

1. PA Wolke aufmalen, kurze Erklärung zu den einzelnen Themen

2. Density estimation problem:

- Parzen window estimator (erklären, Formeln), parametric – nonparametric approach
- Wie wäre das Problem mit Random Forests realisiert (Erklärung, wie es funktioniert)
- Transferfrage: Was ist der Unterschied zwischen den beiden Ansätzen (Vor-, Nachteile)
Meine Antwort: Parzen Window estimator einfach zu berechnen, während der random forest trainiert werden muss (ggf. viele Bäume). Vorteil bei Wäldern, dass sie die Datendichte besser repräsentieren (Flexibilität), während der Parzen kernel nur eine feste Breite h hat.

3. Nach der sanften thematischen Überleitung sind wir nun bei den Random Forests angekommen

- Wie wird der Baum trainiert (Information gain maximiert, Entropie Formel $H(S)$ für den Spezialfall von density forests aufgeschrieben).

4. HMMs:

- Welche Annahme ist grundlegend für die graphischen Modelle: Markov Annahme, für HMM und MRF erklärt
- Hier wurde ich nach "Inferenzproblem" im HMM gefragt, was mich etwas verwirrt hat. Gemeint war das "Wortproblem" $P(o_1, \dots, o_T | \lambda)$ und das Problem der Rekonstruktion der Zustandskette $\max_S P(O, S | \lambda)$
- Trainingsproblem: Baum-Welch-Formeln, das war nicht mehr schwer
- Baum-Welch Formeln skizziert, Viterbi-Algorithmus vollständig aufgeschrieben
- Zu welcher Klasse von Algorithmen gehört Baum-Welch: Gewollt war Expectation-Maximization-Algorithmus (EM), meine beiden Antworten waren greedy und approximativ, was nicht falsch ist aber nicht unmittelbar zum Fach gehörend
- Wenn er greedy ist, dann landet man häufig auf lokalen Maxima, wie kann man dem Problem abhelfen? Leichte Transferfrage: Mehrmals mit unterschiedlichen Initialisierungen laufen lassen und die Performance jeweils auf den Testdaten prüfen
- Wie waren die b_j (zum Zustand S_j gehörige Wahrscheinlichkeitsverteilung) in der Übung: kontinuierlich (Wahrscheinlichkeitsdichte) statt diskret

Die Atmosphäre war sehr entspannt. Vorbereitet hatte ich mich mit den Vorlesungsvideos aus dem Vorjahr, da ich die Vorlesung aus zeitlichen Gründen nicht besuchen konnte. Außerdem habe ich die angegebene Literatur benutzt, dies war sehr hilfreich zum Verständnis der Formeln.