

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 1337

Obrada podataka tehnologijom Apache Spark

Martin Matak

Zagreb, svibanj 2016.

Umjesto ove stranice umetnite izvornik Vašeg rada.
Da bi ste uklonili ovu stranicu obrišite naredbu \izvornik. Na ovoj stranici se

nalazi izvornik.

Zahvala - TODO ...:)

CONTENTS

1. Uvod	1
2. Zaključak	3

1. Uvod

Pametni mobilni uređaji postaju neizostavan dodatak svakog modernog čovjeka.

Većina pametnih mobilnih uređaja u sebi sadrži sljedeće senzore:

akcelerometar - elektromehanička komponenta koja mjeri sile ubrzanja;

barometar - mehanički senzor za mjerenje atmosferskog pritiska (na trenutnoj lokaciji uređaja);

senzor svjetlosti - mjeri intenzitet, tj. jačinu svjetlosti i uglavnom se nalazi s prednje strane uređaja, iznad ekrana;

senzor blizine - u stanju prepoznati situacije kada mu neki objekt stoji u blizini - ovo omogućava automatske pozive prilikom primicanja telefona licu uz zaključavanje telefona da bi onemogućili slučajno prekidanje poziva uhom ili slično;

senzor gestikulacije - prepoznaje kretnje ruke tako što detektira infracrvene zrake koje se reflektiraju - omogućuje nam djelomično upravljanje telefonom bez doticanja ekrana;

žiroskop - uređaj koji se koristi za navigaciju i merenje kutne brzine;

geomagnetski senzor - mjeri okolno geomagnetsko polje za sve tri fizičke osi i u biti služi kao kompas na mobilnim uređajima i

Hall Sensor - magnetski senzor zadužen za prepoznavanje je li maska telefona zatvorena ili otvorena.

Pretpostavimo da je mobilni uređaj spojen na internet i da svakih nekoliko sekundi pošalje vrijednost koju u tom trenutku mjeri pojedini senzor. U samo jednom danu može se skupiti dosta podataka. A što kada to ne bi radili za jedan uređaj nego za sve izdane modele nekog uređaja? Količina podataka bi jako brzo narasla.

Kako količina podataka postaje sve veća, dolazimo do pojma *Velika količina podataka* (engl. *Big Data*). U današnje vrijeme imamo više podataka u digitalnom obliku

nego što smo ikada imali. Jedan od zanimljivijih izazova je kako ih efektivno obraditi i zaključiti nešto iz toga. Kako od te velike količine podataka doći do nekih pametnih zaključaka iz kojih ćemo nešto novo naučiti.

Apache Spark je otvorena (engl. *open source*) tehnologija koja omogućava pisanje programa za obradu podataka u tri programskih jezika: *Java*, *Python*, *Scala*; a nudi i mogućnost interaktivnog rada.

U okviru ovog rada biti će proučene mogućnosti ove tehnologije, razrađeno nekoliko konkretnih primjera obrade podataka te ostvarena programska rješenja koja obavljaju tu obradu koristeći *Apache Spark*.

Svi primjeri će biti napisani u programskom jeziku *Java*.

2. Zaključak

Zaključak.

Obrada podataka tehnologijom Apache Spark

Sažetak

Sažetak na hrvatskom jeziku.

Ključne riječi: Ključne riječi, odvojene zarezima.

Title

Abstract

Abstract.

Keywords: Keywords.