dr hab. Piotr Artiemjew, prof. UWM

Katedra Metod Matematycznych Informatyki,

Wydział Matematyki i Informatyki,

Uniwersytet Warmińsko-Mazurski w Olsztynie

Olsztyn, 20 września 2020

Recenzja rozprawy doktorskiej mgr Tomasza Grębskiego pt. „Podobieństwo melodii i jego rozpoznawanie przez komputer przy użyciu narzędzi automatycznego uczenia się"

Niniejsza recenzja została przygotowana na zlenienie Rady Instytutu Filozofii Katolickiego Uniwersytetu Lubelskiego Jana Pawła II (KUL) w odpowiedzi na prośbę Pana prof. dr hab. Piotra Kulickiego, Dyrektora Instytutu Filozofii KUL, wyrażoną w piśmie WF-520-2/20 z dn. 28 maja 2020

I. **Tematyka rozprawy:**

Celem rozprawy jest matematyczny opis problemu podobieństwa melodii i jego rozpoznawania za pomocą algorytmów samouczących się w sposób zbliżony do detekcji stosowanej przez człowieka. Autor podczas modelowania problemu odkrywania wiedzy związanej z detekcją podobnych melodii wziął pod uwagę (zweryfikował użyteczność) czynników ludzkich, odczuć jakimi kieruje się człowiek. W celu przetestowania możliwości rozpoznawania melodii wyekstrahowano z próbek melodii zestaw zmodyfikowanych danych na potrzeby testowania algorytmów klasyfikacji. Autor zaimplementował pomocniczo autorskie oprogramowanie 'Generator', do tworzenia modyfikacji melodii oraz tworzenia reprezentacji do wczytywania w pakietach Weka i Orange. Wspomniane pakiety posłużyły do analizy danych, oraz weryfikacji poziomu możliwości automatycznego uczenia się. W rozprawie odkryto, że najlepszym algorytmem jest J48, został wytypowany do szerszej analizy. W toku badań okazało się, że rozpoznawanie melodii nie zależy od czynnika ludzkiego. Stworzono model matematyczny, który zasilił testowane modele uczenia maszynowego. Skuteczność systemu sięga 80 procent przy klasyfikacji binarnej oraz 66 procent przy podziale na 5 klas. W rozprawie poruszane są interdyscyplinarne problemy związane z kreowaniem technik inteligencji obliczeniowej, między innymi zagadnienia teorii wiedzy, umysłu. Relacji między matematyką a muzyką, estetyką a porządkiem, percepcją, uczeniem maszynowym i sztuczną inteligencją.

Tematykę rozprawy uważam za aktualną i bardzo ważną w dyscyplinie informatyka i rozwoju technik inteligencji obliczeniowej.

II. **Zawartość, struktura rozprawy**

Rozprawa składa się ze streszczania, spisu treści, wprowadzenia, trzech rozdziałów podzielonych na podsekcje, zakończenia, bibliografii i sekcji informującej o załącznikach. We wstępie autor nakreśla cel pracy i przedstawia informacje o źródłach danych pozyskanych do modelowania postawionego

problemu. Miedzy innymi o materiałach eksperymentalnych pozyskanych do rozprawy, z badań ankietowych, wyników testów automatycznego uczenia się, analizy statystycznej.

Autor w Rozdziale 1, bardzo obszernie wprowadza w problemy kognitywistyki związane problemem uczenia się i rozpoznawania wzorców. W Sekcjach 1.1 i 1.2 Omawiane są problemy z teorii wiedzy i umysłu, uczenia maszynowego, problematyki sztucznej inteligencji oraz relacji matematyki do muzyki. Autor prezentuje state of art technik rozpoznawania podobieństwa melodii. W kolejnych podsekcjach omawia problem percepcji muzyki, liczenia podobieństwa i kategoryzowania. Następnie wprowadza w problematykę binarnej klasyfikacji podobieństwa melodii. Autor omawia takie parametry jak symetria, estetyka, struktura dźwięku, geometria, fraktalne kompozycje, liczba taktów, spektrogram, punkty kontrolne, spektrogramy, chromagramy, szybka transformata Fouriera, geometryczna miara podobieństwa.

W Rozdziale 2, omawiany jest problem rozpoznawania podobieństwa melodii w zależności od czynników ludzkich oraz sposobu wprowadzania zmian w melodii. W tym rozdziale autor prowadzi dyskusję naukową, stawia przed sobą cele, hipotezy, które w kolejnych podsekcjach weryfikuje. Sekcja zawiera obszerne materiały teoretyczne i poprawnie przeprowadzone testy praktyczne. W tym rozdziale autor przeprowadza badania inicjacyjne. Metodą ankietową badano wpływ czynników ludzkich na rozpoznawanie podobieństwa melodii. Po przeanalizowaniu wyników autor podjął próbę formalizacji zebranych informacji, matematycznego opisu.

Autor stawia przed sobą hipotezy:

1) Rozpoznawanie podobieństwa melodii zależy od różnych czynników ludzkich takich jak płeć, wiek, zainteresowania, umiejętności gry na instrumentach, itp.
2) Większe podobieństwo melodii stwierdzamy na podstawie jednego długiego fragmentu tych samych kolejnych nut w melodii, niż na podstawie liczby takich samych nut co w melodii wzorcowej,
3) Podobnie rozpoznajemy podobieństwa melodii w próbkach zmodyfikowanych według tej samej zasady niezależnie od melodii wzorcowej.

Do eksperymentów stosuje grupę 153 osób, zastosowano badanie ilościowe, jakościowe oraz sondaż diagnostyczny.

Weryfikacja postawionych hipotez została przeprowadzona narzędziem IBM SPSS Statistics.

W Rozdziale 3 kontynuowane są badania rozpoznawania podobieństwa melodii, autor w tym rozdziale omawia możliwość automatycznego rozpoznawania przez uczenie się technikami data miningowymi. W rozdziale następuje przegląd technik klasyfikacji. W tym wprowadzenie do drzew decyzyjnych, sieci neuronowych, tabel decyzyjnych, k-NN, Naiwnego Klasyfikatora Bayesa, SVM. Testowanie ogólnie dostępnych i autorskiego narzędzia do przeprowadzania uczenia modeli. Autor przeprowadza eksperymenty finalne, dobiera odpowiednie cechy do modeli. Prezentuje finalne testy własnego systemu. Omawia też problemy między innymi związane z zaszumieniem danych. Omawiany jest też model przenoszenia wzorca podobieństwa na inne melodie. Badania polegały na ocenie podobieństwa 266 próbek zmodyfikowanych melodii do melodii oryginalnej. Autor utworzył 8 atrybutów warunkowych opisanych za pomocą wartości numerycznych i jeden atrybut decyzyjny, o wartości (TAK,NIE}. Narzędziem do przeprowadzenia badań był program Cubase, arkusz kalkulacyjny, oprogramowanie Weka, Orange oraz autorski program Generator. Autor ręcznie porównuje zgodność decyzji w excellu. Następnie dane są wprowadzane do programu Weka i dokonywana jest

rekonstrukcja ludzkiego postrzegania. Autor posiłkuje się własnymi publikacjami opisując wyniki rozprawy. W rozprawie rozważane są dwa podproblemy decyzyjne Pierwszy to problem binarny, o którym wcześniej wspominaliśmy. Drugi to problem klasyfikacji do pięciu klas: N=brak podobieństwa, S=małe prawdopodobieństwo, M-średnie podobieństwo, L-duże prawdopodobieństwo, XL duże podobieństwo. Autor omawia w rozprawie zastosowanie macierzy konfuzji (trafności) i wprowadza stosowne parametry. Autor wprowadza dodatkowo definicje parametrów do szacowania problemów regresji, np. MAE, RMSE, RAE, RRSE. Narzędzia badawcze jakie zostały zastosowane to: do przygotowania danych i analizy użyto instrumentu muzycznego Kurzweil PC3x, oraz programów: Cubase v.5.0, Generator (własny program), Excel 2010, Weka v.3.9.1 i Orange.

Struktura rozprawy jest prawidłowa, podział treści jest czytelny.

Uwagi krytyczne:

Zapisy matematyczne nie zawsze są perfekcyjnie zaprezentowane. Np. Zbiór {TAK/NIE}, powinien być postaci {TAK,NIE}.

- autor ze zbyt dużą pewnością omawia zagadnienia filozoficzne związane z AI, sama kwestia problemu ducha w maszynie powinna być szerzej omówiona,

- przykład, że krytycy sztucznej inteligencji uważają, że „komputer zawsze przeszukuje wszystkie możliwości, zaś człowiek od razu reaguje świadomie" jest nieprecyzyjne, ponieważ i być może nietrafione, ponieważ do prowadzenia procesów decyzyjnych w komputerze nie jest potrzeba cała dostępna wiedza, tylko często zindeksowane, kompaktowe reguły działania.

- przedstawiony test Turinga, jest naturalnie w tej chwili przestarzały i płaskie spełnienie wymogu procentowego odróżniania rozmówcy sztucznego od człowieka, nie determinuje pojawienia się AI, nie tylko Silnej ale także słabej.

- też w pracy czasami można odnieść wrażenia, że pewne osiągnięcia inteligencji obliczeniowej są odbierane jako duży krok do stworzenia Strong AI, np. stwierdzenie „komputery potrafią pamiętać więcej niż my? To chyba zbędna personifikacja maszyny,

- czasami sformułowania w pracy nie są standardowym żargonem informatycznym, np. „olbrzymich baz wiedzy",

- brakuje zdefiniowania, co rozumiemy, przez duże bazy danych,

- pojęcie uczenia maszynowego i data miningu jest tu mocną splątane ze sobą, brakuje oddzielnych definicji,

- brakuje szczegółowego omówienia kodu Parsona,

- pomiędzy technikami klasyfikacji, pojawia się SGD, który sam w sobie nie stanowi techniki klasyfikacji, tylko optymalizuje proces uczenia, techniki są zbyt krótko omówione,

- brakuje definicji MacroF1score, przy omawianiu zastosowania precision i recall do szacowania jakości klasyfikacji wieloetykietowej,

- brakuje definicji miary MAPE i Rsquare, gdy mowa o parametrach stosowanych w regresji,

- równania matematyczne czasami są przedstawione w sposób nadmiarowy,

- nie do końca trafna jest konkluzja porównania zdawalności testu Turinga w jego najprostszej formie do sukcesu wynikającego z możliwości rozpoznawania podobieństwa melodii.

Mimo mojego krytycznego spojrzenia, uważam powyższe uwagi za niemające wpływy na najważniejsze osiągnięcie zaprezentowane w rozprawie i jej główny wątek.

III. Ocena najważniejszych wyników rozprawy

(i) Dużą wartość dodaną ma interdyscyplinarny charakter rozprawy, bardzo szerokie opisywanie problematyki badawczej, wchodząc w wiele pod-dyscyplin kognitywistyki. Rozprawa stanowi bardzo przystępną lekturę dla badaczy nie zajmujących się ściśle technikami uczenia maszynowego.

(ii) Autor podczas modelowania problemu odkrywania wiedzy związanej z detekcją podobnych melodii wziął pod uwagę czynniki ludzkie, zaprezentował postulat, że odczucia jakimi kieruje się człowiek, mogą mieć wpływ na rozpoznawanie melodii, co stanowi nowe podejście do badanego problemu. Ten punkt finalnie został zweryfikowany, jako niekonieczny w procesie rozpoznawania, o czym będzie mowa w kolejnym punkcie.

(iii) Głównym celem rozprawy było stworzenia reprezentacji próbek melodii i utworzenie aplikacji do określania podobieństwa melodii. Moim zdaniem najważniejszym wynikiem niniejszej rozprawy jest sformalizowanie parametrów opisujących ścieżkę melodyczną, zbudowanie modelu i jego weryfikacja eksperymentalna za pomocą utworzonego narzędzia Generator. Przy użyciu programu, autor odkrył, że czynnik ludzki nie ma wpływu na rozpoznawanie podobieństwa melodii. Zaproponował matematyczną koncepcję rozpoznawania dwóch melodii. Na potrzeby eksperymentów stworzył wiele modyfikacji melodii. Generator jest oprogramowaniem napisanym w środowisku Java, główną koncepcją było automatyczne tworzenie modyfikacji melodii. Rozbudowano interfejs graficzny do pokazania melodii oryginalnej i jej modyfikacji. Do programu wprowadzono matematyczne atrybuty na podstawie których program generował plik tekstowy w formacie arff dla odtwarzania w programie Weka. Oprogramowanie generator jest jednym z głównych osiągnięć autora, stanowi innowacyjne narzędzie, o niespotykanej hybrydzie funkcjonalności. Narzędzie współpracuje z pakietem Weka do prowadzenia badań data miningowych.

(iv) Jakość wyników autora jest poparta publikacjami naukowymi:
[1] T.Grębski, K.Pancerz, P.Kulicki, Automatic Recognition of Melody Similarity. In: Rutkowski L., Scherer R., Korytkowski M., Pedrycz W., Tadeusiewicz R., Zurada J. (eds) Artificial Intelligence and Soft Computing. ICAISC 2019. Lecture Notes in Computer Science, vol 11509. Springer, 2019 r.
[2] T.Grębski, K.Pancerz, P.Kulicki, Experiments with Recognition of Melody Similarities Based on Human Independent Parameters, International Conference on Information and Digital Technologies 2019 June 25th - 27th, 2019 Zilina, Slovakia.
[3] T.Grębski, Modelowanie matematyczne w muzyce na podstawie twórczości Iannisa Xenakisa, Roczniki Kulturoznawcze, Tom X (1), Lublin 2019 r.
[4] T.Grębski, O relacjach między matematyką i muzyką, Roczniki Kulturoznawcze 7 (4), Lublin 2016 r.

Sugestie dalszych badań

Naturalnym rozwojem zaproponowanych technik jest sprawdzenie efektywności głębokich sieci neuronowych. Oraz połącznie odkrytych zależności w generowaniu melodii których linia melodyczna jest atrakcyjna dla użytkownika. Ciekawym dalszym wątkiem rozwoju może być zastosowanie odkrytych technik w szukaniu plagiatów piosenek, o czym autor rozprawy wspomniał. Oraz do przeszukiwania bazy utworów dla upewnienia się, że linia melodyczna nie jest obarczona prawami autorskimi. Wytworzone narzędzie nie jest finalnym produktem, stąd warto je rozwijać do osiągnięcia maksymalnej możliwej efektywności.

WNIOSEK KOŃCOWY

W rozprawie doktorskiej zaprezentowano innowacyjną metodę określania podobieństwa melodii z wykorzystaniem technik uczenia maszynowego. Innowację rozumiemy tu jako nową technikę hybrydową. Autor zaprezentował liczne przykłady systemów, które mogą korzystać z jego metody. Poruszony problem analizy podobieństwa strumieni dźwiękowych jest bardzo aktualnym i ważnym problemem w dyscyplinie Informatyka. W rozprawie zaprezentowano oryginalne wyniki autora. Bardzo ważnym osiągnięciem jest implementacja odkrytej metody, weryfikacja skuteczności uformowanych atrybutów w praktyce za pomocą narzędzia Generator.

Oceniając rozprawę doktorską „Podobieństwo melodii i jego rozpoznawanie przez komputer przy użyciu narzędzi automatycznego uczenia się” stwierdzam, że spełnia wymagania określone w Ustawie o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 roku z późniejszymi zmianami. Rozprawa spełnia, moim zdaniem, wszystkie wymagania ustawowe przyjęte w polskim środowisku naukowym stawiane rozprawom doktorskim i wnoszę o jej przyjęcie oraz dopuszczenie do publicznej obrony.

Piotr Artiemjew