Hochschule für Technik und Wirtschaft Berlin
University of Applied Sciences

Project Report

Student Project Business Intelligence – Process Causality Analytics

Field of study: Business Intelligence

Processing period: May 18 2021 - September 17 2021

Fabio Cosimo Andriulo	575311
Timur Burkholz	575308
Maximilian Alexander Reech	575311

Professor:	Prof. Dr. Ingo Claßen
Supervisor:	Benjamin Aunkofer
Partners (Signavio):	Timotheus Kampik, Konstantinos Poulios

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Contents

List of Figures

List of Tables

1 Introduction

This document was created as a report within the context of the project. The project is part of the Master's program in Business Informatics at HTW Berlin. The project has been presented by DATANOMIQ GmbH. The project was accompanied by Signavio GmbH as a cooperation partner.

1.1 Background

As a part of business intelligence, process mining is a way for modern companies to get information about their processes. It is not uncommon for processes to be further adapted and changed over time. The question always arises to what extent the changes have influenced the measurement results of the process.

1.2 Goal

In order to specify this question further, it should be found out whether there is a causal relationship between the change in the process and the change in the measurement result. In the best case, the goal is to find the specific changes from all the changes that also influence the measurement result.

1.3 Structure

In the following, the methodology used is presented first for the sake of clarity. This is to clarify afterwards, for which reasons which kind of data was used. Afterwards, the implementation of the software will be discussed in order to explain its peculiarities, limitations and possibilities.

2 Double Machine Learning

2.1 In general

A data-based causal analysis attempts to measure the effect of a change on observable outcome. The measurement of this effect is possible when the changes are observable. This can be implemented by double machine learning. In general, it tries to estimate the effect of the change on the outcome by the changes.[1] In order to recognize the causal effect of a change, all other influences must be comparable.[2] In addition, it is necessary that the results can arise from the same characteristics, as well as there should be no combination that is found only before the change or after the change.[3]

By using machine learning, the relationships between the changes and their effects can be represented as machine models. It should be noted that the representation as a model is always only an approximately correct estimate of the real relationship.[4]Now we can assume that there are two models. One to calculate the result under condition of the change and one to explain the change. By assuming that a change is calculable if the calculation of the first model is different under the different conditions (changed/not changed), it can be assumed that the representation of the change by the second model explains the difference.[5]

---

[1] see Huber 2020, p. 106.
[2] see ibid., p. 107.
[3] see ibid., p. 109.
[4] see ibid., p. 111.
[5] see ibid., pp. 112 sqq.

2.2 In this case

In this case, however, we are talking about process changes. If there is a suitable representation of the process properties in the flow and its results, then double machine learning can also be used for this. In the following, the idea is explained with an example.

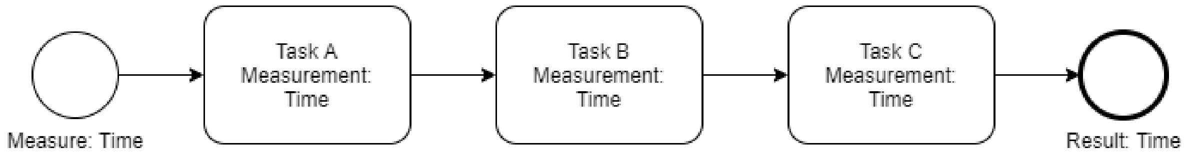

Figure 2.1: Process 1 (Unchanged)

Figure 2.2: Process 2 (Changed)

As can be seen in the figures, Task B was deleted and Task D was introduced in the modified process. As a simple example, we now use time as our measurand. Here, the result of the measurement is the sum of the time used for the tasks. We can therefore set up the following example functions:

p_1/p_2 representing the results, $A..D$ representing the measurements

$$p_1 = A + B + C$$

$$p_2 = A + C + D$$

Now, we can assume that a machine learning model (u) can be trained for the unmodified process as the adornment of the estimate p1 and as the characteristics A..C. Then, we can try to use the model (u) to perform a prediction for our modified process to measure the difference (β) in the outcome.

$$u \approx p_1$$

However, this requires the model to make a prediction (q) for the outcome of the changed process.

$$u \rightarrow p_2 \approx q$$

Now the difference (β) can be calculated.

$$\beta = p_2 - q$$

A second model (m) shows the effect of the change on the difference.

$$m \approx \beta$$

In this case, it is clear because the change in results can only be explained by task D.

$$m \equiv D$$

In more complex examples, it is conceivable that the change in difference may also consist of combination of changed characteristics. In the end, it is only necessary to find out whether the estimation of the difference was successful. A classical regressive measurement would be the mean squared error. The smaller this is (closer to zero) the better the estimation succeeds and thus the presentation as a causal reason.

