



HAL
open science

From Indicators to Predictive Analytics: A Conceptual Modelling Framework

Azadeh Nasiri, Soroosh Nalchigar, Eric Yu, Waqas Ahmed, Robert Wrembel,
Esteban Zimanyi

► **To cite this version:**

Azadeh Nasiri, Soroosh Nalchigar, Eric Yu, Waqas Ahmed, Robert Wrembel, et al.. From Indicators to Predictive Analytics: A Conceptual Modelling Framework. 10th IFIP Working Conference on The Practice of Enterprise Modeling (PoEM), Nov 2017, Leuven, Belgium. pp.171-186, 10.1007/978-3-319-70241-4_12 . hal-01765243

HAL Id: hal-01765243

<https://inria.hal.science/hal-01765243>

Submitted on 12 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

From Indicators to Predictive Analytics: A Conceptual Modelling Framework

Azadeh Nasiri^{1,2,3}, Soroosh Nalchigar³, Eric Yu³, Waqas Ahmed^{1,2}, Robert Wrembel¹, and Esteban Zimányi²

¹ Institute of Computing Science, Poznan University of Technology,
`robert.wrembel@cs.put.poznan.pl`,

² Department of Computer and Decision Engineering, Université Libre de Bruxelles,
`waqas.ahmed, ezimanyi@ulb.ac.be`,

³ Department of Computer Science, University of Toronto,
`nazadeh, soroosh, eric@cs.toronto.edu`

Abstract. Predictive analytics provides organisations with insights about future outcomes. Despite the hype around it, not many organizations are using it. Organisations still rely on the descriptive insights provided by the traditional business intelligence (BI) solutions. The barriers to adopt predictive analytics solutions are that businesses struggle to understand how such analytics could enhance their existing BI capabilities, and also businesses lack a clear understanding of how to systematically design the predictive analytics. This paper presents a conceptual modelling framework to overcome these barriers. The framework consists of two modelling components and a set of analysis that systematically 1) justify the needs for predictive analytics within the organisational context, and 2) identify the predictive analytics design requirements. The framework is illustrated using a real case adopted from the literature.

Key words: Conceptual Modelling, Predictive Analytics, Goal-Oriented Requirements Engineering

1 Introduction

Data analytics is defined as the discovery of meaningful patterns in data [1]. While descriptive analytics provides insights about the past and current business performance, predictive analytics answer what will happen in the future [2]. Results from such analytics trigger proactive actions towards achieving organisational objectives in an optimal and timely fashion [3, 4]. Despite the importance of predictive analytics, only 13% of the organisations utilize it [5]. According to Gartner, over 70% of organisations still use only the traditional business intelligence (BI) which provides descriptive insight by means of indicators. Indicators are measurements that quantify the fulfilment of the business goals. The insights obtained via indicators are retrospective, and answer questions what happened or what is currently happening in terms of goal fulfilment. However, today's fast-paced business environment demands actionable insights to make real-time

optimization adjustments to the business goals. In this regard, descriptive insights provided by indicators should be evolved to future-oriented actionable insights which is provided by predictive analytics.

While many businesses have already invested in BI systems, they are struggling to understand how predictive analytics could enhance their existing BI capabilities. The effective use of predictive analytics requires business stakeholders 1) to identify how to move from descriptive insight to an actionable insight that could answer why and where in the organisations to use predictive analytics, and 2) to identify the important design requirements of predictive analytics from data and computational perspectives.

Conceptual modelling techniques can potentially provide great value towards the effective use of data analytics solutions [6]. Research efforts in this domain have addressed a wide range of data-intensive problems in business domains, from data warehousing [7, 8] to business intelligence [9, 10] to advanced analytics [11, 12]. The existing conceptual modelling approaches do not address challenges such as: How to obtain the right predictive analytics requirements from available descriptive insights? How to make analytics design choices such as what are the available prediction strategies to address the problem at hand and what are the potential algorithms and data asset to focus on?

Consider a retailer from the fashion industry that already uses a descriptive BI system tracking the fulfillment of the business goals. The BI system reveals often deficiency in the inventory-related goals. The insight has driven the retailer to improve the replenishment of the inventory to increase the fulfillment chance of the corresponding goals. It has been decided to estimate the demand and replenish the inventory based on that. The estimation needs to utilize predictive analytics. The challenge is that the retailer does not know how to drive the estimation of demand that responds to the current goal deficiency captured by the traditional BI system. Besides, to support the estimation by predictive analytics, the retailer does not know how to make analytics design choices such as what in the business domain or in its environment drives the demand? What are the potential algorithms and data to use?

This paper presents a conceptual modelling framework to overcome the barriers to adapt predictive analytics in organisations. The framework consists of two modelling components: (1) the **Context Modelling** that extends the Business Intelligence Model (BIM) language [9] to support moving from descriptive insight to an actionable insight (provided by predictive analytics), (2) the **Analytics Design Alternatives Modelling** that adapts the i* framework [13] to support designing the predictive analytics. Furthermore, we present a set of analyses to capture the modelling components from the methodological perspective. The framework is applied to a real case prepared from the literature. Section 2 explains the motivation case from the literature. Section 3 provides an overview of the modelling components and a metamodel of the framework. Section 4 represents the framework which is applied on the motivation case. Section 5 reviews the related works and finally, Section 6 concludes the paper.

2 A Motivating case

In this section, we elaborate on the example from the fashion industry. A retailer orders a certain amount of products from manufacturers and put them in the inventory to meet consumer demands. There are trendy and functional products. Trendy products are sold through the fast fashion retailing process which involves 20 short selling seasons in a year. There is no chance to buy them when a season is over. Functional products are sold through a traditional process with two selling seasons in a year. The retailer wants to improve the replenishment of the inventory by predicting the customer demands and for that, aims at building a predictive analytics solution. We first describe how the BI system of the retailer is set up without use of predictive analytics.

The retailer has already invested in a BI system to track the fulfillment of its business goals by monitoring of various indicators benefiting from the BIM modelling framework. *BIM* is a goal-oriented modeling language that assists enterprises to keep track of their performance and sense how well they are doing with regard to their strategic goals. Modelling with BIM involves building a *goal model* which consists of business concepts such as *strategies*, *business processes*, *indicators* and *influences*. A detailed introduction of BIM is beyond the scope of this paper (refer to [9] in order to use BIM for enterprise modelling). Using BIM, the retailer has formalised the strategies which have been further decomposed into lower level business goals. The relevant business processes, tasks, and resources have been assigned to each business goal to fulfill it. In Fig. 1, the business strategy of *To meet consumer demand* is decomposed into business goals namely, *To increase reliability* and *To increase responsiveness*. The first goal is achieved by executing the *Traditional retailing process*. The second goal is achieved by executing the *Fast fashion retailing process* through which the *trendy products* are sold.

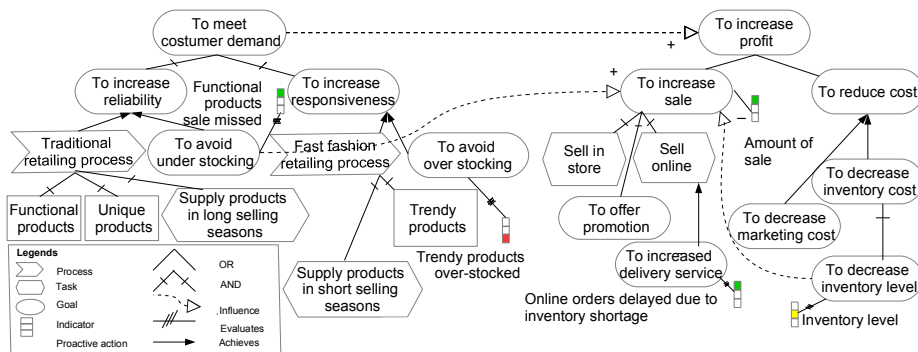


Fig. 1. An excerpt of the BIM schema of the retailer.

Indicators are associated with each business goal to quantify its fulfillment. For example, in Fig. 1, *Functional products sale missed* is an indicator that

measures the amount of unfulfilled demand due to the insufficient product in the inventory. This indicator quantifies the achievement of goal *To avoid under stocking*. At run-time, each indicator obtains its current value from the data sources to which it is linked to. The current value is typically compared with a target, threshold and the worst value. The result of this comparison is visually expressed for business users by color. Depends on the current value at the given time, indicators fall into a green, a yellow, or a red state. The green, yellow and red state of an indicator means that the associating goal is fulfilled, partially fulfilled, or not fulfilled, respectively. At a given time, the state of the various indicators provides a descriptive insight over the current performance of the retailer in terms of goal fulfillment. The insight is historically-oriented and reflects which business goals have been already fulfilled and which are still unfulfilled. However, it does not deal with the desired outcome in terms of goal fulfillment.

The retailer has recently observed inventory-related indicators often in the yellow or red state representing deficiencies in the fulfillment of associating goals. To deal with that, the inventory replenishment should be improved. For that, a predictive analytics solution is needed that estimates the customer demand and based on that the inventory is replenished. The objective of the predictive analytics is to improve the replenishment in a way that the fulfilled goals (with green indicator state) are kept as they are while the fulfillment chance of the deficient goals (with yellow or red indicator state) is increased.

Although BIM simplifies interpreting the descriptive insight over the fulfillment of inventory-related goals using indicators, it does not support a future-oriented insight over the demand that improves the the fulfillment chance of unfulfilled goals. To provide an actionable insight using BIM, there are business and analytics-related challenges. We will go through them in the next two paragraphs. The challenges addressed are synthesized from the literature [14].

Regarding the business challenges, the first issue is to deal with the goals in conflict. It means that the efforts to make one goal fulfilled might negatively impact the fulfillment of another goal. For example, *To decrease the inventory level* which saves inventory cost might negatively impact *To increase sale* because of the insufficient stock to meet demand. The retailer does not know how to effectively replenish the inventory that decreases the cost while meeting the demand. Moreover, the business goals are achieved through different business processes and target various products and services. The retailer needs to take different initiatives when improving the fulfillment of one goal to another. For example, *To avoid over-stocking* is achieved by *Fast fashion retailing process* by which trendy products are sold. To improve such goal the replenishment of *Trendy products* should be improved. Whereas, *To avoid under-stocking* the replenishment of *traditional products* should be addressed with a different replenishment frequency than *Trendy products*. Besides, goal fulfillment undergo changes over time as a result of dynamics in internal and external business environment. For example, *To avoid over stocking* which was fulfilled awhile ago happened to be unfulfilled now because of a sudden drop in sales. Therefore, depending on which inventory-related goals are unfulfilled and which one is met at given time, the

different replenishment and accordingly the different demand estimation is required. It is challenging for the retailer to know how to replenish the inventory that responds to goal deficiencies in an optimal and timely fashion. To conclude, the challenge is to align the predictive analytics that supports the estimation of demand with the descriptive insight provided in BIM.

Along with business challenges to replenish the inventory, there are also challenges to design the predictive analytics for estimating the demand of products. There are different domain-dependant ways to estimate the demand. For example, for trendy products, the retailer deals with the demand of new products which is different from traditional products with a history of sales. To estimate the demand for trendy products, one way is to run a pre-test sale and based on the consumer response, forecast the demand. The other ways are to estimate the demand based on the sale of similar products previously sold, based on user-generated content on social media, or based on weather forecast, etc. Moreover, each way can involve various data as well as various computational techniques. It is challenging for the retailer to make critical choices to design a predictive analytics that supports the estimation of demand. In Sec 3. we introduce new business and analytics-related concepts to make it easier for the retailer to address the above-mentioned challenges.

3 The proposed framework

The proposed framework consists of two components to augment the BIM framework, namely, **Context modelling** and **Analytics design alternatives modelling**. The former justifies the rationale behind the predictive analytics solutions. The latter focuses on how to design the predictive analytics solutions. Next, we explain each component in detail.

3.1 Context modelling

This part aims to justify the need for predictive analytics. This component systematically aligns the predictive analytics with the descriptive insight provided by the indicators in BIM. To support and facilitate this alignment, a set of concepts are introduced and captured in the extended BIM schema using the new notations illustrated in Fig. 3. The concepts are defined as follows:

Optimization goal represents an intention to make a balance among a set of goals that are possibly in conflict. The goals influenced by an optimization goal, called **influenced goals**, have indicators that evaluate to which degree they are fulfilled whereas, optimization goal has no clear-cut fulfillment level. It is considered to be sufficiently fulfilled if all associating goals are fulfilled.

Indicator state depicts visually the performance level of an indicator at run-time using green, yellow and red colors. The colors simplify the interpretation of the fulfillment of the goal. The performance level is obtained from the current value (comes from data sources) of an indicator compared with a target,

threshold, and the worst value [10]. Green, yellow, and red are interpreted as the fully fulfilled, the partially fulfilled, and the unfulfilled goal, respectively.

Business situation is a meaningful combination of indicator states appeared over a group of indicators (belong to the influenced goals of an optimization goal) in a snapshot from BIM at run-time. This combination includes at least one yellow or red indicator state.

Forward-looking insight contributes to satisfy an optimization goal by studying the future. It determines what to know about the future that helps to improve the fulfillment of the goals influenced by an optimization goal. It aims to increase the fulfillment chance of the goals with the yellow or red indicator state while to keep the goals with the green indicator state as they are. There are three main attributes: **focus** as a target object of the forward-looking insight, the **coordinate** as a perspective of the forward-looking insight over the focus, and **time window** as a time frame covered by the forward-looking insight. The **time window** attribute has three types of **short-term**, **mid-term**, and **long-term** whose time horizon can be defined by users, e.g., up to a month, a year, etc.

Proactive action represents an action triggered as a result of a forward-looking insight. It serves to be prepared or intervened in advance to increase the fulfillment chance of deficient goals.

3.2 Analytics design alternatives modelling

This modelling component aims to answer how to design a predictive analytics that supports a certain forward-looking insight. Here, the objective is to systematically capture the requirements that are central to designing a predictive analytics, based on the business analysis provided by the context modelling. Using the goal modelling part of the i^* framework, a goal model is constructed to explore the alternative ways to design predictive analytics. Each alternative is composed of a set of design requirements. To do so, we introduce a set of new concepts adapted from the i^* framework as applied to the domain of predictive analytics. These concepts are generally categorised into analytics goals, tasks, soft-goals, and analytics resources. An **analytics goal** represents the intention to extract insight from data. An **analytics task** is a set of computational steps necessary to achieve an analytics goal. An **analytics resource** is the data required to perform an analytics task. An **analytics soft-goal** [11] is a quality that should sufficiently hold when aiming to achieve an analytics goal or performing an analytics task, or using an analytics resource. A brief description of the new concepts is given below (some of the concepts such as **prediction top goal**, **algorithm**, and **constraint** are adopted from the approach proposed in [11]).

Prediction top goal is an analytics goal with emphasis on prediction. It represents an intention to forecast the future value of an object of interest in the business domain [11]. Each prediction top goal is assigned a **prediction type** as: **predicting a continuous attribute**, **predicting a sequence**, etc. This concept is visualised by the i^* notation to describe a goal with a horizontal line at the top.

