

Review of Master's Thesis

Student: Marušic Marek, Bc.

Title: Deep Neural Networks Used for Customer Support Cases Analysis (id 20853)

Reviewer: Ryšavý Ondřej, doc. Ing., Ph.D., UIFS FIT VUT

- 1. Assignment complexity** **average assignment**
Práce se zaměřuje na řešení praktického problému přidělování úloh/případů pomocí metod umělé inteligence. Student musel nastudovat několik algoritmů a použil je pro vytvoření experimentálního nástroje.
- 2. Completeness of assignment requirements** **assignment fulfilled**
- 3. Length of technical report** **in usual extent**
- 4. Presentation level of technical report** **80 p. (B)**
Práce má logickou strukturu a postupně čtenáře uvádí do problematiky zpracování požadavků zákazníků na podporu společnosti RedHat. Úvod do práce je čtivý, ale místy až velmi pozvolný. Bylo by například vhodné stručně uvést plánované použití AI již v úvodní kapitole. Student dále uvádí přehled různých druhů neuronových sítí a možnosti reprezentace slov pro účely klasifikace textů. Kapitola 4 se zabývá návrhem systému. Na začátku je uvedeno existující řešení a cíle práce. Zde je text poněkud obtížně pochopitelný a chvíli trvá, než se čtenář dozví co je konkrétním cílem práce. Poté ihned následuje návrh samotné aplikace. Podkapitola 4.5 popisuje postup při trénování klasifikátorů. Student zde uvádí nutnost použití GPU bez dalších informací, které by toto tvrzení podpořily. Kapitola 5 se pak celá věnuje experimentům. Prvním experimentem je použití alternativních reprezentací slov. Nicméně z uvedených informací je zřejmé, že dosažené výsledky jsou horší než pro Word2Vec použitý v původním systému.
Další oblastí je návrh klasifikátoru pro výběr skupiny konkrétních možných řešitelů problému. Zde je dosažená přesnost okolo 15% při úloze určení nejvhodnějšího experta pro řešení problému. Dále se pak vyhodnocuje přesnost různých metod pro úlohu nalezení skupiny vhodných expertů různých velikostí, kde je již dosažená přesnost zajímavá pro praktické použití.
- 5. Formal aspects of technical report** **85 p. (B)**
Práce je psána v anglickém jazyce. Některé věty textu jsou poměrně komplikované, ale celkově je práce srozumitelná. Tok textu je na některých místech ovlivněn tím, že autor umísťuje obrázky v textu, čímž vznikají prázdná místa na konci stránek.
- 6. Literature usage** **80 p. (B)**
Autor použil větší množství relevantních studijních zdrojů, týkajících se zejména metod umělé inteligence pro zpracování přirozeného jazyka. Některé informace byly převzaty z online zdrojů. Převzaté informace v textu jsou odlišeny od vlastních výsledků. Student použil dostupné znalosti z oblasti neuronových sítí a NLP při návrhu řešení.
- 7. Implementation results** **75 p. (C)**
Realizačním výstupem je experimentální implementace klasifikátorů požadavků podpory pro různé zvolené metody. Tato implementace byl použita pro provedení experimentů, jenž mají za cíl ukázat možnosti vylepšení stávajícího systému použitého pro SBR. Pro implementaci bylo použito dostupných knihoven s algoritmy strojového učení. Forma realizačního výstupu je vhodná pro experimentální vyhodnocení a částečně jako základ pro finální nástroj. Implementační jazykem je Python, kódy komentované nejsou. Zdrojové kódy jsou doplněny README souborem, kde je uvedeno jak lze nástroj použít.
- 8. Utilizability of results**
Práce se snaží vylepšit stávající systém. Student vytvořil návrh použití algoritmů umělé inteligence pro zpracování požadavků zákazníků na podporu společnosti RedHat. Práce obsahuje důsledné experimentální vyhodnocení, které ukazuje možné kandidáty na vhodné řešení.
- 9. Questions for defence**
 - Na straně 35 je vyhodnocení problému nalezení k-možných expertů pro řešení problému, ve kterém porovnáváte Vámi navržené metody s existujícím SBR. Proč jste použil lineární aproximaci přesnosti pro SBR klasifikátor? Z dat se zdá, že jeho přesnost by ve skutečnosti mohla být vyšší než vypočtená.
 - Máte nějaké vysvětlení nízké přesnosti klasifikace nalezení k-možných expertů pro řešení problému pro k menší než polovina velikosti skupiny SBR?
- 10. Total assessment** **85 p. very good (B)**

Práce je zdařilým dílem, které se snaží vyřešit praktický problém za použití moderních metod. Silnou stránkou práce je experimentální vyhodnocení navržených přístupů. Student zpracoval práci v kontextu reálného prostředí a použil existující data pro návrh a experimentální vyhodnocení klasifikátorů. Díky důkladnému vyhodnocení je možné výsledky práce dále využít pro vytvoření nového SBR klasifikátoru zákaznické podpory společnosti RedHat.

In Brno 7. June 2018

.....
signature