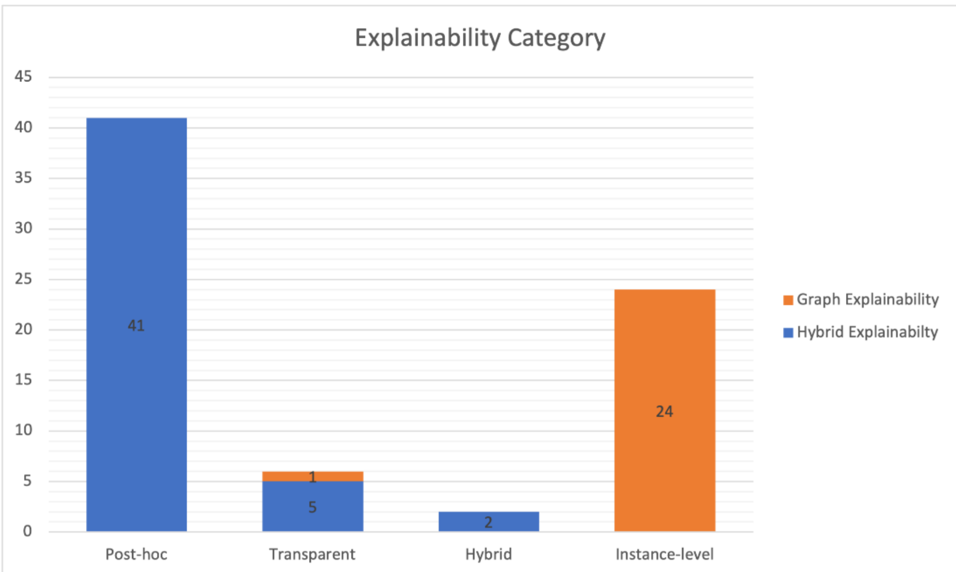


Figure 12 Overview of XAI Baseline Models from an explainability category perspective.

5 Conclusions and Next Steps

Deliverable 4.1 “Draft Catalogue of XMANAI AI and Graph Machine Learning Models” presents an overview of appropriate baseline models for the manufacturing domain and for the different pilot scenarios to be supported by the XMANAI solution, identifying their main structural components, inputs, outputs and potential application scenarios. The document is further elaborating and refining the design of graph machine intelligence algorithms and hybrid ML to be integrated in the XMANAI platform for knowledge extraction, business intelligence and analytics in the manufacturing domains tackled. This work will lead to a scenario where a subset of the models of the draft catalogue are easily integrated within the platform in order to be combined with the rest of the components by means of XAI pipelines.

Based on a general analysis of the manufacturing landscape, this document analyses the 4 manufacturing scenarios based on the demonstrator proposed scenarios. Four different manufacturing scenarios that solve the Use Cases necessities: *Production Optimization, Product demand forecasting, Process/Product quality optimization, Process optimization and Semi-Autonomous Planning*, have been defined. Section 2 focuses on the investigation of different methods that have been proposed for manufacturing, following each potential application scenario. Following the manufacturing insights, there are still insufficient efforts towards the penetration of XAI in industry, being a field yet to be explored, while the real needs of XAI systems were identified. Through the analysis on the sections, it has been shown how Explainable AI (XAI) provides important insights that enhance the interpretability and explainability of the ML applications and redefine the black-boxes approaches as grey-boxes. XAI techniques make the collaboration between humans and artificial intelligence more feasible, advance the human decision-making process and define an upper level of trust that is required for autonomous AI deployment.

The identification of the different manufacturing scenarios (section 2) that fits on the execution of the demonstrators (section 3) has been mandatory in order to establish and elaborate the Baseline Models Catalogue that suits their needs. In Section 3, an overview of the different demonstrators has been presented as well as an identification of the problems to be solved. This task and process aims to understand the real needs of the manufacturers' explainable systems.

Based on the manufacturing scenarios and demonstrators, a functional catalogue with selected baseline algorithms and models has been composed in order to populate the XMANAI Explainable AI Platform and support the needs of the potential manufacturing stakeholders. Following this approach, the versatility of the catalogue has been identified which is not limited to use specific methods for each scenario. Instead, it is possible to work with transversal methods using the same methods to solve different scenarios with common explainability needs.

The next steps related to the XMANAI Machine Learning models will be to deepen the training of the baseline models to cover the real manufacturing scenarios of the demonstrators using their own data sources. In addition to this, the provision of rich explanations not only for data scientists but also for business experts is a key aspect to be taken into account in the next steps of WP4 activities. The next deliverables of this Work Package will also address the integration of the models within the XMANAI platform in order to be trained and tested in the demonstrator use cases. Therefore, the integration process will be developed closely with the other WPs dealing with the implementation of the platform (2, 3 and 5).

References

Adhikari, P. P., Gururaja Rao, H. V. & Buderath, M., 2018. *Machine Learning based Data Driven Diagnostics & Prognostics Framework for Aircraft Predictive Maintenance.* s.l., s.n.

Amruthnath, N. & Gupta, T., 2018. *A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance.* s.l., s.n., pp. 355-361.

Ansari, F., Glawar, R. & Sihn, W., 2020. *Prescriptive Maintenance of CPPS by Integrating Multimodal Data with Dynamic Bayesian Networks.* s.l., Springer Berlin Heidelberg, pp. 1-8.

Arrieta, A. B. et al., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion,* Volume 58, pp. 82-115.

Baldassarre, F. & Azizpour, H., 2019. Explainability Techniques for Graph Convolutional Networks. *ArXiv,* Volume abs/1905.13686.

Benbarrad, T., Salhaoui, M., Kenitar, S. B. & Arioua, M., 2021. Intelligent Machine Vision Model for Defective Product Inspection Based on Machine Learning. *Journal of Sensor and Actuator Networks,* 10(1).

Bousdekis, A., Lepenioti, K., Apostolou, D. & Mentzas, G., 2019. Decision Making in Predictive Maintenance: Literature Review and Research Agenda for Industry 4.0. *IFAC-PapersOnLine,* 52(13), pp. 607-612.

Brito, L. C., Susto, G. A., Brito, J. N. & Duarte, M. A., 2022. An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery. *Mechanical Systems and Signal Processing,* Volume 163, p. 108105.

Burnap, A., Hauser, J. R. & Timoshenko, A., 2019. Design and Evaluation of Product Aesthetics: A Human-Machine Hybrid Approach.. *ArXiv,* Volume abs/1907.07786.

Cao, Q. et al., 2020. Combining chronicle mining and semantics for predictive maintenance in manufacturing processes. *Semantic Web,* Volume 11, pp. 927-948.

Cardoso, D. & Ferreira, L., 2021. Application of Predictive Maintenance Concepts Using Artificial Intelligence Tools. *Applied Sciences,* Volume 11.

Carvalho, T. P. et al., 2019. A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering,* Volume 137, p. 106024.

Chen, W. & Fuge, M., 2019. Synthesizing Designs With Inter-Part Dependencies Using Hierarchical Generative Adversarial Networks. *Journal of Mechanical Design,* 141(11).

Chen, Y. et al., 2021. Surface Defect Detection Methods for Industrial Products: A Review. *Applied Sciences,* 11(16).

Dalzochio, J. et al., 2020. Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. *Computers in Industry,* Volume 123, p. 103298.

Deshwal, S., Kumar, A. & Chhabra, D., 2020. *Exercising hybrid statistical tools GA-RSM, GA-ANN and GA-ANFIS to optimize FDM process parameters for tensile strength improvement.* [Online] Available at: https://doi.org/10.1016/j.cirpj.2020.05.009

Duval, A. & Malliaros, F. D., 2021. *GraphSVX: Shapley Value Explanations for Graph Neural Networks.* s.l., Springer International Publishing, pp. 302-318.

Fang, H. et al., 2021. LEFE-Net: A Lightweight Efficient Feature Extraction Network With Strong Robustness for Bearing Fault Diagnosis. *IEEE Transactions on Instrumentation and Measurement,* Volume 70, pp. 1-11.

Feizabadi, J., 2020. *Machine Learning demand forecasting and supply chain performance.* [Online] Available at: https://doi.org/10.1080/13675567.2020.1803246

Funke, T., Megha, K. & Avishek, A., 2021. Zorro: Valid, Sparse, and Stable Explanations in Graph Neural Networks. *ArXiv,* Volume abs/2105.08621.

Gahm, C. et al., 2022. Applying machine learning for the anticipation of complex nesting solutions in hierarchical production planning. *European Journal of Operational Research,* 296(3), pp. 819-836.

Garre Perez, A., Ruiz, M. & Hontoria, E., 2020. Application of Machine Learning to support production planning of food industry in the context of waste generation under uncertainty. *Operations Research Perspectives,* Volume 7, p. 100147.

Goncalves, J. N. C., Cortez, P., Carvalho, M. S. & Frazao, N. M., 2021. *A multivariate approach for multi-step demand forecasting in assembly industries: Empirical evidence from an automotive supply chain.* [Online]
Available at: https://doi.org/10.1016/j.dss.2020.113452

Gonzalez Rodriguez, G., Gonzalez-Cava, J. M. & Mendez Perez, J. A., 2020. An Intelligent decision support system for production planning based on machine learning. *Journal of Intelligent Manufacturing,* 31(5), pp. 1257-1273.

Guo, L. et al., 2017. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing,* Volume 240, pp. 98-109.

Haarman, M. et al., 2018. *Predictive Maintenance 4.0 - Beyond the Hype: PdM 4.0 delivers results,* s.l.: s.n.

Hooda, N., Ghohan, J. S., Gupta, R. & Kumar, R., 2021. Deposition angle prediction of Fused Deposition Modeling process using ensemble machine learning. *ISA Transactions,* Volume 116, pp. 121-128.

Huang, Q. et al., 2020. GraphLIME: Local Interpretable Model Explanations for Graph Neural Networks. *ArXiv ,* Volume abs/2001.06216.

Kallioras, N. A. & Lagaros, N. D., 2020. DzAIℕ: Deep learning based generative design. *Procedia Manufacturing,* Volume 44, pp. 591-598.

Kharfan, M., Chan, V. W. K. & Efendigil, T. F., 2021. A data-driven forecasting approach for newly launched seasonal products by leveraging machine-learning approaches. *Annals of Operations Research,* 303(1), pp. 159-174.

Konig, R., Johansson, U. & Niklasson, L., 2008. *G-REX: A Versatile Framework for Evolutionary Data Mining.* s.l., s.n., pp. 971-974.

Lauer, T. & Legner, S., 2019. Plain instability prediction by machine learning in master production planning. *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE),* pp. 703-708.

Launay, H., Willot, F., Ryckelynck, D. & Besson, J., 2021. Mechanical assessment of defects in welded joints: morphological classification and data augmentation. *Journal of Mathematics in Industry,* 11(1).

Liu, C., Tang, D., Zhu, H. & Nie, Q., 2021. A Novel Predictive Maintenance Method Based on Deep Adversarial Learning in the Intelligent Manufacturing System. *IEEE Access,* Volume 9, pp. 49557-49575.

Liu, K. et al., 2019. *Steel surface defect detection using GAN and one-class classifier.* s.l., IEEE, pp. 1-6.

Li, X., Makis, V., Zuo, H. & Cai, J., 2018. Optimal Bayesian control policy for gear shaft fault detection using hidden semi-Markov model. *Computers & Industrial Engineering,* Volume 119, pp. 21-35.

Li, Y., Zou, L., Jiang, L. & Zhou, X., 2019. Fault Diagnosis of Rotating Machinery Based on Combination of Deep Belief Network and One-dimensional Convolutional Neural Network. *IEEE Access,* Volume 7, pp. 165710-165723.

Luan, C., Cui, R., Sun, L. & Lin, Z., 2020. *A Siamese Network Utilizing Image Structural Differences For Cross-Category Defect Detection.* s.l., s.n., pp. 778-782.

Lundberg, S. M. & Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions. *ArXiv,* Volume abs/1705.07874.

Madhavi, S., Dowluru, S. & Venkatesh, M., 2017. Evaluation of Optimum Turning Process of Process Parameters Using DOE and PCA Taguchi Method. *Materials Today: Proceedings,* Volume 4, pp. 1937-1946.

Mei, S., Yang, H. & Yin, Z., 2018. An Unsupervised-Learning-Based Approach for Automated Defect Inspection on Textured Surfaces. *IEEE Transactions on Instrumentation and Measurement,* 67(6), pp. 1266-1277.

Morariu, C., Morariu, O., Raileanu, S. & Borangiu, T., 2020. Machine Learning for predictive scheduling and resource allocation in large scale manufacturing systems. *Computers in Industry,* Volume 120, p. 103244.

Nguyen, H. D., Thomassey, S. & Hamad, M., 2021. Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management. *International Journal of Information Management,* Volume 57, p. 102282.

Nguyen, K. T. P. & Medjaher, K., 2019. A new dynamic predictive maintenance framework using deep learning for failure prognostics. *Reliability Engineering & System Safety,* Volume 188, pp. 251-262.

Nobari, A. H., Rashad, M. F. & Ahmed, F., 2021. CreativeGAN: Editing Generative Adversarial Networks for Creative Design Synthesis. *ArXiv,* Volume abs/2103.06242.

Pech, M., Vrchota, J. & Bednář, J., 2021. Predictive Maintenance and Intelligent Sensors in Smart Factory: Review. *Sensors,* 21(4).

Pfrommer, J. et al., 2018. Optimisation of manufacturing process parameters using deep neural networks as surrogate models. *Procedia CIRP,* Volume 72, pp. 426-431.

Pope, P. E. et al., 2019. *Explainability Methods for Graph Convolutional Neural Networks.* s.l., s.n.

Qu, Z. et al., 2018. *PartsNet: A Unified Deep Network for Automotive Engine Precision Parts Defect Detection.* New York, NY, USA, Association for Computing Machinery, p. 594–599.

Rehse, J.-R., Mehdiyev, N. & Fettke, P., 2019. *Towards Explainable Process Predictions for Industry 4.0 in the DFKI-Smart-Lego-Factory.* [Online] Available at: https://doi.org/10.1007/s13218-019-00586-1

Ren, L., Sun, Y., Cui, J. & Zhang, L., 2018. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. *Journal of Manufacturing Systems,* Volume 48, pp. 71-77.

Ribeiro, M. T., Singh, S. & Guestrin, C., 2016. *Model-Agnostic Interpretability of Machine Learning.* s.l., s.n.

Rouniyar, A. & Shandilya, P., 2020. Optimization of process parameters in magnetic field assisted powder mixed EDM of aluminium 6061 alloy. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science,* Volume 235.

Rozanec, J. & Mladenic, D., 2021. *Semantic XAI for contextualized demand forecasting explanations.* s.l.:s.n.

Ruiz-Sarmiento, J.-R.et al., 2020. A predictive model for the maintenance of industrial machinery in the context of industry 4.0. *Engineering Applications of Artificial Intelligence,* Volume 87, p. 103289.

