HODNOCENÍ BAKALÁŘSKÉ PRÁCE

Vedoucí BP

Jméno bakaláře: Martin Stránský **Garantující katedra:** KKY

Název bakalářské práce: Neuronové sítě pro klasifikaci audio signálu

	Předmět hodnocení	Nadprůměrné	Průměrné	Podprůměrné
1	Jazyková a grafická úprava	☐	☐	☒
2	Samostatnost zpracování tématu BP	☐	☒	☐
3	Vhodnost použitých metod	☐	☒	☐
4	Způsob zpracování a vyhodnocení	☐	☐	☒
5	Správnost získaných výsledků	☐	☐	☒
6	Vlastní přínos	☐	☐	☒
7		☐	☐	☐

Doplnění hodnocení, připomínky, dotazy:

Předkládánou bakalářskou práci hodnotím z pozice vedoucího práce, který očekává častou interakci se svým studentem a iterativní postup při řešení práce. Bohužel v případě této práce je hlavní slabinou právě nedostatek této interakce ze strany jejího autora. Ten sice velkou část experimentálních výsledků získal samostatně, nicméně tyto výsledky již s vedoucím nekonzultoval a dále je nerozvíjel.

Nejvíce se vše projevilo v samotném textu bakalářské práce, autor se nevyvaroval vágních a často zavádějících nebo dokonce vyloženě nepravdivých formulací. Příkladem budiž výklad pojmů týkajících se hlubokových neronových sítí, v části 1.1.1. (str. 2) autor tvrdí "každý neuron má obecně weight (váhovou) matici", což je vyloženě chybně. Pokud by připojený obrázek byl vyjádřen pomocí jednoduché rovnice, bylo by zřejmé, že váhová matice se pojí obecně s celou vrstvou neuronové sítě. Obdobně nedostačující je i popis konvolučních neuronových sítí, který se v práci obešel bez jakéhokoli náznaku formálního popisu, nicméně s detailním uvedením popisu nástroje CNN Explainer, který však v rámci řešení nebyl vůbec použit. S tím velmi silně kontrastuje popis LSTM vrstvy, která je formalizována rovnicemi (1.5)-(1.9), bohužel bez popisu jednotlivých veličin. Je však nutné poznamenat, že ani LSTM vrstvy nebyly součástí praktického řešení bakalářské práce. V části popisující CTC ztrátovou funkci jsou uvedeny rovnice (2.7) až (2.11), které jsou však evidentně chybně - postrádají rovnítko, popř. členy na pravé straně.

Je rovněž nutné zmínit, že autor k finální úpravě textu bakalářské práce přistoupil až na poslední chvíli, o to více je zarážející ignorování některých úprav textu navrhovaných vedoucím, např. str. 31 (nedokončená věta v odrážkách), nebo poslední odstavec na str. 35 (nesmyslný text).

Autor práci doplňuje množstvím citovaných zdrojů, kvalita referencí je však velice nevyvážená. Příkladem mohou být zdroje [33] až [35], které odkazují webové stránky, pokaždé však jiným způsobem. Jméno autora ADMIN u zdroje [13] pouze dokresluje malou péči věnovanou této části práce. U některých zdrojů je velice obtížné zjistit, jakým způsobem souvisí text práce s obsahem zdroje, např. [39] cituje rozhlasový příspěvek z kategorie "Životní styl" týkající se Bible kralické, použitý na str. 34.

Při evaluaci dosažených výsledků jsou v práci popsány metriky WER, RIL a WIL, přičemž hodnoty ani jedné z nich nejsou uvedeny v závěrečných tabulkách. Ty obsahují nejen hodnoty slovní přesnosti a správnosti (zavedné v sekci 4.2.1), ale navíc i hodnoty prefixované jako SENT a CHAR, jejichž význam je čtenáři bakalářské práce zatajen.

V části uvádějící experimentální výsledky jsou pro každý experiment uváděny tabulky obsahující vždy dvakrát 30 číselných údajů, přičemž pro hodnocení výsledků je použita pouze hodnota slovní přesnosti

(dva číselné údaje). Ostatní metriky a hodnoty nejsou nijak komentovány. Úsměvně pak působí formulace "Rozdíl není enormní, leč jasně viditelný" v hodnocení výsledků na str. 49.

I přes výše uvedené nedostatky lze konstatovat, že bakalářská práce naplňuje všechny body zadání, proto ji doporučuji k obhajobě s hodnocením dobře.

Do diskuze navrhuji ujasnit následující připomínky a dotazy:

- Proč neuronové sítě používají více vrstev a nelineární aktivační funkce?
- Popište podstatu self-supervised learning (sekce 2.3.1 popisuje cíl tohoto učení, nikoli způsob).
- V práci operujete s termíny token, entita, jednotka. Jakým způsobem spolu souvisí a jak jsou použity v modelu Wav2Vec 2.0?
- V části 3.3 uvádíte, že korpus HHTT má část train obsahující 5240 vět, na další stránce však tvrdíte, že trénovací data sestávají z 70473 audio souborů a jejich anotací. Prosím o objasnění.

Splnění bodů zadání	☒ úplně	☐ částečně	☐ nesplněno	
Doporučení práce k obhajobě	☒ ano	☐ ne		
Celkové hodnocení práce	☐ **výborně**	☐ **velmi dobře**	☒ **dobře**	☐ **nevyhověl**
Jméno, příjmení, titul vedoucího BP: Ing. Jan Švec, Ph.D.				
Pracoviště vedoucího BP: KKY				

23. 8. 2022
Datum

J. Švec
Podpis