

Hodnocení vedoucího diplomové práce

Autor/autorka práce: **Jakub Morávka**

Název práce: **Klasifikace dokumentů s použitím hierarchické reprezentace**

Cílem diplomové práce bylo vyzkoušení metod pro klasifikaci textových dokumentů, které využívají hierarchii v datech. Dokumenty mohou patřit do více tříd zároveň. Dokument je v tomto případě reprezentován pomocí menších logických celků, které mohou být dále členěny. Předpoklad je, že každá část může spadat do jiné klasifikační třídy. Výsledná klasifikace je pak provedena na základě výsledků pro jednotlivé části. Pro reprezentaci částí textu jsou využity distribuované reprezentace. Jako klasifikátor jsou vybrány neuronové sítě. Metody jsou ověřeny na českých a anglických datech.

V rámci práce bylo nutné prostudovat velké množství materiálů. Práci s literaturou považuji za nadstandardní. Přístup k řešení problému byl od počátku aktivní a postup práce byl pravidelně konzultován s vedoucím DP.

Program prezentovaný formou knihovny je napsán v jazyce Python. K implementaci je použita knihovna Keras, která poskytuje přehledné rozhraní pro vytváření neuronových sítí.

Dokument je na velmi vysoké úrovni a je přehledně a logicky členěný. Použité zdroje jsou korektně citovány. Práce neobsahuje pravopisné chyby a překlepy.

V rámci práce bylo provedeno rozsáhlé srovnání různých metod reprezentace dokumentů. Experimenty ukázaly, že pomocí hierarchických reprezentací je možné dosáhnout výsledků srovnatelných s metodami prezentovanými v literatuře. Je zde potenciál pro využití navržených metod praxi. Testování bylo provedeno na standardních datech, kde se nepředpokládá hierarchické uspořádání. Je zde tedy předpoklad, že na určitých typech dat by navržené metody překonaly běžně používané přístupy.

Zadání bylo splněno bez výhrad.

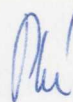
Navrhuji hodnocení známkou **výborně** a práci doporučuji k obhajobě.

V Plzni 8.8.2018

Ing. Ladislav Lenc, Ph.D.



SOUHLASÍ
S ORIGINÁLEM



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

①