

Aufgabe 1 – Gebrauchtwagenpreise

Hinweis: Die Daten werden im bereitgestellten Skript direkt von GitHub geladen.

Im bereitgestellten Datensatz erhalten Sie Informationen zu knapp 100.000 zum Verkauf stehenden Fahrzeugen in UK. Nutzen Sie Ihre Datenanalysefähigkeiten, um die folgenden Fragen zu beantworten:

- a) Sie vermuten einen Zusammenhang zwischen dem Alter und dem Wert eines Fahrzeuges. Stellen Sie den Zusammenhang grafisch dar und quantifizieren Sie ihn.
- b) Der Datensatz enthält weitere interessante Informationen. Probieren Sie den Erklärungsgehalt Ihres Modells zu verbessern und interpretieren Sie die jeweiligen Ergebnisse.

Aufgabe 2 – Basketball

Hinweis: Die Daten werden im bereitgestellten Skript direkt von GitHub geladen.

Im bereitgestellten Datensatz erhalten Sie detaillierte Informationen zu ca. 25.000 Basketballwürfen verschiedener Profispieler. Unser Ziel ist es auf den Daten ein Modell zu trainieren das für zukünftige Würfe Treffer vorhersagt.

- a) Um welche Art von Lernproblem handelt es sich?
- b) Analysieren Sie die Daten und versuchen Sie händisch ein geeignetes Modell für die Vorhersage zu entwickeln.
- c) Trainieren Sie einen Entscheidungsbaum zur Vorhersage der einzelnen Würfe. Achten Sie darauf einen Train-Test Split zu verwenden.

Aufgabe 3 – Counter-Strike Global Offensive

Hinweis: Die Daten werden im bereitgestellten Skript direkt von GitHub geladen.

Im bereitgestellten Datensatz erhalten Sie detaillierte Informationen zu ca. 122.000 Snapshots aus 700 Runden des beliebten Videospiele Counter-Strike Global Offensive. Nutzen Sie die Daten um die folgenden Fragen zu beantworten:

- a) Wie oft gewinnt das Team mit dem meisten Geld?
- b) Welche Variablen sind besonders geeignet den Sieger vorherzusagen?
- c) Wie gut kann ein Entscheidungsbaum den Ausgang einer unbekanntes Runde vorhersagen?