

Übungsblatt 6 (Abgabe 19.12.2006)

1. Beispiel-Evaluation

Die Folien zur Vorlesung zeigen eine Beispielevaluation für eine binäre Klassifikationsaufgabe (Folie 35ff.), allerdings nur für die Klasse Adj.

- Berechnen Sie Precision, Recall und F-Score für die Klasse NAdj.
- Vergleichen Sie die Ergebnisse für NAdj mit den Ergebnissen für Adj. Was fällt Ihnen auf?
- Welche der beiden Evaluationen ist für die tatsächliche Brauchbarkeit des Modells in der Praxis aussagekräftiger? Begründen Sie kurz.

2. Statistische Modellierung (4P)

Herr P. trainiert das statistische Modell zur Adjektivklassifikation aus den Vorlesungsfolien auf einigen hundert Sätzen aus dem TIGER-Korpus und erhält folgende Frequenztabelle:

Selbst Artikel?	Nächstes Wort kapitalisiert?		Adj	NAdj
Falsch	Falsch		66	10533
Falsch	Wahr		787	4966
Wahr	Falsch		0	565
Wahr	Wahr		0	1162

- Wie viele Klassen gibt es, wie heißen sie?
- Wie viele Features gibt es? Wie viele Werte haben die Features jeweils?
- Welches sind die möglichen Ereignisse? Wie groß ist also der Ereignisraum?
- Für welchen Anteil des gesamten Ereignisraumes hat das Modell Trainingsinstanzen gesehen? Was bedeutet das für die Abdeckung des Modells auf neuen Daten?
- Stellen Sie die „Bedeutung“ des Modells dar, indem sie für jedes Ereignis eine Regel formulieren, die ihm eine Klasse zuordnet. Was halten sie von der linguistischen Plausibilität dieser Regeln?
- Welches Problem scheint das naive statistische Modell zu haben, wenn die Häufigkeit der Klassen sehr unterschiedlich ist? Schildern Sie kurz die Ursache.