

# POSUDEK BAKALÁŘSKÉ PRÁCE

**Autor práce:** Vladimír Kunc  
**Název práce:** Zvyšování rozmanitosti slabých klasifikátorů ve složených klasifikátorech pomocí příznakových grafů  
**Vedoucí:** doc. Ing. Jiří Kléma, Ph.D., katedra počítačů, FEL ČVUT

Složené klasifikátory využívají většího počtu dílčích, tzv. slabých, klasifikátorů ke zvýšení klasifikační přesnosti. V nejjednodušší implementaci jsou slabé klasifikátory stejného typu a o konečné klasifikaci hlasují s uniformními vahami. Základními kritérii úspěšnosti složeného modelu jsou přesnosti slabých klasifikátorů a jejich rozmanitost. Rozmanitosti dosahujeme například vzorkováním nebo randomizací vstupních dat. Kritéria přesnosti a rozmanitosti jsou obvykle protichůdná, při přímém tlaku na přesnost slabých klasifikátorů klesá jejich rozmanitost a naopak. Práce pana Kunce se soustředí na využití apriorní znalosti ve formě příznakových grafů ke zvýšení rozmanitosti slabých klasifikátorů při zachování či minimálním poklesu jejich přesnosti. Příznakové grafy chápeme jako vyjádření závislostí mezi příznaky, které je obtížně pozorovatelné přímo z dat. Aplikační doménou je genomika, kde je počet vzorků nízký vzhledem k počtu příznaků a závislosti nelze snadno detekovat na základě korelace. Příznakový graf lze naopak sestavit na základě dříve popsanych regulačních vztahů.

Studentovi se cíle práce podařilo beze zbytku naplnit. Seznámil se s problematikou složených klasifikátorů a jejich rozmanitosti. Pro experimentální účely zobecnil existující metodu Network-constrained forest (NCF) a aplikoval ji na dodaná data. Experimenty dokumentují vliv parametru délky procházky příznakovou sítí na rozmanitost slabých klasifikátorů a přesnost výsledného modelu. Práce nabízí také srovnání odlišných typů algoritmů učení a ukazuje, že jiné způsoby učení mohou v konkrétních doménách překonat původní rozhodovací stromy. Práce zkoumá i dopady na srozumitelnost výsledného modelu. Zadání nebylo implementačně složité. Za obtížné naopak pokládám navržení a systematické uspořádání experimentů včetně volby nutného počtu jejich opakování pro odstranění náhodné složky. Bakalář zvládl i krok nejtěžší, vyhodnocení na první pohled nejednoznačných výsledků a jejich zasazení do kontextu současného teoretického poznání. Musím pochválit i výsledný text bakalářské práce. Autor si osvojil zahuštěný vědecký styl s velkým počtem citací. Práce odkazuje 208 zdrojů, mohu potvrdit, že nejde jen o formální odkazy a student s literaturou aktivně pracoval.

Spolupráce s panem Kuncem byla bezproblémová. Jde o talentovaného a cílevědomého studenta, který ve standardním čase vypracoval velmi nadprůměrnou bakalářskou práci s publikačním potenciálem. Formálně konstatuji, že byly splněny všechny body zadání, pan Kunc prokázal schopnost samostatné práce. Bakalářskou práci hodnotím známkou

**A — výborně.**

V Praze, dne 5. června 2015

doc. Ing. Jiří Kléma, Ph.D., vedoucí