Schnake, T. et al., 2021. Higher-Order Explanations of Graph Neural Networks via Relevant Walks. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*

Schneckenreither, M., Haeussler, S. & Gerhold, C., 2020. Order Release Planning with Predictive Lead Times: A Machine Learning Approach. *International Journal of Production Research.*

Serradilla, O., Zugasti, E. & Zurutuza, U., 2020. Deep learning models for predictive maintenance: a survey, comparison, challenges and prospect. *ArXiv.*

Simonyan, K., Vedaldi, A. & Zisserman, A., 2014. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps.. *CoRR,* Volume abs/1312.6034.

Smirnov, P. & Sudakov, V., 2021. *Forecasting new product demand using machine learning.* [Online] Available at: doi:10.1088/1742-6596/1925/1/012033

Springenberg, J. T., Dosovitskiy, A., Brox, T. & Riedmiller, M. A., 2015. Striving for Simplicity: The All Convolutional Net.. *CoRR,* Volume abs/1412.6806.

Syafrudin, M., Alfian, G., Fitriyani, N. L. & Rhee, J., 2018. Performance Analysis of IoT-Based Sensor, Big Data Processing, and Machine Learning Model for Real-Time Monitoring System in Automotive Manufacturing. *Sensors,* 18(9).

Tercan, H., Deibert, P. & Meisen, T., 2022. Continual learning of neural networks for quality prediction in production using memory aware synapses and weight transfer. *Journal of Intelligent Manufacturing,* 33(1), pp. 283-292.

Usuga Cavadid, J. P. et al., 2020. Machine Learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. *Journal of Intelligence Manufacturing,* 31(6), pp. 1531-1558.

Vu, M. N. & Thai, M. T., 2020. *PGM-Explainer: Probabilistic Graphical Model Explanations for Graph Neural Networks.* Vancouver, Canada, s.n.

Wang, J. et al., 2018. Big data driven cycle time parallel prediction for production planning in wafer manufacturing. *Enterprise Information Systems,* 12(6), pp. 714-732.

Wang, Y., Du, W., Wang, H. & Zhao, Y., 2021. Intelligent Generation Method of Innovative Structures Based on Topology Optimization and Deep Learning. *Materials,* 14(24).

Wang, Y. et al., 2020. A smart surface inspection system using faster R-CNN in cloud-edge computing environment. *Advanced Engineering Informatics,* Volume 43, p. 101037.

Xu, Y., Sun, Y., Liu, X. & Zheng, Y., 2019. A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning. *IEEE Access,* Volume 7, pp. 19990-19999.

Ying, Z. et al., 2019. *GNNExplainer: Generating Explanations for Graph Neural Networks.* s.l., Curran Associates, Inc..

Yuan, H., Yu, H., Gui, S. & Ji, S., 2021. Explainability in Graph Neural Networks: A Taxonomic Survey. *arXiv.*

Zhang, H. et al., 2017. Optimization of process parameters for minimum energy consumption based on cutting specific energy consumption. *Journal of Cleaner Production,* Volume 166.

Zhou, C. & Tham, C.-K., 2018. *GraphEL: A Graph-Based Ensemble Learning Method for Distributed Diagnostics and Prognostics in the Industrial Internet of Things.* s.l., s.n., pp. 903-909.

Zhu, J. & Wu, P., 2021. A Common Approach to Geo-Referencing Building Models in Industry Foundation Classes for BIM/GIS Integration. *ISPRS International Journal of Geo-Information,* Volume 10, p. 362.

Zhu, X., Ninh, A., Zhao, H. & Liu, Z., 2021. *Demand Forecasting with Supply-Chain Information and Machine Learning.* [Online]
Available at: 10.1111/poms.13426

Zilke, J. R., Mencía, E. L. & Janssen, F., 2016. *DeepRED - Rule Extraction from Deep Neural Networks..* s.l., Springer International Publishing.

Zonta, T. et al., 2020. Predictive maintenance in the Industry 4.0: A systematic literature review. *Computers & Industrial Engineering,* Volume 150, p. 106889.

