

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Student: Martin Masař

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

Studijní program: Informační technologie

Specializace: Softwarové inženýrství

Akademický rok: 2024/2025

Téma diplomové práce: Clustering populace evolučních algoritmů

Hodnocení práce:

Diplomová práce se zabývá aktuální a relevantní problematikou analýzy populační dynamiky evolučních algoritmů pomocí shlukovacích metod. Řešený úkol implementovat a porovnat různé clusteringové metody na datech z evolučních algoritmů za účelem analýzy jejich konvergence, stagnace a úniku z lokálních optim je komplexní a pro diplomovou práci adekvátně náročný. Zvolení implementace evolučních a shlukovacích algoritmů od základů v jazyce Python bez využití specializovaných knihoven zvyšuje obtížnost úkolu, ale zároveň poskytuje hlubší vhled do fungování algoritmů. Práce pokrývá jak teoretické základy, tak praktickou implementaci a analýzu, včetně definice a sledování relevantních metrik vývoje shluků napříč generacemi. Z tohoto pohledu stanovené cíle splněny.

Autor pojal úkol systematicky. V teoretické části poskytuje solidní přehled o shlukování a evolučních algoritmech. Praktická část popisuje implementaci dvou evolučních algoritmů (DE, GA) a následně implementaci a aplikaci tří shlukovacích algoritmů (K-Means, agregační, OPTICS) pro analýzu populace.

Téma je zpracováno do hloubky, od teoretického základu přes implementaci až po detailní analýzu a interpretaci výsledků. Hlavním přínosem diplomanta je rozsáhlá praktická implementace evolučních a shlukovacích algoritmů od základů a následná komplexní analýza dynamiky populace prostřednictvím shlukových metrik. Autor prokazuje schopnost aplikovat teoretické znalosti k řešení praktického problému a interpretovat získaná data v kontextu chování evolučních algoritmů. Analýza odhaluje zajímavé rozdíly v chování DE a GA z pohledu shlukové struktury – tendenci DE rychle konvergovat k dominantnímu shluku versus udržování vyšší diverzity a delší fáze explorační u GA. Autor v závěru správně konstatuje, že shlukování slouží jako cenný nástroj k hlubšímu pochopení evoluční dynamiky

Co se týká nedostatků, uvádím, že struktura práce odpovídá standardu, citace jsou uváděny formou [i] a seznam literatury vychází z normy ČSN ISO 690:2022. Obrázky a tabulky jsou číslovány, opatřeny popisky a odkazováno je na ně v textu. V práci se však vyskytují následující chyby:

V prohlášení autora je v jedné větě zmíněno "diplomové práce" a v jiné "bakalářské práce". Jedná se zřejmě o přehlédnutí a zkopírování textu. V textu se opakovaně vyskytují chybné interní odkazy na kapitoly/subkapitoly. Například odkaz na popis eukleidovské a manhattanské vzdálenosti v 1.1.1 směřuje na "subkapitole 3.1.2", místo 1.1.2. Odkaz na K-Means v 1.2 směřuje na "kapitole 3.4.1", místo 1.3.1. Odkaz na schéma rand/1 v 3.1.2 (Mutace) směřuje na "kapitole 4.3.2", místo 2.3.2. Dále se v textu práce vyskytují chyby v odkazech na obrázky v sekci analýzy. Například v sekci 5.2.1 (Genetický algoritmus, Sphere) se text odkazuje na "obrázku 7", který ve skutečnosti zobrazuje Vývoj hodnoty WCSS při shlukování metodou K-Means (DE, Sphere), zatímco v daném kontextu by měl být odkaz na relevantní obrázek pro GA Sphere K-Means. Podobně je v téže sekci na straně 91 odkazováno na "obrázku 7", jehož popis nepatří odkazu). Formule pro Simple Matching Coefficient je uvedena chybně – je zde zopakována formule pro Jaccardův index, ačkoli text vysvětluje SMC správně. V ukázkách kódu pro GA se zdá, že jedinci jsou reprezentováni jako seznamy nebo pole (např. individual[i], population[best], parent1[1:]), zatímco ukázka kódu pro DE křížení (která je v textu

uvedena v sekci DE) používá přístup s klíči (`parent["genes"][i]`), což odpovídá reprezentaci jako slovník použité v ukázce `generate_individual` pro DE. Tato nekonzistence v reprezentaci jedinců napříč ukázkami kódu může být matoucí, ačkoli je možné, že v plné implementaci je konverze mezi formáty ošetřena.

5. Dotazy k obhajobě:

Metoda mapování shluků pro GA je založena na určení dominantního rodiče pomocí eukleidovské vzdálenosti od potomka a následném přiřazení ID historického shluku. Jak robustní je tento přístup pro sledování shluků v komplexních nebo vícemodálních prostorech, kde geny určující blízkost mohou být rozptýleny nebo kde rodiče mohou být prostorově vzdáleni, ale vedou k podobně kvalitní oblasti řešení? Zkoušel jste alternativní metody mapování pro GA?

Předložená diplomová práce Bc. Martina Masaře je kvalitní prací s jasně definovanými cíli a systematickým přístupem k jejich řešení. Autor prokázal dobré teoretické znalosti v oblasti evolučních algoritmů a shlukové analýzy, a především významné praktické dovednosti při implementaci algoritmů od základů a provádění komplexní analýzy populační dynamiky. Práce obsahuje sice několik formálních a technických chyb (chybné interní odkazy, chyba ve vzorci, chybné odkazy na obrázky v jedné sekci, drobná nekonzistence v kódu, nicméně tyto nedostatky nezpochybují hlavní přínos a kvalitu provedené analýzy.

Na základě předložené práce a s ohledem na výše uvedené body doporučuji práci k obhajobě a navrhuji klasifikaci:

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

C - dobře.

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

Datum

6. 6. 2025

Podpis oponenta diplomové práce