

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Ondřej Hübsch

Název práce Redukce počtu parametrů v konvolučních neuronových sítích

Rok odevzdání 2021

Studijní program Informatika **Studijní obor** Teoretická informatika

Autor posudku Mgr. Martin Pilát, Ph.D. **Role** vedoucí

Pracoviště KTIML MFF UK

Text posudku:

Diplomová práce Ondřeje Hübsche se zabývá důležitým problémem zmenšování počtu parametrů v hlubokých neuronových sítích. Cílem bylo najít architektury, které jsou menší (s menším počtem parametrů), ale zároveň stále dosahují dobrých výsledků. Tohoto cíle se podařilo dosáhnout, student navrhl algoritmus, který nahrazuje existující moduly v neuronové síti za jejich menší varianty a následně používá ještě techniky pro prořezání vah v neuronové síti pro další snížení počtu parametrů.

Celá práce je rozdělena do pěti kapitol (včetně úvodu a závěru). V prvních třech kapitolách student napřed popisuje motivaci a cíle práce, vysvětluje hlavní koncepty z oblastí neuronových sítí, evolučních algoritmů a zpětnovazebního učení a nakonec popisuje existující přístupy pro hledání architektur neuronových sítí a pro jejich prořezávání. Tato část práce je dobře napsána a obsahuje všechny informace důležité pro zbytek práce. Vybraná literatura je aktuální a relevantní k řešení daného problému.

Na základě těchto informací student navrhuje hned několik technik pro řešení tohoto problému založených na evolučních algoritmech a zpětnovazebním učení. Navržené algoritmy, na rozdíl od těch existujících, nehledají architekturu sítě od začátku, ale začínají s existující architekturou, kterou se snaží zmenšit nahrazením některých modulů menšími moduly. To výrazně zjednodušuje problém hledání architektury a umožňuje se soustředit na její zmenšování.

Algoritmy jsou následně vyhodnoceny a porovnány v kapitole páté. Student provádí experimenty s celou řadou neuronových sítí a s několika variantami navržených algoritmů. Ukazuje se, že výsledné neuronové sítě splňují požadavky – mají stejnou, nebo podobnou přesnost klasifikace a zároveň mají mnohem méně parametrů než původní neuronové sítě. Při prořezávání potom jejich přesnost klesá pomaleji (v závislosti na počtu parametrů) než tomu je u původních sítí. Takové výsledky jsou důležitým přínosem v oblasti hledání architektur neuronových sítí a zmenšování počtu parametrů v neuronových sítích. Experimenty jsou provedeny i vyhodnoceny velmi pečlivě

a dobře ukazují silné i slabé stránky navržených přístupů.

Celá práce je relativně pěkně napsána, obsahuje všechny potřebné informace. Výsledky experimentů ukazují, že navržené metody fungují dobře a lze je úspěšně použít pro redukci počtu parametrů v hlubokých neuronových sítích. Student ukázal, že je schopný kvalitní samostatné práce a podařilo se mu splnit vytyčené cíle.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 9. června 2021

Podpis: