

## Posudek oponenta bakalářské práce

**Student:** Bezák Adam  
**Téma:** Analýza postojů v oblasti automobilového průmyslu (id 19608)  
**Oponent:** Otrusina Lubomír, Ing., UPGM FIT VUT

- 1. Náročnost zadání** průměrně obtížné zadání
- 2. Splnění požadavků zadání** zadání splněno s drobnými výhradami

Zadání bylo splněno v minimální možné míře. Drobné výhrady mám ke splnění bodu č. 1. Student sice v práci zmiňuje některé přístupy a metody počítačové analýzy postojů. Tyto metody jsou však v práci popsány nedostatečně a z práce není zřejmé, jestli je student správně pochopil.
- 3. Rozsah technické zprávy** je v obvyklém rozmezí

Rozsah technické zprávy je v obvyklém rozmezí. Všechny části technické zprávy jsou informačně bohaté a pro práci skutečně nezbytné.
- 4. Prezentační úroveň předložené práce** 55 b. (E)

Prezentační úroveň technické zprávy je špatná. Práce je v některých místech hůře pochopitelná pro čtenáře. Jednotlivé kapitoly na sebe většinou navazují. Kapitola 2, která popisuje rozbor řešené problematiky by měla obsahovat podstatně detailnější popis algoritmů pro analýzu sentimentu. Kapitulu 5 obsahující příklady použití systému a kapitolu 6 obsahující vyhodnocení systému by také stálo za to rozšířit. Práce také obsahuje mnohá nepravdivá, nepřesná či nepodložená tvrzení.
- 5. Formální úprava technické zprávy** 55 b. (E)

Po typografické a jazykové stránce práce obsahuje několik prohřešků. Práce obsahuje chyby v anglickém abstraktu. Na mnohých místech v textu chybí mezery. Některé názvy kapitol jsou zvoleny nevhodně, např. První příklad, Druhý příklad atd. Odkazy na kapitoly, obrázky či tabulky jsou někdy chybně zakomponovány do textu.
- 6. Práce s literaturou** 60 b. (D)

Zdroje použité v práci jsou většinou smysluplné. Nicméně v práci není vždy zcela jasné odkud student čerpal (např. odkud převzal informace o klasifikátoru SVM). Student také uvádí, kdy daný zdroj navštívil online i u zdrojů, kde to nedává příliš smysl, např. u vědeckých článků publikovaných ve sborníku.
- 7. Realizační výstup** 50 b. (E)

Kvalita realizačního výstupu je velmi nízká. Student v práci zmiňuje několik algoritmů, které lze použít k analýze sentimentu. Nicméně ve výsledné aplikaci použije již hotový systém VADER (tedy nepoužívá žádný vlastní natrénovaný klasifikátor). O tomto systému navíc v práci nejsou téměř žádné zmínky, a to ani v sekci popisující návrh systému. Obdobně pak v návrhu zmiňuje využití vizualizačního nástroje Kibana, který ale stejně nakonec nepoužije. Vyhodnocení vzniklého systému je diskutabilní. Názvy souborů, proměnných a skriptů jsou často voleny nepřehledně.
- 8. Využitelnost výsledků**

Jedná se o práci kompilačního charakteru. V aktuálním stavu není vytvořený systém pravděpodobně příliš použitelný.
- 9. Otázky k obhajobě**
  - Z práce vyplývá, že jste pro stahování dat z Twitteru použil Twitter Search API. Proč jste nepoužil Twitter Streaming API, které je k tomuto účelu vhodnější?
  - V práci uvádíte, že jste se pokoušel natrénovat vlastní Naivní Bayesův klasifikátor na ručně vytvořené datové sadě 1000 tweetů. Kvůli nízké úspěšnosti jste ale od tohoto záměru ustoupil. Můžete komisi popsat, na jaké datové sadě jste ten klasifikátor trénoval, jaké příznaky jste k trénování používal a jakých výsledků jste dosáhl?
- 10. Souhrnné hodnocení** 55 b. dostatečně (E)

Zadání práce bylo splněno v minimální možné míře. Technická zpráva je v některých místech nepřehledná a obsahuje mnohé chyby. Kvalita realizačního výstupu je velmi nízká. Vyhodnocení implementovaného systému je diskutabilní. Z těchto důvodů navrhuji hodnocení stupněm E.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 1. června 2017

.....  
podpis