Automated Prediction of Relevant Key Performance Indicators for Organizations

Ünal Aksu[1,2(✉)], Dennis M. M. Schunselaar[2], and Hajo A. Reijers[1,2]

[1] Utrecht University, Utrecht, The Netherlands
{u.aksu,h.a.reijers}@uu.nl
[2] Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
d.m.m.schunselaar@vu.nl

Abstract. Organizations utilize Key Performance Indicators (KPIs) to monitor whether they attain their goals. For this, software vendors offer predefined KPIs in their enterprise software. However, the predefined KPIs will not be relevant for all organizations due to the varying needs of them. Therefore, software vendors spend significant efforts on offering relevant KPIs. That relevance determination process is time-consuming and costly. We show that the relevance of KPIs may be tied to the specific properties of organizations, e.g., domain and size. In this context, we present our novel approach for the automated prediction of which KPIs are relevant for organizations. We implemented our approach and evaluated its prediction quality in an industrial setting.

Keywords: Key Performance Indicators · Prediction · Relevance

1 Introduction

Organizations measure the performance of their business processes to determine whether they attain their goals. As a means for that, Key Performance Indicators (KPIs) are used [20]. *Average duration of product delivery* is a KPI that organizations use to monitor their product delivery processes. By tracking this KPI, organizations can predict how much staff must be assigned to their product delivery processes to keep the duration of a product delivery below a certain threshold, e.g., on average 3 days.

To support organizations in process performance measurement, software vendors offer predefined KPIs in their software products. With this, they aim to provide the maximal set of KPIs that may be relevant for most organizations. However, predefined KPIs will not work successfully in all organizations because they want relevant KPIs aligned to their specific goals [20]. For example, *taken*

Supported by the NWO AMUSE project (628.006.001): a collaboration between Vrije Universiteit Amsterdam, Utrecht University, and AFAS Software in the Netherlands. This work is a result of the AMUSE project. See amuse-project.org for more information.

© Springer Nature Switzerland AG 2019
W. Abramowicz and R. Corchuelo (Eds.): BIS 2019, LNBIP 353, pp. 283–299, 2019.
https://doi.org/10.1007/978-3-030-20485-3_22

leave per day is a relevant KPI for a production organization, whereas it may not be relevant for a university, which has similar number of employees. For this reason, software vendors include Business Intelligence (BI) functionality into their software products and let organizations develop custom KPIs. Although organizations may do this, custom development of KPIs still requires a significant effort both from software vendors and organizations [20].

Numerous studies have been conducted for determining relevant KPIs for organizations [3,4,14,15,28]. In these studies, relevant KPIs are either defined from scratch or selected from a set of KPIs (e.g., a KPI library) for each organization. Moreover, in these studies, the identified reasons that make certain KPIs relevant for one organization are not usually reusable at determining the KPIs for another. Therefore, for current approaches tailoring KPIs is a manual endeavor that needs to be repeated for each organization, and requires a significant effort both from software vendors and organizations.

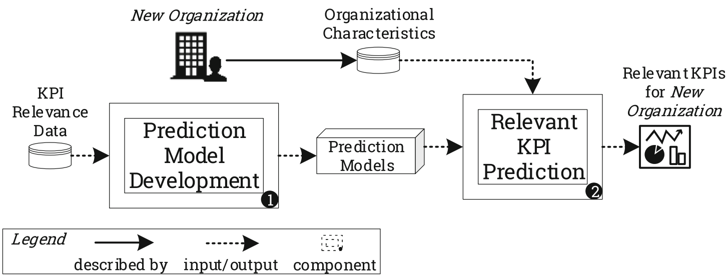

Fig. 1. Our approach for predicting relevant KPIs for organizations

Within this paper, we propose a novel approach for the automated prediction of relevant KPIs for organizations. The approach takes a set of prediction models aimed at predicting the relevance of KPIs and the characteristics of a new organization, e.g., domain, location, and number of employees. By checking the given organizational characteristics of that organization against the relevance factors of KPIs encoded in the prediction models, the approach predicts which KPIs are relevant to the organization (see ❷ in Fig. 1). To determine the relevance factors of KPIs and develop prediction models (see ❶ in Fig. 1), the approach uses the known relevance values of a set of KPIs for a number of organizations, which needs to be given in the form of a specific input, KPI Relevance Data. By means of the automatically determined relevance factors of KPIs, we automate the prediction of relevant KPIs for organizations, which is manually repeated for every organization in current approaches. Thus, our approach sets itself apart from the state of the art. We evaluate the prediction quality of our approach by applying it in a real-life setting at a Dutch ERP software vendor. In this context, we discuss the results that we obtained.

In Sect. 2, we present our approach aimed at the automated prediction of relevant KPIs for organizations. The details of the implementation of the approach are given in Sect. 3. Afterwards, in Sect. 4, we evaluate the prediction quality of our approach by applying it in a real-life, industrial setting and present the results obtained in the application. Section 5 is devoted to the discussions on the implications of the obtained results. In Sect. 6, we provide an overview of related work on providing relevant KPIs to organizations. Finally, we present our conclusions and potential directions for future work in Sect. 7.

2 Approach

In this section, we explain the details of our approach on the automated prediction of relevant KPIs for organizations. As introduced in Sect. 1, there are two tasks: predicting relevant KPIs and developing prediction models. They are taken care of by the components Prediction Model Development and Relevant KPI Prediction. The former uses prediction models and the organizational characteristics of a new organization as inputs; the latter uses KPI Relevance Data as the only input. For the sake of simplicity, first, the definitions of organizational characteristics, prediction models, and KPI Relevance Data are listed below. Then, the details of each component is given.

Definition 1 ***Organizational Characteristics*** *contain the values of a set of characteristics (e.g., domain, location, and number of employees) by which organizations can be characterized. For example, Organization o1* = {*domain* = *Retail* ∧ *location* = *Amsterdam* ∧ *numberOfEmployees*= *[10–19]* ∧ *doesExport* = *Yes* ∧ *industryClassification-MainGroup* = *47* ∧ *industryClassification-SubGroup* = *8109* }.

Definition 2 ***Prediction Models*** *are aimed at predicting the relevance of KPIs. Each prediction model encodes a KPI, the factors that are the determinants to what extent the KPI will be relevant for organizations, and a prediction modeling technique, which outperforms predicting the relevance value of the KPI for organizations using those relevant factors.*

Definition 3 ***KPI Relevance Data*** *is a* 2*-tuple: (1) the relevance values of a set of KPIs for a number of organizations where that KPI set is considered as the comprehensive set from which a sub-set will be selected and (2) the key characteristics of these organizations with their values, i.e., Organizational Characteristics. For example, in Fig. 3, an excerpt from a sample KPI Relevance Data is depicted.*

In our approach, the relevance value of a KPI can be a numeric value from a scale, namely KPI Relevance Scale. As the KPI Relevance Scale, we use a five-points Likert-type scale: $[1, 5]$, where a higher value denotes a higher relevance. The reason for using a five-points scale is that it has been recommended by many researchers [5,10,27] as the optimal number of relevance categories.

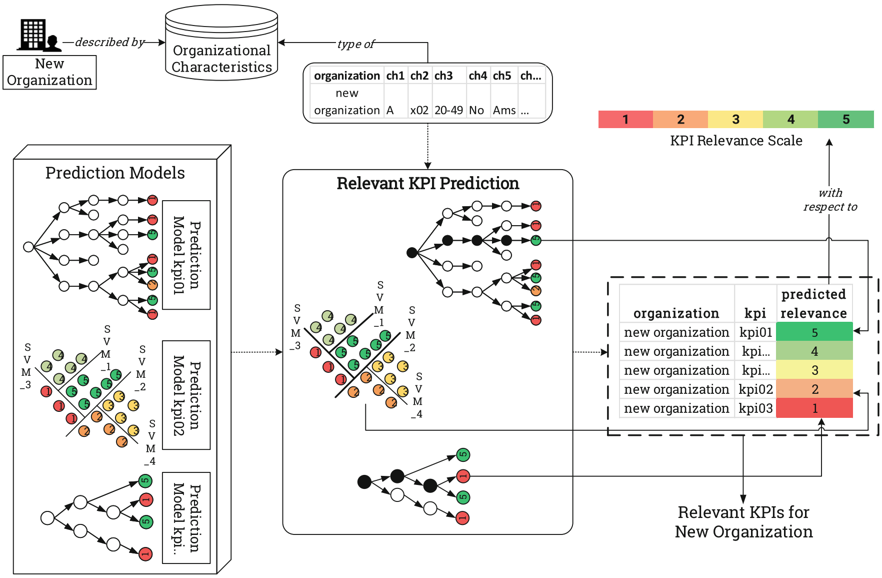

Fig. 2. Predicting relevant KPIs for a new organization

Relevant KPI Prediction: To predict which KPIs are relevant for a new organization, two inputs are required: the organizational characteristics of that organization and prediction models. The prediction modeling technique encoded in each prediction model is executed with the given organizational characteristics. Thus, a predicted relevance value will be obtained for the KPI. In the output, the obtained relevance values are sorted from highest to lowest. Afterwards, the KPIs that have the highest predicted relevance value, a value of 5 in the KPI Relevance Scale used in our approach (see in Fig. 2), are marked as the set of relevant KPIs for the new organization. However, this marking is flexible and one can say that a value of either 4 or 5 may be presented as the set of relevant KPIs for the new organization. For this, to what extent a KPI is used for making decisions about the related business process in an organization may be a reason.

Prediction Model Development: This component takes KPI Relevance Data as input. For each KPI in the input, an analysis task is performed to determine what organizational characteristics are the determinants of the relevance value of a KPI for organizations. The reason for performing the task per KPI is that relevance factors may vary from one KPI to another. For example, "number of employees" may be the only factor that makes a KPI relevant for organizations, whereas for those organizations the relevance of another KPI may be dependent on both "number of employees" and "organization type", e.g., whether it is a non-profit organization.

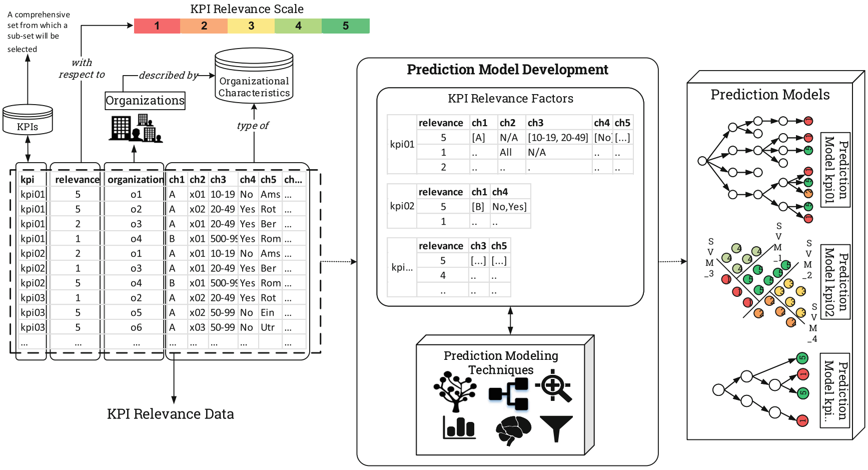

Fig. 3. Creation of the prediction models for predicting relevant KPIs

Since the organizational characteristics in the given KPI Relevance Data are raw data, they need to be transformed into features to better represent the underlying patterns in the given KPI Relevance Data. That transformation is done by the encoders within this component. More specifically, a one-hot encoder is used for each feature. Then, a feature subset is selected for each prediction modeling technique employed within the component–employed techniques are listed in the implementation documentation of the approach[1]. This feature subset selection helps to make sense of the features for prediction modeling techniques. To keep the best performing features in subsets, the worst performing feature at each iteration is eliminated, and then the dependencies between features are uncorrelated by a dimensionality reduction.

Afterwards, the component trains and tests each prediction modeling technique to find out the best performing prediction modeling technique for each KPI. The reason for that is a prediction modeling technique may not outperform for predicting the relevance values of all KPIs since the relevance values in a given KPI Relevance Data may not be the same for all KPIs. Moreover, each prediction modeling technique has its own noise handling mechanism. For example, while Random Forest may be the best for an imbalanced set of relevance values, other prediction modeling techniques, e.g., Ada Boost may perform poorly. While training a prediction modeling technique, the component chooses a set of appropriate hyperparameters to discover the parameters that may result in more accurate predictions. To do so, the component uses a cross-validated grid-search algorithm. As a result of the train and test, the component identifies the best performing prediction modeling technique at finding the relevance factors of each KPI. For this, the balanced accuracy metric [9,25] is used.

[1] The implementation of our Automated Relevant KPI Determination Approach is available at http://amuse-project.org/software/.

By doing so, we aim to deal with the relevance values of KPIs that may have an imbalanced distribution in a given KPI Relevance Data. When the relevance factors and best performing prediction modeling techniques are determined for all KPIs, the component creates the prediction models for the KPIs. In particular, the relevance factors of a KPI, the selected prediction modeling technique for identifying them, and the KPI itself are encoded in the form of a prediction model.

In the next section, we give the details of the implementation of the approach.

3 Implementation

In this section, we give the details of the implementation (see Footnote 1) of our approach. On the one hand, the implementation is a constructive proof of the approach. On the other hand, it shows the applicability of the approach. In Fig. 4, the technical details of the implementation are depicted.

As explained, to predict relevant KPIs, the approach requires prediction models as input. This is taken care of the Prediction Model Development component within the approach. It takes KPI Relevance Data as input and develops prediction models. To accurately capture the knowledge in the given KPI Relevance Data, as shown in Fig. 4, a nested (two-level) stratified cross validation is used: (1) for all KPIs and (2) per KPI. More specifically, both model development and testing will be carried out n-times, which is specified in each stratified cross validation block, using a different sample dataset of the given KPI Relevance Data. By doing so, we aim to develop the prediction models that both capture the patterns in the given KPI Relevance Data, but also generalizes well to unseen organizational characteristics of new organizations.

