ЗАТВЕРДЖЕНО

Наказ Міністерства освіти і науки, України

29.03.2012 № 384

**Форма № Н-9.02**

Національний університет «Полтавська політехніка ім. Ю.Кондратюка»

(повне найменування вищого навчального закладу)

Навчально-науковий інститут інформаційних технологій та робототехніки

(повна назва інституту)

Кафедра комп’ютерних та інформаційних технологій і систем 

(повна назва кафедри)

**Пояснювальна записка**

**до дипломного проекту (роботи)**

бакалавра 

(рівень вищої освіти)

на тему

Нейромережеве забезпечення передпродажної підготовки нерухомості.

Виконав: студент 4 курсу, групи 402-ТК

спеціальності

\_\_\_\_\_\_\_\_\_\_123 Комп’ютерна інженерія\_\_\_\_\_\_\_\_\_\_\_\_

(шифр і назва спеціальності)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_Терещенко А.В.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(прізвище та ініціали)

Керівник \_\_\_\_Альошин С. П. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(прізвище та ініціали)

Рецензент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(прізвище та ініціали)

Полтава – 2021 року

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ПОЛТАВСЬКА ПОЛІТЕХНІКА ІМ. Ю.КОНДРАТЮКА»**

**НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА РОБОТОТЕХНІКИ**

**КАФЕДРА КОМП’ЮТЕРНИХ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ І СИСТЕМ**

**КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА**

**спеціальність 123 «Комп’ютерна інженерія»**

**на тему**

**«Нейромережеве забезпечення передпродажної підготовки нерухомості.»**

**Студента групи 402-ТК Терещенко А.В.**

Керівник роботи

кандидат технічних наук,

доцент Альошин С.П.

Завідувач кафедри

кандидат технічних наук,

доцент Головко Г.В.

Полтава – 2021**РЕФЕРАТ**

Кваліфікаційна робота бакалавра: 61 сторінки, 31 рисунок, 5 джерел, 1 додаток.

**Об’єкт дослідження:** Інформаційне забезпечення передпродажної підготовки нерухомості.

**Предмет дослідження:** Нейромережеве забезпечення передпродажної підготовки нерухомості.

**Мета роботи**: дослідження нейромережевих технологій забезпечення передпродажної підготовки нерухомості у задачі створення нейромережевого забезпечення.

**Методи**: математичний аналіз, нейромережевий аналіз, класифікація, регресія, кластеризація.

**Ключові слова**: нейромережеве забезпечення, інтелектуальний класифікатор, простір ознак, нейромережеві технології.

**ABSTRACT**

Bachelor's thesis: 55 pages., 31 pictures., 15 sources, 1 appendix.

**Object of research**: Information support of pre-sale preparation of real estate

**Subject of research**: Neural network provision of pre-sale preparation of real estate

**Purpose**: research of neural network technologies to provide pre-sale preparation of real estate in the task of creating neural network support.

**Methods**: mathematical analysis, neural network analysis, classification, regression, clustering.

**Keywords**: neural network support, intelligent classifier, feature space, neural network technologies.

**ЗМІСТ**

[ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ ТА ТЕРМІНІВ 6](#_Toc75043800)

[ВСТУП 7](#_Toc75043801)

[РОЗДІЛ 1. 10](#_Toc75043802)

[Теоретична частина 10](#_Toc75043803)

[1.1 Перспективи технології Data Mining 10](#_Toc75043804)

[1.2 Методи і задачі Data Mining 14](#_Toc75043813)

[1.3 Постановка проблеми. 16](#_Toc75043814)

[РОЗДІЛ 2. 24](#_Toc75043815)

[Практична частина 24](#_Toc75043816)

[2.1. Технологія реалізації проекту за допомогою штучних нейронних мереж 24](#_Toc75043817)

[РОЗДІЛ 3 33](#_Toc75043818)

[3.1. Етап розробки 33](#_Toc75043819)

[3.2. Постановка задачі. 40](#_Toc75043820)

[3.3. Розв’язання задачі. 40](#_Toc75043821)

[ВИСНОВКИ 46](#_Toc75043822)

[СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ 47](#_Toc75043823)

[Додаток А 49](#_Toc75043824)

# ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ ТА ТЕРМІНІВ

**ШІ –** штучний інтелект, здатність інженерної системи (англ. engineered system) здобувати, обробляти та застосовувати знання та вміння.

**ШНМ (ANN)** **–** це обчислювальні системи, натхнені біологічними нейронними мережами, що складають мозок тварин.

**ЕОМ –** електронна обчислювальна машина, загальна назва для обчислювальних машин, що є електронними.

**Data Mining –** процес напівавтоматичного аналізу великих баз даних з метою пошуку корисних фактів.

**ЕС –** Експертна система, це методологія адаптації алгоритму успішних рішень однієї сфери науково-практичної діяльності в іншу.

# ВСТУП

Останніми роками впровадження нових технологічних процесів інтенсивно зросло у всіх сферах людської діяльності. Це стосується різноманітних галузей промисловості, сільського господарства, досліджень та галузей, пов’язаних із підтримкою людського життя (медичної, соціальної, домашньої).

Ці процеси відбуваються в контексті швидких і динамічних змін у навколишньому середовищі та високої швидкості інформації. Стратегічне планування та прийняття управлінських рішень важливі для забезпечення ефективної діяльності та поступового розвитку кожного суб'єкта господарювання в таких умовах.

На цьому етапі розробки вирішення різноманітних набагато складніших завдань управління пояснюється тим, що, крім усіх основних функціональних можливостей складної системи, об’єкти та процеси мають такі можливості: Існує проблема. :велика кількість слабоформалізованих і часто суперечливих цілей функціонування, їх мінливість у часі;

* Суперечливий багатогранний характер взаємозв'язку між елементами всередині об'єкта (процесу) під сильним впливом людських факторів та навколишніх об'єктів (процесу).
* В основному не відповідає концептуальному характеру початкового пояснення обмежень, що враховуються умовами експлуатації.

Створення та використання системи підтримки прийняття управлінських рішень у щоденних операціях є однією з найважливіших умов успіху діяльності економічних агентів. Не дивно, що пріоритетами прийняття рішень є ті, хто має здатність мислити стратегічно та передбачати нові події.

Однак одним із недоліків людського інтелекту є те, що він не пристосований для виконання великих обсягів обчислень при аналізі складних процесів та систем, що складаються з ланцюгів взаємозв’язків. Отже, на ефективність управління значний вплив має обмежена здатність людей обробляти складну та різну в часі інформацію.

З огляду на ці властивості, неможливо побудувати точну формальну модель об’єкта, орієнтованого на дані, і вирішення проблеми було розроблене іншими способами, особливо в теорії штучного інтелекту (ШІ). в класі методу. ).

Штучний інтелект - відносно молода наукова галузь. Історично склалося два підходи до розробки систем ШІ. Проектування нейронних мереж та розробка програмного забезпечення, що імітує розумову діяльність людини при вирішенні неформалізованих проблем класу. Останнім часом найбільш розроблений другий підхід.

Під штучним інтелектом розуміються складні наукові напрямки. Його метою є створення та використання програмного та апаратного забезпечення, яке може моделювати процес людського мислення (окремі функції творчої діяльності) та забезпечувати взаємодію людей з комп’ютерами їхньою природною мовою.

Протягом десятиліть дослідники обговорювали визначення галузі знань, яка називається штучним інтелектом. Однак глобальною метою досліджень і розробок тут є створення штучних систем, що виявляють інтелектуальну поведінку. Наукові інтереси стосуються розробки як теоретичних, так і формальних пристосувань для побудови моделей інтелектуальної діяльності людини, а також штучних систем, в яких теоретичні моделі впроваджуються та перевіряються. Це створює багато місцевих цілей моделювання інтелектуальної поведінки людини, які часто формулюються з точки зору інтелектуальних проблем. Беручи до уваги розвиток штучного інтелекту за останні 50 років і протягом короткого періоду, ви можете визначити ключові цілі за періодами.

- До 1970 р. - підтвердження можливості виявлення інтелектуальної поведінки комп’ютерів.

- До 1990 року - Неминуче створення спеціальних засобів штучного інтелекту, включаючи комп'ютери (спеціалізація полягає в імітації поведінки людини при вирішенні складних додатків).

- До 2010 року - побудова інтелектуальних комп’ютерних систем з інтегрованою інтелектуальною поведінкою. Його важливою властивістю є пристосованість до змін у навколишньому середовищі.

Штучний інтелект - Технічні (у всіх сучасних випадках - комп’ютери) системи, що мають специфічні можливості інтелекту, тобто вони можуть:

* Розуміти та розпызнавати;
* Приймати рішення та знаходити спосіб досягнення результату ;
* Навчатися.

Насправді, наявність лише неповного знання функції мозку перешкоджає побудові його наближеної інформаційної моделі для моделювання найскладніших процесів мислення, включаючи творчі, в ECOM, у тому числі й творчі.

# РОЗДІЛ 1

# Теоретична частина

1.1 Перспективи технології Data Mining

Нещодавно в Україні було відкрито філію Всесвітнього центру обробки даних та Національної мережевої інфраструктури. Це дозволило вітчизняним вченим та професіоналам очікувати збільшення даних з різних областей, що обробляються інтегрованою мережею вітчизняних кластерів. З розвитком методів запису та зберігання даних обсяг зібраної та проаналізованої інформації швидко зростає. Потреба в такому аналізі дуже очевидна, оскільки ці "необроблені дані" - це знання, які можна використовувати для прийняття рішень, але обсяг даних настільки великий, що неможливо проаналізувати дані самостійно.

Видобуток даних використовується для автоматичного аналізу даних. Це нова технологія інтелектуального аналізу даних, яка визначає приховані закономірності у вигляді ключових особливостей, кореляцій, тенденцій та шаблонів. Сучасні системи обробки даних використовують штучний інтелект, заснований на методах вираження та інтерпретації. Це не показує, що він розчиняється в терабайтах сховища, але він може знайти дуже цінну інформацію. Насправді в процесі видобутку даних ми говоримо про те, що система забезпечує себе на основі аналізу, а не на попередньо запропонованих гіпотезах.[1]

Існує багато визначень інтелектуального аналізу даних, але загалом вони узгоджуються у виборі чотирьох ключових особливостей. Згідно визначенню, Г. Пиатецкого-Шаниро (G. Pia-tetsky Shapiro,GTE Labs), одного з ведучих світових експертів в даній області, Data Mining — дослідження і виявлення алгоритмами, засобами штучного інтелекту в "сирих даних" прихованих структур, шаблонів або залежності, яка:

* не були відомі;
* нетривіальні;
* практично корисні;
* доступні для інтерпретації людиною і необхідні для ухвалення рішень в різних сферах діяльності.

Деталі останніх вимог до продуктивної обробки інформації такі:

* дані мають необмежений обсяг;
* дані є різнорідними (кількісними, якісними, текстовими);
* результати повинні бути конкретний і зрозумілий;
* інструменти для обробки "сирих даних" повинні бути прості у використовуванні. [4]

Традиційна математична статистика, яка тривалий час претендувала на роль головного інструменту в аналізі даних, не вирішила проблеми. Основною причиною є концепція усереднення вибірки, яка веде до маніпулювання фіктивними величинами. Математичні статистичні методи виявились корисними насамперед для перевірки заздалегідь сформульованих гіпотез та для "грубого аналізу інтелекту", який лежить в основі оперативної аналітичної обробки даних OLAP. Сучасна технологія аналізу даних базується на концепції шаблонів, що відображають фрагменти багатогранних взаємозв’язків у даних. Ці шаблони унікальні для підмножини даних і можуть бути представлені компактно у зрозумілому для людини форматі. Пошук шаблону виконується таким чином, що не обмежується попередніми припущеннями про структуру вибірки та тип розподілу аналізованих метричних значень. Причини популярності інтелектуального аналізу даних:

* стрімке накопичення даних (рахунок йде вже на экзабайти);
* загальна комп'ютеризація бізнес-процесів;
* проникнення Інтернет у всі сфери діяльності;
* прогрес в області інформаційних технологій: вдосконалення СУБД і сховищ даних; прогрес в області виробничих технологій: стрімке зростання продуктивності комп'ютерів, об'ємів накопичувачів, впровадження Grid систем.[2]

Алгоритми, що використовуються при інтелектуальному аналізі даних, вимагають великої кількості розрахунків. Раніше це заважало широкій комерціалізації інтелектуального аналізу даних, але сучасні вдосконалення продуктивності процесорів усунули серйозність цієї проблеми. Тепер ви можете виконати якісний аналіз сотень тисяч і мільйонів записів за розумний проміжок часу. Видобуток даних - міждисциплінарні дисципліни, що виникають і розробляються на основі таких наук, як прикладна статистика, розпізнавання образів, штучний інтелект та теорія баз даних, дивитися рисунок 1.1. [3]

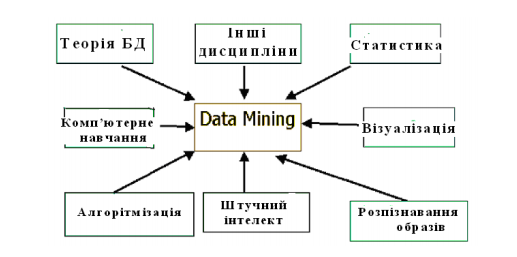


Рисунок. 1.1 – Data Mining як міждисциплінарна галузь

Потенціал інтелектуального аналізу даних дає «зелене світло» для розширення сфери застосування цієї технології. Щодо перспектив інтелектуального аналізу даних, можливі наступні напрямки розробки:

* Вибір предметного типу галузі та відповідної евристики. Його формалізація сприяє вирішенню супутніх питань з видобутку даних, що стосуються цих галузей;
* Створення формальної мови та логічні засоби формалізують умовивід, а його автоматизація буде інструментом для визначення проблем видобування даних у конкретних предметних областях;
* Створення методів інтелектуального аналізу даних, які можуть не тільки витягувати шаблони з даних, а й формувати деякі теорії на основі емпіричних даних;

Подолання великого розриву між можливостями засобів обробки даних та теоретичними результатами цієї галузі. []

Дивлячись на майбутнє видобутку даних у короткостроковій перспективі, стає очевидним, що розвиток цієї технології найбільше зосереджений на галузі, пов'язаній з мережевими системами електронної науки. Особливості електронної науки характеризують обчислювальну інфраструктуру, що складається з трьох концептуальних рівнів (рисунок 1.2) :

# C:\Users\Artemka\Desktop\dfh.pngРисунок 1.2 – Трирівнева архітектура Grid-сервісів

* Дата / обчислювальні послуги. Цей рівень містить інформацію про те, як розташовані обчислювальні ресурси, коли вони планують виконувати обчислювальні засоби та засоби передачі даних між різними обчислювальними ресурсами. Цей рівень може обробляти великі обсяги даних, забезпечує високошвидкісну мережу та різноманітні ресурси як єдиний метакомп’ютер.
* Інформаційна служба. Цей рівень показує, як інформація продається, зберігається, і хто і як може до неї отримати доступ. Тут інформація розуміється як цінні дані. Наприклад, він розпізнає, що рядок - це ім’я людини, і виявляє ціле число як зображення температури процесу реакції.
* Послуги знань. Цей рівень надає користувачам можливість набувати, використовувати, шукати та публікувати знання, щоб допомогти їм досягти своїх конкретних цілей. Знання представлені тут як інформація, яка використовується для досягнення цілей, вирішення проблем або прийняття рішень. Як приклад, існує процедура, при якій оператор знає за температури реакції, скільки часу потрібно завершити.[6]

Розглянуті поняття є невід’ємною частиною так званої інформаційної піраміди, на якій є дані, наступним рівнем є інформація, потім є рішення, а рівень знань завершує піраміду. По мірі просування інформаційної піраміди обсяг даних стає цінністю прийняття рішень, цінністю знань. Як ви можете бачити на малюнку 2, цей процес є циклічним. Для прийняття рішень потрібна інформація на основі даних. Дані містять інформацію, яка підтримує рішення тощо.

Кожна сіткова система, яка вже побудована або буде побудована, містить деякі елементи всіх трьох рівнів. Важливість використання цих рівнів залежить від вас. Тому в деяких випадках обробка великих обсягів даних може бути головним завданням, тоді як підтримка знань може бути основною проблемою.

На сьогоднішній день більшість досліджень у сфері електромереж зосереджено на рівнях даних / обчислень та інформації. Водночас є ще багато відкритих питань щодо управління широкомасштабними розподіленими обчисленнями та ефективного доступу та розповсюдження інформації з неоднорідних джерел. Вважається, що максимізація потенціалу обчислювальних мереж вимагає максимізації можливостей, що надаються рівнем знань, тому цей рівень забезпечує прямий та легкий доступ до операцій та діалогу.[9]

1.2 Методи і задачі Data Mining

Основною функцією видобутку даних є поєднання широкого спектру математичних інструментів (від класичного статистичного аналізу до нових кібернетичних методів) та останніх досягнень інформаційних технологій. Суто формальні та неформальні методи аналізу, кількісний та якісний аналіз даних, гармонійно поєднуються в технології аналізу даних.

Методи та алгоритми видобутку даних включають: Штучні нейронні мережі, дерева прийняття рішень, символічні правила, найближчі сусіди та k-найближчі сусіди, векторні еталонні методи, байєсівські мережі, лінійна регресія, кореляційний регресійний аналіз. Ієрархічні методи кластерного аналізу, неієрархічні методи кластерного аналізу, включаючи k-середні та k-медіанні алгоритми. Як шукати асоціативні правила, включаючи алгоритм Апріорі. Обмежені методи пошуку, еволюційне програмування та генетичні алгоритми, різні методи візуалізації даних та багато інших методів. [11]

Більшість аналітичних методів, що використовуються в техніках аналізу даних, є відомими математичними алгоритмами та методами. Нові функції в цих додатках - це потенціал для нових апаратних та програмних функцій, які дозволяють використовувати їх для вирішення конкретних проблем. Слід зазначити, що більшість методів видобутку даних були розроблені в рамках теорії штучного інтелекту.

Не існує єдиної думки щодо того, які завдання слід віднести до аналізу даних. До найбільш надійних джерел належать класифікація, кластеризація, прогнозування, асоціація, візуалізація, аналіз та виявлення відхилень, оцінка, аналіз взаємозв’язків та узагальнення. Розглянемо деякі з них.[7]

**Класифікація** (Classification). Це найпростіша і найпоширеніша задача інтелектуального аналізу даних. Можливість охарактеризувати групу об’єктів у досліджуваному наборі даних в результаті вирішення класифікаційної задачі-класу. З цих причин нові об’єкти можуть бути віднесені до певного класу. Для вирішення проблеми класифікації можна використовувати такі методи: Найближчий сусід. k-найближчі сусіди; байєсівська мережа; індукція дерева рішень; нейронна мережа.[14]

**Кластеризація** (Clustering) Кластеризація є логічним продовженням концепції класифікації. Це завдання є більш складним. Особливістю кластеризації є те, що клас об’єктів спочатку не визначений. В результаті кластеризації об'єкти поділяються на групи. Прикладом методу кластеризації є спеціальний тип нейронної мережі (карта Кохонена), яка самоорганізовується без контрольованого навчання.

**Асоціація** (Associations). У процесі вирішення проблеми пошуку асоціативного правила ви знайдете регулярність між пов’язаними подіями у наборі даних. Відмінності між попередніми двома завданнями асоціації та видобутку даних: Пошук шаблонів здійснюється не на основі властивостей об’єкта, а між кількома подіями, що відбуваються одночасно. Найвідоміший алгоритм вирішення проблеми пошуку правил асоціації-алгоритмів Apriori.

**Послідовність** (Sequence), Крім того, ви можете використовувати послідовні послідовності асоціацій для пошуку часових зразків між транзакціями. Проблема послідовності подібна до асоціації, але мета полягає у встановленні закономірності між подіями, які пов'язані в часі (тобто, що відбуваються через певні інтервали часу), а не між подіями, що відбуваються одночасно. Видобуток даних для цього завдання також відомий як проблема пошуку послідовності. візерунок (sequential pattern).Правило послідовності: після події X через певний час відбудеться подія У.[10]

**Прогнозування** (Forecasting). В результаті визначення проблем прогнозу на основі характеристик існуючих даних оцінюються пропущені або майбутні значення цільових числових показників. Для вирішення цих проблем широко використовуються такі методи, як математична статистика та нейронні мережі.

**Візуалізація** (Visualization, Graph Mining) Візуалізація створює графічне зображення проаналізованих даних. Для вирішення задач візуалізації використовують графічний метод відображення присутності шаблонів у даних. Прикладами методів візуалізації є відображення даних у 2D та 3D.

**Підведення підсумків** (Summarization) – завдання мета якого є опис певної групи об’єктів, наприклад, з аналізованого набору даних.

Залежно від використовуваної моделі, завдання аналізу даних можуть бути описовими та передбачувальними. В результаті вирішення описової задачі аналітик отримує шаблон, що описує інтерпретовані дані. Ці завдання пояснюють загальну концепцію аналізованих даних та визначають корисні та кінцеві характерні ознаки даних. Завдання прогнозування базуються на аналізі даних, моделюванні, прогнозуванні тенденцій або властивостях нових або невідомих даних.[12]

1.3 Постановка проблеми

Спираючись на приналежність до нейромережевої системи, можна розробити перспективний напрямок корпоративної діяльності на основі математичного моделювання прогнозованого розміру прибутку торгової компанії, за умови оцінки впливу внутрішніх факторів та зовнішніх ризиків. Суб'єктивна (експертна) модель оцінки та прогнозування.

Сьогодні штучні нейронні мережі вирішують великі класи завдань, обробку інформації, насамперед за рахунок ідентифікації, емуляції, інтелектуального управління, передбачення часових рядів довільної природи та умов структурної та параметричної невизначеності. Нейронні мережі - це перспективна обчислювальна технологія, яка забезпечує нові припущення для вивчення контрольно-аналітичних завдань у діяльності комерційних підприємств. [8]

У складному економічному аналізі з використанням методів застосування нейронних мереж оцінюється поточний стан торгових компаній, паралельно також проводиться причинно-наслідковий аналіз змін основних показників під впливом впливових факторів. У той же час у процесі аналізу фінансово-економічного становища комерційних компаній продемонстровано ключові показники з урахуванням короткострокових та довгострокових прогнозів, і їх цінність є комерційними компаніями у високо конкурентне середовище.

Тому необхідно побудувати динамічний інструмент управління (штучна нейронна мережа). Це може враховувати вплив адміністративних рішень на прибутковість та вартість комерційних компаній. За необхідності можна збільшити кількість метрик вхідних даних. Це розширює та покращує інформаційну підтримку в процесі контролю та аналізу, який управляє діяльністю торгових компаній для оцінки поточної ситуації, а також для вибору конкурентних стратегій.

