Information Fusion 81 (2022) 91–102

Contents lists available at ScienceDirect

Information Fusion

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

Knowledge graph-based rich and confidentiality preserving Explainable Artificial Intelligence (XAI)

Jože M. Rožanec [a,b,c,*], Blaž Fortuna [a,b], Dunja Mladenić [a]

[a] Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
[b] Qlector d.o.o., Rovšnikova 7, 1000 Ljubljana, Slovenia
[c] Jožef Stefan International Postgraduate School, Jamova 39, 1000 Ljubljana, Slovenia

ARTICLE INFO

Keywords:
Explainable Artificial Intelligence
Knowledge Graph
Demand forecasting
Smart manufacturing
Confidentiality
Privacy

ABSTRACT

The paper proposes a novel architecture for explainable artificial intelligence based on semantic technologies and artificial intelligence. We tailor the architecture for the domain of demand forecasting and validate it on a real-world case study. The explanations provided result from knowledge fusion regarding concepts describing features relevant to a particular forecast, related media events, and metadata regarding external datasets of interest. The Knowledge Graph enhances the quality of explanations by informing concepts at a higher abstraction level rather than specific features. By doing so, explanations avoid exposing sensitive details regarding the demand forecasting models, thus preserving confidentiality. In addition, the Knowledge Graph enables linking domain knowledge, forecasted values, and forecast explanations while also providing insights into actionable aspects on which users can take action. The ontology and dataset we developed for this use case are publicly available for further research.

1. Introduction

The increasing digitalization of manufacturing in the context of Industry 4.0 provides a growing amount of data describing assets and operations. This data increases the transparency of production processes, accelerates the information flow through the company, and is a valuable asset to build predictive models. Demand forecasting is an essential component for any manufacturing company. A growing body of research has addressed techniques that identify demand-type and develop statistical and Artificial Intelligence (AI) models to predict future demand. In critical operations, such as demand forecasting, it is crucial to provide a forecast, the associated uncertainty and convey some explanations to the user. In this way, the user can make an informed decision, which may have far-reaching consequences, and avoid costly mistakes. Such explanations enhance the AI models, increase trust in the system, and help identify errors and performance issues.

Researchers developed several approaches related to AI models' explainability, dividing the models into glass-box (the AI algorithm can explain its prediction) and black-box (we need another algorithm to obtain an explanation of a model). Such explanations can be provided on a global level (the forecasting model in general) or at a local level (for every prediction instance). Some authors also envisioned that the inclusion of semantic technologies could be used to determine the semantic closeness of concepts encoded in data features or integrate knowledge graph embeddings to a forecasting model to produce explanations along with the forecasts [1,2].

Explanation of an AI model should provide cues on factors that influenced a forecast, provide means to discern relevant context, supporting information, and emphasize actionable aspects to assist decision-making. Our research focuses on local-level explanations so that planners can still weigh cues provided for each case. We develop a domain-specific ontology and knowledge graph to provide domain knowledge and context for creating explanations. The knowledge graph allows us to identify which cues correspond to semantically close concepts and variables people may influence, supporting information from the dataset and external systems. In particular, we integrate a Media Event Retrieval System to match provided cues to relevant news and identify other concepts that may be relevant to the case. The system then proposes to the user such concepts and potential datasets that could be used to enrich the existing data.

The main scientific contributions of the presented research are the following:

1. a modular architecture to provide demand forecasting explanations, which consider semantic proximity, highlight factors that can be influenced by the user, relevant context, and opportunities for data enrichment;

* Corresponding author at: Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia.
E-mail addresses: joze.rozanec@ijs.si (J.M. Rožanec), blaz.fortuna@ijs.si (B. Fortuna), dunja.mladenic@ijs.si (D. Mladenić).

https://doi.org/10.1016/j.inffus.2021.11.015
Received 15 December 2020; Received in revised form 12 November 2021; Accepted 15 November 2021
Available online 29 November 2021

2. a unified ontology to capture domain knowledge regarding demand forecasting, events, and other elements relevant to the explanations;
3. the architecture implementation and validation on a real-world use case in manufacturing;
4. publicly available dataset containing: features relevant to forecasts, retrieved media events, metadata regarding external datasets, and a subset of annotated media events and external datasets metadata. Such a dataset can be a valuable resource for further research on identifying, ranking relevant, and summarizing entries to explain forecasts better.

To evaluate the overall architecture, we manually annotate a set of 528 media events, 401 datasets metadata entries, and 168 forecast explanations and assess their goodness based on metrics and criteria described in Section 6.

The rest of this paper is structured as follows: Section 2 presents related work, Section 3 describes the proposed architecture in the context of demand forecasting, Section 4 describes the ontology we used to construct the knowledge graph, Section 5 describes the use case for which we implemented the architecture, Section 6 provides the results we obtained, and Section 7 details some limitations and improvement opportunities regarding the work described in this paper. Finally, in Section 8, we provide our conclusions and outline future work ideas.

2. Related work

2.1. Demand forecasting

Demand forecasting is a crucial component for any manufacturing company. Accurate forecasts translate to production planning, influencing raw materials purchases, production line schedules, and inventory. It is thus essential that planners, who create the demand forecasts, understand the demand of the products they sell. The intrinsic properties of the product influence such demand. Among other factors affecting the demand, we must mention the market they target, the economic context, and customer expectations [3].

It has been recognized that AI models can play a pivotal role in enhancing demand forecasting [4–6]. To develop such models, it is critical to gather data from a wide range of sources that enable the model to learn patterns and correlations latent in the data. The increasing digitalization of industry, fostered by the decreased cost of sensors and connectivity [7], and the introduction of new technologies and paradigms (e.g., Industrial Internet of Things, Cloud Computing, Cyber–Physical Systems, or Digital Twins [8–11]), have produced an explosive growth of data in the manufacturing domain [12], which can be leveraged to develop such models. Most authors, thus, take advantage of enriching demand data with information regarding exogenous factors that can influence demand, achieving more accurate predictions. In particular, authors addressing demand forecasting in the automotive industry, reported using data regarding the unemployment rate [13], people's personal income [14], the inflation rate [15], fuel prices [16], or the Gross Domestic Product [17], among others.

Authors have observed that demand can adopt different patterns. Furthermore, they have developed criteria to categorize demand into specific demand types. For example, Williams [18] introduced the concept of demand variance during lead times, Johnston and Boylan [19] described the average demand interval (relationship between total time buckets and those with demand occurrence), and Syntetos et al. [20] introduced the concept of coefficient of variation (extent of demand size variability to the mean). Studies conducted by multiple researchers confirmed that demand types relate to demand distribution shapes [21,22].

While demand forecasting can be conceived as a time series forecasting problem, it can also be framed as a supervised regression learning problem [23]. Doing so enables learning from past demand while considering additional information regarding exogenous factors or demand shapes described above. Among frequently used machine learning models for demand forecasting in the automotive industry, we find the Multiple Linear Regression [13,24], Support Vector Machine [13,25], and the Artificial Neural Networks [25–27].

It has been observed that AI models can outperform human forecasting. While human forecasts are subject to a wide range of biases, AI models are not prone to such biases and provide better means towards identifying systematic patterns and integrating a wide range of sources of information related to forecasting processes [28–31]. Nevertheless, the planners are responsible for making judgmental adjustments and record such adjustments when vital information about forthcoming events is available to them, but not to the forecasting models [28]. Furthermore, they are accountable for decisions made based on such forecasts.

Demand forecasts with their associated uncertainty do not provide enough ground for planners' decision-making. While authors consider that justifying decisions is a precondition for responsible decision-making [32], and such justifications will be considered ever more critical in the future [33], it has been shown that explanations enhance users' acceptance and satisfaction, and provide means to understand if causality is captured by the model [32]. Furthermore, model explanations have been found to influence human decision-making [34]. Therefore, providing such explanations to planners helps them understand the reasons driving such a forecast (e.g., features' contribution to the forecast [35]). Furthermore, enriching such explanations with domain context [32] enables the planners to evaluate the forecasts' soundness, make sure they agree on it, and eventually correct the forecast [36] before making a decision [37]. Such explanations can be achieved through different means, studied by the subfield of AI known as Explainable Artificial Intelligence (XAI).

2.2. XAI: black-box models and semantic technologies

Human-understandable explanations in the demand forecasting domain are essential as they help the planner to (i) understand which of the model attributes are strong drivers in determining the expected demand, (ii) identify some bias in the model, and (iii) understand how the model can be improved.

Several taxonomies have been proposed to classify the XAI techniques and ease their study. Adadi and Berrada [38] proposed classifying the XAI methods into three categories, based on the complexity of the method (the degree up to which the method is interpretable), their scope (local or global), and whether the methods are specific to a model or agnostic. Inherently interpretable models are known as *white-box models*, while opaque models are known as *black-box models* [39]. A similar classification was developed by [40], who divided XAI approaches into three categories: scope (whether the explanation targets a single forecast or attempts to explain the whole model), methodology (if focused on the input data or model parameters), and usage (if integrated to the model or applied to any model in general). Liao et al. [41] divided XAI techniques into four categories, whether they explain the model (global), explain the forecast (local), inspect the counterfactual, or provide an example. Finally, Angelov et al. [42] proposed a wider taxonomy, taking into consideration whether the explanation is local or not if the models are transparent or opaque, if the techniques are model-specific or model-agnostic, and whether explanations are created by simplification, conveyed through visualizations, or based on feature relevance. In this research, we focus on black-box explanation methods, since (i) white-box models, due to their complexity, are not intelligible to most users who are not AI experts [37], and given that (ii) black-box models can be applied to any kind of AI models [35]. By doing so, we decouple the forecast and explainability dimensions and provide greater flexibility in choosing the best approach for each of them.

Most of the XAI approaches for time series were developed for deep learning models [43]. Those approaches are conditioned by the deep learning model's architecture, which determines how they extract features from data and learn, thus requiring different ways to convey such information to the users. One such method is the *Gradient*Input*, applied to convolutional networks, which computes the partial derivative of a layer against the input, and multiplies the result by the input, to obtain an insight into how do the filters activate [44]. ConvTimeNet [45] follows another approach occluding time series sub-segments to measure how the forecast is affected when compared against one produced by the original sequence, and thus understand the contribution of each time series' sub-sequence to the forecast. Tonekaboni et al. [46] leveraged the same concept. However, instead of occluding time series sub-segments, he replaced them with synthetically generated ones. In the same line, Ozyegen et al. [47] suggests replacing time series' sub-segments with a local or global mean, while Mercier et al. [48] developed a novel loss to generate such patches using a patch generative network. In recurrent neural networks, insights provided by an attention mechanism are considered a source of explainability since they provide information regarding the temporal dynamics learned by the model [49].

Time series forecasting can also be framed as a supervised learning problem [23]. Doing so enables using a different set of explainability techniques appropriate for supervised learning models. Among most frequently cited posthoc model-agnostic and feature-relevance technique for local explanations [40,42] we find LIME [50], and its variants (e.g.: k-LIME [51], DLIME [52], and LIMEtree [53]), Anchors [54], Local Foil Trees [55], or LoRE [56]. These approaches build surrogate models for each prediction sample, learning the reference model's behavior on the particular case of interest by introducing perturbations to the feature vector variables. By doing so, they can provide a local feature importance estimate, which is considered an indirect method to explain a model [32]. In particular, we opted for the feature relevance explanation above other post-hoc model agnostic techniques (e.g., rule extraction) due to the simplicity of implementation. While other explanations can also be built considering the ontology abstraction, care must be put to ensure these remain valid. E.g., in the case of rule extraction, the non-contradiction principle must be enforced since while the rule can be valid at a feature level, the same is not guaranteed at higher abstraction levels. While such approaches do not consider the temporal dimension *per se*, features' metadata can be considered to capture it and enrich the forecast explanations [57].

When designing a system providing explanations for AI models, it must be considered the explanations must serve multiple stakeholders (and thus target different user profiles [58]), serve different purposes, and that their effectiveness must be assessed quantitatively (through algorithmic validation) and qualitatively (estimating user satisfaction, trust, and other factors) [59]. Good explanations should convey meaningful information, resemble a logic explanation [60], focus on actionability [61], and if possible, provide some counterfactuals. Since different explanations on the reasons behind a models' forecast can be deduced depending on the users' background knowledge, the system must provide enough information to the user so that the user can unequivocally understand the models' rationale and enable users' responsible decision-making [62]. Given the recipient-dependent nature of the explanations, it is important that those are designed following a human-centered approach [41,63]. When focusing on a specific user profile, it is important to adapt the language and content of relevant information that should or can be shown to it [64]. The actionable dimension relates to the ability to distinguish mutable features that can be influenced by the user [65,66].

Multiple authors envisioned the usage of semantic technologies in the explainability domain. Doctor XAI [2] develops an agnostic XAI technique for ontology-linked data classification by training a surrogate model and extracting rules from it. Gaur et al. [1] make use of a Knowledge Graph to feed deep learning models with it to enhance their explainability. Samek and Müller [58] envision Knowledge Graphs can be used to compact large tree models by combining nodes into unique probabilistic concepts. In addition to the opportunities mentioned above, we consider semantic technologies shall (A) provide background knowledge which can be leveraged to provide semantic meaning to dataset features, (B) inform their characteristics (e.g., valid ranges, if they are mutable or immutable if the user can influence their values), or (C) issue explanations with adequate language and context, considering if certain information shall be or not informed to the user.

A modular architecture that decouples forecasts from their explanations is required to provide the planners with adequate explanations regarding machine demand forecasts. Such architecture has the advantage of allowing to iterate both components at their own pace. Black-box explainability models mostly lack domain knowledge and inform relevant forecasting features without a broader context and a more profound interpretation. A third module can provide such context and interpretation, taking into account black-box explanations and domain-knowledge encoded in an ontology and instantiated in a Knowledge Graph to create a better explanation for the end-user.

2.3. Domain specific ontologies

Demand Forecasting is an essential activity in manufacturing, and manufacturing domain ontologies provide some related concepts. E.g., Ameri and Dutta [67] developed the Manufacturing Service Description Language ontology as a formal representation of manufacturing services, and Lemaignan et al. [68] described knowledge regarding individual operations in MASON.

Since our architecture links concepts from datasets used to train the AI models that issue the forecasts, this should also be encoded in the unified ontology. To promote the reuse of concepts, following the MIREOT principle [69], we analyzed ontologies related to the artificial intelligence domain. Cannataro and Comito [70] developed the DAMON (Data Mining Ontology for Grid Programming), which provides a reference model for data mining tasks, methodologies, and available software. A heavyweight ontology was developed by Panov et al. [71,72], who provide means to represent data mining entities, inductive queries, and data mining scenarios. Diamantini et al. [73] developed KDDONTO, focusing on data mining algorithms discovery.

The explanation must provide information regarding features that are most influential to a forecast. This view shall be complemented with context, describing events that illustrate possible reasons behind the values captured in the features and point to similar events observed in the present. Such events shall help the planner draw conclusions that exceed the mere forecast provided. The concept of *Event* can be found in different ontologies [74,75], having a prominent place in the YAMATO ontology [76] as one of the core concepts.

The ontologies mentioned above provide a solid ground to create a unified ontology that serves the purpose of semantically enhancing explanations of demand forecasts issued by AI models.

2.4. Methodologies for ontology design in manufacturing

Multiple general methodologies were proposed to build an ontology. Uschold [77],Uschold and King [78],Fernández-López et al. [79] suggest first identifying the purpose of the ontology, and Uschold [77],Kim et al. [80] stress the importance of defining its scope and level of formality. Next, Kim et al. [80] suggests defining a problem statement and competency questions, which provide common ground to elicit knowledge from multiple sources, identify key concepts and relationships. When defining terms to be used in the ontology, an effort should be made to integrate pre-existing ontologies where possible.

Some authors identified steps specific to the creation of manufacturing domain ontologies. Ameri et al. [81] developed a four-step methodology using a Simple Knowledge Organization System framework to develop a thesaurus of concepts, and then identify relevant classes and

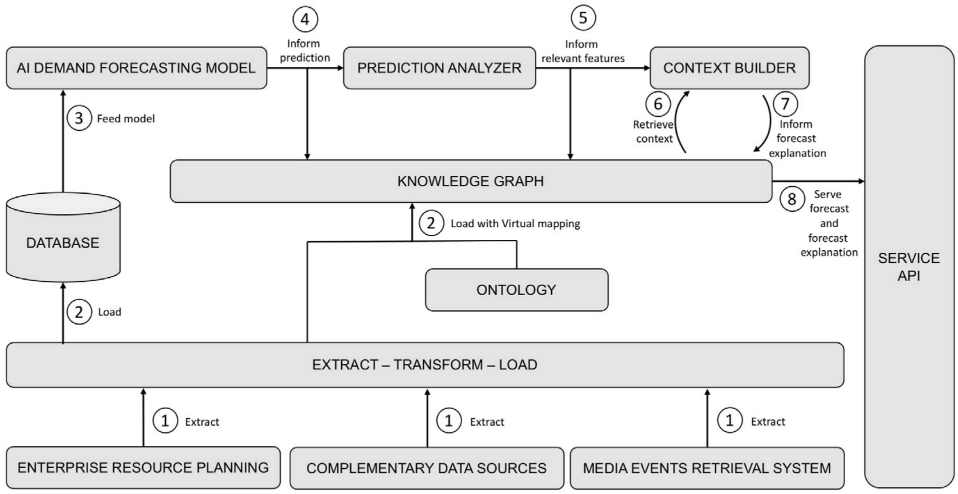

Fig. 1. Semantic XAI architecture for demand forecasting.

provide logical constraints and rules. A similar approach was developed by Chang et al. [82], with an emphasis on manufacturing design.

The methodologies we mentioned above provide us valuable guidance when developing our domain-specific ontology, aiming to integrate domain-specific knowledge regarding demand forecasting with the demand forecasting models, predictions, media events, external datasets, and explanations.

3. The proposed architecture

It has been noted in the literature that AI explanations can be enhanced by complementing insights on models' inner workings with domain knowledge [83,84]. Furthermore, the XAI's ability to explain models can be used to enhance data fusion capabilities, discovering relevant data sources that further enhance the model and deepen the users' domain understanding [32]. However, we found no author addressed the intersection between data fusion, confidentiality, and model explainability from the literature we reviewed. We thus propose a modular architecture to provide forecast explanations that address the challenges mentioned above. The main novelty is combining semantic technologies and media events to build the context and provide an informed prediction explanation while also recommending additional data sources that could be used to enrich the existing dataset.

3.1. Architecture overview

The proposed supporting architecture enables us to realize the explanations for each prediction. The architecture integrates predictions, gathers insights on relevant features, incorporates domain knowledge and context to each prediction, and provides a forecast explanation to the end-user. The architecture (see Fig. 1) comprises the following components:

- **Enterprise Resource Planning**: software used by the manufacturing company to keep track of operations at different levels. For the specific use case of demand forecasting, it provides data regarding strategic sales planning, past and open sales, buyers risk assessment, and past demand.
- **Complementary data sources**: sources of data that provide valuable complementary data to train demand forecasting models.
- **Media Event Retrieval System**: is a system that keeps track of events reported in the media. Media Event Retrieval System uses Natural Language Processing to identify entities and concepts reported in media news and identify which news report about the same events. Wikification is applied to identify concepts, cluster news, and link events.
- **Ontology**: encodes domain knowledge regarding demand forecasting, demand forecasting models, events retrieved from the Media Event Retrieval System, and available datasets. The Extract-Transform-Load module uses it to guide a virtual mapping procedure and instantiate the Knowledge Graph.
- **Knowledge Graph**: is the instantiation of the ontology, based on data provided by the Extract-Transform-Load module and outcomes informed by multiple architecture components. The Knowledge Graph keeps track of data up to a configurable time window.
- **Extract-Transform-Load**: provides means to interface with the Enterprise Resource Planning and Media Events Retrieval System to retrieve relevant data, transform it as required, and ingest it to a database or into the Knowledge Graph. A virtual mapping procedure specifies how data pieces are related to concepts defined in the ontology and how they should be instantiated to the Knowledge Graph. The Extract-Transform-Load module consists of a series of batch processes. The processes are executed regularly to ensure that the KG and database information is updated regarding the Enterprise Resource Planning software and Media Event Retrieval System.
- **Database**: stores data relevant to the AI models, which can be used to train them.
- **AI Demand Forecasting model**: such a model can be built with a variety of algorithms, depending on the demand type and data available. Such models can be machine learning regression [13,85], deep learning [4,6] or probabilistic models [86]. Demand forecasting models shall provide a demand estimate for future points in time together with some uncertainty estimation. This information is stored in the Knowledge Graph.
- **Prediction Analyzer**: interfaces with the **AI Demand Forecasting model** to determine which features are most important to a given prediction. To that end, it may use different black-box explainability techniques to assess relevant features under different criteria. The analysis outcomes are stored in the Knowledge Graph.
- **Context Builder**: has the responsibility to gather required pieces of data to issue good explanations. Interfaces with the Prediction Analyzer to get relevant features to each prediction. With the Knowledge Graph to obtain context regarding feature values, related events reported in the media, and complementary data that may be used to enhance the model. The Knowledge Graph also provides information regarding high-level concepts that describe the features (see Fig. 3). These are used to inform relevant features at a higher abstraction level, avoid exposing model features, and enable more concise explanations.

Table 1
Data samples retrieved from the Media Event Retrieval System (MERS), EU Open Data Portal, and the Enterprise Resource Planning (ERP) software. While the actual entries contain more fields and data, we summarized them to the most relevant fields. While the names of the ERP fields are accurate, the row of data we display does not correspond to a genuine entry.

Data source	Data type	Data sample
MERS	JSON	{'articleCounts': {'eng': 15}, 'categories': [{'label': 'news/Business', 'uri': 'news/Business', 'wgt': 92}, …], 'concepts': [{'label': {'eng': 'Car finance'}, 'score': 100, 'type': 'wiki', 'uri': 'http://en.wikipedia.org/wiki/Car_finance'}, …], 'eventDate': '2019-10-01', 'location': None, 'relevance': 4, 'sentiment': 0.06666666666666665, 'summary': {'eng': 'About a third of auto loans for new vehicles taken in the first half of 2019 ...'}, 'title': {'eng': "The Seven-Year Auto Loan: America's Middle Class Can't Afford Their Cars'}, 'totalArticleCount': 15, 'uri': 'eng-5135186', 'wgt': 307584000}
EU Open Data Portal	JSON	{'result':{'results':[{'dataset':{ 'identifier_dcterms':[{'datatype': None, 'lang': None, 'type': 'literal', 'value_or_uri': 'http://data.europa.eu/89h/jrc-eplca-744e255f-3b81-4eed-b2d3-7794353e0efb'}], 'identifier_dcterms':[{'datatype': None, 'lang': 'en', 'type': 'literal', 'value_or_uri': 'Raw data for polymerization and intermediate products ...'}], 'distribution_dcat':[{'accessURL_dcat': [{'uri': 'http://eplca.jrc.ec.europa.eu/ELCD3/'}], 'format_dcterms': {'uri': 'http://publications.europa.eu/resource/authority/file-type/ZIP'}, 'license_dcterms': [{'uri': 'http://publications.europa.eu/resource/authority/licence/OP_DATPRO'}], ...}] }}]}}
ERP	CSV	VBELN,POSNR,MATNR,LFIMG,WADAT_IST 0000000000,000010,000000000000001111,100.000,20190909

- **Service API**: serves predictions and their explanations to the end-users. It hides explanation details based on the end-user profile to expose only facts that are relevant to the user.

Among the modules described above, three are data sources from which data is extracted. In Table 1 we provide data samples that correspond to each of them.

The proposed architecture considers several sources of information to build the explanations (e.g., features relevance, their mapping to concept abstractions provided by a local ontology, relevant media events, and suggestions regarding potentially beneficial datasets to be used for further data enrichment). We consider that a similar integration could be followed to extend the information sources to different semantic databases and federated ontologies. While multiple link discovery and ontology matching approaches have been studied in the literature [87–90], we consider at least two complementary ways that could be applied in our case to execute such integration. First, external semantic databases and federated ontologies could be integrated based on a common ground established through wikified concepts, wikifying the keywords associated with each feature [91,92]. An alternative could be to consider embeddings semantic meanings and integrate the sources based on concepts' embedding proximity. The embeddings approach could be used either (a) to create mappings from the model features to external ontologies and semantic databases using text embeddings [93–95], or (b) to automatically align instances from the local ontology to the semantic databases and federated ontologies using text or graph embeddings [96,97].

3.2. Context building

The goal of the architecture is to support semantically enhanced explanations for demand forecasting models. A key role is played by the *Prediction Analyzer* and *Context Builder* components. The *Prediction Analyzer* provides an abstraction layer on top of black-box explainability models to inform on features that are most important to certain predictions. As mentioned in Section 2.2, good explanations should convey useful information, target a specific user profile, and focus on actionability. In this line, the *Context Builder* holds the responsibility to provide explanations that (I) convey the predicted demand and uncertainty, (II) provide insights on what concepts were considered to create the prediction, (III) provide relevant context, and (IV) target a specific user. The predicted value and the associated uncertainty are queried from the Knowledge Graph.

To realize (II), the *Context Builder* interfaces with the *Prediction Analyzer* and the *KG* to obtain information regarding relevant features to certain prediction. These features are mapped to a hierarchy of attribute concept abstractions (see Fig. 3). At least one level of abstraction is considered to avoid exposing sensitive information regarding features used in the demand forecasting model and provide more concise explanations. To determine the level of abstraction exposed to the user, we followed the heuristic from Listing 1. The procedure could be further enhanced by weighting the concepts based on the underlying features' individual contribution towards the particular forecast. It is important to highlight that attribute concept abstractions are considered actionable if and only if all of their underlying concepts are actionable.

Listing 1: Algorithm used to compute attribute abstractions.

```
# ontology: an ontology with attribute concept abstractions.
# relevant_features:  list  of relevant  features to the given forecast
# n_concepts: maximum number of concepts to be shown to the user
given ontology, relevant_features, n_concepts

def get_abstraction(ontology, concept_id):
    abstraction = ontology.get_parent(concept_id)
    if abstraction == None:
        return concept_id
```

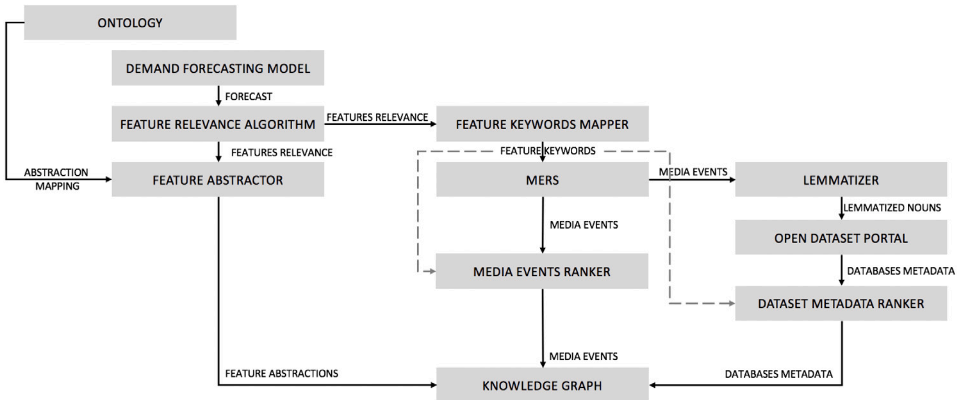

Fig. 2. Flow diagram detailing the procedure followed to provide insights on what concepts were considered to create the prediction and enrich it with relevant context (media events and suggested external datasets).

```
    return abstraction

def include_concepts_until_desired_limit ( abstractions_set ,
    sorted_candidate_concepts, n_concepts):
    idx = 0
    while abstractions_set .length() < n_concepts and idx <
        sorted_candidate_concepts.length():
        abstractions_set .add(sorted_candidate_concepts[idx])
        idx++

def compute_feature_abstractions(ontology, relevant_features, n_concepts):
    feature_abstractions = [get_abstraction(ontology, feature) for feature in
        relevant_features]

    final_abstractions  = set()
    sorted_commons = Counter(feature_abstractions).most_common()

    most_relevant_concepts = [x[0] for x in sorted_commons if x[1]>1]
    remaining_relevant_concepts = [x[0] for x in sorted_commons if x
        [1]==1]

    include_concepts_until_desired_limit ( final_abstractions ,
        most_relevant_concepts, n_concepts)
    include_concepts_until_desired_limit ( final_abstractions ,
        remaining_relevant_concepts, n_concepts)

    return  final_abstractions

compute_feature_abstractions(ontology, relevant_features, n_concepts)
```

To realize (III), it is necessary to consider specific data used to build the feature vector, related events, and additional knowledge we may gain from them. To that end, we provide a set of keywords that describe each feature. The Extract-Transform-Load process considers the keywords to issue queries against the Media Event Retrieval System. To do so, we consider the point in time each feature in the feature vector refers to. Based on it, we associate related events that are temporally close to it. Natural Language Processing techniques are applied to retrieved media events to extract novel keywords relevant to the demand forecasting problem. Those new keywords are used to query for external datasets that could enrich the existing demand forecasting model. Media events, keywords we extracted from them, and metadata regarding external datasets are persisted to the Knowledge Graph and linked to specific features from feature vector instances. We summarize the procedure described above in Fig. 2.

Finally, (IV) is realized by providing a multi-part explanation, where components are displayed or hidden, based on what information is relevant to the end-user or confidentiality policies. This responsibility is delegated to the *Service API*. E.g., a target audience could be planners interested in the demand forecast, associated uncertainty, and some context regarding feature and media events related to them. Another target audience could be experts who develop demand forecasting models. They would benefit from the data mentioned above and the possibility to drill down from described concepts to specific model features. Such an explanation could be enriched with insights on frequent keywords found in related events reported in the media and open datasets that may be useful to the model.

4. Ontology

Following related work elaborated in Section 2.3, we created an ontology that describes concepts related to demand forecasting. Specifically, it describes the use of data and AI models to predict demand, demand forecasts, and forecast explanations. The forecast explanations provide context regarding feature relevance, potentially related media events, and external datasets that could be used to enrich future models. The ontology was published and is available online [98].

Upper ontologies provide a guide on how to think about the target domain when building a domain ontology. We use BFO as our upper ontology, which supports a world's snapshot view. Our entities can be of two types: occurrent (entities that occur) or continuent (entities that persist in time).

Regarding manufacturing, we reused the concepts of *Product*, and *Event* from the ontologies developed by Leitão and Restivo [99],Borgo and Leitão [100], and Kourtis et al. [101]. In Cannataro and Comito [70],Panov et al. [72], we found the concepts of *Dataset Specification*, *Dataset*, *Algorithm*, *Datamining Algorithm*, *Regression Algorithm*, *Predictive Model*, *Regression Model*, and *Prediction*. Gottschalk and Demidova [75] provided us the concept of *Information Provenance* to represent data sources. Finally, we introduced the following concepts:

- *Media Reported Event*: relates to some *Event* that was reported in news media;
- *Media Reported Event Keyword*: keyword obtained from some *Media Reported Event*;
- *External Dataset Metadata*: a piece of metadata describing some external dataset;
- *Forecast Explanation*: a description providing reasons for a certain forecast. Relates to a specific *Prediction*, *Media Reported Event*, and *External Dataset Metadata*;
- *Attribute*: a quantity describing an instance;
- *Attribute abstraction*: a high-level concept describing some aspect of a dataset attribute;
- *Feature Vector*: a vector containing attribute values describing certain an instance of data;

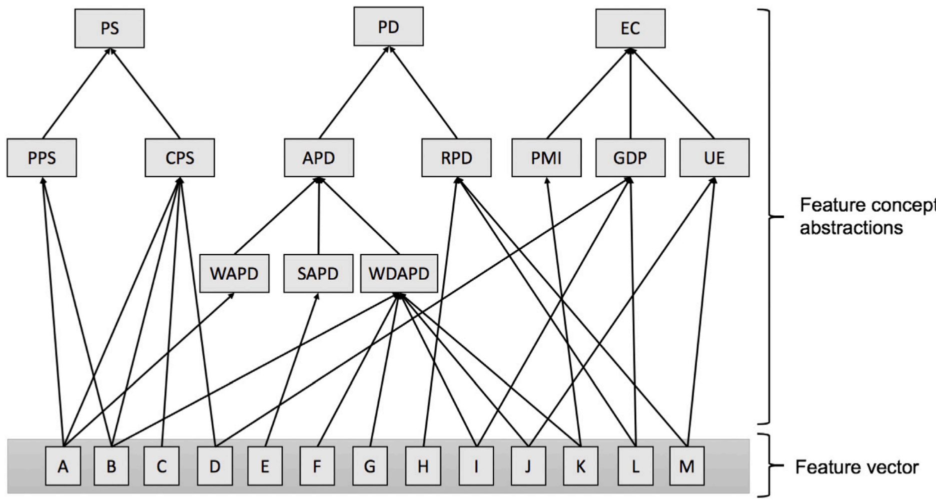

Fig. 3. Feature vector attributes and high-level concepts hierarchy associated with them. Considered attribute abstractions are PS (Planned Sales), PPS (Past Planned Sales), CPS (Current Planned Sales), PD (Past Demand), APD (Adjusted PD), SAPD (Scaled APD), WAPD (Weighted APD), WDAPD (Working Day APD), RPD (Raw Past Demand), EC (Economic Context), PMI (Purchasing Managers' Index), GDP (Gross Domestic Product), UE (Unemployment Rate). Features A-M are defined in Table 2.

Table 2
Sample feature vector and keywords considered for the use case. We use the following abbreviations: *sp* (sales plan), *GDP* (Gross Domestic Product), *PMI* (Purchasing Managers' Index), *wdp* (monthly demand averaged per amount of working days in that month), *demand* (demand in a given month), *lagNm* (N months lag), *pastwavg* (weighted average for a given month in past years), *scaled* (value scaled between 0–1).

Actionable?	Feature ID	Feature definition	MERS query keywords
YES	A	$sp \cdot \frac{demand_{pastwavg}}{sp_{pastwavg}}$	car sales demand new car sales vehicle sales car demand automotive industry
	B	$wdp_{lag12m} \cdot \frac{sp}{sp_{pastwavg}}$	Same as A
	C	sp	Same as A
	D	$sp \cdot \frac{GDP_{lag3m}}{GDP_{lag15m}}$	global GDP projection global economic outlook economic forecast + keywords listed in A
NO	E	$demand_{lag3m-scaled}$	Same as A
	F	wdp_{lag3m}	Same as A
	G	wdp_{lag8m}	Same as A
	H	$demand_{lag3m}$	Same as A
	I	$wdp_{lag3m} \cdot \frac{GDP_{lag3m}}{GDP_{lag15m}}$	Same as D
	J	$wdp_{lag5m} \cdot \frac{UE_{lag3m}}{UE_{lag15m}}$	unemployment rate unemployment numbers unemployment report employment growth long-term unemployment + keywords listed in A
	K	$wdp_{lag12m} \cdot \frac{PMI_{lag3m}}{PMI_{lag4m}}$	purchase managers' index + keywords listed in A
	L	$demand_{lag3m} \cdot \frac{GDP_{lag3m}}{GDP_{lag15m}}$	Same as D
	M	$demand_{lag4m} \cdot \frac{UE_{lag3m}}{UE_{lag15m}}$	Same as J

5. Case study

Among the related research presented in Section 2, we did not find scientific contributions focused on the explainability of AI demand forecasting models. The architecture we present in this research aims to provide the main building blocks required to provide adequate explanations for demand forecasts at a local level. Furthermore, it has been noted that XAI is foreseen to play a critical role in preserving confidentiality and that its ability to extract knowledge from AI models can be used to discover new relevant sources of knowledge [32]. We tackle both challenges, preserving confidentiality by obfuscating features relevance information behind a domain-specific ontology and using XAI insights to search for databases that could enrich the existing dataset.