$$m \rightarrow \beta \equiv \beta_m$$

$$causality = mean((\beta - \beta_m)^2)$$

3 Data acquisition

3.1 Real Data

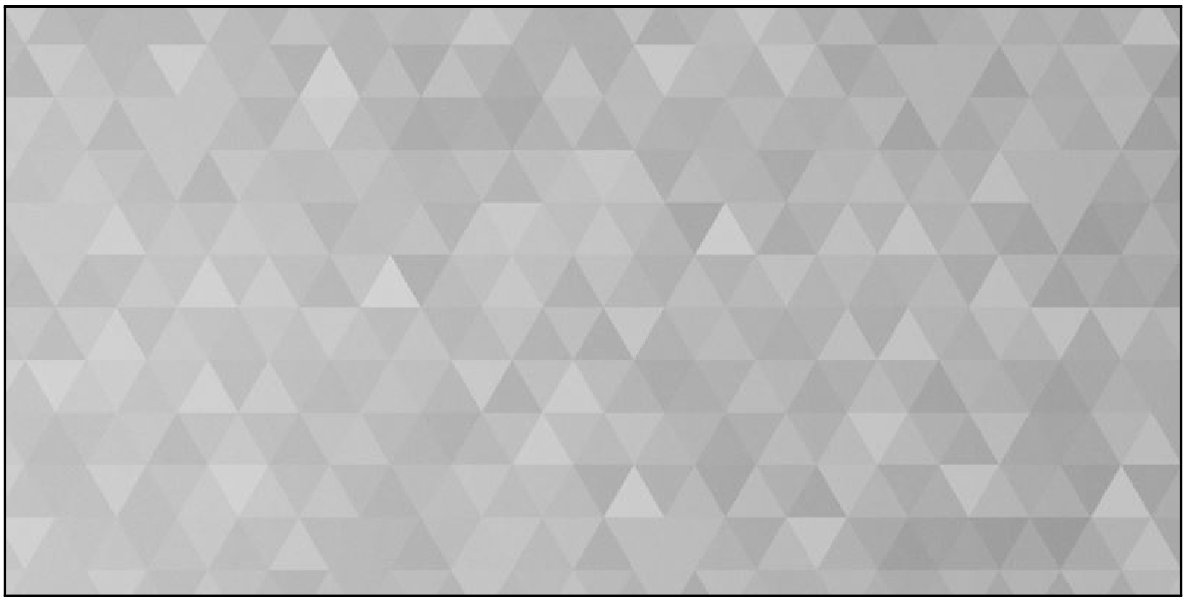

Figure 3.1: Bildunterschrift

Erstellt in Eigenarbeit in Anlehnung an see Unternehmen Sonstige GmbH 2020

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. (siehe Abbildung 3.1)

3.2 Simulation

Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.
Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.
Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.
Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.
Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.

Table 3.1: Tabellenunterschrift

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. (siehe Tabelle 3.1)

4 Implementation

4.1 Simulation

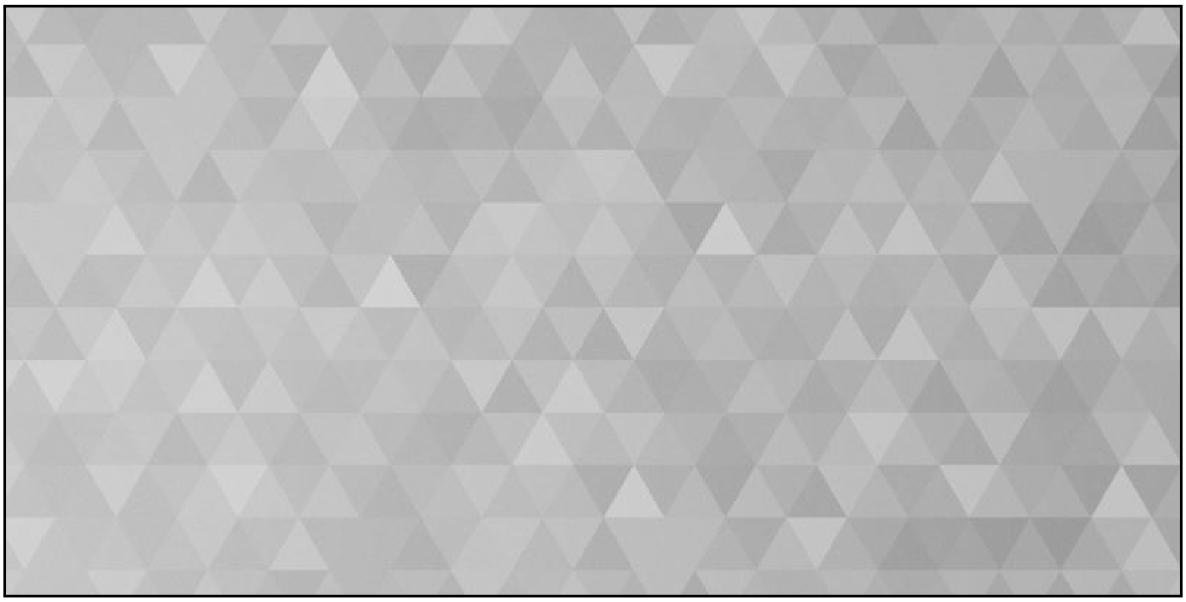

Figure 4.1: Bildunterschrift

Erstellt in Eigenarbeit in Anlehnung an see Unternehmen Sonstige GmbH 2020

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. (siehe Abbildung 4.1)

4.2 Data Transformation

Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.
Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.
Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.
Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.
Inhalt	Jemand musste Josef K. verleumdet haben, denn ohne dass er etwas Böses getan hätte, wurde er eines Morgens verhaftet. “Wie ein Hund!” sagte er, es war, als sollte die Scham ihn überleben.

Table 4.1: Tabellenunterschrift

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. (siehe Tabelle 4.1)

4.3 Double Machine Learning

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.[1]

[1] see Literatur 2020, p. 1.

5 Conclusion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Literaturverzeichnis

Literatur Quellen

Huber, Martin (2020). *Kausalanalyse mit maschinellem Lernen*. Springer Fachmedien Wiesbaden GmbH.

Literatur, Vorname (2020). *Titel Literatur*. Springer Vieweg. ISBN: 123-4-567-89101-1.

Sonstige Quellen

Unternehmen Sonstige GmbH (2020). *Titel Sonstige*. URL: https://www.google.com (visited on 10/14/2020).