

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 22.03.2022. године именовала нас је у Комисију за преглед и оцену мастер рада дипл. инж. **Бориса Мајића** под насловом „**Методе за поређење симулираних и геномских секвенци биолошког порекла**”. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Борис Мајић је рођен 02.01.1995. године у Београду. Завршио је основну школу у Ваљеву као вуковац, а потом и Ваљевску Гимназију са одличним успехом. Универзитет у Београду – Електротехнички факултет уписао је 2013. године. Дипломирао је на Одсеку за сигнале и системе 2020. године са укупном просечном оценом од 8,04. Мастер академске студије на Електротехничком факултету у Београду уписао је октобра 2020. године на Модулу за сигнале и системе. Положио је све испите са просечном оценом 9,6.

2. Извештај о студијском истраживачком раду

Борис Мајић је као припрему за израду мастер рада урадио истраживање релевантне литературе која се односи на области мастер рада: науке о подацима, биоинформатике, машинског учења и биомедицинске технике. Конкретно, анализирао је литературу која се односи на порекло и методе симулирања и анализе биолошких секвенци.

Посебан акценат је стављен на преглед алата за симулацију геномских података, синтезу обележја за поређење геномских секвенци, методе за анализу обележја и на технике поређења геномских узорака. Анализом литературе, одабрана су три симулатора *VarSim*, *pIRS* и *Neat* за синтетисање геномских података и три методе (наивна Бајесова класификација, логистичка регресија и метода насумичних шума) за квантификовање разлика између биолошких и синтетичких узорака.

3. Опис мастер рада

Мастер рад обухвата 80 страна од чега прилози 1-5 обухватају 5 страна, са укупно 33 слика, 2 табеле и 92 референце. Рад садржи увод, 3 поглавља и закључак (укупно 5 поглавља), списак коришћене литературе, списак скраћеница, списак слика, списак табела и 5 прилога.

Прво поглавље представља увод у геномске секвенце биолошког порекла. Описан је процес секвенцирања биолошких узорака и начини за представљање геномских секвенци. Дат је кратак опис постојећих метода за симулацију геномских секвенци и представљени су предмет и циљ рада. Поголавље садржи и кратак преглед два приступа анализи секвенци: 1) поравнање (енг. *alignment*) и 2) склапање (енг. *assembly*).

У другом поглављу је дат преглед свих метода које су коришћене у раду за симулацију геномских секвенци, предобраду података, издвајање обележја, анализу обележја, као и за поређење синтетичких и доступних биолошких узорака.

У трећем поглављу је детаљно представљен начин прикупљања и синтезе података биолошких и синтетисаних узорака. Посебан акценат је дат на одговарајућој конверзији формата и техници синтезе коначног скупа обележја.

Четврто поглавље садржи детаљан преглед свих добијених резултата и дискусију анализе обележја. На основу стечених знања из добијених резултата поређења секвенци, поглавље описује и симулацију додатних узорака са предефинисаним параметрима и резултате успешности ове накнадне анализе.

У оквиру петог поглавља дат је закључак у коме су резимирани резултати поређења секвенци. Представљена су ограничења симулатора на које је примењена метода поређења указала и дате су смернице за даља унапређења.

4. Анализа рада са кључним резултатима

Мастер рад дипл. инж. Бориса Мајића се бави развојем метода за поређење секвенци геномских података добијених коришћењем одабраних симулатора у односу на биолошке секвенце. Употреба симулатора за генерисање синтетичких геномских података одговарајуће сличности са биолошким секвенцама би у великој мери решила проблем доступности података за тестирање софтверских алата за обраду геномских података, посебно имајући у виду проблем приватности геномске секвенце и строге протоколе за приступ подацима. Симулирани подаци омогућавају тестирање софтверских алата у контролисаним условима и са унапред познатим параметрима, што их чини веома важним у свим областима анализе сигнала.

Резултат мастер рада је софтверска апликација реализована у Пајтону са радним токовима за аутоматизацију биоинформатичке анализе који су писани у CWL (енг. *Common Workflow Language*) језику која има за циљ поређење генерисаних секвенци и секвенци биолошког порекла. Посебно су анализирани резултати симулатора са подразумеваним параметрима за три одабрана симулатора, али и параметри добијени на основу стеченог знања са циљем опонашања биолошких узорака.

Основни доприноси рада су: 1) приказ и методологија поређења синтетичких и биолошких секвенци; 2) софтверска апликација за примену нових и одабраних техника за издвајање карактеристичних обележја; 3) могућност наставка рада и коришћења резултата за оптимизацију постојећих и реализацију нових симулатора геномских секвенци.

5. Закључак и предлог

Кандидат **Борис Мајић** је у свом мастер раду успешно решио проблем поређења обележја добијених из симулираних и биолошких геномских података и развио софтверски алат за издвајање обележја, поређење обележја и приказ резултата поређења обележја. Резултати поређења могу допринети значајном унапређењу постојећих и развоју нових симулатора, али и развоју нових метода за поређење геномских секвенци.

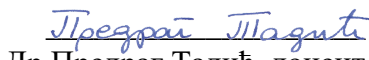
Борис Мајић је исказао велики ентузијазам за област обраде геномских података, али и самосталност, иновативност и систематичност у раду.

На основу изложеног, Комисија предлаже Комисији за студије II степена Електротехничког факултета у Београду да рад дипл. инж. Бориса Мајића прихвати као мастер рад и кандидату одобри јавну усмену одбрану.

Београд, 23.05.2022. године

Чланови комисије:


Др Надица Миљковић, ванредни професор


Др Предраг Тадић, доцент