

Hodnocení vedoucího bakalářské práce

Autor/autorka práce: **Filip Kupilík**

Název práce: **Automatická identifikace revizí textových dokumentů**

Aktivita studenta

Student zejména z počátku potřeboval významnou pomoc od vedoucího práce. Ke konci práci přišel s několika vlastními nápady a ty dostatečně konzultoval.

Spolupráce s vedoucím

Student se řídil pokyny vedoucího, na schůzky se dostavoval včas a pravidelně.

Původnost práce a práce související

Student samostatně provedl zadané experimenty a učinil odpovídající závěry. Ve své práci spoléhá zejména na nástroje *Apache Lucene* a *Apache Tika*, vlastní experimenty jsou ale původním dílem. Student dále navrhl a implementoval nástroj pro řízení experimentů, což hodnotím kladně.

Kvalita řešení

Student úspěšně provedl i poměrně náročné experimenty založené na *Kullback-Leiblerově divergenci* a *rozpoznávání pojmenovaných entit*. Pro měření dosažených výsledků využívá křížovou validaci. Po spuštění program automaticky provede všechny experimenty. Při testování odevzdané práce vše fungovalo naprosto korektně.

Programová realizace není příliš dobrá. Zejména pokulhává dekompozice problému. Kdyby student lépe využil mechanismus dědičnosti, mohl se vyvarovat velkého množství duplicitního kódu¹. Experimentální systém mohl být navržen jednodušeji a univerzálněji, kdyby student využil mechanismu *reflexe* v Javě. Student míchá anglické a české názvy symbolů.

V textu práce se objevují některé nepřesné nebo zavádějící formulace, krkolomné formulace (“Ve strojovém učení se KL-divergence označuje míra, kolik nové informace obsahuje. . .”), nevhodná slova (například “ováhaného”) a nevhodně použité anglické termíny, u kterých existuje český ekvivalent (například *cross validace*), nedefinované konstanty a operátory ve vzorcích (například vzorec 5.4) a další méně či více závažné prohřešky.

¹Automatická inspekce zdrojových kódů odhalila 30 opakujících se bloků.


Využitelnost dosažených výsledků

Dosažené výsledky jsou zajímavé a dle mého názoru mají publikační potenciál. Bude však třeba experimenty ještě rozšířit a dopracovat.

Splnění zadání

Všechny body zadání byly bezesbytku splněny.

Navrhuji hodnocení známkou “**velmi dobře**” a práci doporučuji k obhajobě.



Ing. Miloslav Konopík, Ph.D.
(vedoucí BP)

V Plzni 22. srpna 2017