



## Posudek oponenta závěrečné práce

**Student:** Lukáš Litvan  
**Oponent práce:** Ing. Kamil Dedecius, Ph.D.  
**Název práce:** Automatické analýzy big data v projektu sdílených vozidel Uniqway  
**Obor:** Znalostní inženýrství

**Datum vytvoření:** 19. 8. 2020

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – následující škálou 1 až 4:</i>
<b>1. Splnění zadání</b>	<b><u>1=zadání splněno,</u> 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</b>
<i>Popis kritéria:</i> Posuďte, 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, posuďte 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.	
<i>Komentář:</i> V širším smyslu považuji zadání bakalářské práce za splněné, student provedl velké množství (offline) analýz dostupných dat. Nedokážu ovšem posoudit, zda tyto analýzy byly “automatické”, jak stojí v názvu zadání.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>2. Písemná část práce</b>	<b>95 (A)</b>
<i>Popis kritéria:</i> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, 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. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, 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.	
<i>Komentář:</i> Bakalářská práce je napsána v českém jazyku o dobré čitelnosti a s pouze minimálním množstvím překlepů a gramatických chyb (“pole, které mohou mít...”, “byly tedy vynechána...” aj.). Z hlediska struktury je rozdělena do 7 číslovaných kapitol, plus dvě nečíslované - úvod a závěr. Grafická úprava práce je pěkná. Citace se zdají být v pořádku.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>3. Nepísemná část, přílohy</b>	<b>100 (A)</b>
<i>Popis kritéria:</i> 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	
<i>Komentář:</i> Bez připomínek.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>4. Hodnocení výsledků, jejich využitelnost</b>	<b>85 (B)</b>
<i>Popis kritéria:</i> 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.	

**Komentář:**

Z hlediska vlastní práce s daty mám dojem, že dochází ke "klasickému" big data fenoménu: velké množství heterogenních nasbíraných dat snižuje jejich perceptibilitu a srozumitelnost. Student sice v úvodu zmiňuje metodiku CRISP-DM, ale nemám pocit, že by se jí výrazněji držel. Provedené analýzy jsou v mnoha ohledech užitečné, ale je těžké se ubránit pocitu jejich ad hoc charakteru. Chybí například vícerozměrné analýzy. Podobně postrádám smysl ve shlukové analýze neseparovaných dat (obr. 6.2, 6.3 aj.), kde bych osobně viděl potenciál pro jiné typy modelů. Oceňuji, že se student pokoušel data vhodně transformovat a reprezentovat, např. v případě tzv. cirkulárních dat (kap. 6.2.2). Stejně tak kap. 7 má do značné míry ad hoc charakter. Uvedené nedostatky bych nicméně nepřičítal k tíži studentovi, neboť daný problém je poměrně značně komplexní.

*Hodnotící kritérium:*

*Způsob hodnocení – nehodnotí se*

**5. Otázky k obhajobě**

*Popis kritéria:*

Uvedte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

**Otázky:**

Pojednejte, jaká data by bylo možné využít pro predikci dostupnosti vozidla (vozidel) v dané uživatelem vybrané zóně.

*Hodnotící kritérium:*

*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*

**6. Celkové hodnocení**

90 (A)

*Popis kritéria:*

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.

**Text hodnocení:**

Práce je dobře napsaná a jediné mé výhrady se týkají míry komplexnosti práce s daty. Práci doporučuji k obhajobě.

Podpis oponenta práce: