

Applied Analytical Data Science

- Übung 5 -

Dr. Jörg-Uwe Kietz,
Übung zur Vorlesung,

Mittwoch, 14-16 Uhr Vorlesung,
16-18 Uhr Übung (alle 2 Wochen)

<http://www.kietz.ch/AADS/>

Vorhanden von den letzten Übungen

- 0) Der Datensatz „CleanedEvalData“
- 1) Ein Prozess zum Preprocessing mit allen Modellen
- 2) Ein Prozess zum Klassifikationslernen (mit CPE), ob jemand spendet,
- 3) Ein Prozess zum Regressionslernen, wieviel er wohl spendet
- 4) Eine Evaluation (Profitchart) auf Testdaten, ob 2) oder 3) oder die Kombination von 2) und 3) am besten ist.

5. Aufgabe: Anwenden auf Anwendungsdaten

Wendet

- das Preprocessing und
- das Klassifikationsmodell und
- das Regressionsmodell

auf die CleanedEvalData an.

5. Aufgabe: Result tabelle

Produziert und exportiert eine Tabelle mit den **genau den Feldern in genau dieser Reihenfolge** für **alle Datensätze in CleanedEvalData**:

- CONTROLN
- Anschreiben (integer): 1 (Ja) / 0 (Nein)
- CPE-Order (double) (confidence_true der Klassifikation)
- Reg-Order (double) (predicted, der Regression)
- Combined-Order (double) (Produkt von CPE- und Reg-Order)

und exportiert sie als grX.csv Datei (','-separiert, Spaltennamen in der ersten Zeile).

Wer eine Tabelle mit fehlenden, falschen oder zuvielen oder falsch benannten oder angeordneten Feldern oder records produziert ist (wegen unkontrolliertem Pfusch) durchgefallen.

Abgabe

- Prozess und exportierte Tabelle via E-mail (bis end of 22.5.)
- Businessvortrag (ca 15min) für den Endkunden über euer “Projekt” und alle erzielten/verkaufbaren Ergebnisse (am 30.5.).