

# Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

**Autor práce** Marek Židek

**Název práce** Controlled Music Generation with Deep Learning

**Rok odevzdání** 2020

**Studijní program** Informatika **Studijní obor** Umělá inteligence

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

**Pracoviště** Ústav formální a aplikované lingvistiky

## Text posudku:

Cílem předložené práce je vytvořit metodu generování hudebních kompozic, která produkuje koherentní hudební skladby většího rozsahu prakticky bez omezení na instrumentaci a hudební texturu a umožňuje uživatelům značně ovlivnit výslednou podobu skladby.

Hlavní myšlenka práce spočívá v explicitním zacházení s *hudební formou*, definovanou abstraktně jako rozvržení podobných a nepodobných hudebních úseků. Práce volí jako minimální hudební úsek jeden takt, což je poměrně jemná a zároveň rigidní granularita, a trénuje model pro generování podobného taktu a naopak taktů kompatibilních v čase nebo harmonicky, ale nepodobných. Trénovací příklady získává automaticky z dat pomocí klastrování. Explicitně se modely strojového učení podmiňují také parametry jako hustota, modalita (durová či mollová), znějící nástroj, a další. V práci je toto ovládání realizováno převzetím příslušných nastavitelných parametrů ze skutečných skladeb, nicméně zvolená metoda umožňuje tyto parametry specifikovat také „od stolu“ (ačkoliv příslušné rozhraní součástí implementace není – práce je primárně výzkumná, proof-of-concept, nikoliv softwarová). Strojové učení je základní předpoklad novoty, překvapivosti vygenerovaného hudebního materiálu.

Práce dle mého úsudku své cíle splnila, minimálně v míře očekávatelné od diplomové práce na otevřené výzkumné téma. Přináší metodu automatického generování libovolně dlouhých skladeb pro libovolná obsazení (od klavíru po smyčcový orchestr), s natrénovanými modely, která – alespoň z mého hlediska informovaného profesionální hudební praxí, vč. kompozice – vytváří zhruba jednou z dvaceti až padesáti běhů komplexní skladbu, která je bez dalších úprav zajímavá na poslech. Toto považuji za dostatečnou demonstraci toho, že postup kombinující hudebně-teoretické východisko se strojovým učením byl zvolen dobře, a že i další návrhová rozhodnutí (např. granularita formy na úrovni taktů, volba podobnostní funkce pro takty a její adaptace pro vícehlasou hudbu, výběr clusterovacího algoritmu pro extrakci formy, reprezentace taktů pro učení jako sekvence hudebních rozhodnutí místo matice) nebyla chybná. Jednotlivé komponenty by jistě

bylo možné dále zlepšovat, nicméně to by práce byla rozsahem a náročností spíše disertační než diplomová. Nejsem si vědom podobného hybridního přístupu ani srovnatelných výsledků v literatuře. Řešitel se pokusil navíc o automatickou evaluaci vygenerovaných skladeb – adaptuje z úlohy generování obrázků metodu tzv. Fréchet Inception Distance (FID), která ještě v kontextu generování skladeb použita nebyla, ačkoliv příslušné pokusy mohly být robustnější (viz s. 86).

Slabší stránkou práce je software – ačkoliv téma a výsledky k vytvoření aplikace vybízí, náročnost a množství experimentálních aspektů práce prostor na vývoj softwaru nedaly, a k práci je tím pádem přiložena pouze minimální sada programů nutných k replikování výsledků práce (nicméně zadání softwarovou stránku práce nespecifikuje). Nejpodstatnější problém samotných vygenerovaných skladeb je dle mého stále vysoká míra disonantnosti a místy exotické, nepravidelné rytmické struktury.

Po formální stránce je práce v pořádku, ačkoliv ne dokonalá. Angličtina je místy těžkopádná, ale zůstává srozumitelná, pouze sekce 9.3 o FID nese známky urychleného dokončování. Citace a bibliografické údaje jsou použity v souladu s oborovými zvyklostmi, bibliografie je bohatá a aktuální.

Dotazy k obhajobě:

Výstupy učení jsou disonantnější než trénovací data. Co byste zkusil, aby generátory variant a dalších či doprovodných taktů produkovaly “čistší” harmonické struktury?

Obdobně: jak byste zkusil změnit trénování, případně předzpracování dat, aby byla průzračnější výsledná rytmická struktura?

**Práci doporučuji k obhajobě.**

**Práci nenavrhuji na zvláštní ocenění.**

V Praze dne 26. 1. 2020

Podpis: