

Programmierabgabe 2 - Gruppe 1

Regression zur Bewertung von Weinqualität mit Deep Learning

Datenset

Das Wine Quality Data set beinhaltet 11 chemische Eigenschaften für roten (1599 Samples) oder weißen (4898 Samples) "Vinho Verde" Wein aus Portugal. Jedem Sample ist ein Qualitäts-Score als Label zugeordnet. Beachten Sie, dass es sich bei dem Datenset um ein unbalanced Datenset handelt, d.h. dass sie dies bei Ihrem Splitting in Trainings- und Testdaten sowie den gewählten Evaluationsmethoden berücksichtigen müssen. Führen Sie alle Aufgabenschritte sowohl für weißen als auch für roten Wein durch.

Sie können das Datenset unter <https://archive.ics.uci.edu/ml/datasets/wine+quality> downloaden.

Aufgabe

Implementieren Sie ein ML-System zur Regression für die Vorhersage der Weinqualität mithilfe Deep Learning. Nutzen Sie dafür Implementierungen in Keras/Tensorflow 2. Bearbeiten Sie dabei die folgenden Schritte:

1. Laden Sie die Daten und führen Sie ein Splitting in Trainings- und Testdaten (Verhältnis 75:25) durch.
2. Überlegen Sie sich eine geeignete Netzwerkarchitektur für die Regressionsaufgabe
3. Implementieren Sie Ihr Netzwerk in Keras/Tensorflow.
4. Wählen Sie eine geeignete Loss-Funktion und einen geeigneten Optimizer. Begründen Sie Ihre Auswahl in Ihrem Kurzreport.
5. Evaluieren Sie Ihr Deep-Learning-System für beide Datensets und stellen Sie die Ergebnisse in geeigneter Weise dar. Nutzen Sie zur Evaluation eine Kreuzvalidierung (3-fold). Vergleichen Sie die Ergebnisse für roten und weißen Wein.