Titel und Untertitel der Arbeit

Bachelorarbeit / Masterarbeit

Hochschule der Medien Stuttgart

Fachbereich Information und Kommunikation

Studiengang Wirtschaftsinformatik und digitale Medien

(x. Semester)

Erstbetreuer: Prof. Dr. XYZ

Zweitbetreuer: N.N.

Sommersemester 2021

Vorgelegt von: N. N.

Matrikelnummer: 123456

Stuttgart, 1.4.2021

# Ehrenwörtliche Erklärung C:\Users\Martin Engstler\AppData\Local\Microsoft\Windows\INetCache\Content.Word\hdm-logo-4c-office.jpg

|  |  |  |  |
| --- | --- | --- | --- |
| Name: |  | Vorname: |  |
| Matrikel-Nr.: |  | Studiengang: |  |

Hiermit versichere ich, *<Vorname> <Nachname>* ehrenwörtlich, dass ich die vorliegende *(Bachelorarbeit)* mit dem Titel: *„<Titel der Arbeit wie auf dem Deckblatt angegeben>“* selbständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrecht­lichen Folgen (§ 26 Abs. 2 der Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. §19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

.................................................. ..........................................................

Ort, Datum Unterschrift

# Kurzfassung

Text

[Thematische Einordnung, Forschungsziel/-methodik, wesentliches Ergebnis)

Schlagworte:

[4-6 Stichworte in alphabetischer Reihenfolge]

# Abstract

Text

...

Keywords:

…

# Inhaltsverzeichnis

[Ehrenwörtliche Erklärung I](#_Toc104286596)

[Kurzfassung II](#_Toc104286597)

[Abstract II](#_Toc104286598)

[Inhaltsverzeichnis III](#_Toc104286599)

[Abbildungsverzeichnis V](#_Toc104286600)

[Tabellenverzeichnis VI](#_Toc104286601)

[Abkürzungsverzeichnis VII](#_Toc104286602)

[1 Introduction 1](#_Toc104286603)

[2 State of Research 1](#_Toc104286604)

[2.1 Design Science Research 1](#_Toc104286605)

[2.1.1 Overview 1](#_Toc104286606)

[2.1.2 The Cycles 1](#_Toc104286607)

[2.2 Recommender Systems 1](#_Toc104286608)

[2.2.1 Overview 1](#_Toc104286609)

[2.2.2 Deep & Cross Network Recommender System 1](#_Toc104286610)

[2.2.3 TensorFlow Recommenders (TFRS) 1](#_Toc104286611)

[2.3 MLOps 1](#_Toc104286612)

[2.3.1 Overview 1](#_Toc104286613)

[2.3.2 Concept Drift 1](#_Toc104286614)

[2.3.3 TFX (Maybe in Artefact) 1](#_Toc104286615)

[3 Artifact Design 1](#_Toc104286616)

[3.1 Herausforderung 1](#_Toc104286617)

[3.2 Auswahl relevanter Methoden 1](#_Toc104286618)

[3.3 Darstellung der (eigenen) Forschungsmethode 1](#_Toc104286619)

[4 Anwendung des Forschungsansatzes 1](#_Toc104286620)

[4.1 Beschreibung des Anwendungsfalls 2](#_Toc104286621)

[4.2 Umsetzung der Methode 2](#_Toc104286622)

[4.3 Ergebnis und Interpretation 2](#_Toc104286623)

[5 Conclusion and Outlook 2](#_Toc104286624)

[Quellenverzeichnis 3](#_Toc104286625)

[Anlagen 4](#_Toc104286626)

# Abbildungsverzeichnis

[Abbildung 1: Neues Musterbild 2](#_Toc450042050)

[Abbildung 2: Beispieldarstellung 2](#_Toc450042051)

# Tabellenverzeichnis

[Tabelle 1: Bewertungsansatz 3](#_Toc430360163)

# List of abbreviations

ML Machine Learning

GPU Graphics Processing Unit

CPU Central Processing Unit

RS Recommender System

FANG Facebook, Amazon, Netflix, Google

CD Concept Drift

TFX TensorFlow Extended

TFRS TensorFlow Recommenders

DD Data Drift

STEM Academic disciplines of science, technology, engineering and mathematics

PoC Proof of Concept

IS Information System

# Introduction

Over the last two decades Machine Learning (ML) has become one of the fastest growing technical fields with an estimated growth of 21% in 2022 compared to its previous year. (Rimol, 2021)

It managed to secure a position as one of the top fields in computer science for scientific research as well as enterprise adoption. ML combines concepts of linear algebra and statistics and applies them to large datasets to find patterns and generalizations in the data, which can be used to make predictions or classifications.

