

Posudek vedoucího diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Bc. Kristýna Neumannová

Název práce German Compounds in Transformer Models

Rok odevzdání 2023

Studijní program Informatika **Studijní obor** Jazykové technologie a počítačová lingvistika

Autor posudku doc. RNDr. Ondřej Bojar, Ph.D. **Role** vedoucí

Pracoviště Ústav formální a aplikované lingvistiky

Text posudku:

Diplomová práce Kristýny Neumannové studuje specifický jev: produkci složenin v překladu z angličtiny do němčiny. Složeniny jsou pro němčinu typické, ale rozhodnutí, zda obsah vyjádřit složeninou nebo běžným opisem, záleží na nepřilíh uvědoměném stylistickém záměru. V úloze strojového překladu je proto zajímavé sledovat, jak moderní model Transformeru tento jev zvládá, a zda se jeho výstupy nějak systematicky liší od překladů lidských.

Kristýna ukázala, že v počtu vyprodukovaných složenin i nejlepší současné překladače za lidmi zaostávají, přičemž to platí i pro systémy, které jsou podle všeobecného hodnocení kvality velmi blízko překladům lidským. Druhé důležité pozorování je, že při identifikaci složenin se není možné spolehnout na jakkoli aktuální seznamy složenin nebo trénovací korpusy. Složeniny jsou příkladem produktivního jevu a i ve výstupech Transformeru jsme prokázali výskyt zcela nových složenin. Při omezení seznamem se tedy kvalita systému v tomto ohledu neprojeví. Na druhou stranu počet složenin dobře koreluje s celkovou kvalitou výstupu.

Kristýniny pokusy přimět Transformer k vyššímu počtu složenin na výstupu byly dobře motivované, bohužel však k významnému efektu nevedly. Práce tak končí v situaci, kdy víme, že (1) strojový překlad je (soudě podle složenin) méně propracovaný než lidský překlad, (2) stroje jsou schopny vyrábět nové a zcela akceptovatelné složeniny, (3) počet výskytů složenin nelze měřit pevným seznamem, i když (4) standardně měřená kvalita s počtem složenin koreluje. Tuto pozici shledávám jako ideální začátek pro další výzkum chování Transformeru: proč trénování vede k této určité nedostatečnosti, když principiálně Transformer složeniny tvořit umí? Jsem velmi rád, že Kristýna tuto situaci dobře zachytila a doufám, že i její příspěvek zaslaný na MT Summit bude přijat a podnítí další ke studiu této oblasti.

V průběhu celé spolupráce jsem u Kristýny oceňoval její spolehlivost, metodický přístup a samostatnost, a to jak na poli lingvistickém (validace kompozit), tak zejména na poli strojového překladu (trénování vlastních modelů a rozbor jejich chování).

Výsledná práce je psána velmi dobrou angličtinou, kvalitně vysázena a s jen nepatrným množstvím chyb nebo jiných nedostatků. Pro úplnost shrnuji obsah: Po úvodu první kapitola přináší výstižnou rešerši a poukazuje na to, že pro modely typu Transformer oblast kompozit studována nebyla. Následuje přehled použitých dat, nástrojů a modelů. Třetí, čtvrtá a pátá kapitola je těžištěm práce. Postupně zde Kristýna popisuje kompozita v lidských překladech a ve výstupech cizích systémů, následně vlastní trénování základního modelu a variant, které měly za cíl počet kompozit zvýšit. Závěr v šesté kapitole shrnuje všechna pozorování.

S Kristýninou prací jsem celkově velmi spokojen a doporučuji, aby byla přijata.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 5. 6. 2023

Podpis: