

Białystok, 31 marca 2014 roku

Prof. dr hab. Jarosław Stepaniuk

Katedra Systemów Informatycznych i Sieci Komputerowych

Wydział Informatyki, Politechnika Białostocka

Recenzja Rozprawy Doktorskiej Mgra Krzysztofa Sopyły
pt. „Metody akceleracji procesu klasyfikacji z wykorzystaniem
Support Vector Machines”
dla Rady Wydziału Informatyki Polsko-Japońskiej Wyższej
Szkoły Technik Komputerowych

Niniejsza opinia została przygotowana w odpowiedzi na pismo Pani Dr Aldony Drabik Dziekan Wydziału Informatyki Polsko-Japońskiej Wyższej Szkoły Technik Komputerowych. Moja opinia, przygotowana zgodnie z obowiązującymi przepisami, dotyczyć będzie oceny rozprawy doktorskiej mgra Krzysztofa Sopyły zatrudnionego na stanowisku asystenta w Katedrze Metod Matematycznych Informatyki Wydziału Matematyki i Informatyki Uniwersytetu Warmińsko – Mazurskiego w Olsztynie.

Autor określił, na stronie 13 rozprawy, że jej celem jest przyspieszenie procesu uczenia algorytmu Support Vector Machines dla dużych i wielowymiarowych zbiorów danych. Zagadnienia naukowe rozpatrzone w rozprawie są sformułowane dostatecznie jasno przez Autora. Rozprawa ma charakter zarówno teoretyczny jak i doświadczalny – wyniki rozprawy zostały uzupełnione szczegółową prezentacją wytworzonego oprogramowania.

Rozprawa została napisana w języku polskim, liczy 168 stron. Składa się ze spisu rysunków, spisu tablic, wstępu, sześciu rozdziałów, załącznika i bibliografii. Pokróćce przedstawię zawartość poszczególnych rozdziałów. „Wstęp” zawiera cel, zakres i układ rozprawy. W Rozdziale 2 omówiono zagadnienia związane ze statystyczną teorią uczenia zapoczątkowaną przez Vladimira Vapnika. W Rozdziale 3 Autor przedstawił, na podstawie literatury, zagadnienia architektury jednostek graficznych oraz paradygmat programowania masywnie

równoległego implementowanego w GPU (ang. Graphics Processing Unit). W Rozdziale 4 omówione zostały, na podstawie literatury, osiągnięcia w dziedzinie przeprowadzania treningu algorytmu SVM (ang. support vector machine, pl. maszyna wektorów nośnych). W Rozdziale 5 przedstawiono wkład Doktoranta w dziedzinę uczenia SVM na GPU polegający na dostosowaniu wybranych algorytmów do obliczeń na GPU. Opisano wyniki testów na publicznie dostępnych zbiorach danych oraz dokonano porównania z wynikami uzyskanymi innymi, znanymi z literatury, metodami. Rozdział 6 zawiera opis metod i eksperymentów w ramach badań nad uczeniem wersji liniowej SVM. W Rozdziale 7 opisano możliwości praktycznego zastosowania opracowanych metod oraz nakreślono kierunki przyszłych badań. Na stronach 161 – 168 rozprawy zamieszczono uporządkowaną alfabetycznie literaturę (119 pozycji).

W recenzowanej rozprawie przeprowadzono w sposób właściwy analizę źródeł dotyczących omawianej tematyki. Doktorant wykazał umiejętność poprawnego przedstawienia wniosków z przeglądu literatury.

Za najciekawsze wyniki zamieszczone w rozprawie uważam:

- Opracowanie metod przyspieszenia treningu nieliniowego SVM z wykorzystaniem akceleratorów graficznych. W badaniach wykorzystano funkcje jądra RBF, Chi^2 , ExpChi^2 . Zaproponowano delegowanie kosztownych obliczeniowo sekcji na GPU.
- Opracowanie mniej wrażliwych na dobór parametrów początkowych metod uczenia algorytmu SVM z jądrem liniowym.
- Samodzielny projekt i implementacja systemu oprogramowania wraz ze szczegółowym opisem technicznym oraz przeprowadzenie szerokich badań eksperymentalnych.

Również, za zalety rozprawy uważam:

- Zamieszczenie przykładów związanych z prezentowanymi pojęciami i definicjami.
- Bogatą listę literatury związanej z prezentowanymi w rozprawie zagadnieniami.

Nie zgłaszam uwag krytycznych do strony metodologicznej rozprawy. Natomiast, z obowiązku recenzenta zamieszczam informacje na temat pewnych mankamentów rozprawy związanych ze stroną prezentacyjną.

- Nadmierna długość rozprawy, np. część wykresów oraz pseudokod algorytmów ze stron 156 – 160 rozprawy wystarczyło by załączyć w wersji elektronicznej.

- W Rozdziale 4.3 pt. „Równoległe algorytmy uczenia SVM na GPU” mamy tylko jeden podrozdział 4.3.1 pt. „Porównanie dotychczasowych rozwiązań”, podczas gdy zwyczajowo, jeżeli wyróżniamy podrozdziały, to powinny być wyróżnione co najmniej dwa. Analogicznie w Rozdziale 4.4 mamy również wyróżniony tylko jeden podrozdział.
- Wykresy zamieszczone od strony 132 do strony 139 rozprawy nie są zbyt czytelne.

Wspomniane wyżej drobne nieścisłości nie zmieniają pozytywnego obrazu rozprawy.

Rozprawa zawiera podsumowanie wyników opublikowanych w pracach Doktoranta, głównie we współautorstwie z dr Pawłem Drozdą, dr Przemysławem Góreckim i dr Piotrem Artiemjewem. Spośród kilkunastu publikacji Doktoranta, wymienionych na stronie internetowej <http://wmii.uwm.edu.pl/~ksopyla/research/> zwraca szczególną uwagę siedem prac opublikowanych w materiałach konferencji międzynarodowych i indeksowanych w znanej informatycznej bazie DBLP: <http://dblp.uni-trier.de/pers/hd/s/Sopyla:Krzysztof.html>. Szkoda, że ciekawe badania Doktoranta nie zostały jeszcze zaprezentowane szerszemu gronu badaczy w postaci publikacji w renomowanym czasopiśmie o zasięgu międzynarodowym.

Podsumowanie

Oceniana rozprawa wymagała przeprowadzenia badań teoretycznych i eksperymentalnych. Autor wykazał bardzo dobre opanowanie warsztatu badawczego. Dlatego też, stwierdzam, że oceniana rozprawa spełnia wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy i wnoszę o dopuszczenie Pana mgra Krzysztofa Sopyłę do dalszych faz przewodu doktorskiego.

Jarostaw Stepaniuk