

# Posudek oponenta bakalářské práce

**Název práce:** Nástroj pro statistické zkoumání velkých toků dat

**Autor:** Alexandra Vecherskaya

**Vedoucí práce:** doc. Ing. Miroslav Virius, CSc.

**Oponent:** doc. Ing. Jaromír Kukal, Ph.D.

Předložená práce se zabývá statistickou analýzou toků dat s využitím umělých neuronových sítí a entropie. V teoretické části nalezneme stručnou informaci o neuronových sítích s důrazem na rekurentní sítě, LSTM, BLSTM a softmax včetně jejich učení. Práce má vzornou matematickou sazbu. K učení ANN jsou navrženy gradientní metody: jednoduchá, adadelta a backpropagation. Zde bych chtěl upozornit na fakt, že gradientní metody jsou schopné nalézt pouze lokální minimum pokutové funkce. V rámci obhajoby se chci zeptat, zda to je aplikační nevýhoda a jak jí čelit. Dále by mne zajímalo, zda by danou úlohu bylo možné řešit jinými jednoduššími technikami a kterými. V navazující části práce nalezneme popis předzpracování dat s využitím entropie, oken, shlukování a DTW.

Počínaje šestou kapitolou lze již hovořit o praktické části práce. Návrh a implementace v Pythonu jsou popsány velmi stručně. Navazující uživatelská příručka popisuje metody jednotlivých tříd. Výsledky zpracování dat jsou přehledně uvedeny ve formě grafů a tabulek. Chybí mi zde alespoň stručná interpretace výsledků z hlediska dané aplikace. Co například znázorňuje obr. 8.6 s nepopsanými osami?

Práce podle mého názoru splnila zadání, je napsána přehledným způsobem a obsahuje původní výsledky. Je škoda, že práce je napsána velmi úsporným způsobem, ze kterého není zřejmé, zda výsledkem práce jsou pouze grafy, hodnoty uvedené v tabulkách a návrh systému, nebo i původní zdrojové texty programů. Práci navrhuji hodnotit známkou B – velmi dobře.