Leveraging these complex algorithms with the computational power of modern GPUs and CPUs, ML has seen application in a large variety of sectors ranging from medicine for diagnostics, to transportation for self-driving cars and e-commerce for shopping cart optimization. (Choy et al., 2018) The latter sector employs so called Recommender Systems (RS) with the goal of suggesting products that coincide with taste of the customer. With the advent of e-commerce, RSs have gained increasing interest from academia and especially the enterprise sector. (Singh, Choudhury, Dey, & Pramanik, 2021) RSs serve a major role for large tech corporations in engaging, retaining and enticing the user-base of their platform. (Jannach & Zanker, 2022) Netflix for example uses its own RS to suggest its users what movies they might be interested in. In order to incentivize research in the field of Recommender Systems, Netflix introduced the Netflix Prize in 2007: A dataset was made public with users and their movie ratings. The goal was to create a RS that would beat Netflix’ own RS at that time. For this challenge a prize pool of $1.000.000 was written out. (Bennett, Lanning, & others, 2007) To this day datasets of movie ratings remain a popular way to benchmark RSs.

Despite the wide use and success of Recommender Systems and Machine Learning in general, it still is a relatively new field with a lot of research opportunity. (Jordan & Mitchell, 2015) While Recommender Systems are considered integral to many online-platforms, their precision and accuracy often lack in comparison to other ML fields. This is, among other things, due to the nature of the data that Recommender Systems work with, which is often sparse. (Khusro, Ali, & Ullah, 2016) Consequently Recommender Systems are especially susceptible to bad data quality and therefore could profit from comprehensive data curation and monitoring. This lends itself to take a data-centric approach when building, deploying and maintaining a RS, which is one of the subject matters that the field of MLOps sets out to tackle. (Miranda, 2021)

MLOps emerged from the paradigm of DevOps and seeks to apply an automated and standardized approach to the lifecycle of ML applications, similar to what DevOps does for conventional Software. MLOps is attuned to the specific needs and problems of Machine Learning, such that its practices vary from those of DevOps, while still sharing the same goal of rapid and frequent deployment of Software. (Makinen, Skogstrom, Laaksonen, & Mikkonen, 2021) The effect of data quality on the ML model presupposes that data quality management is an integral part of every MLOps system, since data quality affects all aspects of the Machine Learning lifecycle. (Renggli et al., 2021) Detrimental data to the ML system’s performance can manifest itself in different ways.

One manifestation is Concept Drift (CD), which describes a changing outcome *y* to a constant input *x* over time. (Lu et al., 2018) Real world examples of CD could be changing house prices due to a fluctuating house market, or people changing their taste in movies because of aging or genre trends. Deteriorating RS performance due to CD can directly impact the health of the online platform it is used on, as the results of the RSs are reciprocated back to the user experience. For instance, if a movie streaming platform stops recommending appropriate movies to a user because it failed to adapt to the change in taste, the user might stop watching movies on that platform and eventually cancel their subscription. Issues of Concept Drift need to be addressed and mitigated to ensure user-base retention for online services. Additionally, it needs to be incorporated into a MLOps system to benefit from the maintainability, consistency and automation of a unified process.

The product of this work, called an artifact, will be the implementation of a Concept Drift-aware MLOps pipeline for a RS. CD-awareness meaning, that it possesses the ability to account for potential CD in the data.

This paper serves as a thorough documentation of the design of the artifact, which is based on a comprehensive dissection of scientific literature touching the topics of Recommender Systems, MLOps and Concept Drift. The result will then be qualitatively evaluated and discussed.

The MLOps pipeline will be realized with TensorFlow Extended (TFX), a package with various tools to orchestrate and monitor the ML lifecycle. (Baylor et al., 2017) A Deep & Cross Network RS will be implemented with the TensorFlow Recommenders Python API (TFRS). As the dataset, MovieLens 25M will be used, a collection of 25 million movie scores with 62.000 movies and 162.000 users. (Harper & Konstan, 2016)

In the scope of this work, one solution to CD will be designed that is derived from the literature. It is not a comparison and evaluation of various implementations of CD-awareness. This work does not entail a quantitative evaluation of the artifact, as the focus lies on a qualitative analysis of the prototype.

Lastly, the concept of Data Drift (DD) is not subject of this work and thus only will be touched in the context of Concept Drift.

The artifact will be built using the Design Science Research (DSR) methodology from Alan R. Hevner. (Hevner, March, Park, & Ram, 2004)

# State of Research

## Design Science Research

### Overview

Design Science is a research paradigm that stemmed out as a differentiation to Natural Science in STEM. Natural science, also referred to as Behavioral Science, is associated with fields like mathematics, physics, biology and chemistry. Its research methodology follows the objective of uncovering facts and theories about a persistent reality. Juxtaposed to the Natural Science lies the Design Science. Instead of uncovering rules and theories about the nature of reality, design science sets out to engineer and create artifacts with tools from scientific literature. Design Science is predominantly represented in the engineering and computer science fields, where proof of concepts (PoC) and prototypes are the result of a lot of academic works. Both Behavioral Science and Design Science have distinguished approaches on how to conduct research. Design Science Research contains a set of frameworks and best practices to manage academic work in the Design Science department. One of the more prominent methodologies is Alan R. Hevner’s “three cycles” of DSR. (Hevner et al., 2004) Hevner originally designed his framework to involve the research aspect more closely to the development process of Information Systems (IS). It consists of 3 cycles which are closely related to each other and serve to build an artifact. The three cycles are what Hevner argues separates Design Science from other research paradigms. (Hevner, 2007) The artifact is the eventual product of the academic work using DSR. Since its first publication in 2004, DSR has found application in a wide variety of fields that surpasses conventional engineering and computer science. This means that the term “artifact” has a broad definition and is consequently hard to delimit. Generally, an artifact means anything that emerges from Design Science Research. It could range from a theoretical model that was derived from other academic work, to a physical prototype or a production ready software system.

### The Cycles

## Recommender Systems

### Overview

### Deep & Cross Network Recommender System

### TensorFlow Recommenders (TFRS)

## MLOps

### Overview

### Concept Drift

### TFX (Maybe in Artefact)

# Artifact Design

(….)

## Herausforderung

(….)

## Auswahl relevanter Methoden

(….)

## Darstellung der (eigenen) Forschungsmethode

(….)

# Anwendung des Forschungsansatzes

(….)

## Beschreibung des Anwendungsfalls

(….)

## Umsetzung der Methode

(….)

|  |  |  |
| --- | --- | --- |
|  | Aspekt 1 | Aspekt 2 |
| Kriterium 1 |  |  |
| Kriterium 2 |  |  |
| Kriterium 3 |  |  |
| Kriterium 4 |  |  |

Quelle: Autor (2015), S. 1

Tabelle 1: Bewertungsansatz

(…)

## Ergebnis und Interpretation

(….)

# Conclusion and Outlook

(…)

# Quellenverzeichnis

Autor, A. (2015): Musterliteratur für wissenschaftliche Arbeiten, 2. Aufl., Stuttgart: Hochschulverlag

...

# Anlagen

(optional)

References

Baylor, D., Breck, E., Cheng, H.‑T., Fiedel, N., Foo, C. Y., Haque, Z., . . . Zinkevich, M. (2017). TFX: A TensorFlow-Based Production-Scale Machine Learning Platform. In S. Matwin, S. Yu, & F. Farooq (Eds.), *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1387–1395). New York, NY, USA: ACM. https://doi.org/10.1145/3097983.3098021

Bennett, J., Lanning, S., & others (2007). The netflix prize. In *Proceedings of KDD cup and workshop.* Symposium conducted at the meeting of Citeseer.

Choy, G., Khalilzadeh, O., Michalski, M., Do, S., Samir, A. E., Pianykh, O. S., . . . Dreyer, K. J. (2018). Current Applications and Future Impact of Machine Learning in Radiology. *Radiology*, *288*(2), 318–328. https://doi.org/10.1148/radiol.2018171820

Harper, F. M., & Konstan, J. A. (2016). The MovieLens Datasets. *ACM Transactions on Interactive Intelligent Systems*, *5*(4), 1–19. https://doi.org/10.1145/2827872

Hevner, March, Park, & Ram (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75. https://doi.org/10.2307/25148625

Hevner, A. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, *19*.

Jannach, D., & Zanker, M. (2022). Value and Impact of Recommender Systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 519–546). New York, NY: Springer US. https://doi.org/10.1007/978-1-0716-2197-4\_14

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science (New York, N.Y.)*, *349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Khusro, S., Ali, Z., & Ullah, I. (2016). Recommender Systems: Issues, Challenges, and Research Opportunities. In K. J. Kim & N. Joukov (Eds.), *Lecture Notes in Electrical Engineering. Information Science and Applications (ICISA) 2016* (Vol. 376, pp. 1179–1189). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-10-0557-2\_112

Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., & Zhang, G. (2018). Learning under Concept Drift: A Review. *IEEE Transactions on Knowledge and Data Engineering*, 1. https://doi.org/10.1109/TKDE.2018.2876857

Makinen, S., Skogstrom, H., Laaksonen, E., & Mikkonen, T. (2021). Who Needs MLOps: What Data Scientists Seek to Accomplish and How Can MLOps Help? In *2021 IEEE/ACM 1st Workshop on AI Engineering - Software Engineering for AI (WAIN)* (pp. 109–112). IEEE. https://doi.org/10.1109/WAIN52551.2021.00024

Miranda, L. J. (2021). Towards data-centric machine learning: a short review. *Ljvmiranda921. Github. Io*.

Renggli, C., Rimanic, L., Gürel, N. M., Karlaš, B., Wu, W., & Zhang, C. (2021, February 15). *A Data Quality-Driven View of MLOps*. Retrieved from http://arxiv.org/pdf/2102.07750v1

Rimol, M. (2021). Gartner Forecasts Worldwide Artificial Intelligence Software Market to Reach $62 Billion in 2022. Retrieved from https://www.gartner.com/en/newsroom/press-releases/2021-11-22-gartner-forecasts-worldwide-artificial-intelligence-software-market-to-reach-62-billion-in-2022

Singh, P. K., Choudhury, P., Dey, A. K., & Pramanik, P. K. D. (2021). Recommender systems: an overview, research trends, and future directions. *International Journal of Business and Systems Research*, *15*(1), 14. https://doi.org/10.1504/ijbsr.2021.10033303