To build and test the architecture, we worked with real-world data related to the automotive industry, provided by partners from European Horizon 2020 FACTLOG and STAR projects. The partners provided data for 56 different materials. Since time series forecasting can be framed as a regression problem [23], we did so to take advantage of exogenous sources and leverage data fusion to create a single dataset. Among

Table 3
Data sources grouped by their purpose.

Purpose	Data	Source
Demand Forecasting	Car Sales	International Organization of Motor Vehicle Manufacturers
	Copper Price	London Metal Exchange
	Crude Oil Price	World Bank
	Demand (deliveries)	Provided by EU H2020 FACTLOG and STAR partners
	Gross Domestic Product	World Bank
	Purchasing Managers' Index (PMI)	Institute of Supply Chain Management
	Sales Plan	Provided by EU H2020 FACTLOG and STAR partners
	Unemployment Rate	World Bank
Potentially relevant datasets	Databases Metadata	EU Open Data Portal
Media events relevant to XAI	Media Event Entries	Event Registry

the exogenous sources, we considered three relevant to the automotive sector (GDP, unemployment rates, and oil price), and three suggested by experts (Purchasing Managers' Index (PMI), copper prices, and sales plans). The PMI index summarizes purchasing managers' expectations regarding how the market will behave in the future (e.g., if it will expand, contract, or stay the same, and how strong the growth or contraction are expected to be). Copper prices are relevant to this specific demand forecasting scenario since the product prices are tied to the copper price, and thus they indirectly influence the products' demand. Finally, we incorporated the sales plans created on a quarterly and yearly basis by the strategic sales department, given the experts' suggestions and claims in the scientific literature on how do purchase intentions contribute to the forecasts' accuracy [102,103]. We summarize the data sources by their purpose in Table 3.

We considered a window of three years of data to build demand forecasting models for them. We issued predictions over the last six months of data and built explanations for all materials in the last three months. Our demand forecasting models use the Support Vector Regression algorithm [104]. We used nested cross-validation [105] to evaluate the models and make sure to use the latest data available for each target month. The models achieve an $R^2_{adjusted}$ of *0,9212* and a Mean Absolute Scaled Error [106] of *0,2600*. Finally, we used LIME [50] to implement the *Prediction Analyzer*.

The *Prediction Analyzer* provides a ranking of features for each forecast. For each of those features, we consider the query keywords presented in Table 2 to query the Media Events Retrieval System and retrieve media events that provide context to feature values. We used Event Registry [107], a well-established Media Events Retrieval System that has monitored mainstream media since 2014. Our queries consider the keywords that describe each feature and the point-in-time in which each feature component took place. E.g., suppose a feature considers (I) GDP values three and fifteen months before the forecasting horizon and (II) demand four months before the forecasting horizon. In that case, the query will ask (I) for media events related to GDP at three and fifteen months before the forecasting horizon and (II) media events regarding demand four months before the forecasting horizon.

We considered only media events reported in the English language. Once obtained, we processed them to identify and lemmatize nouns that provide insights on the most important facts related to the features. We also use these nouns to query for open datasets that may enrich the current model. In particular, we queried the EU Open Data Portal [108], a portal that provides access to open data published by EU institutions and bodies. For each matching external dataset metadata entry (provided or translated to the English language), we processed their title and description to create a word embedding with the Word2Vec algorithm [109], using Google News 300 pre-trained model [110]. Next, we create another embedding with the feature keywords from Table 2, and the keywords obtained from the media events we retrieved. To rank datasets metadata relevance, we compute the word movers distance [111] between both embeddings. Finally, to increase the diversity of displayed media event entries and datasets' metadata, we random sample them among top candidates for each case.

To create the ontology, we used Protege [112]. The ontology served as a guideline to implement virtual mapping functions to map data to the knowledge graph. We implemented the knowledge graph in Neo4j [113].

We present a sample demand forecast explanation in Fig. 4. The explanations display the following information:

- forecasted demand and associated uncertainty, for a given material at a certain point in time;
- main factors driving the forecast, which are obtained considering most important features to the forecast as provided by the *Prediction Analyzer*, and abstracted using domain knowledge from the *Knowledge Graph*;
- a highlight regarding an actionable aspect, if it exists;
- media events associated with the most important features of the particular forecast instance;
- media events' keywords frequently found in media events related to the forecast instance;
- external dataset that may be used to enrich the existing demand forecasting model.

While the user interface lists the primary factors driving the forecast, sorted based on the features' relevance and the abstractions' occurrence frequency, it does not provide insights on their relative impact on the forecast. However, the abstractions' relative impact can be computed considering the features' relative impact while considering the features' correlation and the abstractions' occurrence frequency. Computing and informing the abstractions' relative impact on the forecast remains a subject of future work.

6. Evaluation

In our research, we develop an architecture that aims to provide explanations for demand forecasting. Since media events play an essential role in our forecast explanations, we were interested in assessing (A) if media events describe features context and (B) if retrieved external datasets were related to the features for which we queried them. Regarding the explanations, we assessed metrics related to listed external datasets, events, and keywords to understand if the provided entries were accurate and diverse. To measure diversity, we computed the ratio of diverse entries (RDE) defined in Eq. (1). It is important to note that media events, keywords, and external datasets listed on explanations do not repeat themselves in a single explanation. However, the same media event, keyword, or external dataset reference may be found in another explanation. We consider that while explanations for similar data points are expected to be similar (e.g., provide similar feature ranking), multiple media events likely report a specific aspect (e.g., economic growth or job losses), or many datasets can be good candidates for data enrichment. Thus, we consider the best possible case is to find completely different but accurate media events, keywords, and external datasets across all explanations. Such a scenario would maximize end-users' learning.

In particular, for listed external datasets, we computed the accuracy and RDE. At the same time, we measured the average precision@K

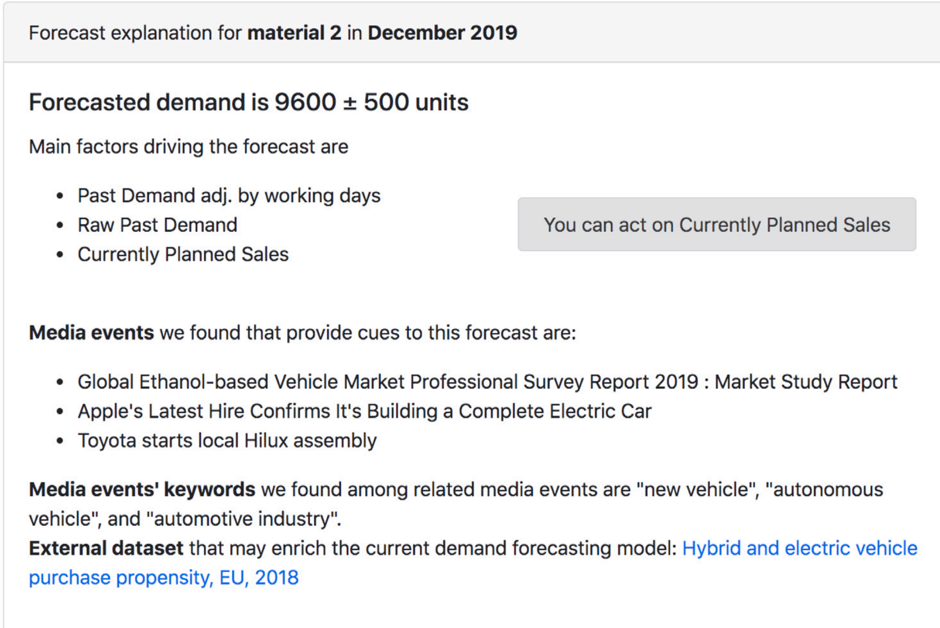

Fig. 4. An example of a demand forecast explanation that is displayed to the end user based on data retrieved from the *Service API*.

and RDE@K with K=1,3 over all of the 168 forecast explanations for media events and related media events keywords. We selected K=1,3 since three media events and media events' keywords are listed in each explanation. The choice regarding listing at most three media events and media events per explanation was made to avoid cluttering the user interface with too many features. The limit is arbitrary and could be changed if such a need arises. Average precision@K is computed as the average of the precision@K we computed for each forecast explanation. RDE@K is computed as RDE across all 168 explanation entries, considering only the first K listed media event or media events' keywords entries.

$$RDE = \frac{Unique\ Entries}{Total\ Listed\ Entries} \tag{1}$$

We evaluated these aspects by manually annotating a random sample of retrieved media events, retrieved external datasets metadata, and all media events, keywords, and external datasets proposed in demand forecast explanations.

We retrieved 128226 media events related to the features of interest and manually annotated a random sample of 528 media events. Regarding (A), we found that 69% of retrieved media events were meaningful to features explaining a specific forecast. In addition, all media events were processed to search and retrieve information regarding external datasets, as described in Section 5. From the 401 external dataset entries we retrieved and annotated, we found that (B) 64% were relevant to our use case.

We manually annotated 168 forecast explanations collected over three months. Each forecast explanation listed three media events, three media events keywords, and an external dataset recommendation. We manually annotated all of them. Results of their evaluation are presented in Table 4.

We compiled relevant features to each forecast, associated media events, their annotated subset, and the annotated suggested datasets into a dataset made publicly available at [114].

When annotating the explanations, we found some interesting cases. One of them was the suggestion of the ozone pollution dataset to enrich the automotive demand forecasting model. After some research, we found that while ozone is not emitted directly by automobiles, it is formed in the atmosphere due to a complex set of chemical reactions involving hydrocarbons, oxides of nitrogen, and sunlight. Another interesting case was a news media entry describing wireless charging.

Table 4
Results we obtained from the analysis of forecast explanations issued for 56 products over three months. Media Events, Media Events' Keywords, and External Datasets correspond to contextual information displayed for each forecast explanation.

	Metric	Value
Media Events	average precision@1	0,97
	average precision@3	0,97
	RDE@1	0,30
	RDE@3	0,11
Media Events' Keywords	average precision@1	0,77
	average precision@3	0,78
	RDE@1	0,14
	RDE@3	0,09
External Datasets	accuracy	0,56
	RDE	0,41

This media event seemed unrelated to our use case, but we confirmed that electric vehicles' demand drives research and development of a segment of wireless charging solutions. Finally, the most frequently listed media event among our forecast explanations was *"Number of people in work surges again to a new record high of 32.3m"*. We consider this event to relate to the economic context and be relevant to our demand forecasting case.

The results and examples presented above show that the current architecture and knowledge fusion procedures provide comprehensive explanations to the users while preserving confidentiality. Nevertheless, it must be noted that such an approach requires extracting relevant concepts and associating them to the specific features used to generate the forecasts, as described in Section 3.2.

7. Limitations and improvement opportunities

The architecture and knowledge fusion techniques presented in this paper assume that relevant concepts present in the features that are used to generate the forecasts are identified, mapped, and encoded into a knowledge graph. In addition, a set of keywords characterizing such concepts must be provided, as they are used to obtain additional information by querying datasets and media events. This approach requires that the features used to train the model have a particular meaning, a criterion usually satisfied by handcrafted features. Further research is required to understand how such an approach can be tailored to models where features are not handcrafted, like in deep learning models. In particular, promising advances have been made regarding the extraction of deep learning models' features semantic meaning [115,116], the incorporation of semantic aspects to deep learning models [117–119], and the use of deep learning for automated ontology development and alignment [117]. Such advancements must be considered to enable the integration of deep learning models into the architecture presented in this work.

We have identified some improvement opportunities from the experience we obtained through the experiments and results described in this paper. While the method we have used to create local explanations and find the most meaningful features to a particular forecast LIME is widely adopted, it is not deterministic. Changes in the feature ranking computations can result in different forecast interpretations. Replacing LIME for some deterministic variant (e.g., DLIME), would provide a more consistent experience to the end-user and enhance scenarios reproducibility for researchers and engineers. Another possible improvement is to highlight current events related to the factors that are driving the forecast. Insights into past and current events would provide the user an enhanced understanding of the overall context, having insights into the past events that affect the models' prediction and relevant current events that can enhance judgment and derive better decision-making.

8. Conclusion

This research work presented an architecture that supports building semantically enhanced explanations for AI models' demand forecasts. We evaluated the architecture and explanations on a real-world use case developed with data obtained from European Horizon 2020 projects FACTLOG and STAR.

Current explanations provide demand forecast values, the associated uncertainty, high-level description of features that influenced the prediction, related media events, and a reference to some external dataset, providing context and means to improve the demand forecasting model. In addition, the high-level description of features informs the main factors influencing the forecast and avoids exposing sensitive details regarding the demand forecasting models.

Our future work will focus on further pre-processing media events, media events' keywords, and external datasets, to filter out those not relevant to the demand forecasting context and increase the forecast's explanation quality. At least three additional research directions can be pursued. First, explore means to increase the diversity of recommended media events and associated keywords displayed on forecast explanations. Second, evaluate the impact of incorporating suggested external datasets to the performance of existing demand forecasting models. Third, we envision enhancing the explanations by including meaningful current events reported by the media. This way, in addition to a good understanding of the past context, the explanations will provide information on events that are likely to influence future demand so that the user can gain a better judgment and perspective when making a decision. Finally, we would like to research the integration of semantic databases and federated ontologies, and how explanations can be enhanced with them.

CRediT authorship contribution statement

Jože M. Rožanec: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Blaž Fortuna:** Resources, Supervision, Project administration, Funding acquisition. **Dunja Mladenić:** Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Slovenian Research Agency, Slovenia and the European Union's Horizon 2020 program project FACTLOG under grant agreement H2020-869951, and European Union's Horizon 2020 program project STAR under grant agreement number H2020-956573.

References

[1] M. Gaur, K. Faldu, A. Sheth, Semantics of the Black-Box: Can knowledge graphs help make deep learning systems more interpretable and explainable? 2020, arXiv preprint arXiv:2010.08660.
[2] C. Panigutti, A. Perotti, D. Pedreschi, Doctor XAI: an ontology-based approach to black-box sequential data classification explanations, in: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 2020, pp. 629–639.
[3] J.M. Rožanec, B. Kažič, M. Škrjanc, B. Fortuna, D. Mladenić, Automotive OEM demand forecasting: A comparative study of forecasting algorithms and strategies, Appl. Sci. 11 (2021) 6787.
[4] F. Lolli, R. Gamberini, A. Regattieri, E. Balugani, T. Gatos, S. Gucci, Single-hidden layer neural networks for forecasting intermittent demand, Int. J. Prod. Econ. 183 (2017) 116–128.
[5] J.J. Bergman, J.S. Noble, R.G. McGarvey, R.L. Bradley, A Bayesian approach to demand forecasting for new equipment programs, Robot. Comput.-Integr. Manuf. 47 (2017) 17–21.
[6] M. Babai, A. Tsadiras, C. Papadopoulos, On the empirical performance of some new neural network methods for forecasting intermittent demand, IMA J. Manag. Math. (2020).
[7] T. Benbarrad, M. Salhaoui, S.B. Kenitar, M. Arioua, Intelligent machine vision model for defective product inspection based on machine learning, J. Sensor Actuator Netw. 10 (2021) 7.
[8] R. Rajkumar, I. Lee, L. Sha, J. Stankovic, Cyber-physical systems: the next computing revolution, in: Design Automation Conference, IEEE, 2010, pp. 731–736.
[9] M.W. Grieves, Virtually intelligent product systems: digital and physical twins, 2019.
[10] P. Tubaro, A.A. Casilli, Micro-work, artificial intelligence and the automotive industry, J. Ind. Bus. Econ. 46 (2019) 333–345.
[11] R.D. Raut, A. Gotmare, B.E. Narkhede, U.H. Govindarajan, S.U. Bokade, Enabling technologies for Industry 4.0 manufacturing and supply chain: concepts, current status, and adoption challenges, IEEE Eng. Manag. Rev. 48 (2020) 83–102.
[12] F. Tao, Q. Qi, A. Liu, A. Kusiak, Data-driven smart manufacturing, J. Manuf. Syst. 48 (2018) 157–169.
[13] B. Brühl, M. Hülsmann, D. Borscheid, C.M. Friedrich, D. Reith, A sales forecast model for the german automobile market based on time series analysis and data mining methods, in: Industrial Conference on Data Mining, Springer, 2009, pp. 146–160.
[14] F.-K. Wang, K.-K. Chang, C.-W. Tzeng, Using adaptive network-based fuzzy inference system to forecast automobile sales, Expert Syst. Appl. 38 (2011) 10587–10593.
[15] A. Vahabi, S.S. Hosseininia, M. Alborzi, A sales forecasting model in automotive industry using adaptive neuro-fuzzy inference system (anfis) and genetic algorithm (GA), Management 1 (2016) 2.
[16] J. Gao, Y. Xie, X. Cui, H. Yu, F. Gu, Chinese automobile sales forecasting using economic indicators and typical domestic brand automobile sales data: A method based on econometric model, Adv. Mech. Eng. 10 (2018) 1687814017749325.
[17] N.Z. Ubaidillah, A study of car demand and its interdependency in sarawak, Int. J. Bus. Soc. 21 (2020).
[18] T. Williams, Stock control with sporadic and slow-moving demand, J. Oper. Res. Soc. 35 (1984) 939–948.
[19] F. Johnston, J.E. Boylan, Forecasting for items with intermittent demand, J. Oper. Res. Soc. 47 (1996) 113–121.
[20] A.A. Syntetos, J.E. Boylan, J. Croston, On the categorization of demand patterns, J. Oper. Res. Soc. 56 (2005) 495–503.
[21] A.A. Syntetos, M.Z. Babai, N. Altay, On the demand distributions of spare parts, Int. J. Prod. Res. 50 (2012) 2101–2117.

