

Posudek oponenta diplomové práce

Autor/autorka práce: **Kateřina Štollová**

Název práce: **Rozpoznávání hudebních žánrů**

Obsah práce

Cílem práce je na návrh a implementace algoritmu, který bude klasifikovat digitální akustický signál se záznamem hudby v MP3 do jedné z několika definovaných tříd (žánrů). Jedná se o téma relativně náročné, se kterým si však studentka velice dobře poradila. Předložená diplomová práce je po obsahové stránce přehledná. Za jediný trochu nelogický krok považuji zařazení kapitoly 2.4 s názvem *Strojové učení* mezi kapitolu 2.3 *Metody zpracování digitálního signálu* a kapitolu 2.5 *Extrakce příznaků*. Rozsah práce splňuje požadovaný standard pro diplomovou práci (cca 53 stran textu bez příloh a seznamu použité literatury).

Kvalita řešení a dosažených výsledků


Po grafické stránce má práce vcelku dobrou úroveň, je psaná relativně přehledně a srozumitelně, vytvořená aplikace pro klasifikaci žánrů je funkční a dobře zdokumentovaná příloze práce. Bylo by ale vhodné se v práci zmínit o době zpracování jednotlivých hudebních souborů, popř. by spuštěná aplikace mohla nějakým způsobem informovat uživatele, že „žije“. Po spuštění aplikace se sice vypíše, že běží klasifikace, ale delší dobu (cca 5 min na počítači s procesorem i7 3.4 GHz) se nic neděje a uživatel může mít pocit, že aplikace „zamrzla“. Rozbor výsledků práce je na dobré úrovni, dosažené výsledky jsou prezentovány formou tabulek a grafů a jsou diskutovány v závěru práce.

Formální úroveň

Stylisticky je práce na dobré úrovni. Počet chyb a překlepů je akceptovatelný. Příložené DVD má logicky uspořádaný obsah, struktura disku je popsána jak v příloze práce, tak v souboru README na DVD. Předložená práce je logicky rozdělena do třech částí. V první části (kap. 1-2) diplomantka popisuje teorii potřebnou k úspěšnému vyřešení úlohy, tj. provádí rozbor, jak klasifikovat hudební žánry na základě frekvenčních parametrů vstupního signálu, dále se zabývá popisem metod extrakce příznaků a algoritmy pro klasifikaci vstupního signálu. Druhá část práce (kap. 3) je věnována vlastní realizaci. Diplomantka zde popisuje jednotlivé datasety, použité metody výběru příznaků a klasifikace, dále popisuje technologie použité k řešení úlohy a problémy, na které při řešení práce narazila. Ve třetí části práce (kap 4-5) jsou prezentovány a diskutovány dosažené výsledky. Součástí textu diplomová práce jsou 2 přílohy. V příloze A diplomantka prezentuje výsledky referenčního testu, kdy uživatelé hodnotili žánr skladeb, které byly použity pro testování. Příloha B obsahuje uživatelskou příručku aplikace.

Práce s literaturou

Studentka pracovala jak s tištěnou literaturou (knihy, články), tak s internetovými zdroji. Použitou literaturu považuji za relevantní.


**SOUHLASÍ
S ORIGINÁLEM**

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

Dotazy a připomínky k práci :

1. Na str. 13 používáte termín trojúhelníkové filtry bez dalšího vysvětlení, popř. odkazu na literaturu. Vysvětlete, o co se jedná.
2. Na str. 14 je v souvislosti s extrakcí MFCC příznaků zmínka o jakýchsi koeficientech 2-13. O jaké koeficienty se jedná? Vysvětlete.
3. Na str. 16 je v podkapitole *Umělá neuronová síť* uvedeno: „Umělá neuronová síť je adaptivní systém, který během učení mění svojí strukturu.“ Co máte v tomto případě na mysli změnou struktury? Vysvětlete.
4. U popisu neuronové sítě na str. 17 by mohl být uveden obrázek modelu jednoduchého neuronu pro lepší pochopení dalšího textu.
5. Na str. 17 v předposledním odstavci tvrdíte: „Obecně platí, že síť s více neurony ve skryté vrstvě bude mít lepší výsledky...“. Co máte na mysli „lepšími“ výsledky? Obecně s tímto tvrzením nelze souhlasit. Pokud rozsáhlou síť natrénuji malou trénovací množinou, tak pravděpodobně dosáhnu nízké hodnoty chyby v rámci trénovací množiny, ale taková síť bude pravděpodobně špatně zobecňovat tj. špatně reagovat na vstupy, které neviděla při trénování. Zkuste vysvětlit proč. Existuje nějaký vztah mezi velikostí trénovací množiny a velikostí sítě?
6. Dále zde tvrdíte (poslední odstavec na stránce), že: „Na základě trénovacího vzorku na vstupu se aktivují všechny neurony v síti.“ To také nemusí být pravda, na čem to závisí?
7. Co je $x[n]$ a co je $z[n]$ ve vztazích 2.16 a 2.17 na str. 21. V popisu jsem nenašel, o co se jedná.
8. Vysvětlete co je License Creative Common na str. 24. V práci by mohl být uveden alespoň odkaz, kde lze znění licence dohledat.
9. Co je na svislé ose grafu na obr. 4.4 str. 51?

Splnění zadání


Zadání bylo splněno bez výhrad, diplomantka prokázala, že je schopná řešit i relativně složité inženýrské problémy. Implementovaná aplikace je plně funkční a dobře zdokumentovaná v příloze. Práci proto doporučuji k obhajobě a vzhledem k uvedeným připomínkám navrhuji klasifikovat známkou

„velmi dobře“.

V Plzni 3. 6. 2016

Ing. Pavel Mautner, Ph.D.




SOUHLASÍ
S ORIGINÁLEM

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky
①