
Posudek oponenta diplomové práce

Pavel Průcha
Kategorizace zákaznické zpětné vazby nezávislá na jazyce

Obsah práce

Cílem práce je vytvořit datovou sadu recenzí a implementovat metodu, která dovoluje klasifikovat kategorie a polaritu těchto recenzí ve více jazycích.

V první části práce je popsána teorie NLP a základní techniky zpracování textu a jeho reprezentace, včetně stručného popisu distribuční sémantiky a modelů používajících architekturu Transformer. Bohužel modely skip-gram, CBOW a fastText nejsou matematicky vůbec popsány, což považuji za velký nedostatek vzhledem k tomu, že jsou použity při implementaci a jsou důležitou součástí práce. Dále jsou poměrně podrobně popsány algoritmy strojového učení pro klasifikaci textu.

Diplomant se následně věnuje možnostem mapování vektorových prostorů mezi jazyky. V textu chybí analýza vybraného řešení (resp. použitého modelu) i relevantních a podobných prací. Diplomant pouze krátce zmíní, že pro použití novějších metod je potřeba velký výkon a paměť. To nepovažuji za správné odůvodnění, protože studenti mají k dispozici výpočetní zdroje organizace MetaCentrum. Popis použitých modelů je velmi strohý. Pro modely CNN a LSTM mi v textu chybí jejich vizualizace a detaily architektury a procesu trénování (např. počet parametrů, počet epoch, batch size). Z textu není příliš zřejmé, jak je řešena samotná klasifikace (resp. architektura), zda jako multi-label nebo multi-class.

Kvalita řešení a dosažených výsledků

Pro stažení recenzí, je použita existující služba. To považuji jako vhodné řešení, protože implementovat vlastní stahování recenzí z Google Maps je poměrně komplikované. Recenze jsou následně manuálně anotovány pouze diplomantem, ale malá množina dat byla také anotována vedoucím a spočtena metrika udávající shodu. Toto ověření považuji za správné z pohledu ověření kvality vytvořeného datasetu.

Pokud jsem správně pochopil řešenou úlohu, tak se jedná o tzv. aspektově orientovanou analýzu sentimentu (angl. aspect-based sentiment analysis – ABSA), ale autor to nijak nezmiňuje, nevím zda si to neuvědomuje nebo zda je to záměr. Pro tuto úlohu již existují české datasety.

Provedené experimenty jsou relevantní. Výsledky dosahují poměrně dobré úspěšnosti, ale bohužel jsou rozděleny do několika tabulek a je náročné je mezi sebou porovnat (např. model SVM a LSTM). Zdrojové kódy bylo možné spustit. Kód je komentován a poměrně dobře dekomponován do jednotlivých modulů a tříd.

Formální a jazyková úroveň

Formální úroveň textu je v pořádku. V textu se vyskytuje minimum překlepů a gramatických chyb a diplomant si dal na textu zjevně záležet. Bohužel většina obrázků je rastrová a ve špatném rozlišení.

Práce s literaturou

Velké nedostatky vidím v citovaných zdrojích. Autor velmi často cituje internetové stránky obsahující návody a popisy řešené problematiky místo výzkumných článků a jiných důvěryhodných publikací. V diplomové práci tohoto typu by internetové stránky neměly být primárním zdrojem literatury. Autor necituje modely word2vec a fastText.

Z textu práce a ze zdrojových souborů mám dojem, že diplomant se inspiroval bakalářskou prací nazvanou *Detekce polarity textu s využitím mezijazyčné transformace* a článkem *Linear Transformations for Cross-lingual Sentiment Analysis*, které ale nikde v práci necituje ani nijak nezmiňuje.

Splnění zadání

I přes výše uvedené nedostatky považují zadání práce za splněné.

Dotazy k práci

1. Jaká je architektura použitých modelů CNN a LSTM a na základě čeho jste ji navrhnul?
2. Jak je implementována klasifikace kategorií a sentimentu? Kategorie a sentiment se klasifikují odděleně nebo společně?
3. Jak se liší řešená úloha od aspektově orientované analýzy sentimentu (angl. aspect-based sentiment analysis)?

Navrhuji hodnocení známkou **dobře** a práci doporučuji k obhajobě.

V Plzni dne 2. června 2023

Ing. Pavel Přibán
(oponent DP)