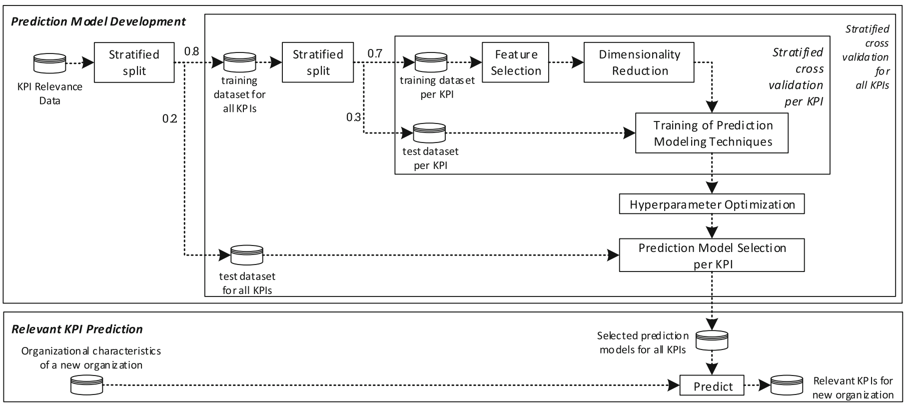

Fig. 4. Technical details of the implementation of our approach

Within the *stratified cross validation block for all KPIs*, prediction model development and testing for all KPIs will be done in 5-folds. This means that in each fold, an 80%/20% stratified split [18] is done to divide the given input into training and test datasets. By doing so, the approach can develop prediction models and test them using the sample datasets that are preserving the percentage of the data points for each KPI and class (i.e., relevance values). As a result, the approach creates two different dataset in each fold: training dataset for all KPIs and test dataset for all KPIs. Similarly, within the *stratified cross validation block per KPI*, the approach does the prediction model development and testing for each KPI in 5-folds. A training dataset per KPI and a test dataset per KPI will be generated in each fold. The aforementioned prediction modeling techniques will be trained and tested using these datasets. In order to avoid over-fitting to the training data, a 70%/30% split is preferred [12].

The approach utilizes feature selection and dimensionality reduction [19] to select the organizational characteristics in the training dataset per KPI that contribute most to the prediction variable (the relevance value of a KPI). Then, the approach applies feature scaling to have standardized range of the values of the selected organizational characteristics. By doing so, we aim to make that each selected organizational characteristic may equally influence the prediction variable.

Afterwards, the approach trains each prediction modeling technique contained in it. Meanwhile, the approach tunes the parameters of each prediction modeling technique to determine the parameter set with which each trained prediction modeling technique performs best. When all prediction modeling techniques are trained, the approach tests them using test dataset per KPI mentioned above. Using the balanced accuracy metric, the approach selects the best performing prediction model for a KPI, from the set of the prediction models that are created in all folds of the *cross-validation block per KPI* and tested. When the best performing prediction model for each KPI is selected, then the approach completes the prediction model generation process. In other words, the selected prediction models are ready for predicting relevant KPIs for organizations.

To achieve a high quality at predicting relevant KPIs for organizations, while developing prediction models within the approach, we use 3 types of meta-algorithms: stacking, boosting, and bagging [29]. In stacking, a meta-technique tries to learn the best combination of the prediction models of the primary prediction modeling techniques, which are combined as a stack. In boosting, the same prediction modeling technique is applied in a chain to learn and fix the prediction errors of prior prediction models developed in the chain. Different sub-samples of the training dataset are taken and multiple prediction models are generated in bagging. Then, these models are aggregated to form a final prediction model, which has a better accuracy value.

Since most properties of organizations are categorical data types and the scale we used for relevance values (KPI Relevance Scale) has multiple points, it is required to support multi-class classification [12] prediction modeling techniques from the machine learning discipline within our approach. Accordingly, our approach employs the prediction modeling techniques that are listed in the implementation documentation of the approach (see Footnote 1). However, our approach is flexible to support continuous (numeric) data types. This can be indicated in the configuration where the approach learns the organizational characteristics contained in a given KPI Relevance Data. Moreover, our approach is extensible to support regression prediction modeling techniques in the case that one may want to predict a decimal value for the relevance value of KPIs instead of a numeric value from a KPI Relevance Scale.

In addition, to obtain better predictive performance, the approach reduces the problem of multi-class classification to multiple binary classification problems while developing prediction models. In this regard, we apply the following strategies: one-vs-rest and one-vs-one [12]. The former involves training a single prediction modeling technique per class, whereas in the latter a particular prediction modeling technique is trained for each different pair of classes.

In the following section, we describe how we evaluate our approach in a practical use of its implementation.

4 Evaluation

In this section, we demonstrate the use of the proposed approach in an industrial setting and evaluate how accurately it predicts relevant KPIs for organizations. In this regard, in the following subsections, we describe how we develop prediction models and use them at predicting relevant KPIs for organizations in a case study.

4.1 Data Collection

In order to develop prediction models, the approach requires KPI Relevance Data. However, although software vendors usually know the key characteristics of the organizations that they deliver their software products to, they are typically not aware of the relevance of the KPIs that they offer to those organizations. Therefore, we investigated whether we could identify a proxy for this type of data. KPI Usage Logs are typical data sources in which software vendors typically keep track of how KPIs are being used by organizations. In general, software vendors either record these logs using the software product in which they offer KPIs for organizations or using a third-party BI tool (e.g., Qlik Sense and Microsoft Power BI), which they use as a means for enabling organizations to develop custom KPIs. KPI Usage Logs is a data source from which one can obtain information on the usage of KPIs. For example, how many times a particular KPI is used in an organization, when that KPI is used, and how much time has spent using the KPI can be obtained from KPI Usage Logs. The obtained information on the usage of KPIs can be seen as the interest of organizations in KPIs. Moreover, one can interpret the interest of organizations in KPIs as the relevance values of KPIs for those organizations. Thus, as a primary proxy for known relevance values of KPIs for organizations, KPI Usage Logs are determined.

A Dutch ERP software vendor, the case study company, records KPI Usage Logs for the KPIs that the company offers to its customers. In the company, we had a training session on the KPIs, which are offered to its customers within its ERP software product. In particular, we examined the KPIs in the Human Resource Management (HRM) area. The reason for that is that human resources form a key asset in any organization. As such, the availability of the employees in an organization is essential in performing its business processes to attain its goals. However, due to various reasons, employees may not be always available, for example, sickness, injury, maternity leave, or vacations. Absence and leave are the two sub areas in HRM that concern the unavailability of employees in organizations. The former deals with the unexpected reasons of unavailability of employees. The latter focuses on the unavailability of employees resulting from statutory rights as granted by labour laws. In this regard, together with two experts who manage the KPIs in the company, we selected 13 KPIs from the absence and 6 KPIs from the leave sub area. While selecting the KPIs, our main consideration was the wide usage of the KPIs by organizations to get sufficient data points such that our approach can predict relevant KPIs for a new organization accurately. The selected 19 KPIs are commonly used by more than 2000 client organizations of the software vendor. Afterwards, the experts defined a set of metrics for transforming the usage of the KPIs into the relevance values of them. Since the defined metrics require a minimum of one year usage of the KPIs, the relevance values of the selected 19 KPIs are obtained for approximately 1100 organizations, which use those KPIs at least for a year.

In addition to the obtained relevance values, the characteristics of those 1100 organizations were required to create KPI Relevance Data and develop the prediction models for the selected KPIs. To determine which characteristics and their values for those organizations are available in the company, we arranged three meetings with various experts. The first meeting was with the following experts: a director–the CIO (Chief Information Officer) of the company–who has knowledge on scoping the KPIs offered to the organizations and a senior product developer who is an expert on designing and developing KPIs. These experts explained the characteristics of the organizations that they consider while scoping and developing the KPIs offered to the organizations. A marketing manager participated in the second meeting and described the characteristics of the organizations that often request adjustments for the offered KPIs. In the last meeting, together with a product manager, we analyzed the data about the organizations to identify what characteristics and their values are available within the company. As a result, we selected the characteristics shown in Table 1. All these characteristics are categorical and the values of them were available for 750 out of the aforeselected 1100 organizations. Moreover, the names of these characteristics are translated from their original definitions in Dutch.

Table 1. Selected characteristics of organizations

Characteristic	Explanation with some example values
Legal form - Main group	The main group of the legal form of an organization, e.g., with or without legal entity
Legal form	The legal form of an organization, e.g. private limited company, foundation or association
Non-profit	A non-profit organization uses the money it earns to help people. However, a profit organization invests the money it earns on developing new products or services to sell them and make more money
Industry classification - Main group	The main group to which the organization is assigned by the Chamber of Commerce within the Netherlands. For example, construction is the main group to which general civil and utility construction organizations are assigned. Transportation and storage is another example for main group
Industry classification - Sub group	The sub group to which the organization is assigned by the Chamber of Commerce within the Netherlands. Construction of residential buildings and construction of railways are two example sub groups of the construction main group. Similarly, passenger transport and freight transport are two example sub groups of the transport and storage main group
Province	The province where the organization is registered
Number of - Employee range	The range of the total number of employees in the organization. For example, 10–19, 20–49, and 50–99
Import	Whether the organization does import
Export	Whether the organization does export
Has subsidiary organizations	Whether the organization has subsidiary organizations

By combining the obtained relevance values of the selected KPIs 19 with the organizational characteristics of the aforementioned 750 organizations, we created the KPI Relevance Data for our evaluation. This means that the required input for developing the prediction models for the KPIs is ready. Accordingly, the approach created 19 prediction models for predicting the relevance of the selected 19 KPIs.

4.2 Applying the Automated Relevant KPI Determination Approach

We predicted relevant KPIs for a set of organizations using the developed prediction models. Since the developed prediction models are for the KPIs in the absence and leave sub areas, the set of the organizations for which relevant KPIs are predicted are accordingly selected. In particular, the organizations are selected from the client organizations of the case study company that are not purchased and not use to the selected 19 KPIs, but use the related functionalities, i.e., absence and leave within the ERP software product of the company. As a result, we selected 261 organizations and predicted relevant KPIs for them.

To determine the prediction accuracy of our approach in the case study, we collaborated with the CIO of the company and a senior product manager in the case study company. The reasons for collaborating with these two experts is that these experts have extensive knowledge both on the organizations for which relevant KPIs are predicted and on the expected relevance values of the selected KPIs for those organizations. Then, we calculated the prediction accuracy of our approach by comparing the predicted relevance values of the KPIs for the organizations with the expected relevance values of the KPIs for these organizations, which are provided by the aforementioned two experts. In Fig. 5, the prediction accuracy of our approach in the case study is depicted.

