

Data Exploration Project

M. Sc. Daniel Wehner
SAP SE / DHBW Mannheim
Sommersemester 2020

Inhaltsverzeichnis zur Vorlesung

Planung und Organisatorisches	3
Grundsätzliches zum Projekt	3
Bearbeitung und Abgabe	4
Anwesenheitspflicht	7
Bewertung des Projekts	8
Zeitlicher Ablauf des Projekts	9
 Themen	 10
Eigene Themen	10
Themenvorschläge	11

Planung und Organisatorisches

Grundsätzliches zum Projekt

- Name der Veranstaltung: *Data Exploration Project*
- Laut [⇒ Modulkatalog](#) beträgt der Workload des Projekts **pro Person**:
 - Präsenzzeit: 27 Stunden
 - Selbststudium: 47 Stunden
- Definition des Projekts aus dem [⇒ Modulkatalog](#):

„Anwendung von Methoden und Verfahren des maschinellen Lernens auf eine vorgegebene Datenbasis unter Laborbedingungen. Verwendung von üblichen Repositorien wie Hadoop/Spark/Flink/Mahout, Python-RASBT, R, etc. **Ein besonderer Fokus soll auf einer ganzheitlichen wirtschaftsinformatischen Betrachtung liegen.** Es soll dabei neben der informatischen Betrachtung auch der betriebswirtschaftliche Nutzen, z. B. anhand eines Use Cases, betrachtet werden.“

Bearbeitung und Abgabe

Bearbeitung

- Das Projekt ist in Gruppen von **maximal vier Studierenden** zu bearbeiten (bei Bedarf können auch kleinere Gruppen gebildet werden, größere Gruppen sind jedoch nicht erlaubt).
- Jede Gruppe bearbeitet ein anderes Thema.
- Mögliche Projektthemen: siehe ...
- Bzgl. der zu verwendenden Technologien werden keine Einschränkungen gemacht.



Bei der Themenwahl ist darauf zu achten, den Umfang weder zu gering noch zu groß zu wählen!

Zwischenpräsentation

- Die Präsentation beträgt **maximal 10 Minuten** und dient dem Zweck, die grobe Konzeption des Projektvorhabens darzulegen und zu präsentieren.
- Sie geht **nicht** mit in die Endwertung ein.
- Von Seiten der Projektgruppen sind zu diesem Termin keine Arbeitsergebnisse oder Dokumente einzureichen.

Projektabschluss am Ende des Semesters

- Am Ende des Semesters sind eine Abschlusspräsentation und ein Projektreport anzufertigen.
- Die Abschlusspräsentation beträgt **25 Minuten**.
- Der Projektreport ist gemäß den Regeln des wissenschaftlichen Arbeitens anzufertigen und weist einen Umfang von **minimal 3 und maximal 4 Seiten** (ohne Abbildungen) auf.
- Die Termine können der \Rightarrow [table 1](#) entnommen werden.

- **Tipp:** Online gibt es bereits diverse \LaTeX -Vorlagen für die Anfertigung von Projektreports; z. B.:
 - ShareLaTeX: \Rightarrow [Vorlage American Institute of Physics \(AIP\)](#)
 - Overleaf: \Rightarrow [Report Vorlagen Overleaf](#)
- Der Projektreport deckt mindestens folgende Bestandteile ab:
 - Thema und Motivation
 - Related Work (*welche wissenschaftlichen Publikationen gibt es zu diesem Thema?*)
 - Verwendete Technologien und Bibliotheken (z. B. `scikit-learn`, `tensorflow`, ...)
 - Anmerkungen zum Quellcode (*wie ist der Code auszuführen und was gibt es zu beachten?*)
 - Präsentation der Ergebnisse
 - Kritische Bewertung der Ergebnisse („*lessons learned*“)



Die komplette Abgabe besteht aus dem Projektreport (*.pdf), der Abschlusspräsentation (*.pdf), sowie dem im Rahmen des Projekts erstellten Quellcode.

Anwesenheitspflicht

- Mit wenigen Ausnahmen ist die Anwesenheit der einzelnen Projektgruppen an den Terminen **freiwillig**.
 - Es sollte jedoch regelmäßig Rücksprache bzgl. des Zwischenstands gehalten werden.
 - Bitte melden Sie sich **rechtzeitig** und **eigenverantwortlich** (z. B. per E-Mail: daniel.wehner@sap.com), falls von Ihrer Seite aus Diskussionsbedarf besteht.
- **Anwesenheitspflicht** besteht an folgenden Terminen (siehe ⇒ [table 1](#)):
 - Zwischenpräsentation
 - Finale Präsentation und Abgabe

Bewertung des Projekts

- Nur rechtzeitig eingereichte Dokumente können bewertet werden.
- Die einzelnen Bestandteile werden folgendermaßen gewichtet:
 - Quellcode und Ergebnisse (70 %)
 - Projektreport (20 %)
 - Abschlusspräsentation (10 %)



Wichtig: Das Fehlen einer Teilabgabe führt zu erheblichem Punkteabzug (unter Umständen auch zum Nichtbestehen der gesamten Veranstaltung)!

Zeitlicher Ablauf des Projekts

Termin	Bemerkung	Anwesenheitspflicht
xx.xx.2020	Einführung	ja
xx.xx.2020	<i>Bedarf bitte anmelden</i>	nein
xx.xx.2020	Zwischenpräsentation (alle Gruppen)	ja
xx.xx.2020	<i>Bedarf bitte anmelden</i>	nein
xx.xx.2020	<i>Bedarf bitte anmelden</i>	nein
xx.xx.2020	<i>Bedarf bitte anmelden</i>	nein
xx.xx.2020	<i>Bedarf bitte anmelden</i>	nein
xx.xx.2020	Abschlusspräsentationen und Abgabe (alle Gruppen)	ja

Tabelle 1:

Ablaufplan des Projekts

Themen

Eigene Themen

- Laut [⇒ Modulkatalog](#) soll der Fokus auf einer „*ganzheitlichen wirtschaftsinformatischen Betrachtung*“ liegen, und auch dem betriebswirtschaftlichen Aspekt Rechnung getragen werden.
- Es ist grundsätzlich erlaubt und auch erwünscht, **eigene Themenvorschläge** einzubringen.
- Eigene Themen müssen natürlich vorher genehmigt werden.
- Auf der nächsten Folie sind einige Projektvorschläge aufgelistet, falls einzelne Gruppen kein eigenes Thema finden sollten.

Themenvorschläge

- Sentiment Analyse von Kundenrezessionen (kommt das Produkt beim Kunden gut oder schlecht an?)
- Vorhersage von Aktienkursen (falls Sie reich werden möchten, ist das ein guter Anfang)
- Recommender Systems (z. B. *Collaborative Filtering*)
- Baue deinen eigenen Chatbot (z. B. zur automatischen Beantwortung von Kundenfragen)
- Speech Recognition
- Analyse medizinischer Scans zur Krankheitsdiagnose