

ИЗВЕШТАЈ

о прегледу мастер рада

„Имплементација и евалуација метода класификације за велике скупове података“

кандидат: Дарко Стошић

Одлуком Наставно-научног већа Математичког факултета која је донета на 333. редовној седници одржаној 24. јуна 2016. године именовани смо за чланове комисије за преглед и оцену мастер рада под насловом „Имплементација и евалуација метода класификације за велике скупове података“ кандидата Дарка Стошића, студента мастер студија на студијском програму Информатика, на Математичком факултету.

Област рукописа

Рукопис „Имплементација и евалуација метода класификације за велике скупове података“ припада областима Истраживање података и Машинско учење. Додатно, рад се може сврстати и у подобласт Аналитике великих података.

Структура рукописа и кратак приказ

Рукопис се састоји од 44 стране које су организоване у 6 поглавља, укључујући и библиографију.

Прво поглавље је фокусирано на појашњавање концепта Великих података (енг. Big data). Машинско учење је други битан концепт у овом поглављу, а поред тога аутор се специфично осврће на релевантне алгоритме Машинског учења које надаље користи у раду, логистичку регресију и вештачке неуронске мреже.

Друго поглавље даје преглед актуелних технологија које се користе у Аналитици великих података, у питању су *Hadoop* и *Apache Spark* технологије за дистрибуирану обраду података, обе засноване на тзв. *MapReduce* парадигми.

Треће поглавље садржи предложене дистрибуиране имплементације логистичке регресије и вештачких неуронских мрежа помоћу поменуте две платформе. Због даљег испитивања ефикасности предложених дистрибуираних имплементација, кандидат је имплементирао и секвенцијалне верзије алгоритама.

Четврто поглавље описује експерименталне резултате и њихов упоредни однос. Кандидат је демонстрирао ефикасност дистрибуираних алгоритама тако што их је упоредио међу собом, али и са секвенцијалним верзијама алгоритама. Поређењем добијених тачности класификације на јавно доступним скуповима података, кандидат је показао да не постоји значајна разлика у квалитету између секвенцијалних и дистрибуираних алгоритама, као и да су добијене тачности значајно боље од оне која се добија применом наивног алгорита класификације.

Закључна разматрања су дата у претпоследњем поглављу. Исто поглавље садржи и предлоге за могућа унапређења.

Поглавље Литература садржи списак од 17 библиографских јединица које је кандидат користио приликом писања рада.

Анализа рукописа

У рукопису који смо анализирали кандидат је описао проблем класификације података над Великим подацима и применио ефикасне дистрибуиране алгоритме за његово решавање. Кандидат је демонстрирао разумевање проблематике Великих података и показао завидно програмерско знање када је у питању примена *MapReduce* парадигме зарад прилагођавања неких метода Машинског учења.

Закључак и предлог

Реализацијом овог рада и пратећом имплементацијом, кандидат Дарко Стошић је показао висок степен стручног знања. На основу свега наведеног, Комисија предлаже да се рукопис под насловом:

„Имплементација и евалуација метода класификације за велике скупове података ”

прихвати као мастер рад и да се закаже његова јавна усмена одбрана.

Београд, 22. октобар 2019.

Комисија:

др Александар Картељ, доцент, ментор

др Владимир Филиповић, ванредни професор

др Младен Николић, доцент