

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	User Behavior Clustering and Behavior Modeling Based on Clickstream Data
Jméno autora:	Jan Žaloudek
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Vedoucí práce:	Ing. Jan Drchal, Ph.D.
Pracoviště vedoucího práce:	Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
K úspěšnému splnění zadání je třeba dobrá orientace v problematice zpracování přirozeného jazyka. Je třeba počítat s vysokou výpočetní a paměťovou náročností.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Zadání je splněno bez výhrad.	

Aktivita a samostatnost při zpracování práce	C - dobře
<i>Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatně tvůrčí práce.</i>	
Se studentem jsme se sešli na pouhých několika schůzkách, tudíž jsem neměl na výslednou formu práce zásadní vliv. Text byl dokončován na poslední chvíli, přesto musím konstatovat, že výsledek je nadprůměrný.	

Odborná úroveň	B - velmi dobře
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Práce je na vysoké odborné úrovni. Je zřejmé, že se student v problematice NLP dobře orientuje a byl tak schopen použít příslušné metody v rozdílném kontextu modelování uživatelů webových aplikací. Dílčí navrhovaná řešení jsou dostatečně vysvětlena s výjimkou sekce 8.2, kde nejsou zřejmé detaily toho, jak byly získány labely pro metody semi-supervised learning. V sekci 7.2 není zřejmé, jaká metoda shlukování byla použita. V sekci 6.5.3 postrádám podrobnější diskuzi, proč vykazuje „baseline“ metoda lepší výsledky.	

Formální a jazyková úroveň, rozsah práce	B - velmi dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psána kvalitní angličtinou s minimem překlepů. Úvodní kapitoly práce jsou pěkným úvodem do NLP. Student používá v dostatečném rozsahu matematický formalismus, v mnoha vzorcích jsou bohužel vysázeny chybně dolní indexy. Text je doplněn řadou přehledných tabulek a obrázků.	

Výběr zdrojů, korektnost citací	A - výborně
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i>	

K výběru zdrojů a jejich citacím nemám výhrady.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Za mimořádně přínosnou považuji analýzu embeddingu založeného na neuronové síti, kdy jej student diskutuje z hlediska sémantiky. Elegantní je i způsob, jakým vyřešil předzpracování časové složky dat.

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

Předloženou práci hodnotím klasifikačním stupněm **B - velmi dobře**.

Datum: 6.6.2018

Podpis: