



# Hodnocení vedoucího závěrečné práce

**Student:** MUDr. Tomáš Mazel Ph.D.  
**Vedoucí práce:** RNDr. Petr Škoda, CSc.  
**Název práce:** Cloud-Based Platform for Active Learning of Astronomical Spectra  
**Obor:** Webové a softwarové inženýrství

**Datum vytvoření:** 13. 6. 2020

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – následující škálou 1 až 4:</i>
<b>1. Splnění zadání</b>	<b>1=zadání splněno, 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> Zadání práce bylo poměrně značně náročné, neboť představovalo nutnost integrace výsledků vlastní práce studenta (spočívající zejména v interaktivním vizualizačním prostředí pro miliony astronomických spekter) s databází Elasticsearch a cizím kódem pro strojové učení do infrastruktury komplexního cloudového systému VO-CLOUD vyvíjeném od roku 2011 na FIT ČVUT ve spolupráci s Astronomickým ústavem AVČR.  Student se musel dobře orientovat v poměrně různorodém prostředí kombinujícím části v Pythonu, Jupyter notebooku, Java EE a Javascriptu i v instalaci a administraci databázového prostředí Elasticsearch. Ve výsledku vznikl flexibilní systém realizující specifický a poměrně inovátorský workflow pro hluboké aktivní učení astronomických spekter.  Všechny body zadání byly splněny. V zadání navrhovaný způsob realizace systému v podobě Docker kontejneru byl však nakonec zavrhnut po konzultaci s autorem VO Cloudu, protože by znamenal nutnost přepsání většiny současné infrastruktury.	
<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> Práce v dostačujícím celkovém rozsahu 75 stran (z toho vlastní text bez příloh činí 59 stran) je členěná do čtyř na sebe logicky navazujících částí. Kromě standardního infromatického obsahu (funkční a nefunkční požadavky, popis realizace, diskuse implementace a návrhy budoucích vylepšení) obsahuje i stručný úvod do astronomické spektroskopie, včetně její historie, i do aktivního učení. Velmi užitečnou součástí práce je i stručný uživatelský návod k provádění aktivního učení pro klasifikaci spekter.  Práce je psána dobrou angličtinou bez překlepů. Po typografické stránce má standardní LaTeXovou úpravu a je doprovázena kromě faktických grafů a kopií obrazovek i nápaditými ilustracemi procesu aktivního učení či klíčovými kusy kódu a konfiguračních souborů. Rozsáhlá bibliografie obsahuje 65 položek, včetně řady korektně značených elektronických odkazů s DOI, a jasně ukazuje studentovu publikační zkušenost. Velmi mě překvapily korektně uvedené citace na historické práce o spektroskopii včetně Newtonovy práce Opticks či publikací Fraunhofera, Plancka či N.Bohra. Všechny použité SW balíky jsou typu Open Source.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>

### 3. Nepísemná část, přílohy

95 (A)

#### Popis kritéria:

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

#### Komentář:

Nedílnou součástí práce je v souladu se zvyklostmi FIT i CD-ROM se zdrojovými texty a zdrojové kódy SW modulů vytvořených studentem. Pro úplnost je (se souhlasem autora) doplněn i Python kód pro aktivní učení O. Podstávka. Navíc, a to cením nejvíce, je celý kód práce dokumentován na GitHubu v komunitě vývojářů systému VO-CLOUD včetně podrobného návodu k použití. (konkrétně <https://github.com/vodev/vocloud-active-learning>). Výsledek práce, modul pro aktivní učení, je rutinně nasazen na systému VO-CLOUD na serverech Astronomického ústavu v Ondřejově. (<https://vocloud-dev.asu.cas.cz>, položka menu - Active learning)

#### Hodnotící kritérium:

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

### 4. Hodnocení výsledků, jejich využitelnost

95 (A)

#### Popis kritéria:

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:

Práce má velký publikační potenciál inovativního pojetí cloudového systému pro aktivní učení astronomických spekter, popř. jiných vědeckých dat v podobě 1D vektoru (např. DNA analýza, hmotové spektra, časové řady). Funkční demonstrace již proběhla na virtuálním zasedání pracovní skupiny pro astronomické strojové učení v rámci (telekonferenčně pořádaného) workshopu Mezinárodní aliance pro Virtuální observatoř v květnu 2020. Modul bude v nejbližší době použit i pro plnění jednoho z cílů skupiny HPC v rámci grantu RCI na FIT ČVUT.

#### Hodnotící kritérium:

Způsob hodnocení – následující škálou 1 až 5:

### 5. Aktivita a samostatnost studenta

5a:

**1=výborná aktivita,**  
2=velmi dobrá aktivita,  
3=průměrná aktivita,  
4=slabší, ale ještě dostatečná aktivita,  
5=nedostatečná aktivita

5b:

**1=výborná samostatnost,**  
2=velmi dobrá samostatnost,  
3=průměrná samostatnost,  
4=slabší, ale ještě dostatečná samostatnost,  
5=nedostatečná samostatnost

#### Popis kritéria:

V souvislosti s průběhem a výsledkem práce posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven (5a). Posuďte schopnost studenta samostatně tvůrčí práce (5b).

#### Komentář:

Student postupoval při vývoji velmi samostatně, chodil na pravidelné konzultace a často konzultoval e-mailem dílčí výsledky. Důležitých bylo i několik vzájemných schůzek s autory metody aktivního učení O. Podstávkem a autorem infrastruktury VO-CLOUDu J. Kozou, s kterými pak student udržoval pravidelný kontakt při řešení problémů integrace systému. Aktivně si vyhledával literaturu i "best practices" pro SW integraci. Po odevzdání písemné části práce ještě stále pracoval na vylepšení funkcionality i komfortu ovládání.

#### Hodnotící kritérium:

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

### 6. Celkové hodnocení

97 (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:

Z mého pohledu je modul aktivního učení v cloudovém systému velmi kritickou součástí pro explorativní analýzu velkých astronomických dat. Zatím je optimalizován pro práci s miliony spekter projektu LAMOST. Velmi usnadňuje primární přiřazování tříd (labelling) na základě vizuální podoby, ale hlavně je kritický ve fázi aktivního učení, kdy je třeba výsledky předpovídané klasifikátorem vizuálně verifikovat v každé iteraci a to i se znalostí dostupných metadat o kandidátském objektu. Studentova realizace je velmi robustní, snadná na použití a přitom dostatečně flexibilní pro adaptaci na jiné typy dat. GUI je propracované a zohledňuje i potřebu únavného opakování stejných postupů (např. zavádí rychlé klávesové zkratky). Proto práci celkově hodnotím vysoko známkou A

Podpis vedoucího práce: