



Review report of a final thesis

Reviewer: Ing. Pavel Hrabák, Ph.D.
Student: Pavel Chudomel
Thesis title: Quantification of uncertainty of precipitation nowcasting
Branch / specialization: Knowledge Engineering
Created on: 12 June 2023

Evaluation criteria

1. Fulfillment of the assignment

- [1] assignment fulfilled
- ▶ [2] assignment fulfilled with minor objections
- [3] assignment fulfilled with major objections
- [4] assignment not fulfilled

Hlavní body práce jsou splněné až na bod 2) "Understand and describe the used dataset," kterému se student v práci z mého pohledu věnuje nedostatečně. Dataset je zevrubně popsán v kapitole 3.1 (strana 19), chybí rozměr jednotlivých vstupů (velikost obrázků), interpretace hodnot z intervalu [0,1] pro jednotlivé pixely, jak byla data získána apod. Jsou-li porozumění a popis dat explicitně zmíněny v zadání, je to zřejmě pro práci důležité, proto to zde zmiňuji.

2. Main written part

75 /100 (C)

Práce má logickou strukturu, nadstandardně pracuje se zdroji, text vychází a kombinuje velké množství původních zdrojů (nejedná se tak o výtah z jedné přehledové publikace), což velmi oceňuji. Práce je psaná srozumitelnou angličtinou, občasná česká stavba věty a lehce vyprávěcí styl ("Then we have done ...") je odpustitelná. Rešerše používaných metod patří k lepším mezi bakalářskými pracemi, neomezuje se jen na prostý výčet metod, ale snaží se i popsat jejich princip.

Nedostatkem práce je jistá strohost popisu v některých důležitých pasážích. Kapitola 1 popisující architekturu PhyDNet je poměrně těžko čitelná pro někoho, kdo se v ní neorientuje. Oceňuji, že student se snaží popsat princip pomocí rovnic, ale bez dalšího kontextu jsou často těžko čitelné. Velmi by pomohlo, kdyby student například osvětlil, co je vstupem, skrytým stavem, apod. v konkrétním případě, který řeší jeho práce - buď v úvodní kapitole, nebo odkazem v kapitole týkající se popisu dat. Tato výtka souvisí i s (z mého hlediska) nedostatečným popisem datasetu.

V některých klíčových pasážích nejsou dostatečně vysvětlené použité symboly. respektive, nejsou zcela zřejmě navázané na řešený problém. Například:

* Na straně 16 je vzorec pro konfidenční interval, ale není vysvětleno, co jsou jednotlivá pozorování X_i , tedy není jasné, co je v tomto kontextu chápáno jako náhodný výběr.

* Při definici SSIM v kapitole 3.5.1 není zřejmé, co jsou proměnné x , y a jak souvisejí s řešenou úlohou, co jsou v tomto kontextu uvedené hodnoty x_{-i} a y_{-i} .

Celkově mi v práci chybí nějaký popis vytvořené pipeline a kapitola, která by popsala její implementaci, aby bylo zřejmé, které části student převzal, které modifikoval, a které vytvořil.

3. Non-written part, attachments

85 /100 (B)

Praktická část se skládá z implementace dvou metod umožňujících analyzovat neurčitost ve zkoumaném modelu a provedení řady experimentů, které tuto neurčitost analyzují. Praktická část mi byla předvedena a vysvětlena, vše se zdá funkční a implementováno. Vzhledem k absenci nějaké dokumentace, či rámcového vysvětlení je nicméně obtížné posoudit rozsah této části praktické části.

Provedené experimenty jsou v práci popsány a jednotlivé metody porovnány pomocí deklarovaných metrik. Ocenil bych lepší specifikaci datového vzorku, velikosti datových sad, velikosti použitých náhodných souborů v případě MC Dropout metody a podobně.

4. Evaluation of results, publication outputs and awards

80 /100 (B)

Prezentovaná studie je zajímavá a jistě cenná pro analýzu neurčitosti v předpovědích a vhodnosti různých metrik. Jedná se o prozkoumání dosud neprobádané (nebo ne zcela probádané) oblasti. Věřím, že výstupy mají velký potenciál využití. Je ale záhodno specifikovat blíže chybějící informace zmíněné výše.

The overall evaluation

83 /100 (B)

Prezentovaná práce je nadprůměrná, student si osvojil netriviální metody, které využil v následné analýze. Práci doporučuji k obhajobě.

Questions for the defense

1. V sekci 2.5 týkající se kvantilové regrese na stránce 17 píšete: "The rest is straightforward: ...". Můžete toto více rozepsat v souvislosti s vámi řešenou úlohou? Výstupem modelu (pokud jsem to správně pochopil) je 6 matic hodnot mezi $[0, 1]$. Jak v tomto případě chápete medián, kvantil apod.?

2. Vysvětlíte, co berete jako měření X_{-i} při odhadu konfidenčního intervalu v kapitole 2.4.3, jak velký je vzorek (tj. počet měření n) a jaká je interpretace "skutečné střední hodnoty" jejíž CI sestavujete. Jakým způsobem pak pomocí těchto intervalů kvantifikujete neurčitost?

3. Vysvětlíte, jaký je význam proměnných (x, y) v popisu $SSIM(x, y)$ v kapitole 3.5.1, jak souvisí s hodnotou SSIM, kterou uvádíte pro jednotlivé modely (jedno číslo pro jeden model).

Instructions

Fulfillment of the assignment

Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Main written part

Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies?

Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 52/2021, Art. 3.

Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Non-written part, attachments

Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Evaluation of results, publication outputs and awards

Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

The overall evaluation

Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.