

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Patrik Romanský

Název práce Automatická detekcia fake-news na slovenských textoch

Rok odevzdání 2023

Studijní program Informatika **Studijní obor** Umělá inteligence

Autor posudku David Mareček **Role** vedoucí

Pracoviště ÚFAL MFF UK

Text posudku:

Hlavním cílem práce je porovnání různých metod pro řešení konkrétního úkolu zpracování přirozeného jazyka (v tomto případě detekce fake-news) na jazyce, pro který existuje menší množství datových zdrojů (v tomto případě na Slovenštině). V práci je srovnáno a vyhodnoceno velké množství možnosti z hlediska využití různých jazykových modelů a trénovacích dat, tj. automatický překlad do angličtiny a použití velkých anglických modelů, použití multilinguálních modelů, nebo použití slovenských modelů bez nutnosti překladu, dále pak použití anglických nebo slovenských dat pro fine-tuning těchto modelů, případně jejich kombinace.

Práce obsahuje celkem třináct kapitol. Po úvodu do problematiky následuje shrnutí souvisejících prací, historie fake-news, a metody detekce. Následují metody přípravy dat, způsoby evaluace tohoto problému. Dále pak celkem rozsáhlý popis všech možných způsobů strojového učení od klasických, jako je naive-bayse nebo random forest, přes neuronové sítě (perceptron, vícevrstvé, konvoluční, rekurentní), až po moderní velké neuronové sítě s architekturou Transformer (BERT, RoBERTa). Další kapitola pak popisuje slovenské a anglické datasey fake-news, potřebné k trénování a evaluaci metod. Kapitola 10 si pak všímá toho, že mezi nejvýznamějšími rysy pro detekci fake news jsou často metadata, jako jméno autora nebo zdroj článku. Hlavním úkolem práce je však rozpoznat fake-news ze samotného textu. Poslední tři kapitoly jsou věnovány experimentům, a to s použitím základních metod strojového učení, konvolučních a rekurentních neuronových sítí a sítí s architekturou Transformer. V závěru je pak shrnut celý příběh práce a vyzdvihnuty hlavní výsledky.

Hodnocení práce:

Práce je celkem rozsáhlá, obsahuje 108 stránek čistého textu a je psána slovensky. Práce není úplně vyvážená, zbytečně mnoho prostoru je v ní věnováno základním metodám a experimentům s nimi (celkem 55 stránek), zatímco Transformerům, o kterých jsme od začátku věděli, že budou

nejlepší, je věnováno pouze celkem cca 25 stran. Důvodem je zejména to, že Patrik se dlouho věnoval datům a základním metodám, dlouho měl napsanu velkou část práce, ale experimenty s Transformerem, tedy tu nejzajímavější a nejdůležitější část dělal až jako poslední. S tím souvisí i některé nedostatky textu, jako například v Kapitole 4 (Příprava vstupních dat), kde uvádí metody lematizace a vypouštění stop-words jako důležitou součást předzpracování dat, což je značně zavádějící, protože se tyto metody pro sítě Transformer vůbec nepoužívají. Dále pak je v textu značné množství překlepů, nedokončených vět, nevysvětlených pojmů a zkratk a jiných nejasností. Jako velmi dobře napsaný hodnotím závěr práce, který celou práci pěkně vysvětluje a přehledně shrnuje všechny poznatky. Věřím, že kdyby takto, tedy až po všech provedených experimentech napsal celou práci, byla by daleko kvalitnější a konzistentnější.

Z práce není příliš patrné to, že velké množství času věnoval Patrik vytvoření datové sady, tedy čištění, filtrování a kontrole datasetu SlovakCovid19 FN. Podle mého názoru jde o kvalitní data, navíc na jiném než anglickém jazyce, které mohou být užitečná pro následné multilinguální evaluace.

Co se týká experimentální části, jsem spokojen s množstvím experimentů, co se týká kombinací dostupných dat a modelů. Chybí tu ale řada dalších experimentů potvrzujících správnost výsledků. Například, fine-tuning trvá dle práce vždy 20 epoch. Je to dost? Jak se liší výsledky v závislosti na počtu epoch? Ideální množství epoch by také mělo záviset i na použitém datasetu a fine-tunovaném modelu. Dále pak chybí nějaká analýza výsledků. Jak moc se například Transformer naučil rozpoznávat jméno autora? Pokud by začal publikovat fake-news nějaký nový autor, rozpoznal by to?

Prezentace výsledků je rovněž slabá, například výsledky jednotlivých experimentů jsou v tabulkách na různých stránkách a pro porovnání čísel je třeba listovat. Chybí nějaké porovnávací grafy, ze kterých by byly patrné závěry.

Závěr:

V práci byl vytvořen dataset a bylo provedeno rozumné množství experimentů odpovídající úrovni diplomové práce. Po prezentační stránce je ale práce slabší, obsahuje velké množství chyb, nedodělků a výsledky jsou prezentovány velmi nepřehledně.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 6. června 2023

Podpis: