

# Posudek oponenta bakalářské práce

**Autor práce:** Roman Zeleník

**Název práce:** Framework pro prezentaci NLP algoritmů

## **Obsah práce:**

Bakalářská práce se zabývá návrhem a implementací webové prezentace pro vyzkoušení NLP algoritmů. Hlavní motivace je usnadnění práce při přidávání nových algoritmů.

## **Kvalita řešení a dosažených výsledků:**

Framework je prezentován na třech NLP úlohách: sumarizace, rozpoznávání pojmenovaných entit a rozpoznávání jazyka. Vše působí přehledným dojmem - jednoduché a intuitivní ovládání. Framework navíc ukládá základní statistiky o přístupu uživatelů do databáze a umožňuje vizualizaci těchto statistik.

Ve vytvořeném frameworku mi chybí administrace dem, tzn. něco, co by umožňovalo uživatelsky přívětivé vytvoření nového či editaci již vytvořeného dema a veškeré jeho nastavení. Ve frameworku tak jak je vytvořen je nutné ručně vytvořit či upravit několik konfiguračních souborů, což nepovažuji za šťastné. Navíc zde není žádná kontrola chyb. To, že uživatel všechny soubory upravil tak jak měl, pozná až podle toho, že demo funguje nebo nefunguje.

Dále bych očekával více péče se serverovou částí frameworku. Je napsána v jazyce PHP, což nepovažuji za vhodné neboť všechny dema, která v NLP skupině máme, jsou napsána v jazyce Java. Navíc bych očekával nějaké rozhraní pro dema umožňující sdílet nějakou společnou funkcionalitu. Serverová část dem tak jak je ve frameworku navrhnutá, jsou naprosto oddělené PHP skripty. Neboli při přidávání nového dema je nutné vytvořit jeho serverovou část kompletně znova.

## **Formální úroveň:**

Bakalářské práce se skládá ze 42 stran vlastního textu (67 stran včetně příloh a úvodních stránek), což považuji za více než dostatek. Práce je logicky strukturovaná. Na první pohled působí dokument přehledně a harmonicky. Práce je vysázena v LaTeXu. Při bližším zkoumání jsem však narazil na řadu překlepů a gramatických chyb.

Mám pár připomínek k několika kapitolám. Celá kapitola 2 NLP by zasloužila přepsat. Je zde spousta nepřesností (většinou se jedná o špatný překlad z angličtiny). Některé NLP úlohy jsou špatně zařazeny a nepřesně popsány. Sekce 2.3 Korpusy textů mi přijde naprosto zbytečná. Není třeba vysvětlovat na celou stránku co je korpus. Navíc se jedná o kopii ze stránek na wikipedii. Kapitola 3.1 Google Analytics mi taktéž přijde naprosto zbytečná. Není nutné dopodrobna vysvětlovat, co tato služba umožňuje ani jak jí provoznit. Stačila by zmínka, že něco podobného existuje.

## **Práce s literaturou:**

U reference číslo 4 a 5 chybí zdroj. Autor cituje převážně wikipedii, což nepovažuji za šťastné. Všechny tyto reference by se daly nahradit nějakou publikací. Autor vždy cituje na konci odstavce a to až po ukončení věty tečkou. Citace by měly být na konkrétních místech v textu.

## **Splnění zadání:**

Práce splňuje zadání. Navrhuji hodnocení známkou **dobře** a práci doporučuji k obhajobě.

V Plzni 8. 8. 2016

Ing. Tomáš Brychcín, Ph.D.

