

Information Mining - Wintersemester 2014

Dipl.-Inform. Vu Tran, LF 139

vtran@is.inf.uni-due.de

Übungsblatt 9

Aufgabe 20: Praktische Übung: Klassifikation

Auf *uClassify*¹ kann man Textklassifizierer unterschiedlicher Art direkt auf der Webseite ausprobieren. Darunter sind auch die folgenden zu finden:

- *Text Language*²,
- *Topics*³
- *Sentiment*⁴

Schaue dir die Klassifizierer mit den oben genannten Bezeichnungen genauer an.

- Führe jeweils Klassifikationen mit geeigneten Beispielen durch.
- Wonach werden die Texte jeweils klassifiziert?
- Wie könnten diese Klassifizierer jeweils funktionieren?

Aufgabe 21: Support-Vector-Machines

Support-Vector-Machines (SVMs) haben sich als erfolgreicher Ansatz zur Klassifikation etabliert.

- Skizziere geometrisch die Grundidee hinter SVMs.
- Schreibe die Klassifikationsfunktion von SVMs auf und erläutere diese.
- Bei SVMs ist eine Überanpassung unwahrscheinlich. Warum?
- Mit SVMs lassen sich hervorragend 2-Klassen-Probleme lösen. In unseren bisherigen Daten hatten wir es aber oft mit mehreren Klassen zu tun. Wie kann man Mehrklassenprobleme mit SVMs lösen?

¹<http://www.uclassify.com/>

²<http://www.uclassify.com/browse/uClassify/Text-Language>

³<http://www.uclassify.com/browse/uClassify/Topics>

⁴<http://www.uclassify.com/browse/uClassify/Sentiment>