Expert Systems With Applications 46 (2016) 249–261

Contents lists available at ScienceDirect

Expert Systems With Applications

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

A comparative study of Artificial Intelligence methods for project duration forecasting

Mathieu Wauters[a], Mario Vanhoucke[a,b,c,*]

[a] Faculty of Economics and Business Administration, Ghent University, Tweekerkenstraat 2, Gent 9000, Belgium
[b] Technology and Operations Management, Vlerick Business School, Reep 1, Gent 9000, Belgium
[c] UCL School of Management, University College London, Gower Street, London WC1E 6BT, United Kingdom

ARTICLE INFO

Keywords:
Project management
Earned Value Management (EVM)
Prediction
Artificial Intelligence

ABSTRACT

This paper presents five Artificial Intelligence (AI) methods to predict the final duration of a project. A methodology that involves Monte Carlo simulation, Principal Component Analysis and cross-validation is proposed and can be applied by academics and practitioners. The performance of the AI methods is assessed by means of a large and topologically diverse dataset and is benchmarked against the best performing Earned Value Management/Earned Schedule (EVM/ES) methods. The results show that the AI methods outperform the EVM/ES methods if the training and test sets are at least similar to one another. Additionally, the AI methods report excellent early and mid-stage forecasting results. A robustness experiment gradually increases the discrepancy between the training and test sets and demonstrates the limitations of the newly proposed AI methods.

1. Introduction

With the advent of the Critical Path Method (CPM, Kelley, 1961; Kelley and Walker, 1959) and the Program Evaluation and Review Technique (PERT, Fazar, 1959), project planning increasingly became a separate research discipline. While the former method focuses on the construction of a baseline schedule, the latter turns the attention to the relation between the project duration and activity duration variability. The construction of the baseline plan should be accompanied by an identification of its weak spots, namely those activities that are most likely to have the largest and possibly detrimental impact on the project outcome. As such, regardless of the presence of limited resources, the baseline schedule should be seen as a point of reference with which actual performance can be compared and contrasted. By relating the actual performance of the project to the planned performance, corrective actions can be triggered as soon as the project is deemed to be out-of-control. These three aspects, baseline scheduling, schedule risk analysis and project control are coined *dynamic scheduling* (Uyttewael, 2005; Vanhoucke, 2012; 2014).

The emphasis of this paper lies on one of the dynamic scheduling aspects, namely project control. A popular methodology for tracking project progress is Earned Value Management (EVM). It originated at the US Department of Defense (DoD) in the 1960s and aids project managers in controlling the projects' time and cost by using three key metrics that form the foundation of a number of performance indicators. An overview of the fundamentals of EVM can be found in Fleming and Koppelman (2005). While initial studies were dominated by the cost objective, the work of Lipke (2003) proved to be a turning point as Earned Schedule (ES) allowed project managers to track progress in units of time rather than monetary units. Along with the inception of Earned Schedule, academic studies shifted to the time dimension. A popular topic involved forecasting the final duration (or previously the project's real cost) based on progress data. The importance of an accurate estimate for the project's end cannot be underestimated. Not only does it allow a project manager to look ahead, it also implies an implicit call for action. Forecasting methods can be embedded in decision support systems that trigger a warning once the expected duration exceeds a user-defined threshold. Obviously, the validity of this warning signal greatly depends on the trustworthiness of the underlying forecasting method. The performance of three Planned Value, three Earned Duration and three Earned Schedule methods has been investigated on simulated data (Vanhoucke & Vandevoorde, 2007) as well as on real-life projects (Vandevoorde & Vanhoucke, 2006). Vanhoucke (2010b) and Vanhoucke (2011) advocated the integration of top-down control systems such as EVM with bottom-up sensitivity indicators which result from schedule risk analysis. This was followed up by Elshaer (2013) who proposed an adaptation of the Earned Schedule forecasting method using activity sensitivity information.

* Corresponding author at: Faculty of Economics and Business Administration, Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium. Tel.: +32 92643569.
E-mail addresses: mathieu.wauters@ugent.be (M. Wauters), mario.vanhoucke@ugent.be (M. Vanhoucke).

http://dx.doi.org/10.1016/j.eswa.2015.10.008
0957-4174/

Recent years witnessed a spike in contributions studying the application of statistical techniques and concepts to project control. Confidence intervals were introduced by Lipke, Zwikael, Henderson, and Anbari (2009) and a similar concept, known as (stochastic) S-curves, was introduced by Barraza, Back, and Mata (2000, 2004). S-shaped curves were also used by Narbaev and De Marco (2014a) in which growth models were utilized for cost forecasting. This cost forecasting problem was also approached from a regression point of view ((Narbaev & De Marco, 2014b)). Other applications include fuzzy Earned Value Management (Naeni, Shadrokh, & Salehipour, 2011; Salari, Bagherpour, & Reihani, 2015) and detecting and quantifying abnormal variation (Colin & Vanhoucke, 2014; Pajares & López-Paredes, 2011). The reader is referred to Willems and Vanhoucke (2015) for a more comprehensive literature overview on EVM and project control.