Кількісна та якісна параметризація на основі фактично обґрунтованого аналізу вибору стратегії, розробки показників та бізнес-планів, формування довгострокових стратегій ділового партнерства та оцінки реалізації стратегії відбувається в рамках двох процесів. Перший - це процес формулювання. Це пізнавальний та рішучий процес, спрямований на визначення ключових стратегічних цілей компанії та визначення ключових перешкод нашляху їх досягнення. Другий - процес впровадження. Розробка контрзаходів. Його застосування призведе до досягнення запланованих стратегічних цілей економічних (високий рівень дохідності та прибутко‑ вості, економічного потенціалу та ефективних джерел фінансування, наявність платоспроможності);

* Технічно-технічні (наявність передових систем управління програмним забезпеченням, низький рівень зносу основних та нематеріальних активів, високий рівень інформаційного забезпечення завдяки використанню інформаційних технологій);
* Організаційно-адміністративні (високий рівень корпоративної культури, ефективна організаційна структура, ступінь мотивації персоналу, управлінський досвід);
* Маркетинг (наявність ефективних стратегій просування продукції на ринку, оптимальна цінова політика, розвинені мережі продажів та агентств, впізнавані бренди, позитивний діловий імідж);соціально-психологічних (психологічний настрій споживачів товарів, довіра з боку населення тощо);
* Географічна (географічна присутність комерційних компаній у районах з розвиненою ринковою інфраструктурою та платоспроможним споживчим попитом).

Аналіз останніх досліджень та публікацій. Питання аналітичного обґрунтування конкурентних стратегій знайшло своє відображення у багатьох наукових розробках вітчизняних та зарубіжних вчених, зокрема: Гордієнко, І. Ансофф, Г.Л.Кармін, Н. І.Ко Уолл, Л.А. Костирко, Н. В. Куденко, К. І. Редченко, В. Д. Нємцов, Є.М. Сич, С.І.Шкарабан та ін. Однак у цих завданнях аналітична підтримка виділяється лише з точки зору вибору та демонстрації стратегії діяльності або щодо оперативного аналізу запланованих регульованих заходів. Діагностика можливих операцій з досягнення стратегічних цілей, оцінка підприємницької стратегічної та тактичної інтеграції, аналіз ризиків тактичних рішень за критеріями досягнення стратегічних цілей тощо не розглядаються. Вирішення цих питань надає можливість досягти прозорості, переконливості у стратегічному та тактичному управлінні та мінімізувати ризик досягнення цілей.

Тому пропонується використовувати математичні інструменти теорії нечітких множин, що використовуються у дослідженнях такими дослідниками, як Круглов В.В., Матвійчук А.В., Божич В.І., Лебедєв О.Б., Шніцер Ю.Л., Зоріна О.І. [1–9] тощо. Тому метою цієї статті є вивчення деталей формування конкурентних стратегій, заснованих на використанні нейромережевих технологій для забезпечення конкурентоспроможності та прибутковості комерційних компаній.[9]

Презентація основних дослідницьких матеріалів. Використання методів прогнозування корпоративної діяльності на основі використання нейронних мереж характеризується багатьма сильними і слабкими сторонами. Перевага полягає в тому, що нейронні мережі дозволяють досліджувати залежність передбачуваних значень від незалежних змінних на основі числових та текстових даних. Однак лише в тому випадку, якщо невідома закономірність аналізу не потребує вирішення проблеми взаємозалежності між вхідними показниками. Аналітикам не потрібно знати про передові технічні можливості нейронних мереж. Це дозволяє припустити, що майбутні продажі залежатимуть від таких параметрів:

* продажі в останній період;
* продажі у передостанній період кількість робочих днів тощо.

Однак потрібно враховувати неконтрольовані фактори навколишнього середовища, особливо сезонність, активність конкурентів у корпоративному домені, кількість покупців та час доставки товарів. Нейронні мережі дозволяють врахувати основні фактори для складання короткострокових прогнозів. Ви можете використовувати архітектуру нейронної мережі (персептрон з одним прихованим шаром) та базу даних (роздрібні продажі та інші дані із зовнішнього та внутрішнього середовищ), щоб отримати ефективну систему прогнозування. Потрібно враховувати зовнішні параметри, включаючи відповідні входи до нейронної мережі. Він використовує алгоритм для визначення важливості та важливості вхідних змінних, за винятком параметрів, які мають незначний ефект. Перевага нейронних мереж полягає в тому, що експерти не вибирають математичних моделей поведінки часових рядів.

Побудова моделі нейронної мережі виконується адаптивно без участі експертів у процесі навчання. Недоліком цієї техніки є те, що для неї потрібні спеціалізовані програмні засоби. Значима складність інтерпретації та невизначеність нейронних мереж. Це так звана "чорна скринька", де логіка прийняття рішень нейронних мереж прихована від експертів. Тому ця модель не може чітко та прозоро визначити внесок кожного показника у покращення або погіршення фінансового стану торгової компанії. Для цього існує алгоритм «вилучення знань з нейронних мереж». Він формалізує перелік логічних правил та створює експертну систему на основі мережі. [10]

Однак ці алгоритми не вбудовані в нейромережеві пакети, і набір правил, що генеруються такими алгоритмами, дуже великий. Тому для вирішення проблеми аналітичного забезпечення прогнозування конкурентної стратегії комерційних підприємств ми є частиною запропонованої системи методів оцінки та прогнозування діяльності комерційних підприємств. Пропонується використовувати метод. Нейрони, інтегровані в систему певної архітектури, називаються нейронними мережами, тип яких визначається завданнями, що стоять перед підприємством.

Важливими параметрами для побудови нейронної мережі є алгоритми оптимізації та здатність до навчання. Алгоритми оптимізації дуже важливі для побудови нейронних мереж, оскільки вони можуть згладжувати помилки в процесі формування навчального набору та прискорювати навчання. В результаті алгоритм виключається, а вхідний компонент виключається (видаляються дані, що погіршують загальний результат). ). Поведінка нейронних мереж оцінюється на основі реакції відносно простих елементів одного типу. Для кожного нейрона синапси (односпрямовані вхідні сигнали x1, ... xn, пов'язані з іншими прихованими нейронами) та аксони (початкові значення y1, ... ym, sv відносяться як до прихованих, так і до вхідних нейронів). Під час підготовки до прогнозування весь набір даних розділений на три підвибірки. Перший - це навчання. Цей тренінг відбирає набір прикладів для ефективної роботи мережі. Кожен приклад містить пару даних для входу xi та виводу yj. Дані з серії спостережень xi використовуються для навчання персептрону. У цьому випадку кожен нейрон прихованого шару отримує сигнал від нейрона вхідного шару.

Після виконання операції над сигналом ваги нейронів змінюються, і його вихід надсилається всім нейронам наступного шару, забезпечуючи одну пряму (пряму) передачу на вихід yj та центрифугування. Забезпечуються гендерні зв’язки. Друга під вибірка вважається валідацією, оскільки вона покликана дати можливість оцінити прогнози та визначити оптимальну складність моделі. Остаточна під вибірка використовується для оцінки обґрунтованості та реалістичності запропонованої моделі. Вона перевіряє мережу після тренувань.

Існує чотири етапи розробки та використання нейронних мереж за допомогою нейронних інструментів. Крок 1. Дані, що використовуються Neural Tools, визначаються у наборі даних, який ви готуєте. Менеджер наборів даних дозволяє налаштувати ваш набір даних для повторного використання при проектуванні мережі. Вхідні дані були введені перед моделюванням нейронної мережі. Цей крок дозволив нам позбутися значного діапазону даних і отримати налаштований діапазон змінних. Варіант був обраний, якщо стандартизація метрики базувалася на змінній шкалі. [8]

Під час навчання нейронна мережа створюється на основі набору даних. Адже на момент прогнозування конфігурація нейронної мережі зчитується з бази даних.

