

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Bc. Patrícia Březinová

Název práce Computational analysis and synthesis of song lyrics

Rok odevzdání 2021

Studijní program Informatika **Studijní obor** Umělá inteligence

Autor posudku Mgr. Martin Popel, PhD **Role** vedoucí

Pracoviště ÚFAL MFF UK

Text posudku:

Hlavním tématem práce je automatická analýza písňových textů, zejména detekce rýmů, rozlišení různých typů rýmů a jejich kvality, detekce rýmových schémat a zhodnocení kvality (tzv. ratingu) celé písně z hlediska rýmů. Původním plánem bylo využít pro tyto úlohy existující nástroj SPARSAR, který se ale v průběhu práce ukázal nedostatečným z několika příčin popsaných v kapitole 3.1, a to navzdory nemalé snaze autorky o jejich vyřešení. Tím se těžiště práce přesunulo do vývoje vlastního nástroje pro tyto úlohy a nezbylo tolik času pro automatické generování písňových textů zmíněné v původním zadání. Krom toho se ukázalo, že problematika detekce a analýzy rýmů je složitější, než naznačovala úvodní rešerše a původní plány. Autorka v diplomové práci vyvinula vlastní řešení úloh, které ostatní (mně známé) nástroje řeší neadekvátně či je neřeší vůbec, například určení optimálního rýmového schématu při daném ohodnocení rýmové kvality všech dvojic veršů (kapitola 3.6.3), výpočet ratingu (kapitola 3.7) nebo návrh metriky pro porovnání dvou rýmových schémat (typicky predikovaného a ručně anotovaného, tedy vyhodnocení kvality detekce schématu; kapitola 4.1.2). Ve zpětném pohledu se navrhnutá řešení mohou zdát intuitivní a „samozřejmá“, ale u všech tří zmíněných příkladů jsme původně při konzultacích zvažovali několik alternativ, které se nakonec ukázaly jako horší. Celkově hodnotím zpracování tohoto hlavního tématu práce velmi kladně, jak z hlediska návrhu a implementace, tak z hlediska popisu v textu diplomové práce.

Autorka nad rámec původního zadání jednak vytvořila webový nástroj pro vizualizaci detekce rýmů a jejich kvality (kapitola 5), jednak zanalyzovala rozdíly v kvalitě, kvantitě a typech rýmů napříč pěti hudebními žánry ve velkém korpusu písní (kapitola 4.2). Obojí považuji za velmi přínosné.

Kapitola 6 o automatickém generování písňových textů má jen dvě stránky, což přibližně odpovídá i množství práce věnované této oblasti. Autorka původně zkoušela natrénovat vlastní model, ale nakonec převzala předtrénovaný model GPT-2 z projektu THEaiTRE a zaměřila se

na jednoduché prozkoumání vlivu délky vstupního textu (tzv. *primer*). Jak jsem uvedl výše, tato část práce je mnohem menší, než bylo plánováno v původním zadání (kde zřejmě bylo cílem natrénovat i specializované modely pro jednotlivé žánry), což ale považuji za adekvátní zvýšené obtížnosti a rozsahu zbylých částí práce. Tato změna priorit byla ovlivněna i změnou vedoucího (původní vedoucí nemohl práci dále vést, já jsem po seznámení se s dosaženými výsledky doporučil upřednostnit detekci rýmů, aby alespoň jedna část práce měla patřičnou hloubku), tedy okolností, které autorka nemohla ovlivnit.

Jako slabinu práce hodnotím to, že se analýza písňových textů omezila jen na rýmy (a jejich typy, kvalitu, schémata), ale opoměla metrickou strukturu (rytmus, typy stop), analýzu lexikální (poetičnost či bohatost jazyka) i morfologickou (libozvučnost i jinde než koncích řádků relevantních pro rýmy) a analýzu struktury celé písně (opakování slov, témat, veršů, slok/refrénu; délky veršů a slok). Přitom slabičná segmentace a detekce přízvuků je již implementována v detekci rýmu, tedy pro jednoduchou analýzu metra by stačilo vzít počet slabik a pozice přízvuků v každém verši, čímž by šlo zpřesnit i detekci rýmových schémat.

Z informatického hlediska by bylo zajímavé podrobněji popsat a prozkoumat problematiku hledání optimálního rýmového schématu vzhledem ke zvolenému způsobu výpočtu ratingu celé písně. Zejména by bylo zajímavé formálně dokázat, zda hladový algoritmus navržený v kapitole 3.6.3 vede vždy k maximalizaci ratingu. Z praktického hlediska by bylo zajímavé zkoumat, jak často vede optimalizace ratingu též k optimálnímu rýmovému schématu vzhledem ke zvolené metrice shody s ručně anotovaným schématem. Krom těchto dvou drobností (spíše námětů pro další výzkum než nedostatků) jsem s textem práce spokojen – je psána srozumitelnou a čtivou angličtinou s minimem překlepů. Oceňuji zejména kapitolu 1 s přiměřeně stručným přehledem typů rýmů a rešersí existujících nástrojů.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 23. 8. 2021

Podpis: