

Übungsblatt 6

Machine Learning (WS 16/17)
Stefan Edelkamp

12. Januar 2017

Sämtliche Aufgaben sind von der Gruppe selbstständig zu lösen. Die Verwendung von Hilfsmitteln und Quellen außerhalb der Vorlesungsunterlagen gilt es in expliziter Weise zu dokumentieren.
Abgabe ist am Donnerstag, den 2.2.2017 oder per E-Mail an: edelkamp@tzi.de.

1 Begriffsdefinition

1. Welche VC-Dimension haben Intervalle auf dem reellen Zahlenstrahl? (5 P)
2. Was ist die Realisierbarkeitsannahme (für PAC)? (5 P)
3. Vervollständigen Sie den Satz: *Eine Hypothesenklasse \mathcal{H} ist PAC-lernbar, falls es ein Lernverfahren L und eine Funktion $m_{\mathcal{H}}$ abhängig von ϵ und δ gibt, so dass für alle Verteilungen D über \mathcal{X} und jede Klassifizierungsfunktion $f : \mathcal{X} \rightarrow \{0, 1\}$ unter der Realisierbarkeitsannahme und $m \geq m_{\mathcal{H}}(\epsilon, \delta)$ unter D identisch und unabhängig verteilten und mit f beschrifteten Beispielen gilt, ...* (5 P)
4. Wie besagt der *Fundamentalsatz der Lerntheorie* nach Blumer, Ehrenfeucht, Haussler und Warmuth (four-Germans-paper)? (5 P)
5. Was ist *No-Free Lunch* (in der Lerntheorie)? (5 P)
6. Was ist das Spielprinzip in Sokoban? Wie würden Sie nicht mehr aufzulösende Teilstellungen (wie vier im 2×2 Feld stehende Kisten) effizient speichern und suchen? (5 P)
7. Finden sie die *Prinzipielle Komponentenanalyse* als Vorverarbeitungsschritt in dem Weka-Werkzeug und wenden sie diese auf eine selbst generierte Datenmenge oder eine der in ml/code zur Verfügung gestellten Datenquellen an. (5 P)

2 A Priori-Algorithmus

Betrachten sie folgende Tabelle von Transaktionen in einem Supermarkt.

Transaktionsnummer	Einkaufsliste
t_0	Brot, Bier
t_1	Schokolade, Cola, Chips
t_2	Wasser
t_3	Saft, Cola, Bier, Wein
t_4	Cola, Bier
t_5	Saft, Wasser
t_6	Saft, Cola, Wein
t_7	Saft, Cola, Bier
t_8	Schokolade, Schinken, Brot

Minimaler Support sei 2 und der minimale Konfidenzwert liege bei 75%

10 P Wenden sie den Apriori-Algorithmus zur Entdeckung von Assoziationsregeln an.

10 P Dokumentieren Sie die einzelnen Schritte des Verfahren.

5 P Diskutieren Sie Verbesserungen des grundlegenden Algorithmus.

3 General Video Game - MCTS

Zum Schluss der Vorlesung wollen wir maschinellen Lernverfahren in der Praxis testen und gegeneinander antreten lassen. Sie sollen in dieser Aufgabe einen Allgemeinen Video-Spieler schreiben, der mehrere vorher unbekannte Video-Spiele spielen kann, die in einer allgemeinen textuellen Eingabe spezifiziert werden. GVGAI steht dabei für *The General Video Game AI*. Mehr Infos finden Sie auf www.gvgai.net.

Das Beispiel-Set von 10-20 Videospielen beinhaltet Arcade-Klassiker wie *Frogger*. Alle 40ms muss von der KI ein Zug gefunden werden, der bei dem Server zu eine Zustandsänderung führt.

Sie finden unter auf der Vorlesungsseite unter `ml/code` ein funktionierendes GVGAI-Framework (mitunter in einer nicht aktuellen Version). Beispiel-Controller finden sie dann unter `src/controller`. Als Startpunkt Ihrer Implementierung bietet sich von den angebotenen Controllern MCTS-Controller `sampleMCTS` an.

50 P Schreiben Sie einen leicht veränderten (verbesserten) MCTS-Controller für GVGAI und testen Sie diesen. Erklären Sie kurz Ihren Ansatz.