ИЗВЕШТАЈ О ОЦЕНИ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

ПОДАЦИ О КАНДИДАТУ

Презиме, име једног родитеља и име	**Стевановић Марислав Никола**
Датум и место рођења	**17.08.1992, Ниш**

Основне студије

Универзитет	**Универзитет у Нишу**
Факултет	**Природно-математички факултет**
Студијски програм	**Информатика**
Звање	**Информатичар**
Година уписа	**2011**
Година завршетка	**2014**
Просечна оцена	**10.00**

Мастер студије, магистарске студије

Универзитет	**Универзитет у Нишу**
Факултет	**Природно-математички факултет**
Студијски програм	**Рачунарске науке**
Звање	**Мастер информатичар**
Година уписа	**2014**
Година завршетка	**2016**
Просечна оцена	**10.00**
Научна област	**Развој софтвера**
Наслов завршног рада	**Препознавање хијерархијских релација у векторском простору репрезентација речи**

Докторске студије

Универзитет	**Универзитет у Нишу**
Факултет	**Природно-математички факултет**
Студијски програм	**Рачунарске науке**
Година уписа	**2016**
Остварен број ЕСПБ бодова	**150**
Просечна оцена	**10.00**

НАСЛОВ ТЕМЕ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Наслов теме докторске дисертације	**Вештачке неуронске мреже за детекцију веб напада**
Наслов теме докторске дисертације на енглеском језику	**Artificial neural netowrk for web attack detection**
Име и презиме ментора, звање	**др Бранимир Тодоровић, ванредни професор**
Број и датум добијања сагласности за тему докторске дисертације	**НСВ број 8/17-01-010/22-014, 05.12.2022**

ПРЕГЛЕД ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Број страна	**102**
Број поглавља	**7**
Број слика (шема, графикона)	**13**
Број табела	**29**
Број прилога	**--**

ПРИКАЗ НАУЧНИХ И СТРУЧНИХ РАДОВА КАНДИДАТА
који садрже резултате истраживања у оквиру докторске дисертације

Р. бр.	Аутор-и, наслов, часопис, година, број волумена, странице	Категорија
1	**Stevanović, N.**, Todorović, B. & Todorović, V. Web attack detection based on traps. *Appl Intell* **52**, 12397–12421 (2022). https://doi.org/10.1007/s10489-021-03077-9 *У овом раду су коришћене замке за прикупљање података о новим веб нападима. Ти подаци су комбиновани са подацима из регуларног саобраћаја, како би се креирао корпус за тренирање и евалуацију. Тестиран је и адаптиран велики број плитких и дубоких модела машинског учења на проблем детекције напада. Модели користе једноставне улазне податке, као што су карактери и комбинације узастопних карактера. Тестиране су перформансе модела и приликом детекције напада нултог дана. Један од кључних проблема у машинском учењу је катастрофално заборављање. У овом раду је понуђено више начина инкременталног учења, који су прилагођени проблему детекције малициозних веб захтева.*	M21
2	**Nikola Stevanović,** Character and word embeddings for phishing email detection; *Computing and Informatics, 41*(5), 1337–1357. (2022)https://doi.org/10.31577/cai_2022_5_1337 *Рад обрађује тему детекције фишинг мејлова. Фишинг напади су неки од најучесталијих малициозних активности на интернету, и у њима нападачи презентују себе као особу или организацију од поверења, не би ли од жртве добили неку поверљиву информацију или остварили илегалну материјалну добит. За разлику од ранијих приступа који су користили ручно осмишљене улазне атрибуте, у овом раду је коришћена уградња карактера и речи из текстова мејлова у векторски простор. Предложена је комплексна архитектура неуронске мреже као модел детекције фишинг мејлова. За екстракцију глобалних и локалних шаблона, архитектура користи рекурентне и конволуционе слојеве. Приликом евалуације је потврђена њена висока ефикасност.*	M23
3	**Nikola Stevanović,** Population-based feature selection for intrusion detection; International Conference on Applied Artificial Intelligence (SICAAI); 2022 *Рад обрађује тему детекције напада на основу анализе разлика у параметрима мреже у току напада и иначе. Проблем који се решава је одређивање параметара који су најкорисниији за откривање напада. Предложен је метод селекције корисних параметара мреже базиран на популацији. Метод је комбинован са више класификатора, и одликује га способност да избор параметара мреже прилагоди класификатору који се одабере. У току евалуације је потврђена ефикасност метода. Успео је да повећа тачност свих тестираних класификатора помоћу избацивања параметара мреже који немају велики допринос у откривању напада. Поред веће тачности, смањењем броја улазних параметара се повећава и брзина класификовања.*	M34
4	**Nikola Stevanović,** *Embedding and weighting of Website Features for Phishing Detection, Facta Universitatis, Series: Mathematics and Informatics, Vol. 40, No 1, 013-031, (2025) https://doi.org/10.22190/FUMI221113002S* *Рад обрађује тему избора и уградње карактеристика „фишинг“ напада. Током „фишинг“ напада, нападачи користе разне техничке и трикове социјалног инжењеринга како би покушали да намаме жртве на „фишинг“ веб локацију. Веб локација изгледа као да припада поузданој организацији, али је заправо воде нападачи и користе је да обману жртве да открију своје лозинке, бројеве кредитних картица или друге поверљиве информације. У раду је коришћен метод дискретне описне карактеристике веб локације, на основу које се открива да ли је веб локација „фишинг“ или легитимна. Креиран је прилагођени слој за уграђивање, посебно дизајниран за ове врсте карактеристика, као и механизам за софт селекцију атрибута. Предлажен је модел, заснован на конволуционој неуронској мрежи, за откривање „фишинг“ веб локација и демонстрирана његова ефикасност на три скупа података. Са стопама тачности до 97,56%, модел ради упоредо са или боље од тренутних најсавременијих приступа на разматраним скуповима података.*	M51

ИСПУЊЕНОСТ УСЛОВА ЗА ОДБРАНУ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Кандидат испуњава услове за оцену и одбрану докторске дисертације који су предвиђени Законом о високом образовању, Статутом Универзитета и Статутом Факултета.

[redacted] **НЕ**

Кандидат је у својој дисертацији разматраио веома актуелан проблем примене вештачких неуронских мрежа и других алгоритама и модела машинског учења у области сајбер безбедности, а посебно проблем инкременталног учења беза заборављања и селекције релеватних атрибута за препознавање напада. Сви резултати приказани у дисертацији су нови и оригинални, а добар део тих резултата је публикован у међународним и домаћим часописима и представљен на међународој конференцији. Вредност дисертације представља јасна дефиниција проблема, предлог и имплементација решења. Тестирања, извршена на различитим корпусима примера, су показала да предложена решења значајно редукују заборављање претходно стеченог знања приликом тренирања

модела на новим примерима. Такође, тренирање са селекцијом релевантних атрибута даје моделе са већим тачношћу препознавања напада на тест примерима, од модела који су тренирани без селекције релеватних атрибута.

ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Кратак опис појединих делова дисертације (*до 500 речи*)

У првом поглављу су описане теоријске основе, слојеви вештачких неуронских мрежа(линеарни, слој уграђивања, конволуциони слој, слој избора, слој нормализације и рекурентни слој ЛСТМ), као и основе примене машинског учења у детекцији веб напада.

У другом поглављу је описана детекција малициозних веб захетва помоћу замки. Замке представљају дигиталне клонове разних уређаја, постављене на јавним веб адресама. Коришћене замке су резултат истраживања фирме АСТ д.о.о. Ниш, са којом је кандидат сарађивао при изради докторске дисертације. Са ових замки је прикупљен скуп података за тренирање модела машинског учења, описаних у дисертацији. Поред овог корпуса је коришћен и велики корпус примера веб захтеве FWAF, који је креирала компанија Fsecurity. Укратко је описан поступак креирања корпуса за тренирање, на основу података добијених са замки, као и поступак издвајања карактеристика на основу којих ће бити извршена класификација веб захтева. Описани су улази тзв. плитких модела машинскох учења као и начин представљања улаза тзв. Дубоких модела, тј. Неуронских мрежа. У анализи успешности детекције напада тестирани су следећи, тзв. плитки. модели машинског учења: логистичка регресија, класификатори са максималном маргином (енгл. Support Vector Machines SVM), пасивно-агресивни класификатори и класификатори случајне шуме. Поред плитких, разматрана су и три тзв. дубока модела: LSTM, TextCNN и CNNLSTMCNN. Као критеријумска функција за дубоке моделе коришћена је бинарна крос ентропија. Описан је поступак тренирања и тестирања и успех модела у детекцији напада нултог дана. Такође је приказано време тренирања и тестирања у циљу анализе применљивости разматраних модела у реалном времену. Приликом тренирања, због изразито небалансираног скупа података коришћене су технике подузорковања и преузорковања. Резултати показују да сви модели имају сличне резултате на тест примерима, при чему плитки модели не заостају, напротив врло често, поготову пасивно агресивни класификатор и логистичка регресија, дају боље резултате од дубоких модела. При томе су на тест скупу примера, подузорковањем добијена боља тачност него пре-узорковањем. Што се тиче напада нултог дана, бољу тачност су постизали дубоки модели. Очекивано, плитки модели су бржи и при тренирању и при тестирању.

У трећем поглављу је разматран проблем инкременталног учења модела на основу ограниченог броја нових примера, при чему је посебна пажња посвећена проблему катастрофалног заборављања. Проблем катастрофалног заборављања је добро познат али, до сада, није представљено опште решење. Приликом тренирања на новим примерима и плитки и дубоки модели, у мањој или већој мери заборављају већ стечено знање и грабљиво се прилагођавају новим примерима. Овакво понашање, поготову у случају модела за одбрану од веб напада, је неприхватљиво. Од таквих модела се очекује да се брзо прилагођавају новим нападима, тј. да уче брзо само на основу нових примера, без потребе да буду тренирани на свим доступним примерима, и да том приликом не забораве већ стечено знање. Као корпус примера за тренирање коришћена су оба корпуса из поглавља 2. У 3. поглављу је предложен модел, који представља надоградњу модела учења са мањим заборављањем и модели који користе бафер карактеристичних примера из претходних скупова података. Ова унапређења примењена су само на дубоке моделе LSTM, TextCNN и CNNLSTMCNN. Најбољи резултати у учењу без заборављања добијени су комбиновањем унапређеног алгоритма учења са мањим заборављањем уз постојећи бафер за претходне примере.

У четвртом поглављу разматран је проблем избора атрибута, тј. карактеристика на основу којих ће бити извршена класификација веб захтева. Разматран је алгоритам избора карактеристика заснован на популацији. Елементи популације су подксупови скупа свих атрибута. Сваки елемент популације је представљен вектором нула (атрибут није изабран) и јединица (атрибут јесте изабран). Свака генерација популације је представљена вектором вероватноћа димензија броја атрибута, чији елементи представљају вероватноћу да је атрибут одабран. Скор који се постиже на валидационом скупу примера, после тренирања са одабраним скупом атрибута је коришћен у алгоритму промене вероватноћа избора атрибута. При свакој провери селектован је подскуп инстанци из сваке генерације. Овај приступ избора атрибута тестиран је са моделима «адабоост», стабло одлуке, к-најближих суседа, логистичка регресија, неуронске мреже и случајне шуме. На свим овим моделима показано је унапређење тачности са бројем популација.

У петом поглављу је разматран проблем детекције «фишинг» мејлова. Предложено решење засновано је на примени неуронске мреже са адаптивном уградњом карактера и речи као карактеристика на основу којих ће бити извршена класификација. За тренирање и тестирање су коришћена два корпуса SpammAssassin и Nazario. Предложени модел има две паралелне гране, грану обраде карактера и грану обраде речи. Сваки од ових грана се додатно дели на две гране обраде: двосмерну LSTM и конволуциону мрежу. После примене нормализације слоја на излазе из ове две гране, њихови излази си комбинују конкатенацијом. Слична логика је примењена и на две главне гране обраде карактера и обраде речи. Резултати добијени на тест делу корпуса су упоредиви или бољи од

резултата тестирања тренутно најбољих објављених решења.

У шестом поглављу је описан модел машинског учења за детекцију «фишинг» напада, односно «фишинг» веб сајтова заснован на конволуционој неуронској мрежи. Неуронска мрежа поседује адаптивни слој уградње карактеристика сајтова у векторски простор. Додатном адаптацијом утицаја-тежина појединих карактеристика-атрибута у односу на друге, извршена је такозвана софт селекција карактеристика атрибута. Коришћена су три корпуса са UCI репозиторијума корпус за тренирање и тестирање модела машинског учења за детекцију «фишинг» сајтова. Корпуси се састоји од примера представљених почетним скупом атрибута подељеним је у четри подскупа. То су атрибути адресне линије са укупно дванаест атрибуте, атрибути абнормалности са укупно шест атрибута, HTML и JavaScript атрибути, њих пет, и атрибути домена којих има седам. Сваки пример има укупно тридесет атрибута. За тренирање је коришћена критеријумска функција бинарне крос ентропије. Коришћена је L2 регуларизација параметара модела, L1 регуларизација параметара слоја софт селекције карактеристика атрибута и регуларизација којом се максимизира варијанса ових параметара. Тестови су показали да су тачност, прецизност и сензитивност предложеног модела са адаптивном софт селекцијом атрибута боље него без адаптивне селекције за разлику од 0.2 до 0.8 процената на разматраним корпусима.

У седмом поглављу дат је кратак преглед дисертације.

ВРЕДНОВАЊЕ РЕЗУЛТАТА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Ниво остваривања постављених циљева из пријаве докторске дисертације (*до 200 речи*)

Остварени су сви научни циљеви постављени у пријави докторске дисертације. У дисертацији је разматрана примена алгоритама и модела вештачке интелигенције, плитких и дубоких, у одбрани од сајбер напада. Разматран је проблем прикупљања података за тренирање, способност различитих архитектура да уче на основу примера напада, инкрементално учење на основу малог броја примера без заборављања као и избор релаватних атрибута од стране алгоритама учења у циљу повећања тачности и ефиксности модела машинског учења. Предложена су, имплементирана и тестирана два приступа за инкрементално учење. Резултати на тест примрима су показали да је значајно смањено заборављање претходно стеченог знања, поготову ако се два предложена приступа комбинују у току учења.

Разматран је проблем избора релевантних атрибута за класификацију веб захтева применом алгоритма заснованог на анализи популације примера и модификације вероватноће селекције одређених атрибута, тренирања са селектованим атрибутима и избора оптималног подскупа атрибута. Овај приступ избора атрибута тестиран је на моделима «адабоост», стабло одлуке, к-најближих суседа, логистичка регресија, неуронске мреже и случајне шуме. На свим овим моделима показано је унапређење тачности са бројем популација. Предложен је, имплементиран и тестиран модел машинског учења за детекцију «фишинг» напада, односно «фишинг» веб сајтова, заснован на конволуционој неуронској мрежи. Предложен је, имплементиран и тестиран алгоритам софт селекције релеватних атрибута примера за тренирање. Резулати на тест примерима су показали да су тачност, прецизност и сензитивност предложеног модела са адаптивном софт селекцијом атрибута боље него без адаптивне селекције за разлику од 0.2 до 0.8 процената на разматраним корпусима.

Вредновање значаја и научног доприноса резултата дисертације (*до 200 речи*)

Кандидат је у својој дисертацији разматрао веома актуелан проблем примене вештачке интелигенције у области сајбер безбедности. Посебно актуелан је проблем инкременталног учења без заборављања и селекције релеватних атрибута за препознавање напада. Решавање ових проблема омогућав моделима мапшинског учења и вештачке интелигенције да уче и раде на проблемима из области сајбер безбедности у реалном времену. Кандидат је предложио, имплементирао и тестирао решења ових проблема, и резултати на тестовима су показали унапређења у односу на постојеће алгоритме и моделе машинског учења и вештачке интелигенције.

Резултати дисертације публиковани су у међународном часопису Appllied Intelligence ***52****, 12397–12421 (2022) (категорија 21), међународном часопису Computing and Informatics, 41(5), 1337–1357, (категорија 23) и домаћем часопису Facta Universitatis, Series: Mathematics and Informatics, 2025, Vol. 40, No 1, 013-031, (категорија 51).*

Оцена самосталности научног рада кандидата (*до 100 речи*)

Током израде своје докторске дисертације кандидат је показао да може самостално да се бави научним радом и долази до вредних научних резултата. Самостално је одабрао област примене вештачке интелигенције и алгоритама машинског учења у одбрани од веб напада. У договору са ментором дефинисао проблеме којима ће се конкретно бавити: инкрементално учење без заборављања и селекција релеватних атрибута. Самостално је дошао до предлога решења, самостално имплементирао моделе и алгоритме тренирања, самостално тестирао на разноврсним скуповима података. Експерименти су показали да предложена решења значајно редукују заборављање претходно стеченог знања приликом тренирања модела на новим примерима. Такође, тренирање са селекцијом релевантних атрибута даје моделе са већим тачношћу препознавања напада на тест примерима, од модела који су тренирани без селекције релеватних атрибута.

ЗАКЉУЧАК (*до 100 речи*)

Кандидат је у својој дисертацији разматраио веома актуелан проблем примене вештачких неуронских мрежа и других алгоритама и модела машинског учења у области сајбер безбедности, а посебно проблем инкременталног учења беза заборављања и селекције релеватних атрибута за препознавање напада.
Сви резултати приказани у дисертацији су нови и оригинални, а добар део тих резултата је публикован у међународним и домаћим часописима и представљен на међународој конференцији. Резултат аутоматске процене оригиналности дисертације показује понављања само опште познатих математичких израза, имена атрибута јавно доступних скупова података за тренирање и тестирање модела и математичких израза који су већ објављени у радовима каднидата.
Вредност дисертације представља јасна дефиниција проблема, предлог и имплементација решења. Тестирања, извршена на различитим корпусима примера, су показала да предложена решења значајно редукују заборављање претходно стеченог знања приликом тренирања модела на новим примерима. Такође, тренирање са селекцијом релевантних атрибута даје моделе са већим тачношћу препознавања напада на тест примерима, од модела који су тренирани без селекције релеватних атрибута.
На основу свега овог Комисија са задовољством предлаже Наставно-научном већу Природно-математичког факултета Универзитета у Нишу да прихвати докторску дисертацију ***Николе Стевановића*** *под насловом* ***Вештачке неуронске мреже за детекцију веб напада (енгл. Artificial neural netowrk for web attack detection)*** *и да одобри њену јавну одбрану.*

КОМИСИЈА

Број одлуке Научно-стручног већа за природно математичке науке о именовању Комисије	**НСВ број 817/01/3-25/14**
Датум именовања Комисије	**07.04.2025**

Р. бр.	**Име и презиме, звање**		**Потпис**
	др Мирослав Ћирић, редовни професор	председник	
1.	Рачунарске науке	Природно математички факултет, Универзитет у Нишу	
	(Научна област)	(Установа у којој је запослен)	
	др Бранимир Тодоровић, ванредни професор	ментор, члан	
2.	Рачунарске науке	Природно математички факултет, Универзитет у Нишу	
	(Научна област)	(Установа у којој је запослен)	
	Др Марко Миладиновић, ванредни професор	члан	
3.	Рачунарске науке	Природно математички факултет, Универзитет у Нишу	
	(Научна област)	(Установа у којој је запослен)	
4.	др Александар Трокицић, доцент	члан	
	Рачунарске науке	Природно математички факултет, Универзитет у Нишу	
	(Научна област)	(Установа у којој је запослен)	
	др Драган Јанковић, редовни професор	члан	
4.	Рачунарство и информатика	Електронски факултет, Универзитет у Нишу	
	(Научна област)	(Установа у којој је запослен)	

Датум и место:

19.05. 2025, Ниш