

Process Mining: Data science in Action

Технический университет Эйндховена

Автор: Wil van der Aalst, Professor dr.ir. Department of Mathematics & Computer Science

Платформа Coursera

Ссылка на курс: <https://ru.coursera.org/learn/process-mining>

Process mining is the missing link between model-based process analysis and data-oriented analysis techniques. Through concrete data sets and easy to use software the course provides data science knowledge that can be applied directly to analyze and improve processes in a variety of domains. Data science is the profession of the future, because organizations that are unable to use (big) data in a smart way will not survive. It is not sufficient to focus on data storage and data analysis. The data scientist also needs to relate data to process analysis. Process mining bridges the gap between traditional model-based process analysis (e.g., simulation and other business process management techniques) and data-centric analysis techniques such as machine learning and data mining. Process mining seeks the confrontation between event data (i.e., observed behavior) and process models (hand-made or discovered automatically). This technology has become available only recently, but it can be applied to any type of operational processes (organizations and systems). Example applications include: analyzing treatment processes in hospitals, improving customer service processes in a multinational, understanding the browsing behavior of customers using booking site, analyzing failures of a baggage handling system, and improving the user interface of an X-ray machine. All of these applications have in common that dynamic behavior needs to be related to process models. Hence, we refer to this as "data science in action". The course explains the key analysis techniques in process mining. Participants will learn various process discovery algorithms. These can be used to automatically learn process models from raw event data. Various other process analysis techniques that use event data will be presented. Moreover, the course will provide easy-to-use software, real-life data sets, and practical skills to directly apply the theory in a variety of application domains. This course starts with an overview of approaches and technologies that use event data to support decision making and business process (re)design. Then the course focuses on process mining as a bridge between data mining and business process modeling. The course is at an introductory level with various practical assignments. The course covers the three main types of process mining. 1. The first type of process mining is discovery. A discovery technique takes an event log and produces a process model without using any a-priori information. An example is the Alpha-algorithm that takes an event log and produces a process model (a Petri net) explaining the behavior recorded in the log. 2. The second type of process mining is conformance. Here, an existing process model is compared with an event log of the same process. Conformance checking can be used to check if reality, as recorded in the log, conforms to the model and vice versa. 3. The third type of process mining is enhancement. Here, the idea is to extend or improve an existing process model using information about the actual process recorded in some event log. Whereas conformance checking measures the alignment between model and reality, this third type of process mining aims at changing or extending the a-priori model. An example is the extension of a process model with performance information, e.g., showing bottlenecks. Process mining techniques can be used in an offline, but also online setting. The latter is known as operational support. An example is the detection of non-conformance at the moment the

deviation actually takes place. Another example is time prediction for running cases, i.e., given a partially executed case the remaining processing time is estimated based on historic information of similar cases. Process mining provides not only a bridge between data mining and business process management; it also helps to address the classical divide between "business" and "IT". Evidence-based business process management based on process mining helps to create a common ground for business process improvement and information systems development. The course uses many examples using real-life event logs to illustrate the concepts and algorithms. After taking this course, one is able to run process mining projects and have a good understanding of the Business Process Intelligence field. After taking this course you should: - have a good understanding of Business Process Intelligence techniques (in particular process mining), - understand the role of Big Data in today's society, - be able to relate process mining techniques to other analysis techniques such as simulation, business intelligence, data mining, machine learning, and verification, - be able to apply basic process discovery techniques to learn a process model from an event log (both manually and using tools), - be able to apply basic conformance checking techniques to compare event logs and process models (both manually and using tools), - be able to extend a process model with information extracted from the event log (e.g., show bottlenecks), - have a good understanding of the data needed to start a process mining project, - be able to characterize the questions that can be answered based on such event data, - explain how process mining can also be used for operational support (prediction and recommendation), and - be able to conduct process mining projects in a structured manner.

This course is aimed at both students and professionals. A basic understanding of logic, sets, and statistics (at the undergraduate level) is assumed. Basic computer skills are required to use the software provided with the course (but no programming experience is needed). Participants are also expected to have an interest in process modeling and data mining but no specific prior knowledge is assumed as these concepts are introduced in the course.

Introduction and Data Mining

This first module contains general course information (syllabus, grading information) as well as the first lectures introducing data mining and process mining.

Process Models and Process Discovery

In this module we introduce process models and the key feature of process mining: discovering process models from event data.

Different Types of Process Models

Now that you know the basics of process mining, it is time to dive a little bit deeper and show you other ways of discovering a process model from event data.

Process Discovery Techniques and Conformance Checking

In this module we conclude process discovery by discussing alternative approaches. We also introduce how to check the conformance of the event data and the process model.

Enrichment of Process Models

In this module we focus on enriching process models. We can for instance add the data aspect to process models, show bottlenecks on the process model and analyse the social aspects of the process.

Operational Support and Conclusion

In this final module we discuss how process mining can be applied on running processes. We also address how to get the (right) event data, process mining software, and how to get from data to results.