

Arbeitsauftrag: Projekt

Als Data Scientist ist es eure Aufgabe, Probleme durch Erstellen eines Klassifikationsmodells zu lösen. Dazu steht euch ein Datensatz zur Verfügung, der den Ausgangspunkt für euer KI-Modell darstellt.

Folgt dazu folgenden Schritten:

1. Erstellt ein neues Orange3-Projekt.
 - Wenn euer Datensatz bereits in Orange3 enthalten ist, zieht das *Datasets*-Widget auf die Leinwand und wählt den entsprechenden Datensatz mit einem Doppelklick aus.
 - Wenn euer Datensatz nicht in Orange3 enthalten ist, öffnet ihn wie den Äffchendatensatz mithilfe des *File*-Widgets.
2. Verschafft euch anschließend einen Überblick über den Datensatz mithilfe des *Data Table*-Widgets! Notiert, welche Informationen in den Spalten gespeichert sind! Was sind die möglichen Klassen? Welche Merkmale stehen euch zur Verfügung?
3. Teilt die Daten mithilfe des *Data Sampler*-Widget in Trainings- und Testdaten auf.
4. Trainiert mehrere Modelle, indem ihr unterschiedliche Merkmale für das Training auswählt und verschiedene Parameter des Trainingsprozesses festlegt!
5. Testet und bewertet eure Modelle mithilfe einer geeigneten Metrik.
6. Überlegt euch, welche Personengruppen wie von eurem Modell profitieren und welche Personengruppen wie von eurem Modell benachteiligt werden könnten!
7. Gestaltet ein Poster, das euren Datensatz, euer Ziel und eure Ergebnisse knapp und präzise zusammenfasst! Was findet ihr an eurer Auswertung überraschend?



Datasets



File