

ИЗВЕШТАЈ О ОЦЕНИ МАСТЕР РАДА

<b>I ПОДАЦИ О КОМИСИЈИ</b>
<p><b>1. Датум и орган који је именовео Комисију</b> 09.05.2019. Веће Департмана за математику и информатику Природно-математичког факултета Универзитета у Новом Саду</p> <p><b>2. Састав Комисије са назнаком имена и презимена сваког члана, звања, назива уже научне области за коју је изабран у звање, датума избора у звање и назив факултета, установе у којој је члан комисије запослен:</b></p> <ul style="list-style-type: none"><li>• др Душан Јаковетић, доцент Природно-математичког факултета у Новом Саду, ужа научна област: математичко моделирање, изабран у звање: 15.11.2015. – председник</li><li>• др Сања Брдар, научни сарадник БиоСенс Института, ужа научна област: информационе технологије, изабрана у звање: 26.10.2017. – ментор</li><li>• др Татјана Лончар-Турукало, ванредни професор Факултета техничких наука у Новом Саду, ужа научна област: телекомуникације и обрада сигнала, изабрана у звање: 26.04.2017. - члан</li></ul>
<b>II ПОДАЦИ О КАНДИДАТУ</b>
<p><b>1. Име, име једног родитеља, презиме:</b> Мирјана, Миленко, Самарцић</p> <p><b>2. Датум рођења, општина, република:</b> 26.04.1990., Нови Сад, Србија</p> <p><b>3. Година уписа на дипломске академске студије, смер/усмерење:</b> 2016, Мастер математичар-примењена математика (модул: наука о подацима)</p>
<b>III НАСЛОВ МАСТЕР РАДА</b>
“Редукција димензионалности над подацима људског микробиома“
<b>IV ПРЕГЛЕД МАСТЕР РАДА</b>
Мастер рад “Редукција димензионалности над подацима људског микробиома“ је написан на 53 стране. Рад се састоји од 8 поглавља: 1. Увод 2. Микробиом подаци

3. Редукција димензионалности
4. Класификација и евалуација алгоритама
5. Имплементација алгоритама
6. Резултати
7. Закључак
8. Литература

Рад почиње: апстрактном, списком табела, графикона и скраћеница коришћених у раду и уводом, а на крају се налази закључак и литература. Рад садржи 4 табеле и 16 графика. На крају, приложен је и списак коришћене литературе сачињен од 47 референци. У раду је дат опис података, теоријске основе и практична имплементација алгоритама за редукцију димензионалности. За сваки експеримент извршено је детаљно тестирање квалитета класификације–тачношћу, прецизношћу, одзивом и Ф1 мером.

## **V ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА МАСТЕР РАДА**

Прво поглавље представља увод у тематику рада, кратак опис студије људског микробиома, представљен је значај анализе микробиома за људско здравље. У уводном делу су такође наведене примењене методе, алатке и алгоритми које се користе за анализу и обраду података.

Друго поглавље описује тип биоинформатичких података, људских микробиома (микробиоми мушкараца и жена узетих са пет различитих делова тела) и поступак претпроцесирања података. Описане су методе које су биле потребне да се примене над сировим ДНК подацима у циљу припреме података за примену алгоритама машинског учења и неуронских мрежа. У мастер раду за ту сврху је коришћен QIIME2 софтверски пакет за микробиомску анализу, а почетни подаци су у форми демултиплексираних секвенци података.

Треће поглавље даје опис проблема високе димензионалности података, а надаље обрађује методе које се користе у раду за редукцију димензионалности (СВД метода факторизације матрице и аутоенкодер техника неуралних мрежа). Поглавље садржи теоријске и математичке основе алгоритама СВД декомпозиције. Дат је кратак увод о неуронским мрежама са дефиницијом и структуром мрежа и теоријским основама, описан је алгоритам пропагације уназад и аутоенкодер алгоритам.

Четврта глава садржи опис евалуације алгоритама и мера за евалуацију алгоритама. Представљене су технике за тестирање квалитета класификациј (тачност, прецизност, одзив и Ф1 мера).

Пета глава даје опис практичне имплементације алгоритама СВД декомпозиције и аутоенкодера као и објашњење круцијалних параметара за имплементацију. Сви алгоритми су извршени у Python програмском језику, коришћена је Scikit Learn package библиотека за машинско учење и Keras, Python библиотека за Deep Learning. Шеста глава представља преглед и дискусију добијених резултата у складу са претходно описаним мерама за евалуацију алгоритама. Висока прецизност (0.95, за прецизност оригиналних података 0.96) је постигнута применом Random Forest Classifier-а над подацима над којим је претходно извршен аутоенкодер алгоритам.

## **VI ЗАКЉУЧЦИ ОДНОСНО РЕЗУЛТАТИ ИСТРАЖИВАЊА**

Разумевање домена људских микробиомских података је од великог значаја за даљу анализу која може довести до веома важних медицинских открића. Подаци о људским микробиомима постају веома важни у одржавању здравља људи и новим медицинским третманима, укључујући персонализовану медицину.

Велика димензионалност података људског генома је једна од главних карактеристика. Применом алгоритама машинског учења и неуронских мрежа за редукују димензионалности смо показали да тачност редукованих података са добијеном вредношћу од 95% не одступа много од добијене вредности над целокупним подацима од 96%

## **VII КОНАЧНА ОЦЕНА МАСТЕР РАДА**

Мастер рад је у потпуности урађен у складу са одобреном темом. Садржај рада показује знање кандидаткиње да алгоритме редукује димензионалности и машинског учења примени на сложен реалан проблем из домена биоинформатике. У мастер раду су коришћени актуелни подаци преузети из опсежне микробиом студије. Рад је прегледно и добро написан, а резултати су детаљно анализирани и приказани.

## **VIII ПРЕДЛОГ**

На основу укупне оцене Комисија предлаже да се мастер рад прихвати, а кандидаткињи Мирјани Самарцић одобри одбрана.

Нови Сад,  
24.06.2019.

ПОТПИСИ ЧЛАНОВА КОМИСИЈЕ

др Душан Јаковетић,  
доцент ПМФ-а, председник

---

др Татјана Лончар-Турукало,  
ванредни професор ФТН-а, члан

---

др Сања Брдар,  
научни сарадник БиоСенс Института, ментор

---