



# Posudek oponenta závěrečné práce

**Oponent práce:** doc. Ing. Pavel Kordík, Ph.D.  
**Student:** Bc. Tomáš Stanovčák  
**Název práce:** Extraktor informací o firmách z webových zdrojů  
**Obor / specializace:** Znalostní inženýrství  
**Vytvořeno dne:** 31. května 2022

## Hodnotící kritéria

### 1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Práce se zaměřuje na extrakci informací o firmách z webu. Slibuje i použití metod strojového učení pro zlepšení identifikace jednotlivých elementů.

### 2. Písemná část práce

84 /100 (B)

Práce obsahuje všechny důležité komponenty. Zlepšit by se dala metodika experimentů, zejména pak vyhodnocení použití metod strojového učení tak, aby bylo dobře patrné, jaké zlepšení přinášejí. Jak přeně vypadá zakódování vstupních atributů, cílová proměnná.

### 3. Nepísemná část, přílohy

82 /100 (B)

Poměrně velká část práce je věnovaná akvizici dat a strukturám pro jejich uložení. Z hlediska ZI jsou pak podstatné experimenty a porovnání různých metod. Zde mi chybí podrobnější analýza chyb, jakých se jednotlivé metody dopouštějí a jaké chyby případně redukuje metody strojového učení.

### 4. Hodnocení výsledků, jejich využitelnost

89 /100 (B)

Je pěkné, že byl projekt uvolněn jako open source. Zvážil bych i uvolnění datasetu se seznamem nápadů, jak se dá využít a vylepšit. Určitě by bylo zajímavé zkusit nějaké předtrénované transformery například na zlepšení rozpoznání adresy.

## **Celkové hodnocení**

88 /100 (B)

Práce je zajímavá, ale ML část není úplně excelentní nejen z důvodu nutnosti ruční anotace dat, ale rezervy jsou i v metodice experimentů a jejich interpretaci.

## **Otázky k obhajobě**

Kdo myslíte bude uživatel výsledků Vaší práce a jakým směrem by se mohla dále rozvíjet?

## **Instrukce**

### **Splnění zadání**

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

### **Písemná část práce**

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

### **Nepísemná část, přílohy**

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

### **Hodnocení výsledků, jejich využitelnost**

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

### **Celkové hodnocení**

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.