



## Posudek oponenta závěrečné práce

**Student:** Bc. Michal Bajer  
**Oponent práce:** Ing. Jaroslav Kuchař, Ph.D.  
**Název práce:** Neural Autoencoders in Recommender Systems  
**Obor:** Znalostní inženýrství

**Datum vytvoření:** 3. 6. 2019

<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení – následující škálou 1 až 4:</b>
<b>1. Splnění zadání</b>	<b>1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</b>
<b>Popis kritéria:</b> Posuďte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.	
<b>Komentář:</b> Zadání je z oblasti aktuálně řešených problémů. V práci lze najít část ke každému bodu zadání. Některé části ale nejsou úplně dotazeny do konce. Práce jako taková obsahuje poměrně rozsáhlou teoretickou část popisující obecně teorii doporučovacích systémů. Očekával bych ale více řešerši problémů spojených s autoencoderem v oblasti doporučování jak bylo zmíněno v zadání. Jakým způsobem je využívají jiné již existující přístupy apod. Dle zadání se také měl využít autoencoder pro minimálně dva problémy spojené s doporučováním. Práce se ale zabývá aplikací na dvou datových sadách.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>2. Písemná část práce</b>	<b>65 (D)</b>
<b>Popis kritéria:</b> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytl-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišené od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.	
<b>Komentář:</b> S přihlédnutím ke zmíněným nedostatkům ze splnění zadání, je práce na hranici doporučeného rozsahu. Práce tedy obsahuje např. teoretickou část zaměřenou dost obecně. Po jazykové stránce práce obsahuje překlepy a zvláštní formulace. Text je ale čitelný a pochopitelný. Po věcné stránce jsem nenarazil na zásadní nedostatky. Jak již bylo řečeno, práce by zasloužila rozsáhlejší řešerši z oblasti autoencoder v doporučovacích systémech. Tím by se zlepšila i situace s použitými zdroji, aktuálně je seznam poměrně strohý. U popisu experimentů není úplně objasněno jakým způsobem došlo k rozdělení a proč není využita časová informace při rozdělení na trénovací, testovací a validační množinu. Bylo by také vhodné lépe popsat některé magické konstanty jako proč zrovna 4 a výše znamená 1 pro předzpracování, proč se používá konkrétní velikost dat apod.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>3. Nepísemná část, přílohy</b>	<b>70 (C)</b>
<b>Popis kritéria:</b> Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů	
<b>Komentář:</b> Součástí zadání i samotné práce je množství experimentů, které řeší např. ideální nastavení parametrů jednotlivých přístupů. Jednotlivé experimenty jsou popsány a přílohou je i kompletní implementace všech experimentů. Zdrojové kódy by ale zasloužily podrobněji dokumentovat jednotlivé kroky.	

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>4. Hodnocení výsledků, jejich využitelnost</b>	<b>75 (C)</b>
<i>Popis kritéria:</i> Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.	
<i>Komentář:</i> Výstupem práce je sada experimentů, která může být zajímavým vstupem pro další návazné práce. Experimenty jsou ale omezené na konkrétní řešený problém a určitě by bylo vhodné práci rozšířit o rešerši v oblasti doporučovacích systémů a na práci navázat řešením dalších problémů.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – nehodnotí se</i>
<b>5. Otázky k obhajobě</b>	
<i>Popis kritéria:</i> Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odřázkami).	
<i>Otázky:</i> - Jaký vliv by mohla mít časová informace při rozdělení na trénovací a testovací množinu? Nemá současné rozdělení problém s tím, že se bere v úvahu i položka, která vznikne až v budoucnosti? - Víte o nějakém dalším využití autoencoderu nebo variational autoencoderu v oblasti doporučování? - V rámci závěru je navrženo využití v A/B testech. Máte nějaký přehled o časové náročnosti jedné predikce?	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>6. Celkové hodnocení</b>	<b>65 (D)</b>
<i>Popis kritéria:</i> Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.	
<i>Text hodnocení:</i> Student se seznámil s teorií doporučovacích systémů a s využitím autoencoderu pro doporučení. Práce obsahuje sadu experimentů zaměřených na nejlepší nastavení pro konkrétní sadu ze dvou datových sad. Práce má nedostatky v textové části a celkově řeší především základní problematiku. Zasloužila by rozšíření. Doporučuji ale k obhajobě.	

Podpis oponenta práce: