

## Posudek vedoucího diplomové práce

Jméno a příjmení autora posudku: Iveta Mrázová

Jméno a příjmení autora práce: Peter Zvirinský

Název práce: Social network and data mining

---

Vlastní text (sem prosím napište text posudku, délka textu posudku není omezena):

Práce Bc. Petra Zvirinského "Social networks and data mining" se zabývá vysoce aktuální problematikou efektivní analýzy rozsáhlých veřejně dostupných dat (big data). Cílem předkládané diplomové práce přitom bylo zrekapitulovat různá paradigmatu použitelná pro dobývání znalostí (mj. techniky pro předzpracování dat, extrakci příznaků, klastrování, klasifikaci a extrakci pravidel). Dále měl diplomant popsat a navzájem porovnat známé techniky používané při analýze sociálních sítí pro detekci význačných uzlů a identifikaci jejich role v dané sociální síti (např. PageRank, HITS, klastrovací metody). Věnovat se měl i možnostem interpretace a vizualizace extrahovaných znalostí.

Těžiště práce tedy spočívá v aplikaci studovaných přístupů při analýze obrovského množství sociologických dat volně dostupných z Insolvenčního rejstříku ČR. Údaje o insolvenčních řízeních autor zkoumal z nejrůznějších pohledů (např. z hlediska geografického, demografického a mnoha dalších). Především z pohledu sociálně právního se pak snažil zmapovat jak vzájemné vztahy věřitelů a dlužníků, tak také průběh vlastního procesu insolvenčního řízení. Ke zpracování analyzovaných dat uchazeč využil veřejně dostupné nástroje pro dobývání znalostí (např. systém Weka, knihovnu SPMF, software Gephi aj.), ale i vlastní kód přiložený k práci na CD.

Nemalé úsilí Bc. Zvirinský věnoval otázkám vizualizace a adekvátní interpretace dosažených výsledků. V tomto směru autor prokázal mimořádné schopnosti pro interdisciplinární výzkum – řešení totiž předpokládalo intenzivní diskusi dílčích výsledků s kolegy z právnické fakulty a z fakulty sociálních věd. Vzhledem k obrovskému množství proběhlých, resp. probíhajících insolvenčních řízení, které v současné době přesahuje 100 000 případů, se tak Bc. Petrovi Zvirinskému podařilo dosáhnout významných výsledků, a to jak v odborném, tak také v sociálně společenském kontextu.

Práce sama je poměrně rozsáhlá a má hezkou grafickou úpravu. Získané výsledky jsou dokumentované celou řadou ilustrativních příkladů, grafů a sumarizačních tabulek. Seznam použité literatury čítá úctyhodných 54 položek a potvrzuje orientaci autora v řešené problematice, bylo by však vhodnější seřadit jednotlivé položky lexikograficky. Text práce je psaný anglicky a jen zřídka se v něm vyskytují gramatické chyby a překlepy (např. 'by' namísto 'be' a 'now' namísto 'know' na str. 31 nahoře, záměna 'its' a 'it's' na str. 7, zdvojená předložka 'of' na str. 16 a další). I přes výše uvedené drobné nedostatky však práce splňuje svůj původní cíl. Diplomant pronikl dostatečně hluboko do problematiky dobývání znalostí a analýzy sociálních sítí a prokázal schopnost samostatně analyzovat rozsáhlá data z Insolvenčního rejstříku.

Doporučení k obhajobě:

Z výše uvedených důvodů práci *doporučuji* k obhajobě.

Vynikající práce vhodná pro soutěž studentských prací	ANO <input type="checkbox"/>
---	------------------------------

Seznam soutěží studentských prací, viz <http://www.mff.cuni.cz/studium/bcmgr/prvzoryace/>

Pokud jste výše zaškrtnli ANO, zdůvodněte prosím svůj návrh, případně uveďte konkrétní soutěž, pro kterou je práce vhodná (rámeček lze nechat prázdný, pokud za dostatečné zdůvodnění považujete text posudku):

**Výsledky, kterých uchazeč dosáhl, byly oceněny Novinářskou cenou 2013 – Cenou Googlu pro mladé talenty datové žurnalistiky.**

V Praze dne: 19. května 2014

Podpis:\*\*

\* *nehodící se škrtněte (vymažte)*

\*\* *do SISu vkládejte formulář nepodepsaný (ve formátu PDF), podpis je potřeba doplnit až na vytištěný posudek.*