This paper focuses on a research branch known as statistical learning (Hastie, Tibshirani, & Friedman, 2009). More specifically, a new class of methods is implemented to construct project duration estimates. These methods hail from the domain of Artificial Intelligence (AI) which is dedicated to learning the relation between inputs and outputs and applying that relation for classification or prediction purposes. To the best of our knowledge, only four works deal with statistical learning involving EVM metrics. Cheng, Peng, Wu, and Chen (2010) deploy a Support Vector Machine (SVM) to estimate the final cost of two construction projects. The parameters are tuned with a fast messy genetic algorithm. The combination of a basic metaheuristic as a tuning mechanism and the SVM was united with fuzzy logic. Cheng and Roy (2010) tested this system for function approximation and cost estimation. Wauters and Vanhoucke (2014) applied Support Vector Regression for project control time and cost forecasting. The authors compared its performance with the best performing EVM and ES methods on a large dataset and revealed the pitfalls in a robustness experiment. The computational experiment revealed that SVMs outperform the current EVM forecasting methods when the training set is equal or at least similar to the test set. More recently, Acebes, Pereda, Poza, Pajares, and Galan (2015) illustrated the use of a number of statistical learning techniques and Anomaly Detection algorithms by means of a case study.

The goal of this paper consists of introducing Artificial Intelligence methods for constructing better forecasts to predict the final duration of a project. The contribution of this paper to the body of project control literature is fourfold. First, we propose five Artificial Intelligence methods for predicting a project's final duration. AI methods use historical or simulated data to learn the relation between inputs and one or multiple outputs. In a project control environment, the simulated data contains information with regard to the progress of the project. The proposed methods learn how the performance indicators are related to the project's Real Duration (RD). This knowledge is then applied to new data to come up with an estimate of the project's final duration. Secondly, a generally applicable methodology is put forward starting with the generation of project and progress data. This serves two purposes. The first purpose lies in the nature of the AI methods which learn a relation from existing data. Second, a wide array of data allows us to reach general conclusions. Apart from the data generation, a decision needs to be made on which progress data will be fed to the AI methods. These progress data contain periodic measurements of EVM performance indicators, as well as EVM forecasting estimates. The high volume of data may be correlated and some data may be irrelevant. In order to alleviate these problems, Principal Component Analysis is applied. Once the relevant data is selected, the AI methods need to be fine-tuned, since their performance is highly dependent on the chosen parameters. The third contribution lies in assessing the general performance of the AI methods and comparing it against the EVM/ES methods. Incidentally, the impact of the project network's topology and percentage complete is identified. Finally, the main experiment assumes that the project manager can provide an accurate estimate of the variability. As a result, the training and test set are drawn from a distribution with identical parameters. In real-life situations, it is extremely difficult to appraise the variability affecting project activities. In line with this reasoning, a robustness experiment in which the training and test sets no longer coincide is set up.

The outline of this manuscript is as follows. In Section 2, an overview of the Artificial Intelligence methods is supplied. Section 3 goes over the process of data generation, the EVM progress data, data pre-processing with PCA and training, validation and test sets to construct a forecast. The steps of the methodology are revisited in Section 4 which proffers specific settings. The results found in Section 5 first elaborate on fine-tuning the parameters of the AI methods (Section 5.1.1) and the desired level of explained variation for the principal components (Section 5.1.2). Section 5.2 discusses the general accuracy, the impact of the topology and the percentage complete. The sensitivity analysis varies the mean and standard deviation of the underlying distribution and proves how data-intensive methods conform to the well-known "garbage-in, garbage-out" principle. Section 6 concludes this paper by sharing the main insights of our research.

2. Artificial Intelligence methods

In this section, a general overview of the employed Artificial Intelligence methods will be given. All of these methods will be used to construct a forecast of the final project's duration. In Section 3, it will be shown how these techniques are embedded in the presented methodology.

2.1. Decision tree learning

2.1.1. Decision trees

Decision tree learning finds its roots in the seminal work by Morgan and Sonquist (1963). In this paper, the authors deal with automated interaction detection and propose a new procedure for data analysis and regression, which is now known as decision tree learning. Inspired by this new research direction, Breiman, Friedman, Olshen, and Stone (1984) and Quinlan (1993) independently came up with algorithms that are comprised of two phases. In the first phase, the solution space is partitioned using a binary (Breiman et al., 1984) or multi-way (Quinlan, 1993) split after which, in the second phase, a constant model is applied to each node of the partition. These algorithms are known as Classification And Regression Trees (CART) and C4.5, respectively. The approach of these well-known techniques are subject to two pitfalls, namely overfitting and selection bias. The overfitting problem results from the lack of statistical significance, as noted by Mingers (1987). Even though some information measure is maximized in order to make a split in the decision tree, there is no way of establishing whether this split is significant. The selection bias follows from the fact that attributes with more split points are preferred. This issue was raised by Breiman et al. (1984) but no remedy was provided.

The solution was found by Hothorn, Hornik, and Zeileis (2006b), who proposed a conditional inference framework that makes use of permutation tests developed by Strasser and Weber (1999). These permutation tests look for dependence between the outcome and the different predictors, after which the predictor with the smallest p-Value is selected for splitting. Consequently, the conditional inference framework meets the need for a statistical approach to recursive partitioning, as demanded by White and Liu (1994).

2.1.2. Bagging and Random forest

The principal shortcoming of decision trees lies in their instability when small changes in learning data occur. Variable selection and selection of the cutpoint for the selected variable(s) highly depends on