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

**Prohlašuji, že**

* beru na vědomí, že odevzdáním diplomové práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
* beru na vědomí, že diplomová práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně;
* byl/a jsem seznámen/a s tím, že 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;
* beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
* beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou práci nebo poskytnout licenci k jejímu využití jen připouští-li tak licenční smlouva uzavřená mezi mnou a Univerzitou Tomáše Bati ve Zlíně s tím, že vyrovnání případného přiměřeného příspěvku na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše) bude rovněž předmětem této licenční smlouvy;
* beru na vědomí, že pokud bylo k vypracování diplomové práce  
  využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými  
  subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu  
  využití), nelze výsledky diplomové práce využít ke komerčním  
  účelům;
* beru na vědomí, že 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 samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
  + že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně, dne …………………….

podpis studenta

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 teoretický prehľad histórie a vývoja technológie STT, klasifikáciu modelov na základe používaných technológií a podrobnú analýzu vybraných základných modelov.

Praktická časť je zameraná na zostavenie metodiky na porovnanie modelov a vykonanie 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 aplikácie 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:

ABSTRACT

Abstrakt ve světovém jazyce

Keywords:Poděkování, motto a čestné prohlášení, že odevzdaná verze diplomové práce a verze elektronická, nahraná do IS/STAG jsou totožné ve znění:

Prohlašuji, že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

OBSAH

[Úvod 3](#_Toc120888191)

1. [TEORETICKÁ ČÁST 3](#_Toc120888192)

[1 Nadpis 3](#_Toc120888193)

[1.1 Podnadpis 3](#_Toc120888194)

[1.2 Podnadpis 3](#_Toc120888195)

[1.2.1 Podpodnadpis 3](#_Toc120888196)

[2 Nadpis 3](#_Toc120888197)

[2.1 Podnadpis 3](#_Toc120888198)

[2.1.1 Podpodnadpis 3](#_Toc120888199)

1. [Praktická část 3](#_Toc120888200)

[3 Nadpis 3](#_Toc120888201)

[3.1 Podnadpis 3](#_Toc120888202)

[3.2 Podnadpis 3](#_Toc120888203)

[4 Nadpis 3](#_Toc120888204)

[4.1 Podnadpis 3](#_Toc120888205)

[Závěr 3](#_Toc120888206)

[Seznam použité literatury 3](#_Toc120888207)

[Seznam použitých symbolů a zkratek 3](#_Toc120888208)

[Seznam obrázků 3](#_Toc120888209)

[Seznam tabulek 3](#_Toc120888210)

[Seznam Příloh 3](#_Toc120888211)

Úvod

text

|  |  |
| --- | --- |
|  | Teoretický prehľad technológie Speech-to-text |

# TEORETICKÝ PREHĽAD TECHNOLÓGIE SPEECH-TO-TEXT

text

## História a vývoj STT technológií

text

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

text

## Klasifikácia STT modelov podľa technológií a použití

text

|  |  |
| --- | --- |
|  | Výber a charakteristika modelov |

# Výber a charakteristika modelov

text

## Kritériá výberu dostupných modelov (open-source, komerčné riešenia, podpora jazykov)

text

## Kritériá výberu dostupných modelov (open-source, komerčné riešenia, podpora jazykov)

|  |  |
| --- | --- |
|  | Metodológia porovnávania |

# Popis metodiky testovania a porovnávania

text

## Kritériá hodnotenia modelov

text

## Popis nástrojov a prostredia použitých na testovanie

|  |  |
| --- | --- |
|  | Praktické testovanie modelov |

# Príprava datasetov

text

## Postup testovania jednotlivých modelov

text

## Zber a spracovanie výsledkov

|  |  |
| --- | --- |
|  | Analýza a prezentácia výsledkov |

# Porovnanie modelov na základe zvolených metrík

text

## Vizualizácia výsledkov

text

## Rozbor silných a slabých stránok modelov

|  |  |
| --- | --- |
|  | Rozbor a reálne aplikácie technológie |

# Možnosti využitia STT technológií v praxi

automatizácia, prístupnosť, vzdelávanie

## Budúci vývoj a trendy v oblasti STT

text

## Závěr

text

Seznam použité literatury

[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 použitých symbolů a zkratek

|  |  |  |
| --- | --- | --- |
| ABC |  | Význam první zkratky. |
| B |  | Význam druhé zkratky. |
| C |  | Význam třetí zkratky. |
|  |  |  |

Seznam obrázků

[Obrázek 1. Ukázkový obrázek 3](#_Toc56699409)

Seznam tabulek

[Tabulka 1. Ukázková tabulka 3](#_Toc56699398)

Seznam Příloh

Příloha P I: Název přílohy