

HODNOCENÍ ZÁVĚREČNÉ KVALIFIKAČNÍ PRÁCE POSUDEK VEDOUCÍHO

Autor závěrečné práce: Bc. Ondřej Vacek

Vedoucí práce: prof. Ing. Jan Nouza, CSc.

Název práce: Segmentace řečových trénovacích dat pro učení systému ASR

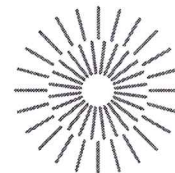
- A. Úplnost abstraktu, klíčová slova odpovídají náplni práce Výborně (1)
- B. Kvalita zpracování rešerše Výborně (1)
- C. Řešení práce po teoretické stránce Výborně minus (1-)
- D. Vhodnost, přiměřenost použité metodiky Výborně (1)
- E. Úroveň zpracování výsledků a diskuse Výborně (1)
- F. Vlastní přínos k řešené problematice Výborně (1)
- G. Formulace závěru práce Výborně (1)
- H. Splnění zadání (cílů) práce Splněno
- I. Skladba, správnost a úplnost citací literárních údajů Výborně (1)
- J. Typografická a jazyková úroveň (vč. pravopisu) Výborně (1)
- K. Formální náležitosti práce Výborně (1)
(struktura textu, řazení kapitol, přehlednost ilustrací)
- L. Přístup studenta k řešení (samostatnost, aktivita, ...) Výborně (1)

Komentáře či připomínky:

K práci nemám explicitní připomínky, protože student většinu mých průběžných připomínek a poznámek zapracoval do finální verze DP.

Oceňuji velkou motivaci studenta, jeho samostatnost a v neposlední řadě i nadprůměrný rozsah textové části DP.

... pokračuje na straně 2



Celkové zhodnocení:

Práce řeší aktuální téma, kterým je vyhledávání a příprava dat pro trénování moderních systémů rozpoznávání řeči typu End-to-end. Protože trénovací data jsou alfou a omegou těchto systémů, musí splňovat řadu požadavků, počínaje co největším objemem dat a jejich různorodostí (řečníci, témata, kvalita signálu), přes technické parametry (délka, zásady pro střih, atd). Zásadním požadavkem je ale maximální souhlas mezi tím, co bylo řečeno v nahrávce a co je uvedeno v jejím textovém přepisu. Student se proto zabýval metodami, jak automaticky získat velké objemy takových dat z veřejně přístupných zdrojů (parlament, televize, audioknihy + eknihy, atd.) Jím navržený systém se skládá z několika dílčích modulů, z nichž každý byl řešen variantním způsobem, experimentálně porovnáván a následně optimalizován. V průběhu řešení vznikla potřeba (daná mezinárodním projektem) připravit co nejvíce dat pro vývoj systému rozpoznávání mluvené dánštiny. Student to přijal jako velkou výzvu, kvůli které byl ochoten se naučit základy dánštiny a připravit pro tento jazyk moduly zvyšující celkovou výtěžnost dat. Jím vytvořený systém byl úspěšně nasazen na data z parlamentu, a po odevzdání práce i na další zdroje, jako jsou televizní pořady s titulky, Youtube videa, či podcasty. Jeho program se ukázal jako efektivnější pro tvorbu trénovacích dat než nástroje dosud využívané v týmu SpeechLab. V nejnovějším modelu pro dánštinu tvoří data získaná jeho systémem více než třetinu (cca 550 hodin) a další budou tímto způsobem připravována.

Výsledky práce hodnotím jako velice přínosné pro praxi a navíc použitelné a rozšiřitelné i v dalším výzkumu



Otázky k obhajobě:

V průběhu ověřování systému jste se setkal s různou mírou výtěžnosti trénovacích dat z různých zdrojů. Na čem všem tato míra závisí a jakými dalšími postupy by se dala ještě zvýšit?

Kontrola plagiátů:

Míra shody podle STAG 0 % (viz [www IS/STAG](http://www.IS/STAG))

Komentář v případě shody nad 5 %:

Celková klasifikace a doporučení k obhajobě:

Navrhuji tuto práci klasifikovat stupněm: Výborně (1)

Podpisem současně potvrzuji, že nejsem v žádném osobním vztahu k autorovi práce

V Liberci

dne 14.06.2023

.....
podpis vedoucího práce