
Posudek oponenta bakalářské práce

Jakub Šmíd
Detekce polarity textu s využitím mezijazyčné transformace

Cílem bakalářské práce Jakuba Šmída bylo provést vícejazyčné experimenty na úloze detekce polarity textu s využitím klasických klasifikátorů a základních modelů neuronových sítí. Práce se dále věnuje možnostem mezijazyčného přenosu modelu pomocí lineárních transformací.

Práce je psaná velmi srozumitelnou formou a velice dobře se čte. Práce má dobrou logickou strukturu. Z textu je patrné, že student problematice rozumí velice dobře. Mojí jedinou výtku k obsahu práce je, že text je poměrně nevyvážený. Autor se na začátku věnuje základním modelům strojového učení (logistická regrese, dopředná neuronová síť), které popisuje velice dopodrobna a text zde hloubkou odpovídá spíše diplomové práci. Naproti tomu autor pokročilejší metody dále v textu popisuje jen velice stručně (konvoluční a rekurentní sítě). Také autor v některých případech sáhodlouze popisuje matematický vztah, který je jasný ze vzorce. Práce také obsahuje několik drobných terminologických nepřesností. Kromě výše zmíněných drobných výhrad mi v práci nic nechybí, ani nepřebývá.

V práci se vyskytuje naprosto minimální množství chyb a překlepů. Její formální úroveň je vysoce nadstandardní. Grafickému zpracování dokumentu nelze vůbec nic vytknout.

Student ve své práci cituje velké množství relevantních zdrojů, jak tištěných, tak webových. Citované zdroje jsou aktuální a popisují současné state-of-the-art metody.

V praktické části student provedl dlouhou řadu experimentů. Jejich metodice nelze nic vytknout. Student navíc dosáhl vynikajících výsledků, kde na českém jazyce výrazně překonal model, který byl donedávna state of the art. Se současným SotA dosahuje srovnatelných výsledků, avšak s využitím výrazně jednoduššího modelu. Student také dosáhl velice dobrých výsledků při přenosu modelu, natrénovaného na Anglických datech, do Češtiny. Toto je hlavní přínos práce a rozhodně není zanedbatelný. Drobnou výtku mám k nedostatečné přehlednosti konfigurací provedených experimentů, kvůli které není jednoduché experimenty replikovat.

Práce splňuje zadání ve všech bodech. Přes výše zmíněné výtky práci považuji za vynikající. Proto práci doporučuji k obhajobě a hodnotím klasifikačním stupněm

„výborně“.

Doplující otázky:

1. V monolingálních experimentech na češtině dosahujete překvapivě vysokých výsledků (výrazně lepších, než bylo do nedávna SotA), čemu to přisuzujete?

Ing. Ondřej Pražák
(oponent BP)

V Plzni 31. května 2021