



Posudek oponenta závěrečné práce

Student: Georgii Korostii
Oponent práce: Ing. Daniel Vašata, Ph.D.
Název práce: Machine learning for earthquake prediction and forecasting
Obor: Znalostní inženýrství

Datum vytvoření: 9. 6. 2019

Hodnotící kritérium:	Způsob hodnocení – následující škálou 1 až 4:
1. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: 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.	
Komentář: První výhradou je nepříliš podrobná analýza existujících metod predikce zemětřesení. Prvním třem bodům zadání tak v zásadě odpovídají pouze strany 5 a 6 v práci s několika útržkovitými zmínkami o nejnovějších pracech v části 1.10.4. Další výhradou je odlišná terminologie v zadání a v práci u hlavních cílů práce. V zadání je uvedeno, že se má student zabývat úlohou "predikce", která je specifikována jako hledání času, lokace a síly zemětřesení, a úlohou "předpovídání", která je specifikována jako určení pravděpodobnosti s jakou nastane zemětřesení. V práci studenty tyto pojmy změnil, ovšem nevysvětlil dostatečně k jakým úlohám se vlastně vážou. Každopádně podle uvedených výsledků spočívá úloha "predikce" v binární klasifikaci (nejspíše tedy odpovídá "předpovídání" ze zadání a úloha "předpovídání" pak spočívá v predikci spojité veličiny (ovšem v práci není zřejmé jaké). Sám student v práci k těmto pojmům uvádí větu "Forecasting in our case is making a prediction about the future based on past and/or present data."	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	65 (D)
Popis kritéria: 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.	
Komentář: Práce obsahuje občasné typografické chyby, jako například vynechávání tečky na koncích vět, které obsahují výčty. Hlavní výtka je ale informační rozsah práce. Z celkových 37 stran (včetně literatury) je cca 15 stran práce věnováno popisu známých metod strojového učení. V jedné z částí tohoto popisu (konkrétně ve druhé půlce části 1.10.4) je bez jakéhokoliv označení vnořen velmi krátký popis nejnovějších výsledků predikce zemětřesení pomocí neuronových sítí. Také je zcela nevhodné tvrzení, že úloha "předpovědi" (tj. dle výše uvedeného nejasného zavedení tohoto pojmu studentem se jedná o predikci spojité veličiny) je neřešitelná standardními algoritmy strojového učení. Potom se ale student tuto úlohu pokouší řešit pomocí neuronových sítí, které mezi standardní modely v dnešní době zcela jistě patří. Co se týče citací použitých zdrojů, vytkl bych především formát citací. Například u zdroje 15 nejsou uvedeny žádné bibliografické údaje, ačkoliv se jedná o časopisecký zdroj. Dále student často cituje blogové příspěvky i u témat, pro které existují knižní nebo časopisecké zdroje. Asi nejvýznamnější výtka je ale uvedena výše v bodu Splnění zadání. Z textu práce není zřejmé, co přesně student predikuje v úloze "předpovídání".	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	85 (B)

Popis kritéria:

Die 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:

Nepísemnou částí práce bylo experimentální porovnání vybraných metod predikce zemětřesení. V příloze práce jsou uvedeny zdrojové kódy, které jsou relativně přehledné a umožňují případnému zájemci všechny provedené experimenty zreplikovat.

Hodnotící kritérium:

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

4. Hodnocení výsledků, jejich využitelnost

68 (D)

Popis kritéria:

Die 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:

Vzhledem k nepříliš obsáhlé rešerši jsou výsledky v teoretické části práce spíše podprůměrné. Obecné popisy metod strojového učení, které tvoří její největší část jsou totiž zcela standardní. Experimenty v praktické části práce již nějaké zajímavé výsledky nabízejí, ovšem vzhledem k špatnému popisu úloh, které byly řešeny, nejsou bez dalšího doplnění využitelné.

Hodnotící kritérium:

Způsob hodnocení – nehodnotí se

5. Otázky k obhajobě

Popis kritéria:

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

Otázky:

Co jste predikoval v úloze "předpovídání"?

Používal jste při trénování také data, která jsou časově vzdálená od jakéhokoliv zemětřesení? Nebo pouze data těsně před zemětřesením, jak jsem pochopil z části 2.1.1.

Hodnotící kritérium:

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

6. Celkové hodnocení

65 (D)

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:

Vzhledem k výše uvedeným nedostatkům navrhuji hodnocení stupněm D.

Podpis oponenta práce: