



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

doc. RNDr. Václav Kučera, Ph.D.
Katedra numerické matematiky MFF UK
Sokolovská 83, 186 75 Praha 8
Tel.: 951 553 362
e-mail: kucera@karlin.mff.cuni.cz

Posudek vedoucího bakalářské práce

Práce: Optimalizační algoritmy pro trénování neuronových sítí
Řešitel: Jiří Szotkowski

Práce se zabývá popisem a implementací neuronových sítí, se zaměřením na optimalizaci parametrů sítě (tzv. trénování). Toto téma je vysoce aktuální, s tím, že volba vhodných optimalizačních algoritmů pro trénování sítě hraje klíčovou roli pro dosažení dobrých výsledků v, pokud možno, co nejkratším čase. V současné době existuje celá řada v praxi používaných optimalizačních algoritmů. Ty jsou většinou jednoduchou či složitější modifikací metody spádových směrů. Cílem práce bylo porovnat několik známých a doporučených algoritmů na akademických příkladech i praktických úlohách trénování neuronových sítí.

Jiří Szotkowski odvedl vynikající práci. Aby plně pochopil danou problematiku, rozhodl se sám o sobě implementovat neuronové sítě od začátku, tedy bez použití jakýchkoli knihoven či již existujících implementací. Dále implementoval metodu na výpočet derivací sítě vzhledem k jejím parametrům, tzv. *backpropagation*, a šest optimalizačních algoritmů v základní verzi, plus jejich stochastické varianty. Toto vše představuje obrovský kus práce, který je o to úctyhodnější, že byl proveden v rámci bakalářské práce. Všechny tyto optimalizační algoritmy, včetně jejich variant, byly porovnány na akademických příkladech z klasické optimalizační literatury a na praktických problémech trénování neuronových sítí (rozpoznání písma a klasifikace ropných vrtů).

Dále musím ocenit Kapitulu 2, kde jsou neuronové sítě formulovány matematicky. V dostupné literatuře se zavedení neuronových sítí často neobejde bez většího či menšího množství „máčení rukama“ nebo nejasného či značně přetěžovaného značení s tím, že představa získaná z obrázku neuronové sítě je poměrně dostačující. Pak se naráží na to, že například derivace vzhledem k parametrům (*backpropagation*), což je vlastně derivace mnohonásobně složené nelineární funkce vzhledem k mnoha různým parametrům, už není tak intuitivně zřejmá a jednoznačně zapsatelná vágním značením. Pan Szotkowski se velmi dobře zhostil úkolu neuronové sítě formalizovat velmi rigorózně, kdy pak není prostor pro nejednoznačnost.

Při porovnání optimalizačních algoritmů bylo dosaženo některých překvapivých zjištění, které do jisté míry odporují doporučením v literatuře. Často například vítězila základní *stochastic gradient descent method* nad sofistikovanějšími metodami. Dále tato metoda fungovala překvapivě dobře (či dokonce nejlépe) i při nejmenší velikosti tzv. *minibatche* (velikosti 1). Silně zde ovšem závisí na konkrétním příkladu.

Jiří Szotkowski odvedl obrovský kus samostatné práce, zejména implementační, která vyžadovala mnoho času, úsilí a studia různých nesouvisejících oblastí. Práci považuji ze velice zdařilou. O to větší je v tomto kontextu zásluha Jiřího Szotkowského, že jde o práci bakalářskou a tedy vše zmíněné bylo dosaženo v podstatně kratším čase a s nutností většího samostudia, než by bylo u práce diplomové. **Práci doporučuji uznat jako bakalářskou.**

V Praze dne 3. 6. 2024



doc. RNDr. Václav Kučera, Ph.D.