

10. Übungsblatt - Abgabe: 14.01.2014

Aufgabe 10.1 - WSD mit Bayes-Klassifizier

Mit Hilfe von Kontextwörtern soll ein Klassifikator zur Word-Sense-Disambiguierung für die beiden Lesarten von *Schloss* aus einem Trainingskorpus mit 100 Dokumenten je Lesart gelernt werden.

Dabei ermittelt man die folgenden Kontextwortfrequenzen:

	Schloss ₁	Schloss ₂
Tür	5	23
Graf	22	3
Fahrrad	11	15
Neuschwanstein	17	1
Schlüssel	2	33
Ausflug	35	5

Außerdem gelten a-priori-Wahrscheinlichkeiten von $P(\text{Schloss}_1) = 0,4$ und $P(\text{Schloss}_2) = 0,6$.

Beobachtet wird nun ein Auftreten von Schloss mit Merkmalsmuster v_i :

	v_i
Tür	0
Graf	1
Fahrrad	1
Neuschwanstein	0
Schlüssel	0
Ausflug	1

Bestimmen Sie die wahrscheinliche Lesart!

Feedback

Bitte beantworten Sie kurz folgende Fragen:

- Wie schwer fanden Sie den Stoff der aktuellen Vorlesung? (sehr einfach, einfach, durchschnittlich, schwer, sehr schwer)
- Wie viel Zeit haben Sie mit der Bearbeitung des aktuellen Übungsblattes verbracht?
- Haben Sie sonstige Fragen oder Anmerkungen zur aktuellen Vorlesung? Wenn ja, welche?

Abgabe in Gruppen von bis zu drei Studierenden bis **14.1.2014** 10 Uhr entweder als Email- Anhang im pdf-Format an **i2cl@coli.uni-sb.de** mit dem Betreff **i2cl homework 10** und Ihrer Matrikelnummer oder auf Papier im Briefkasten an der Tür von Raum 1.04 in C7.2.