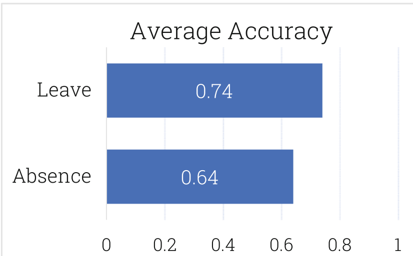

Fig. 5. Prediction accuracy of our approach in the case study

As depicted in Fig. 5, in the case study, our approach achieves a 74% balanced accuracy at predicting the relevance of 6 KPIs in the leave sub area for 261 organizations. Similarly, a 64% balanced accuracy is achieved at predicting the relevance of 13 KPIs in the absence sub area for the same organizations. The weighted average of the prediction quality values for these two sub areas will show the prediction quality of our approach for the HRM area, which is 67%. In the following section, we discuss the implications of those results.

5 Discussion

To consider a certain prediction quality as good, one should look into the context of the application where that quality is measured [8,23,26]. In this regard, we discuss the prediction accuracy of our approach. Since no other study has so far focused on the automated prediction of relevant KPIs for organizations, there is no exact reference to compare the prediction quality of our approach with. However, we think that the prediction quality of our approach shown in Fig. 5 is reasonable [26].

As mentioned in Sect. 3, the problem that our approach tries to solve is a multi-class classification. Having a balanced distribution of each class has a significant effect on prediction quality in multi-class classification. However, this was not the case for the KPIs that we used in the case study. Notably for the lower relevance values (e.g., 1 and 2 with respect to the used KPI Relevance Scale), there were fewer data points than for the higher relevance values (e.g., 4 and 5 with respect to the used KPI Relevance Scale). As a result, the approach was inclined to predict higher relevance values for some KPIs, which are expected to have lower relevance values by the experts in the case study company. This indicates that the prediction quality of the approach has negatively affected due to lacking data points.

One of the possible reasons for a lower prediction quality value in the absence sub area is that there were fewer data points in the used KPI Relevance Data for each KPI in the absence sub area than for each KPI in the leave sub area. In particular, there was a limited variety of small organizations in the known relevance values of the KPIs in the absence sub area. This was because in small organizations, the management of absence-related data is ad-hoc, i.e., absence-related data may not be stored day-to-day. Therefore, there was missing usage information in the KPI Usage Logs for those KPIs for small organizations. By contrast, leave operations in organizations are mostly recorded day-to-day since there are regulations defined by law to keep the data related these leave operations up-to-date. In addition, leave is a type of operation that can be planned ahead, whereas absence has a more unpredictable nature.

As a result of having fewer data points for the KPIs in the absence sub area, as shown in Table 2, the prediction modeling techniques that use linear separation method for input data outperformed at predicting the relevance values of the KPIs in this sub area. However, tree/forest based prediction modeling techniques were majority in the leave sub area since there were more data points for the KPIs this sub area.

Table 2. Outperformed predicting modeling techniques for the KPIs in the HRM area

	Logistic regression	Support Vector Machines (SVMs)	Decision tree	Random forest	Stacked (Decision Tree & SVMs)
Absence	7	3	2	1	0
Leave	0	0	2	2	2

We also had the idea to apply our approach in the finance area. Using the KPI Usage Logs of 109 KPIs in the finance area, we created KPI Relevance Data.

Then, we analyzed this data before applying the approach on it. We found out that for more than 60% of the KPIs, there are fewer data points for 3 out of 5 known relevance values. We decided against actually using the approach although the data was not good enough, i.e., containing fewer data points. Unfortunately, the approach performed worse than predicting the relevance of the KPIs in the HRM area. We examined the failing predictions for the KPIs in the finance area with the two experts together with whom we determined the prediction accuracy of the approach in the HRM case–a director and a senior product developer. These experts pointed out that the expected relevance of the KPIs in the finance area are mostly dependent on various financial characteristics of organizations such as debt, revenue, payment periods of both the customers and suppliers of these organizations, and how the products and services are sold by these organizations. Although our approach is extensible to new organizational characteristic, however; unfortunately, these organizational characteristics are not available in the case study company since these are mostly sensitive data about organizations.

Software vendors that focus on automatically predicting relevant KPIs for their customers and operate various domains can apply our approach. However, if these software vendors may want to predict relevant KPIs for their customers using a different set of KPIs and organizational characteristics than we demonstrated in the case study, they need to provide their KPI Relevance Data to our approach and develop prediction models using the approach. Then, these software vendors can predict relevant KPIs for their customers by executing the approach with the developed prediction models.

6 Related Work

Due to the high interest in both academia and business, there is a broad literature in the field of organizational performance measurement. Notably researchers proposed various approaches dealing with determining relevant KPIs for organizations since KPIs are widely used as a means for measuring the performance of organizations. Within these approaches, creating relevant KPIs afresh for any organization or choosing KPIs from a reference set of KPIs (e.g., a KPI library) as the relevant set for a particular organization are the two common ways of determining relevant KPIs. In this section, we list some of the works, which cover the following question that we are interested in: how are relevant KPIs determined for organizations?

Much work has been conducted on defining relevant KPIs from scratch for organizations in various domains. Granberg and Munoz develop KPIs for airport managing organizations [11]. Similarly, to monitor the performance of airports, a set of KPIs are proposed in [7]. Kaganski et al. [15] describe the development of KPIs for small and medium-sized enterprises (SMEs). While a set of KPIs for the organizations that have highly diverse product families are defined in [24], Elliot et al. [6] specify a set of KPIs for a large pediatric healthcare organization. Since the development of KPIs in the aforementioned works is from scratch and manual, in each work, it is required to have an intensive technical knowledge of the organization to which relevant KPIs are determined. Thus, a significant effort is required to obtain that knowledge.

Apart from the aforementioned works, del Río-Ortega et al. present a meta-model [4] as a basis for working with KPIs. Using the language proposed as part of the meta-model one can model KPIs within the process models of the processes in an organization. Then, the values of the modeled KPIs can be derived from the execution logs of the process models. However, this still requires each organization to determine relevant KPIs for itself and model them using the proposed meta-model. Therefore, this will require a significant effort of each organization.

In some studies [13,16,21,22], researchers focus on selecting a subset from a set of KPIs to determine the relevant set of KPIs for organizations. Within that selection process, researchers mostly consider the sector of an organization or a set of business processes in an organization. However, due to the varying needs of organizations, a KPI subset that is selected as the relevant set for one organization may not be relevant for all other organizations, which are in the same sector or perform similar business processes with that organization. Therefore, that KPI subset selection process needs to be repeated for many organizations. To deal with that, Analytic Network Process (ANP) is utilized [3,14,17,28]. In particular, certain characteristics of KPIs such as reliability, comparability, and understandability are taken into account to determine the priorities of a set of existing KPIs in organizations. This is mostly done together with specific experts in organizations. Then, the KPIs that have the highest priorities are selected as the relevant KPIs for organizations. However, on the one hand, since the considered characteristics of KPIs are subjective to experts, the priority of a KPI may vary from one organization to another. On the other hand, ANP is a time-consuming and complex multi-criteria decision-making method, and therefore requires a significant effort from organizations.

7 Conclusion and Future Work

In this paper, we presented a novel approach aimed at the automated prediction of relevant KPIs for organizations. A set of prediction models aimed at predicting the relevance of KPIs and the organizational characteristics of a new organization are the required inputs by the approach. The approach determines which of the KPIs that are encoded in the prediction models are relevant for that new organization using the relevance factors of the KPIs. To identify these factors automatically and develop prediction models, the approach employs prediction modeling techniques and applies them on the known relevance values of KPIs for organizations, which should be given in the form of a specific input, KPI Relevance Data.

To show the accuracy of our approach, we implemented it and demonstrated in a case study at a Dutch ERP software vendor. Within the case study, together with experts in the company, we selected 19 KPIs from the HRM area that are offered to organizations by the company in its ERP software product. The known relevance values of the selected KPIs were not available in the company. Therefore, we identified KPI Usage Logs as a proxy for known relevance values of KPIs and subsequently we created KPI Relevance Data and developed the prediction models for the selected 19 KPIs together with the experts in the company. Afterwards, the relevance values of the KPIs were predicted for 261 organizations, which are new to those KPIs. Finally, we evaluated the prediction quality of the approach by comparing the predicted relevance values of the KPIs against the expected relevance values of those KPIs, which are provided by two experts in the company. The prediction quality of the approach was of sufficient quality to show the practical usage of the approach. As a result, we automate the selection of relevant KPIs for every organization. For current approaches, this is a manual endeavor that needs to be repeated for every single organization. Thus, we believe that our approach lowers the efforts of software vendors for determining relevant KPIs for their client organizations or the efforts of these organizations doing this themselves.

In future work, we want to extend our approach for determining relevant KPIs for different roles in organizations since the relevance of a KPI might vary from one role to another in organizations. For example, there may be a significant difference in the relevance value of a KPI on daily stock changes between a CEO and for a warehouse employee. Furthermore, sales, purchasing, and logistics are the areas to which we envision extending our approach since their commonality among organizations in addition to the less sensitivity of data for organizations in these areas in comparison to other areas, e.g., finance and accounting. Besides, we plan to develop a decision graph aimed at identifying which visualization best suits for particular KPIs. Thus, engaging dashboards comprising relevant KPIs can be built automatically. Moreover, the approach in this paper and the approach that we presented in [2] are part of our Cross-Organizational Process Mining Framework, which we introduced in [1], and will be together incorporated into the framework. With this, we aim to provide recommendations for organizations using the benchmarks that are developed utilizing relevant KPIs.

References

1. Aksu, Ü., Schunselaar, D.M.M., Reijers, H.A.: A cross-organizational process mining framework for obtaining insights from software products: accurate comparison challenges. In: 18th IEEE Conference on Business Informatics (2016)
2. Aksu, Ü., Schunselaar, D.M.M., Reijers, H.A.: An approach for automatically deriving key performance indicators from ontological enterprise models. In: 7th International Symposium on Data-driven Process Discovery and Analysis (2017)
3. Carlucci, D.: Evaluating and selecting key performance indicators: an ANP-based model. Measuring Bus. Excellence **14**(2), 66–76 (2010)
4. del-Río-Ortega, A., Resinas, M., Cabanillas, C., Cortés, A.R.: On the definition and design-time analysis of process performance indicators. Inf. Syst. **38**(4), 470–490 (2013)
5. Eisenberg, M., Hu, X.: Dichotomous relevance judgments and the evaluation of information systems. Proc. Am. Soc. Inform. Sci. **24**, 66–69 (1987)

6. Elliot, C., Mcullagh, C., Brydon, M., Zwi, K.: Developing key performance indicators for a tertiary children's hospital network. Aust. Health Rev. **42**(5), 491–500 (2018)
7. Eshtaiwi, M., Badi, I., Abdulshahed, A., Erkan, T.E.: Determination of key performance indicators for measuring airport success: a case study in Libya. J. Air Transp. Manag. **68**, 28–34 (2018)
8. Frost, J.: Making predictions with regression analysis, March 2019. https://statisticsbyjim.com/regression/predictions-regression
9. García, V., Mollineda, R.A., Sánchez, J.S.: Index of balanced accuracy: a performance measure for skewed class distributions. In: Araujo, H., Mendonça, A.M., Pinho, A.J., Torres, M.I. (eds.) IbPRIA 2009. LNCS, vol. 5524, pp. 441–448. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-02172-5_57
10. Gluck, M.: Exploring the relationship between user satisfaction and relevance in information systems. Inf. Process. Manage. **32**(1), 89–104 (1996)
11. Granberg, T.A., Munoz, A.O.: Developing key performance indicators for airports. In: 3rd ENRI International Workshop on ATM/CNS (2013)
12. Gravetter, F.J., Wallnau, L.B.: Essentials of statistics for the behavioral sciences (2013)
13. Ioan, B., Nestian, A.S., Tita, S.M.: Relevance of key performance indicators (KPIs) in a hospital performance management model. J. Eastern Europe Res. Bus. Econ. **2012**, 15 (2012)
14. Kachitvichyanukul, V., Luong, H., Pitakaso, R.: A hybrid MCDM approach to KPI selection of the coordination problems of production and sales departments-an empirical study of iron and steel industry of China and Taiwan. In: 13th Asia Pacific Industrial Engineering and Management Systems Conference (2012)
15. Kaganski, S., Paavel, M., Lavin, J.: Selecting key performance indicators with support of enterprise analyse model. In: 9th International DAAAM Baltic Conference "Industrial Engineering" (2014)
16. Kaganski, S., Snatkin, A., Paavel, M., Karjust, K.: Selecting the right KPIs for smes production with the support of PMS and PLM. Int. J. Res. Soc. Sci. **1**(3), 69–76 (2013)
17. Kucukaltan, B., Irani, Z., Aktas, E.: A decision support model for identification and prioritization of key performance indicators in the logistics industry. Comput. Hum. Behav. **65**, 346–358 (2016)
18. Parsons, V.L.: Stratified Sampling (2017)
19. Padmaja, D.L., Vishnuvardhan, B.: Comparative study of feature subset selection methods for dimensionality reduction on scientific data. In: IEEE 6th International Conference on Advanced Computing (2016)
20. Parmenter, D.: Key Performance Indicators: Developing, Implementing, and Using Winning KPIs. Wiley, New York (2015)
21. Peral, J., Maté, A., Marco, M.: Application of data mining techniques to identify relevant key performance indicators. Comput. Stand. Interfaces **54**, 76–85 (2017)
22. Pinna, C., Demartini, M., Tonelli, F., Terzi, S.: How soft drink supply chains drive sustainability: key performance indicators (KPIs) identification. In: 51st CIRP Conference on Manufacturing Systems (2018)
23. Saitta, S.: What is a good classification accuracy in data mining? March 2019. http://www.dataminingblog.com/what-is-a-good-classification-accuracy-in-data-mining
24. Schmidt, M., Schwöbel, J., Lienkamp, M., et al.: Developing key performance indicators for variant management of complex product families (2018)

25. Sokolova, M., Lapalme, G.: A systematic analysis of performance measures for classification tasks. Inf. Process. Manage. **45**(4), 427–437 (2009)
26. Steyerberg, E.W., et al.: Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology **21**(1), 128 (2010)
27. Tang, R., Shaw Jr., W.M., Vevea, J.L.: Towards the identification of the optimal number of relevance categories. J. Am. Soc. Inform. Sci. **50**(3), 254–264 (1999)
28. Van Horenbeek, A., Pintelon, L.: Development of a maintenance performance measurement framework-using the analytic network process (ANP) for maintenance performance indicator selection. Omega **42**(1), 33–46 (2014)
29. Zhou, Z.H.: Ensemble Methods: Foundations and Algorithms. Chapman and Hall/CRC, Boca Raton (2012)