

uebungsblatt03

May 17, 2019

###

Inhaltsbasierte Bild- und Videoanalyse - SoSe 19

Übungsblatt 3

Abgabe: So. 19.05.2018 - 23:59 Uhr

Hinweis: Alle Lösungen sollen in einem IPython Notebook realisiert werden, wobei Teilaufgaben und Zwischenergebnisse ausgegeben bzw. visualisiert werden sollen. Die Übungsblatt sollen in Gruppen von 3 Personen abgegeben werden. Die Datei soll nach folgendem Muster benannt werden: uebung03 {nachname1} {nachname2} {nachname3}.ipynb Für eine Einleitung in IPython siehe z.B.: <http://cs231n.github.io/ipython-tutorial/>

1 Aufgabe 1 - Machine Learning, Definitions, Algorithms (1+1+2 Punkte)

1.1 Aufgabe 1.1: Machine Learning Definition (1 Punkt)

Welche der folgenden Aussagen ist eine sinnvolle Definition für maschinelles Lernen?

- a) Maschinelles Lernen ist die Wissenschaft der Computerprogrammierung.
- b) Maschinelles Lernen ist der Forschungsbereich, der Computern die Fähigkeit vermittelt zu lernen ohne explizit programmiert zu werden.
- c) Maschinelles Lernen ist der Forschungsbereich, der Robotern erlaubt intelligent zu handeln.
- d) Maschinelles Lernen nutzt "gelabelte" Daten zum Lernen.

1.2 Aufgabe 1.2: Classification, Regression (1 Punkt)

Die folgenden Aufgaben sollen mit Machine Learning gelöst werden. Handelt es sich dabei jeweils um ein Regressions- oder Klassifikationsproblem?

- a) Angenommen Sie beschäftigen sich mit dem Problem der Wettervorhersage und wollen vorhersagen ob es morgen um 17:00 Uhr regnen wird.
- b) Gegeben sind Bilder mit Gesichtern und das Alter der darauf gezeigten Person soll geschätzt werden.

1.3 Aufgabe 1.3: Supervised vs. unsupervised Learning (2 Punkt)

Auf welche der folgenden Probleme würden Sie ein überwachtes Lernverfahren anwenden? Gehen Sie davon aus, dass die jeweils benötigten Trainingsdaten vorhanden sind.

- a) Take a collection of 1000 essays written on the US Economy, and find a way to automatically group these essays into a small number of groups of essays that are somehow "similar" or "related".
- b) Examine the statistics of two football teams, and predicting which team will win tomorrow's match (given historical data of teams' wins/losses to learn from).
- c) Examine a large collection of emails that are known to be spam email, to discover if there are sub-types of spam mail.
- d) Given genetic (DNA) data from a person, predict the odds of him/her developing diabetes over the next 10 years.

Aufgabe 2 - CIFAR-10 (4 Punkte)

a) Laden Sie die Python-Version von CIFAR-10 herunter (<https://www.cs.toronto.edu/~kriz/cifar.html>) und machen sie sich mit dem Interface vertraut.

[]:

b) Visualisieren Sie ein zufälliges Bild aus jeder Klasse.

[]: