



ТЕХНИЧЕСКИ УНИВЕРСИТЕТ – ГАБРОВО
Факултет „Електротехника и електроника”

маг. инж. Теодора Валентинова Жорова

**МОНИТОРИНГ НА ТРАФИКА В КОМУНИКАЦИОННИ
СИСТЕМИ С ИЗКУСТВЕН ИНТЕЛЕКТ**

А В Т О Р Е Ф Е Р А Т

на дисертация

за придобиване на образователна и научна степен „доктор”

Област на висше образование: 5. Технически науки

Професионално направление: 5.3. Комуникационна и компютърна техника

Докторска програма: Комуникационни мрежи и системи

Габрово, 2026 г.



ТЕХНИЧЕСКИ УНИВЕРСИТЕТ – ГАБРОВО
Факултет „Електротехника и електроника”

маг. инж. Теодора Валентинова Жорова

**МОНИТОРИНГ НА ТРАФИКА В КОМУНИКАЦИОННИ
СИСТЕМИ С ИЗКУСТВЕН ИНТЕЛЕКТ**

А В Т О Р Е Ф Е Р А Т

на дисертация

за придобиване на образователна и научна степен „доктор”

Област на висше образование: 5. Технически науки

Професионално направление: 5.3. Комуникационна и компютърна техника

Докторска програма: Комуникационни мрежи и системи

Научен ръководител: доц. д-р инж. Ивелина Стефанова Балабанова

Рецензенти: проф. д-р инж. Станимир Михайлов Садинов

проф. д-р инж. Валентина Илиева Маркова

Габрово, 2026 г.

Дисертационният труд е обсъден и насочен за официална защита на заседание на Разширен катедрен съвет на катедра „Комуникационна техника и технологии“ към факултет „Електротехника и електроника“ на Технически университет – Габрово, проведен на 19.03.2026 г.

Дисертационният труд съдържа 193 страници. Научното съдържание е представено в увод, четири глави и включва 151 фигури и 41 таблици. Цитирани са 161 литературни източника. Номерацията на фигурите, таблиците и формулите в автореферата е в съответствие с тази в дисертацията.

Изследванията по дисертационния труд са извършени в катедра „Комуникационна техника и технологии“ към факултет „Електротехника и електроника“ на Технически университет – Габрово.

Официалната защита на дисертационния труд ще се състои на 12.06.2026 г. от 11.00 ч. в зала 2215 на Технически университет – Габрово.

Благодарности: Изказвам своите най-сърдечни благодарности към доц. д-р инж. Ивелина Стефанова Балабанова за помощта и подкрепата при разработване на дисертационния труд.

Автор: маг. инж. Теодора Валентинова Жорова

Заглавие: МОНИТОРИНГ НА ТРАФИКА В КОМУНИКАЦИОННИ СИСТЕМИ С ИЗКУСТВЕН ИНТЕЛЕКТ

Тираж: 5 бр.

Място на отпечатване: Университетско издателство „Васил Априлов“

ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

Актуалност на проблема

През последните години революционното развитие на Изкуственият интелект и Машинното обучение създават основополагаща аналитична рамка, обезпечаваша нарастващата необходимост за подпомагане и облекчаване на трудоемкостта във всички дейности и процеси на технологичното развитие на човешкия фактор в различни сфери. Нараства необходимостта и изискванията към платформите с Изкуственият интелект за ефективно оптимизиране на параметрите на преносната среда за намаляване на смущенията и повишаване на диапазона на покритие. Други изисквания се свързват с подобряване на функционалността на DPI (Deep Packet Inspection) системите в следните направления:

- Обработка на Big Data клъстери;
- Установяване на възникналия прекъсвания в предаван трафик;
- Детектиране на широк спектър от киберзаплахи като DDoS атаки и т.н.;
- Анализирание на характеристиките на пакетно предаване и установяване на информация с нарушен интегритет и т.н.;

Съществен технически аспект се отнася до подобряване на механизмите и етапите на проектиране, експлоатация и оптимизиране на мрежови инфраструктури с висока степен на мащабируемост. Налице е необходимост от подпомагане на дейностите на QA (Quality Assurance) специалисти по отношение на идентифициране на дефекти, установяване на повтарящи се модели при манипулации с информационно съдържание, генериране на прогнозни модели на поведение и т.н.

Посочените факти потвърждават актуалността в търсене на иновативни възможности за функционално надграждане на услугите QoS и QoE чрез интеграция на AI базирани инструменти, приложения и модели в етап "Планиране" при проектиране на комуникационни трасета и ИКТ мрежови инфраструктури.

Цел и задачи на дисертационния труд

Целта на дисертационния труд е „Разработване на методология и хибридни аналитични модули за интелигентен анализ и оценка на качеството на услугите (QoS) и качеството на потребителското възприятие (QoE), базирани на подходи с изкуствен интелект, машинно обучение и дълбоко обучение, предназначени за адаптивен мониторинг, диагностика и прогнозиране на смущенията и трагичните характеристики на преносната комуникационна среда, при едновременно осигуряване на киберсигурността и информационната сигурност на мрежовите и комуникационните ресурси“. В обхвата на изпълнението на заложената цел се поставят следните основни задачи:

1. Да се проведе подбор на технически подходи за извличане на унифицирани информативни признаци при обработка на смущения и предаван мрежови трафик в комуникационна среда в хода на обучение и верификация на модели за разпознаване и прогнозен анализ.
2. Да се синтезират алгоритми за имитационно моделиране на телекомуникационни системи с чакане и Deskriptивен анализ на параметрите, формиращи спецификацията на преносната среда.
3. Да се изведат аналитични модели и се изследва взаимовръзката между факторите при обектен прогнозен анализ индекси на производителност, дефиниращи качествата на преносната среда при предаване на пакетна информация.
4. Да се извърши адаптиране и селектиране на методи и алгоритми за оптимизация при процедури по търсене на оптимум при минимизация на функции, отъждествявани с параметри за повишаване на бързодействието и обема на обслужвания трафик на комуникационната среда.
5. Да се разработят подходи за разпознаване и класификация на възникнали случайни смущения в комуникационните канали за връзка, внасящи негативни ефекти върху

- качеството на предаваната информация, чрез Artificial Intelligence и Machine Learning инструменти.
6. Да се направи подбор и верификация на Artificial Intelligence модели за прогнозиране на амплитудните изменения на шумове, възникващи в структурата на преносната среда при предаване на аналогови и цифрови сигнали в електрониката и комуникациите.
 7. Да се синтезират подходи с включване на Artificial Intelligence инструменти за регресионно моделиране при апроксимация на подобрени индекси на производителност на преносната среда в комуникационни инфраструктури с чакане.
 8. Да се разработят подходи за мониторинг на предавания трафик между мрежови сегменти по отношение на детектиране на състояние на зловредни атаки и мрежови аномалии на основата на Artificial Intelligence, Machine Learning и Deep Learning подходи.
 9. Да се съставят подходи за идентификация на разпределението на мрежови ресурси за ширококолов достъп в урбанизирани градски условия към корпоративни клиенти чрез Artificial Intelligence, Machine Learning и Deep Learning аналитични инструменти.
 10. Да се проведе сравнителен анализ и изготви комплексна оценка на ефективността на подходите и алгоритмите, създадени на основата на Artificial Intelligence, Machine Learning и Deep Learning апарати за диагностика на мрежовия трафик и достъпа до Интернет услуга.

Методи на изследване

Методологията на изследванията включва разработване на аналитични компоненти за клъстерен анализ, класификация и разпознаване, минимизация на функции чрез оптимизационни методи и алгоритми, регресионна диагностика и моделиране на основата на технологиите Artificial Intelligence, Machine Learning и Deep Learning.

Научна новост

Научната новост в приложените изследвания се свързва с адаптиране на класическите Невронни структури, традиционните методи на Машинно обучение, принципите на Дълбокото обучение, конвенционалните и неконвенционалните оптимизационни подходи за създаване на хибридни аналитични инструменти за мониторинг и диагностика на мрежовата среда в комуникационните възли и информационните системи.

Приложимост

Насочеността на предложените хибридни аналитични модули с AI, ML и DL апарати се състои в подпомагане на дейностите на мрежовите администратори в различни фази на планиране и управление на ресурсите, както и подобряване на качеството на обслужване в комуникационните инфраструктури. Комплексното предназначение в обхвата на потенциалните потребители следва да бъде обособено в следните категории:

- центрове за информационно обслужване, мобилни оператори, доставчици на ширококолов достъп до Интернет услуга при мониторинг на преносната среда, свързан с диагностика на смущенията, прогнозиране и оптимизация метриците за повишаване на капацитета и бързодействието на трафичното натоварване;
- индустриални производства и корпоративни мрежови сегменти при подобряване на устойчивостта, ефективността и гъвкавостта на внедрените системи за информационната сигурност срещу киберзаплахи и зловредни прониквания до управлявани дигитални ресурси.

Апробация на дисертационния труд

Основните аспекти на изследванията, запознаващи научната общност с постигнатите резултати в дисертационния труд, са отразени в седем научни труда на национални и международни форуми. Публикациите, изнесени на чуждестранни международни

конференции с реферирание и индексирание на научното съдържание в база данни Scopus, са две в обхвата на "Environment. Technology. Resources" - Латвия, и "Electronics, Engineering Physics and Earth Science" - Гърция. Част от научните трудове, обсъждани на семинари в Катедра "Комуникационна техника и технологии", са представени в организирани симпозиуми към ТУ - Габрово - "Международна научна конференция "УНИТЕХ"" и "Студентска научна сесия", съответно три публикации. Награден от два изнесени доклада на "Научна конференция на Русенски университет "Ангел Кънчев" и Съюз на учените - Русе" към секция "Комуникационна и компютърна техника" е "Internet Traffic Analysis by FFNN, k-Nearest Neighbors and Decision Tree Approach".

Изложените резултати са реализирани във връзка с изпълнение на три вътрешни университетски проекта в специализирани лаборатории за научни изследвания в Катедра "Комуникационна техника и технологии" при ТУ - Габрово.

Структура и обем на дисертационния труд

Дисертационният труд е структуриран в уводна част и четири основни глави, като включва списък с използваните съкращения и абривиатури, научно-приложни и приложни приноси, публикации, участие в проекти, заключителна част и библиографска справка.

В Глава 1 са обобщени научни трудове и изследвания относно анализ на състоянието на проблема, поставен в дисертационния труд. Изучени са съществуващите подходи и системни решения при детектиране на смущения, диагностика на предавания трафик, подходи за оптимизация и киберсигурност. Посочените обекти за анализирани в контекста на решаването на класификационни задачи и задачи и прогнозен анализ чрез технологиите Изкуствен интелект, Машинно обучение и Дълбоко обучение.

В Глава 2 са въведени подходи за детектиране на „смущения в комуникационните канали за връзка“ и „наложени шумови въздействия върху предавани аналогови и цифрови сигнали“. Въз основа на анализ на функционалността на AI и ML инструменти са синтезирани модели за идентификация и прогнозиране на амплитудните вариации на някои от най-често срещаните случайни въздействия, възникващи при пренос на сигнали в мрежовата среда.

В Глава 3 са предложени подходи за регресионно моделиране на основата на Класически регресионен анализ и Изкуствени невронни мрежи спрямо специфични индекси на трафичната производителност „Пропускателна способност“ и „Сумарно време за чакане и обслужване“. Решена е оптимизационна задача с интегриране на набор от алгоритми за нелинейна минимизация и Генетични алгоритми относно метрика „Комплексно време за престой в опашката и обработка на системни заявки“.

В Глава 4 е разработена методика за детектиране на киберзаплахи към предаван UDP трафик между корпоративни клиенти, базирана на AI и DL инструменти. Приложена е методична верификация при сценарии с различни набори от трафични метрики и типове трафик, съответно „Атакуван трафик от тип DoS“, „Нормален трафик“ и „Трафик с установено присъствие на мрежови аномалии“. Предложен е подход за диагностика на WEB трафик към корпоративни клиенти в активни часови зони.

СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

Глава I: Изследване на технически аспекти и подходи за интелигентен анализ, приложими при оценка на качеството на преносната среда

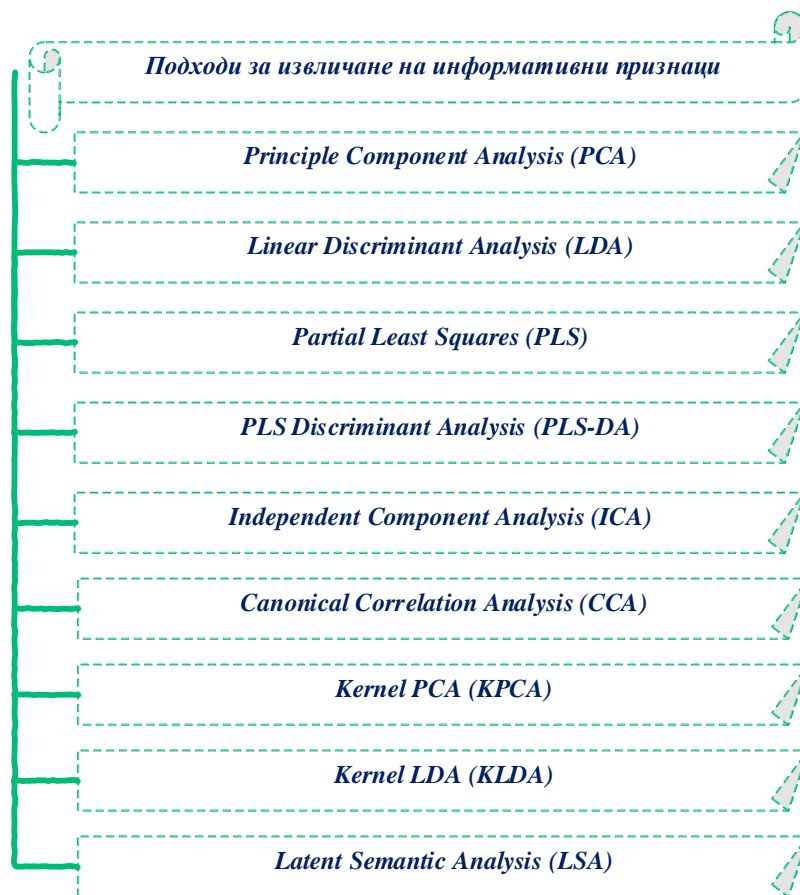
1.1. Data Mining инструменти за интелигентен анализ на данни при обработка на масиви от данни.

Концепцията за технологично развитие на съвременния свят в посока на аритметична прогресия е силно базирана на „Изкуственият интелект или Artificial Intelligence (AI)“. AI навлиза все повече във всички технически сфери на човешкото ежедневие, бизнеса, икономиката, информационните технологии и основни сфери на промишлеността.



Фиг. 1.10. Класификация на методи и алгоритми на Машинно обучение

Machine Learning (ML) на фиг. 1.10 се дефинира като AI подоблас, където са комбинирани информационни клъстери и специализирани алгоритми за провеждане на обучение във връзка с прогнозиране или изработване на набори за многовариантен избор на решение. При ML няма необходимост от конкретни процедури от тип „Явно програмиране“. Аналитичните “Deep Learning (DL)” инструменти или Deep Neural Networks (DNN) се разглеждат като част от ML методите, при които понятието “Deep” се обвързва с „дълбочината на слоевете на невронните мрежи или съвкупността от изчислителни единици в структурните многослойни скрити слоеве“.



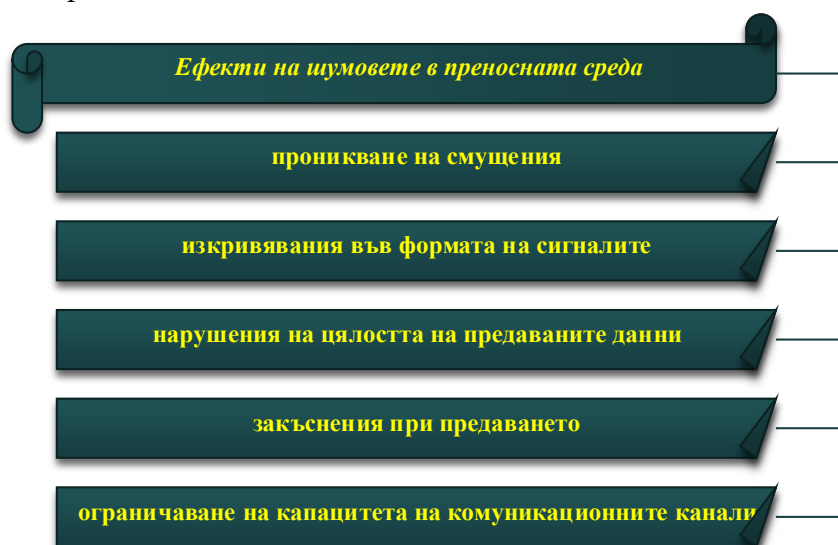
Фиг. 1.13. Подходи за извличане на информативни признаци при обработка на тестови AI и ML обекти

Подборът и прилагането на подходящи подходи за извличане на информативни признаци (фиг. 1.13) при обработка на целеви обекти като една от жизнените фази в AI и ML задачите за класификация, клъстеризация и прогнозен анализ са съществени за постигането на добра техническа рамка с висока ефективност.

1.2. Оценка на въздействието на смуцаващите въздействия в структурната преносна среда и комуникационните канали за връзка.

Шумовете като смуцаващи явления в комуникационните, електронните и електрическите системи могат да бъдат разделени и категоризирани по различни начини, критерии или признаци. Основна част от научните изследвания обособяват две групи “Internal Noises” и “External Noises“. Първата включва: “Thermal Noise” и “Imperfections”, докато към втората спадат: “Natural Origins”; “Electromagnetic Interference”; “Radio Frequency Interference” и “Crosstalk”. Според някои проучвания случайните въздействия в комуникационните канали за връзка могат да се дефинират в специфични категории: “White Noise”; “Pink Noise” и “Blue Noise”. Като различните форми на представяне и изследване на сигналите се отнасят до теоретичните: “Time”; “Spatial”; “Frequency”; “Autocorrelation” и “Wavelet”. Наличието на шумове в различните по вид, предназначение и обхват електронни и комуникационни системи

се свързва с редица ефекти, обобщени на фиг. 1.15. Един от моментите за ограничаване на тези въздействия е прилагането на аналитични инструменти за прогнозен анализ относно нивата на шума, показани на фиг. 1.17.



Фиг. 1.15. Ефекти на шумовете в преносната среда



Фиг. 1.17. Аналитични инструменти за класификация и прогнозен анализ

Широко използвани аналитични инструменти са Изкуствените невронни мрежи. В изследванията особено се акцентира на „принципите на дълбокото обучение“, т.е. DNNs. Популярни в това отношение са Long Short-Term Memory или LSTM DNNs. Често употребата на посочените невронни структури се свързва с хибридни подходи с Конволюционни невронни мрежи.

1.3. Инструменти за интелигентен анализ на състоянието и параметризацията на преносната среда в комуникационните системи.

Комплексният мониторинг на мрежовия трафик е една от основните дейности, регламентирани в политиките за достъп до информационни кълстери посредством WEB пространството на частни и корпоративни клиенти. По дефиниция се определя като многостранен процес за ефективно управление на трафичния поток, извършван от страна на:

- операторите "доставчици на Интернет услуга",
- мрежовите администратори в учреждения;
- компании "клиенти";
- различните регулаторни органи в сферата на информационното обслужване.

В качеството на основни комплексни направления при анализ и оценка на Интернет трафика могат да бъдат дефинирани следните аспекти:

- повишаването на производителността“;
- „ефективните технология за достъп, методи за защита“;
- „обезпечаване на обслужването на големи трафични потоци и буферното натрупване на информационни масиви“.



Фиг. 1.19. Категории интелигентни инструменти за анализ на състоянието на мрежовия трафик

В огромната си част съществуващите проучвания засягат детектиране и установяване на типа аномалии и нарушения в целостта на предаваната информация като за целта се използват различни категории инструменти (фиг. 1.19). Рекурентните невронни мрежи RNN и една от основните им разновидности LSTM мрежите се прилагат при оценка на трафичните характеристики при подмрежовите енкодери. Често използван невронен апарат се явяват също конволюционните невронни мрежи CNN в ролята на способи за предотвратяване на кибератаки.



Фиг. 1.20. *Аспекти при осигуряване на качество на услугата при частни и корпоративни клиенти*

Могат да се посочат и процесите, свързани с диагностика на трафика и детектиране на аномалии от въздействия на зловреден софтуер с помощта на мрежи на базата Gradient Descent и Momentum алгоритми. Проблемът с установяване на нарушения в мрежовия трафик е обект на цялостни системи в правителствени, административни и индустриални центрове по отношение на предлаганите от тях обществени услуги, систематизирани на фиг. 1.20. Друг аспект се свързва с подобряване на скалируемостта в интегрираните системни решения с машинно обучение при онлайн диагностика на мрежовия трафик.

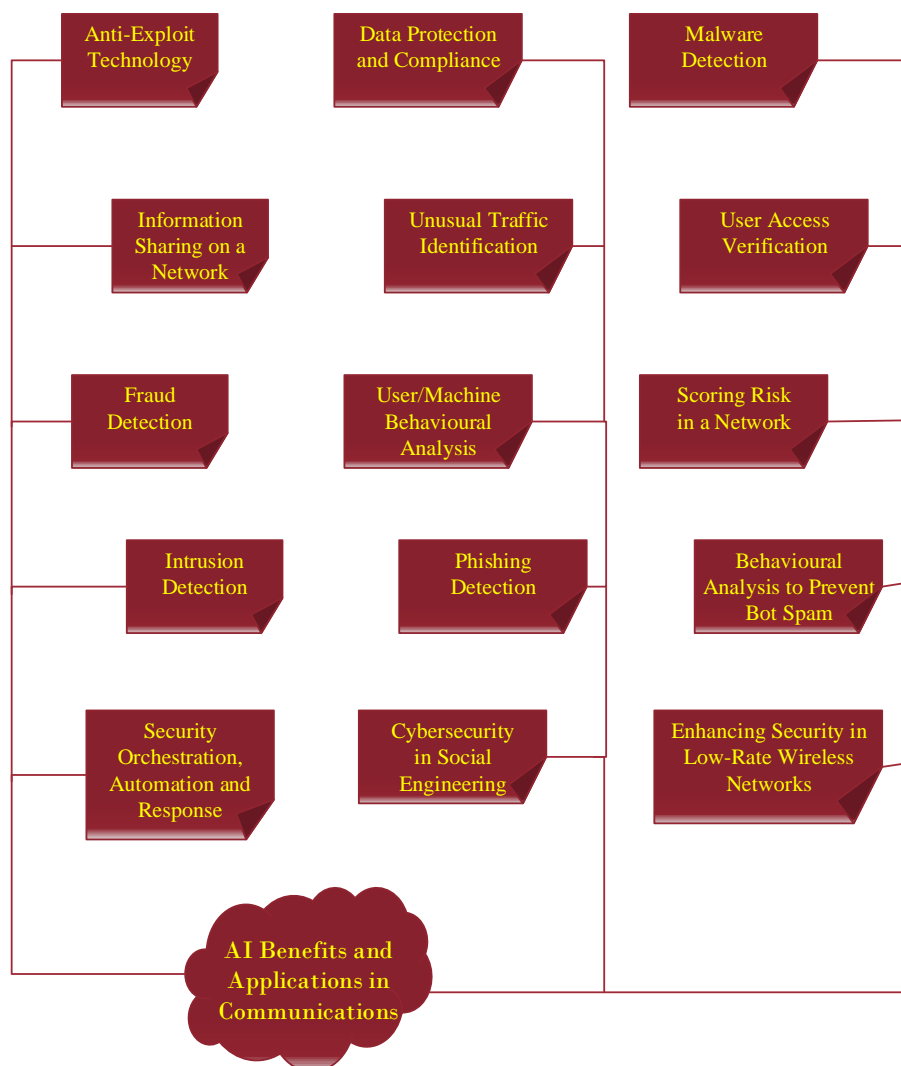
1.4. Ролята на изкуствения интелект в информационната сигурност и киберсигурността.

Понятието „Киберсигурност“ разграничава следните основни направления „Network security“, „Application security“, „Information security“, „Operational security“ и „Ресурсно възстановяване след аварийни и бедствени ситуации“. Интеграцията на инструментите на Изкуственият интелект в сферата на Киберсигурността включва процедурите „детектиране“, „прогнозиране“, „анализ“ и „ограничаване на заплахите“. Изграждането на AI – базирани публични платформи за управление на сигурността се превръщат в основна предпоставка за превенция срещу прилаганите видове атаки при злонамерени прониквания в мрежовите и комуникационните инфраструктури, както показано на фиг. 1.28 и фиг. 1.30.

В обхвата на корпоративната бизнес организация AI може да бъде успешно използван в редица области. Организациите, регламентиращи стандартизиращи рамки, предоставят на корпоративните бизнес организации критерии за изграждане на устойчиви и стабилни системи за сигурност, базирани на следните AI профили:

- Govern – предоставяне на набор от правни и технически аспекти;

- Identity – поддържане на актуална инвентаризация на информационното съдържание;
- Protect – проследяване и персонализация на активностите на потребителите;
- Detect – установяване на натоварвания, критични приложения, ограничения на ресурсите;
- Respond – предприемане на адекватни технически мерки;
- Recover – възстановяване на информационните и комуникационните ресурси при различни по сложност и вид инциденти.

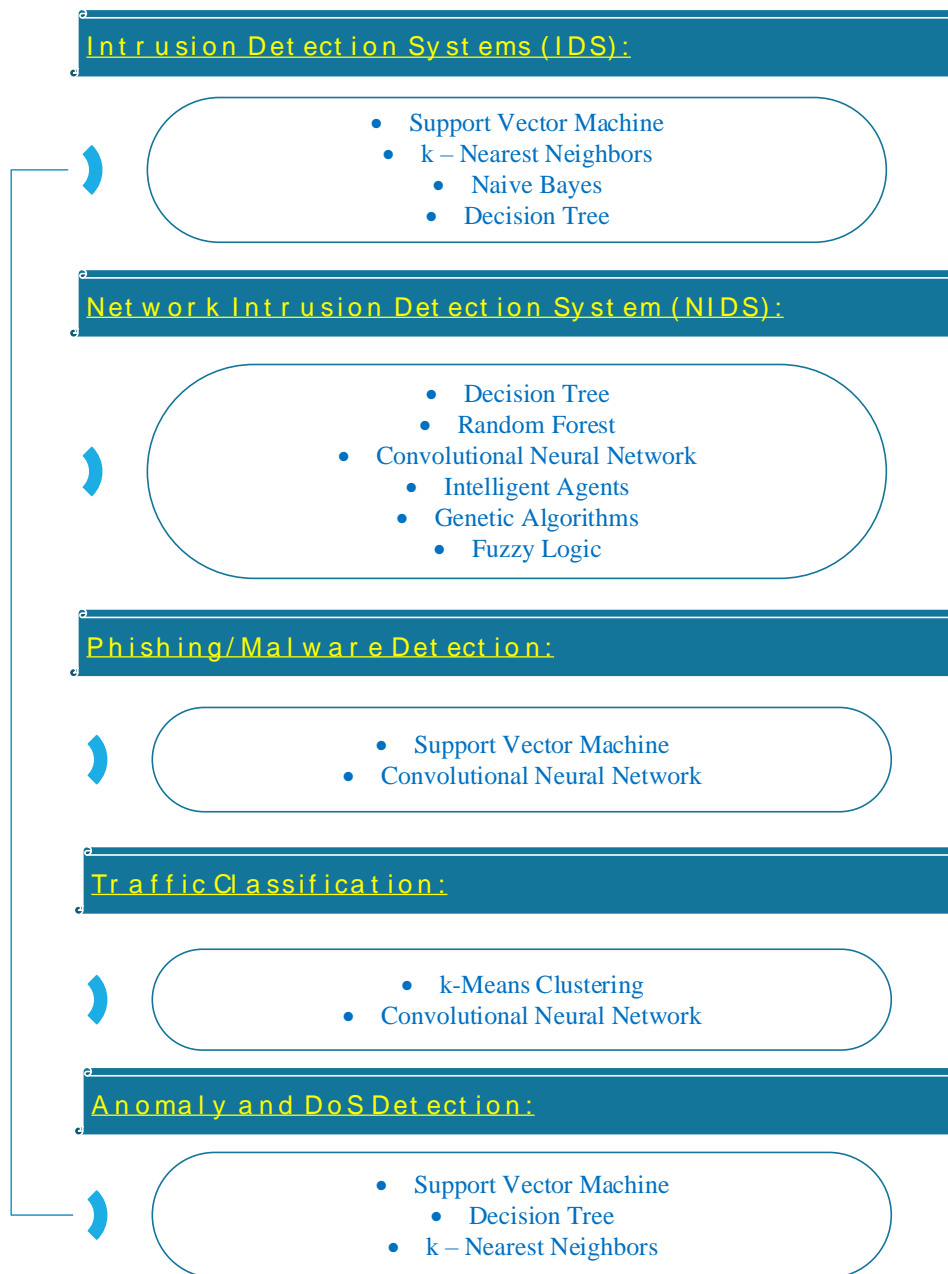


Фиг. 1.28. Ресурсни предимства и приложения при използване на AI в Киберсигурността

По отношение на информационните ресурси и хранилищата за данни, управлявани от AI в хода на създаване на модели за превенция срещу кибератаки в индустриалните ИКТ инфраструктури, съществуват редица рискове. Прямо действията на атакуващите съществуват два основни вида атаки, които могат да бъдат категоризирани, както следва:

- Input Attacks: Едновременно с входните обучаващи данни, атакуващият цели да интегрира модел, който да промени правилната функционалност и да сине AI системата:
 - ✓ Indistinguishable Perturbation: Атакуващите извършват неоткриваеми нарушения;
 - ✓ Content-Preserving Perturbation: Атакуващите извършват промени в информационното разпределение, без да променят съдържанието;
 - ✓ Non-Suspicious Input: Атакуващите могат да променят мястото на проникване, без възникване на подозрение за злонамерен достъп;

- ✓ Content-Constrained Input: Атакуващите могат да променят позицията на проникване толкова дълго, колкото тази процедура е свързана с внасяне на злонамерено натоварване на ресурсите на системата;
- ✓ Unconstrained Input: Атакуващите могат да извършат прониквания от произволни места в системата.
- Poisoning attacks: Тук атакуващият се стреми да разруши AI моделите, които вече са насочени към фазата на обучение.



Фиг. 1.30. AI технически подходи в системите за Киберсигурност

1.5. Изводи към Първа глава.

- Въз основа на анализ е потвърдена необходимостта от интегриране на технически подходи с Изкуствен интелект и Машинно обучение при спомагане на процесите за имитационно моделиране на телетрафични системи при предаване на данни;
- Направена е класификация въз основа на оценка на моделите, методите, алгоритмите и фазите на различните видове инструменти, формиращи концепцията Машинно обучение;

- Систематизирани са подходи за извличане на информативни признаци при обработка на регистрирани количествени параметри на обекти, процеси и системи с насоченост към диагностика чрез AI и ML;
- Извършена е оценка на техническите средства за класификация и прогнозен анализ на шумовете и смущенията в преносната среда в системите за комуникация;
- На базата на анализ е направена категоризация на интелигентните инструменти за анализ на състоянието на мрежовия трафик в информационно-комуникационните инфраструктури;
- Въз основа на анализ е установена необходимостта от конвенционални и неконвенционални методи и алгоритми за оптимизация за подобряване на поведението, устойчивостта и производителността в комуникационните системи;
- Потвърдена е необходимостта от внедряване на инструментите на Изкуственият интелект за диагностика на мрежовата сигурност и установяване на киберзаплахи, насочени към ресурсите на корпоративните информационно-комуникационни системи.

Глава II: Анализ на смущенията в преносната среда при предаване на данни

2.1. Изследване и синтез на SVM и k-NN класификатори за идентификация на Гаусов бял шум и Периодичен случаен шум.

C-SVM се базира на метрика "Capacity", чиито по-малки стойности позволяват контрол на допустимата граница между обособената граница на решение и областта с некоректно класифицирани еталони. В хода на изследванията е демонстрирано влиянието на високите нива на c-parameter. Спрямо отчетен Capacity = 10,000 в конкретния случай за указаните функции Linear, Polynomial и Sigmoid бяха наблюдавани силно занижени крайни точности, респективно 59.500 %, 68.900 % and 50.000 %. Въпреки това при Кернел функция „Radial Basis Function (RBF)“ бяха достигнати удовлетворяващи точности за обучаващия 86.863 % и 87.795 % тестовия набор, избрана с най-висока степен на пригодност.

Nu-SVM е класификационен тип, основаващ се на регулируем Nu-параметър с вариация в интервала от 0 до 1. Указаният индекс се асоциира с долната граница на количеството еталони, представляващи опорните вектори, и горната граница на частта на количеството еталони, разположени от некоректната страна на Хиперповърхнината. За конкретния SVM тип класификационни модели беше констатирана Минимална обща точност 68.900 % (Nu = 0,400) при употреба на Linear функция, последвана от по-доброто отчетено ниво 86.850 % (Nu = 0,500) при Sigmoid Kernel тип. Предимствата на Nu-SVM стават изразени при Polynomial функция и при ясно преимущество на RBF тип, където бяха достигнати силно положителни индикации на Обща точност 95.300 % и 97.900 % при Nu = 0,100 (фиг. 2.4). В последователен ред на дефинираните Kernel функции при C-SVM и Nu-SVM класификатори са установени „1468, 1279, 800 и 1482“ и „603, 168, 158 и 748“ комплексни Опорни вектора.

```

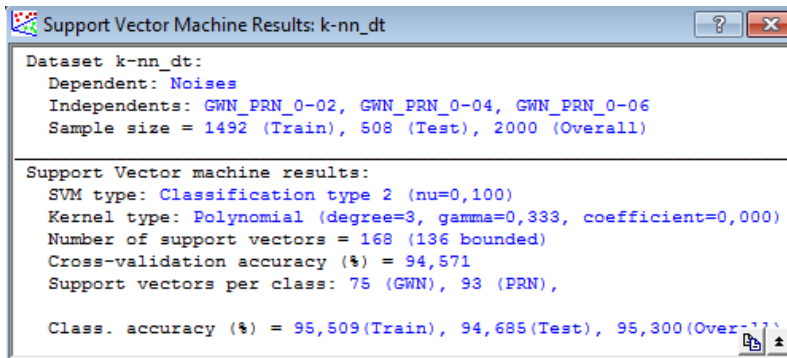
Support Vector Machine Results: k-nn_dt
Dataset k-nn_dt:
  Dependent: Noises
  Independents: GWN_PRN_0-02, GWN_PRN_0-04, GWN_PRN_0-06
  Sample size = 1492 (Train), 508 (Test), 2000 (Overall)

Support Vector machine results:
SVM type: Classification type 2 (nu=0,400)
Kernel type: Linear
Number of support vectors = 603 (603 bounded)
Cross-validation accuracy (%) = 68,164
Support vectors per class: 299 (GWN), 304 (PRN),

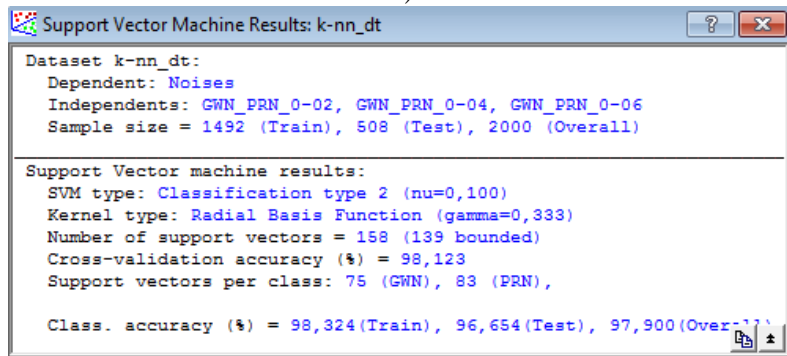
Class. accuracy (%) = 68,834(Train), 69,094(Test), 68,900(Overall)

```

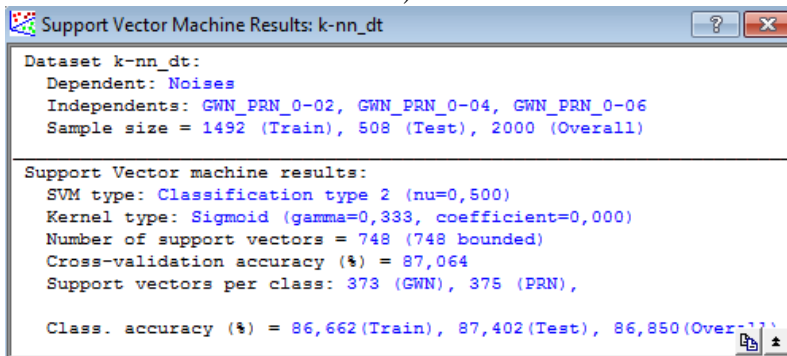
a)



б)

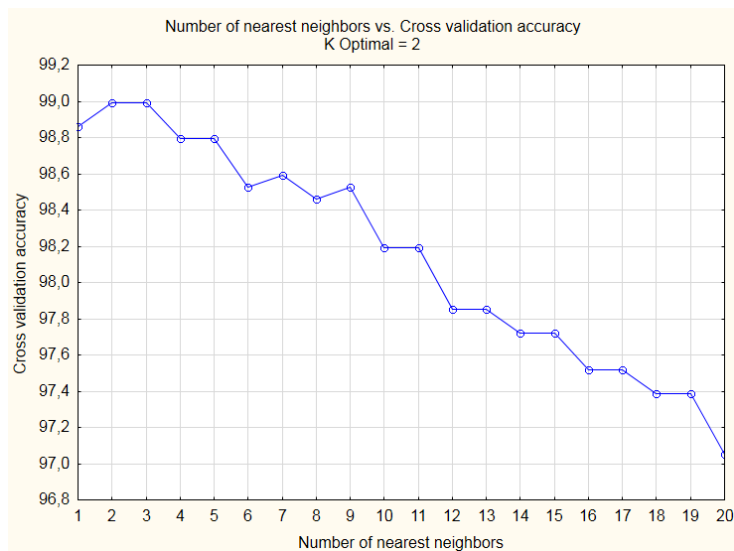


б)

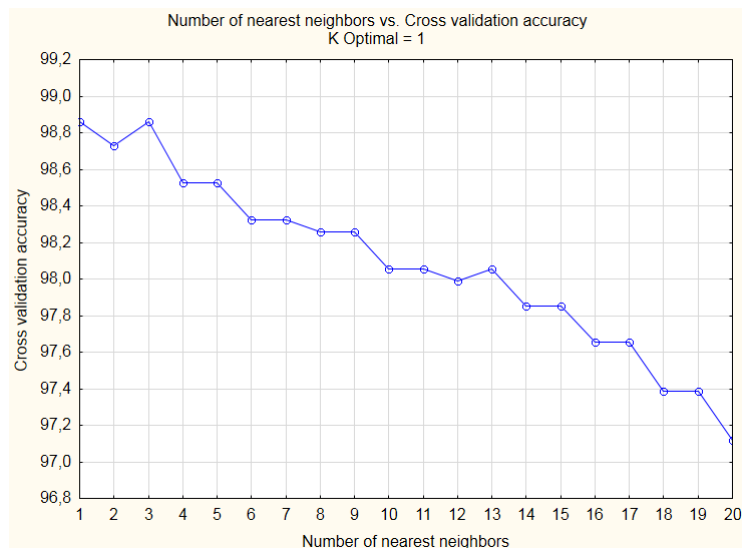


з)

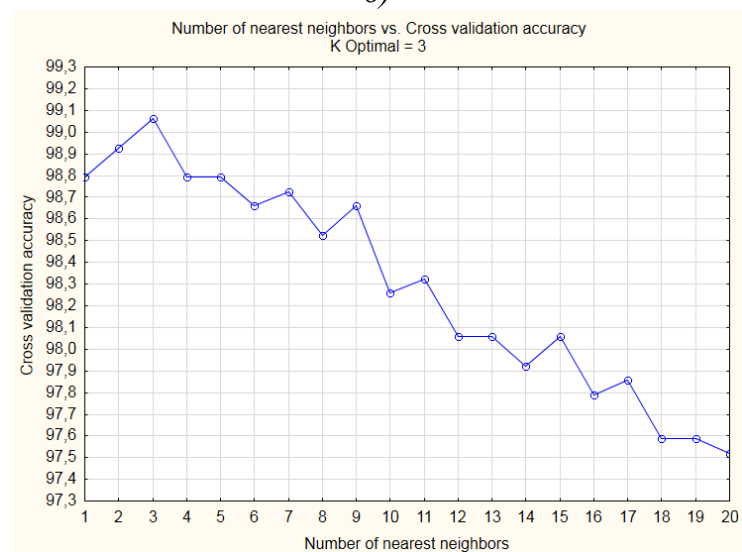
Фиг. 2.4. Nu-SVM класификатори при а) Linear, б) Polynomial, в) RBF и з) Sigmoid тип Кернел функции



а)



б)



в)

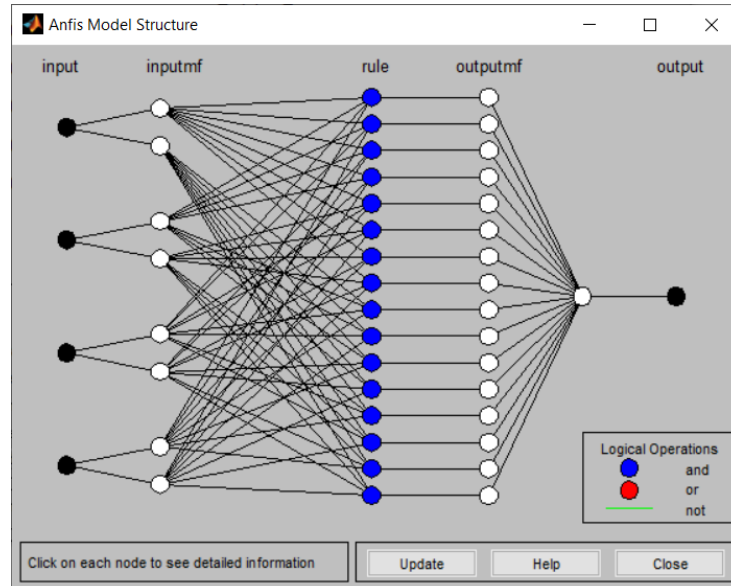
Фиг. 2.6. Итеративно проследяване на поведението на създадените k -NN модели при а) Euclidean, б) Cityblock и в) Chebychev разстояния за разпознаване на GWNs и PRNs

Процедурите по оценка на качеството на създадените k -NN класификационни модели, представени на фиг. 2.6, при зададени в последователен ред метрични единици за разстояние, препоръчителни в научните изследвания при решаването на класификационни задачи за случаите на „ $n < 10$ “. Получените емпирични степени на критерия точност, проследяващ итеративно поведението на приложения Крос-валидиращ подход, определят поведението на k -NN класификационните модели с установени предимства пред SVMs, независимо от използваните. Според представеният анализ е налице известно предимство пред SVM на k -NN, където най-висока адекватност се присвоява на базисното Евклидово разстояние.

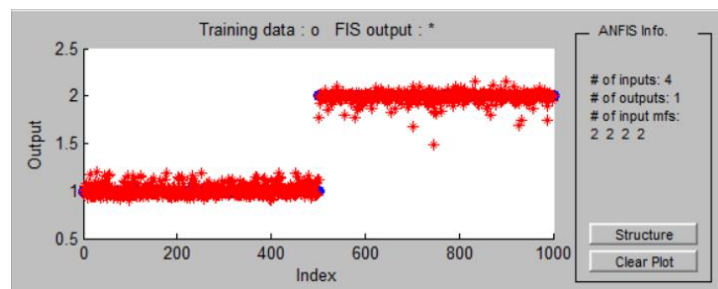
2.2. Изследване на невронно-размити класификатори за идентификация на Гаусов бял шум и Периодичен случаен шум на основата на FFT спектрални характеристики.

