
Hodnocení vedoucího diplomové práce

Michal Seják
Summarization of News Articles

Téma diplomové práce Michala Sejáka je sumarizace novinových článků. Text celé práce je dobře strukturovaný, teoretická část adekvátně popisuje evaluační metriky a problematiku extraktivní sumarizace včetně moderních modelů užívaných v doméně zpracování přirozeného jazyka.

Tabulky, vzorce a diagramy vhodně doplňují text a pomáhají popisovanou problematiku čtenáři lépe pochopit. Autor nekonzistentně užívá velká písmena v nadpisech.

Práce vyčnívá množstvím odvedené práce, přístupem studenta a jeho odborností. Diplomant připravil systém pro automatickou extraktivní sumarizaci textů v českém jazyce nejmodernějšími metodami na něm nově připravené datové sadě, kterou vytěžil z dostupných dat velice důmyslným způsobem. Nejmodernější metody rovněž sám analyzoval, upravil a ukázal přínos jeho úprav. Výsledkem je systém pro automatickou extraktivní sumarizaci na vysoké technické úrovni bez nutnosti velkých investic pro uplatnění v praxi. Rovněž z diplomové práce vzniká odborný článek, který odešleme na konferenci. Student má již publikační historii.

Student pracoval většinu času samostatně, avšak si domluvil konzultaci kdykoli měl zajímavé téma k diskusi. Mezi hlavní přínosy patří:

- Nový přístup pro extraktivní sumarizaci více dokumentů užívající neuronové sítě.
- Nový český dataset pro trénování a vyhodnocování modelů pro extraktivní sumarizaci více dokumentů.
- Sada experimentů, které poslouží jako dobrý základ pro publikaci na odborné konferenci.

Student splnil zadání ve všech bodech. Práci **doporučuji k obhajobě** a navrhuji hodnocení známkou **výborně**.

V Plzni dne 6. 6. 2022

Ing. Jakub Sido
KIV FAV ZČU