

МОЛБА

Школска 2025/26 година

ПОДАЦИ О ПОДНОСИОЦУ МОЛБЕ			
Презиме:	Филиповић	Име:	Милена
Е-пошта:	mr221022@alas.matf.bg.ac.rs		
Број индекса:	1022/2022	Статус:	Самофинансирање
Ниво студија:	Мастер академске студије		

Тип: пријава теме мастер рада
Број молбе:
Катедра: Катедра за рачунарство и информатику
Кластеровање гена - електронска лекција
САДРЖАЈ МОЛБЕ:

Значај теме и области:

Кластеровање је један од најважнијих задатака у савременој биоинформатици. Могу се кластеровати различити биолошки подаци, а један од примера били би подаци везани за експресију гена. Експресија гена је процес којим се информација гена користи за синтезу функционалног генског продукта, а ти производи су најчешће протеини. Подаци о експресији гена стварају велике матрице бројева у којима је задатак идентификовати функционалне групе гена са сличним обрасцима понашања.

Постоји велики број алгоритама за кластеровање - од класичних попут *k-means*, преко хијерархијских (*UPGMA*, *DIANA*), до графовски заснованих метода детекције заједница (*Louvain*, *Leiden*). Сваки алгоритам има своје предности и мане и различито се понаша на стварним подацима.

Специфични циљ рада:

У овом раду биће развијена електронска лекција посвећена кластеровању гена, са фокусом на интерактивно упознавање са различитим алгоритамским приступима. Лекција обухвата теоријско објашњење и визуелизацију десет алгоритама из четири породице: партиционисање (*FarthestFirstTraversal*, *k-means / Lloyd*, *Soft k-means / EM*, *Brute-force k-means*), хијерархијски (*UPGMA* - агломеративно, *DIANA* - дивизивно), густо-засновани (*DBSCAN*) и алгоритми на графу сличности (*CAST*, *Louvain*, *Leiden*).

Биће приказана примена ових алгоритама на податке о експресији гена квасца *Saccharomyces cerevisiae*, при чему ће резултати различитих приступа бити упоређени како би се идентификовале конкретне предности и ограничења сваког од њих у контексту анализе генске експресије.

Корисницима ће бити омогућено да прате извршавање алгоритама корак по корак, чиме се олакшава разумевање процеса кластеровања. Ова интерактивна платформа може служити као едукативни алат за студенте и предаваче у области биоинформатике.

Остале битне информације:

Имплементација алгоритама биће изведена у програмском језику *Python*, искључиво уз употребу стандардне библиотеке (модули *math* и *random*). Овај приступ ставља нагласак на едукативну вредност и потпуно разумевање математичких основа алгоритама.

Поред програмског кода, рад ће садржати: математичке формулације проблема и алгоритама (функције циља, доказе конвергенције, асимптотске сложености); анализу метрика сличности (Еуклидско растојање, *Pearson*-ова корелација) и њихов утицај на резултате кластеровања; упоредне тестове на скупу гена кроз временске тачке; као и дискусију практичних примена за биологе.

Рад је представљен као електронска лекција - интерактивни *Jupyter* notebook, заједно са пратећим *Python* модулима и документацијом, биће јавно доступан на *GitHub*-у као пројекат отвореног кода.

Комисија:

1. Ментор: проф. др Јована Ковачевић

2. Председник: проф. Мирјана Маљковић Ружичић
3. Члан: Невена Ћирић

ПРИЛОЗИ МОЛБИ:
1. образац молбе за одобрење теме мастер рада

Београд, мај 2026.

(потпис)

(Попуњава надлежна особа)

☐ ОДОБРАВА СЕ	мај 2026. (датум, име, презиме и потпис)
☐ НЕ ОДОБРАВА СЕ	

ОБРАЗЛОЖЕЊЕ: