

## I. IDENTIFIKAČNÍ ÚDAJE

<b>Název práce:</b>	Rozpoznávání řeči s dostupnými internetovými moduly
<b>Jméno autora:</b>	Adam Jirkovský
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra řídicí techniky
<b>Oponent práce:</b>	Ing. Tomáš Bořil, Ph.D.
<b>Pracoviště oponenta práce:</b>	Fonetický ústav FF UK

## II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

<b>Zadání</b>	<b>průměrně náročné</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Jedná se o aktuální téma v prudce se rozvíjejícím oboru. Hlavním úkolem zadání bylo zorientovat se v dostupných modulech rozpoznávání řeči, vybrat na základně praktických kritérií nejvhodnější variantu, seznámit se s jejím rozhraním, naprogramovat front-end a otestovat úspěšnost rozpoznávání.	

<b>Splnění zadání</b>	<b>splněno</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>	

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

<b>Odborná úroveň</b>	<b>C - dobře</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
<p>Zhodnocení dosažené úspěšnosti rozpoznávání považuji za velmi důležitou část třetího bodu zadání. Přesto je mu věnována pouze jedna strana práce (str. 20). Postup evaluace je naprosto nedostatečně popsán, chybí zásadní informace pro reprodukovatelnost postupu. Pokud jsou zmiňovány dvě databáze, měly by být uvedeny reference. Nedožíváme se žádné informace o počtu položek, složení mluvčích. Způsob pořízení nahrávek je popsán také velmi vágně („pro každou promluvu dvě nahrávky lišící se kvalitou použitého mikrofону“). Jaké to byly mikrofony? V jaké vzdálenosti od mluvčího? V jakém typu prostředí?</p> <p>Překvapivě opačné výsledky oproti očekávání v tabulce 3.1 autor pouze konstatuje, nijak se nad nimi (ani v závěru) dále nezamýšlí. Vzhledem k tomu, že zde bylo porovnáno několik nastavení rozpoznávače (bez adaptace kontextu a s adaptací, pro niž autor připravoval vstupní data), očekával bych, že proces výběru textů pro adaptaci bude v práci vysvětlen, případně bude vyzkoušeno více variant a vyhodnoceny příslušné úspěšnosti WER za účelem vyšetření překvapivých dopadů.</p>	

<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>	
<p>Obrázky 4.1 až 4.5 jsou nadbytečné, nepřinášejí nové informace oproti textu a velmi výrazně zaplňují prostor práce. Obrázky 4.6 a 4.7 jsou z velké části duplicitní, vhodnější by bylo je sloučit do jednoho obrázku, na kterém by šlo obě záležitosti bez problémů demonstrovat. Obrázky 4.8 až 4.10 jsou opět prakticky zbytečné.</p>	

<b>Výběr zdrojů, korektnost citací</b>	<b>B - velmi dobře</b>
--	------------------------

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.

První bod zadání (seznámení se s principy rozpoznávání spojitě řeči a vypracování stručného přehledu používaných přístupů) realizovaný sekci 2.1 vychází celý téměř výhradně ze zdroje [8], tedy diplomové práce Kristýny Žákové z roku 2020 pod stejným vedoucím. I obrázek 2.1 je z ní zcela přejat (a správně ocitován). Jelikož se jedná o jeden ze tří hlavních bodů zadání, očekával bych, že se nebude jednat pouze o výtah z jedné dříve publikované závěrečné práce, ale bude pracováno s více primárními zdroji.

#### **Další komentáře a hodnocení**

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.

### **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.*

K obhajobě vznáším následující dotazy.

- Na str. 3 v prvním odstavci sekce 2.1 hovoříte o hláskách, které vznikají rezonancemi a artikulačními pohyby upravujícími zvuk kmitajících hlasivek o tzv. základní frekvenci. Existují v češtině hlásky, při kterých hlasivky nekmitají?
- Pro implementaci rozpoznávače jste si zvolil modul společnosti Google. Pokud by někdo chtěl takový systém použít pro vědecké účely, umožňuje tento modul replikovatelnost postupu se získáním stejných výstupů? Jinými slovy – předpokládám, že se jedná o organický systém, který společnost Google neustále vyvíjí a vylepšuje. Je zde nějaká možnost vrátit se k modelům a algoritmům určité konkrétní verze z konkrétního data, zafixovat toto nastavení a nevyužívat tak vždy jen nejnovější verzi modulu, díky které povedou stejné vstupy pravděpodobně postupem času na různé výstupy?
- Vzhledem k tomu, že v sekci výsledků ani v závěru se nedozvídám žádné hypotézy k tomu, proč v případě testované adaptace jazykového kontextu s využitím slovníku vygenerovaného z databáze CzLecDSP došlo k překvapivému zhoršení kvality rozpoznávání, rád bych se zeptal alespoň zde, jaká možná vysvětlení Vás napadají.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře**.

Datum: 30.5.2021

Podpis: