

# Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

**Autor práce** Petr Martišek

**Název práce** Feature Evaluation for Scalable Cover Song Identification Using Machine Learning

**Rok odevzdání** 2019

**Studijní program** Informatika **Studijní obor** Softwarové systémy

**Autor posudku** Jan Hajič jr. **Role** oponent

**Pracoviště** Ústav formální a aplikované lingvistiky

## Text posudku:

Cílem předložené práce je porovnat rysy (features) používané v literatuře pro identifikaci tzv. cover verzí, tj. rozhodnout, zda jsou dvě nahrávky různé verze téže písně. Cover verzi typicky nahrává jiný interpret, s jinou instrumentací, s přesunem do jiného žánru (např. metalový cover původně folkové písně), atd. – jedná se tedy o těžkou úlohu. Pro její řešení byla navržena řada deskriptorů jak na úrovni celých písní, tak formou detailnějších časových řad. Jádro práce spočívá v tom, že autor tyto deskriptory (rysy, features) sesbíral z literatury, porovnal, vybral nejslibnější, a z nich pak sestavil systém pro identifikaci cover verzí, který srovnává se state-of-the-art systémy. Souhlasím s diplomantovým závěrem, že „this is the first rigorous comparison of cover song identification features on this scale and depth” (s. 93, odst. 3).

Selhání plánované dvoustupňové metody (předvýběr pomocí „rychlých” rysů a reranking pomocí „přesných” není vina studenta – jedná se o problém spojený featurami MFCC s datasetem Million Song Dataset, jak je ostatně v práci detailně diskutováno.

Nejsilnější stránkou diplomové práce je robustní práce s poměrně velkou částí metodologie strojového učení. Kromě natrénování jednotlivých modelů diplomant předvedl značnou šíři metod výběru rysů (feature selection), volba hyperparametrů, výběr modelu (model selection), robustní vyhodnocování, atd. Práce je v tomto směru nadprůměrná. Je třeba také vyzdvihnout existenci kapitoly 3, která se samostatně věnuje právě metodám výběru rysů. Práce je obecně velice důkladná: každý pokus je podložen tabulkou, evaluační metody nejsou vybírány slepě, nýbrž dobře diskutovány, atd. Škoda že kvůli potížím s informacemi dostupnými v datasetu nebylo možné z množství provedených pokusů vyvodit zajímavější závěry! Práce je výborně strukturovaná, pouze v závěru bych možná uvítal rekapitulaci vybraných rysů, a krom diskuse toho, proč nefungovaly exact features (která je zcela na místě!), nějakou diskusi toho, proč mohly různé scalable features fungovat lépe či hůře než ostatní – přijde mi, že z množství vynaložené práce by bývalo mohlo být

možné dospět aspoň v této otázce k silnějším závěrům.

Nejslabší stránka práce je asi porovnání se state-of-the-art metodami v sekci 9.2.1, nicméně slabiny tohoto porovnání diplomant opět podrobně diskutuje. Co mi mírně chybí je ovšem analýza chyb: například na základě metadat nalezených kandidátů cover verzí zkusit najít určité hudební jevy, kvůli kterým příslušná feature nebo jejich kombinace může selhávat (avšak vzhledem k tomu, že tuto úlohu a dataset SHSD nijak podrobně neznám, je možné, že taková analýza nedává smysl).

Po formální stránce je práce v pořádku, chyby jsou drobné (např. číslování sekcí v kapitole 7). Citace a bibliografické údaje jsou použity v souladu s oborovými zvyklostmi. Práce je psána prakticky bezchybnou precizní angličtinou.

Dotazy k obhajobě:

1. V závěru je řečeno, že ani kombinování takto pečlivě vybraných rysů nevedlo k celkovému zlepšení. Co by dle Vašeho názoru k takovému zlepšení vést mohlo? Lze říct, co jsou obecné slabiny této sady rysů? V závěru kap. 7 zmiňujete, že je škoda, že Trailovy rysy založené právě na MFCC nelze použít, přestože jsou výrazně jiné než ostatní – navrhl byste pro identifikaci cover verzí nějaké vlastní, typově podobně odlišné rysy?

**Práci doporučuji k obhajobě.**

**Práci nenavrhuji na zvláštní ocenění.**

V Praze dne 3. 6. 2019

Podpis: