

## Hodnocení vedoucího diplomové práce

### Bc. Zdeněk Janeček Generátor podkresové hudby

Diplomová práce Bc. Zdeňka Janečka vychází z oblasti volnočasového zájmu diplomanta o hudbu a zabývá se problémem automatického generování podkresové hudby (tzv. „elevator music“) prostředky umělé inteligence. Takto vytvořenou hudbu by pak bylo lze reprodukovat v různých veřejných prostorech (čekárny, výtahy, obchodní domy, atp.) ke zpestření pobytu v nich a nebyla by zatížena copyrightem.

Autor již má předchozí zkušenosti s obdobným tématem, neboť v rámci své bakalářské práce zpracovával problematiku detekce fundamentální frekvence hudebních záznamů.

V diplomantově přístupu k řešení problému bylo možné sledovat střídání období horečné aktivity s obdobími horečné nečinnosti. V aktivních údobích navštěvoval vedoucího práce poměrně pravidelně a hojně diskutoval zejména o výběru technik řešení zadaného problému. Vyhledal také celou řadu relevantních zdrojů informací: vědeckých článků, knihoven, příkladů zdrojového kódu, atp.

Na konzultace s vedoucím docházel diplomant nepravidelně, podle toho, jak se mu dařilo při vývoji softwaru. Většinou byl velmi dobře připraven a měl dobrý přehled o technikách, algoritmech a paradigmatech užívaných v předmětné oblasti. Naopak jsem byl občas překvapen projevenou neznalostí v oblasti hudební teorie, přestože dobrá orientace v hudební nauce byla jedním z požadavků kladených na vhodného kandidáta pro toto zadání a diplomant ji v prvotních rozhovorech přesvědčivě deklaroval.

Práce je původní, při návrhu řešení autor vycházel z celé řady vědeckých publikací na dané téma z poslední doby; je zřejmé, že se podstatným způsobem inspiroval zejména pracemi skupiny Yoshuy Bengia z Université de Montréal a předložený výsledek je z částečně implementací techniky popsané v článku Nicolase Boulanger-Lewandowskiho, Yoshuy Bengia a Pascala Vincenta z roku 2012, doplněný o modul rekurentní sítě LSTM, která slouží jako model časového vývoje generované skladby (což ale neuvádím jako výtku, jde o legitimní postup – v rámci diplomové práce zcela jistě). Nicméně by bylo jistě vhodné v textu práce zřetelněji explicitně uvést, co je vlastním nápadem autora a co bylo implementováno podle nastudovaných vědeckých prací.

Bibliografie je nadprůměrně rozsáhlá (35 titulů), vysoce odborná a zcela relevantní. Citace v textu i bibliografie na konci práce jsou provedené v souladu s požadavky.

Implementační část předloženého díla je plně funkční a řeší zadaný úkol. Dosažené výsledky – ve smyslu zadání práce – lze těžko zhodnotit (viz níže), protože nebyl proveden žádný systematický test. Několik zkusmo vygenerovaných skladeb mi připadá dost disharmonických, a tedy pro použití k veřejné reprodukci nevhodných.

Použité algoritmy pracují správně a jsou stabilní. Vyvinutý software funguje dobře, ovšem jeho provozování je omezeno na operační systémy UNIXového typu.

K vývoji byl použit jazyk Lua ve spojení s řadou knihoven (Torch, MIDI.lua, ap.). Zdrojový kód programového řešení je zapsán ve shodě se zvyklostmi, je velmi dobře čitelný a přehledný. Je však velmi málo komentovaný (prakticky vůbec) a obsahuje celou řadu magických čísel – různých ad hoc konstant (což by snad šlo přehlédnout, neboť se jedná vlastně o výzkumný prototyp a nikoliv software v produkční kvalitě).

Za nejrozporuplnější skutečnosti týkající se implementační části díla považuji, že autor jen využil algoritmů a technik, které poskytují užité knihovny. Nepokusil se vložit do díla vlastní invenci a např. normalizovat výstup LSTM-RBM generátoru nějakým vlastním algoritmem, což by jistě přispělo k vyšší přijatelnosti vygenerované hudby.

Textová část díla je přiměřeně rozsáhlá – má 55 stran (bez příloh), je napsána dobrou technickou češtinou a vysázena v L<sup>A</sup>T<sub>E</sub>Xu. Autor dost často rezignuje na hledání vhodného českého překladu anglických termínů a uvádí je v anglické podobě, ovšem podléhající české flexi. Na několika málo místech je autorovo vyjadřování poněkud neobratné (str. 30: „... protože dělá méně počítání, ...“).

Grafická úroveň je velmi dobrá a celkově působí dokument harmonickým dojmem. Text práce je srozumitelný, byť místy v teoretické části značně komplikovaný. Některé postupy nejsou vysvětleny zcela pochopitelně (např. odd. 3.1.2);

domnívám se, že je to proto, že autor sám měl problémy jim v příslušné literatuře porozumět dostatečně hluboce (což ale nemusí být nutně jeho chyba).

Gramatické chyby či překlepy se v práci prakticky nevyskytují. Text je doplněn velkým množstvím vzorců (některé nejsou dle mého názoru nutné), obrázků, grafů a výpisů zdrojového kódu, které ho žádoucím způsobem obohacují a přispívají k pochopení vykládané problematiky.

Práce je dobře logicky strukturovaná a poměr jednotlivých částí textu je vyvážený.

Předložené dílo bohužel nelze využít zamýšleným způsobem, neboť přijatelnost automaticky vygenerované hudby je nízká. Nicméně autorem prozkoumané a shromážděné poznatky, stejně jako vyvinutý software, mohou sloužit jako dobrý základ pro budoucí výzkum v předemné oblasti a potenciál pro další rozvoj a zlepšování přijatelnosti výsledků je veliký.

Body zadání byly splněny bez závad, s výjimkou bodu 3 („Experimentálně ověřte přijatelnost vygenerované hudby“). Tento bod předpokládá provedení percepčního testu na skupině posluchačů a v tomto ohledu byl také diplomant instruován. Přesto se svévolně rozhodl experimentální ověření přijatelnosti nahradit objektivním vyjádřením kvality natrénování modelu prostřednictvím několika standardních metrik. Takovéto vyjádření kvality je sice zřejmě z ryze vědeckého hlediska přijatelné a bod 3 zadání tak snad lze z určitého úhlu pohledu a s výhradami považovat za splněný, ovšem „poslouchatelnost“ vygenerované hudby zcela jistě nijak nevystihuje.

Po teoretické stránce se jedná o velmi náročné téma a je tedy třeba ocenit snahu a úsilí autora do této problematiky proniknout, byť výsledný produkt není bezvadný. Autor rozhodně prokázal výzkumnickou odvahu a schopnost řešit inženýrským způsobem komplexní problémy.

Práci i přes uvedené výhrady, výtky k dosaženým výsledkům a nesouhlas s provedením bodu 3 zadání **doporučuji k obhajobě** a hodnotím klasifikačním stupněm

„dobře“.



Ing. Kamil Ekštejn, Ph.D.  
KIV FAV ZČU

V Plzni dne 26. května 2016

**SOUHLASÍ  
S ORIGINÁLEM**

*Ok!*  
**Západočeská univerzita v Plzni**  
Fakulta aplikovaných věd  
katedra informatiky a výpočetní techniky

①