

**МОЛБА  
ЗА ОДОБРАВАЊЕ ТЕМЕ МАСТЕР РАДА**

Молим да ми се одобри израда мастер рада под насловом:

**Аналитика великих података на кластеру коришћењем Apache Spark**  
чији су значај и специфични циљ следећи:

Енормна производња података на вебу - терабајта и петабајта дневно, као и растући захтеви за софистицираним претраживањем, анализом и другим Интернет услугама, захтевају нови приступ њиховом моделирању, складиштењу, приступу и обради.

Најчешће коришћени програмски модел, двоетапни Map-Reduce базиран на диску, као и системи развијени на њему, веома су успешни као подршка апликацијама над великим подацима на кластеру. Ипак, ови системи нису ефикасни за апликације које укључују вишефазна израчунавања и дељење података међу фазама, какви су, на пример, алгоритми машинског учења. Spark је нови софтверски оквир израчунавања на кластеру, оригинално развијен на Калифорнијском универзитету у Берклију, који користи меморијске апстракције складиштења. Због тога је посебно погодан за интерактивну анализу великих података, посебно за извршавање алгоритама машинског учења. Постиже повећање брзине и до 30 пута у итеративним алгоритмима машинског учења.

У раду ће бити приказане главне карактеристике Apache Spark - а, принципи рада, опис његових компоненти са посебним акцентом на библиотеку за машинско учење MLlib. Употреба Apache Spark - а ће бити илустрована у апликацији машинског учења великих размера, која ће користити неки од алгоритама из те библиотеке.

Литература:

Matei Zaharia. Spark: In-Memory Cluster Computing for Iterative and Interactive Applications. Invited Talk at NIPS 2011 Big Learning Workshop: Algorithms, Systems, and Tools for Learning at Scale.

<http://www.spark-project.org>.

Мурат Дацић 1072/2014,  
рачунарство и информатика

\_\_\_\_\_

(име и презиме студ., бр. инд., ознака програма и модула)

\_\_\_\_\_

(својеручни потпис студента)

\_\_\_\_\_

(датум подношења молбе)

Катедра \_\_\_\_\_ је сагласна са предложеном темом.

\_\_\_\_\_

(шеф катедре)

Сагласан ментор  
др Гордана Павловић-Лажетић

\_\_\_\_\_

(својеручни потпис ментора)

Чланови комисије

1. др Саша Малков \_\_\_\_\_

2. др Јелена Граовац \_\_\_\_\_

\_\_\_\_\_

(датум одобравања молбе)