

МОЛБА
ЗА ОДОБРАВАЊЕ ТЕМЕ МАСТЕР РАДА

Молим да ми се одобри израда мастер рада под насловом:

„Скривени Марковљеви модели у биоинформатици – електронска лекција”

Значај теме и области:

Биоинформатика је интердисциплинарна област која се бави применом рачунарских технологија у биологији и сродним наукама. Међу познатом литературом из овог домена издваја се *Bioinformatics Algorithms: An Active Learning Approach*, који служи као уџбеник на преко сто светских факултета. Ту спада и Математички факултет Универзитета у Београду, кроз мастер курс Увод у биоинформатику. Уџбеник обрађује разне занимљиве теме у оквиру биоинформатике, а десета глава бави се скривеним Марковљевим моделима. Ради се о статистичким моделима који се врло упрошћено могу схватити као стохастички коначни аутомат са приватним (скривеним) стањима и јавним (опсервабилним) емисијама. У питању су заправо два повезана процеса, чији се однос моделује. Примене скривених Марковљевих модела су разнорodne, како за надгледано, тако и за ненадгледано машинско учење над секвенцијалним подацима: класификација гена и протеина, моделовање и препознавање људског понашања, гестова, рукописа и говора, обрада звука и сигнала, одређивање врсте речи у тексту или чак моделовање тока пандемије.

Специфични циљ рада:

Актуелна иницијатива на нивоу курса Увод у биоинформатику јесте израда електронског уџбеника, заснованог на поменутој књизи. Стога је идеја обрадити управо претпоследње поглавље *Chapter 10: Why Have Biologists Still Not Developed an HIV Vaccine? – Hidden Markov Models*. Главни резултат била би електронска лекција, која би прво објаснила два важна биолошка проблема – класификацију протеина и проналажење гена (*CG* острва) – те прокоментарисала њихово „наивно” решавање. Затим би увела појам скривеног Марковљевог модела, детаљно га развила и имплементирала од нуле, те за крај прецизно решила изложене биолошке проблеме помоћу уведеног модела. Ток мисли из књиге био би само основа, док би тема била обрађена минуциозно и конкретније, са већим бројем додатних детаља и имплементираних могућности, чиме би био добијен прилично комплетан пакет за рад са скривеним Марковљевим моделима. Предвиђене технологије су програмски језик *Python* и интерактивне *Jupyter* свеске, али би постојала и *HTML* верзија, како би лекцију могао да погледа свако са прегледачем веба.

Лазар Васовић, 1011/2020, Информатика

(име и презиме студента, бр. индекса, модула)

Сагласан ментор доц. др Јована Ковачевић

(својеручни потпис студента)

(својеручни потпис ментора)

<датум>

(датум подношења молбе)

Чланови комисије

1.
2.

Катедра за рачунарство и информатику је сагласна са предложеном темом.

(шеф катедре)

(датум одобравања молбе)