

Posudek oponenta bakalářské práce

Jméno a příjmení uchazeče/ky : *Květa Brázdilová*

Název práce: *Využití strojového učení pro rozpoznávání a počítání objektů v buněčné biologii*

A. Bodové hodnocení jednotlivých aspektů práce (označte právě jednu z možností)

1. Rozsah BP a její členění	
<input checked="" type="radio"/>	A - přiměřené, odpovídají charakteru BP a významu jednotlivých částí
<input type="radio"/>	B - nevyrovnané, členění není logické n. rozsah jednotlivých částí nekoresponduje s jejich významem
<input type="radio"/>	C - uspokojivé, rozsah některých částí nedostačuje
<input type="radio"/>	N - nedostatečné

2. Odborná správnost	
<input type="radio"/>	A - výborná, bez závažnějších připomínek
<input type="radio"/>	B - velmi dobrá, s ojedinělými drobnými závadami (nejasnost výkladu, chyby ve vzorcích nebo chemických názvech, nedokonalý popis metod nebo výsledků)
<input checked="" type="radio"/>	C - uspokojivá, s četnějšími drobnými závadami
<input type="radio"/>	N - nevyhovující, s hrubými chybami

3. Uvedení použitých literárních a j. zdrojů	
<input type="radio"/>	A - bez připomínek, všechny převzaté údaje s citací zdroje, celkový počet citací odpovídá charakteru práce
<input checked="" type="radio"/>	B - uspokojivé, s občasnými neobratnostmi zejm. v umístění odkazů, nebo s celkově nižším počtem citací
<input type="radio"/>	C - s vážnějšími závadami, např. převažují "nestandardní" odkazy na učebnice, přednášky, webové stránky, nebo se ojediněle vyskytuje opominutí odkazu na zdroj převzatých dat
<input type="radio"/>	N - nevyhovující, velmi málo citací, ev. rysy plagiátu (časté opomíjení odkazu na zdroj převzatých dat, popř. opsání velkých částí textu)

4. Jazyk práce	
<input type="radio"/>	A - výborný, práce je napsána čtivě a srozumitelně, bez závažnějších gramatických n. pravopisných chyb
<input checked="" type="radio"/>	B - velmi dobrý, ojedinělé stylistické neobratnosti, gramatické n. pravopisné chyby
<input type="radio"/>	C - uspokojivý, četnější slohové neobratnosti, gramatické n. pravopisné chyby, ojediněle se vyskytují obtížně srozumitelné n. nejednoznačné formulace
<input type="radio"/>	N - nevyhovující, s četnými hrubými chybami

5. Formální a grafická úroveň práce	
<input type="radio"/>	A - výborná, bez překlepů a chyb ve formátování
<input checked="" type="radio"/>	B - velmi dobrá, ojedinělé chyby formátu citací, překlepy, chybějící zkratky apod.
<input type="radio"/>	C - uspokojivá, s ojedinělými většími (např. vynechání stránky) nebo četnějšími drobnými chybami

5. Formální a grafická úroveň práce

N - nevyhovující, s četnými hrubými chybami

Případný slovní komentář k bodům 1. až 5.:

- Velice často jsou v textu použity termíny, které nejsou nejdříve definovány nebo vysvětleny (nebo jsou vysvětleny až o několik stran po prvním použití)
 - Overfitting (str. 7)
 - Activation function /str. 7)
 - Deep learning (str. 8)
 - Double channels (str. 10)
 - Upsampling (str. 10)
 - Progressive deep learning (str. 12)
 - $F_i^0(\cdot)$ (str.16) – je definováno pouze $F(p|w)$ a není jasné, jaký je vůbec význam “.”, zvláště pak v nepřítomnosti vektoru w
 - ...
- Některé části textu obsahují chyby nebo nepřesnosti
 - Na str. 7 se zmiňuje, že gradient descent používá “estimation of the gradient ... to progress”, ale gradient se explicitně počítá, neodhaduje se.
 - U popisu SVM se tvrdí, že SVM se používá k separování objektů v multidimenzionálním prostoru. Nicméně to platí skoro pro všechny ML metody a nejde tedy o specifikum SVM.
 - Na str. 9 se píše, že získání labelovaných dat je extrémně náročné a někdy skoro nemožné. To je sice pravda, ale hodilo by se zmínit, že toto řeší tzv. transfer learning
- Struktura textu je v některých místech nekonzistentní
 - Sekce 2.4 obsahuje krátkou sekci o generickém frameworku WEKA, který je na úrovni přístupů k detekci objektů. Navíc nechápu, proč je tomuto frameworku věnována vlastní sekce, když ani není použit v implementační části.
 - Nekonzistence v míře informací u sekcí, které jsou na stejné úrovni. Např. u CellProfileru je informace o implementaci původně v MATLABu a pak přepsání do Pythonu. U ostatních frameworků tato informace zcela chybí, tj. výsledek působí nekonzistentně. Totéž platí i pro popis ML metod pro řešení detekce.

Obsahuje-li práce i vlastní výsledky uchazeče/ky (nejsou povinnou součástí práce), pak prosíme o Vaše stanovisko k následujícímu:

Jsou řádně stanoveny a vysvětleny cíle experiment?

Je množství experiment adekvátní k cílům?

Je dokumentace výsledků dostačující?

Jsou výsledky diskutovány a zasazeny do kontextu existující literatury?

Případný další slovní komentář k výsledkům autora:

- V sekci o výzvách ML (1.3.) je zmíněn overfitting a jeho řešení formou cross validace. V popisu výsledků jsem ovšem žádnou zmínku o cross validaci nenašel.
- Některé fotky byly vypuštěny kvůli nízké kvalitě a z textu není jasné, kolik fotek to bylo.
- Do datasetu byl přidáno 18 fotek, jejichž provenience je nejasná.
- Vztah k výsledkům existujících metod nebyl testován (viz. dotazy níže).

B. Obhajoba

Dotazy k obhajobě (povinná část posudku)

- Byla provedena cross validace výsledků?
- Proč nebyl použit libovolná z detekčních metod v existujících frameworkcích, které jsou ostatně i v práci zmíněny?
- Proč nebyl software implementován rovnou jako plugin do některých z frameworků (např. CellProfileru).
- Proč nebyly výstupy srovnány s libovolnou stávající metodou?

Stanovisko k opravě chyb v práci:

opravný lístek/oprava v textu **JE** / **NENÍ** (zakroužkujte) podmínkou přijetí práce

C. Celkový návrh

Práci **doporučuji** k přijetí k dalšímu řízení: **ANO** / **NE**

Navrhovaná celková klasifikace: Velmi dobře

Datum vypracování posudku: 7. 7. 2020

Jméno a příjmení, podpis oponenta (SIS): David Hoksza

Instrukce pro vypracování a odevzdání posudku:

- Pro vypracování posudku bakalářské práce použijte tento formulář.
- Posudek můžete sami vložit do SIS, anebo s předstihem zaslat v elektronické podobě na adresu: marian@natur.cuni.cz, a dále zajistit dodání podepsaného originálu (v 1 výtisku, jako součást protokolu o obhajobě) na sekretariát Katedry buněčné biologie PŘF UK (p. Růžičková), Viničná 7, 128 44 Praha 2. Podepsaný originál posudku musí být dodán před vlastní obhajobou, bez něho nesmí být obhajoba zahájena!
- Student by měl být s posudkem seznámem nejméně tři dny před obhajobou – posudek mu můžete poslat Vy nebo Váš přepošleme.