

**I. IDENTIFIKAČNÍ ÚDAJE**

<b>Název práce:</b>	<b>Transfer Learning for Textual Topic Classificaton</b>
<b>Jméno autora:</b>	<b>Janata Pavel</b>
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra kybernetiky
<b>Oponent práce:</b>	Ing. Martin Selecký, Ph.D.
<b>Pracoviště oponenta práce:</b>	Ataccama Software, s.r.o., Sokolovská 685/136f, Invalidovna, 186 00 Praha 8

**II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ**

<b>Zadání</b>	<b>lehčí</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	

<b>Splnění zadání</b>	<b>splněno s většími výhradami</b>
<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>	
<ul style="list-style-type: none"> <li>• První bod zadání „Study the state-of-the-art approaches to transfer learning in the field of NLP” je v práci řešen nedostatečně, pouze na několika málo řádcích sekce 2.3 a nereferencuje nejnovější metody.</li> <li>• Druhý bod zadání “Analyze existing datasets containing textual data...” je v práci řešen jen okrajově a autor se zabývá spíše popisem vybraného datasetu.</li> </ul>	

<b>Zvolený postup řešení</b>	<b>správný</b>
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
-	

<b>Odborná úroveň</b>	<b>A - výborně</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
-	

<b>Formální a jazyková úroveň, rozsah práce</b>	<b>A - výborně</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
-	

<b>Výběr zdrojů, korektnost citací</b>	<b>C - dobře</b>
<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>	
Viz část Splnění zadání	

<b>Další komentáře a hodnocení</b>
<i>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.</i>
Autor práce pracoval s implementací Fast AI knihovny, existujícím ULMFiT modelem a předzpracovanými datasey. Z předložené práce není jasné, co student v rámci jejího řešení implementoval sám.

### III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

*Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.*

Student se v práci zabývá aktuálním tématem Transfer Learningu v oblasti NLP a jeho využitím pro rozpoznávání tématu psaných textů. Konkrétně pak student pozoroval vliv chybějících dat na kvalitu klasifikace ULMFiT modelu. Práce je psaná anglicky, je dobře strukturovaná a postup řešení i dosažené výsledky jsou srozumitelně prezentované.

K práci mám několik připomínek:

- 1) První Bod zadání „Study the state-of-the-art approaches to transfer learning in the field of NLP“ je v práci řešen jen okrajově na několika málo řádcích sekce 2.3 a v této sekci nereferencuje nejnovější metody řešící danou problematiku.
- 2) Druhý bod zadání “Analyze existing datasets containing textual data...” je v práci řešen jen okrajově a spíše autor se zabývá spíše popisem vybraného datasetu.
- 3) Věřím, že aplikace existujících knihoven a předučených modelů nebyla přímočará, ale z textu práce není jasné, do jaké míry student musel zasahovat do existujících zdrojových kódů.

Na druhou stranu oceňuji informativní hodnotu navržených a realizovaných experimentů a kvalitu a přehlednost prezentace jejich výsledků.

Studentovi navrhuji položit následující otázky:

- 1) Jaké změny bylo třeba provést se zdrojovým kódem Fast AI knihovny a samotnými daty, aby bylo možné realizovat a vyhodnotit provedené experimenty?
- 2) Byl při fázi „Target Task LM Fine-Tuning“ rozšířen slovník modelu o slova vyskytující se v cílové doméně, nebo model využíval stále stejný slovník vytvořený při prvotním předtrénování?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

Datum: 3.6.2019

Podpis: