



MATEMATICKO-FYZIKÁLNÍ
FAKULTA
Univerzita Karlova

Posudek bakalářské práce

Posudek oponenta

| | |
|-------------------|--|
| Autor: | Michal Sekerka |
| Název práce: | Kolmogorovovská složitost a Shannonova informace |
| Školitel: | prof. Mgr. Michal Koucký, Ph.D. |
| Oponent: | doc. Mgr. Robert Šámal, Ph.D. |
| Kontaktní e-mail: | samal@iuuk.mff.cuni.cz |
| Rok odevzdání: | 2020 |

Předložená bakalářská práce by rozsahem i obsahem obstála i jako práce diplomová. Pan Sekerka představuje dvě zásadní definice, které popisují ze dvou různých pohledů složitost řetězce. V závěrečné kapitole ukazuje, jak spolu tyto dva pojmy souvisejí.

Všechny zaváděné pojmy jsou pečlivě motivovány: počínaje rozborem Platónovy jeskyně, přes srovnání frekventistického a bayesovského přístupu k pravděpodobnosti až po dramatické historické souvislosti spojené se zaváděním Shannonovy informace, Turingova stroje, atd.

V kapitole 2 autor odvozuje asymptotic equipartition property a vysvětluje její důsledky pro kompresi. V kapitole 3 napřed autor vybuduje základy teorie Turingových strojů (včetně aplikace na halting problem). Na těchto (pracně) vybudovaných základech může konečně formálně přesně vysvětlit pojem Kolmogorovská složitost – a v další kapitole předvést Zvonkinovu větu o vztahu Shannonovy informace a Kolmogorovské složitosti.

Z odborného hlediska jsem nenašel mnoho, co by šlo práci vytknout. Definice zápisu „ $\# a_i$ v a “ není příliš srozumitelná (str. 10). Na str. 13 je matoucí pojem *funkcionální* analýza v souvislosti s vlastnostmi logaritmu. Na str. 20 je možná zavádějící název pro větu 2.2.1 (jako Sluckého věta bývá většinou označováno jemnější tvrzení pro konvergenci v distribuci). Na str. 36 je ve vzorci \min místo $\arg\min$. V lemmatu 2.2.4 je ve vzorci X_1 místo X_i , a pár podobných drobných překlepů. Na závažné matematické chyby jsem nenarazil.

Poněkud slabší byla jazyková kontrola a typografická. Časté jsou překlepy, ale i prohřešky proti gramatice (shoda podmětu s přísudkem!, kostrbaté větné konstrukce). Ve vzorcích je často místo a_1, \dots, a_k psáno a_1, \dots, a_k , místo $\langle a \rangle$ je $< a >$, a podobně. V seznamu literatury jsou názvy článků s mnoha chybami (patrně nesprávné použití bibtexu). Jedná se o drobnosti, které neohrožují pochopitelnost textu – na druhou stranu odvádějí pozornost od hlavního sdělení. Kdyby měl být text použitý jako např. rozšiřující studijní materiál, bylo by vhodné je opravit. Musím však zdůraznit, že jak rozsah, tak i obsah práce je silně nadstandardní, autor ví co píše a ví proč to píše. Jak je uvedeno v úvodu, nejedná se o původní výsledky (pochopitelně, tyto výsledky jsou už několik dekad známy). Jde o kompilaci několika přehledových článků – ale o kompilaci nadstandardně zdařilou a velice „dotáženou“, tj. bez mezer, které by byly ponechány laskavému čtenáři k dokončení, jak bývá častým zvykem.

Doporučení: S ohledem na výše uvedené navrhuji práci přijmout jako bakalářskou práci.

V Burnaby dne 22. června 2020

Robert Šámal