

## Posudek oponenta bakalářské práce

**Student:** Stodůlka Martin  
**Téma:** Aproximace hlubokých neuronových sítí (id 21819)  
**Oponent:** Mrázek Vojtěch, Ing., Ph.D., UPSY FIT VUT

- 1. Náročnost zadání** **značně obtížné zadání**  
Cílem zadání bylo implementovat nepřesné (aproximované) výpočty do frameworku Caffe. Implementačně se jedná o **značně náročnou práci**, protože framework je velmi rozsáhlý a jeho rozšíření vyžaduje znalost principu celé řady jeho komponent.
- 2. Splnění požadavků zadání** **zadání téměř splněno**  
Student přesně splnil body 1 - 4 a 6. Navržené a implementované rozšíření frameworku však bylo vyhodnoceno pouze na jedné úloze - klasifikace ručně psaných čísel v datasetu MNIST. Nebyl prezentováno chování na žádné další síti či datasetu. Je však nutno dodat, že **funkčnost aproximací** v inferenci neuronových sítí **byla dostatečně prezentována**.
- 3. Rozsah technické zprávy** **splňuje pouze minimální požadavky**  
Rozsah předložené práce splňuje pouze minimální požadavky kladené na bakalářskou práci. Zejména kapitola 5 shrnující implementaci a experimentální část, je v poměru k ostatním kapitolám stručnější (viz bod 7).
- 4. Prezentační úroveň předložené práce** **55 b. (E)**  
Práce nepůsobí jednotně. Úvod je velmi stručný, nicméně kapitoly 2, 3 a 4 popisují všechny prerekvizity na dobré úrovni. I v těchto částech však nacházíme velké množství obrázků a výpisů, které nejsou pro čtenáře důležité (např. výpis terminálu při trénování sítě ve frameworku Caffe či definice vrstev v modelu sítě *prototxt*). Ovšem klíčová část práce - vlastní rozšíření o aproximační konvoluční vrstvu - je však popsána na méně než 3 stranách. Výhradu mám k prezentaci výsledků - výsledky jsou vždy prezentovány pouze v grafu (navíc často logaritmickém) a v textu chybí tabulka s přesnými hodnotami výsledků. Není pak možné si udělat přesnou představu o kvalitě výsledků.
- 5. Formální úprava technické zprávy** **65 b. (D)**  
Text je psaný téměř bez chyb, objevují se typografické chyby a nejednotnost v psaní velkých písmen. V práci se vyskytuje několik převzatých obrázků, které však jsou v nižším rozlišení.
- 6. Práce s literaturou** **75 b. (C)**  
Práce využívá základní informační zdroje (z toho několik vědeckých publikací), které jsou řádně ocitovány.
- 7. Realizační výstup** **55 b. (E)**  
Cílem práce je navrhnout rozšíření stávajícího nástroje pro práci s hlubokými neuronovými sítěmi. Autor se úspěšně seznámil s velmi obsáhlým nástrojem Caffe a vytvořil pro něj nový model aproximované vrstvy. Pro efektivní simulaci aproximovaného počítání využil texturovou paměť. Aplikace je funkční, slabinou je však experimentální vyhodnocení. U práce tohoto charakteru bych hlavně očekával vyhodnocení rychlosti výpočtu, které však v práci chybí. Autor patřičně neukázal, že zvolil správný postup (přestože se o správný postup může jednat!).
- 8. Využitelnost výsledků**  
Práce obsahuje implementaci aproximované vrstvy neuronové sítě, které již byly ve vědecké komunitě představeny, nicméně významné rozšíření je přenesení ne příliš efektivního výpočtu z CPU na výkonnější GPU. Po vyhodnocení výkonnosti této implementace se může jednat o zajímavý výsledek pro další výzkum.
- 9. Otázky k obhajobě**
  1. Z jakých paměťových přístupů pro ukládání vyhledávacích tabulek jste volil a proč jste se rozhodl pro použití *texturové paměti*?
  2. Jaké kroky je potřeba provést pro aproximování jiné neuronové sítě?
- 10. Souhrnné hodnocení** **58 b. dostatečně (E)**  
Autor se úspěšně seznámil s nástrojem pro práci s neuronovými sítěmi Caffe a implementoval aproximační konvoluční vrstvy. Implementace byla optimalizována tak, aby byla inference neuronové sítě co nejrychlejší. Funkčnost byla prezentována na jednom benchmarkovém modelu (MNIST na síti LeNet-5). Vzhledem k slabšímu vyhodnocení efektivity a přesnosti klasifikace, testování pouze jednoho modelu neuronové sítě, menšímu rozsahu a nižší prezentační úrovni navrhuji hodnocení 58 - E.

V Brně dne: 28. května 2019

.....  
podpis