Результати прогнозування відображаються користувачеві у вигляді звіту HTML. Коли навчання буде завершено, конфігурація мережі буде збережена в базі даних. Коли програма працює, вона використовує базу даних для зберігання інформації. Як зазначив ОА, Зоріна сказала: "Найважливішим рішенням, яке аналітик повинен прийняти при використанні нейронної мережі, є пояснення змодельованого процесу аналізу фінансово-економічної діяльності компанії. Серія вибору змінних".

Отже, дані базуються на спостереженнях, де відомі значення як залежних, так і незалежних змінних. Згідно з опитуванням, чинниками комунікації першого рівня є чистий прибуток, вартість реалізації, рентабельність продажів на похідному рівні, швидкість обслуговування споживачів, а факторами другого рівня є ціна продажу, асортимент продукції та організація збуту. Ритми, ресурси та їх використання індивідуально.

Під час тестування нейронної мережі ми перевірили нашу здатність прогнозувати початкові значення. Дані, що використовувались для тесту, являли собою конкретний набір ретроспективних даних.

Основні стратегічні показники, розраховані в прогнозі, повинні бути основою вашої стратегії. Тому ми пропонуємо відношення ринкової вартості компанії до вартості її активів як ключовий стратегічний параметр. Цінність компанії з абсолютної точки зору може бути використана для характеристики компанії за її критеріями фінансової стійкості.

Серед інших стратегічних показників ми виділили прибутковість, валовий прибуток, оборотні активи та запаси. Розроблена нейронна мережа використовується для прогнозування невідомих початкових значень. За допомогою інструменту NeuralTools ви можете встановити параметри для автоматичного пошуку найкращої мережі з правильними стратегічними показниками. Згідно з опитуванням, чистий прибуток має тенденцію до зростання приблизно на 20% щороку, що є результатом ефективного управління торговими компаніями при збереженні конкурентних переваг та позицій на ринку. Оцінка суми ліквідних активів та запасів свідчить про збільшення суми дебіторської заборгованості та грошових коштів. Це свідчить про те, що компанія балансує на рівні стратегії зростання. Питання точного та чіткого вибору стратегії є дуже високим, оскільки короткостроковий вибір стратегії та довгострокова реалізація залежать від економічного та політичного середовища, а також юридичної підтримки компаній різних форм організації. [15]

Тому використовуйте стійку, збалансовану корпоративну стратегію зростання, засновану на "ключових факторах успіху", що захищає частку ринку на основі оптимізації структури капіталу управління грошовими потоками та управління фінансовими ризиками на основі даних. Структурні вдосконалення, зміна товару (приватна марка); підтримувати баланс між якістю обслуговування та якістю зв'язку, впроваджувати лідерство в цінах (лідерство витрат), оптимізувати грошовий потік, оцінювати та передавати окремі бізнес-процеси на аутсорсинг.

Головною перевагою цього методу є те, що він збільшує дохід від продажу товарів за рахунок підвищення рівня сервісу, точності постачання та прогнозування попиту на відповідність договірній дисципліні. Скоротити запаси, мінімізувати накладні витрати, складування та продаж, а також трансакційні витрати, покращити використання логістичних можливостей, забезпечити орієнтованість на клієнта бізнес-процесів та відкритість для обміну знаннями між діловими партнерами Підвищити соціальну відповідальність підприємств та врахувати сучасні екологічні проблеми.[14]

# 

# 

# РОЗДІЛ 2

# Практична частина

# 2.1. **Технологія реалізації проекту за допомогою штучних нейронних мереж**

1. Відкрийте файл даних Series\_g із даних у пакеті. Цей файл містить єдину змінну, яка визначає обсяг трафіку за кілька років за допомогою щомісячного протоколу. (Відкривши цей файл, ви побачите деякі таблиці, пов’язані з опціями розумного розв’язування. Ці таблиці на цьому етапі слід закрити, залишивши лише вихідну таблицю).

2. Встановіть тип змінної "вхід-вихід" наступним чином: Клацніть на заголовок таблиці, щоб вибрати змінну, клацніть правою кнопкою миші та виберіть Параметри меню Вхід / Вивід (Вхід / Вивід). Назви змінних виділені зеленим кольором.

3. Використовуйте діалогове вікно Створити мережу, щоб створити нову мережу.

У питаннях прогнозування часових рядів мережа повинна знати, скільки копій змінної потрібно отримати і наскільки далеко слід передбачити значення змінної. У цьому завданні використовується параметр Steps (Тимчасове вікно), рівний 12, і параметр Lookahead (Horizon) -1, оскільки дані є щомісячними спостереженнями.

4. Виберіть Multilayer Perceptron як тип мережі та встановіть кількість мережевих шарів на 3. Потім натисніть кнопку Порада, і програма автоматично встановить кількість нейронів у всіх трьох мережевих шарах. 12-6-1 ( рисунок. 2.1).

5. Коли ви створюєте мережу, програма SNN автоматично призначає перші 12 спостережень із файлу даних типу Ігнорувати (не враховано). Подальше навчання та маніпулювання мережею для завдання аналізу часових рядів включатиме дані, що належать до окремих спостережень, кожна з яких надсилається у вхідний блок даних. Весь блок обумовлений спостереженнями, які включають значення вихідних змінних. Як результат, перші 12 спостережень насправді не ігноруються, але вони є вхідними даними першого блоку даних у часових рядах і відповідають спостереженню №13. Насправді програма створює трансформований набір даних на 12 менших спостережень, але дані для кожного спостереження беруться з 13 послідовних рядків у вихідному файлі. Створена мережа показана в рисунок. 2.1.

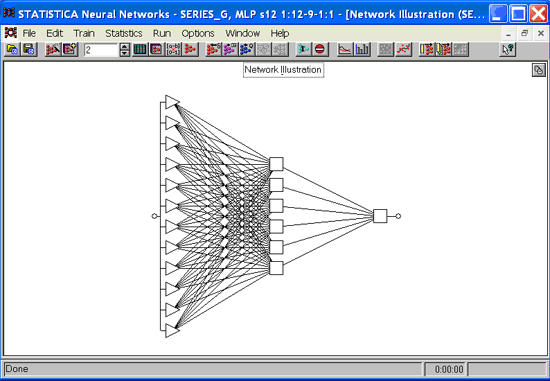


Рисунок. 2.1 – Тришаровий персептрон

6. У Редакторі набору даних у вікні Джерело встановіть 66 навчальних (навчальних) та 66 контрольних (перевірочних) спостережень (Рисунок 2.2) та клацніть рядок перетасування наступним чином: З меню [Редагувати] - [Справа] - [Перемішати] - [Усі] ([Редагувати] - [Справа]) - Перемішати - Усі).

7. Навчіть мережу по-Левенбергу-Маркварду. Потрібно натиснути: Train-Multilayer Perceptron-Levenberg-Marquardt (Teach-Multilayer Perceptron-Levenberg-Marcar). Процедура навчання займає кілька секунд (залежно від типу процесора). Метод Левенберга-Маркара є одним із надійних та швидких алгоритмів навчання, але його застосування має певні обмеження.:

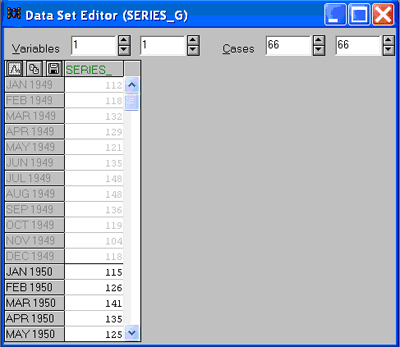


Рисунок. 2.2 – Вікно вихідних даних з розділеними спостереженнями

* Цей метод може бути використаний лише для мереж з одним вихідним елементом.
* Метод - Левенберга-Маркара вимагає пам'яті, пропорційної квадрату кількості ваг у мережі, тому цей метод не підходить для великих мереж. (близько 1000 ваг).
* Цей метод застосовується лише до функції середньоквадратичної помилки.

