

POSUDEK OPONENTA NA DIPLOMOVOU PRÁCI
MARTINA SMOLÍKA
NEURAL MODELLING OF MATHEMATICAL STRUCTURES
AND THEIR EXTENSIONS

Práce popisuje experimenty se strojovým učením. Předmětem učení jsou operace v malých grupách a jejich rozšíření.

Práce má pět částí rozdělených do čtyř kapitol.

- Nejprve jsou vyloženy základy teorie modelů a algebraického rozšíření grup (sekce 1.1.).
- V sekci 1.2 jsou popsány základní myšlenky neuronových sítí.
- Druhá kapitola popisuje způsob aplikace neuronových sítí na rozšíření grup.
- Třetí kapitola předkládá výsledky experimentů.
- Čtvrtá kapitola se nazývá „Comments and discussion“, jedná se o několik poznámek na okraj.

V porovnání s dřívější, neobhájenou verzí práce se výrazně zlepšila sdělnost úvodních pasáží. Čtenář může z práce porozumět základním principům neuronových sítí, způsobům jejich použití a souvisejícím problémům.

Tato úvodní část je převážně kompilační, výjimku představuje autorův vlastní popis algebraického rozšíření grup. Ačkoli tato vlastní pasáž obsahuje několik drobných nejasností, ukazuje, že student problematice rozumí.

Těžiště práce je nicméně zřejmě v experimentální části. U teoretického oboru by si i zde mohl oponent přát nějakou teoretickou reflexi přesahující programátorský výkon a prezentaci získaných dat. Bylo by např. možné diskutovat otázku, jaká je naděje, že optimalizátor najde jednoduchou dvouvrstvou reprezentaci grupy \mathbb{Z}_n z kapitoly 3.1.1. Jsou jí nalezené reprezentace nějak podobné?

Obecně by byla vítána podrobnější interpretace dosažených výsledků (kvůli „human interpretability“ bylo ostatně i podle práce přikročeno ke zjednodušení, viz str. 29 nahoře). Proč je např. „unclear“ jak se chová podivná nalezená jednotka pro \mathbb{Z}_{10} (viz Figure 3.3)? Není možné vypsát celou multiplikační tabulku, kterou software našel? Bylo by případně možné ukázat celou nalezenou neuronovou síť?

Podobné otázky by se jistě vešly do všezahrnujícího oficiálního zadání, chápu ale, že nebylo součástí zadání faktického. Zaměříme-li se tedy na experimenty, je v práci bohužel pro čtenáře stále ještě dost málo informací o tom, v čem tato experimentální práce přesněji spočívala. V příloze se není bohužel příliš snadné orientovat (mimořádně, CD přiložené k oponentské kopii je prázdné), letmý náhled do souborů s koncovkou .py naznačuje spíše rutinní spouštění dostupného optimalizátoru. Nejsem odborník na standardy softwarového inženýrství, nicméně uvítal bych přesnější informaci nejméně v těchto směrech:

- Hardware, na kterém experimenty probíhaly.
- Časová náročnost výpočtu. Práce se omezuje na údaje typu „significantly higher training times“.

- Informace o závislosti výsledků na různých volbách parametrů. V čem se vlastně jednotlivé běhy mezi sebou lišily? Byla používána jediná sada parametrů doporučená softwarem? (Pokud jsem správně pochopil, právě volba parametrů je tvůrčí složkou stavby neuronových sítí.)
- Přesnější výčet provedených pokusů. V kapitole 3.1.2 se např. mluví o tom, co se stalo „most of the time“ a tabulka 3.3 reprodukuje „several different runs“. Bylo jich víc než oněch šest? Nevím také, kde mohu „note that the runs had different lengths“ (str. 35 nahoře).

Pokud jde o úpravu, je přijatelná. Přibylo několik nových překlepů. Z poznámek uvedených v předchozím posudku bohužel nebyly opraveny formule přesahující okraj a stále se používá „as such“ namísto „such that“.

Navzdory uvedeným nedostatkům doporučuji práci přijmout jako diplomovou.

Praha 3. září 2019

Štěpán Holub