

Syllabus
Process Mining: Data science in Action
(3 ECTS)

Author:

Wil van der Aalst, Eindhoven University of Technology (TU/e)

Psychology Programme

Meeting Minute #2.6-06.7/4 dated 28 June 2019

1. Course Description

a. Title of a Course: Process Mining: Data science in Action

b. Pre-requisites:

There are no prerequisites.

c. Course Type (compulsory, elective, optional): elective

d. Abstract: 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".

2. Learning Objectives:

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.

3. Learning Outcomes:

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.

4. Course Plan:

Week 1 - 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.

Week 2 - 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.

Week 3 - 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?

Week 4 - 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.

Week 5 - 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

Week 6 - 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

5. Reading List:

a. Required:

Busemeyer, J., Wang, Z., Townsend, J., & Eidels, A. (Eds.), *The Oxford Handbook of Computational and Mathematical Psychology*. : Oxford University Press,(или более поздние издания). – URL:

<http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199957996.001.0001/oxfordhb-9780199957996>– ЭБС: Oxford Handbooks Online

b. Optional:

AGARWAL, S. Introduction to Neuromarketing and Consumer Neuroscience. *Journal of Consumer*

Little, T. (Ed.), *The Oxford Handbook of Quantitative Methods in Psychology*, Vol. 1. : Oxford University Press, (или более поздние издания). – URL: <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199934874.001.0001/oxfordhb-9780199934874>.– ЭБС: Oxford Handbooks Online

In this course, information on recommended readings will be provided at the beginning of each module. At the end of each module, we will also give a list of additional resources to help further expand your knowledge on the topics discussed.

6. Grading System:

The final grade is calculated as follows: Final Quiz (100%). The grades will be given on a scale of 1 to 10 throughout the class. All grades, having a fractional part greater than 0.5, are rounded up. greater than 0.5, are rounded up.

7. Examination Type:

Sample questions for assessing the quality of knowledge:

«When we talk about a 'spaghetti' process we mean a process that is...

1. ...connected to the internet through many wires.
2. ...describing a process to cook something.
3. ...well structured, like the uncooked spaghetti in the package.
4. ...unstructured like a plate full of spaghetti threads running through each other.

When we talk about "replay", we mean the practice where...

1. we start from both a process model and a collection of observed behavior, e.g. traces, and compare these.
2. we start from a process model and generate behavior, e.g. traces.
3. we start from event data and generate a process model, e.g. a Petri net.»

8. Methods of Instruction:

Blended course: On-line lectures (<https://www.coursera.org/learn/process-mining>) and out-of-class work.

9. Special Equipment and Software Support (if required): PC, internet access