

Vorlesung Information Retrieval

Übungsblatt 7 (Aufgaben 16-17)

Ausgabe: 29.06.2009

Besprechung: 06.07.2009

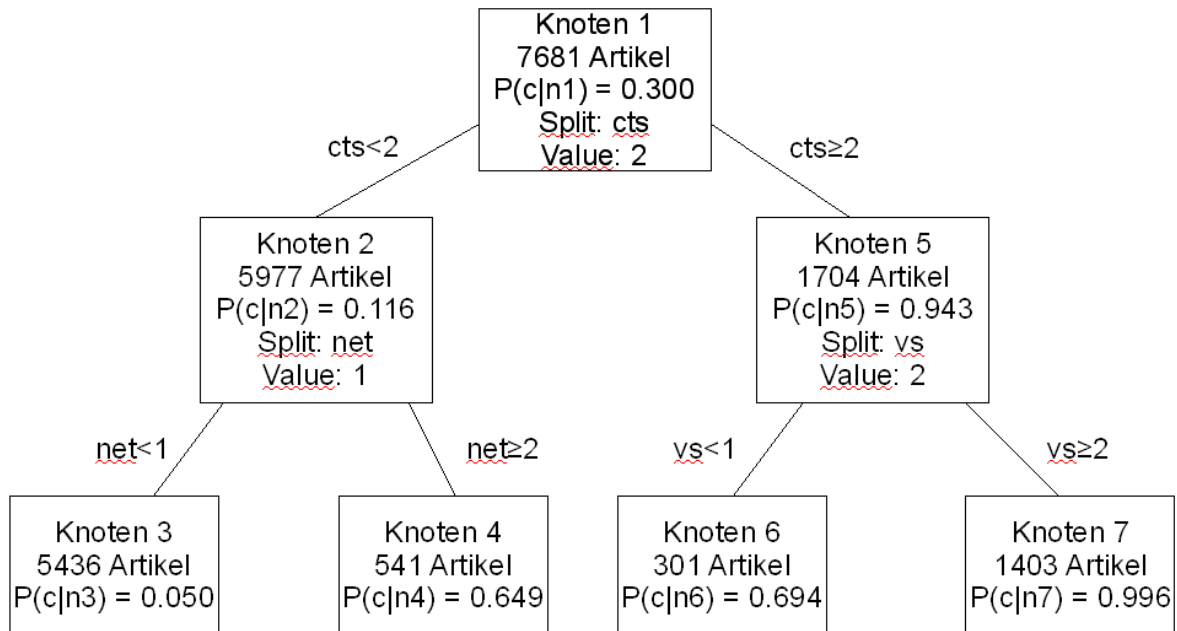


Sommersemester 2009

Aufgabe 18:

Beschreibe die Idee der kNN-Klassifikation. Welche Rolle spielt das Ähnlichkeitsmaß?

Aufgabe 19:



$P(c|n_i)$ gibt die Wahrscheinlichkeit an, mit der ein Dokument an Knoten i zu der Kategorie c gehört.

Bestimme die Entropie an Knoten 1, 2 und 5 und den Information Gain für den Split von Knoten 1 in Knoten 2 und 5.