

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Rozpoznávání řeči s dostupnými internetovými moduly
Jméno autora:	Adam Jirkovský
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra řídicí techniky
Vedoucí práce:	Doc. Ing. Petr Pollák, CSc.
Pracoviště vedoucího práce:	Katedra teorie obvodů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání práce považuji za náročnější. Student se musel orientovat v principech relativně složitých systémů automatického rozpoznávání řeči, byť pro vlastní realizaci použil existující dostupný modul. Pro vlastní jádro práce bylo nutností zvládnout implementaci v jazyce Python a vytvoření výsledné aplikace v OS Linux i MS Windows. Ač bylo pro řešení použito dostupných modulů, na základě potřeby zvládnutí problematiky z různých oblastí nepředstavovalo řešení dané bakalářské práce zcela triviální úlohu.	

Splnění zadání	splněno
<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>	
Zadání bylo jednoznačně splněno v plném rozsahu.	

Aktivita a samostatnost při zpracování práce	A - výborně
<i>Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatně tvůrčí práce.</i>	
P. Jirkovský začal na řešení pracovat ihned po našem prvním kontaktu v minulém semestru v rámci semestrální práce. I v nastalé mimořádné situaci, kdy v důsledku karanténních opatření nebylo možné pravidelně osobně konzultovat, jsme pravidelně konzultovali v on-line módu a p. Jirkovský pracoval velmi samostatně, po konkrétním nasměrování byly z jeho strany jednotlivé úkoly průběžně a pravidelně řešeny a kompletovány, čímž jednoznačně prokázal schopnost samostatného řešení daných úkolů.	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Z mého pohledu předložená práce odpovídá nárokům kladeným na bakalářskou práci, a to jak z hlediska přehledového zvládnutí obecného teoretického základu rozpoznávání řeči, tak zejména pak z hlediska praktické realizace, kde p. Jirkovský prokázal jednoznačně programátorskou zručnost a schopnost rychlé orientace v dostupných nástrojích a v jejich vhodném použití. Kromě samotného použití dostupného Google modulu pro rozpoznávání byla řešena manipulace s audio-video záznamy v různých formátech, extrakce samostatné audio složky, a zejména pak implementace paralelního zobrazení AV záznamu a získaného přepisu pro účely prohlížení, možná korekce automaticky získaných přepisů. Významným výsledkem je samozřejmě i vlastní realizace přepisů souboru přednášek z předmětu Zpracování řeči a jejich příprava pro zveřejnění zapsaným studentům. Důležitým výsledkem je i orientační vyhodnocení úspěšnosti rozpoznávání při použití zvoleného modulu.	

Formální a jazyková úroveň, rozsah práce	A - výborně
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Formální a jazykovou úroveň práce považuji celkově také za velmi zdařilou. Výsledkem je dle mého názoru přehledný a srozumitelný popis řešené problematiky. Ocenil jsem sazbu textu práce v systému LaTeX.	

Výběr zdrojů, korektnost citací

A - výborně

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.

Diplomant aktivně a samostatně pracoval s dostupnými prameny, všechny relevantní prameny jsou řádně citovány dle zvyklostí.

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.

Kromě již výše zmíněného oceňuji u předložené bakalářské práce skutečnost, že vznikla funkční aplikace pro přepis a prohlížení audio-video záznamů, která bude jistě dále zdokonalována. Kromě možných dílčích vylepšení funkčnosti aplikace v této podobě je dalším možným cílem implementace prohlížení přes WEBové rozhraní a zejména pak implementace vlastního rozpoznávače spojitě řeči pomocí nástrojů sady KALDI. Toto však již jednoznačně přesahuje rámec řešení na úrovni bakalářské práce.

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

Předloženou bakalářskou práci považuji na základě výše uvedených skutečností za velmi dobrou a hodnotím ji jednoznačně klasifikačním stupněm **A - výborně**.

Datum: Klepněte sem a zadejte datum.

Podpis:

I. IDENTIFIKAČNÍ ÚDAJE

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

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání <i>Hodnocení náročnosti zadání závěrečné práce.</i>	průměrně náročné
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 nejhodnější variantu, seznámit se s jejím rozhraním, naprogramovat front-end a otestovat úspěšnost rozpoznávání.	

Splnění zadání <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>	splněno
---	----------------

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

Odborná úroveň <i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	C - dobře
<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 mikrofonu“). 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>	

Formální a jazyková úroveň, rozsah práce <i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	A - výborně
<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>	

Výběr zdrojů, korektnost citací	B - velmi dobře
--	------------------------

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. Uvedte 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: