FAKULTA
APLIKOVANÝCH VĚD
ZÁPADOČESKÉ
UNIVERZITY
V PLZNI

Ing. Ondřej Pražák
Katedra informatiky a výpočetní techniky
Fakulta aplikovaných věd
Technická 8, CZ-30614 Plzeň
Tel.: (+420) 377 63 2151
E-mail: ondfa@ntis.zcu.cz

Posudek oponenta diplomové práce

Bc. Milan Kuda
Korelace a kauzalita sentimentu extrahovaného z textu a tepové frekvence

---

Cílem práce bylo navrhnout metody pro extrakci sentimentu z tepové frekvence a ověřit, jestli existuje korelace se sentimentem příspěvků, které subjekty psaly během měření na síť Twitter.

Práce je psaná velmi dobrou češtinou, počet gramatických chyb odpovídá rozsahu práce. Bohužel některé kapitoly autor pravděpodobně psal ve spěchu, protože je u nich gramatická úroveň podstatně nižší, než ve zbytku práce (například začátek kapitoly 8). Text je psaný košatou češtinou, kde spousta slov nenese prakticky žádnou podstatnou informaci.

Struktura textu práce není logicky členěna. Práce je rozdělena na teoretickou, analytickou, implementační část a experimenty. Mnoho rozhodnutí o návrhu experimentu je ale prezentováno už v teoretické části a některé části jsou v práci zopakované vícekrát. Kapitolu *Strojové učení* bych doporučil posunout do dřívější části práce, protože například kapitola o extrakci sentimentu se na strojové učení často odkazuje. Poněkud chaotická struktura je podle mě zčásti způsobena příliš širokým zadáním.

Co se týče práce s literaturou, diplomant ve své práci cituje velké množství zdrojů, bohužel některé z nich vůbec nejsou relevantní (například článek *Automatic Indexing: An Experimental Inquiry* jako reference k NBK). To ale není příliš častý jev, takže jde pravděpodobně o chybu z nepozornosti a celkově jsem s prací s literaturou spokojen.

V práci mi také chybí detailnější popis cíle, a toho, jak spolu jednotlivé části souvisí (například proč autor používá nástroje na automatickou extrakci sentimentu z textu a neporovnává metody extrakce z tepové frekvence přímo s lidmi označenými daty). V části zabývající se strojovým učením a NLP je z textu znát pouze povrchní porozumění a někdy dokonce neporozumění danému problému. Některé definice jsou špatně přeložené a je z nich patrné, že autor nepochopil, o čem píše.

Největší nedostatky vidím v prováděných experimentech, na kterých je vidět, že byly prováděny ve velkém spěchu. Často chybí odůvodnění prováděného experimentu a popis souvislostí mezi jednotlivými experimenty. Také proto v některých experimentech nevidím smysl. Většina závěrů práce je podložená nerelevantními nebo nesprávně interpretovanými fakty. To je z největší části způsobeno nevhodnou evaluační metrikou (která dává větší váhu pozitivní třídě) a nevyváženými testovacími daty (zhruba 70:30) při trénovacích datech vyvážených. Kombinace obou problémů způsobuje, že pouhá klasifikace do nejpravděpodobnější třídy by dosáhla úspěšnosti 82.19%, což je přibližně o 4% více, než nejlepší výsledky prezentované v práci. Dále diplomant optimalizoval hyperparametry na testovacích datech, čímž byly výsledky ještě více zkreslené. To, že vysoká hodnota takto spočítané metriky nemusí znamenat lepší výsledky, napovídají například grafy 9.2 a 9.3 nebo tabulka 9.3, kde se jedna z metod s přesností 35.02% (a s F mírou 19.34%) vlastně naučila s 65% přesností předpovídat opačnou třídu. Chybná interpretace F míry je hlavním důvodem chybných závěrů. Kvůli tomu považuji vědecký přínos práce za minimální. Pokud bylo cílem práce zjistit, jestli existuje závislost mezi tepovou frekvencí a sentimentem (například nízká tepová frekvence je častěji spojena s negativním sentimentem, jak je prezentováno v závěru) pak určitě není strojové učení vhodný nástroj k dosažení tohoto cíle. Pokud bylo cílem určit sentiment z tepové frekvence (a

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky
①

případně dalších údajů o subjektech), jak uvádí zadání, potom mi v práci chybí srovnání s nějakým referenčním modelem (baseline).

Ke splnění zadání práce mám velké výhrady. Body 1 a 3 jsou splněné pouze zčásti, protože práce se vůbec nezabývá ani detekcí vzorů v časových řadách ani zpracováním časové řady pomocí metod strojového učení. Kvůli výše zmíněným problémům nebyl správně splněn ani poslední bod zadání.

Softwarové řešení spočívá v použití externích knihoven, jedná se pouze o pár jednoduchých skriptů.

S ohledem na výše zmíněné připomínky práci **nedoporučuji** k obhajobě a hodnotím klasifikačním stupněm

„nevyhověl“.

Doplňující otázky:

1. Proč jste učil klasifikaci sentimentu z tepové frekvence na výstupu z Vaderu a ne na lidmi označených datech?
2. Proč jste očekával F míru kolem 70%, jak v práci uvádíte?

Ing. Ondřej Pražák
(oponent DP)

V Plzni 3. června 2019

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky