

Hodnocení vedoucího bakalářské práce

Autor práce: **Jan ČEDÍK**

Název práce: **Neuronové sítě pro zpracování přirozeného jazyka**

Jazyková a grafická úprava

Průměrné

Samostatnost zpracování tématu

Průměrné

Vhodnost použitých metod

Nadprůměrné

Způsob zpracování a vyhodnocení

Průměrné

Správnost získaných výsledků

Podprůměrné

Vlastní přínos

Průměrné

Doplnění hodnocení, připomínky:

Autor se věnuje velice aktuálnímu tématu využití velkých jazykových modelů založených na Transformer architektuře pro zpracování přirozeného jazyka. Práce byla rozpracována zhruba dva měsíce před vydáním průlomového řešení ChatGPT od společnosti OpenAI. Student v teoretické části rozebral principy tří základních architektur využívající Attention mechanismus - BERT, GPT, a T5. Dále byl sestaven robustní anglický dialogový korpus o celkové velikosti 55K dialogů (727K promluv), který byl navíc strojově přeložen do českého jazyka. Student sám navrhl a realizoval dvě vhodné metody augmentace dat. V experimentální části byly vyhodnoceny metody poskládání vstupních dat a jejich vliv v úloze chatovací platformy založené na T5. Vyhodnocení práce by si zasloužilo pečlivější zpracování a detailnější okomentování některých výsledků. Rovněž je škoda, že se nestihly dopočítat relevantní experimenty využívající všechna augmentovaná data. Původně plánovaná úloha analýzy sentimentu byla rozpracována, ale nebyla do finálního řešení aplikována. I přes tyto výhrady a také vzhledem k objemnému a tematicky náročnému zadání BP hodnotím práci výborně a doporučuji ji k obhajobě.

Dotazy

- Jaké další metody augmentace textových dat jste otestoval a s jakými výsledky?
- Chatovací platforma je klasicky míněna pro aplikaci v reálném čase. Testoval jste rychlost výsledného T5 modelu a máte představu, jak by se případně dal zrychlit?

Splnění bodů zadání

úplně

Doporučení k obhajobě

ANO

Hodnocení: 1 - Výborně

V _____ dne _____

Ing. Martin Bulín, M.Sc.