Prediction strategy is a type of analytics goal that represents an intention to forecast the object of interest using another object. It aims to look for the

possible objects from the business domain or its environment that could relate to the object of interest to forecast. This concept is visualised by the i^* notation to describe a goal with one horizontal line at the top and one at the bottom.

Exploration task is an analytics task that addresses the procedure to fulfill a **prediction strategy**. It captures the steps required to derive the forecast of the object of interest from another object. This concept is visualised by the i^* notation to describe a task with a horizontal line at the top.

Computation task is an analytics task that represents the computation aspect of an **exploration task** by assigning a mining technique. Mining techniques include: **clustering**, **classification**, etc. This concept is visualised by the i^* notation to describe a task with one horizontal line at the top and one at the bottom.

Algorithm is an analytics task that represents how to perform a **computation task** using computer. Therefore, this concept addresses the computer understandability of the computation task. This concept is visualised by the i^* notation to describe a task with one horizontal line at the top and two at the bottom.

Variable is an analytics resource that deals with the data used to perform a **computation task** which can be **structured data** or **unstructured data**. Structured data could be coming from data warehouses which has **fact** and **dimension**. Unstructured data takes a **form** such as **text**, **audio**, etc. This concept is visualised by the i^* notation to describe a resource with a horizontal line at the top.

Constraint is an analytics soft-goal that represents business, computation or data limitations to deal with when designing a predictive analytics. This concept denotes a **constraint type** such as **data constraint** and **business constraint**.

3.3 Metamodel

The concepts introduced in the framework are described and the relationships between them are captured using a metamodel. The metamodel aims to provide an overall perspective over the concepts. The attributes captured in the classes embody the important information to collect for each concept. In Fig. 2 the metamodel is captured using a UML class diagram. In the figure, the concepts introduced for the Context modelling are distinguished from the one in the Analytics design alternatives modelling with a gray shadow on the background.

4 Illustrative example

In this section, we apply the framework on the motivating case. Figure 3 illustrates the **context modelling**. The purpose of this modelling is to facilitate the elicitation of the predictive analytics requirements that is responsive to the descriptive insight provided by indicators. It ensures that deficiencies in the goal fulfilment are addressed effectively. The context modelling uses the notions of **optimization goal**, **indicator state**, **business situation**, **proactive action** and **forward-looking insights** to align the fulfilment of business goals and the predictive analytics solution (new concepts are highlighted by purple in Fig. 3).

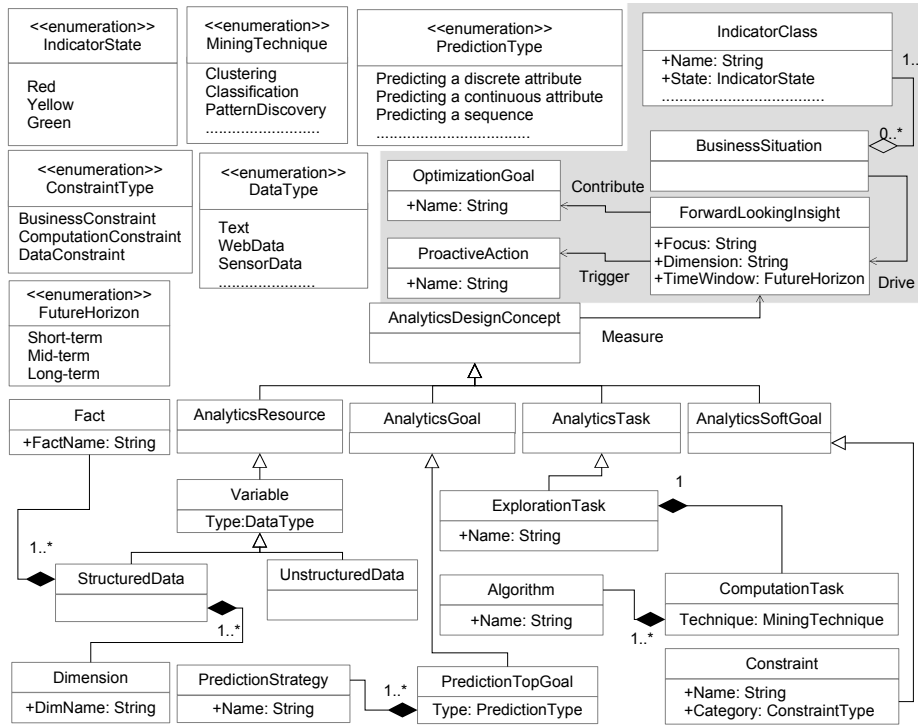


Fig. 2. Metamodel for concepts involved to adapt predictive analytics

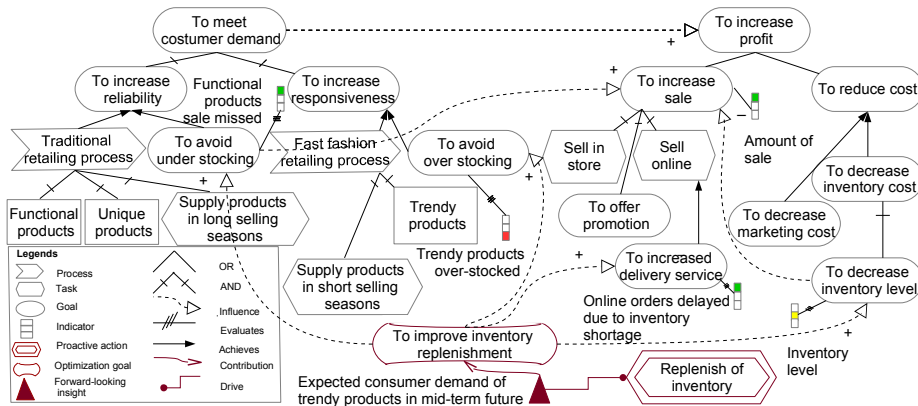


Fig. 3. An excerpt of the context modelling to improve the inventory replenishment.

The model in Fig.3 shows that *To improve inventory replenishment* is an example of an optimization goal which influences various business goals such as *To increase sale*, *To decrease inventory level*, *To avoid under stocking*, etc. Such goals are influenced goal which some of them might be in conflict such as

To increase sale and *To decrease inventory level*. At run time, the color of the indicator of an influenced goal visually expresses to which degree the goal has been fulfilled. Figure 3 represents the *Functional product sale missed* indicator as green which is an example of the **indicator state**. It means that the *To avoid under stocking* goal has been fulfilled. However, at the same time, *Trendy product over stocked* and *Inventory level* have the red and yellow that represents unfulfillment and partially fulfilment of associating goals, respectively. The combination of [green: *Functional products sale missed*, yellow: *Trendy product over stocked*, red: *Inventory level*] captures an example of a **business situation** that the retailer faces at the given time. Such business situation requires a **forward-looking insight** over demand in the context of trendy products which is replenished weekly.

Fig. 3 captures a forward-looking insight over *The expected consumer demand for trendy products in the mid-term future horizon* that helps to increase the chance of unfilled goals by triggering the right proactive action of *Replenish of inventory*¹. At some other point of time the business situation might change to the combination of [green: *Inventory level*, yellow: *Functional product sale missed*, red: *Amount of sale*] which is captured by BIM at run-time. This business situation motivates a different forward-looking insight of *The expected consumer demand for functional products in the long-term future horizon*.

By modelling **optimisation goal** and its influence to other goals, the areas that need support from analytics insights are represented. By capturing the **business situation** composed of **indicator states**, the connection between the predictive analytics and the descriptive insight provided by indicators are ensured. The **forward-looking insight** concept secures the relatedness and alignment of the predictive analytics to the current business situation in terms of goal fulfilment. It frames what is the subject of interest to be estimated by predictive analytics as a result of indicator state. Besides, the concept helps business users to make sense of predictive analytics within organisational context. This facilitates triggering the right **proactive action** that aims to increase the fulfilment chance of goals.

Figure 4 illustrates the **Analytics design alternatives modelling**. The purpose is to capture the alternative ways to design the analytics that supports the **forward-looking insight**. A goal model is developed for each insight that captures the design alternatives in terms of **prediction top goal**, **prediction strategies**, **exploration tasks**, **computation tasks algorithms**, **data variables**, and **constraints**.

In Fig. 4, *To forecast demand for trendy products in mid-term future horizon* is an example of a **prediction top goal**. To achieve this goal, there are various **prediction strategies** which specifies a domain-dependent way to estimate the demand for trendy products. Trendy products are new products for which there is no historical sale data. Therefore the estimation of demand might rely on weather condition, user's comments on the social media, a pilot sale of the new product. Some of these factors such as pilot sale are under the control of the business while some others such as weather condition are not. In this regard, the

¹ The concepts of **forward-looking insight** and **proactive actions** captured in BIM will have the temporal attribute of expiration at run time. The instances of these concepts will be automatically expired and replaced when the business situation change.

prediction top goal is decomposed into *To forecast based on controlled factors* and *To forecast based on uncontrolled factors* which are the examples of a prediction strategy to take. The model in Fig. 4 illustrates the further decomposition of prediction strategies. There are various domain-dependent ways to accomplish *To forecast based on controlled factors*. For example, *To forecast based on pre-sale test* is a way that relates the demand for the trendy products to the sale data of the pilot sales. To achieve the latter prediction strategy, the retailer selects some stores and sells the new trendy products for a short period to test how consumers respond to the new products. Then, the sales of the test period is being used to predict the demand across all stores.

