

## Posudek vedoucího závěrečné práce

Název práce: **Distributed Algorithms for Decision Forest Training in the Network Traffic Classification Task**

Autor: **Radek Starosta**

Typ práce: Bakalářská práce

Katedra kybernetiky, Fakulta Elektrotechnická, ČVUT v Praze

Vedoucí: **Ing. Jan Brabec**

Pracoviště vedoucího: Cisco Systems, Inc.

Student se v rámci své práce zabýval distribuovaným tréninkem náhodných lesů na platformě Apache Spark. Cílem práce bylo zrychlení algoritmu, který je k dispozici v populární knihovně Spark MLlib. Práce byla vypracována ve spolupráci se společností Cisco a evaluace výsledků byla provedena na velkých datech pocházejících ze síťového provozu.

Zadání práce hodnotím jako velmi náročné. Student se musel nejen do hloubky seznámit s náhodnými lesy, ale i s komplikovanými distribuovanými algoritmy pro jejich trénink, a výpočetní platformou Apache Spark. Hlavním cílem práce bylo implementování lokálního trénování podstromů do knihovny MLlib. Splnění tohoto úkolu samo o sobě by stačilo k velmi kvalitní a nadstandardně obtížné bakalářské práci. Studentovi se tento úkol podařilo splnit velmi rychle, a proto zbylo ještě dost času na další vylepšování algoritmu. Výsledkem je implementace, která je výrazně stabilnější a v benchmarcích na datech ze síťového provozu více než stokrát rychlejší, než původní implementace. Studentovu implementaci již používáme v Ciscu pro trénink mnohem větších rozhodovacích lesů mnohem rychleji než jsme byli schopni dřív a budeme se snažit aby nahradila stávající implementaci přímo v knihovně Spark MLlib. Výsledky bychom také rádi publikovali na vědecké konferenci.

Student pracoval svědomitě a na konzultace docházel pravidelně. Velmi oceňuji jeho samostatnost, iniciativu a práci nad rámec zadaných úkolů.

Práce samotná je psaná dobrou angličtinou a jak svým rozsahem přes 40 stran tak obsahem je to velmi nadstandardní bakalářská práce. Práce se zdroji je korektní. Celkem je obsaženo 37 citací. Znovu oceňuji studentovu samostatnost ve shánění nových zdrojů a práci s nimi.

Práce vysoce převyšuje požadavky na bakalářskou práci a její výsledky mají signifikantní dopad a využití. Práci doporučuji k obhajobě, hodnotím jednoznačně známkou: **A (výborně)** a doporučuji zvážit udělení ceny děkana FEL za excelentní bakalářskou práci.

Datum: 3.6.2018

Podpis: