

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 29.08.2023. године, именовала нас је у Комисију за преглед и оцену мастер рада дипл. инж. Ђавида Ђукића под насловом „Анализа утицаја броја опсервација на успешност ране детекције сепсе у медицинској дијагностици”. Након прегледа материјала, Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Давид Ђукић је рођен 18.06.1999. године у Лозници. Гимназију је завршио у Лозници као вуковац и ћак генерације. Електротехнички факултет, Универзитета у Београду је уписао 2018. године. Дипломирао је у септембру 2022. године са просечном оценом на испитима 9,50, а на дипломском 10. Мастер академске студије на Електротехничком факултету, Универзитета у Београду је уписао октобра 2023. на модулу за Сигнале и системе. Положио је све испите са просечном оценом 9,60.

2. Извештај о студијском истраживачком раду

Кандидат Давид Ђукић је као припрему за израду мастер рада, урадио истраживање релевантне литературе, која се односи на област машинског учења и статистичке анализе медицинских података. Конкретно, анализирана су постојећа решења и проблеми у области ране детекције сепсе уз помоћ модела машинског учења, са акцентом на утицај величине обучавајућег скупа на перформансе модела. Истраживањем релевантне литературе, утврђен је кључна улога одређивања одговарајућег броја података у обучавајућем скупу, што је неопходно за постизање висококвалитетних предикција.

3. Опис мастер рада

Мастер рад обухвата 45 страна, са укупно 11 слика, 7 табела и 24 референце. Рад садржи увод, 6 поглавља и закључак (укупно 8 поглавља) и списак коришћене литературе.

Прво поглавље представља увод у коме су описаны предмет, циљ и структура мастер рада. Такође, представљен је значај ране детекције сепсе у медицинској дијагностици, као и проблем величине скупа приликом примене метода машинског учења у раду са медицинским подацима.

У другом поглављу је дат опис коришћене базе података, као и детаљан поступак претпроцесирања података са нагласком на методе које су коришћене за замену недостајућих вредности.

У трећем поглављу су детаљно представљени ансамбли у машинском учењу, њихов начин рада, као и разлози за примену ансамбла за рану детекцију сепсе.

Четврто поглавље детаљно описује начин рада *XGBoost* модела машинског учења, од обучавања модела, додатног обучавања, комплексности модела, одређивања оптималне структуре, тзв. слабог ученика, учења структуре стабла одлучивања, техника за спречавање преобучавања модела до манипулатије недостајућим вредностима.

Пето поглавље пружа увид у методологију рада са посебним нагласком на опис анализе утицаја смањеног броја опсервација у обучавајућем скупу на перформансе модела, као и на одабир хиперпараметара *XGBoost* модела.

У шестом поглављу су дефинисане метрике, које су коришћене за евалуацију модела.

Седмо поглавље представља приказ добијених резултата, као и дискусију добијених резултата. Детаљно је разматрано како смањење броја података утиче на метрике модела на тест скупу података, али и како утиче на промену у расподели атрибута.

Последње поглавље је закључак у оквиру кога су сумирана достигнућа мастер рада. Додатно, резимирани су резултати мастер рада и дефинисана су потенцијална унапређења.

4. Анализа рада са кључним резултатима

Мастер рад дипл. инж. Ђукића се бави анализом утицаја броја опсервација на успешност ране детекције сепсе. Коришћени су подаци из јавно доступне базе података, која садржи податке за преко 16 хиљада испитаника, који су прикупљени у десетогодишњем периоду из две медицинске установе. За сваког испитаника су доступни подаци, који су мерени на сваких сат времена током боравка на интензивној нези, где је циљ анализе да се коришћењем модела машинског учења детектује сепса код пацијента у раним фазама болести. Развијена је софтверска апликација за рану детекцију сепсе и приказана је детаљна анализа утицаја броја пацијената у обучавајућем скупу на успешност предикције.

Основни доприноси рада су: (1) предложена је методологија претпроцесирања података код секвенцијалних података са великим бројем недостајућих вредности, (2) реализована је софтверска апликација за рану детекцију сепсе и (3) спроведена је анализа утицаја броја испитаника на перформансе модела машинског учења.

Део резултата мастер рада је приказан у раду који је представљен на међународној конференцији: S. Stokanović, D. Đukić, and N. Miljković, "Robustness of xgboost algorithm to missing features for binary classification of medical data," in 2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH), pp. 1–6, IEEE, 2024. <https://doi.org/10.1109/INFOTEH60418.2024.10495929>

5. Закључак и предлог

Кандидат Ђавид Ђукић је у свом мастер раду „Анализа утицаја броја опсервација на успешност ране предикције сепсе у медицинској дијагностици“ успешно решио проблем ране детекције сепсе помоћу *XGBoost* модела машинског учења и извршио детаљну анализу утицаја броја опсервација на перформансе модела. Предложена побољшања могу значајно да унапреде будуће примене *XGBoost* модела машинског учења за класификацију медицинских података. Поред тога, представљени резултати отварају пут ка додатном разматрању ограничења метода машинског учења у односу на доступан број опсервација и расподелу атрибута.

Давид Ђукић је исказао самосталност, зрелост и систематичност у своме поступку, као и иновативне елементе у решавању проблематике овог рада.

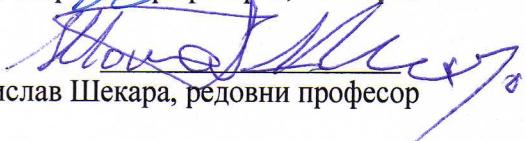
На основу изложеног, Комисија са задовољством предлаже Комисији за студије II степена Електротехничког факултета, Универзитета у Београду да рад дипл. инж. Ђавида Ђукића прихвати као мастер рад и кандидату одобри јавну усмену одбрану.

Београд, 28.08.2024. године

Чланови комисије:

 Нада Мильковић

Др Надица Мильковић, ванредна професорка, менторка

 Томислав Шекара

Др Томислав Шекара, редовни професор