

Informatik im Bau- und Umweltwesen 2

Aufgabenstellung zur 2. Blockübung



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Univ.-Prof. Dr.-Ing. Uwe Rüppel
Sommersemester 2019

04. Juli 2019

Exemplarische Anwendung der Machine-Learning-Methoden an Beispielen aus dem Bau- und Umweltingenieurwesen

Einleitung

Neueste Errungenschaften im Bereich des Maschinellen Lernens (ML) haben einen enormen Beitrag zur Weiterentwicklung von Künstlicher Intelligenz (KI) und kognitiven Systemen geleistet. In Zukunft werden Menschen in vielen Kontexten lernenden Systemen begegnen. Auch in der Bau- und Umweltbranche entstehen durch Digitalisierung täglich große Mengen an Daten, die durch Data Mining und maschinelles Lernen den Ingenieuren beim Entscheidungstreffen helfen können. In dieser Übung bekommen Sie zwei Datensätze "Concrete_Data_Yeh.csv" und "weatherAUS.csv", jeweils aus dem Bauingenieurwesen und der Umweltwissenschaft. Bitte entwickeln Sie anhand der demonstratischen Quellcode und die in der Hörsaalübung vorgestellte Python Pakete jeweils ein neues Modell, das zum Vorhersagen der Ergebnisse den Ingenieuren helfen können.

Python Pakete:

- Numpy, Pandas, scikit-learn
- Matplotlib, Seaborn (optional)

Aufgabe 1 Betonfestigkeit analysieren - Regression

Bitte Recherchieren Sie die Zusammensetzung von Beton und die Grundlage der Festigkeitsentwicklung. Analysieren Sie welche Faktoren die Festigkeitsentwicklung beeinflussen und wenden Sie Ihres Fachwissen bei der Featureauswahl. Achten Sie dabei die Datenqualität und bereiten Sie bitte Ihrer Daten so, dass sie dem Modelltraining vernünftig (Preprocessing). Splitten Sie bitte Ihrer Daten dem Preprocessing in Trainingsdaten (inklusive Daten zur Validierung) und Testdaten. Wählen Sie im Anschluss bitte einen geeigneten Regression-Algorithmus und Validierungsmethod aus der Python Bibliothek "scikit-learn". Optimieren Sie Ihre Modelle anhand der Umgestaltung deren Parameter. Analysieren und evaluieren Sie Ihre Outputs aus den Modellen mit den Testdaten. Entspricht es Ihrer Erwartung?

Aufgabe 2 Regen in Australien vorhersagen - Klassifizierung

Bitte Recherchieren Sie die Einflussfaktoren von Niederlage. Analysieren Sie welche Faktoren tragen dem Regen bei. Wenden Sie bitte Ihres Fachwissen bei der Feature Auswahl. Achten Sie dabei die Datenqualität und bereiten Sie bitte Ihrer Daten so, dass sie dem Modelltraining vernünftigt (Preprocessing). Splitten Sie bitte Ihrer Daten dem Preprocessing in Trainingsdaten (inklusive Daten zur Validierung) und Testdaten. Wählen Sie bitte im Anschluss einen geeigneten Klassifikation-Algorithmus und Validierungsmethode aus der Python Bibliothek "scikit-learn". Optimieren Sie Ihre Modelle anhand der Umgestaltung deren Parameter. Analysieren und evaluieren Sie Ihre Outputs aus den Modellen mit den Testdaten. Entspricht es Ihrer Erwartung?

Aufgabe 3 Bewertung UML-Diagramme

Bewerten Sie die Ihrer Gruppe zugeteilten ERM Diagramm. Verwenden Sie hierfür den zur Verfügung gestellten Bewertungsbogen. Bitte gehen Sie dabei sowohl auf positive als auch auf negative Aspekte ein und geben Sie ggf. Verbesserungsvorschläge.

Allgemeine Hinweise

- Bitte achten Sie beim Anfertigen der Übung auf Portabilität. Die Projekte müssen im Rahmen des Kolloquiums innerhalb kürzester Zeit auf den zur Verfügung gestellten Rechnern lauffähig sein.
- Die Betreuungstermine der Blockübung sind unter http://www.iib.tu-darmstadt.de/lehre_iib/mastermodule/iib2/iib2.de.jsp zu finden.
- Die Abgabe der Übungsergebnisse sowie der Dokumentation muss bis zum **Sonntag, den 21.07.2019 um 23:55 Uhr** über Moodle erfolgen.
- Die Kolloquien zur Blockübung finden voraussichtlich in der 30/31. KW statt, die individuellen Termine werden noch bekannt gegeben.
- Die Durchführung der Blockübung und der Kolloquien erfolgt in Gruppen zu je zwei Studierenden.
- Für die Datensicherung ist jede Gruppe selbst verantwortlich.
- **Übungen, die die vorgegebenen Namenskonventionen nicht beachten, werden nicht korrigiert!**
- Es wird nur die Abgabeversion beim Kolloquium bewertet. Sollte im Kolloquium versucht werden, eine aktualisierte Version zu verwenden, wird die Übung mit 'nicht bestanden' gewertet.
- Während des Kolloquiums stehen nur die PCs im L5|01 Raum 222 zur Verfügung.
- Bei ungenügenden Abgaben können einmalig pro Modul zum Bestehen erforderliche Nachbesserungen verlangt werden, in diesem Fall kann eine Übung jedoch nur mit maximal 1 Punkt bewertet werden.

Abzugeben sind:

- a) Alle Übungsinhalte in **einer** ZIP-Datei
Namenskonvention: **IIB2_UE1_GruppeXX.zip**
- b) Python-Projekte:
Namenskonvention: **IIB2_UE1_GruppeXX**
- c) Dokumentation (doc- oder pdf-Datei)

- Die Übung ist zu dokumentieren!
- Namenskonvention: **IIB2_UE1_GruppeXX_Doku.doc(x)** (bzw. *.pdf)
- Bitte die Formatvorlage der TU Darmstadt verwenden und auf dem Titelblatt Gruppennummer, sowie Namen, Studiengänge und Matrikelnummern der Teilnehmer vermerken.

Hinweise zur Abgabe:

- Nur ein Gruppenpartner muss die Lösung auf Moodle stellen.
- Laden Sie die Abgabe rechtzeitig hoch. Eine verspätete Abgabe ist nicht möglich!
- Bitte überprüfen Sie nach dem Upload, ob dieser erfolgreich war!
- Prüfen Sie vor der Abgabe die Portabilität des Projektes auf andere Computer!

Viel Erfolg!