

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Student: Bc. David Bulawa

Oponent: doc. RNDr. Martin Kotyrba,
Ph.D.

Studijní program: **Inženýrská informatika**
Studijní obor/Specializace: **Softwarové inženýrství**
Akademický rok: **2021/2022**

Téma diplomové práce: **Datamining pro IoT integrační platformu**

Hodnocení práce:

Diplomová práce je svým rozsahem poměrně obsáhlá, rozkládá se na 134 stranách a dělí se celkem na 8 hlavních kapitol.

Tato práce se věnuje aplikaci metod z oblasti data miningu na data pocházející z integrační platformy pro IoT. V teoretické části je nastíněna problematika big dat a získávání znalostí. Praktická část reflektuje poznatky z teoretické části převedením do praxe a ukázce na konkrétním příkladu. Zpracování dat předchází popis aktuální struktury uchovávaných historických hodnot z IoT zařízení v platformě od společnosti MyMight a návrh nového vhodného systému, který reflektuje rychlý přístup k datům v jakémkoli čase. Výsledkem je aplikace data mining algoritmů na zmíněné historické hodnoty. Pro testování byla použita data historických hodnot ze vzorového domu v rádech několika měsíců reálného provozu několika IoT technologií v různých případech užití. Součástí práce je i interpretace výsledků dosažených vybranými algoritmy a zhodnocení využitelnosti získaných informací. Celkové pojetí práce jasně naznačuje autorovu orientaci v problematice a schopnost zvládat koncepci vědeckého pojetí práce.

Cílem diplomové práce bylo navrhnout vhodné prostředí pro uchování hodnot ze senzorů a akčních členů tak, aby bylo možné k nim pohodlně a rychle přistupovat a v neposlední řadě aplikovat data mining na uchovávaná data s cílem identifikovat anomálie a opakující se vzory. Ze struktury a obsahu práce mohu zcela bez diskuze konstatovat, že cíle práce byly naplněny.

Formální pojetí práce je na velmi vysoké úrovni. Práce má všechny náležitosti, které si od kvalitní diplomové práce představuji. Pro konkrétní představu uvádím (kvalitní obrázky, odkazování, číslování vzorců, logická návaznost kapitol, přehlednost práce). Pouze ve třech případech došlo ke špatné provázanosti odkazů se zdrojem, jinak opravdu velmi kvalitní zpracování.

Otázky k obhajobě:

1. Uveďte rozdíl vlivu mezi použitím algoritmu DBSCAN s časovou dimenzí a bez ní?
2. Na základě čeho jste volil právě vybrané algoritmy (model ARIMA, Apriori, DBSCAN..)?
3. Jak se lišily výsledky při implementaci modelu ARIMA v případě počátečního odhadu parametrů p a q ?
4. Uvažujete o nasazení výsledku Vaší práce do reálného provozu?

V práci je použito celkem 37 zdrojů použité literatury a všechny jsou v práci řádně citovány.

Diplomová práce interpretuje výsledky při testování těchto úloh nad reálnými daty ze vzorového domu. Práce taktéž porovnává všechny implementované varianty vybraných algoritmů. Algoritmus Apriori dokázal správně odhalit časté kombinace událostí, které uživatel během dne ve svém domě vykonával. Predikce časových řad pomocí statistického modelu ARIMA se ukázala jako relativně přesná pro jednotky až desítky budoucích hodnot a tím využitelná pro praktické nasazení, přesto bych pro predikci volil jiné techniky z oblasti UI, které se v porovnání s metodou ARIMA ukázaly jako úspěšnější v rámci nasazení na různorodá data.

Práce je svým zaměřením velmi aktuální a využitelná pro praxi a popisuje celý proces tvorby systému s podporou data miningu od návrhu architektury až po implementaci jednotlivých úloh.

Práce doporučuji k obhajobě a hodnotím jako

Celkové hodnocení práce:

Známku uvede oponent dle svého uvážení dle klasifikační stupnice ECTS:

A – výborně, B – velmi dobře, C – dobře, D – uspokojivě, E – dostatečně, F – nedostatečně.

Stupeň F znamená též „nedoporučuji práci k obhajobě“.

Předloženou diplomovou práci doporučuji k obhajobě a navrhuji hodnocení

A - výborně.

V případě hodnocení stupněm „F – nedostatečně“ uveďte do připomínek a slovního vyjádření hlavní nedostatky práce a důvody tohoto hodnocení.

Datum 1. 6. 2022

Podpis oponenta diplomové práce