

Gruppenarbeit KW 44, 30.10.2023-02.11.2023

Im folgenden Praxisprojekt soll eine Gruppenarbeit simuliert werden, wie diese auch in dem IHK-Modul auftreten kann. Ihr werdet für diese Arbeit in Gruppen mit jeweils 4 Personen aufgeteilt, welche dann an dem Projekt arbeiten sollen. Für jede Gruppe stehen alle Informationen frei zu Verfügung, die sich direkt unter dem Link abrufen lassen, eine Kurzzusammenfassung findet ihr auch im weiteren Verlauf dieser Aufgabenstellung. Sollten hierbei Variablen nicht vollständig beschrieben werden oder nicht klar werden, welche Inhalte die Variable umfasst, so muss dafür entweder eine begründete Annahme getroffen werden wie diese interpretiert wird oder eine begründete Annahme erfolgen, warum diese ausgeschlossen werden sollte.

Folgendes soll vorbereitet werden:

Datenanalyse per Python

- Strukturierter Workflow (Datenimport, Datenbereinigung, Datentransformation, Datenanalyse, Datenexport)
- Explorative Datenanalyse zu den importierten Daten
- Datenanalyse per Machine Learning mit Bewertung, Optimierungs- und Stabilitätstests
- Einfache (grafische) Darstellung der Ergebnisse

Datenvisualisierung

- Bericht oder Dashboard mit Visualisierung der Daten
- Interaktive Elemente zur flexiblen Erkundung der Informationen
- Darstellung des Zusammenhangs der Informationen (Welche Geschichte soll erzählt werden?)
- Kurze Dokumentation der Datenbearbeitung (reicht in Power-BI, zB als extra Berichtssseite / Willkommensblatt)

Präsentation

- Eine 30-minütige Präsentation, in welcher die Problemstellung erläutert, das Vorgehen dargestellt und begründet wird und letztlich die Problemstellung mit einem Fazit beantwortet wird.
- In dieser Präsentation dürfen Ausschnitte des Codes verwendet werden, wenn diese nötig sind, um das Vorgehen zu verstehen. Achtung: Geht davon aus, dass der Kunde vor euch keine Ahnung von Code hat!
- Das Power-BI Dashboard darf live gezeigt werden.

Es ist ausdrücklich KEINE Projektbeschreibungsdatei als extra nötig – konzentriert euch in dieser Zeit unbedingt vollkommen auf die Skripte/Boards, die Gruppenarbeit und das Verstehen des Datensatzes.

Zeitplan

30.+31.10.2023 -> Gruppenarbeit

01.11.2023 -> Feiertag und kein Kurs, Abgabe der Unterlagen bis 20 Uhr

02.11.2023 -> Präsentationen

03.11.2023 -> Wiederholungstag mit Ray

Die Projekte

Gruppe 1

Personen: Ingo, Anca, Bernd, Constantin

Thema: Bank Marketing, „The classification goal is to predict if the client will subscribe (yes/no)“

Link: <http://archive.ics.uci.edu/dataset/222/bank+marketing>

Abhängige Variable im Datensatz: y

Gruppe 2

Personen: Nikolaos, Brigitte, Kacper, Rene

Thema: In-vehicle coupon recommendation, „This data studies whether a person will accept the coupon recommended to him in different driving scenarios“

Link: <http://archive.ics.uci.edu/dataset/603/in+vehicle+coupon+recommendation>

Abhängige Variable: Y

Gruppe 3

Personen: Hyun, Stefan, Claudia, Sarah

Thema: Online Shoppers Purchasing Intention Dataset, „Of the 12,330 sessions in the dataset, 84.5% (10,422) were negative class samples that did not end with shopping, and the rest (1908) were positive class samples ending with shopping.“

Link: <http://archive.ics.uci.edu/dataset/468/online+shoppers+purchasing+intention+dataset>

Abhängige Variable: Revenue

Gruppe 4

Personen: Benjamin, Klaus, Svetlana, Viktor

Thema: South German Credit, „700 good and 300 bad credits with 20 predictor variables. Data from 1973 to 1975.“

Link: <http://archive.ics.uci.edu/dataset/573/south+german+credit+update>

Abhängige Variable: kredit