Nirdizati: A Web-Based Tool for Predictive Process Monitoring

Kerwin Jorbina[1], Andrii Rozumnyi[1], Ilya Verenich[1,2], Chiara Di Francescomarino[3], Marlon Dumas[1], Chiara Ghidini[3], Fabrizio Maria Maggi[1], Marcello La Rosa[3], and Simon Raboczi[3]

[1] University of Tartu, Estonia
{kerwin.jorbina,andriiro,marlon.dumas,f.m.maggi}@ut.ee
[2] Queensland University of Technology, Australia
{ilya.verenich,m.larosa,simon.raboczi}@qut.edu.au
[3] FBK IRST, Trento, Italy
{dfmchiara,ghidini}@fbk.eu

Abstract. This paper introduces Nirdizati: A web-based application for generating predictions about running cases of a business process. Nirdizati is a configurable full-stack web application that supports users in selecting and tuning prediction methods from a list of implemented algorithms and enables the continuous prediction of various performance indicators at runtime. The tool can be used to predict the outcome, the next events, or the remaining time of each case of a process. For example, in a lead-to-order process, Nirdizati can predict which customer leads will convert to purchase orders and when. In a claims handling process, it can predict if a claim decision will be made on time or late. The results of the predictions, as well as the real-time summary statistics about the process executions, are presented in a dashboard that offers multiple visualization options. Based on these predictions, process participants can act proactively to resolve or mitigate potential process performance violations. The target audience of this demonstration includes process mining researchers as well as practitioners interested in exploring the potential of predictive process monitoring.

Keywords: Predictive Process Monitoring, Machine Learning

1 Introduction

Predictive Process Monitoring [1] is an emerging paradigm based on the continuous generation of predictions about the future values of user-specified performance indicators of a currently running process execution. In this paradigm, a user defines the type of predictions she is interested in and provides a set of historical execution traces. Based on the analysis of these traces, the idea of predictive monitoring is to continuously provide the user with predictions and estimated values of the measure of interest. Such predictions generally depend both on: (i) the sequence of activities executed in a given case; and (ii) the values of data attributes after each activity execution in the case.

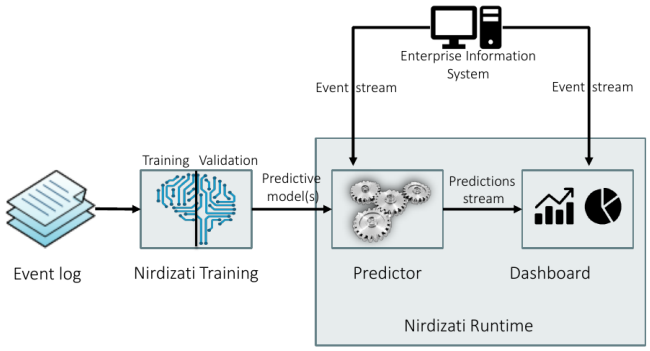

Fig. 1. A high-level overview of Nirdizati.

As an example, consider a business process for managing loan applications; the applicant can be advised about the combinations of loan amount and duration that are the most likely to lead to an acceptance of the application, given some contextual information about the application and the personal data of the applicant (e.g., age, salary, etc.).

Several approaches have been proposed in the literature to tackle common predictive process monitoring tasks. However, so far, these approaches have largely remained in the academic domain and have not been widely applied in real-time scenarios, in which users require a continuous predictive support.

In this paper, we present Nirdizati, a pioneering open-source Web-based predictive monitoring tool, which is able to fill this gap between research and practice, by providing business analysts with a highly flexible instrument for the selection, generation and analysis of different predictive models, and end-users with continuous runtime predictions.

Nirdizati consists of two components: *Nirdizati Training* and *Nirdizati Runtime* (Fig. 1). Nirdizati Training takes as input a business process event log and produces one or more prediction models, which can then be deployed in Nirdizati Runtime. Once a model is deployed, Nirdizati Runtime listens to a stream of events related to the process, and produces a stream of predictions. These predictions are then visualized in a continuously updated Web dashboard.

2 Nirdizati Training

Nirdizati Training is the component of Nirdizati that allows users to produce predictive models later used by Nirdizati Runtime for making predictions on streams of events. It provides several algorithms for generating predictive models suitable for different types of predictions and tailored to each specific dataset. For example, it is able to build predictive models for predicting remaining time, the next activity to be performed, whether a certain outcome will be achieved or not and even the overall workload per day. To this aim, the training component of

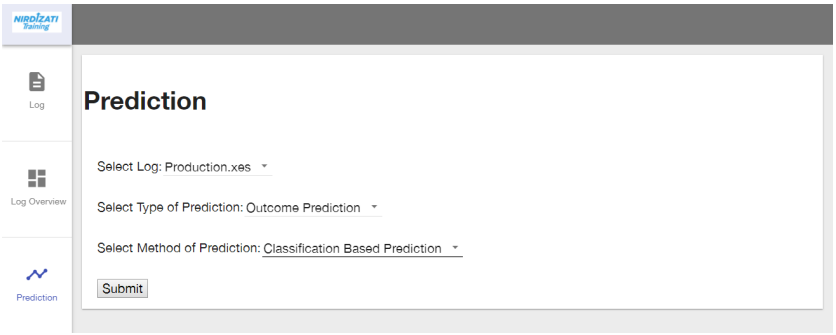

Fig. 2. The front-end of Nirdizati Training.

Nirdizati relies on two phases: a training and a validation phase. In the former, one or more predictive models are built; in the latter, their suitability to the specific dataset is evaluated, so as to support the user in selecting the predictive model that ensures the best results.

Nirdizati Training is composed of a front-end application (Fig. 2), which allows users to configure the predictive models and to assess the goodness-of-fit of the built models, and a back-end application responsible for the actual training and validation. The back-end application is in turn composed of four submodules shown in Fig. 3. A Log Manager is in charge of managing the logs. Uploading and retrieving the logs are the basic operations of this module. The Encoder is responsible for parsing the logs, labeling them according to the desired type of prediction, and preparing the data for the training phase. In the Training submodule, the encoded data are split into training and validation set used for evaluation purposes, and the predictive models are built from the training data.

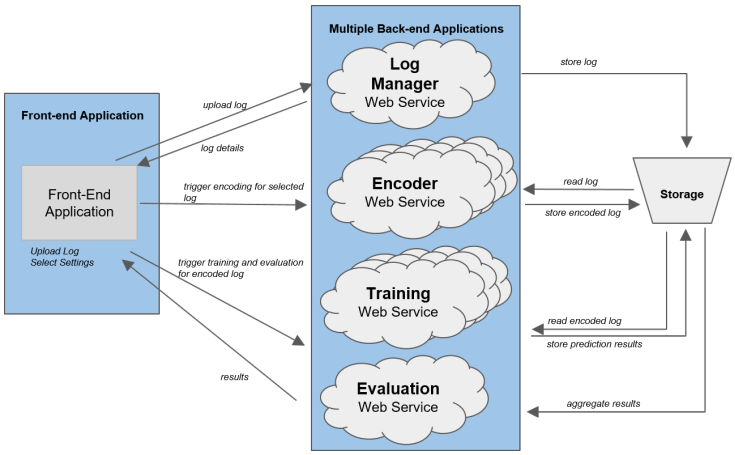

Fig. 3. A high-level overview of Nirdizati Training.

Finally, the Evaluation submodule tests the validation data against the model(s) created to get accuracy measures with respect to the ground truth, available from the complete traces in the validation data. The back-end application comes with a storage module for saving the uploaded logs and the built predictive models.

3 Nirdizati Runtime

Once the predictive models have been trained, they are used by the Runtime component to make predictions for ongoing cases. As an input, Nirdizati Runtime takes a stream of events produced by information systems that register tasks performed in the organization (Fig. 4). Next, the produced stream is consumed by a web server which is responsible for storing data and further information processing. The web server, in turn, communicates with a predictive engine which applies models obtained via Nirdizati Training for a business process and returns all the results of calculations. After web server receives the outputs from each predictive model, it writes them to a database and updates corresponding clients. From the user's side, the results are visualized via a dashboard-based interface which is capable of sending queries requesting various visualization options.

