



# Posudek oponenta závěrečné práce

Oponent práce:	Ing. Tomáš Čejka, Ph.D.
Student:	Karel Mudruňka
Název práce:	Klasifikace typu zařízení ze síťového provozu ISP sítě pomocí shlukovacích metod
Obor / specializace:	Umělá inteligence 2021
Vytvořeno dne:	10. června 2024

## Hodnotící kritéria

### 1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Cílem závěrečné práce bylo analyzovat dostupné shlukovací algoritmy použitelné pro zpracování časových řad síťového provozu. V rámci práce proběhla velká sada experimentů s využitím Metacentra a byla vyvinuta prototypová implementace. Zadání bylo splněno.

### 2. Písemná část práce 60/100 (D)

Práce obsahuje řadu typografických nedostatků, reference na obrázky a výpisy jsou místy matoucí - uvedeno je pouze číslo. Text obsahuje drobné překlepy a nepřesné formulace. Přesto je text práce stále ještě dobře čitelný a struktura je logicky členěna.

### 3. Nepísemná část, přílohy 100/100 (A)

Výsledkem je implementace nástrojů pro analýzu časových řad pomocí shlukovacích algoritmů a rozpoznání několika typů zařízení na síti. Výhodou vyvinutého výsledku je možnost využití rozsáhlých výpočetních zdrojů Metacentra.

### 4. Hodnocení výsledků, jejich využitelnost 90/100 (A)

V rámci závěrečné práce vznikla implementace softwarového nástroje v jazyce Python pro analýzu síťového provozu pomocí časových řad a shlukovacích algoritmů. Výsledky vypadají slibně a mají publikační potenciál. Využití příznaku TTL bude s velkou pravděpodobností málo přenositelné, dosažené výsledky však ukazují, že lze

natrénované modely úspěšně použít k dlouhodobému rozpoznávání vybraných typů zařízení na základě časových řad z jejich síťového provozu. Vyvinuté technologie a vytvořené pracovní procesy trénování a vyhodnocování proto vypadají využitelně v praxi.

## Celkové hodnocení

89 /100 (B)

Celkově jsou výsledky práce kvalitní, avšak textová část práce obsahuje řadu nedostatků a nabízí tak široký prostor ke zlepšení. Implementované výsledky vypadají zajímavě s publikačním potenciálem.

## Otázky k obhajobě

1. Na str. 20 je uvedeno: "Experiment také ukázal, že přesnost predikce se výrazně liší pro různé komunikační protokoly, model byl například schopen lépe predikovat komunikaci pomocí TCP než komunikaci přes UDP (User Datagram Protocol)." Podařilo se odhalit, co způsobuje rozdíly ve schopnosti predikovat komunikaci?
2. Jaké kroky by bylo potřeba udělat pro produkční nasazení vyvinutých výsledků v reálné síti?
3. Jak by bylo možné u vyvinutého softwarového řešení změnit sadu vstupních příznaků (feature vector) nebo sadu podporovaných výsledků klasifikace?

## **Instrukce**

### **Splnění zadání**

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

### **Písemná část práce**

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

### **Nepísemná část, přílohy**

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

### **Hodnocení výsledků, jejich využitelnost**

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

### **Celkové hodnocení**

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.