

Einführung in die künstliche Intelligenz

Aufgabe “ML Regression”

Prof. Dr. A. del Pino

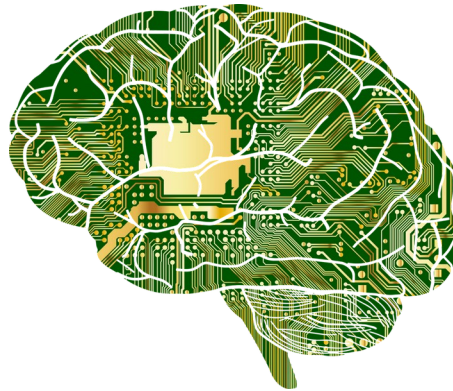


Bild: https://en.wikipedia.org/wiki/Machine_learning#/media/File:Anatomy-1751201_1280.png

Aufgabe “ML Regression”

In dieser Aufgabe geht es darum, ML Regression im Rahmen des Kaggle Wettbewerbs “TMDB Box Office Prediction” <https://www.kaggle.com/competitions/tmdb-box-office-prediction> anzuwenden. Bearbeiten Sie diese Aufgabe in den festgelegten 2er-Teams.

- Falls noch nicht getan, dann lesen Sie die Abschnitte 2.1 bis 2.4 des Buchs “Applied Artificial Intelligence – An Engineering Approach” mit Schwerpunkt auf die Abschnitte, welche relevant für diesen Kurs sind.
- Beantworten Sie den Mini-Test “ML Regression” in Moodle
- Laden Sie die Trainings- und Testdaten, sowie die Beispiel-Datei für eine Wettbewerbs-Einreichung von <https://www.kaggle.com/competitions/tmdb-box-office-prediction/data> herunter.
- Analysieren Sie den Datensatz
- Legen Sie ein neues Jupyter Notebook “TMDB.ipynb” an.
- Kopieren Sie die Dateien in ein Unterverzeichnis “data” und implementieren Sie Ihre Lösung für den Kaggle Wettbewerb in dem Jupyter Notebook.

Aufgabe “ML Regression” ...

- Führen Sie eine Kreuzvalidierung durch. Gegebenenfalls können Sie versuchen, Ihr Ergebnis zu verbessern, z.B. durch andere Vorverarbeitung, andere ML Ansätze, andere ML Parameter...
- Reichen Sie Ihren Beitrag zum Kaggle Wettbewerb ein
<https://www.kaggle.com/competitions/tmdb-box-office-prediction> (Knopf “Submit Predictions”).
- Überprüfen Sie die Position Ihres Beitrag in dem leaderboard
<https://www.kaggle.com/competitions/tmdb-box-office-prediction/leaderboard> (Knopf “Jump to your leaderboard position”).
- Gegebenenfalls können Sie versuchen, Ihr Ergebnis zu verbessern (andere Vorverarbeitung, andere ML Ansätze, andere ML Parameter, ...) und nochmals einreichen.
- Machen Sie einen Screenshot von dem leaderboard mit Ihrem Beitrag.
- Laden Sie den Screenshot und Ihre Jupyter Notebook Datei “TMBD.ipynb” in Moodle hoch