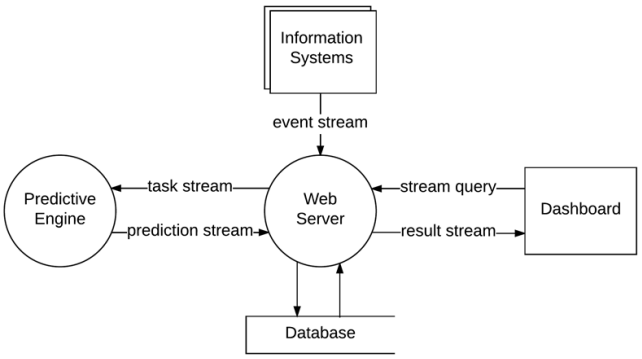

Fig. 4. High-level data flow diagram of Nirdizati Runtime

When a user logs into the system, she initially sees a table listing both currently ongoing cases (colored in gray) as well as completed cases (colored in green), as shown in Fig. 5. For each case, we provide a range of summary statistics including the number of completed events in a given case, case start time and the time of the latest event completion. For ongoing cases, we also estimated their completion time and a range of other process-specific indicators, such as the probability that a customer will not accept an offer from a salesperson. For completed cases, instead, we show the actual completion time and the actual indicators.

Process workers and operational managers – typical users of such system – can set the process performance targets and subscribe to a stream of warnings

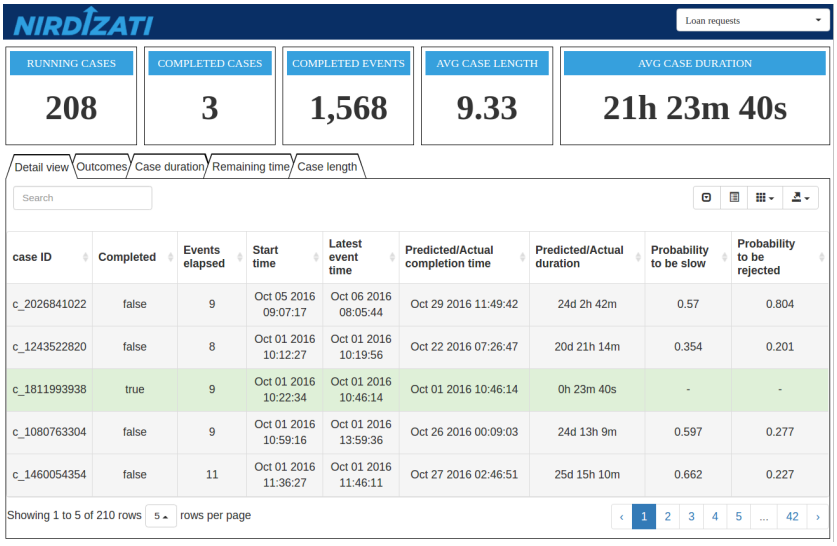

Fig. 5. Main view of Nirdizati Runtime

and alerts generated whenever these targets are predicted to be violated. Thus, process workers will be capable of making more informed, data-driven decisions to get a better control of the process execution. This is especially beneficial for processes where process workers have more leeway to make corrective actions, for example, in a lead management process.

4 Conclusion

The developed solution, named *Nirdizati*, is a configurable full-stack web application that supports users in selecting the preferred prediction method from the list of implemented methods and enables the continuous prediction of various performance indicators at runtime. The results of the predictions, as well as the real-time summary statistics about the process execution, are presented in a dashboard that offers multiple visualization options.

A video demo of Nirdizati can be found at `http://youtu.be/0nr14lX04-I`. A public release of Nirdizati Training and Nirdizati Runtime are available at `http://training.nirdizati.com` and `http://dashboard.nirdizati.com` respectively. The source code is available under the Lesser GNU Public License (LGPL) at `http://github.com/nirdizati`.

References

1. Maggi, F., Di Francescomarino, C., Dumas, M., Ghidini, C.: Predictive monitoring of business processes. In: Proc. of the International Conference on Advanced Information Systems Engineering (CAiSE). Springer (2013)