

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Data clustering in Hilbert spaces - Shlukování dat v Hilbertových prostorech
Jméno autora:	Diana Varšiková
Typ práce:	bakalářská práce
Fakulta:	Fakulta jaderná a fyzikálně inženýrská (FJFI)
Katedra:	Katedra matematiky
Oponent práce:	Quang Van Tran
Pracoviště oponenta práce:	Katedra softwarového inženýrství

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Náročnost zadání odpovídá úrovni tohoto typu kvalifikační práce. Má potenciální využití v praxi	

Splnění zadání	splněno
<i>Posuďte, 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>	
Předložená práce splňuje všech 5 bodů taxativně vyjmenovaných v zadání.	

Zvolený postup řešení	vhodný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Postup řešení je uspokojivý, tedy vhodný.	

Odborná úroveň	výborná
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Celková úroveň práce je velmi dobrá, literární rešerše by se mohly mírně rozšířit. Získání dat je standardní.	

Formální a jazyková úroveň	průměrná
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Celkově formální úroveň práce je kvalitní. Práce je napsaná dobrou angličtinou. Přimluvil bych se za to, aby pseudokódy byly očísleny s následující odvolávkou v textu. Některé popisky v grafech jsou nečitelné. Tabulka 5.2. nemá odkaz v textu.	

Výběr zdrojů, korektnost citací	výborné
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i>	
Výběr zdrojů byl adekvátní, citace byla v pořádku, citační etika nebyla porušena. V seznamu použité literatury by mohla být lepší editace u některých zdrojů.	

Další komentáře a hodnocení
<i>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.</i>
Jelikož se jedná o práci s experimentálním zaměřením, kde důraz je kladen na úspěšnou implementaci programových nástrojů na řešení daného problému, tento cíl je zcela naplněn. Autorka vytvořila fungující

knihovnu funkcí v Pythonu, kterou potom použila klasifikaci shlukovou analýzou. Ohledně publikačního potenciálu, od práce tohoto typu se obvykle neočekává a tato práce není výjimkou. Možná bych tam přivítal, kdyby autorka přidala do hodnocení dosažených výsledků vysvětlení, proč výsledky jsou takové, jaké jsou. A protože použila pro svou práci standardní datasety, mohla by provést srovnání svých výsledky s výsledky získanými jinými metodami, aniž by musela tyto metody vyzkoušet sama.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

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

Už z jednotlivých dílčích hodnocení je patrné, že odevzdaná bakalářská práce splnila zadání, tudíž požadavky na práci tohoto typu. Pokud byly v ní suboptimální zjevy, jsou to problémy spíše lokálního charakteru. Proto ji hodnotím známkou uvedenou níže. Otázky

- V Tabulce 5.2. jsou uvedeny tři metriky zvané acc. Jak se počítají a k čemu slouží?
- Na konci strany 30 jsou uvedeny výsledky klasifikace podle přesnosti pro známý tzv. iris dataset. V čem vidíte příčinu případného zlepšení nebo zhoršení výsledku klasifikace z jednotlivých přístupů?
- Pokud srovnáte svoje nejlepší výsledky s výsledky dosaženými jinými metodami pro použité datasety, kde stojíte?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře**.



Datum: 15.8.2023

Podpis: Quang Van Tran