|  |  |
| --- | --- |
|  | |
| **Srovnání současných modelů Speech-to-Text** | |
|  | |
| Bc. Juraj Králik | |
|  | |
|  |  |
| Diplomová práce  2025 |  |
|  |  |





**PROHLÁŠENÍ AUTORA DIPLOMOVÉ PRÁCE**

**Beru na vědomí, že**

* odevzdáním diplomové práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb., v platném znění bez ohledu na výsledek obhajoby;
* diplomová práce bude uložena v elektronické podobě v univerzitním informačním systému a bude dostupná k nahlédnutí;
* jedno vyhotovení diplomové práce v listinné podobě bude ponecháno Univerzitě Tomáše Bati ve Zlíně k uložení;
* na moji diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
* podle § 60 odst. 1 autorského zákona má Univerzita Tomáše Bati ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
* podle § 60 odst. 2 a 3 mohu užít své dílo – diplomovou práci – nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
* pokud bylo k vypracování diplomové práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tj. k nekomerčnímu využití), nelze výsledky diplomové práce využít ke komerčním účelům;
* pokud je výstupem diplomové práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá; neodevzdání této součásti může být důvodem k neobhájení práce.

**Prohlašuji, že**

* jsem na diplomové práci pracoval(a) samostatně a použitou literaturu jsem řádně citoval(a); v případě publikace výsledků budu uveden(a) jako spoluautor;
* odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou obsahově totožné.

Ve Zlíně, dne .............................. ...............................................................

podpis autora

ABSTRAKT

Diplomová práca sa zaoberá problematikou a porovnaním súčasných modelov technológie Speech-to-Text (STT). Cieľom práce je analyzovať dostupné modely z hľadiska ich presnosti, rýchlosti a schopnosti adaptácie na rôzne jazyky a prízvuky. Práca obsahuje teo-retický prehľad histórie a vývoja technológie STT, klasifikáciu modelov na základe použí-vaných technológií a podrobnú analýzu vybraných základných modelov.

Praktická časť je zameraná na zostavenie metodiky na porovnanie modelov a vyko-nanie testovania na vybraných datasetoch. Výsledky testov sú analyzované s dôrazom na porovnanie silných a slabých stránok jednotlivých modelov.

V závere práce sú zhrnuté hlavné poznatky, vrátane návrhov na možné reálne apli-kácie a odporúčaní pre ďalší výskum v oblasti Speech-to-Text technológií. Práca prispieva k lepšiemu pochopeniu možností súčasných modelov STT a ich využitia v rôznych oblastiach, ako sú automatizácia, prístupnosť alebo lingvistika.

Klíčová slova:

Klíčová slova: klíčové slovo, klíčové slovo

ABSTRACT

Text abstraktu ve světovém jazyce (angličtině).

Keywords: keyword, keyword

Zde je místo pro případné poděkování, popř. motto, úryvky knih atp.

OBSAH

Nebyla nalezena položka obsahu.

Úvod

Pro základní text jsou v šabloně připravené dva styly: Odstavec 1 a Text práce. Styl „Odstavec 1“ je bez odstavcové zarážky a používá se pro první odstavec pod nadpisem (nastaví se automaticky, jakmile nadpis ukončíte klávesou Enter), obrázkem, tabulkou nebo citátem. Styl je nastaven jako styl následujícího odstavce po nadpisových stylech. Následující odstavce jsou naformátované pomocí stylu „Text práce“, který je vyznačen odstavcovou zarážkou a je určen pro ostatní odstavce kromě prvního. Je nastaven jako styl následujícího odstavce po stylu Odstavec 1.

Šablona je navržena tak, aby bylo možné práci vytisknout oboustranně: má zrcadlové okraje a zrcadlová záhlaví. V šabloně se využívají oddíly, pomocí kterých jsou úvodní strany (zadání, prohlášení, abstrakt, obsah a úvod) nastavené tak, aby byly vždy liché, proto se může stát, že po liché stránce následuje opět lichá stránka. Není to chyba. Při tisku Word (či jeho alternativa) automaticky odešle na tiskárnu prázdnou stránku mezi lichými stranami. Obdobně to platí při exportu do PDF. Pokud se některé úvodní části upravují, doporučujeme zobrazit si v dokumentu všechny znaky (**tlačítko ¶ nebo Ctrl + Shift + 8**), aby se zabránilo rozhození ostatních stránek.

Pro bezproblémovou práci se šablonou doporučujeme prostudovat si „Manuál pro šablonu závěrečné práce na UTB ve Zlíně“, který je dostupný na univerzitním webu v sekci <https://www.utb.cz/student/dokumenty-a-sablony/>. Obsahuje pokyny pro:

* úpravu závěrečné práce (např. formální a typografická úprava práce),
* tisk práce,
* tvorbu seznamu použité literatury,
* vkládání obrázků a tabulek (a jejich citování), rovnic či zdrojových kódů,
* převod práce do formátu PDF/A.

Šablona vychází z aktuální směrnice rektora „Pravidla pro zadávání a zpracování bakalářských, diplomových a rigorózních prací, jejich uložení, zpřístupnění a kontrola původnosti“, která je dostupná na <https://www.utb.cz/univerzita/uredni-deska/vnitrni-normy-a-predpisy/smernice-rektora/>. V případě nejasností doporučujeme prostudovat tuto směrnici, případně příslušnou vnitřní normu součásti na webových stránkách součásti. Manuál a šablony byly aktualizovány v září 2024.

|  |  |
| --- | --- |
|  | teoRetická část |

Teoretický prehľad technológie speech-to-text

V této kapitole se nachází ukázka zápisu nadpisů pro různé úrovně.

História a vývoj STT technológií

Text podkapitoly.

Prehľad základných princípov a metód (akustické modely, jazykové modely, neurónové siete)

Text podkapitoly.

Klasifikácia STT modelov podľa technológií a použití

Text podkapitoly.

|  |  |
| --- | --- |
|  | praktická část |

Výber a charakteristika modelov

text

Kritéria výberu dostupných modelov

text

Príprava modelov

text

Metodika testovania a porovnávania

Text čtvrté hlavní kapitoly.

Kritéria hodnotenia modelov

Text podkapitoly.

Nástroje a prostredie použité na testovanie

Text podkapitoly.

Praktické testovanie modelov

Text čtvrté hlavní kapitoly.

Príprava datasetov

Text podkapitoly.

Postup testovania jednotlivých modelov

Text podkapitoly.

Zber a spracovanie výsledkov

Text podkapitoly.

Analýza a prezentácia výsledkov

Text čtvrté hlavní kapitoly.

Porovnanie modelov na základe zvolených metrík

Text podkapitoly.

Vizualizácia výsledkov

Text podkapitoly.

Rozbor silných a slabých stránok modelov

Text podkapitoly.

Možnosti využitia STT technológií v praxi

Text čtvrté hlavní kapitoly.

Budúci vývoj a trendy v oblasti STT

Text podkapitoly.

ZávEr

Text závěru práce.

Zoznam použitej literatúry

[1] JURAFSKY Daniel, MARTIN James. *Speech and Language Processing*. Pearson, Harlow. 2014. ISBN 9781292025438.

[2] GOODFELLOW, Ian. *Deep Learning*, The MIT Press. Cambridge, MA. 2017. ISBN 9798841205937.

[3] RABINER Lawrence, JUANG Biing-Hwang. *Fundamentals of speech recognition*. Pearson Education, Delhi, 2005. ISBN 9788129701381.

[4] YU Dong, DENG Li. *Automatic speech recognition : a deep learning approach*. Springer, London. 2015. ISBN 9781447157786.

[5] KAMATH Uday, LIU John, WHITAKER James. *Deep learning for NLP and speech recognition.* Springer, Cham, Switzerland. 2019. ISBN 9783030145965.

Seznam obrázků

[Obrázek 1 Popisek obrázku 17](#_Toc173623594)

Seznam tabulek

[Tabulka 1 Popisek tabulky 17](#_Toc173262262)

Seznam použitých symbolů a zkratek

STT Speech-to-text

Seznam příloh

Příloha P I: Název první přílohy

Příloha P I: Název první přílohy

Obsah první přílohy.