

## SPOROČILO ZA JAVNOST

### Prva doktorska disertacija na FIŠ

**Novo mesto, 2. 6. 2015 - Doktorand Andrej Kastrin je danes, 2. 6. 2015, izvedel zagovor prve doktorske disertacije na Fakulteti za informacijske študije v Novem mestu z naslovom Uvrščanje in diskretizacija mnogorazsežnih mikromrežnih DNA podatkovij. Zagovor je potekal pred komisijo, ki so jo sestavljali doc. dr. Zoran Levnajić (predsednik), izr. prof. dr. Janez Povh (mentor, član), doc. dr. Bernard Ženko (član) in izr. prof. dr. Ljupčo Todorovski (član).**

Doktorand je zagovor opravil zelo uspešno in dobil pohvale za visoko kakovostno opravljeno delo, ki bo lahko osnova za nadaljnje raziskovanje na tem področju v slovenskem in svetovnem merilu. Komisija za zagovor je ocenila, da je njegov izdelek celovit, predstavitev disertacije ter odgovori pa so bili podani na visokem nivoju. Avtorjevi objavljeni članki kažejo na relevanten prispevek k sodobni znanosti in na perspektivno nadaljevanje kariere. Prvemu doktorju znanosti s področja informacijske družbe s Fakultete za informacijske študije v Novem mestu čestitamo in želimo veliko uspeha na nadaljnji karierni poti.

**POVZETEK VSEBINE:** Tehnologija DNA mikromrež je danes dostopna v vsakem bolj opremljenem biomedicinskem laboratoriju. Uvedba tehnologije DNA mikromrež pa je v zadnjih dveh desetletjih pomembno zaznamovala tudi statistično znanost, saj je statistika neločljivo povezana s celotnim postopkom načrtovanja, priprave, analize in interpretacije mikromrežnih podatkov. Še posebej bi izpostavili problematiko analize mnogorazsežnih podatkovij, kjer imamo po pravilu opraviti s podatkovnimi tabelami, ki po stolpcih združujejo nekaj 10 000 genov, po vrsticah pa le nekaj desetih primerov.

Pomembno vlogo v biomedicinski praksi predstavlja problem uvrstitve vzorca v ustrezen razred. V primeru mikromrež bi to npr. pomenilo, da znamo pravilno napovedati ali profil meritev večjega števila genov ustreza skupini pacientov s Parkinsonovo boleznijo ali kontrolni skupini z zdravimi posamezniki. Tako je avtor v doktorski nalogi raziskoval vedenje različnih klasifikatorjev v problemski nalogi uvrščanja primerov v dva naprej podana razreda. Pri tem ga je še posebej zanimala kakovost uvrščanja ob hkratni uporabi metod za krčenje mnogorazsežnih podatkov ter vpliv diskretizacije zveznih spremenljivk na uvrščanje.

Kot ugotavlja avtor, se pri uvrščanju realnih mikromrežnih podatkovij najbolj odreže logistična regresija s kaznijo, najslabše pa nevronske mreže. Nad sintetičnimi podatkovji po kakovosti izstopa metoda podpornih vektorjev. Med metodama krčenja podatkovne matrike glede na kakovost uvrščanja ni statistično značilnih razlik (z izjemo ploščine pod ROC krivuljo). Med metodami diskretizacije se glede na uvrščanje najbolj odrežeta metodi MDLP in ChiMerge. Po našem védenju in dostopni empirični evidenci gre za prvo raziskavo na tako velikem številu mikromrežnih podatkovij in tako predstavlja nov in izviren prispevek k razvoju tega področja.