

Posudek oponenta diplomové práce

Autor práce: **Michal Patočka**

Název práce: **Metody strojového učení pro analýzu sentimentu**

Práce se zabývá metodami strojového učení a jejich aplikováním na úlohu analýzy sentimentu v textu. Dále popisuje možnosti předzpracování dokumentů a metody selekce příznaků.

Práce je velmi kvalitní. Rozsah experimentů je nadprůměrný a diplomová práce se blíží vědecké publikaci. Množství nastudované literatury ukazuje dobrou orientaci v úloze analýzy sentimentu.

Formálně je práce v pořádku až na některé detaily a divná vyjádření:

- str. 1: „Nás ovšem zajímá vyjádření **objektivní**“
- str. 6: „klasifikace dokumentů do **tříd kategorií**“
- str. 26: recall bych označil raději R než TP
- str. 34: „hloupější předzpracování“
- přetékání konců odstavců na další stranu

Zadání bylo splněno bez výhrady.

Dotazy k práci:

- Klasifikátor MaxEnt byl použit téměř ve všech experimentech (selekce příznaků, předzpracování). Dokážete odhadnout, zda by závěry byly podobné při použití ostatních diskutovaných klasifikátorů?
- Podle popisu na str. 45 je přísnost procento vyfiltrovaných příznaků. Potom mi není jasný graf na straně 47 a překvapivé chování chi square na straně 46. Prosím, objasněte to.
- Naive Bayes – překvapivě velmi negativní výsledek. Dokážete vysvětlit proč?


Navrhuji hodnocení známkou **výborně** a práci doporučuji k obhajobě.

V Plzni 31.5.2013

Ing. Josef Steinberger, Ph.D.



**SOUHLASÍ
S ORIGINÁLEM**



Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky
①