Направена е оценка на ефективността на невронно-размити класификатори при идентични и функции на принадлежност на входните променливи за процесите на обучение посредством „Алгоритъм на обратно разпространение на грешката“ и „Хибриден подход с алгоритми на най-малките квадрати и обратно разпространение на грешката“. Особен акцент е поставен на показателя “Root Mean Square Error (RMSE)”. Опитно е регистрирано предимство на

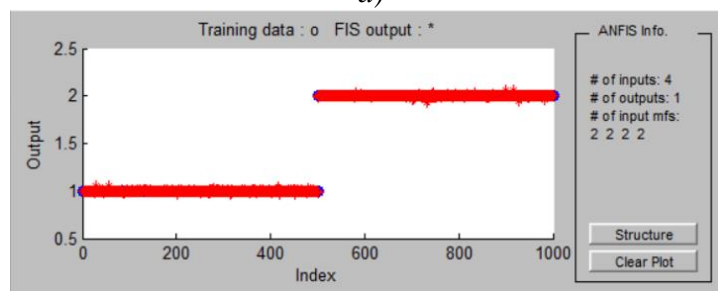
хибридният подход в сравнение с метода на обратно разпространение на грешката. Минимални показания на критерия за ефективност 0.020000 и 0.007005 са намерени при фиксиран „dsigmf“ за „Подход с обратно разпространение на грешката“ и „камбановиден“ тип на функциите на принадлежност на приложените спектрални информативни признаци за Hybrid обучение. Най-високи нива на грешката 0.048338 и 0.058642 са установени при функции на принадлежност „trimf“ в рамките на двата обучаващи подхода.



Фиг. 2.21. Невронно-размит класификатор за идентификация на GWN и PRN



а)



б)

Фиг. 2.24. Тестови процес на невронно-размити класификатори с а) най-ниски и б) най-добри качествени показатели при Hybrid и Backpropagation алгоритми на обучение

Коректността на избор на невронно-размития модел на фиг. 2.21 с „камбановидна“ функция на принадлежност на входните променливи ясно се потвърждават от значително по-високата степен на групиране на еталоните от верификация на фиг. 2.24, потвърждаващо се значително по-ниската степен на отчетената грешка 0.0069997. Спрямо посочената невронно-размита структура е постигната класификационна точност, съответстваща на 100.0 %.

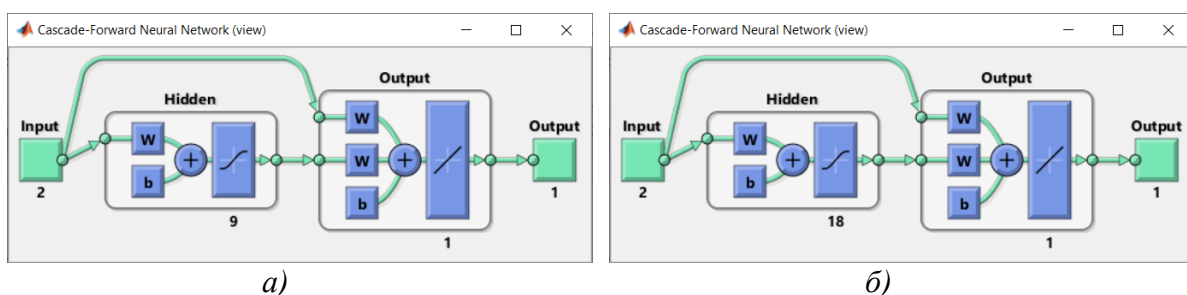
2.3. Синтез на CFNNs структури за прогнозен амплитуден анализ на Периодичен случаен шум на основата на извлечени статистически индикатори „RMS“ и „SD“.

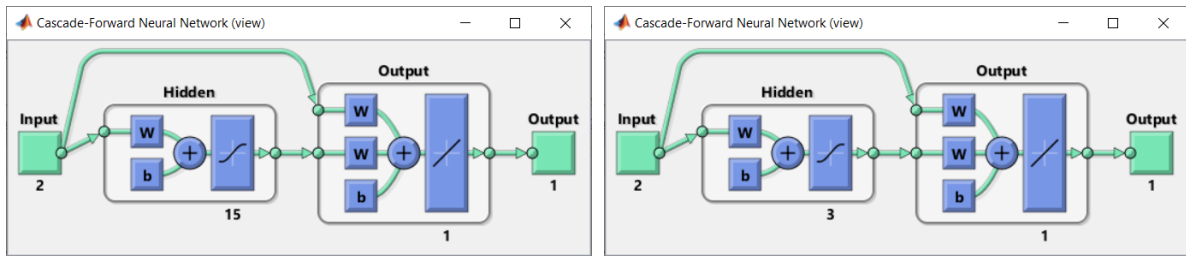
Направен анализ на качеството на „Cascade-Forward Neural Networks (CFNNs)“ прогнозни модели при двойка входни аналитични индекси “RMS и SD”, извлечени при първична обработка на Синусоидални, Правоъгълни, Триъгълни и Трионообразни сигнали с добавен Периодичен случаен шум. Според таблица 2.8 относно първите два тестови сигнали – Sine и Square сигнали, с присъствие на PRN е наблюдавани най-висока стойности на грешката от порядъка на “10⁻⁴”, респективно 1.7728e-04 и 1.5808e-04 за архитектури със съдържание на 18 и 8 неврона в междинните мрежови слоеве. Докато при процедурите във връзка на амплитудните вариации на Периодичен случаен шум, добавен към Triangle и Sawtooth сигнали, са отчетени по-добри максимални нива на MSE при съответна степен “10⁻⁵”. Тук са постигнати най-високи нива на грешката, MSE = 2.3040e-05 за CFNN при включване на 17 и MSE = 2.4337e-05 за прогнозен модел със съдържание на 16 скрити структурни неврона.

Таблица 2.8. Резултати при подбор на CFNNs за прогнозен анализ на амплитудните нива на PRN към аналогови и цифрови сигнали

Скрити неврона	MSE за Sine с PRN	MSE при Square с PRN	MSE относно Triangle с PRN	MSE при Sawtooth с PRN
3	4.5050e-07	1.6201e-05	2.7307e-06	3.7010e-07
4	1.1481e-04	2.2175e-05	6.1414e-07	1.7660e-06
5	1.4655e-06	5.1361e-06	8.5244e-07	1.6841e-06
6	1.2891e-06	1.3236e-05	2.4050e-06	3.7491e-06
7	5.0314e-06	1.4740e-05	2.5050e-06	4.2153e-07
8	3.2265e-07	1.5808e-04	2.6565e-06	9.8543e-06
9	2.5875e-07	4.2886e-06	1.4770e-07	4.9312e-06
10	3.6273e-07	2.6949e-06	1.2471e-05	5.1917e-06
11	4.9705e-07	6.9126e-07	2.0770e-06	3.3378e-06
12	1.9296e-06	7.1500e-06	3.1081e-06	7.4629e-06
13	7.5288e-07	6.0046e-06	3.4242e-06	9.9613e-06
14	2.4697e-06	6.2643e-06	2.2480e-07	9.4128e-07
15	1.6811e-05	3.7231e-06	1.1510e-07	1.3057e-05
16	1.2469e-06	2.0272e-05	9.6236e-06	2.4337e-05
17	1.6221e-06	2.1184e-05	2.3040e-05	4.6365e-06
18	1.7728e-04	6.6411e-07	1.0771e-06	5.7096e-06
19	9.1631e-06	1.3891e-04	3.7065e-06	5.0842e-06
20	3.0148e-06	2.0975e-04	1.2703e-05	2.4016e-05

Относно поставената задача за прогнозен анализ бяха определени като най-подходящи крайните CFNNs с включени 9, 18, 15 и 3 изчислителни единици в скритите слоеве, показани на фиг. 2.31. Невронните архитектури относно обектните PRN въздействия се отличават с постигнати по-добри в низходяща степен нива на базисният MSE показател в сравнение с получените при сигнали при UWN. В последователен ред бяха регистрирани удовлетворяващи стойности MSE = 2.5875e-07, MSE = 6.6411e-07, MSE = 1.1510e-07 и MSE = 3.7010e-07 за изследваните Sine, Square, Triangle и Sawtooth сигнали с регистрирано присъствие на PRNs.





Фиг. 2.31. Cascade-Forward Neural структури за прогнозен амплитуден анализ при сигнали с влияние на PRN - а) "Sine с PRN", б) "Square с PRN", в) "Triangle с PRN" и г) "Sawtooth с PRN"

2.4. Изследване на FFNNs за прогнозиране на амплитудните изменения на Постоянен бял шум на базата на извлечени статистически индикатори „Min“ и „Max“.

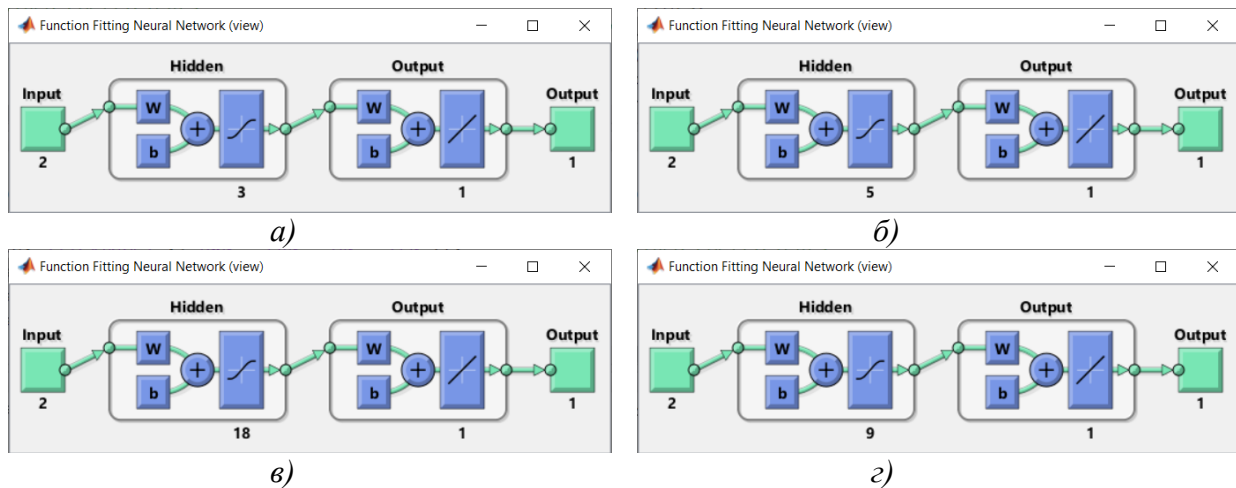
Извършена е селекция на FFNNs при обучение с Bayesian Regularization алгоритъм. Създадени са идентични невронни структури при фиксирана „тангенс-сигмоидална“ и „линейна“ активация на невроните в междинния и изходния слой във фиксиран диапазон от 3 до 20, чиито резултати са обобщени в таблица 2.9.

Таблица 2.9. Изследване на FFNN при BR обучение за прогнозен анализ на амплитудните нива на UWN към Sine и Square, Triangle и Sawtooth сигнали

Скрити неврони	MSE при тестови наблюдения		Скрити неврони	MSE при тестови наблюдения	
	Sine с UWN	Square с UWN		Triangle с UWN	Sawtooth с UWN
3	1.49445e-5	5.01191e-6	3	4.31698e-4	6.57935e-6
4	3.15040e-5	2.26269e-6	4	4.65051e-4	1.55920e-5
5	3.15040e-5	7.69753e-7	5	6.64184e-3	1.87699e-5
6	3.62525e-5	1.34544e-6	6	3.73203e-4	2.29972e-5
7	2.57697e-5	7.03414e-6	7	1.14863e-4	1.14334e-5
8	2.92002e-5	4.79208e-6	8	2.01085e-3	1.16126e-5
9	5.44664e-5	8.07133e-6	9	1.61300e-4	7.32980e-6
10	4.38610e-5	4.62541e-6	10	3.78390e-3	1.41347e-5
11	4.58814e-5	9.06271e-7	11	3.93365e-2	1.13889e-5
12	1.77804e-5	4.76656e-6	12	3.71264e-3	1.93207e-5
13	2.27808e-5	1.75893e-6	13	1.75104e-3	1.04982e-5
14	7.67937e-5	9.01756e-7	14	7.20549e-4	3.08972e-5
15	2.11002e-5	6.76589e-6	15	9.85018e-4	1.85419e-5
16	2.13257e-5	3.32387e-6	16	1.49009e-4	9.21799e-6
17	2.81509e-5	7.71967e-6	17	8.78740e-4	8.00949e-6
18	4.43218e-5	1.57173e-6	18	9.86971e-5	1.26585e-5
19	3.91506e-5	3.59555e-6	19	7.50651e-4	1.13836e-5
20	2.33375e-5	3.40613e-6	20	6.29804e-3	2.96661e-5

В обхвата на проведените дейности беше постигната минимизация на приетия базисен критерий при подбор на модели, както следва:

- $MSE = 1,49445 \cdot 10^{-5}$ при структура с начален брой изчислителни единици относно Sine с UWN;
- $MSE = 7,65753 \cdot 10^{-7}$ при модел с наличие на 5 скрити неврона за Square с UWN;
- $MSE = 1,49445 \cdot 10^{-5}$ при невронна структура с включени 18 междинни неврона във връзка с Triangle с UWN;
- $MSE = 1,49445 \cdot 10^{-6}$ за прогнозен модел с фиксирано количество изчислителни единици в скрития слой относно Sawtooth с UWN.



Фиг. 2.34. FFNN за прогнозен анализ на вариациите на Постоянен бял шум към а) Sine, б) Square, в) Triangle и г) Sawtooth сигнали

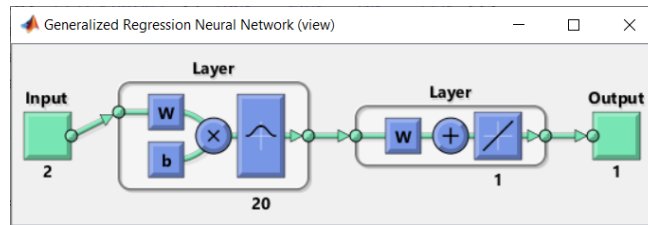
Максимални показания на грешката бяха констатирани при прогнозни невронни структури със съдържание на 14, 9, 11 и 14 скрити неврона в последователен ред на декларираните категории сигнали, дадени на фиг. 2.34.

2.5. Оценка на GRNNs при прогнозен анализ на амплитудните вариации на Периодичен случаен шум при извлечени статистически метрики „SD“ и „RMS“.

Следвайки подобна аналогия, както приложената при подбор на „Feed-Forward Neural Networks (FFNNs)“, е проведена селекция на „Generalized Regression Neural Networks (GRNN)“ за регресионна диагностика в таблица 2.10. Обща тенденция за всички анализирани GRNN модели е наблюдаваното постепенно, преминаващо към експоненциално, нарастване на критерия MSE. Независимо от категорията на изследваните сигнали беше приета минимална „широчина“ на Кернел функцията във вторите структурни слоеве. Тук достигнатите минимални MSE нива се равняват на $9.5632 \cdot 10^{-7}$, $9.1807 \cdot 10^{-7}$, $1.4335 \cdot 10^{-6}$ и $7.6517 \cdot 10^{-7}$ за сигнали със синусоидална, правоъгълна, триъгълна и трионообразна форма с отразено влияние на PRN.

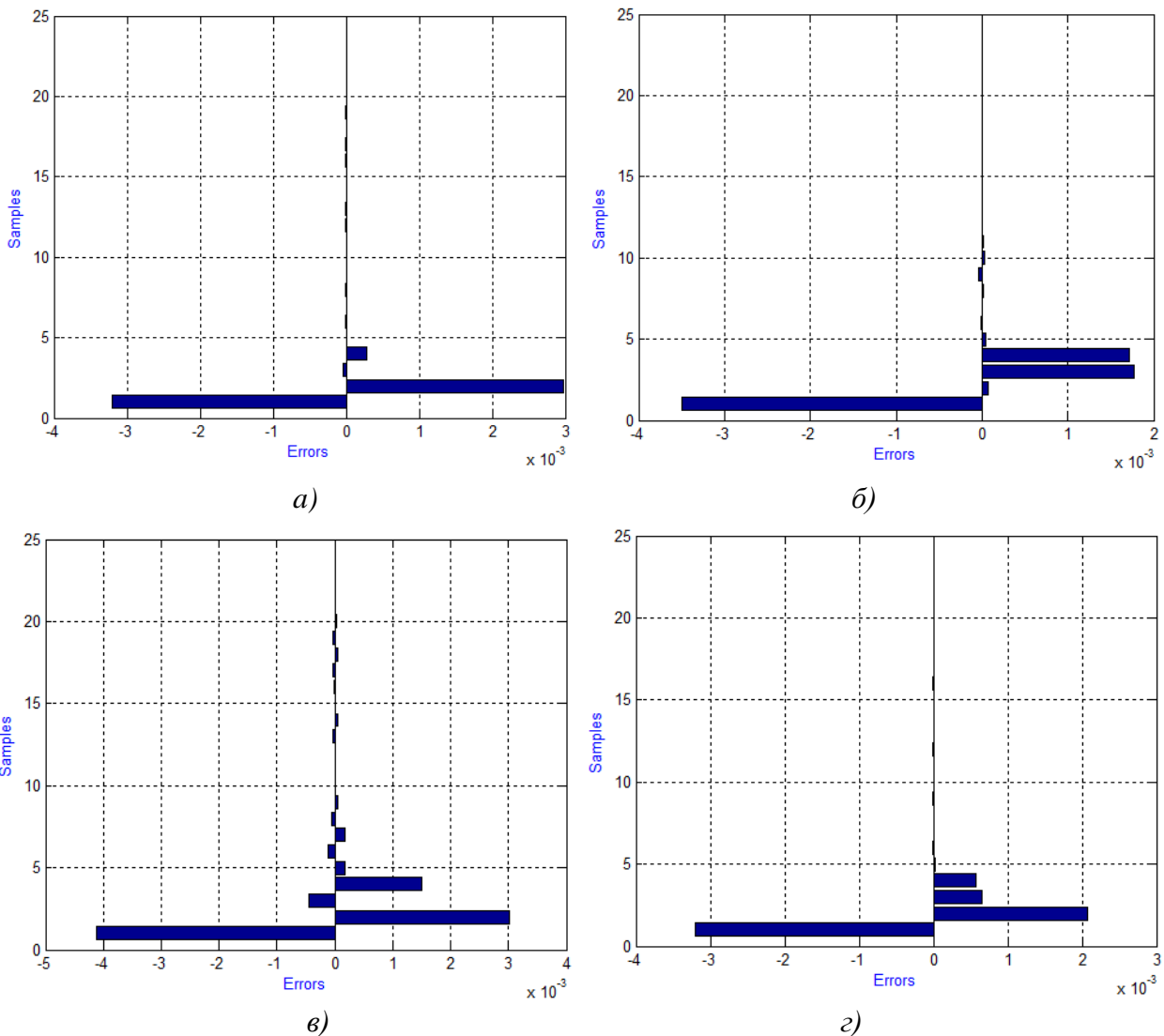
Таблица 2.10. Анализ на GRNN архитектури за регресионно моделиране относно амплитудните изменения Sine и Square, Triangle и Sawtooth сигнали с насложен PRN

Spread	MSE при наблюдения в тестовия поднабор		Spread	MSE при наблюдения в тестовия поднабор	
	Sine с PRN	Square с PRN		Triangle с PRN	Sawtooth с PRN
0.050	9.5632e-07	9.1807e-07	0.050	1.4335e-06	7.6517e-07
0.075	1.4661e-06	3.1523e-06	0.075	2.2448e-06	1.6920e-06
0.100	2.1596e-06	6.3159e-06	0.100	4.0874e-06	3.1820e-06
0.125	3.3734e-06	9.6002e-06	0.125	6.5746e-06	5.1112e-06
0.150	4.9466e-06	1.2671e-05	0.150	9.4332e-06	7.2785e-06
0.175	6.7363e-06	1.5503e-05	0.175	1.2665e-05	9.5648e-06
0.200	8.6902e-06	1.8148e-05	0.200	1.6357e-05	1.1926e-05
0.225	1.0815e-05	2.0757e-05	0.225	2.0595e-05	1.4378e-05
0.250	1.3142e-05	2.3498e-05	0.250	2.5440e-05	1.6978e-05
0.275	1.5706e-05	2.6492e-05	0.275	3.0935e-05	1.9793e-05
0.300	1.8542e-05	2.9806e-05	0.300	3.7121e-05	2.2876e-05
0.325	2.1672e-05	3.3473e-05	0.325	4.4038e-05	2.6264e-05
0.350	2.5115e-05	3.7514e-05	0.350	5.1730e-05	2.9984e-05
0.375	2.8886e-05	4.1951e-05	0.375	6.0241e-05	3.4054e-05
0.400	3.2999e-05	4.6810e-05	0.400	6.9618e-05	3.8494e-05
0.425	3.7468e-05	5.2122e-05	0.425	7.9910e-05	4.3324e-05
0.450	4.2306e-05	5.7924e-05	0.450	9.1170e-05	4.8565e-05
0.475	4.7529e-05	6.4249e-05	0.475	1.0345e-04	5.4240e-05
0.500	5.3153e-05	7.1132e-05	0.500	1.1683e-04	6.0373e-05



Фиг. 2.37. Модел на GRNN структури за регресионно моделиране във връзка с вариациите на PRN към Sine, Square, Triangle и Sawtooth сигнали

Във връзка с допълнителен анализ на функционалността на модела на фиг. 2.37 е построено семейство линейни регресионни зависимости за мрежовите изходи. Констатирано е пълно припокриване на опитната и теоретичната линия на регресия за всички групи сигнали. Получени са „Корелационни коефициенти“ $R = 0.99986$ при Sine и Square, $R = 0.99978$ за Triangle и $R = 0.99989$ при Sawtooth сигнал с присъствие на PRN.



Фиг. 2.39. Диаграма на разликите между теоретичните и прогнозните нива на амплитудните вариации на PRN към а) Sine, б) Square, в) Triangle и г) Sawtooth сигнали с приложение на GRNNs

Построени са Диаграми на остатъците на фиг. 2.39 по отношение на тестовите сигнали е наблюдавано лимитирано изменение на мрежовите грешки, , респективно „-0.0032 до 0.0030“, „-0.0035 до 0.0018“, „-0.0041 до 0.0030“ и „-0.0032 до 0.0021“. Като обща тенденция се наблюдава пренебрежимо ниските нива на „Остатъците“. Изключение е констатирано за

първи, втори, трети и четвърти еталон от верификация на GRNNs за прогнозен анализ на амплитудните вариации на Постоянен бял към аналогови и цифрови сигнали.

2.6. Изводи към Втора глава.

- Доказана е ефективността на C-SVM, Nu-SVM и k-NN методи на Машинно обучение за идентификация на Гаусов бял шум и Периодичен случен шум в имитирана комуникационна преносна среда с установено предимство на качеството на k-NN;
- Потвърдена е ефективността на избран класификационен модел с най-висока надеждност PNN пред Невронно-размити класификатори, CFNNs, Вероятностни изкуствени невронни мрежи и Алгоритъм на Naïve Bayes за разпознаване на GWN и PRN;
- Разработен е подход за идентификация на GWN и PRN въздействия чрез Метод на опорните вектори, ANFIS при хибриден алгоритъм на обучение на основата на получени FFT спектрални характеристики;
- Въз основа на анализ е доказана ефективността на Обобщените регресионни невронни мрежи за разпознаване на GWN и PRN шумове на основата на получени FFT спектрални характеристики при различни прозоречни функции;
- Постигната е минимизация на MSE при предложен подход за прогнозиране на амплитудни нива на UWN и PRN, насложени към аналогови и цифрови сигнали, с предварителна статистическа обработка и модели за регресионно моделиране на основата на CFNNs;
- Доказана е ефективността на FFNN и GRNN модели за прогнозиране на амплитудните вариации на UWN и PRN към аналогови и цифрови сигнали с ограничаване на диапазоните на изменение между теоретичните и прогнозните нива на смущенията.

Глава III: Прогнозен анализ на производителността на преносната среда в симулационни ИКТ-базирани инфраструктури

3.1. Извеждане на аналитични модели за прогнозиране на индекси на производителност на преносната среда на основата на Регресионен анализ.

Разгледана е задачата за извеждане на регресионни модели на избрани индекси на производителност на преносната среда “System Throughput (ST)” и “System Response Time (SRT)” при входящ индекс “System Demand (SD)” спрямо моделираната ИКТ инфраструктура посредством MoM алгоритъм. Приложената регресионна диагностика се състои в оценка на адекватността на математически регресионни модели от линеен (3.1) и полиномиален вид (3.2) с продукт STATISTICA като са въведени следни показатели:

- x_1 – управляем фактор SD, s;
- y_1 и y_2 – изходни зависими променливи ST и SRT.

$$y = b_0 + b_1x_1 \quad (3.1)$$

$$y = b_0 + b_1x_1 + b_{11}x_1^2 \quad (3.2)$$

Проверката на пригодността на теоретични модели (3.1) и (3.2) на фиг. 3.3 и фиг. 3.4 показва високи стойности на приетата базисна метрика за избор на модел „Коефициент на определеност R^2 “. Анализът спрямо линейния модел за трафични индекси ST и SRT се отъждествява с постигнати $R^2 = 0.99573178$ и $R^2 = 0.99395010$ на фиг. 3.3. Регресионните процеси спрямо полиномиалния модел на фиг. 3.4 се характеризират с минимално превишаване на наблюдавания R^2 , респективно $R^2 = 0.99999789$ при y_1 и $R^2 = 0.99998982$ за фактор y_2 . Всички опитни коефициенти на регресия b_i са оценени като значими, спрямо прието равнище на значимост $\alpha = 0.05$.

Regression Summary for Dependent Variable: y (system_throughput)						
R= ,99786361 R ² = ,99573178 Adjusted R ² = ,99564286						
F(1,48)=11198, p<0,0000 Std.Error of estimate: ,03997						
N=50	b*	Std.Err. of b*	b	Std.Err. of b	t(48)	p-value
Intercept			45,77744	0,029338	1560,353	0,00
x1	-0,997864	0,009430	-3,35135	0,031670	-105,820	0,00

a)

Regression Summary for Dependent Variable: y (system_response_time)						
R= ,99697046 R ² = ,99395010 Adjusted R ² = ,99382406						
F(1,48)=7886,0 p<0,0000 Std.Error of estimate: ,00026						
N=50	b*	Std.Err. of b*	b	Std.Err. of b	t(48)	p-value
Intercept			0,217351	0,000192	1132,956	0,00
x1	0,996970	0,011227	0,018391	0,000207	88,803	0,00

б)

Фиг. 3.3. Регресионни резултати при проверка на адекватността на модел (3.1) за индекс на производителност а) Пропускателна способност и б) Комплексно време за чакане и обслужване

Regression Summary for Dependent Variable: y (system_throughput)						
R= ,99999894 R ² = ,99999789 Adjusted R ² = ,99999780						
F(2,47)=1113E4 p<0,0000 Std.Error of estimate: ,00090						
N=50	b*	Std.Err. of b*	b	Std.Err. of b	t(47)	p-value
Intercept			44,68508	0,003606	12393,19	0,00
x1	-0,253550	0,002425	-0,85155	0,008143	-104,57	0,00
x11	-0,747174	0,002425	-1,37503	0,004462	-308,16	0,00

a)

Regression Summary for Dependent Variable: y (system_response_time)						
R= ,99999491 R ² = ,99998982 Adjusted R ² = ,99998939						
F(2,47)=2309E3 p<0,0000 Std Error of estimate: ,00001						
N=50	b*	Std.Err. of b*	b	Std.Err. of b	t(47)	p-value
Intercept			0,224490	0,000043	5162,663	0,000000
x1	0,111348	0,005324	0,002054	0,000098	20,915	0,000000
x11	0,889026	0,005324	0,008986	0,000054	166,991	0,000000

б)

Фиг. 3.4. Регресионни резултати при оценка на пригодността на аналитичен модел (3.2) за индекс на производителност а) ST и б) SRT

Изведени са крайни регресионни модели за прогнозиране на потенциалните индекси на производителност за имитационно моделираната ИКТ инфраструктура на основата на модел от нулева (3.3) и (3.4) и от втора степен (3.5) и (3.6).

$$y_1 = 45.77744 - 3.35135x_1 \quad (3.3)$$

$$y_2 = 0.217351 + 0.018391x_1 \quad (3.4)$$

$$y_1 = 44.68508 - 0.85155x_1 - 1.37503x_1^2 \quad (3.5)$$

$$y_2 = 0.224490 + 0.002054x_1 + 0.008986x_1^2 \quad (3.6)$$

Изследвана е обратната регресионна задача за получаване на модел за прогнозиране на вариациите на специфичен индекс на производителността на преносната среда „System Demand или „у““ при входящи управляеми фактори ST или „x₁“ и SRT или „x₂“. Резултатите от извършената линейна регресия на фиг. 3.7 показват постигната висока стойност на Коефициента на определеност R² = 0.99994031, показваща, че 99.994031 % от изменението на параметър „у“ се дължи на въздействието на управляемите фактори „x₁“ и „x₂“. Всички изходни коефициенти на регресия, както следва b₀ = 144.905, b₁ = -1.835 (отразяващ влиянието

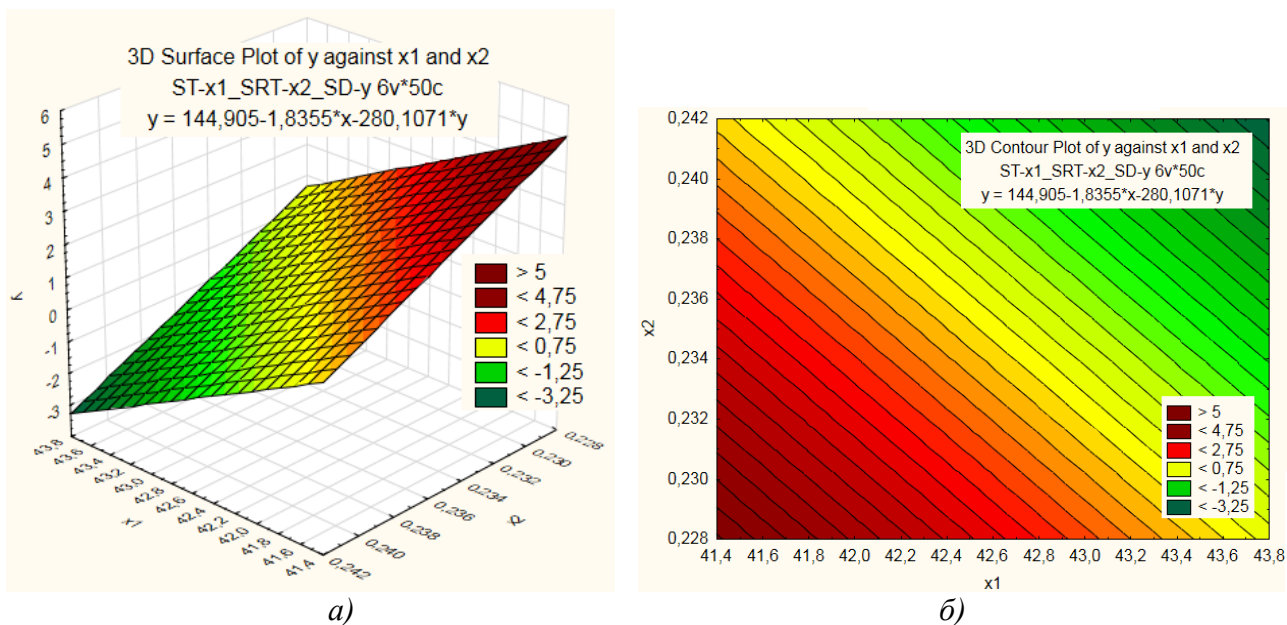
на x_1) и $b_2 = -280.170$ (свързан с оценка на въздействието на x_2), се определят като значими спрямо констатираната вероятност “p level“ много по-малка от прието равнище на значимост $\alpha = 0.05$.

Regression Summary for Dependent Variable: y (QoS)						
R= ,99997016 R?= ,99994031 Adjusted R?= ,99993777						
F(2,47)=3937E2 p<0,0000 Std.Error of estimate: ,00142						
N=50	b*	Std.Err. of b*	b	Std.Err. of b	t(47)	p-value
Intercept			144,905	2,280890	63,5300	0,000000
x1	-6,16449	0,089758	-1,835	0,026725	-68,6791	0,000000
x2	-5,16703	0,089758	-280,107	4,865810	-57,5664	0,000000

Фиг. 3.7. Резултати от линейна регресия при анализ на адекватността на модел (3.7) за прогнозиране на индекс на производителност SD

Крайният вид на синтезираният аналитичен модел за прогнозиране на вариациите на „Комплексна метрика на постъпванията и обслужването на пакети“ следва да бъде записан последният начин:

$$y = 144.905 - 1.835x_1 - 280.170x_2 \quad (3.8)$$



Фиг. 3.9. Повърхнина на отклика а) и Линии на еднакъв отклик $y = f(x_1, x_2)$ б) за индекс на производителност на преносната среда SD

Генерирани са „Повърхнината на отклика“ $y = f(x_1, x_2)$ и полученото сечение на повърхнината с равнина, успоредна на равнината x_1Ox_2 , т.е. „Линиите на еднакъв отклик“ $y = f(x_1, x_2)$, показани на фиг. 3.9. Посочените зависимости на отклик потвърждават, че значителен дял от възходящите изменения на изходен параметър SD се получават при „средни и високи нива на фактор x_1 “ и „ниски и средни нива на фактор x_2 “.

3.2. Оптимизационна задача за търсене на екстремум относно минимизация на индекс на производителност „Комплексно време за чакане в опашката и обслужване“.

Един от съществените аспекти за подобряване на качеството на обслужване на информационно-комуникационните звена се свързва с постигане на изискване за оптималност чрез минимизиране на комплексния индекс на производителност “SRT”. В тази връзка могат да бъдат приложени алгоритми и методи за оптимизация за установяване на оптимум, удовлетворяващ постигането на посочения критерий. В рамките на конкретния случай е разгледан регресионен модел за калкулиране на прогнозни нива на индекс SRT спрямо

входящо въздействие на управляем фактор SD, [s]. По дефиниция от теория на оптимизацията, изведеният аналитичен модел се приема за "целева функция или $f(x)$ " обект на нелинейна минимизация.

В резултат от проведената оптимизация с последователно приложение на "Interior point", "SQP" и "Active-set" алгоритми беше констатирано идентично решение „-0.1143“, представляващо условие за оптимум, изразено в „кодирани“ стойности на заложения управляем фактор SD. Намереното оптимално решение гарантира постигане на изисквани минимални нива на целевия трафичен индекс на производителност SRT в приблизителен порядък „0.224373“ за всеки от оптимизационните подходи. Спрямо съвпадащите положителни степени на индекс "exitflag" при „Interior point“, "SQP" и "Active-set" алгоритмите е потвърдена коректността на реализираните оптимизационни процеси.

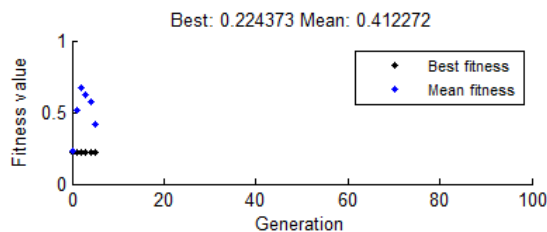
С оглед на разширяване на обхвата на методиката за нелинейна оптимизация с ограничения на приложените подходи „Interior point“, "SQP" и "Active-set" са проведени процедури по верификация за търсене на други оптимизационни инструменти, потвърждаващи аналогична или близка ефективност. Въведена е оптимизация с изследване на „Genetic Algorithm (GA)“ в качеството на неконвенционален подход за нелинейна минимизация на трафичен индекс на производителност „SRT“ спрямо управляем фактор „SD“.

Получена е спецификация на оптимизационния процес относно приложението на GA, както и намерената най-добра (най-ниска) „Best: 0.224373“ и средна стойност „Mean: 0.412272“ на целевата функция – обект на минимизация, е представена на фиг. 3.16. В рамките на процедурите по достигане на оптимално решение са генерирани пет популации. Първоначално създаденото случайно поколение се състои от 2650, докато последната популация, представлява съвкупност от 13050 точки в оптимизационното параметрично пространство.

```

GA_SRT.m x +
1 function [x,fval,exitflag,output,population,score] = GA_SRT(nvars)
2 %% This is an auto generated MATLAB file from Optimization Tool.
3
4 %% Start with the default options
5 options = gaoptimset;
6 %% Modify options setting
7 options = gaoptimset(options,'Display','iter');
8 options = gaoptimset(options,'PlotFcns',{ @gaplotbestf @gaplotbestindiv @gaplotc
9 [x,fval,exitflag,output,population,score] = ...
10 ga(@SRT,nvars,[],[],[],[],[],[],@ogrSRT,[],options);
  
```

а)



б)

Command Window				
Generation	f-count	Best f(x)	max constraint	Stall Generations
1	2650	0.224439	0	0
2	5250	0.224381	0	0
3	7850	0.224375	0	0
4	10450	0.224375	0	1
5	13050	0.224373	0	0

в)

Фиг. 3.16. Спецификация а), достигати количествени индикатори б) и ход на оптимизационния процес в) с приложение на Genetic Algorithm

Намереното оптимално решение предвижда „кодирана стойност: -0.1208“, която е по-ниска спрямо установеното кодирано ниво „ $x = -0.1143$ “, за управляем фактор SD. Докато по отношение на констатираната „най-добра стойност на Fitness function: 0.2244“, е налице съвпадение с изведената в хода на минимизацията „най-добра натурална стойност на трафичен индекс SRT“ при подходи „Interior point“, „SQP“ и „Active-set“. Потвърдена е коректността на извършените оптимизационни процедури – свидетелство, за което е положителната стойност на променливата „exitflag“ на фиг. 3.17. Включени са петдесет функционални оценки на множеството от точки в последната генерирана популация, където е установена най-малката стойност на Fitness функция, която клони към оптимално решение.

Field	Value	Min	Max
x	-0.1208	-0.1208	-0.1208
fval	0.2244	0.2244	0.2244
exitflag	1	1	1
output	1x1 struct		
population	50x1 double	-13.4622	15.0550
score	50x1 double	0.2244	2.2921

Фиг. 3.17. Изходни показатели от минимизация на целевата Fitness функция с приложение на Genetic Algorithm

Извършената верификация показва ефективност на GA подход, успешно съпоставима с постигната при „Interior point“, „SQP“ и „Active-set“. Достигнати са оптимуми („ $x = -0.1143$ “ и „ $x = -0.1208$ “), при които се гарантира теоретично минимално ниво на трафичен индекс System Response Time. Прогнозната стойност от минимизация на SRT се доближава до получените най-ниски „Observed values“ при симулационно моделиране. Този факт може да бъде маркиран като индикация за висока ефективност на всички приложени подходи във връзка с разглежданата оптимизационна задача. Постигнати са целите на оптимизацията за гарантиране на SRT нива, удовлетворяващи изискването за повишаване на капацитета на обслужване на системни потребители в обсега на имитираната ИКТ инфраструктура.

3.3. Прогнозиране за индекси на производителност на преносната среда с FFNNs при две входни въздействия и различни обучаващи алгоритми.

Проведени са процедури по изследване и подбор на FFNNs при анализ на приложимостта на следните обучаващи алгоритми: Levenberg-Marquardt (LM); Bayesian Regularization (BR) и Scaled Conjugate Gradient (SCG), обобщени в таблица 3.4 и таблица 3.5, при приет базисен критерий (MSE). Въведени са следните входни и изходни параметри в процесите на синтез на прогнозни модели:

- Входни променливи – 1) SD, и 2) Стойностите на телетрафичния показател System demand, повдигнати на втора степен;
- Изходни параметри, съответно ST и SRT.

Относно първият заложен трафичен индекс бяха регистрирани максимални показания на грешката, респективно $MSE = 2.02598e-4$, $MSE = 1.64856e-10$ и $MSE = 2.43299e-3$ последователно при LM, BR и SCG обучение за структури с 9, 13 и 10 междинни изчислителни единици. Удовлетворяващите минимални MSEs са получени при FFNN структури със зададени 4 ($MSE = 3.17825e-10$), 10 ($MSE = 7.79211e-11$), и 13 ($MSE = 3.54521e-5$), скрити неврона. Според обобщените резултати избраният невронен модел с BR алгоритъм се оценява с най-висока степен на адекватност за прогнозиране на вариациите на трафична характеристика „Пропускателна способност“. Сравнявайки приложените обучаващи алгоритми беше установено, че SCG подход се характеризира с най-ниска ефективност спрямо конкретния анализиран параметър на трафика.

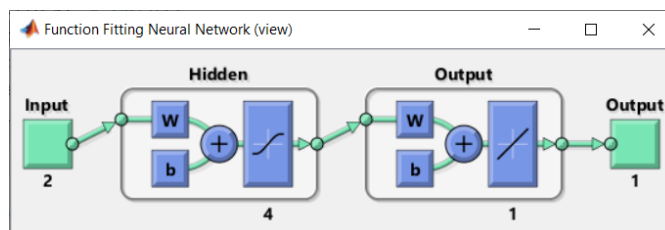
Таблица 3.4. Резултати при изследване на невронни мрежи за прогнозиране на индекс System Throughput

Скрити неврони	LM алгоритъм	BR алгоритъм	SCG алгоритъм
	MSE при тестване	MSE при обучение	MSE при обучение
3	1.14099e-9	1.57589e-10	7.01570e-5
4	3.17825e-10	1.19221e-10	7.82394e-5
5	2.09195e-9	1.20101e-10	6.55992e-5
6	2.09897e-6	1.25709e-10	7.80464e-5
7	8.42537e-9	1.53955e-10	8.78858e-4
8	5.79434e-7	9.02992e-11	9.09759e-4
9	2.02598e-4	1.41492e-10	4.08480e-4
10	2.67695e-7	7.79211e-11	2.43299e-3
11	1.00936e-6	1.21158e-10	1.52216e-3
12	1.89889e-4	1.19690e-10	1.70738e-4
13	3.52341e-6	1.64856e-10	3.54521e-5

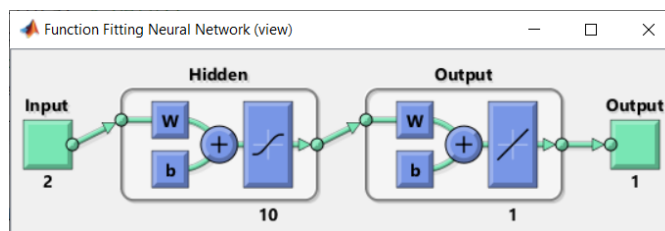
Спрямо синтеза на прогнозни модели при вторият заложен телетрафичен показател, обхващащ „Времената на престой в опашката и обработка на постъпилите заявки“, относително беше наблюдавана известна тенденция на по-малка ефективност при SCG в сравнение с останалите използвани алгоритми. Но съблюдавайки изходните степени на критерия за качество MSE, тук се отличава изразена адаптивност на SCG спрямо подбора на невронни модели при “Системната пропускателна способност”, показани на фиг. 3.25.

Таблица 3.5. Резултати при изследване на невронни мрежи за прогнозиране на индекс System Response Time

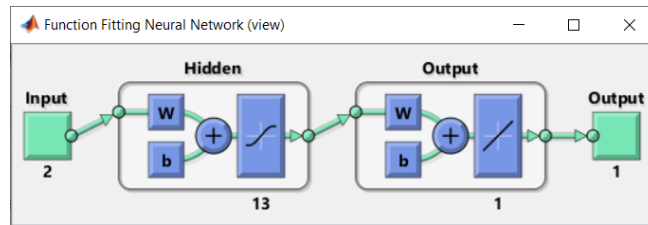
Скрити неврони	LM алгоритъм	BR алгоритъм	SCG алгоритъм
	MSE при тестване	MSE при обучение	MSE при обучение
3	5.00144e-9	1.44758e-9	7.20825e-9
4	6.88759e-10	1.06745e-9	4.29007e-7
5	5.46010e-8	2.18784e-9	3.83399e-7
6	2.51603e-10	5.80534e-9	2.16496e-7
7	5.91742e-9	2.67628e-9	9.91093e-8
8	5.50524e-9	6.38057e-10	2.55897e-7
9	5.41969e-11	1.65099e-9	3.60012e-8
10	7.73573e-9	1.15110e-9	2.06796e-7
11	1.20022e-9	4.25863e-10	1.38354e-7
12	5.58102e-10	6.71894e-10	1.37807e-7
13	1.36464e-11	1.61493e-9	6.66151e-8



a)

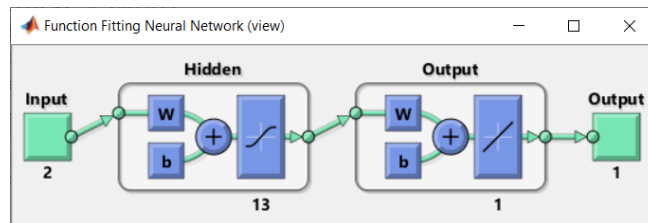


б)

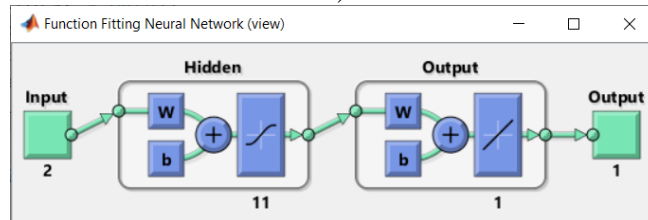


в)

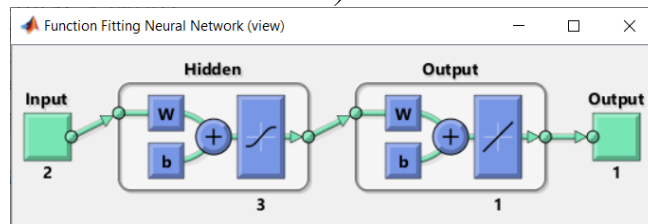
Фиг. 3.25. Селектирани прогнозни FFNNs за индекс на производителност System Throughput при а) LM, б) BR и в) SCG обучаващ алгоритъм



а)



б)



в)

Фиг. 3.26. Селектирани прогнозни FFNNs за индекс на производителност System Response Time при а) LM, б) BR и в) SCG обучаващ алгоритъм

Най-големите грешки $MSE = 5.46010e-8$, $MSE = 5.80534e-9$ и $MSE = 4.29007e-7$ бяха констатирани при FFNNs със съдържание на 5, 6 и 4 скрити неврона, респективно при LM, BR и SCG алгоритми. Постигачи изискването за минимизация на грешката са невронни архитектури при наличие на 13 за LM, 11 при BR и 3 структурни междинни неврона относно SCG обучение – $MSE = 1.36464e-11$, $MSE = 4.25863e-10$ и $MSE = 7.20825e-9$, представени на фиг. 3.26. Съпоставяйки избраните крайни FFNN модели за прогнозиране на Комплексното време за престой в опашката и обслужване се откроява предимството на LM алгоритъм на обучение пред BR и SCG подход.

3.4. Прогнозен анализ на индекси на производителност на преносната среда при комбинация от две входни въздействия с използване на RBNN, GRNN и CFNN.

Въведена е концепция за регресионно моделиране с обучение и оценка на ефективността на Radial Basis Neural Networks (RBNNs). Анализирани са регистрираните изходни показатели MSE, MAE и RMSE при оценяване на RBNNs, едновременно обобщени за трафични индекси ST и SRT в таблица 3.7. Проведеният анализ на поведението на указаните критерии обхваща RBNNs при стъпково нарастване на SI в диапазона 0.1 – 0.9. Установени са удовлетворяващи целите на изследването степени:

- „ 10^{-11} “ за средноквадратичната и „ 10^{-6} “ при средната абсолютна грешка и средното квадратично отклонение;
- „ 10^{-15} “ при MSE и „ 10^{-8} “ за MAE и RMSE;

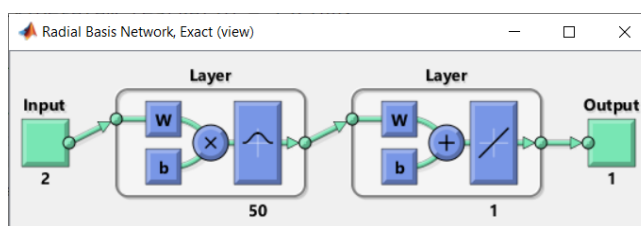
на приложените входни данни при обучаващите процедури. Регистрираните стойности на MSE, MAE и RMSE се характеризират с експоненциално нарастване до нива на $SI = 0.4$. Наблюдавано е известно установяване на количествените критерии с вариращи колебания:

- „ $8.1400e-11$ до $8.9000e-11$ “ при MSE, „ $7.7000e-6$ до $8.2000e-6$ “ за MAE и „ $9.0000e-6$ до $9.4200e-6$ “ относно RMSE при невронен подбор за ST;
- „ $2.5600e-15$ до $2.7100e-15$ “ относно MSE, „ $4.3500e-8$ до $4.5100e-8$ “ за MAE и „ $5.0000e-8$ до $5.2100e-8$ “ относно RMSE при избор на прогнозен модел за SRT;

до крайна широчина на радиално базисната функция $SI = 0.9$. Вследствие на извършените дейности са селектирани идентични RBNN архитектури на фиг. 3.29 за прогнозиране на индекси на производителност ST и SRT при минимална заложена широчина на радиално базисната функция $SI = 0.1$, в случаите на постигане на $MSE = 7.7309e-12$ и $MSE = 2.2659e-16$.

Таблица 3.7. RBNN при прогнозиране на изменението на индекси на производителност ST и SRT при две входни променливи

SI	MSE	MAE	RMSE	MSE	MAE	RMSE
	Пропускателна способност			Комплексно време за чакане и обслужване		
0.10	7.7309e-12	2.0281e-06	2.7805e-06	2.2659e-16	1.0985e-08	1.5053e-08
0.15	2.3484e-11	4.0773e-06	4.8460e-06	7.0014e-16	2.2364e-08	2.6460e-08
0.20	5.9231e-11	6.0761e-06	7.6962e-06	1.7846e-15	3.3327e-08	4.2245e-08
0.25	6.4694e-11	6.2573e-06	8.0433e-06	1.7846e-15	3.3327e-08	4.2245e-08
0.30	7.1730e-11	7.0615e-06	8.4693e-06	2.1609e-15	3.8694e-08	4.6485e-08
0.35	7.7241e-11	7.3046e-06	8.7887e-06	2.3329e-15	4.0039e-08	4.8300e-08
0.40	8.1960e-11	7.7676e-06	9.0532e-06	2.4844e-15	4.2629e-08	4.9844e-08
0.45	8.1520e-11	7.7370e-06	9.0288e-06	2.4698e-15	4.2454e-08	4.9697e-08
0.50	8.4402e-11	7.9536e-06	9.1870e-06	2.5663e-15	4.3712e-08	5.0659e-08
0.55	8.4309e-11	7.9503e-06	9.1820e-06	2.5637e-15	4.3698e-08	5.0633e-08
0.60	8.4188e-11	7.9454e-06	9.1754e-06	2.5581e-15	4.3630e-08	5.0578e-08
0.65	8.6694e-11	8.0451e-06	9.3110e-06	2.6415e-15	4.4207e-08	5.1396e-08
0.70	8.8591e-11	8.1885e-06	9.4123e-06	2.7021e-15	4.5046e-08	5.1981e-08
0.75	8.8577e-11	8.1878e-06	9.4115e-06	2.7017e-15	4.5041e-08	5.1978e-08
0.80	8.8570e-11	8.1875e-06	9.4112e-06	2.7010e-15	4.5034e-08	5.1971e-08
0.85	8.8576e-11	8.1883e-06	9.4115e-06	2.7006e-15	4.5010e-08	5.1968e-08
0.90	8.8721e-11	8.1890e-06	9.4192e-06	2.7064e-15	4.5021e-08	5.2023e-08

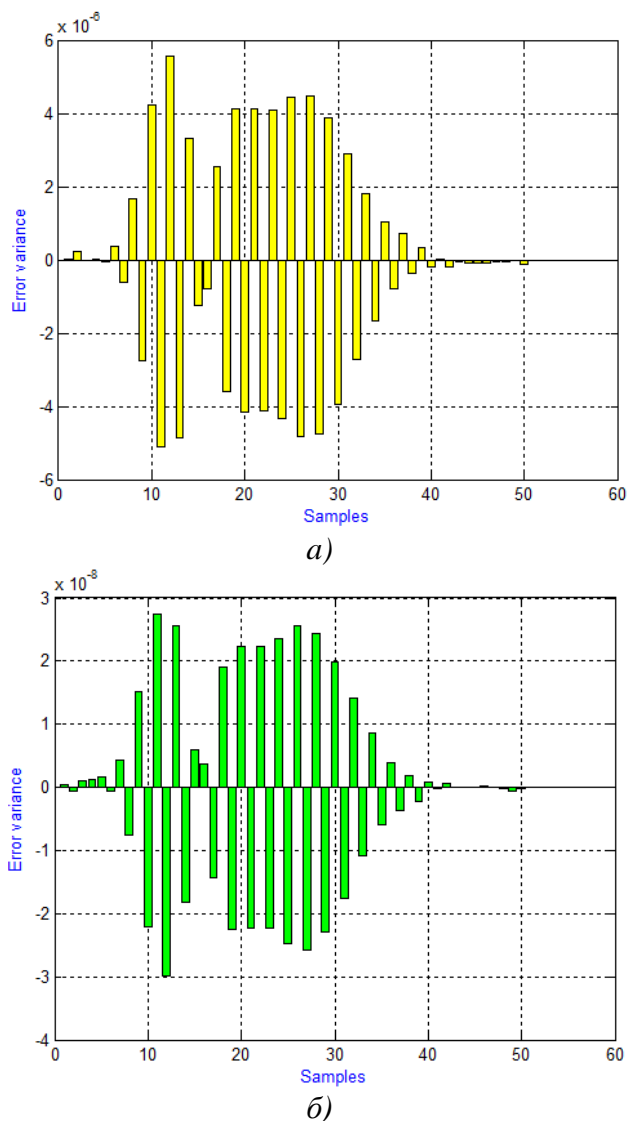


Фиг. 3.29. RBNN за прогнозен анализ на индекси на производителност ST и SRT

Получени са Диаграми на остатъците за синтезираните RBNN прогнозни модели при задаване на минимална широчина в структурния радиално-базисен слой на фиг. 3.30, чиито вариации попадат в следните граници:

- „ $-5.1208e-06$ “ до „ $5.5636e-06$ “ при мониторинг на индекс ST;
- „ $-2.9878e-08$ “ до „ $2.7410e-08$ “ спрямо метрика SRT.

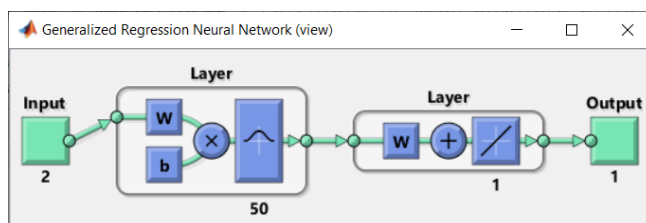
Регистрираните лимити показват високо ниво на ефективност при използване на RBNNs в качеството на аналитични инструменти за регресионна диагностика на параметризацията на преносната среда.



Фиг. 3.30. Вариации на мрежовите грешки при RBNN при $SI = 0.1$ за прогнозен анализ на индекси на производителност ST и SRT

Следваща фаза от подхода интегрира четиристойните GRNNs при залагане на функция на активация в радиално-базисния слой от Гаусов тип на фиг. 3.31. Подборът на прогнозни GRNN модели се базира на оценка на критериите MSE и MAE при експоненциално стъпково увеличаване на SI метриката в радиално-базисния мрежови слой, показано в таблица 3.8. Направена е оценка на регистрираните стойности на MSE и MAE за двата целеви телетрафични индекси на производителност. При двата прогнозни индикатора беше установена обща тенденция на експоненциално нарастване на грешките при възходяща плавна промяна на функционалния параметър SI от 0.050 до 0.300. Най-високи нива на критериите по групи за ST, съответно $MSE = 0.02114$ и $MAE = 0.1144$, и SRT, както следва $MSE = 6.4204e-07$ и $MAE = 6.2290e-04$, са намерени при $SI = 0.300$. При минимална зададена степен $SI = 0.050$, определена с най-голяма степен на пригодност, на функциите в радиално-базисния структурен слой на GRNN са постигнати най-ниски нива $MSE = 1.01008e-04$, $MAE = 0.0042$ за „Пропускателната способност“ и $MSE = 2.9993e-09$, $MAE = 2.21640e-05$ при трафична характеристика „Сумарно време за чакане и обслужване“.

Третият модул от подхода се състои в изследвания, насочени към подбор на CFNNs за целите на прогнозната задача в таблица 3.9. Този категория невронни мрежи се разглеждат като разновидност на FFNNs при съответни структурни различия, свързани с наличието на директна връзка между входния и изходния слой и включването на втора тегловна матрица към изходния слой. Подобни обучаващи процедури бяха приложени с използване на LM алгоритъм при структурно изменение на скритите неврони в познатия диапазон от 5 до 15 изчислителни единици по отношение на комплексен анализ на индикатори за качество MSE и MAE.



Фиг. 3.31. Структура на GRNN модел за прогнозен анализ на индикатор на производителност ST и SRT

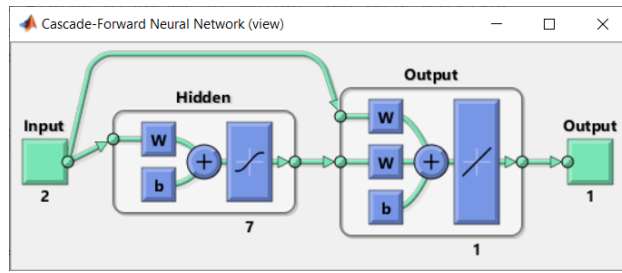
Таблица 3.8. Резултати от оценка на качеството на GRNN модели за прогнозиране на нивата на индикатори на производителност ST и SRT

Spread индикатор	System Throughput		System Response Time	
	MSE	MAE	MSE	MAE
0.050	1.0108e-04	0.0042	2.9993e-09	2.2164e-05
0.075	3.4460e-04	0.0091	1.0252e-08	4.8472e-05
0.100	8.1530e-04	0.0156	2.4299e-08	8.3690e-05
0.125	0.0016	0.0236	4.7319e-08	1.2712e-04
0.150	0.0027	0.0331	8.1474e-08	1.7830e-04
0.175	0.0043	0.0438	1.2892e-07	2.3670e-04
0.200	0.0064	0.0558	1.9185e-07	3.0184e-04
0.225	0.0091	0.0689	2.7244e-07	3.7342e-04
0.250	0.0124	0.0831	3.7294e-07	4.5111e-04
0.275	0.0165	0.0983	4.9550e-07	5.3435e-04
0.300	0.0214	0.1144	6.4204e-07	6.2290e-04

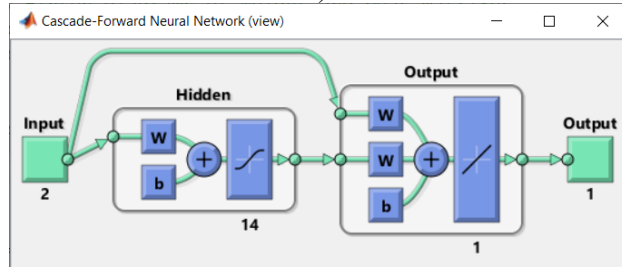
Таблица 3.9. Резултати от изследване на CFNNs за прогнозен анализ на индикатор на производителност System Throughput и System Response Time

Скрити неврони	System Throughput		System Response Time	
	MSE	MAE	MSE	MAE
5	1.6415e-09	2.0126e-05	7.5662e-10	2.0018e-05
6	1.0022e-09	1.4178e-05	6.0152e-11	5.7154e-06
7	1.0782e-10	8.5520e-06	1.3590e-10	8.1408e-06
8	2.2607e-10	9.8568e-06	2.0148e-12	1.2498e-06
9	1.6997e-10	9.0878e-06	5.3233e-11	5.0956e-06
10	5.8023e-09	5.7095e-05	4.0863e-09	5.3444e-05
11	1.8411e-10	1.1316e-05	6.3765e-12	1.5975e-06
12	1.3511e-10	9.4508e-06	7.2580e-10	2.2891e-05
13	1.7786e-09	1.7127e-05	3.3792e-10	1.5585e-05
14	7.3851e-10	1.8198e-05	1.1041e-12	8.0021e-07
15	3.0354e-09	3.3020e-05	6.6437e-10	2.0675e-05

Намерени са невронни архитектури, постигащи изискванията за минимизация на анализирания показател, със залагане на 7 за трафична характеристика ST и 14 междинни неврона при индекс SRT, дадени на фиг. 3.34. Постигнатите минимални нива на грешките попадат в последователни порядъци “1,0782.10⁻¹⁰ – MSE за у₁”, “8,5520.10⁻⁶ – MAE при у₁”, “1,1041.10⁻¹² – MSE при у₂” и “8,0021.10⁻⁷ – MAE за у₂”. Установени са най-високи стойности на критериите за качество групите MSE = 5.8023e-09, MAE = 5.7095e-5 при индекс ST и MSE = 4.0863e-09, MAE = 5.3444e-5 за “Сумарното време за престои в опашката и обслужване”.

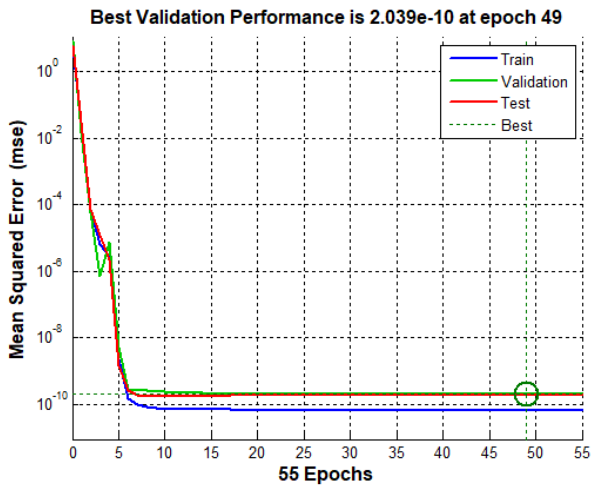


a)

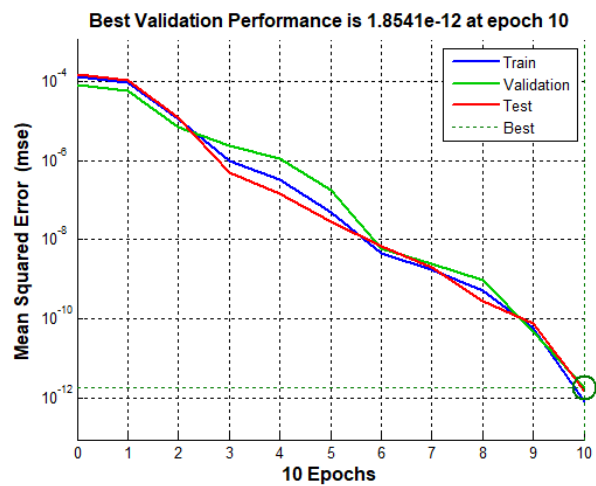


б)

Фиг. 3.34. CFNNs за прогнозиране на нивата на индикатор на производителност
а) Пропускателна способност и б) Комплексно време за чакане и обслужване

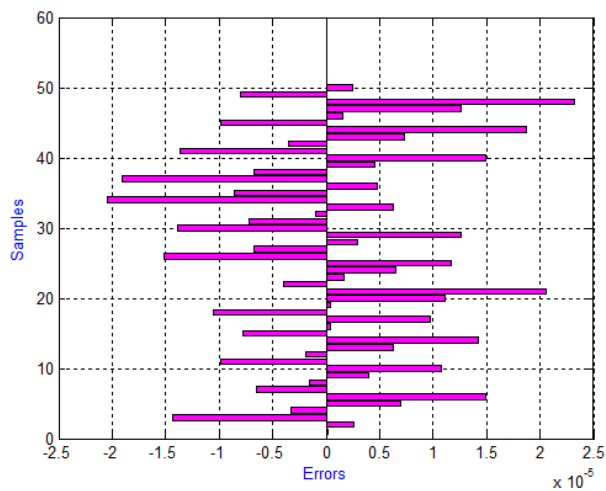


a)

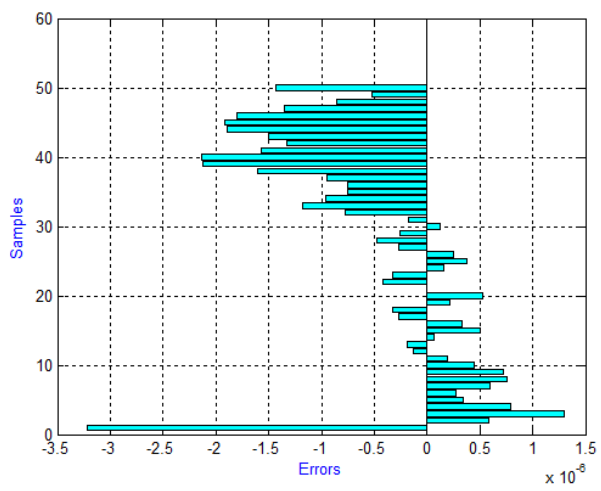


б)

Фиг. 3.35. MSE за синтезираните CFNNs за прогнозиране на индекс ST и б) SRT



a)



б)

Фиг. 3.36. Диаграма на грешките за CFNN за прогнозиране на индекс на производителност
Пропускателна способност а) и Сумарно време за чакане и обслужване б)

Анализиран е критерия MSE при обучение, валидиране и тестване на селектираните крайни CFNNs на фиг. 3.35. Налице е обща тенденция на сходство в направлението на изменение на изходните характеристиките между мрежовите процеси. Обучаващите процедури протичат в рамките на 55 при модела за прогнозиране на ST и 10 итерации за CFNN архитектурата, асоциирана с мониторинг на индекс SRT. Най-добра „Валидираща мрежова производителност“ $2.039e-10$ и $1.8451e-12$ е достигната при 49-ти и 10-ти цикъл от обучение, респективно за мрежите за прогнозен анализ на трафични характеристики „Пропускателна способност“ и „Комплексно време за чакане и обслужване“. Отчитайки диаграмите на остатъците на фиг. 3.36, както и сравнявайки цялостното поведение на последните два аналитични инструмента за прогнозен мониторинг, се отчита, че синтезираните прогнозни Cascade-Forward модели превъзхождат в значителна степен GRNNs.

3.5. Изводи към трета глава.

- Предложен е алгоритъм на имитационно моделиране на ИКТ мрежова инфраструктура с опашкова организация на мониторинг и администриране на мрежовия трафик;
- Въз основа на регресионна диагностика е установено предимство на моделите от 2-ри ред при оценка на адекватността на изведени линейни и полиномиални модели за прогнозен анализ на трафични индекси ST, SRT и SD на имитационно моделирана ИКТ инфраструктура с чакане;
- Установена е ефективност на предложен подход за оптимизация с Interior-point, SQP, Active-set и Genetic алгоритми за намиране на оптимум при минимизация на трафичен параметър Комплексно време за престой в опашката и обслужване с цел подобряване на бързодействието при обработка на заявки;
- Установена е необходимост от включване на втори управляем фактор спрямо единично входно въздействие на управляем фактор SD при синтезиран подход с интеграция на GRNNs и FFNNs за регресионно моделиране на трафични параметри ST и SRT;
- Въз основа на проведена верификация на мрежовата производителност е доказано предимството на LM пред BR и SCG обучаващи алгоритми при подход с Feed-forward невронни мрежи и две входящи въздействия за прогнозиране на ST и SRT;
- Доказана е ефективността на подход за регресионно моделиране във връзка с вариациите на трафични индекси ST и SRT с установено предимство за мрежова ефективност при интегриране на RBNN в сравнение с GRNN и CFNN.