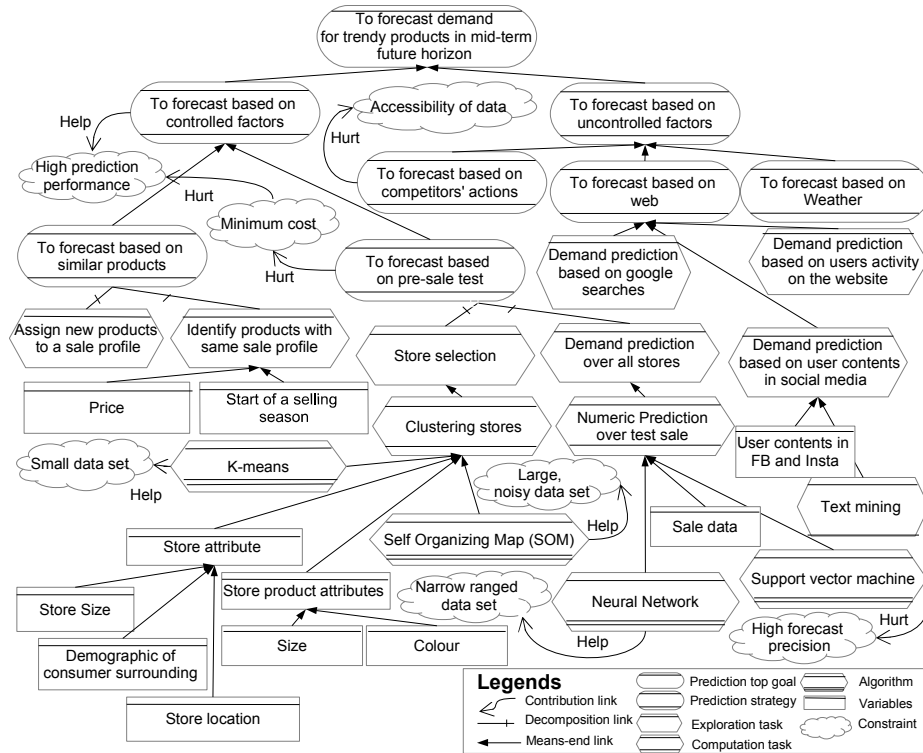


Fig. 4. A portion of the Analytics design alternatives modelling

The model in Fig. 4 shows that *To forecast based on pre-sale test* requires to take two courses of actions *Select stores* and then *Forecast demand over all stores*². These tasks are the examples of a **exploration task** that is necessary to accomplish a prediction strategy. The tasks address the steps of a procedure to drive the prediction of the object of interest from other objects in the business

² To capture the order of analytics tasks, readers can refer to proposals in the literature to deal with the temporal aspects of goal models.

domain or its environment. To accomplish *Select stores*, a **computation task** needs to be applied which is *Cluster stores*. There are also computer understandable algorithms to accomplish a computation task. For example, *K-Means* is an algorithm to address the *Cluster stores*. We remark that **computation task**, and **algorithm** do not necessarily address forecasting. For example, *Select stores* is only a step to fulfill a prediction strategy that forecast the demand for trendy products based on the pre-sale technique. It should not be mistaken that *Select stores*, alone, represents a task to perform the forecast. The same case applied to **computation task**, and **algorithm**.

Figure 4 also shows that the *Store selection*, as a **computation task**, can be either performed over the *Store attributes* including *Store location*, *Store size*, *etc* or over the *Store products attributes* such as *Size* and *Colour*. These are the examples of **variables** to use in a predictive analytics. Figure 4 also represents some of the **constraints** to fulfill a goal or perform a task or use a resource in the goal model. For example, *Low-cost prediction* is a constraint to take into account to design a predictive analytics that supports a certain **forward-looking insight**. The constraints guide to choosing a certain alternative among analytics goals, tasks, or resources. For example, *Limited accessibility to data* might be considered to choose the alternative of *To forecast based on competitor's action* when the competitor's information is not available.

By modelling **prediction top goal**, the objective of a forecast to support a forward-looking insight is specified. A prediction top goal connects the analytics design concepts with the business concepts captured in BIM. Modelling the **prediction strategy** ensures that all of the possible drivers in the business environment are taken into accounts when designing a predictive analytics. The **exploration task** makes it explicit how to relate and drive the forecast of the object of interest from another object. The **computation task** clarifies the quantitative aspect of relating the object of interest to forecast to other objects using a certain data mining technique. The quantitative aspect of the forecast is further clarified by the **algorithms** that are understandable by computers. Modelling the **variables** gives the possibility to explore the data set where an algorithm can be performed on to run the forecast. It makes sure that the data asset are used effectively. Modelling the **constraints** simplify to make analytics design choices.

5 Analysis with the proposed framework

The proposed framework suggests five type of analyses which support the framework from a methodological perspective. Each analysis provides a set of steps required to systematically build the models discussed in Sect. 3.

The **Insight analysis** aims to drive a forward-looking insight from a descriptive insight provided by the indicators in BIM. It systematically captures the **Context modelling** by eliciting the concepts of **optimisation goal**, **indicator state**, **business situation**, **forward-looking insight**, and **proactive action**. The **Analytics goal analysis**, **Analytics task analysis**, **Analytics resource analysis**, and **Analytics soft-goal analysis** aim to design a predictive analytics that supports the forward-looking

insight. The Analytics goal analysis systematically elicits and captures the prediction top goal and prediction strategy concepts in the Analytics design alternatives modelling. The Analytics task analysis elicits the tasks required to accomplish the analytics goals. It systematically captures the exploration task, computation task, and algorithm concepts in the Analytics design alternatives modelling. The Analytics resource analysis elicits the analytics resources to perform analytics tasks by systematically capturing the variable concept in the Analytics design alternatives modelling. The Analytics soft-goal analysis elicits the analytics soft-goals by systematically capturing the constraints concepts in the Analytics design alternatives modelling. The analyses are described as follows:

(1) Insight analysis: Identify an optimization goal and represent it using the new notation in the BIM schema developed for monitoring purposes. Then, identify which business goals of BIM are influenced by the optimisation goal and name these as influenced goals. Next, link the optimization goal with the corresponding influenced goals by means of a influence link of BIM. After that, identify the meaningful business situations. To do so, analyse the indicators associated with the influenced goals at the run-time. Determine which meaningful combination of color over various indicators in a snapshot of BIM would shape a meaningful business situation that needs an improvement. Various combinations of indicator states constitute various business situations. Afterwards, translate each business situation into a forward-looking insight that aims to keep the green states of the influenced goal as they are and improve the yellow and red states to green. Furthermore, define precisely the insight by specifying the attributes of focus, dimension, and time window (indicators with a red or yellow state can provide a guide). Next, capture the forward-looking insight in the BIM with the new notation introduced and link it with the optimization goal using a contribution link of the BIM language. Identify the proactive action triggered by the forward-looking insight and capture it in BIM. Also connect the proactive action to the forward-looking insight using the new link drive.

(2) Analytics goal analysis: Start by developing a goal model for each forward-looking insight identified in the insight analysis. To do so, represent a prediction top goal that addresses an intention to design a predictive analytics in order to support the forward-looking insight. Define the prediction top goal based on the attributes of the forward-looking insight. Then, identify prediction type by denoting the predicting a discrete variable or predicting a continues variable etc. Next, decompose the prediction top goal into prediction strategies. Determine the strategies to drive the forecast from the objects in the business domain or its environment. Decompose prediction strategies till a certain object is determined.

(3) Analytics task analysis: Analyse the lowest level goal of the prediction strategies in order to identify the tasks to achieve them. Determine the exploration tasks which relates the object of interest to forecast to an object in the business or its environment to drive the predictive analytics. Link the exploration tasks to the corresponding prediction strategy using decomposition links or means-end links. Next, decompose each task until it is not possible to address the computation aspect of the task by a mining technique. Then, capture a computation

task that addresses the mining technique and link it to the corresponding exploration task using a means-end link. Afterwards, capture the alternatives of machine understandable **algorithms** to perform a computation task. Link the algorithms to the corresponding computation tasks with a means-end link. Refer to data mining catalogue introduced in [11] to see how to assign algorithms to a computation task. While exploration tasks and computation tasks are defined at business level, the algorithms address a generic concept in the advanced analytics domain. It is important to note that exploration task, computation task, and algorithm, individually, do not necessarily address forecasting. Instead, they represent a task in a prediction procedure that aims to forecast the target object.

(4) Analytics resource analysis: Analyse the exploration task in order to find a relevant data set to perform the computation task. Capture the **variables** representing the data set and link the variable with a means-end links to the corresponding exploration task. Next, decompose variables if it is necessary. Also, determine the attribute of **fact** and **dimensions** if the variables are accessed from a data warehouse repository. Determine the **form** of the data set if the variables are accessed from a source of unstructured data by denoting **audio**, **text**, etc.

(5) Analytics soft-goal analysis: Identify the **constraints** that should be taken into account when designing a predictive analytics. Capture the **constraints** in the goal model. Identify analytics goals, tasks, and resource that contribute to each **constraint**³. Link the analytics goals, analytics tasks or analytics resources to constraints that contribute to satisfy the **constraints** by means of the contribution links of the i* framework. Next, link the constraints among each other when satisfying one affects the other. Such links guide (or restrict) the selection among alternatives of analytics goals, tasks, and resources captured in the goal model. Next determine the attribute of **constrain type** by specifying **business constraint**, **data constraint**, or **computation constraint**.

6 Related work

Conceptual modelling research plays a critical role in addressing challenges of designing and implementing data-intensive solutions in business domains [15]. The work in [11] proposes a modelling framework for requirements analysis and design of business analytics solutions. It captures analytics requirement in terms of business goals, decision goals and questions goals and connects them to advanced analytics system design. The main difference of this work is that the predictive analytics requirements are derived from descriptive analytics sensed from **indicators states**, whereas in [11] requirements derived from business goals. It is also different in that it focuses on the predictive type of analytics and captures requirements in terms of **forward-looking insight**. The approach in this paper can take advantage of the algorithms catalogue in [11] to support the analyses. In [16] authors extend the notion of indicators into measures, key performance

³ Satisfied means sufficiently satisfied and it is taken from the definition of the soft goal in the i* framework.

indicators, and key result indicators and use that to drive data mining (e.g., time series analysis). The work in this paper is different as it captures the analytics solutions design and associated trade-offs.

Previous research also provides modelling support for design and development of BI systems. The work in [17] uses a goal-oriented approach to derive the analytics requirements as a form of reports demanded by business users and then defines the relation of the analytics component with other BI components such as data warehouse, indicators, data sources, and ETL at macro level. There are frameworks [9, 10, 18] that discuss indicators as a BI component used to track enterprise performance. These approaches model indicators as a BI components in relation to business concepts, such as goals, processes, and roles in both formal [18] and semi-formal [9, 10] ways. Indicators in their approach are captured to provide a descriptive insight over the business goal fulfillment. The framework in this paper extends the BIM language [9] with forward-looking insight to reveal predictive analytics requirements and connects them to a solution design.

Another body of related work proposes conceptual modelling approaches for requirements design of data warehouses [19]. For example, the work in [7] proposes GRANd as a goal-oriented approach for requirement analysis of data warehouses. The work in [20] proposes the Goal-Decision-Information (GDI) model from requirements analysis of data warehouses. The work in this paper is different in that it aims to support designing the predictive analytics solutions that generate forward-looking insights, as opposed to descriptive analytics.

7 Conclusion and future work

In this paper, we presented a conceptual modelling framework to adopt predictive analytics in organisations. In the future, we plan to enrich the presented framework to cover all the steps of the design science methodology [21] including: problem motivation, objectives of a solution, design and development, demonstration, evaluation, and communication. We have already represented the business and technical challenges of adopting predictive analytics in organisations. The objective is to develop a dedicated requirements engineering (RE) framework by which a predictive analytics is effectively and efficiently adopted to enhance the organisational performance. The design artefact which we represented in this paper is a model-based RE framework demonstrated with the help of an inventory problem. The in-depth evaluation of the framework is going to be represented in the future work. For the evaluation, we are conducting a case study in which business users from various domains are utilizing the proposed framework. Besides, the design and development step of the design science methodology which involves creating the artefact is still an ongoing work. In this step, we are extending the framework to quantitatively measure the influence of the predictive analytics on the business goals. Moreover, we are also precisely representing the temporal aspect of the BIM schema at the run time. We also plan to develop tools that support the framework.

References

1. I. Vaisman and E. Zimányi. *DW Systems: Design and Implementation*. 2014.
2. R. Sharda, D. Delen, and E. Turban. *Business Intelligence: A managerial perspective on analytics*. Prentice Hall Press, 2013.
3. J. R. Evans. An exploratory study of performance measurement and relationships with performance results. *Journal of Operations Management*, 22(3):219–232, 2004.
4. D. Parmenter. *Key performance indicators: developing, implementing, and using winning KPIs*. John Wiley & Sons, 2015.
5. T.W. Oestreich. Gartner magic quadrant for BI and analytics platforms. 2016.
6. V. C. Storey and I. Song. Big data technologies and management: What conceptual modeling can do. *Data & Knowledge Engineering*, pages 50–68, 2017.
7. P. Giorgini, S. Rizzi, and M. Garzetti. Grand: A goal-oriented approach to requirement analysis in data warehouses. *Decision Support Systems*, 45(1):4–21, 2008.
8. N. Tryfona, F. Busborg, and Borch C. starER: A conceptual model for DW design. In *Proc. of Int. Workshop on Data warehousing and OLAP*, pages 3–8, 1999.
9. J. Horkoff, D. Barone, L. Jiang, E. Yu, D. Amyot, A. Borgida, and J. Mylopoulos. Strategic business modeling: representation and reasoning. *Software & Systems Modeling*, 13(3):1015–1041, 2014.
10. A. Pourshahid, D. Amyot, L. Peyton, S. Ghanavati, P. Chen, M. Weiss, and A. Forster. Business process management with the user requirements notation. *Electronic Commerce Research*, 9(4):269–316, 2009.
11. S. Nalchigar, E. Yu, and R. Ramani. A conceptual modeling framework for business analytics. In *Proc. of Int. Conf. on Conceptual Modeling*, pages 35–49, 2016.
12. S. Nalchigar and E. Yu. Conceptual modeling for business analytics: A framework and potential benefits. In *Proc. of Int. Conf. on Business Informatics (CBI)*, volume 1, pages 369–378. IEEE, 2017.
13. E Yu. *Modelling strategic relationships for process reengineering*. 2011.
14. M. Nenni, L. Giustiniano, and L. Pirolo. Demand forecasting in fashion industry: a review. *International Journal of Engineering Business Management*, 5:37, 2013.
15. V. Storey, J. Trujillo, and S. Liddle. Research on conceptual modeling: Themes, topics, and special issues. *Data Knowledge Engineering*, 98:1–7, 2015.
16. A. Maté, J. Trujillo, and J. Mylopoulos. Key performance indicator elicitation and selection through conceptual modelling. In *Proc. of Int. Conf. on Conceptual Modeling*, pages 73–80. Springer, 2016.
17. C. Burnay, I. J. Jureta, I. Linden, and S. Faulkner. A framework for the operationalization of monitoring in business intelligence requirements engineering. *Software & Systems Modeling*, 15(2):531–552, 2016.
18. V. Popova and A. Sharpanskykh. Modeling organizational performance indicators. *Information systems*, 35(4):505–527, 2010.
19. A. Nasiri, R. Wrembel, and E. Zimányi. Model-based requirements engineering for data warehouses: From multidimensional modelling to KPI monitoring. In *Proc. of Int. Workshop on Conceptual Modeling*, pages 198–209. Springer, 2015.
20. N. Prakash and A. Gosain. Requirements driven data warehouse development. In *roc. of Int. Conf. on CAiSE*, volume 252, 2003.
21. K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee. A design science research methodology for information systems research. *Journal of management information systems*, 24(3):45–77, 2007.