List of Acronyms/Abbreviations

Acronym/ Abbreviation	Description
AE	Auto-Encoder
AI	Artificial Intelligence
ANN	Artificial Neural Network
ANOVA	Analysis Of Variance
ARIMA	Auto-Regressive Integrated Moving Average (univariate)
ARIMAX	Auto-Regressive Integrated Moving Average (multivariate)
BP	Back-Propagation
B2B	Business-to-Business
CAM	Class Activation Mapping
CBLOF	Clustering Based Local Outlier Factor
CBM	Condition-Based Maintenance
CLSC	Closed-Loop Supply Chain
CMS	Condition Monitoring System
CNN	Convolutional Neural Network
CT	Cycle Time
DAG	Directed Acyclic Graph
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DCNN	Diffusion- Convolutional Neural Network
DFE	Demand Forecast Effectiveness
DGI	Deep Graph Infomax
DL	Deep Learning
DNN	Deep Neural Network
DP-RBFN	Density Peak based Radial Basis Function Network
DT	Decision Tree
D2C	Direct-to-Consumer
EDM	Electrical Discharge Machining
EWR	Electrode Wear Rate
FastABOD	Fast-Angle-Based Outlier Detector
FCM	Fuzzy Cognitive Map

FE	**Finite Element**
FDM	**Fused Deposition Modelling**
FIS	**Fuzzy Inference System**
FTT	**First Time Through**
GA	**Genetic Algorithm**
GAN	**Generative Adversarial Network**
GAT	**Graph Attention Networks**
GBM	**Gradient Boosting Machine**
GCN	**Graph Convolutional Network**
GCNN	**Graph Convolutional Neural Network**
GCRN	**Graph Convolutional Recurrent Network**
GNN	**Graph Neural Network**
Grad-CAM	**Gradient-Weighted Class Activation Mapping**
HBOD	**Histogram-Based Outlier Detection**
HSIC	**Hilbert-Schmidt Independence Criterion**
IF	**Isolation Forest**
IHL	**Industrial Hospital Laundry**
kNN	**k-Nearest Neighbours**
LDOA	**Lightweight Detector Of Anomalies**
LOF	**Local Outlier Factor**
LRP	**Layer-wise Relevance Propagation**
LSTM	**Long Short-Term Memory**
MAD	**Mean Absolute Deviation**
MCD	**Minimum Covariance Determinant**
MER	**Material Erosion Rate**
MFAPM-EDM	**Magnetic Field Assisted Powder Mixed Electrical Discharge Machining**
ML	**Machine Learning**
MLP	**Multi-Layer Perceptron**
MSE	**Mean Squared Error**
NC	**Numerical Control**
NN	**Neural Network**
OCSVM	**One-Class Support Vector Machine**
ODP	**Operational Demand Plan**
PCA	**Principal Component Analysis**

PGM	**Probabilistic Graphical Modelling**
PdM	**Predictive Maintenance**
PHM	**Prognosis and Health Monitoring**
PoF	**Physics-of-Failure**
PR-AUC	**Precision-Recall Area Under the Curve**
RF	**Random Forest**
R2F	**Run-To-Failure**
RMSE	**Root Mean Squared Error**
RNN	**Recurrent Neural Network**
RUL	**Remaining Useful Life**
SA	**Sensitivity Analysis**
SD	**Standard Deviation**
SHAP	**Shapley Additive Explanations**
SKU	**Stock-Keeping Unit**
SSD	**Single Shot multi-box Detector**
SSIM	**Structural Similarity Index**
SVC	**Support Vector Classifier**
SVM	**Support Vector Machine**
SVR	**Support Vector Regressor**
SWFS	**Semiconductor Wafer Fabrication System**
TGN	**Temporal Graph Networks**
TLBO	**Teacher-Learning-Based optimization**
VGAE	**Variational Graph Auto-Encoder**
WMAPE	**Weighted Mean Absolute Percentage Error**
WMPE	**Weighted Mean Percentage Error**
XAI	**eXplainable Artificial Intelligence**
XGBoost	**eXtreme Gradient Boosting**