Глава IV: Мониторинг на състоянието на мрежовия трафик към корпоративни клиенти с AI и ML инструменти

4.1. Диагностика на състоянието на UDP мрежовия трафик с трислойни и многослойни невронни архитектури.

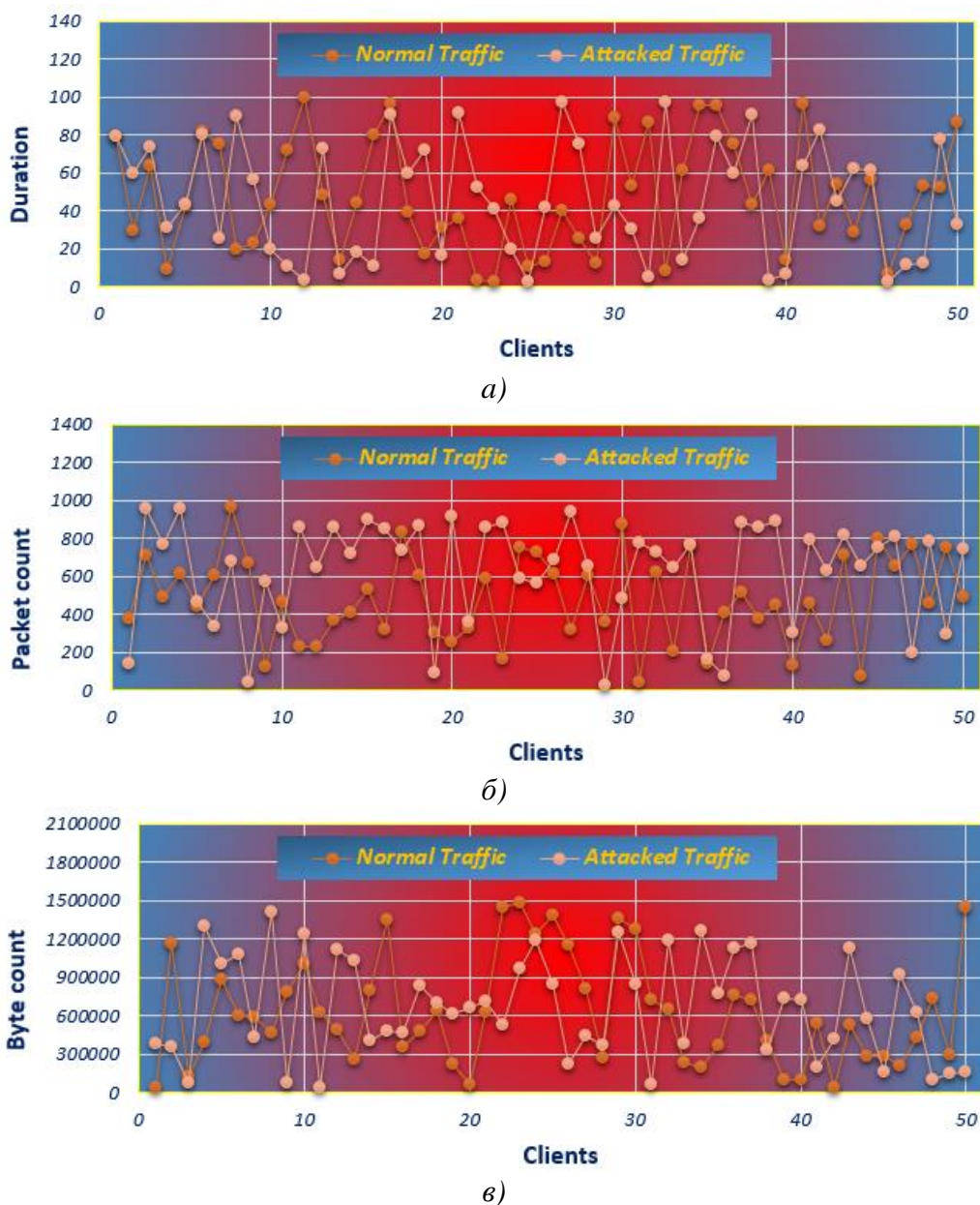
4.1.1. Параметризация на регистриран UDP трафик.

Съвременната концепция за качество на услугата в комуникациите при предаване на пакети между мрежови сегменти се основава на прилагането на адекватен интелигентен анализ на състоянието на преносната среда. Основен акцент се поставя на мрежовата диагностика за установяване на “присъствието на атаки”. Една от възможностите за това е процедурата да се отнася до внедряване на AI за анализ на състоянието на предавания трафик на някои често използвани мрежови стандарти - TCP, UDP или ICMP. Дефинирана е задачата за разпознаване на категории мрежови UDP трафик към корпоративни клиенти, както следва:

- Normal traffic;
- Attacked traffic,

на базата на регистрирани трафични метрики (фиг. 4.1), както следва:

- Duration;
- Packet count;
- Byte count.



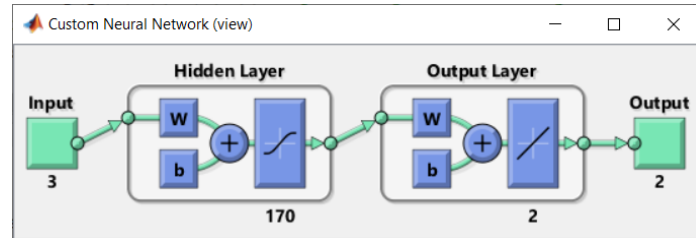
Фиг. 4.1. Параметри при предаване на UDP (User Datagram Protocol) трафик – а) Duration, б) Packet count и в) Byte count

4.1.2. Мониторинг на състоянието на UDP мрежови трафик с интегриране на Feed-Forward, Cascade-Forward мрежи с LM обучаващ алгоритъм и Probabilistic Neural Networks.

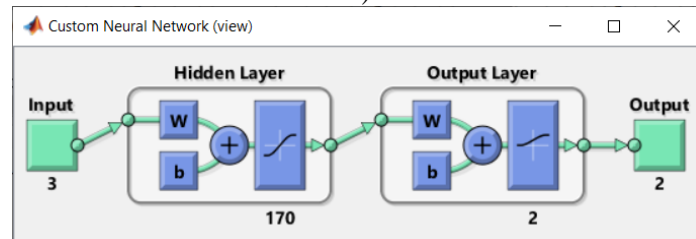
Разработена е методика за диагностика на състоянието на предавания UDP трафик към корпоративни клиенти с включване на изкуствени невронни мрежи с количество изчислителни неврони във вторите структурни слоеве идентично с установеното при предходната базисна невронна архитектура, респективно 170 градивни невронни единици. Процедурите по изследване интегрират следните невронни апарати:

- Feed-Forward структура при замяна на обучаващия SCG алгоритъм с LM обучение и използване на “линейна” функция на активация в изходния слой (фиг. 4.6.а);
- Feed-Forward модел с LM алгоритъм при прилагане на “логаритмична” активационен тип на невроните в изходния слой (фиг. 4.6.б);
- Feed-Forward невронна мрежа при LM обучение при задаване на “тангенс-сигмодална” функция на активация в изходния слой (фиг. 4.6.в);
- Cascade-Forward Neural Network на основата LM обучаващ алгоритъм (фиг. 4.6.г).

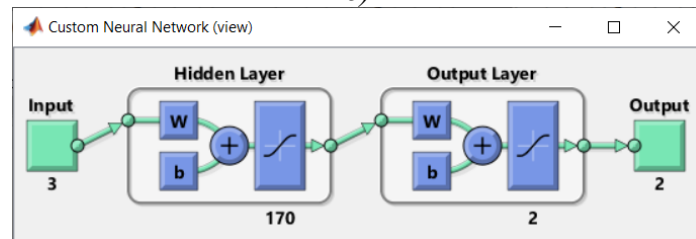
В хода на експерименталните изследвания на тестовите невронни модели в последователен ред са получени количествени еквиваленти на „средноквадратичната грешка“, съответно $MSE = 0.1882$, $MSE = 0.1916$, $MSE = 0.1355$ и $MSE = 0.3015$. Отчетени са сравнително близки класификационни точности за тестовите невронни модели за разпознаване, равняващи се на 90.0 % при „purelin“, 85.0 % за „logsig“, 84.0 % при „tansig“ активация за FFNNs и отново 90.0 % при CFNN.



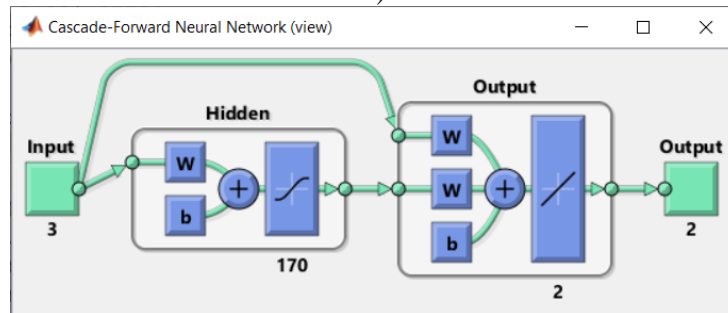
a)



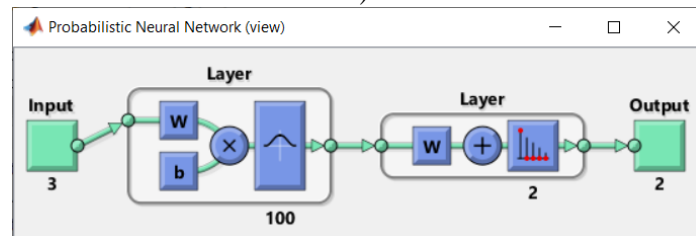
б)



в)



г)



д)

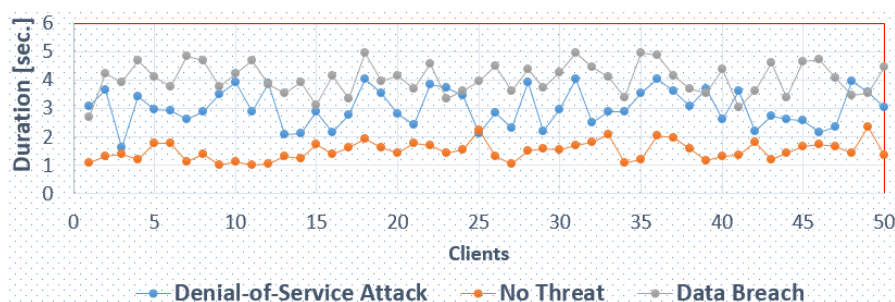
Фиг. 4.6. FFNNs при а) „purelin“, б) „logsig“ и в) „tansig“ изходна активация, г) CFNN и д) PNN за детектиране на атаки при предаване на UDP трафик

Въпреки достигнатия праг на разпознаване 90.0 %, нивата на регистрираните стойности на показателя „MSE“ остават значително високи. Този факт дава основание за търсене на друг аналитичен инструмент с по-добра производителност, каквато беше установена за случая при

включване на четирислойните Probabilistic Neural Networks на фиг. 4.6.д) при средни нива на „широчината на Кернел функцията в радиално-базисния слой“. Тук е констатирано значително минимизиране на критерия „средноквадратична грешка“ до числовия еквивалент $MSE = 0.0100$ и повишаване на класификационната точност до 99.0 %

4.2. Детектиране на DoS атаки, състояние на коректност и аномалии в предавания мрежови трафик за корпоративни клиенти.

Проведена е верификация в условията на лабораторна среда за потвърждаване на функционалността на многослойните FFNNs при диагностика на мрежовия трафик за „Детектиране на киберзаплахи и мрежови аномалии“. Приложена е спецификация със запазване на трафичен индекс “Duration” и въвеждане на специфичните параметри на преносната среда “Network Latency” и “Data Transfer Rate”, показано на фиг. 4.12.



а)



б)



в)

Фиг. 4.12. Параметри на преносната среда а) “Duration“, б) “Network Latency” и в) “Data Transfer Rate” при ситуации на „DoS киберзаплаха“, „нормално състояние на трафика“ и „аномалии в мрежовия трафик“

Изследвана е ефективността на създадени многослойни FFNNs и допълнително въведени CFNNs невронни структури в таблица 4.4 и таблица 4.5. След анализ на количествените индикатори „средноквадратична грешка“ и „класификационна точност“ са подбрани невронни архитектури със съдържание на 9, 16, 20 и 17 неврона в междинните структурни слоеве, представени на фиг. 4.13. Постигнати са точности 90.0 %, 86.7 % и 91.7 % за feed-forward

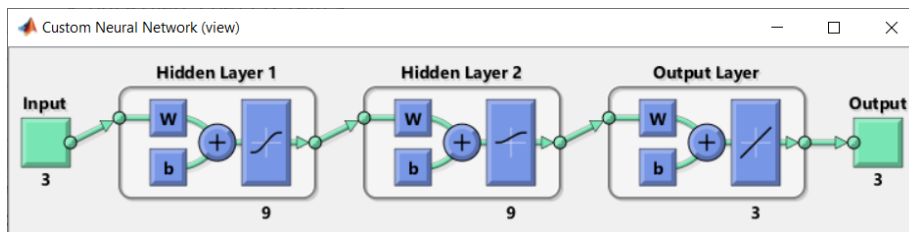
мрежи с “purelin”, “logsig“ и “tansig” изходни активационни функции, и 88.3 % за CFNN с най-добра валидираща производителност. В последователност на указания ред са регистрирани минимални показания MSE = 0.0628, MSE = 0.2056, MSE = 0.0681 и MSE = 0.0714.

Таблица 4.4. Метрики при оценка на многослойни FFNNs за диагностика на киберзаплахи, състояние на нормално предаване на трафика и наличие на аномалии

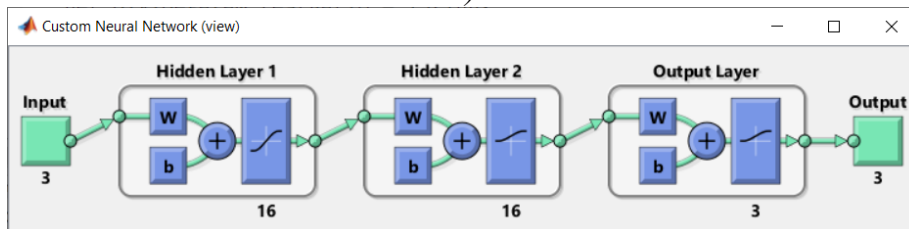
НЕВРОНИ В ПЪРВИ СКРИТ СЛОЙ	НЕВРОНИ ВЪВ ВТОРИ СКРИТ СЛОЙ	MSE	ТОЧНОСТ, %
“PURELIN” ИЗХОДНА АКТИВАЦИЯ			
5	5	0.0814	76.7
6	6	0.0966	81.7
7	7	0.1276	66.7
8	8	0.1068	76.7
9	9	0.0628	90.0
10	10	0.1364	68.3
11	11	0.1076	75.0
12	12	0.1045	70.0
13	13	0.1001	85.0
14	14	0.0861	81.7
15	15	0.0936	75.0
16	16	0.0978	76.7
17	17	0.1361	75.0
18	18	0.1120	73.3
19	19	0.1338	70.0
20	20	0.0929	85.0
“LOGSIG” ИЗХОДНА АКТИВАЦИЯ			
5	5	0.2151	80.0
6	6	0.2176	70.0
7	7	0.2269	60.0
8	8	0.2155	65.0
9	9	0.2325	70.0
10	10	0.2189	63.3
11	11	0.2265	
12	12	0.2112	80.0
13	13		78.3
14	14	0.2099	80.0
15	15	0.2073	83.3
16	16	0.2056	86.7
17	17	0.2240	60.0
18	18	0.2113	78.3
19	19	0.2237	58.3
20	20	0.2364	60.0
“TANSIG” ИЗХОДНА АКТИВАЦИЯ			
5	5	0.0855	80.0
6	6	0.0838	83.3
7	7	0.1532	66.7
8	8	0.1595	63.3
9	9	0.0855	85.0
10	10	0.0837	80.0
11	11	0.0929	
12	12	0.1269	73.3
13	13	0.0964	75.0
14	14	0.0954	81.7
15	15	0.0595	85.0
16	16	0.0945	83.3
17	17	0.0946	85.0
18	18	0.0738	90.0
19	19	0.0913	81.7
20	20	0.0681	91.7

Таблица 4.5. Метрики при оценка на многослойни CFNNs за детектиране на DoS атаки, състояние на нормално предаване на трафика и мрежови аномалии

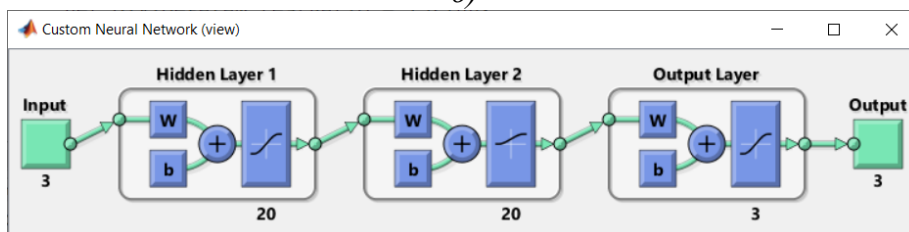
Неврони в първи скрит слой	Неврони във втори скрит слой	Средноквадратична грешка	Точност, %
5	5	0.0971	75.0
6	6	0.0920	83.3
7	7	0.0803	81.7
8	8	0.0856	83.3
9	9	0.0799	81.7
10	10	0.1152	71.7
11	11	0.0928	81.7
12	12	0.0882	80.0
13	13	0.1066	78.3
14	14	0.1026	
15	15	0.0879	
16	16	0.0820	
17	17	0.0714	88.3
18	18	0.0869	83.3
19	19	0.0984	81.7
20	20	0.0914	85.0



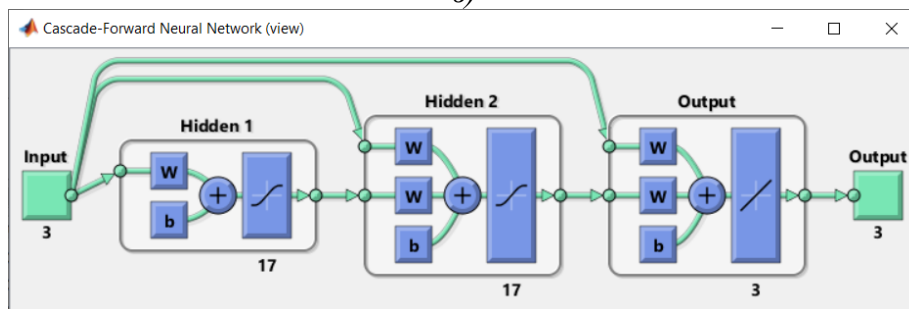
a)



б)



в)



г)

Фиг. 4.13. Многослойни FFNNs с а) “purelin”, б) “logsig”, в) “tansig” изходна активация и г) CFNN за детектиране на „DoS“, „Отсъствие на киберзаплаха“ и „Мрежови аномалии“

4.3. Диагностика на разпределение на WEB съдържание към корпоративни клиенти в урбанизирани зони през активен работен часови пояс.

4.3.1. Параметризация на регистриран WEB трафик към корпоративни клиенти.

Разглежда се задачата за идентификация на географски зони на потребление на Интернет трафик за корпоративни клиенти според данни от активен обеден часови пояс за райони на гр. Чикаго, САЩ. Анализът се базира на оценка основни параметри на WEB трафика на фиг. 4.17, както следва:

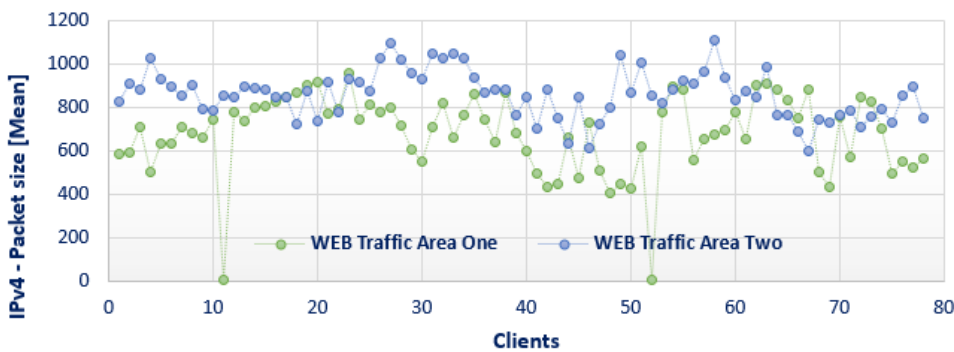
- Количество информационни потоци за единица време, flows/s (kilo-);
- Средна големина пакетите относно протокол IPv4;
- Средна големина пакетите за IP протокол версия 6;
- Средна скорост на предаване на пакети за единица време, pkts/s (kilo-).

Дефинирани са географски зони на потребление Интернет услуги, съответно:

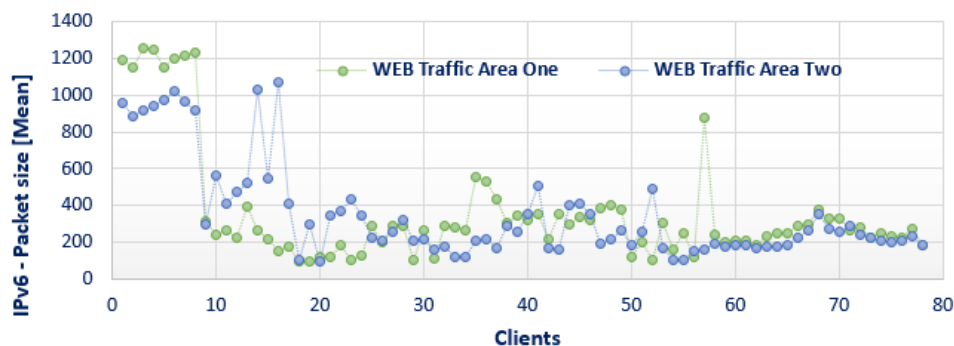
- Категория №1: Chicago city area “1”;
- Категория №2: Chicago city area “2”.