[22] D. Lengu, A.A. Syntetos, M.Z. Babai, Spare parts management: Linking distributional assumptions to demand classification, European J. Oper. Res. 235 (2014) 624–635.
[23] R. Saluja, A. Malhi, S. Knapič, K. Främling, C. Cavdar, Towards a rigorous evaluation of explainability for multivariate time series, 2021, arXiv preprint arXiv:2104.04075.
[24] A. Dwivedi, M. Niranjan, K. Sahu, A business intelligence technique for forecasting the automobile sales using adaptive intelligent systems (ANFIS and ANN), Int. J. Comput. Appl. 74 (2013).
[25] X. Wang, D. Zeng, H. Dai, Y. Zhu, Making the right business decision: Forecasting the binary NPD strategy in Chinese automotive industry with machine learning methods, Technol. Forecast. Soc. Change 155 (2020) 120032.
[26] D.S. Farahani, M. Momeni, N.S. Amiri, Car sales forecasting using artificial neural networks and analytical hierarchy process, DATA ANALYTICS 2016 (2016) 69.
[27] K.K. Chandriah, R.V. Naraganahalli, RNN/LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting, Multimedia Tools Appl. (2021) 1–15.
[28] R. Fildes, P. Goodwin, Stability in the inefficient use of forecasting systems: A case study in a supply chain company, Int. J. Forecast. 37 (2021) 1031–1046.
[29] P. Corredor, E. Ferrer, R. Santamaria, Is cognitive bias really present in analyst forecasts? The role of investor sentiment, Int. Bus. Rev. 23 (2014) 824–837.
[30] R.M. Hogarth, S. Makridakis, Forecasting and planning: An evaluation, Manage. Sci. 27 (1981) 115–138.
[31] J.H. Barnes Jr., Cognitive biases and their impact on strategic planning, Strat. Manag. J. 5 (1984) 129–137.
[32] A.B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, et al., Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, Inf. Fusion 58 (2020) 82–115.
[33] O. Biran, C. Cotton, Explanation and justification in machine learning: A survey, in: IJCAI-17 Workshop on Explainable AI (XAI), vol. 8, 2017, pp. 8–13.
[34] A. Malhi, S. Knapic, K. Främling, Explainable agents for less bias in human-agent decision making, in: International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems, Springer, 2020, pp. 129–146.
[35] X. Wang, M. Yin, Are explanations helpful? A comparative study of the effects of explanations in AI-assisted decision-making, in: 26th International Conference on Intelligent User Interfaces, 2021, pp. 318–328.
[36] J.J. Ferreira, M. Monteiro, The human-AI relationship in decision-making: AI explanation to support people on justifying their decisions, 2021, arXiv preprint arXiv:2102.05460.
[37] O. Biran, K.R. McKeown, Human-centric justification of machine learning predictions, in: IJCAI, vol. 2017, 2017, pp. 1461–1467.
[38] A. Adadi, M. Berrada, Peeking inside the black-box: a survey on explainable artificial intelligence (XAI), IEEE Access 6 (2018) 52138–52160.
[39] O. Loyola-Gonzalez, Black-box vs. white-box: Understanding their advantages and weaknesses from a practical point of view, IEEE Access 7 (2019) 154096–154113.
[40] A. Das, P. Rad, Opportunities and challenges in explainable artificial intelligence (xai): A survey, 2020, arXiv preprint arXiv:2006.11371.
[41] Q.V. Liao, D. Gruen, S. Miller, Questioning the AI: informing design practices for explainable AI user experiences, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–15.
[42] P.P. Angelov, E.A. Soares, R. Jiang, N.I. Arnold, P.M. Atkinson, Explainable artificial intelligence: an analytical review, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 11 (2021) e1424.
[43] T. Rojat, R. Puget, D. Filliat, J. Del Ser, R. Gelin, N. Díaz-Rodríguez, Explainable artificial intelligence (XAI) on TimeSeries data: A survey, 2021, arXiv preprint arXiv:2104.00950.
[44] S.A. Siddiqui, D. Mercier, M. Munir, A. Dengel, S. Ahmed, Tsviz: Demystification of deep learning models for time-series analysis, IEEE Access 7 (2019) 67027–67040.
[45] K. Kashiparekh, J. Narwariya, P. Malhotra, L. Vig, G. Shroff, ConvTimeNet: A pre-trained deep convolutional neural network for time series classification, in: 2019 International Joint Conference on Neural Networks (IJCNN), IEEE, 2019, pp. 1–8.
[46] S. Tonekaboni, S. Joshi, D. Duvenaud, A. Goldenberg, Explaining time series by counterfactuals, 2019.
[47] O. Ozyegen, I. Ilic, M. Cevik, Evaluation of local explanation methods for multivariate time series forecasting, 2020, arXiv preprint arXiv:2009.09092.
[48] D. Mercier, A. Dengel, S. Ahmed, P2exnet: Patch-based prototype explanation network, in: International Conference on Neural Information Processing, Springer, 2020, pp. 318–330.
[49] B. Lim, S.O. Arık, N. Loeff, T. Pfister, Temporal fusion transformers for interpretable multi-horizon time series forecasting, Int. J. Forecast. (2021).
[50] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should I trust you?" Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.
[51] P. Hall, N. Gill, M. Kurka, W. Phan, Machine learning interpretability with h2o driverless ai, 2017, H2O. Ai. URL: http://docs.h2o.ai/driverless-ai/latest-stable/docs/booklets/MLIBooklet.pdf.
[52] M.R. Zafar, N.M. Khan, DLIME: A deterministic local interpretable model-agnostic explanations approach for computer-aided diagnosis systems, 2019, arXiv preprint arXiv:1906.10263.
[53] K. Sokol, P. Flach, LIMEtree: Interactively customisable explanations based on local surrogate multi-output regression trees, 2020, arXiv preprint arXiv:2005.01427.
[54] M.T. Ribeiro, S. Singh, C. Guestrin, Anchors: High-precision model-agnostic explanations, in: AAAI, 18, 2018, pp. 1527–1535.
[55] J. van der Waa, M. Robeer, J. van Diggelen, M. Brinkhuis, M. Neerincx, Contrastive explanations with local foil trees, 2018, arXiv preprint arXiv:1806.07470.
[56] R. Guidotti, A. Monreale, S. Ruggieri, D. Pedreschi, F. Turini, F. Giannotti, Local rule-based explanations of black box decision systems, 2018, arXiv preprint arXiv:1805.10820.
[57] J. Rožanec, E. Trajkova, K. Kenda, B. Fortuna, D. Mladenić, Explaining bad forecasts in global time series models, 2021.
[58] W. Samek, K.-R. Müller, Towards explainable artificial intelligence, in: Explainable AI: Interpreting, Explaining and Visualizing Deep Learning, Springer, 2019, pp. 5–22.
[59] R. Srinivasan, A. Chander, Explanation perspectives from the cognitive sciences-a survey, in: IJCAI, 2020, pp. 4812–4818.
[60] D. Pedreschi, F. Giannotti, R. Guidotti, A. Monreale, L. Pappalardo, S. Ruggieri, F. Turini, Open the black box data-driven explanation of black box decision systems, 2018, arXiv preprint arXiv:1806.09936.
[61] S. Verma, J. Dickerson, K. Hines, Counterfactual explanations for machine learning: A review, 2020, arXiv preprint arXiv:2010.10596.
[62] D. Doran, S. Schulz, T.R. Besold, What does explainable AI really mean? A new conceptualization of perspectives, 2017, arXiv preprint arXiv:1710.00794.
[63] M. Chromik, Human-Centric Explanation Facilities, (Ph.D. thesis), lmu, 2021.
[64] M. El-Assady, W. Jentner, R. Kehlbeck, U. Schlegel, R. Sevastjanova, F. Sperrle, T. Spinner, D. Keim, Towards XAI: Structuring the processes of explanations, in: ACM Workshop on Human-Centered Machine Learning, 2019.
[65] M.T. Keane, B. Smyth, Good counterfactuals and where to find them: A case-based technique for generating counterfactuals for explainable AI (XAI), in: International Conference on Case-Based Reasoning, Springer, 2020, pp. 163–178.
[66] M.T. Keane, E.M. Kenny, E. Delaney, B. Smyth, If only we had better counterfactual explanations: five key deficits to rectify in the evaluation of counterfactual XAI techniques, 2021, arXiv preprint arXiv:2103.01035.
[67] F. Ameri, D. Dutta, An upper ontology for manufacturing service description, in: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 42578, 2006, pp. 651–661.
[68] S. Lemaignan, A. Siadat, J.-Y. Dantan, A. Semenenko, MASON: A proposal for an ontology of manufacturing domain, in: IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06), IEEE, 2006, pp. 195–200.
[69] M. Courtot, F. Gibson, A.L. Lister, J. Malone, D. Schober, R.R. Brinkman, A. Ruttenberg, MIREOT: The minimum information to reference an external ontology term, Appl. Ontol. 6 (2011) 23–33.
[70] M. Cannataro, C. Comito, A data mining ontology for grid programming, in: Proc. 1st Int. Workshop on Semantics in Peer-To-Peer and Grid Computing, Citeseer, 2003, pp. 113–134.
[71] P. Panov, S. Džeroski, L. Soldatova, OntoDM: An ontology of data mining, in: 2008 IEEE International Conference on Data Mining Workshops, IEEE, 2008, pp. 752–760.
[72] P. Panov, L. Soldatova, S. Džeroski, Ontology of core data mining entities, Data Min. Knowl. Discov. 28 (2014) 1222–1265.
[73] C. Diamantini, D. Potena, E. Storti, Kddonto: An ontology for discovery and composition of kdd algorithms, in: Third Generation Data Mining: Towards Service-Oriented Knowledge Discovery (SoKD'09), Citeseer, 2009, pp. 13–24.
[74] V. Ermolayev, N. Keberle, W.-E. Matzke, An ontology of environments, events, and happenings, in: 2008 32nd Annual IEEE International Computer Software and Applications Conference, IEEE, 2008, pp. 539–546.
[75] S. Gottschalk, E. Demidova, Eventkg: A multilingual event-centric temporal knowledge graph, in: European Semantic Web Conference, Springer, 2018, pp. 272–287.
[76] R. Mizoguchi, YAMATO: yet another more advanced top-level ontology, in: Proceedings of the Sixth Australasian Ontology Workshop, 2010, pp. 1–16.
[77] M. Uschold, Building ontologies: Towards a uni ed methodology, in: Proceedings of 16th Annual Conference of the British Computer Society Specialists Group on Expert Systems, Citeseer, 1996.
[78] M. Uschold, M. King, Towards a Methodology for Building Ontologies, Citeseer, 1995.
[79] M. Fernández-López, A. Gómez-Pérez, N. Juristo, Methontology: from ontological art towards ontological engineering, 1997.

[80] H.M. Kim, M.S. Fox, M. Gruninger, An ontology of quality for enterprise modelling, in: Proceedings 4th IEEE Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE'95), IEEE, 1995, pp. 105–116.
[81] F. Ameri, C. Urbanovsky, C. McArthur, A systematic approach to developing ontologies for manufacturing service modeling, in: Proceedings of the Workshop on Ontology and Semantic Web for Manufacturing, vol. 14, 2012.
[82] X. Chang, R. Rai, J. Terpenny, Development and utilization of ontologies in design for manufacturing, J. Mech. Des. 132 (2010).
[83] R. Weber, M. Shrestha, A.J. Johs, Knowledge-based XAI through CBR: There is more to explanations than models can tell, 2021, arXiv preprint arXiv:2108.10363.
[84] S.R. Islama, W. Eberleb, Implications of combining domain knowledge in explainable artificial intelligence, 2021.
[85] R. Sharma, A.K. Sinha, Sales forecast of an automobile industry, Int. J. Comput. Appl. 53 (2012).
[86] D. Salinas, V. Flunkert, J. Gasthaus, T. Januschowski, DeepAR: Probabilistic forecasting with autoregressive recurrent networks, Int. J. Forecast. 36 (2020) 1181–1191.
[87] M. Nentwig, M. Hartung, A.-C. Ngonga Ngomo, E. Rahm, A survey of current link discovery frameworks, Semantic Web 8 (2017) 419–436.
[88] M.A. Khoudja, M. Fareh, H. Bouarfa, Ontology matching using neural networks: survey and analysis, in: 2018 International Conference on Applied Smart Systems (ICASS), IEEE, 2018, pp. 1–6.
[89] J.P. McCrae, P. Buitelaar, Linking datasets using semantic textual similarity, Cybernet. Inform. Technol. 18 (2018) 109–123.
[90] E. Thiéblin, O. Haemmerlé, N. Hernandez, C. Trojahn, Survey on complex ontology matching, Semantic Web 11 (2020) 689–727.
[91] L.-W. Chen, Extending wikification: Nominal discovery, nominal linking, and the grounding of nouns, 2017.
[92] J. Brank, G. Leban, M. Grobelnik, Annotating documents with relevant wikipedia concepts, in: Proceedings of SiKDD.
[93] R.C. Fernandez, E. Mansour, A.A. Qahtan, A. Elmagarmid, I. Ilyas, S. Madden, M. Ouzzani, M. Stonebraker, N. Tang, Seeping semantics: Linking datasets using word embeddings for data discovery, in: 2018 IEEE 34th International Conference on Data Engineering (ICDE), IEEE, 2018, pp. 989–1000.
[94] R.J. Miller, Open data integration, Proc. VLDB Endow. 11 (2018) 2130–2139.
[95] X.L. Dong, T. Rekatsinas, Data integration and machine learning: A natural synergy, in: Proceedings of the 2018 International Conference on Management of Data, 2018, pp. 1645–1650.
[96] E. Demidova, A. Zaveri, E. Simperl, Matching domain and top-level ontologies exploring word sense disambiguation and word embedding, in: Emerging Topics in Semantic Technologies: ISWC 2018 Satellite Events, vol. 36, IOS Press, 2018, p. 27.
[97] A. Assi, H. Mcheick, A. Karawash, W. Dhifli, Context-aware instance matching through graph embedding in lexical semantic space, Knowl.-Based Syst. 186 (2019) 104925.
[98] J.z.M. Roˇ zanec, D. Mladenić, Ontology for contextualized demand forecasting explanations, 2020, http://dx.doi.org/10.7910/DVN/GIREAY.
[99] P. Leitão, F. Restivo, ADACOR: A holonic architecture for agile and adaptive manufacturing control, Comput. Ind. 57 (2006) 121–130.
[100] S. Borgo, P. Leitão, Foundations for a core ontology of manufacturing, in: Ontologies, Springer, 2007, pp. 751–775.
[101] G. Kourtis, E. Kavakli, R. Sakellariou, A rule-based approach founded on description logics for Industry 4.0 smart factories, IEEE Trans. Ind. Inf. 15 (2019) 4888–4899.
[102] J.S. Armstrong, V.G. Morwitz, V. Kumar, Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy? Int. J. Forecast. 16 (2000) 383–397.
[103] V.G. Morwitz, J.H. Steckel, A. Gupta, When do purchase intentions predict sales? Int. J. Forecast. 23 (2007) 347–364.
[104] H. Drucker, C.J. Burges, L. Kaufman, A.J. Smola, V. Vapnik, Support vector regression machines, in: Advances in Neural Information Processing Systems, 1997, pp. 155–161.
[105] M. Stone, Cross-validatory choice and assessment of statistical predictions, J. R. Stat. Soc. Ser. B Stat. Methodol. 36 (1974) 111–133.
[106] R.J. Hyndman, et al., Another look at forecast-accuracy metrics for intermittent demand, Int. J. Appl. Forecast. 4 (2006) 43–46.
[107] G. Leban, B. Fortuna, J. Brank, M. Grobelnik, Event registry: learning about world events from news, in: Proceedings of the 23rd International Conference on World Wide Web, 2014, pp. 107–110.
[108] P.O. of the European Union, EU open data portal: The official portal for European data, 2020, https://data.europa.eu (Accessed 15 December 2020).
[109] T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, in: Advances in Neural Information Processing Systems, 2013, pp. 3111–3119.
[110] Google, Word2Vec: an efficient implementation of the continuous bag-of-words and skip-gram architectures for computing vector representations of words, 2015, https://code.google.com/archive/p/word2vec (Accessed 15 December 2020).
[111] M. Kusner, Y. Sun, N. Kolkin, K. Weinberger, From word embeddings to document distances, in: International Conference on Machine Learning, 2015, pp. 957–966.
[112] N.F. Noy, M. Crubézy, R.W. Fergerson, H. Knublauch, S.W. Tu, J. Vendetti, M.A. Musen, Protégé-2000: an open-source ontology-development and knowledge-acquisition environment, in: AMIA... Annual Symposium Proceedings. AMIA Symposium, 2003, p. 953.
[113] Neo4j, Neo4j: The graph database platform for today's intelligent applications, 2020, https://neo4j.com (Accessed 15 December 2020).
[114] J.z.M. Roˇ zanec, D. Mladenić, Data for contextualized demand forecasting explanations, 2020, http://dx.doi.org/10.7910/DVN/FPXYCM.
[115] A. Gudi, Recognizing semantic features in faces using deep learning, 2015, arXiv preprint arXiv:1512.00743.
[116] F.D. de Souza, S. Sarkar, G. Cámara-Chávez, Building semantic understanding beyond deep learning from sound and vision, in: 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE, 2016, pp. 2097–2102.
[117] X. Huang, C. Zanni-Merk, B. Crémilleux, Enhancing deep learning with semantics: an application to manufacturing time series analysis, Procedia Comput. Sci. 159 (2019) 437–446.
[118] E. Coviello, A.B. Chan, G. Lanckriet, Time series models for semantic music annotation, IEEE Trans. Audio Speech Lang. Process. 19 (2010) 1343–1359.
[119] Y. Dong, H. Su, J. Zhu, B. Zhang, Improving interpretability of deep neural networks with semantic information, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4306–4314.