

Schriftliche Prüfung zur Vorlesung
Statistische Methoden in der maschinellen Sprachverarbeitung
WS 2018/19
Dozent: Helmut Schmid

Aufgabe 1) Wie sind die Erwartungswerte $E(X)$ und $E(\log \frac{1}{p(X)})$ der Zufallsvariablen X definiert? (2 Punkte)

Wie wird der zweite Erwartungswert üblicherweise genannt? (1 Punkt)

Aufgabe 2) Wie definiert ein Naive-Bayes-Modell die gemeinsame Wahrscheinlichkeit einer Klasse (bzw. Wortlesart) c und eines Textes (bzw. Wortkontextes) w_1, \dots, w_n ? (2 Punkte)

Wie berechnen Sie konkret die Wahrscheinlichkeit der Klasse *Spam* für den Text: *“Diese Mail ist kein Spam !”* (2 Punkte)

Aufgabe 3) Wie lautet die Formel für die Berechnung geglätteter bedingter Wahrscheinlichkeiten nach dem normalen interpolierten Backoff-Verfahren? (2 Punkte)

Gegeben sind die Buchstaben-N-Gramm-Häufigkeiten $f(a,a)=1$, $f(a,b)=2$, $f(b,a)=0$ und $f(b,b)=1$.

Berechnen Sie den Discount für die Glättung der Bigramm-Wahrscheinlichkeiten. (1 Punkt)

Berechnen Sie dann die Unigramm-Häufigkeiten $f(a)$ und $f(b)$ einmal nach dem Standard-Backoff-Verfahren und einmal nach dem Kneser-Ney-Verfahren. (1 Punkt)

Berechnen Sie aus den Kneser-Ney-Häufigkeiten die ungeglätteten Unigramm-Wahrscheinlichkeiten $p(a)$ und $p(b)$. (1 Punkt)

Aufgabe 4) Nach welchen Formeln berechnen Sie die Forward- und Backward-Wahrscheinlichkeiten mit einem Hidden-Markow-Modell? (4 Punkte)

Wie berechnen Sie aus den Forward- und Backward-Wahrscheinlichkeiten die Wahrscheinlichkeit des Tags t an Position i bzw. die Wahrscheinlichkeit der Tags t und t' an den Positionen $i-1$ und i ? (In der Vorlesung wurde diese Werte mit γ bezeichnet.) (3 Punkte)

Wie ändern sich die Formeln für die Berechnung der Forward- und Backward-Wahrscheinlichkeiten, wenn Sie mit einem Conditional-Random-Field-Tagger arbeiten? (2 Punkte)

Aufgabe 5) Geben Sie Pseudo-Code an für das Gradientenanstiegs-Training (2 Punkte)

Welche Zielfunktion wird beim Training eines CRF-Taggers maximiert? (1 Punkt)

Wie ist der Gradient dieser Zielfunktion definiert? (1 Punkt)

Mit welchem Algorithmus wird der Gradient berechnet? (1 Punkt)

Aufgabe 6) Wie ist beim Conditional Random Field die Wahrscheinlichkeit einer Tagfolge definiert? (4 Punkte)

(30 Punkte insgesamt)

Viel Erfolg!