

Data Science 2

Projektphase

Prof. Dr. Christian Bockermann

23. November 2022

Dieses Dokument beschreibt die Projektphase des Kurses *Data Science 2*, sowie die Bearbeitungs- und Abgabemodalitäten.

1 Zielsetzung der Projektphase

Der Kurs *Data Science 1+2* gibt eine Einführung in die Datenanalyse und umfasst sowohl die grundlegende Programmierung in Python mit Fokus auf die Datenanalyse, als auch die Vermittlung (in einem Überblick) der theoretischen Grundlagen zum Maschinellen Lernen (Lernaufgaben, Algorithmen).

Das Ziel des 2-semesterigen Kurses ist insbesondere auch die praktische Umsetzung von einfachen Analysen zu erlernen. Dies umfasst die Datenvorverarbeitung und Exploration, sowie die Modellierung von Daten mit Hilfe der vermittelten Lernverfahren.

In der **Projektphase** sollen die Studierenden das erlernte Wissen an einem zuvor im Kurs nicht behandelten Datensatz vertiefen. Dabei geht es sowohl darum, die Daten einzulesen und zu explorieren, als auch die Fragestellung, die mit Hilfe der Analyse der Daten beantwortet werden soll, selbstständig zu entwickeln.

Dies dient zum einen der Vermittlung eines analytischen Vorgehens und schult zum anderen ein strukturelles Arbeiten. Die Bearbeitung erfolgt zudem in Gruppen zu 4-5 Studierenden, so dass ein (interdisziplinärer) Austausch zwischen den Teilnehmern/Teilnehmerinnen gefördert wird.

2 Gruppenarbeit und Themenwahl

Die Projekte werden in Gruppen bearbeitet. Die Einteilung der Gruppen findet zu Beginn der Projektphase durch die Studierenden selbst statt.

Jede Gruppe wählt aus einer Liste von vorgeschlagenen Datensätzen/Projekten den Datensatz, den sie in der Projektphase bearbeiten will. Die Gruppe hat dabei die folgenden Aufgaben:

1. Definition/Ableitung einer Fragestellung zur Analyse
2. Exploration des Datensatzes
3. Analyse der Daten in Hinblick auf die Fragestellung.

Durch die offene Aufgabenstellung können die Studierenden der Gruppen auch mehrere Fragestellungen erarbeiten und die Daten dahingehend analysieren.

3 Ablauf und Abgabe der Projektarbeit

Nach Bekanntgabe der Themen/Datensätze ca. in der Mitte der Vorlesungsreihe bearbeiten die Studierenden die Projektaufgabe als Gruppe bis zum Ende des Semesters. Während der Vorlesungszeiten finden Vorlesungen kurze Status-Treffen zum aktuellen Fortschritt der Projekte statt. Hier (und in den zugehörigen Übungen) bietet sich die Gelegenheit, Probleme zu diskutieren und Unterstützung zu bekommen.

Die Abgabe der Projektarbeit erfolgt als Gruppenabgabe, d.h. jede Gruppe gibt zum Ende der Bearbeitungszeit ein gemeinsames Jupyter-Notebook als Analyse Ergebnis ab. Dieses Notebook soll:

- Den Datensatz und die Fragestellung vorstellen
- Explorationen der Daten zeigen
- die Schritte der Datenvorverarbeitung und Analyse zeigen

Zusätzlich zum Notebook erstellt jede Gruppe eine gemeinsame Präsentation (z.B. mit PowerPoint, LaTeX Beamer, PDF) die das Vorgehen der Gruppe bei der Bearbeitung des Projektes und die Ergebnisse zeigt. Die Präsentation wird von der Gruppe am Präsentationstag gehalten und stellt den Abschluss der Projektarbeit dar.

Neben der Präsentation des Projektes sollte die Präsentation zu Beginn auch eine kurze Vorstellung der Gruppe enthalten. Gerade im Rahmen reiner Online-Veranstaltungen ist es hilfreich nochmal kurz zu sehen, wie sich die Gruppen zusammensetzen (Namen der Teilnehmer, ggf. Fachsemester) und aus welchen verschiedenen Fächern die Gruppenmitglieder kommen.

4 Bewertung

Die **Bewertung** der Projektarbeit erfolgt für jede Gruppe zusammenfassend, d.h. jeder Teilnehmer/jede Teilnehmerin einer Gruppe bekommt in der Regel die gleiche Note. Dies setzt jedoch voraus, dass die Studierenden in der gemeinsamen Präsentation die Beiträge der einzelnen Gruppenmitglieder mit herausstellen.

Sofern der Beitrag einzelner Teilnehmer/Teilnehmerinnen zum Gruppenergebnis nicht klar erkennbar ist oder weit unter dem durchschnittlichen Beitrag in der Gruppe liegt,

kann die Einzelbewertung von der Gruppenbewertung abweichen. Hierfür ist ggf. ein Gespräch/mündliches Kolloquium mit dem Prüfer im Anschluss an die Gruppenpräsentationen erforderlich.