

**I. IDENTIFIKAČNÍ ÚDAJE**

<b>Název práce:</b>	Distribuované algoritmy pro trénink rozhodovacích lesů v úloze klasifikace síťového provozu
<b>Jméno autora:</b>	Radek Starosta
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra kybernetiky
<b>Oponent práce:</b>	Ing. Vojtěch Cvrček
<b>Pracoviště oponenta práce:</b>	Katedra kybernetiky

**II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ**

<b>Zadání</b>	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání vyžaduje dobré porozumění implementačním detailům metod určených pro distribuované trénování rozhodovacích stromů. Zároveň vyžaduje schopnost metody modifikovat, pracovat s velkými daty a navrhopvat/testovat paralelní algoritmy. Zadání považuji za náročnější.	

<b>Splnění zadání</b>	splněno s menšími výhrada
<i>Posudte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Zadání považuji za splněné. Drobné výhrady mám k experimentální části. Autor tvrdí: „This performance improvement allows us to train random forests on larger datasets, which can significantly improve classification predictive performance.“. Postrádám standardní rozdělení dat na trénovací, validační a testovací podmnožiny a demonstrace vlivu velikosti datové sady na klasifikátor.	

<b>Zvolený postup řešení</b>	vynikající
<i>Posudte, zda student zvolil správný postup nebo metody řešení.</i>	
Zvolená metoda řešení vychází z dobrého porozumění problému. Přínos jednotlivých částí metody je doložen v experimentální části. Postup řešení považuji za vynikající.	

<b>Odborná úroveň</b>	A - výborně
<i>Posudte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Práce je výborně odborné úrovně. Vhodně kombinuje znalosti získané studiem a praxí k řešení komplexní praktické úlohy.	

<b>Formální a jazyková úroveň, rozsah práce</b>	A - výborně
<i>Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku.</i>	
Nejsem spokojen s odkazováním a číslováním vzorců na straně 22 odvozujícím paměťové nároky lokálního tréninku. Číslování chybí také u prediktoru délky lokálního tréninku na straně 32, zápis funkce je navíc příliš neformální. Samotný text je napsán srozumitelně a na výborné úrovni. Práce je dostatečně rozsáhlá. Celkově hodnotím formální úroveň jako výbornou.	

<b>Výběr zdrojů, korektnost citací</b>	A - výborně
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posudte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od</i>	

vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Autor prokázal schopnost pracovat s rozličnými zdroji. Převzatá tvrzení jsou doložena citacemi a snadno je lze odlišit od vlastní práce. Vzhledem k technickému charakteru práce považuji za přijatelné citace repositářů a dokumentací. Nedostatkem je chybějící datum přístupu. Práci se zdroji hodnotím jako výbornou.

#### **Další komentáře a hodnocení**

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

### **III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE**

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Na práci zejména oceňuji pečlivé experimentální ověření relevance jednotlivých částí metody, srozumitelnost textu, praktické zaměření a dosažení výrazného zrychlení a stabilizace tréninku náhodných lesů.

V trénovací množině je velký nepoměr mezi pozitivními/negativními vzorky. Jaký vliv má tento poměr na výsledný klasifikátor?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm .

Datum: 4.6.2018

Podpis: