
Hodnocení vedoucího diplomové práce

Jakub Šmíd

Cross-lingual Aspect-Based Sentiment Analysis (Mezijazyčná aspektově orientovaná analýza sentimentu)

Obsah práce

Práce je zaměřena na úlohu aspektově orientované analýzy sentimentu (angl. aspect-based sentiment analysis – ABSA) ve více jazycích. Práce je výzkumného charakteru s potenciálním využitím výsledků v praxi. Cílem práce bylo vytvořit metodu založenou na moderních modelech architektury Transformer. Vytvořená metoda by měla umožňovat provést úlohu ABSA na více jazycích s využitím dat pouze z jednoho jazyka.

V první části práce student poskytuje ucelený a přehledný seznam jednotlivých ABSA úloh, kterých existuje velké množství a je poměrně náročné se v nich napoprvé zorientovat. Následně je popsána architektura Transformer, s důrazem na nejnovější tzv. *sequence-to-sequence (text-to-text)* modely. Dále diplomant zvládl nastudovat novou techniku, tzv. *prompting* aktuálně využívanou při řešení NLP úloh.

S využitím techniky *prompting* a *sequence-to-sequence* modelů, student navrhl a vytvořil novou metodu, která řeší více ABSA úloh najednou. Provedené experimenty ukazují, že tato metoda překonává výsledky dosažené pomocí standardního postupu a to jak pro jednojazyčnou tak i pro vícejazyčnou verzi řešené úlohy. Pokud je to možné, dosažené výsledky jsou porovnány s výsledky z existujících publikací. V některých případech diplomant dosáhl nových nejlepších tzv. *state-of-the-art* výsledků.

Původnost práce a kvalita řešení

Práce je zcela původní a je psaná v anglickém jazyce s minimem gramatických chyb a překlepů. Formální úroveň práce je v pořádku. Dokument bakalářské práce je vysázen v typografickém systému L^AT_EX a je použita vektorová grafika. Práce je implementována v jazyce Python a jsou použity standardní knihovny pro neuronové sítě. Zdrojový kód je přehledný, vhodně rozdělen do tříd a samostatných souborů a bylo možné jej bez problému spustit.

Přestože to z textu práce není možná přímo zřejmé nebo je to zmíněno pouze okrajově, diplomant provedl velké množství potřebných experimentů, které často nevedly ke zlepšení výsledků a k požadovanému řešení. Z výzkumného pohledu bylo potřeba tyto experimenty realizovat. Jejich provedení a vyhodnocení vyžadovalo velké množství času.

Využitelnost dosažených výsledků

Diplomant navrhl a vytvořil novou metodu, která umožňuje řešit ABSA úlohu i v jazycích pro které nejsou dostupná anotovaná data a označení těchto dat je obvykle velmi drahé a časově náročné. Jako vedlejší produkt práce vznikl i nový český dataset pro úlohu ABSA. Práce přesahuje rozsah standardní diplomové práce a je možné ji rozšířit na dvě až tři publikace. Z toho jedna publikace už byla zaslána na konferenci RANLP 2023.

Spolupráce s vedoucím a aktivita studenta

Spolupráce probíhala naprosto bezproblémově. Student se pravidelně účastnil konzultací a vždy včas plnil dohodnuté úkoly. Student pracoval samostatně a aktivně se snažil vnést do práce své vlastní nápady potřebné pro řešení problémů, které se v průběhu práce objevily.

Splnění zadání

Zadání bylo kompletně splněno. S prací jsem byl nad očekávání spokojen. Práce svou kvalitou, rozsahem a obtížností daného tématu přesahuje standardní diplomové práce.

Navrhuji hodnocení známkou **v ý b o r n ě** a práci doporučuji k obhajobě.

V Plzni dne 2. června 2023

Ing. Pavel Přibán
(vedoucí DP)