**Name of student:** Mickael LIU

**Name of your Level:** L0

**Source (e.g. scholars.google.com):** <https://ieeexplore.ieee.org/document/8103817>

**Paper title:** Predictive Monitoring of Business Processes: A Survey

**Keywords specific to the paper:**

Business process management, process mining, predictive monitoring, process indicators

**Summary of the main contributions:**

Nowadays, process mining is becoming a growing area of interest in business process management. Process mining consists of extracting information from the event logs of a business process. Modern information systems that support complex business processes typically maintain significant amounts of process execution data, in particular event records corresponding to the execution of activities (event logs). One of the applications of Process Mining is predictive monitoring of business processes. We want to analyze these event logs to predictively monitor objectives during business process execution. At any time during the execution of a process, the user can define business objectives in the form of linear temporal logic rules. When an activity is running, the framework identifies the input data values that are likely to lead to the achievement of each business objective. Unlike more traditional compliance monitoring approaches that detect violations only after they have occurred, this predictive monitoring approach provides early guidance so that users can redirect or improve current process executions if necessary. The goal is the achievement of business objectives. In other words, risks are predicted (and potentially avoided) rather than simply detected.

The goal of these techniques is the prediction of quantifiable metrics of a running process instance with the generation of predictive models. There are different types of predictive computing methods that we want to categorize.

Business process prediction is defined as an execution method aimed at generating predictive models that can be used to predict a particular value of a current process based on input values which may be event logs, or l current state of the process.

This value can be Boolean, categorical or numerical depending on the prediction object, for example the remaining time of a process (numeric) or the accomplishment of a certain objective (Boolean).

Thus, the development of mechanisms to predict these values based on the execution processing of event streams exchanged between different information systems is very interesting from a practical point of view.

Over the past six years, various different approaches to predictive monitoring have emerged. They were developed to predict different types of measurements, approached the problem from different angles, and were applied to different domains. However, despite their differences, they all share many commonalities. Thus, a joint analysis of all these approaches can provide an overall vision as well as identify new challenges in this area. This is the main objective of this survey, which collects and analyzes a compilation of execution monitoring prediction approaches on Business Processes.

These methods include techniques based on machine learning approaches, statistical methods, annotated transition systems, and hybrid methods.

Machine learning is employed for predictive monitoring in the context of business processes to make predictions about the performance and behavior of ongoing process instances. The document mentions several ways machine learning is utilized for predictive monitoring:

1. Classification Methods: The document lists two process-aware classification methods that are based on machine learning. Additionally, some works employ machine learning-based methods for predictive monitoring, specifically using artificial neural networks (a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain) in experimentation to analyze the accuracy of prediction and the complexity of the tree for each representation.
2. Probabilistic Models: Probabilistic models are used in predictive monitoring. These models include stochastic elements (Stochastic elements allow you to explicitly represent uncertainty in the input data for your model) and are fitted according to the event log using an expectation maximization algorithm (makes it possible to find the maximum likelihood parameters of a probabilistic model) to predict the next event of running process instances.
3. Feature Selection and Domain Knowledge: The document highlights the need for the use of domain knowledge about a process event log to build new features that improve the prediction power of models. It also refers to the selection of key attributes and discusses how little work has been done in this area concerning the machine learning aspect of predictive monitoring.

By comparing the performance of different approaches, we cannot draw clear conclusions about which methodology is better. It depends, among other things, on the dataset used, the input features of the machine learning algorithm.

Although this work has been mainly focused in Business processes, predictive monitoring can also be applied to other types of industrial processes such as production, manufacturing or case handling processes. For example the cost of a manufacturing process can be analyzed and predicted. A process model-enhanced cost, as well as predictions of volume and remaining times, are used as process mining techniques in this case.

Some challenges related to predictive monitoring can be identified as a result of this study. We then detail them grouped into 5 main categories :

1. Application of new technologies, powerful software is required to process a large amount of data
2. Evaluate and compare the different models offered
3. Predict the number of measures that will need to be taken to resolve all the problems and not just predict one by one
4. The quality of the prediction
5. Strategy for updating the model, knowing when to improve the product and taking the time necessary to avoid losing quality
6. Practical predictions, most proposals are focused on improving the accuracy of predictions, but little attention has been given to providing recommendations and explaining the prediction values to the users so that they can determine the best way to act upon.