

# Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

**Autor práce** Marek Nagy  
**Název práce** Methods of genetic programming for classification  
**Rok odevzdání** 2024  
**Studijní program** Informatika      **Studijní obor** Umělá inteligence

**Autor posudku** Roman Neruda      **Role** Vedoucí  
**Pracoviště** ÚI AV ČR, v.v.i.

## Text posudku:

V předkládané práci „Methods of genetic programming for classification“ se Marek Nagy zabývá metodami genetického programování, které se aplikují na problém klasifikace z oblasti strojového učení s učitelem. Genetické programování je specifickou oblastí evolučních algoritmů, která používá populační prohledávání k nalezení modelu reprezentující program řešící úlohu charakterizovanou trénovacími daty s příklady vstupů a požadovaných výstupů. Výhodou metod genetického programování je jejich jednoduchost a vysvětlující schopnosti např. v porovnání s metodami hlubokých neuronových sítí. Cílem práce tedy bylo aplikovat metody stromového a karteziánského genetického programování na problém klasifikace, vytvořit jejich přizpůsobenou implementaci a ověřit efektivitu algoritmů na benchmarkových klasifikačních úlohách.

Struktura práce je následující. V úvodu autor předkládá motivaci práce a seznamuje s cíli a strukturou textu. V první kapitole se zavádí základní definice a pojmy z oblasti evolučních algoritmů a klasifikačních problémů. Je představen obecný evoluční algoritmus a poté dvě varianty genetického programování – stromové a karteziánské – které autor dále používá ve své práci.

Druhá a třetí kapitola jsou jádrem autorovy samostatné práce a popisuje se v nich návrh a implementace algoritmů genetického programování. Nejprve je popsán karteziánský přístup, který reprezentuje algoritmus jako rovinné nakreslení grafu v mřížce s uzly odpovídajícími výpočetním jednotkám, poté se popíše stromová varianta, která reprezentuje algoritmus jako strom se vstupními jednotkami v listech a výstupem ve svém kořeni. Součástí popisu je detailní zavedení většího množství operátorů selekce, mutace a křížení specifických pro daný přístup a také pro řešený problém klasifikace.

Čtvrtá kapitola je praktická a ukazuje, jak daný systém s výpočetním jádrem realizovaným v jazyce C# instalovat a spouštět. Pátá kapitola je věnována experimentům, kde se prakticky demonstruje výkon klasifikátorů založených na stromovém a karteziánském genetickém programování na několika malých a středně velkých úlohách (Iris, Wine, Breast cancer a MNIST). Závěr práce shrnuje dosažené výsledky a ukazuje další možné směry práce. Součástí práce je i rozsáhlá softwarová implantace uvedených algoritmů.

Za hlavní klady práce považují implementaci daných algoritmů v prostředí C# společně se zevrubným popisem implementace a použití. Další silnou stránkou práce je návrh variant algoritmů, které jsou specializované na problém klasifikace, student musel řešit volbu vhodného zakódování, operátorů, návrh fitness. V případě stromových genetických programů implementoval navíc i ansámbl jedno-třídivých klasifikátorů tak, aby bylo možno řešit problém klasifikace do více tříd.

Slabší stránkou práce je – z mého pohledu – experimentální část, která by si zasloužila více experimentů demonstrujících relativní výkonnost jednotlivých variant algoritmů. Ve svém návrhu i implementaci autor realizoval poměrně velké množství genetických operátorů, ale v experimentech využívá jen jejich část. Ty experimenty, které jsou v práci, jsou provedené a statisticky vyhodnocené dobře a často přinášejí zajímavý vhled do efektivity a explainability jednotlivých metod. Zhodnocení výhody či nevýhod některých operátorů by si bývalo zasloužilo podrobnější rozbor opřený o další experimenty.

Předkládaná práce přináší dobré výsledky v oblasti klasifikace pomocí genetického programování společně s praktickou implementací navržených algoritmů, proto práci doporučuji k obhajobě.

**Práci doporučuji k obhajobě.**

**Práci nenavrhuji na zvláštní ocenění.**

*Pokud práci navrhuje na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).*

**Datum** 31. 5. 2024

**Podpis**