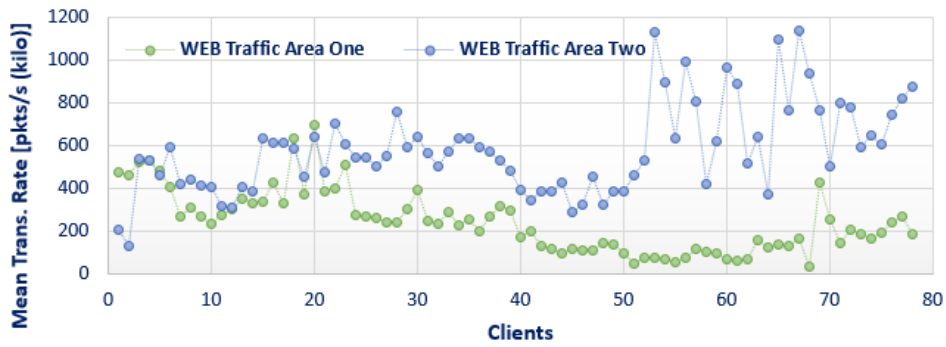
a)



б)



в)



г)

Фиг. 4.17. Регистрирани мрежови параметри при пакетно предаване
 а) информационни потоци, б) големина на пакетите при IPv4,
 в) големина на пакетите при IPv6 и г) средна скорост на предаване на данни

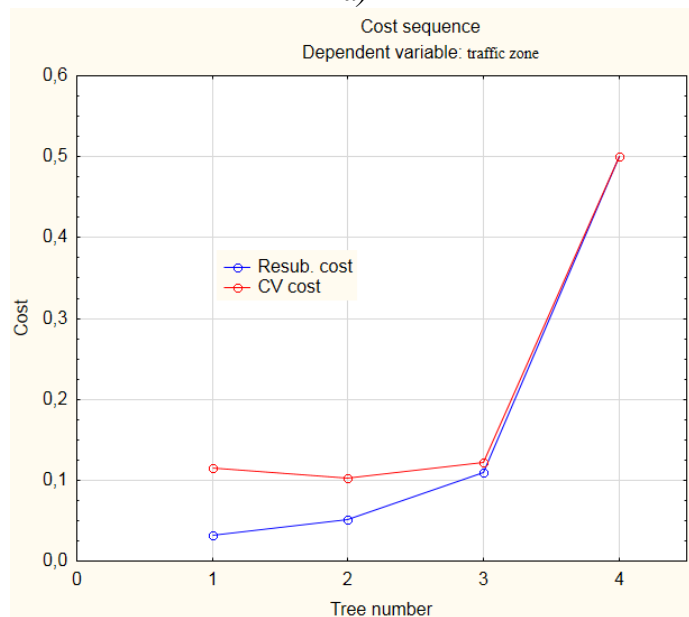
4.3.2. Изследване на модели по метод Дърво на решенията за идентификация на зони на Интернет потребление.

Построени са модели за многовариантен избор на решение за трафична категоризация. Оценени са изведени изходните нива частта на некоректна класификация от техническите подходи V-fold крос-валидиране и Ресубституция на фиг. 4.23 за генерираната базова “Tree 1 (модел при 7 крайни възела)” и архитектури с редуцирани структурни възли:

- “Tree 2 (модел при 4 крайни възела)”;
- “Tree 3 (модел при 2 крайни възела)”;
- “Tree 4 (модел с 1 краен възел)”.

Tree sequence (trafficdanni_zones)					
Dependent variable: traffic zone					
Optimal tree denoted by *					
	Terminal nodes	CV cost	CV std. error	Resubstitution cost	Node complexity
Tree 1	7	0,115385	0,025579	0,032051	0,000000
Tree 2	4	0,102564	0,024291	0,051282	0,006410
*Tree 3	2	0,121795	0,026185	0,108974	0,028846
Tree 4	1	0,500000	0,040032	0,500000	0,391026

а)

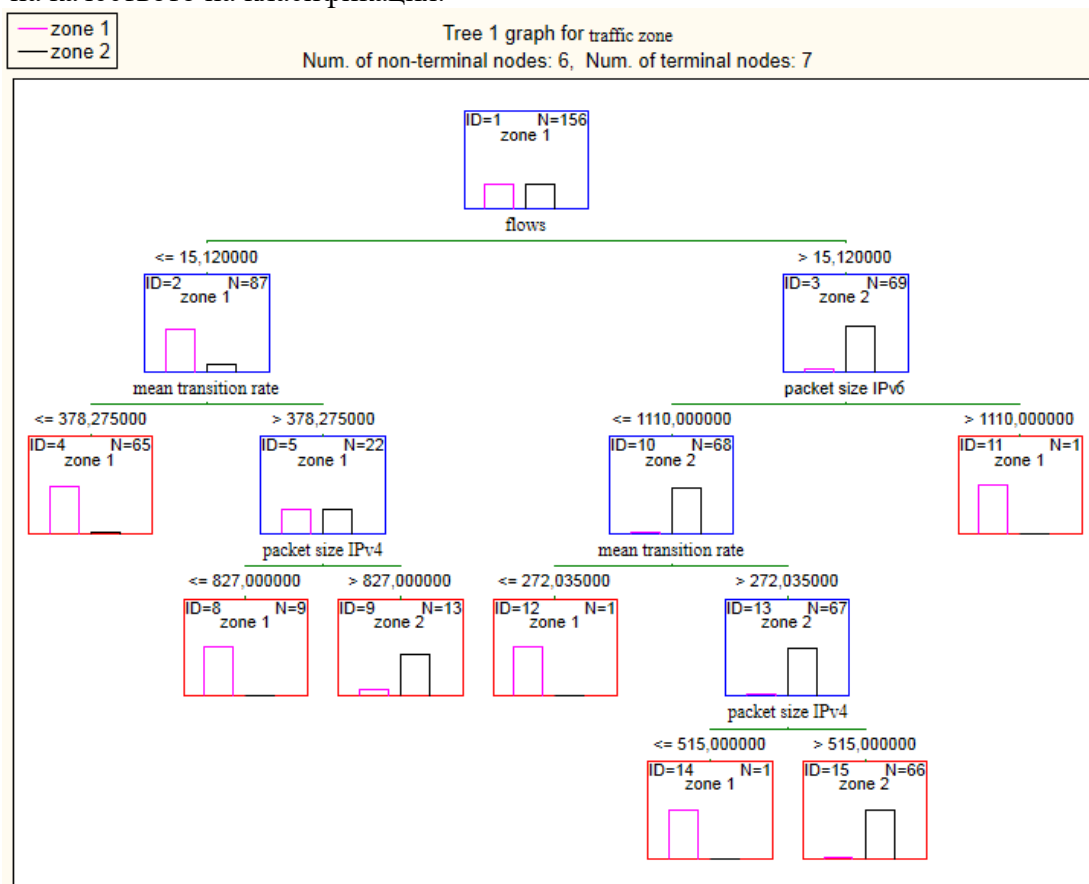


б)

Фиг. 4.23. Оценка на моделите за многовариантен избор на решение чрез крос-валидиране и ресубституция – а) таблично и б) графично представяне

Според изложените резултати е установена тенденция на по-голямо завишаване на точността от V-fold валидация спрямо подход ресубституция с изключение на последната структура. Калкулирани са следните нива на точност за разглежданите модели по метод Дърво на решенията, респективно:

- “Tree 1 (базов модел без редукция на структурни възли)” – от V-fold валидиране 88.4615 % и 96.7949 % при ресубституция;
- “Tree 2” – 89.7436 % чрез крос валидиране и 94.8718 с помощта на ресубституция;
- “Tree 3” – от V-fold валидиране 87.8205 % и 89.1026 % при ресубституция;
- “Tree 4” – точност, равняваща се едва 50.00 % при двата приложени подхода за оценка на качеството на класификация.



Фиг. 4.24. Структура за многовариантен избор на решение при категоризация на зони на Интернет потребление на основата на статистически трафични параметри –

На база на посочените количествени показатели може да бъде дадена „приблизително очаквана точност“ при категоризация на основата на трафични данни, които не са били използвани при обучение и тестване на моделите – 92.6282 % за “Tree 1”, 92.3077 % при архитектура “Tree 2” и 88.46155 % относно определеният като оптимален модел “Tree 3”. Разглеждайки получените крайни нива и съблюдаването на изискването за максимална точност беше избрана генерираната базисна дървовидна архитектура без премахване на разклонения за категоризация на зони на Интернет потребление на фиг. 4.24. Структурата съдържа шест вътрешни и седем крайни възела като е построена без изключване на някой от входните информативни признаци след предварителна оценка на тяхната значимост.

4.3.3. Анализ и оценка на модели по метод Support Vector Machine за разпознаване на зони на Интернет потребление.

Проведен е анализ на „Метода на опорните вектори“ в контекста на задачата за диагностика на региони на потребление на достъпно WEB съдържание в условията на градска

среда. Адаптирани са две категории SVM класификатори, респективно C-SVM и Nu-SVM. при “Linear”, “Polynomial”, “RBF” и “Sigmoid” тип Кернел функции на фиг. 4.25 и фиг. 4.26.

```
Support Vector Machine Results: Internet_Traffic_Zones
Dataset Internet_Traffic_Zones:
Dependent: WEB Traffic Area
Independents: Flows, Packets IPv4, Packets IPv6, Mean Transition Ra
Sample size = 117 (Train), 39 (Test), 156 (Overall)

Support Vector machine results:
SVM type: Classification type 1 (capacity=4,000)
Kernel type: Linear
Number of support vectors = 20 (18 bounded)
Cross-validation accuracy (%) = 99,145
Support vectors per class: 10 (AreaOne), 10 (AreaTwo),

Class. accuracy (%) = 100,000(Train), 100,000(Test), 100,000(Overall)
```

a)

```
Support Vector Machine Results: Internet_Traffic_Zones
Dataset Internet_Traffic_Zones:
Dependent: WEB Traffic Area
Independents: Flows, Packets IPv4, Packets IPv6, Mean Transition Ra
Sample size = 117 (Train), 39 (Test), 156 (Overall)

Support Vector machine results:
SVM type: Classification type 1 (capacity=10,000)
Kernel type: Polynomial (degree=3, gamma=0,250, coefficient=0,000)
Number of support vectors = 50 (48 bounded)
Cross-validation accuracy (%) = 91,453
Support vectors per class: 25 (AreaOne), 25 (AreaTwo),

Class. accuracy (%) = 92,308(Train), 92,308(Test), 92,308(Overall)
```

б)

```
Support Vector Machine Results: Internet_Traffic_Zones
Dataset Internet_Traffic_Zones:
Dependent: WEB Traffic Area
Independents: Flows, Packets IPv4, Packets IPv6, Mean Transition Ra
Sample size = 117 (Train), 39 (Test), 156 (Overall)

Support Vector machine results:
SVM type: Classification type 1 (capacity=8,000)
Kernel type: Radial Basis Function (gamma=0,250)
Number of support vectors = 20 (18 bounded)
Cross-validation accuracy (%) = 99,145
Support vectors per class: 10 (AreaOne), 10 (AreaTwo),

Class. accuracy (%) = 100,000(Train), 100,000(Test), 100,000(Overall)
```

в)

```
Support Vector Machine Results: Internet_Traffic_Zones
Dataset Internet_Traffic_Zones:
Dependent: WEB Traffic Area
Independents: Flows, Packets IPv4, Packets IPv6, Mean Transition Ra
Sample size = 117 (Train), 39 (Test), 156 (Overall)

Support Vector machine results:
SVM type: Classification type 1 (capacity=5,000)
Kernel type: Sigmoid (gamma=0,250, coefficient=0,000)
Number of support vectors = 28 (28 bounded)
Cross-validation accuracy (%) = 98,291
Support vectors per class: 14 (AreaOne), 14 (AreaTwo),

Class. accuracy (%) = 98,291(Train), 97,436(Test), 98,077(Overall)
```

г)

Фиг. 4.25. C-SVM класификатори за разпознаване на зони на разпределение на WEB съдържание

Аналитично са детерминирани необходимите „Опорни вектори“ за построяване на „Хиперповърнината“, разделяща целевите изходни групи в класификационното пространство, съответно „20“, „50“, „20“ и „28“. Установена е най-ниска крайна точност 92.308 за C-SVM модел с приложена Кернел функция “Polynomial”. Постигнато е пълно коректно разпознаване 100.0 % на еталоните, участващи в извадката за обучение и тестване, съвпадащо за случаите на “Linear” и “RBF”, докато при “Sigmoid” Кернел тип крайната отчетена точност е 98.077 %.

```
Support Vector Machine Results: Internet_Traffic_Zones
Dataset Internet_Traffic_Zones:
  Dependent: WEB Traffic Area
  Independents: Flows, Packets IPv4, Packets IPv6, Mean Transition Ra
  Sample size = 117 (Train), 39 (Test), 156 (Overall)

Support Vector machine results:
  SVM type: Classification type 2 (nu=0,100)
  Kernel type: Linear
  Number of support vectors = 12 (10 bounded)
  Cross-validation accuracy (%) = 99,145
  Support vectors per class: 6 (AreaOne), 6 (AreaTwo),

  Class. accuracy (%) = 100,000(Train), 100,000(Test), 100,000(Over...
```

a)

```
Support Vector Machine Results: Internet_Traffic_Zones
Dataset Internet_Traffic_Zones:
  Dependent: WEB Traffic Area
  Independents: Flows, Packets IPv4, Packets IPv6, Mean Transition Ra
  Sample size = 117 (Train), 39 (Test), 156 (Overall)

Support Vector machine results:
  SVM type: Classification type 2 (nu=0,100)
  Kernel type: Polynomial (degree=3, gamma=0,250, coefficient=0,000)
  Number of support vectors = 12 (10 bounded)
  Cross-validation accuracy (%) = 100,000
  Support vectors per class: 6 (AreaOne), 6 (AreaTwo),

  Class. accuracy (%) = 99,145(Train), 100,000(Test), 99,359(Over...
```

б)

```
Support Vector Machine Results: Internet_Traffic_Zones
Dataset Internet_Traffic_Zones:
  Dependent: WEB Traffic Area
  Independents: Flows, Packets IPv4, Packets IPv6, Mean Transition Ra
  Sample size = 117 (Train), 39 (Test), 156 (Overall)

Support Vector machine results:
  SVM type: Classification type 2 (nu=0,100)
  Kernel type: Radial Basis Function (gamma=0,250)
  Number of support vectors = 12 (10 bounded)
  Cross-validation accuracy (%) = 99,145
  Support vectors per class: 6 (AreaOne), 6 (AreaTwo),

  Class. accuracy (%) = 100,000(Train), 100,000(Test), 100,000(Over...
```

в)

```
Support Vector Machine Results: Internet_Traffic_Zones
Dataset Internet_Traffic_Zones:
  Dependent: WEB Traffic Area
  Independents: Flows, Packets IPv4, Packets IPv6, Mean Transition Ra
  Sample size = 117 (Train), 39 (Test), 156 (Overall)

Support Vector machine results:
  SVM type: Classification type 2 (nu=0,100)
  Kernel type: Sigmoid (gamma=0,250, coefficient=0,000)
  Number of support vectors = 12 (10 bounded)
  Cross-validation accuracy (%) = 99,145
  Support vectors per class: 6 (AreaOne), 6 (AreaTwo),

  Class. accuracy (%) = 100,000(Train), 100,000(Test), 100,000(Over...
```

г)

Фиг. 4.26. Nu-SVM класификационни модели за идентификация на зони на Интернет потребление за корпоративни клиенти

С известно предимство е оценена възможността за адаптиране на Nu-SVM класификационни модели по отношение на аналогични типове Kernel функции. Максимален праг на разпознаване 100.0 % е достигнат при „Linear“, „Radial Basis Function“ и „Sigmoid“ функции. Получената крайна точност при Nu-SVM модел със заложен на „Polynomial“ Кернел тип се равнява на 99.359 %.

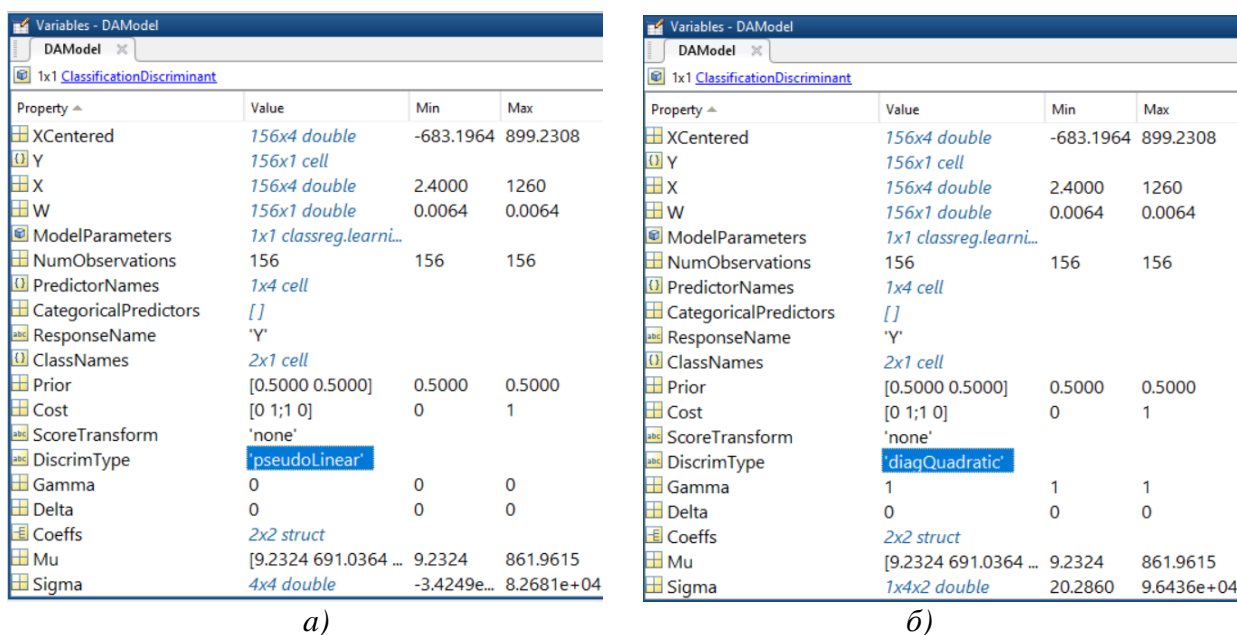
4.3.4. Анализ на ефективността на линейни и квадратични типове дискриминантни класификатори за диагностика на WEB трафик към корпоративни клиенти.

Направена е диагностика на ефективността на „Линейни“ и „Квадратични“ типове дискриминантни класификатори, обобщена в таблица 4.12. Изброени в последователен ред са оценени „Linear“, „Diagonal Linear“, „Pseudo-Linear“, „Quadratic“, „Diagonal Quadratic“ и „Pseudo-Quadratic“, създадени класификационни модели. Извършена е оценка на резултатите за изследваните дискриминантни модели, показващи добра и близка постигната ефективност над нива 85.0 %. Най-ниска класификационна точност 87.82 % е отчетена за случаите на „Diagonal Linear“ и „Diagonal Quadratic“ дискриминантен тип. Най-добра адекватност спрямо спецификата на класификационната задача е установена при „Pseudo-Linear“ модел, изразена с количествена оценка 90.38 %.

Таблица 4.12. Метрики при оценка на Дискриминантни класификатори за идентификация на зони на разпределение на WEB съдържание

Дискриминантен тип класификатор	Грешка при крос-валидиране	Точност при крос-валидиране, %
Linear	0.1026	89.74
DiagLinear	0.1218	87.82
PseudoLinear	0.0962	90.38
Quadratic	0.1154	88.46
DiagQuadratic	0.1218	87.82
PseudoQuadratic	0.1026	89.74

Изведена е съвкупност от специфични статистически калкулирани параметри в хода на изследване на Дискриминантните класификатори на фиг. 4.31, като с най-голяма („Pseudo-Linear“) и най-ниска постигната ефективност („Diagonal Quadratic“). Анализирани са съставът от специфични променливи, включващи „Gamma“, „Delta“, „Sigma“, „Mu“ и т.н.



Фиг. 4.31. Статистически параметри при изследване на а) „Pseudo-Linear“ и б) „Diagonal Quadratic“ дискриминантен модел

4.3.5. Синтез на многослойни FFNNs при „тангенс-сигмоидална“ активационна функция в изходните неврони за диагностика на WEB трафик.

Проведена е систематизация и оценка на вариациите на избраните критерии „точност“, „MSE“ и „MAE“ при анализ на ефективността на FFNNs на основата на приложени съотношения между изчислителните единици в скритите структурни слоеве, респективно „1:1“, „1:2“ и „2:1“, представено от таблица 4.17 до таблица 4.19.

Във връзка с първото приложено съотношение между невроните в първи и втори скрит слой са регистрирани минимално 56.5 % и максимално показание за точността 100.0 % за модели с 10 и 7 до 9 изчислителни междинни единици. При средноквадратичната грешка бяха констатирани най-ниска 0.0017 и най-висока стойност 0.2628, респективно при зададени 8 и 10 скрити неврона. Намерените минимална и максимална MAE се равняват на 0.0200 и 0.2684, установени при същите FFNNs относно предходния индикатор за качество. Съобразно изложените резултати е избрана невронна архитектура с наличие на 8 неврона в структурните междинни слоеве, показана на фиг. 4.40.а).

Таблица 4.17. Изследване на FFNNs с дълбоко обучение за идентификация на зони на потребление на Интернет съдържание при съотношение между невроните в скритите слоеве 1:1

№	Скрит слой №1	Скрит слой №2	Точност, %	MSE	MAE
1.	3	3	95.7	0.0397	0.1007
2.	4	4	78.3	0.1134	0.1821
3.	5	5	65.2	0.1855	0.2070
4.	6	6	95.7	0.0393	0.0627
5.	7	7	100.0	0.0151	0.0436
6.	8	8	100.0	0.0017	0.0200
7.	9	9	100.0	0.0222	0.1061
8.	10	10	56.5	0.2608	0.2684
9.	11	11	95.7	0.0463	0.0669

Таблица 4.18. Изследване на FFNNs с дълбоко обучение за идентификация на зони на потребление на Интернет съдържание при съотношение между невроните в скритите слоеве 1:2

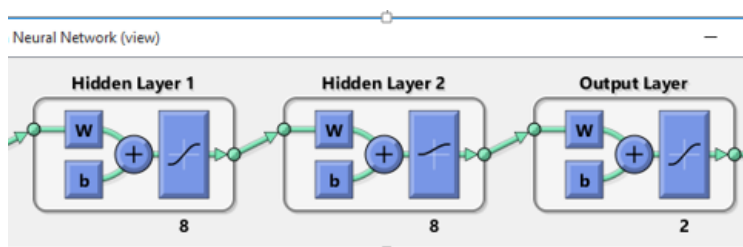
№	Скрит слой №1	Скрит слой №2	Точност, %	MSE	MAE
1.	3	6	95.7	0.0531	0.0609
2.	4	8	95.7	0.0451	0.0641
3.	5	10	47.8	0.2681	0.2848
4.	6	12	95.7	0.0558	0.0900
5.	7	14	87.0	0.1142	0.1811
6.	8	16	43.5	0.4134	0.4874
7.	9	18	100.0	0.0263	0.1145
8.	10	20	95.7	0.0244	0.0616
9.	11	22	95.7	0.0317	0.1211

По отношение на второто разгледано съотношение са намерени две архитектури, при които е получена точност под ниво от 50.0 %, както следва 43.5 % при 8 в първи и 16 неврона във втори скрит слой и 47.8 % за FFNN с 5 в първи и 10 скрити неврона във втори слой. Констатирано е еднократно достигане на пълно коректно разпознаване на обработваните информационни трафични еталони – точност 100.0 %, при структура с 9 и 18 неврона в първия и втория междинен слой. В рамките на описания случай бяха регистрирани и минимални MSE

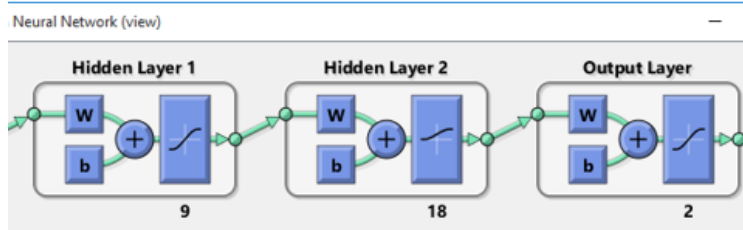
= 0.0263 и MAE = 0.1145. Максимални нива на средноквадратичната грешка 0.4134 и средната абсолютна грешка 0.4874 са наблюдавани за невронна архитектура при „8 в първия и 16 неврона във втория междинен слой“. Анализът на резултатите показва предимство на FFNN със съдържание на девет в първи и осемнадесет междинни неврона във втори скрит слой (фиг. 4.40.б)). Но имайки предвид в пъти по-ниските степени на MSE и MAE при първия разгледан случай в процедурния синтез, по-подходяща се оказва неговата употреба.

Таблица 4.19. Изследване на FFNNs с дълбоко обучение за идентификация на зони на потребление на Интернет съдържание при съотношение между невроните в скритите слоеве 2:1

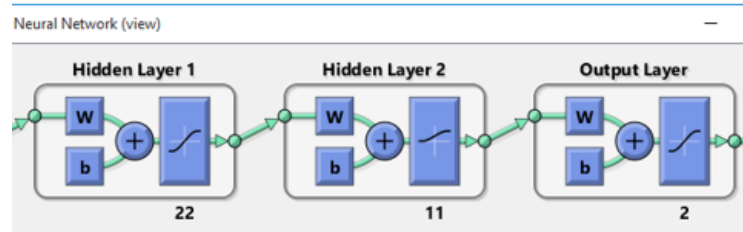
№	Скрит слой №1	Скрит слой №2	Точност, %	MSE	MAE
1.	6	3	52.2	0.2394	0.2468
2.	8	4	87.0	0.0774	0.1402
3.	10	5	39.1	0.3102	0.3378
4.	12	6	95.7	0.0327	0.0792
5.	14	7	91.3	0.0859	0.1137
6.	16	8	95.7	0.0506	0.1303
7.	18	9	91.3	0.0671	0.1124
8.	20	10	82.6	0.1600	0.2364
9.	22	11	100.0	0.0073	0.0345



a)



б)



в)

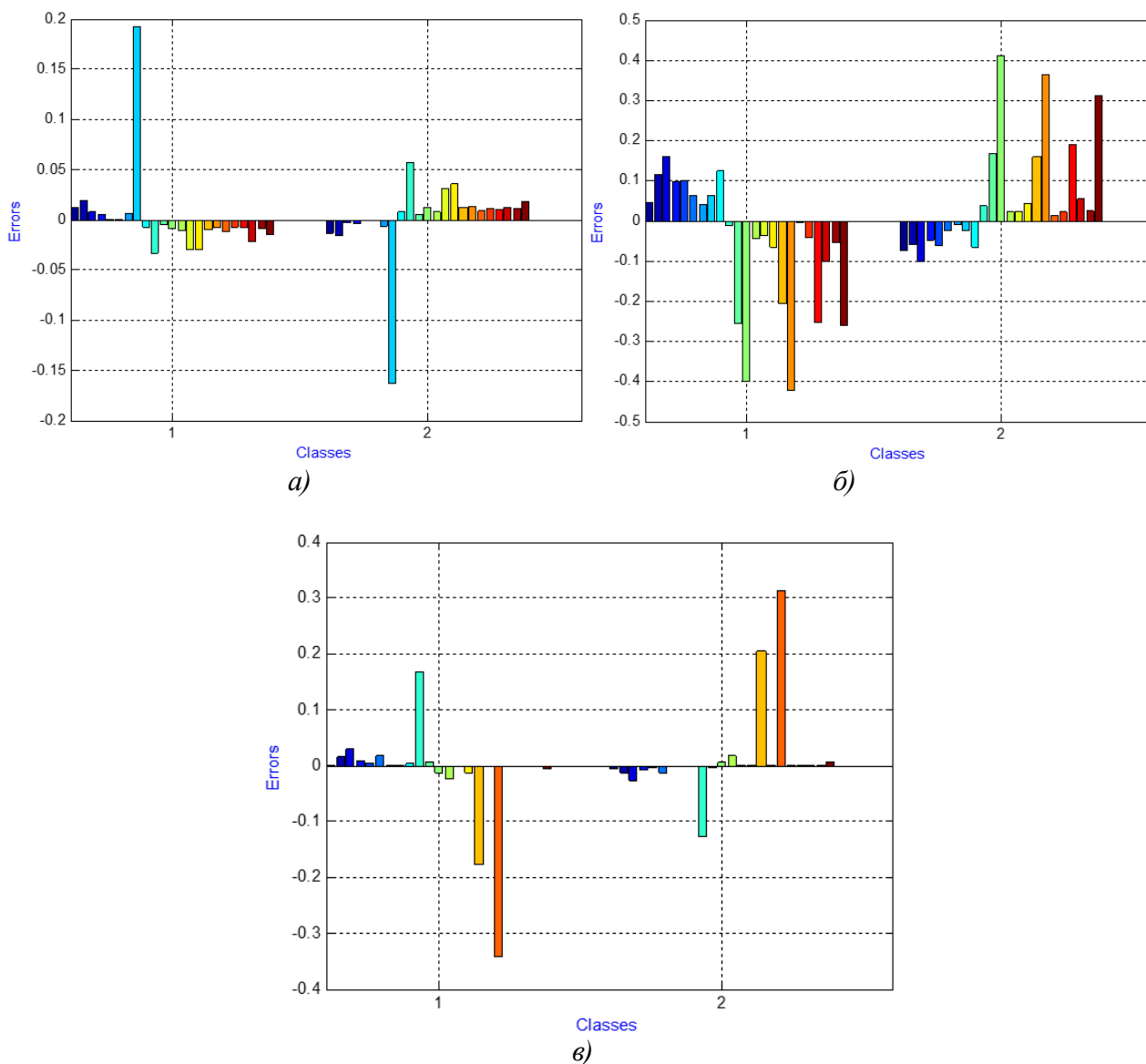
Фиг. 4.40. Селектирани FFNNs за идентификация на географски зони на корпоративно Интернет потребление при съотношения между невроните в скритите слоеве а) „1:1“, б) „1:2“ и в) „2:1“

Съгласно последното приложено количествено съотношение между невроните в скритите слоеве отново е установена една комбинация, при която е постигната коректна идентификация

на целевите зони на корпоративно Интернет потребление. Това е FFNN с наличие на 22 и 11 неврона, респективно в първи и втори скрит слой на фиг. 4.40.в). По отношение на указаната невронна структура бяха регистрирани най-ниски нива на $MSE = 0.0073$ и $MAE = 0.0345$. Най-възходящи вариации на грешките $MSE = 0.3102$ и $MAE = 0.3378$ са намерени в архитектура, при която са използвани 10 изчислителни единици за първия и 5 междинни неврона за втория скрит слой. От разгледаните случаи на невронен синтез тук е констатирана най-ниска класификационна точност, попадаща в порядъка 39.1 %.

Сравнявайки нивата на грешките, може да се каже, че FFNNs при съотношение „2:1“ притежават по-добри качества спрямо тези, изградени на основата на съотношение „1:2“. Оценявайки синтезираните крайни модели на идентификация на трафични зони, архитектурата с идентично количество невронни изчислителни единици в междинните слоеве се определя като модел с най-висока мрежова производителност.

Генерирани са Диаграмите на мрежовите грешки на фиг. 4.42, представляват вариациите на разликите между теоретично заложените и калкулираните посредством синтезираните многослойни модели резултати при манипулации с данните, включени в състава на тестовия поднабор.



Фиг. 4.42. Диаграми на грешките за селектираните FFNNs за идентификация при съотношения между невроните в скритите слоеве а) „1:1“, б) „1:2“ и в) „2:1“

При анализираниите многослойни невронни модели за идентификация на зони на корпоративно Интернет потребление са установени следните вариационни диапазони:

- -0.1627 до 0.1928 при FFNN с приложено съотношение между невроните в междинните слоеве „1:1“;
- -0.4241 до 0.4119 относно модела при задаване на 9 в първия и 18 изчислителни единици във втория скрит слой;
- -0.3428 до 0.3130 за многослойната невронна архитектура с използвано съотношение между невроните в скритите слоеве „2:1“.

Според оценка на констатираните вариации на мрежовите грешки се откроява предимството на модела при идентично количество неврони в двата междинни слоя, следвано от невронни архитектури с приложено съотношение „1:2“ и „2:1“. Представените грешки потвърждават предимствата на първата алтернатива за разпределение на невронните изчислителни единици относно разглежданата задача за идентификация.

4.4. Изводи към четвърта глава.

- Предложен е подход за диагностика на мрежовия трафик и установяване на атаки при предаване на UDP пакети към корпоративни клиенти с включване на Feed-Forward при SCG и LM обучение при различни изходни активации, Cascade-Forward и Probabilistic невронни апарати с установено предимство на PNN;
- Анализирани е ефективността на въведени подходи за Машинно обучение на основата на Дискриминантен анализ, Алгоритъм на Naïve Bayes, Дърво на решенията и k – най-близки съседи при детектиране на атаки към предаван UDP трафик с констатиран сравнително нисък праг на точност около 67.0 %;
- Направена е оценка на Дълбокото обучение при създадени Feed-Forward архитектури с идентично количество неврони в скритите слоеве при различни изходни активационни функции за разпознаване на UDP трафик, подложен на атаки, с установено значително предимство пред използване на ML инструменти;
- Проведена е верификация на адекватността на принципите на „Дълбокото обучение“ при въведен подход за диагностика на обслужвания мрежови трафик между корпоративни клиенти с цел детектиране на киберзаплахи и мрежови аномалии;
- Проведен е синтез на FFNN, k-NN, CART Decision tree, SVM и Naïve Bayes модели чрез STATISTICA и MATLAB за класификация на географски зони на Интернет потребление в урбанизирана градска среда с постигната най-висока ефективност при Алгоритъма на Naïve Bayes;
- Изследвани са създадени линейни и квадратични дискриминантни класификатори, Feed-Forward с различни изходни активации и Probabilistic невронни мрежи за анализ на разпределението на WEB съдържание към корпоративни клиенти с достигната най-висока точност 100.0 % при PNN;
- Предложен е подход при различни съотношения между невроните в междинните слоеве при Дълбоко обучение на Multilayer Feed-Forward за идентификация на зони на достъп до Интернет услуга към корпоративни клиенти в урбанизирана градска среда с установена най-висока ефективност при съотношение 1:1 при тангенс-сигмоидална и линейна функция на изходна активация.

ЗАКЛЮЧЕНИЕ

Дисертационният труд предлага създадена комплексна стратегия за мониторинг и диагностика на смущенията, параметризацията и мрежовите ресурси на преносната среда, както и оценка на състоянието и потреблението на обслужваните потоци от предавани пакетни данни. Проведени са систематични дейности по адаптиране, модифициране и разработване на хибридни аналитични модули за извличане на знания от данни с внедряване на значим спектър от инструменти на основата на технологиите AI, ML и DL.

Предложената стратегия се характеризира с широка потенциална интеграция, както в локални системи с относително автономно функциониране, така и в мащабни телекомуникационни системи и информационно-комуникационни инфраструктури за администриране на мрежовия трафик от етап планиране до експлоатация. Отсъствието на регламентирана обвързаност със специфични технически изисквания към: „структурните комуникационни сегменти“; „типа на преносната среда“; „дължината на трасето“; „стандартите за управление на мрежовата комуникация“; „количеството обслужвани потребители“ и „технологията за широколентов достъп“, формира основа за висока степен на техническо внедряване.

Налице са значими възможности за интеграция с насоченост към малки, средни и големи корпоративни клиенти и административни центрове и клъстери в сферата на информационното обслужване доставчици на Интернет услуги и мрежови оператори. Бъдещите технически перспективи за подобряване на обхвата на изследванията могат да бъдат обобщени в следните направления:

- включване на допълнителни обекти и процеси в контекста на задачите за класификация и прогнозен анализ по отношение на управление на киберсигурността при обезпечаване на достъпа до електронни ресурси и услуги;
- разширяване на обсега на използваните оптимизационни методи и алгоритми за търсене на екстремуми във връзка с комплексни индекси на производителност на преносната среда и оценка на трафичното натоварване на обслужващите мрежови сървъри, разпределителни звена и системни модули.

ПРИНОСИ НА ДИСЕРТАЦИОННИЯ ТРУД

Научно-приложни приноси:

- Доказани са предимствата на разработения подход с C-SVM, Nu-SVM, k-NN, ANFIS, CFNN, PNN и Naïve Bayes класификатори за разпознаване на Гаусов бял шум и Периодичен случен шум с FFT и без предварителен спектрален анализ в симулационно моделирана преносна среда;
- Установено е предимството на Изкуственият интелект, включващ Cascade-Forward Neural Networks, Feed-Forward Neural Networks и Generalized Regression Neural Networks архитектури, при прогнозиране на амплитудните вариации на Постоянен бял шум и Периодичен случаен шум към аналогови и цифрови сигнали в симулационна комуникационна среда;
- Предложени са подходи за прогнозен анализ на трафични параметри Пропускателна способност и Комплексно време за престой в опашката и обслужване с Регресионен анализ и GRNN, FFNN при LM, BR и SCG алгоритми на обучение, RBNN и CFNN при доказана ефективност с два управляеми фактора;
- Разработена, оценена и внедрена е методика относно ефективността на Interior-point, Sequential Quadratic Programming, Active-set и Genetic алгоритми за оптимизация на параметър на мрежовата среда Комплексно време за чакане и обработка на заявки в имитационно моделирана ИКТ инфраструктура;
- Доказано е превъзходството на Feed-Forward с SCG и LM обучение, CFNN, PNN, DA, Naïve Bayes, DT, k-NN и Multilayer FFNNs инструменти за диагностика на нормален и подложен на атаки мрежови трафик на основата на UDP комуникационен протокол;
- Синтезирани и разработени са методи за внедряване на аналитични инструменти FFNN, k-NN, CART Decision Tree, SVM, Naïve Bayes, PNN и Multilayer FFNNs при различни съотношения между невроните в скритите и активационни функции в изходните слоеве за идентификация на WEB трафик в урбанизирана градска среда.

Приложни приноси:

- Установена е ефективността на приложени аналитични инструменти със спектрален анализ и дескриптивна статистика за извличане на информативни признаци при синтез на модели за класификация и прогнозиране;
- Предложена е програмна реализация и експериментални изследвания с алгоритъм за симулационно моделиране и е доказана производителността на телетрафични системи с опашкова организация на обслужването на потребители;
- Потвърдена е ефективността на Дескриптивния анализ за оценка на метрики на постъпващ мрежови трафик при диагностика на пакетното предаване на данни.

СПИСЪК НА ПУБЛИКАЦИИТЕ ПО ДИСЕРТАЦИОННИЯ ТРУД

1. Balabanova I., **Zhorova T.**, Georgiev G., “Internet Traffic Analysis by ANFIS, k-Nearest Neighbors and Decision Tree Approach“. Proceedings of University of Ruse, vol. 61, book 3.2, FRI-2G.302-1-CCT2-01, pp. 44-51, 2022. ISSN: 2603-4123 (online), ISSN: 1311-3321 (print), ROAD Database. Best Paper Award!
2. Balabanova I., **Zhorova T.**, “Traffic Information Analysis Using Deep Learning Artificial Neural Networks“. Proceedings of International Scientific Conference UNITECH’22, 18 – 19 November 2022, Gabrovo, Bulgaria, pp. I-173 - I-178, 2022.
3. Balabanova I., **Zhorova T.**, Georgiev G., “Amplitude Noise level Prediction in Signals with Noises by Statistics Extraction and CFNN Techniques“. AIP Conference Proceedings - 2nd International Conference on Electronics, Engineering Physics and Earth Science, EEPES’23, vol. 3063, issue 1, pp. 1-10, 2024. ISSN: 0094243X, ISBN: 978-073544856-8, DOI: 10.1063/5.0196146, Scopus SJR 0.152 (2023).
4. Balabanova I., **Zhorova T.**, Georgiev G., “Internet Traffic Zone Identification by Backpropagation and Probabilistic Neural Networks“. Proceedings of 14th International Scientific and Practical Conference “Environment. Technologies. Resources”, vol. 2, pp. 12-17, 2023. ISSN: 2256-070X, ISBN: 978-171387488-1, DOI: 10.17770/etr2023vol2.7265, Scopus.
5. Balabanova I., Georgiev G., **Zhorova T.**, “WEB Traffic Monitoring by Descriptive Statistics and Multilayer Neural Networks“. Proceedings of University of Ruse, 27-28 October 2023, Ruse, Bulgaria, vol. 62, book 3.2, FRI-2G.302-1-CCT2-08, pp. 125-133, 2023. ISSN: 2603-4123 (online), ISSN: 1311-3321 (print), ROAD Database
6. **Zhorova T.**, „Synthesis of Predictive Models for Network Traffic with FFNNs and Training Algorithms Selection“. Proceedings of International Scientific Conference UNITECH’ 23, 17 - 18 November 2023, Gabrovo, pp. I-211 - I-216, 2023.
7. **Жорова Т.**, Косев К., Петкова А., Пенчев И., Ганчев Ц., „Мониторинг на параметри на мрежовия трафик с изкуствен интелект“. Сборник доклади на „Студентска научна сесия СНС’24“, 18 Октомври, Габрово, Том I, 78-83 стр., 2024, ISSN: 1313-3055.

УЧАСТИЕ ВЪВ ВЪТРЕШНИ ПРОЕКТИ

1. Проект № **НИП2025-6** „ИЗСЛЕДВАНЕ НА ПРЕНОСНАТА СРЕДА И МОДЕЛИРАНЕ НА БИЗНЕС ПРОЦЕСИ НА ПРОЕКТИ С МАШИННО ОБУЧЕНИЕ В КОМУНИКАЦИИТЕ“ към Факултет „Електротехника и електроника“ с научен ръководител доц. д-р инж. Ивелина Стефанова Балабанова за 2025г.
2. Проект № **2406Е** „МОНИТОРИНГ НА МРЕЖОВИЯ ТРАФИК И ДОСТЪП НА ИНФОРМАЦИЯ ЧРЕЗ БИОМЕТРИЧНА ДИАГНОСТИКА С ИЗКУСТВЕН ИНТЕЛЕКТ“ към Факултет „Електротехника и електроника“ с научен ръководител гл. ас. д-р инж. Георги Иванов Георгиев за 2024г.
3. Проект № **2305Е** „РАЗРАБОТВАНЕ НА ИКТ-БАЗИРАНИ СИСТЕМИ ЗА ИЗСЛЕДВАНИЯ И МОНИТОРИНГ НА ТРАФИКА И ПОТРЕБИТЕЛСКИЯ ДОСТЪП С ИЗКУСТВЕН ИНТЕЛЕКТ“ към Факултет „Електротехника и електроника“ с научен ръководител гл. ас. д-р инж. Георги Иванов Георгиев за 2023г.

TRAFFIC MONITORING IN COMMUNICATION SYSTEMS WITH ARTIFICIAL INTELLIGENCE

Mag. Eng. Teodora Valentinova Zhorova

ABSTRACT

Dissertation integrates complex methods for adapting Artificial Intelligence instruments in the context of considered scenarios for monitoring and diagnostics of the network resources in communication systems. The scientific research scope is aimed at Hybrid analytical models developing by Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) for noise impact analysis, performance indices, transmission media security and traffic services assessment. Objects of research about the state of communication channels are Gaussian White Noise, Periodic Random Noise and Uniform White Noise as independent and present effects to Analog and Digital signals. The specified types of interference are analyzed on the basis of synthesized AI and ML classification and predictive models about their Noise amplitude variance analysis in the conditions of a simulated network environment. The transmission media state in ICT infrastructures simulated with queue management systems of Incoming and Outgoing traffic streams is evaluated with the inclusion of Statistics, AI and ML techniques for Regression modeling and diagnostics. Predictive analysis and Non-linear optimization objects are the specific metrics of the network environment such as "System Throughput" and "System Response Time". Technical approaches with AI and DL modules for ensuring Cybersecurity and Information Security are proposed when servicing UDP network traffic between corporate clients. The verification of the developed modules for network diagnostics is based on the states of "Denial-of-Service Attack", "No Threat" and "Data Breach". A methodology for monitoring the state when prioritizing WEB traffic in active time zones for corporate clients based on hybrid AI, ML and DL modules, applicable to the structures of Media Service Providers and WEB Traffic Data Providers has been introduced.

Keywords

Artificial Intelligence; Machine Learning; Deep Learning; Analytics; Noise Effects Assessment; Traffic Diagnostics; Optimization of Transmission Media; Cybersecurity and Information Security.