

Hodnocení vedoucího diplomové práce

Autor/autorka práce: **Zuzana Mikolášová**

Název práce: **Analýza dat v sociálních sítích**

Práce je věnována analýze obsahu příspěvků v síti Twitter a jejím cílem je umožnit výběr takových příspěvků, které pro zvolené téma obsahují nejhodnotnější informace. Diplomantka je příznivkyní této sociální sítě, proto o téma projevovala velký zájem. Ochotně dostudovala problematiku, která překračovala náplň standardní výuky. Jelikož téma diplomantku zaujalo, byla radost s ní spolupracovat. Konzultací se zúčastňovala pravidelně.

Námět pro diplomovou práci jsem našel v jedné publikaci ACM z roku 2015 (je citována v referencích pod č.[1]). Ta nám zpřístupnila anotovaná data, jejichž vytvoření by bylo zcela mimo naše možnosti. I když základní myšlenka není původní, její rozvinutí a interpretace je vlastním přínosem diplomantky. Inspiraci z citované publikace nelze chápat jako nedostatek. Právě naopak, schopnost rozvinout dílo z prestižní publikace, svědčí o dobré připravenosti diplomantky řešit komplexní úlohy.

Předložené řešení je funkční a forma průvodní zprávy odpovídá požadovanému standardu. Prvé kapitoly popisují vlastnosti Twitteru a použitých dat. Kap. 4 pojednává o výběru a vlastnostech klasifikačních metod. Významná je kap. 5, která popisuje možnost hodnocení významnosti příspěvků na grafovém principu. Pro názornost uvádí i zjednodušený příklad (str. 22). Část konstrukce grafu je nejasná a nezdá se být správně. Následující výpočty se stochastickými maticemi jsou ale v pořádku. Kap. 6 je technickým popisem řešení. Závěrečné kapitoly jsou věnovány vyhodnocení a experimentování. Experimenty byly z časových důvodů provedeny ve zjednodušené podobě. Výsledky jsou náznakem možného dalšího bádání – zjištění důležitosti zpracovávaných charakteristik tweetů na jejich výsledné pořadí. Pokud se týká jazykové úrovně, tak počty překlepů nepřekračují obvyklou míru. Někdy autorka používá nesprávné pádové vazby.


Vytvořený systém, po jeho dopracování do podoby nad rámec školního díla, ocení uživatelé Twitteru při výběru příspěvků s nejvyšší informativní hodnotou. Vzhledem k významnosti vyhledávání a filtrace informací, systémy tohoto typu jsou velmi potřebné.

Zadání bylo splněno.

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

V Plzni 30.5.2016

Prof. Ing. Karel Ježek, CSc.



SOUHLASÍ
S ORIGINÁLEM



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