Алгоритм Левенберга-Маркара призначений для мінімізації функції середньоквадратичної помилки. Близько точки мінімуму це припущення зроблено настільки точно, що алгоритм рухається дуже швидко. Далеко від мінімуму, це припущення може бути не вірним, тому цей метод знаходить компроміс між лінійною моделлю та методом найкрутішого спуску.

Цей крок робиться лише тоді, коли помилка зменшується, і при необхідності для забезпечення прогресу використовується досить малий кроковий метод найкрутішого спуску.

Діалогове вікно методу Левенберга-Маркара показано на рисунок 2.3.

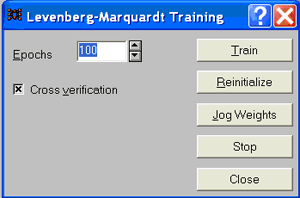


Рисунок 2.3 – Діалогове вікно методу Левенберга-Маркара

Основні елементи вікна:

* **Epochs** (Число епох) - вказує кількість епох, протягом яких буде виконуватися алгоритм. У кожну епоху весь набір тренувань проходить мережу, і ваги регулюються.
* **Cross-Verification** (Крос-перевірка) - у зазначеному місці якість, видана отриманою мережею, перевіряється в кожну епоху в наборі управління (якщо налаштовано). Якщо позиція вимкнена, контрольні спостереження ігноруються, навіть якщо вони існують у файлі даних.
* **Train** (Навчити) - кожного разу, коли ви натискаєте кнопку, алгоритм рухається через певну кількість епох.
* **Reinitialize** (Переустановити) - розмір мережі буде відновлений випадковим чином, тому вам потрібно буде натиснути кнопку перевстановлення перед початком нового тренінгу.
* **Jog Weights** (струшування ваг) - Цей параметр додає невелике значення кожній вазі, якщо алгоритм зафіксовано на локальному мінімумі.

8. Створіть проекцію часового ряду. Використовуйте проекцію серії виконання (Run-Time Series Projection), щоб відкрити відповідне вікно (рисунок 2.4).

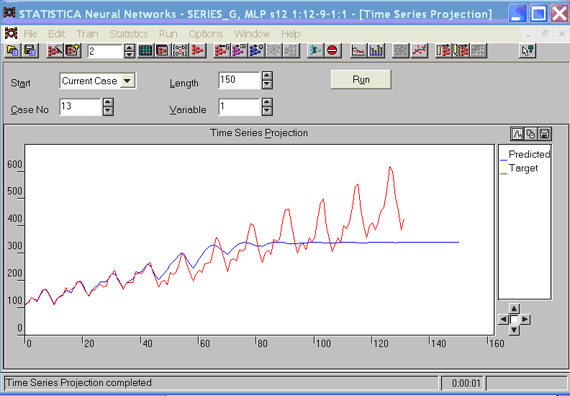


Рисунок 2.4 – Вікно проекції тимчасового ряду

Опис діалогового вікна

* **Start** (Початок) - Вказує, чи прогноз часових рядів починається з певної кількості спостережень (номерів справ) у файлі даних або з одного спостереження.
* **Case No** (Номер спостереження) - При проектуванні часового ряду з файлу даних номер спостереження позначається початковим значенням для початку.
* **Length** (Довжина) - Кількість передбачених кроків.
* **Variable** (Змінна) - Вказується змінна, яку потрібно проектувати.

9. Ви можете спроектувати часовий ряд за допомогою навченої мережі. Спочатку мережа обробляє перші 12 вхідних значень і передбачає наступні. Потім передбачене значення повертається на вхід мережі разом з попередніми 11 вхідними значеннями, остання забезпечує прогнозування наступного значення.

Єдиним контрольним параметром, який потрібно вибрати, є довжина проекції. У цьому прикладі є загалом 144 спостереження, 12 з яких видаляються під час попередньої обробки, тому ви можете порівняти результати до 132 кроків. Однак, хоча ви можете розробити більше, ніж доступні дані, результати порівнювати ні з чим.

Показує поведінку передбачуваних значень різної довжини. За допомогою кнопки Виконати можна спостерігати зміни цільових і вихідних значень серії..

На рисунку 2.4 показує, що крива прогнозування (синя на екрані монітора) недостатньо навчена, оскільки між початковою серією та прогнозованою серією існує значне відхилення, починаючи з приблизно 70 спостережень.

10. Використовуйте розумний вирішувач (третя кнопка ліворуч від верхнього рядка), щоб передбачити рядок. У цьому випадку вам потрібно відповісти на деякі запитання в інтерактивному режимі:

* Вибрати основну версію (рисунок. 2.5) і натиснути Next.

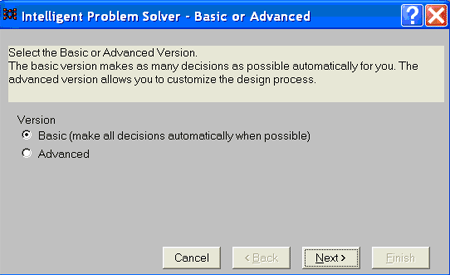


Рисунок 2.5 – Вибір основної версії

* Визначає тип завдання (стандартний або часовий ряд). Тут потрібно звернути увагу на часові ряди (рисунок 2.6).

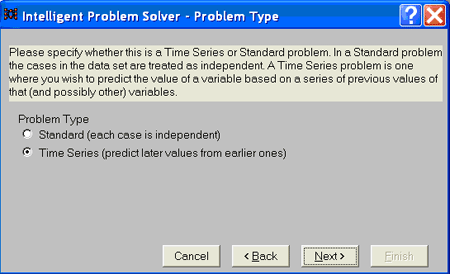


Рисунок 2.6 – Вибір типу завдання

* Встановіть період спостереження до 12 місяців (рисунок 2.7).

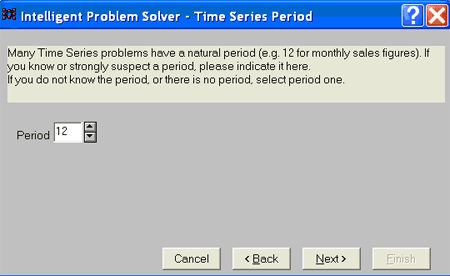


Рисунок 2.7 – Як встановити інтервал спостережень

* Виберіть залежні та незалежні змінні, що знаходяться в одному ряді змінних.
* Встановіть час процедури оплати до 2 хвилин (рисунок 2.8).

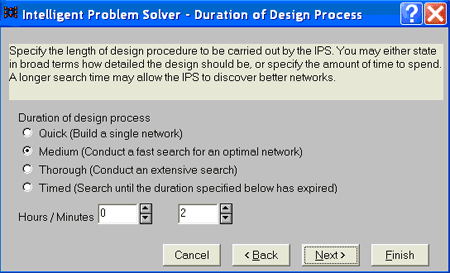


Рисунок 2.8 – Установка часу розрахункової процедури

* Зберегти, вказавши кількість мереж та дії, які потрібно зберегти

(рисунок 2.9).

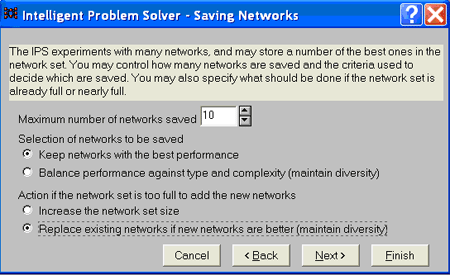


Рисунок 2.9 – Дії по вибору мереж

* Виберіть формат відображення результатів (рисунок 2.10) і натисніть кнопку Готово.

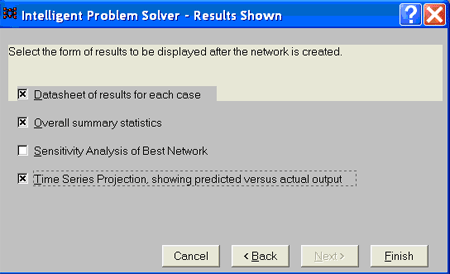


Рисунок 2.10 – Вибір форми представлення результатів

В результаті використання інтелектуального вирішувача передбачення набагато точніші, оскільки навчальна мережа набагато ближча до вихідного діапазону. (рисунок. 2.11).

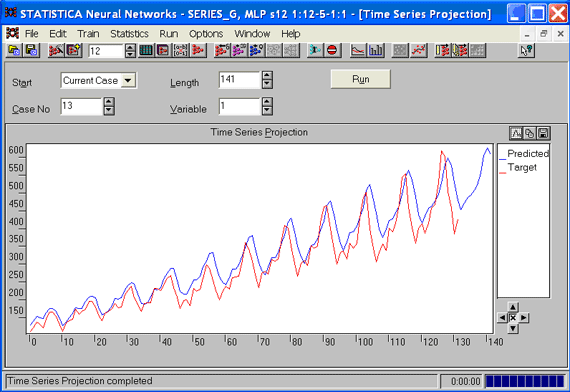


Рисунок 2.11 – Прогноз за допомогою інтелектуального вирішувача

# РОЗДІЛ 3

# 3.1. Етап розробки

Завдання прогнозування часових рядів знаходять практичне застосування в багатьох сферах бізнесу. Спроба передбачення різних розподілених в часі параметрів може принести хороші результати за умови, що завдання сформульована розумно і немає спроби вирішити усі виклики, які проблеми тільки за допомогою прогнозування. Суттєвими моментами в даній області є наступні:

* Хороший прогноз можна отримати практично тільки для детермінованого ряду, розвиток якого і так відомо.
* Прогноз передбачає, що в майбутньому не відбудеться яких-небудь істотних змін факторів, які можуть вплинути на ряд.
* Як правило, віддача від прогнозування досягається при застосуванні комплексних рішень, що включають не тільки методи прогнозу, а й засоби оптимізації.

Застосовуйте методи прогнозування для розгляду прикладів проблем, з якими стикаються виробники: завдання прогнозування продажів деяких товарів. Важко очікувати, що початковий графік (розподілений під час продажу) буде детермінованим. На продажі впливає багато факторів, але врахувати їх майже неможливо. Серія, яку ми вважаємо, сформувалася за два з половиною роки. У цей час реєструвались щотижневі відвантаження напівфабрикатів зі складу виробника.

Побудувати змодельований тимчасової ряд з пакета Statistica можна наступним чином:

* Створити новий файл, що складається з 20 рядків і 2 стовпців.
* Через меню Data - Variable Specs (Дані - опис змінної) ввести у вікно формул вираз = vnormal (rnd (1); 1; 3).
* Змоделювати 20 значень випадкової нормально розподіленої величини з математичним очікуванням, рівним 1, і середньоквадратичним відхиленням, рівним 3. Ці 20 значень визначають змінну Var 1. Перевести їх до цілого типу даних, встановивши у вікні опису змінної в якості Type значення Integer.
* Перейти до змінної Var 2 наступним чином: перше значення Var 2 одно першому значенню змінної Var 1; друге значення Var 2 дорівнює сумі перших двох значень змінної Var 1; третє значення змінної Var 2 дорівнює сумі перших трьох значень змінної Var 1 і т. д.
* Скопіювати змінну Var 2 і перейти в пакет SNN, розмістивши скопійовані дані в новому створеному файлі.
* Провести прогнозування отриманого ряду за допомогою нейронної мережі.

Вихідний часовий ряд був представлений у вигляді таблиці excel. Дані були успішно імпортовані в STATISTICA.

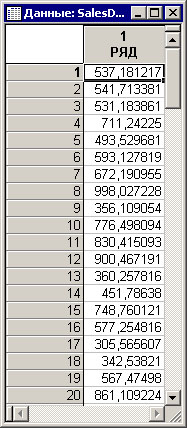


Рисунок 3.1 – Вихідні дані

Дані являють собою щотижневі продажу напівфабрикатів. Всього 116 спостережень. Таким чином ми вже маємо "спущену зверху" сезонну структуру. Експертно встановлено, що ряд схильний високочастотним коливанням невеликої амплітуди з періодом в 4 вимірювання. Такі коливання властиві будь-якому виробнику, збувати продукцію, зокрема оптом. окупателі планують закупівлі з періодичністю в місяць.

Необхідно побудувати прогноз продажів досліджуваного товару - короткостроковий і середньостроковий. Для вирішення цього завдання буде використаний пакет STATISTICA, що володіє, зокрема, можливостями для побудови прогнозів часових рядів.

Побудова прогнозу за допомогою нейронних мереж

Інструментом для побудови прогнозу подібного ряду були обрані нейронні мережі. Інші методи виявилися неефективними з кількох причин:

* Ряд досить короткий для використання сезонних методів.
* Даний ряд був отриманий в результаті згладжування вихідного, в результаті чого з нього були вилучені практично всі значущі автокорреляции.

Крім того, використання нейронних мереж позбавить нас від необхідності дотримуватися якоїсь певної сезонної або трендового моделі.

Крок 1. Виберіть Аналіз - Нейронні мережі. З'явиться діалогове вікно настройки роботи з нейронними мережами. Необхідно встановити тип завдання на Тимчасові ряди.

Крок 2. Натисніть на кнопку Змінні і задайте змінну, що відповідає за згладжений ряд як вихідну і вхідну безперервну. Натисніть на ОК для повернення на основне вікно. Встановіть опцію Інструмент на Майстер рішень. Потім натисніть ОК для переходу на діалогове вікно налаштувань Майстра.

Крок 3. На вкладці Тимчасові ряди встановіть параметри тимчасового вікна від 1 до 30. Цей параметр використовується як підстава для визначення періоду сезонної компоненти.

Крок 4. На вкладці Тип мережі встановіть все галочки, щоб дозволити системі тестувати всі можливі типи нейронних мереж.

Крок 5. На вкладці Швидкий виберете навчання 50 нейронних мереж і збереження 50 кращих. Правильно налаштована панель буде виглядати наступним чином:

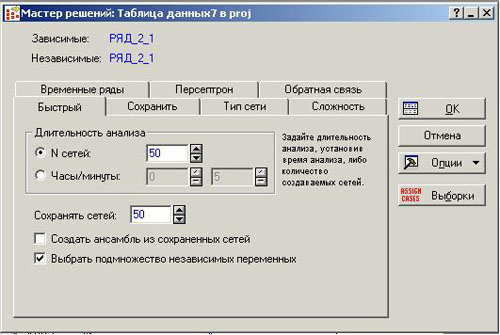


Рисунок 3.2 – Налаштування нейронної мережі

Крок 6. Натисніть на кнопку ОК для початку навчання нейронних мереж. Через деякий час буде відображена панель аналізу результатів.

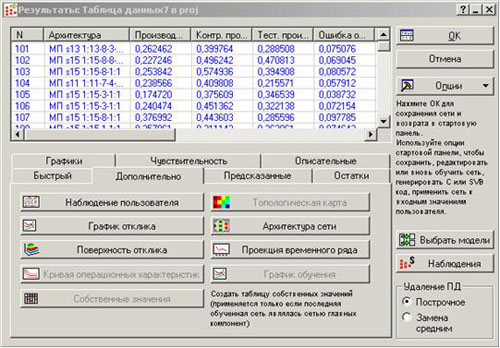


Рисунок 3.3 – Панель результатів навчання нейронних мереж

З навчених 50 нейронних мереж необхідно вибрати кілька з найменшою тестової продуктивністю. Це можна зробити за допомогою кнопки Підсумки моделей на вкладці Швидкий. Всі мережі необхідно впорядкувати за параметром Тестова продуктивність і вибрати кілька верхніх мереж.

Безпосередній прогноз виходить прогоном спостережень через нейронну мережу - будується проекція часового ряду. Це може бути зроблено на вкладці Додатково. Натисніть на кнопку Проекція часового ряду, буде відображено наступне діалогове вікно:

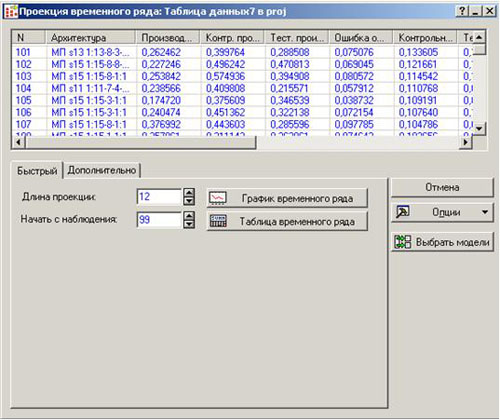


Рисунок 3.4 – Налаштування проекції тимчасового ряду

Крок 7. Задайте проекцію на 12 спостережень (3 місяці) і натисніть на кнопку Графік. З'явиться графік подовження нашого тимчасового ряду.

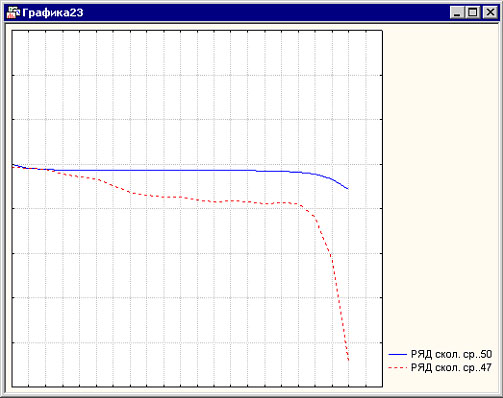


Рисунок 3.5 – Проекція тимчасового ряду

При наложении на исходный график он принимает вид:

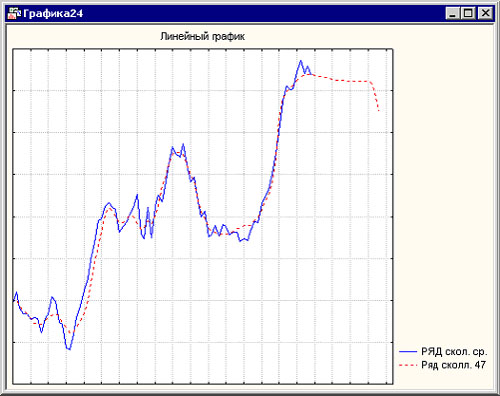


Рисунок 3.6 – Графік прогнозу трендової компоненти ряду

Таким чином, ми отримали прогноз частини вихідного згладженого ряду. Його можна відновити в початковому стані перебував середнє і лінійний тренд. Результат наведено нижче.

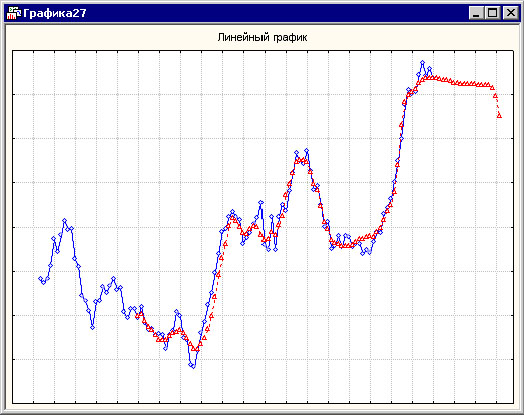


Рисунок 3.7 – Графік ряду і прогнозів згладженого ряду

Таким чином, протягом найближчих декількох місяців можна чекати спад рівня продажів. Червоним кольором заданий спостережуваний згладження ряд, синім - побудований прогноз. Зауважимо, що прогнозний ряд Коротше спостережуваного на деяке число спостережень. Це пов'язано з тим, що певна кількість спостережень було втрачено, як наслідок формування вікна для нейронної мережі.

Стабільна негативна тенденція у розвитку даного часового ряду є незвичною. Справді, прогнозне зниження становить близько половини від загального рівня продажів, а такого за всю історію не було ніколи. Отже, раціональним буде скоротити подовження ряду до 2 місяців, так зниження в цей період буде гарантованим.

# 3.2. Постановка задачі

Дана вибірка Нерухомість. Вона містить в собі 6 стовпців вхідних даних:

* Площа;
* Спальні кімнати;
* Ванні кімнати;
* Покупці;
* Додаткове утеплення;
* Район.

Стовбець вихідних даних - Ціна. Він визначається заданими вище факторами та напряму залежить від них.

Задано 130 спостережень.

Необхідно розробити модель лінійної регресії, досягнувши критерію мінімальних помилок та максимальної продуктивності, а також досягти встановлення правильної залежності ціни від інших величин.

# 3.3. Розв’язання задачі

Для розв’язання поставленої задачі запускаємо вибірку у програмному середовищі STATISTICA 12.0. На Рисунок3.8 зображено фрагмент навчальної вибірки.

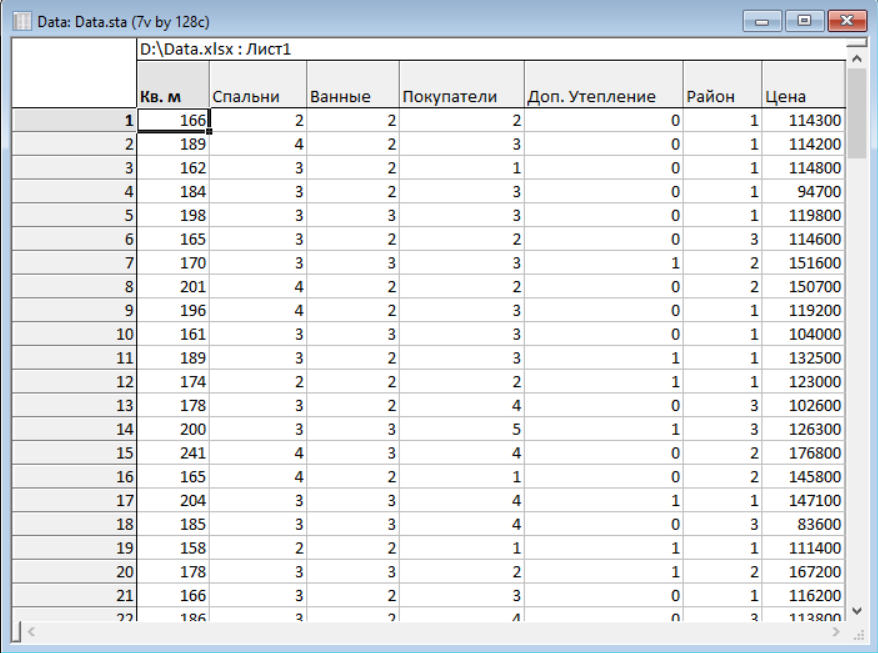


Рисунок 3.8 – Фрагмент навчальної вибірки

Запускаємо модуль Нейронні мережі STATISTICA. На стартовому вікні вибираємо тип завдання: *Regression (Регресія).* Натискаємо *Ок.*

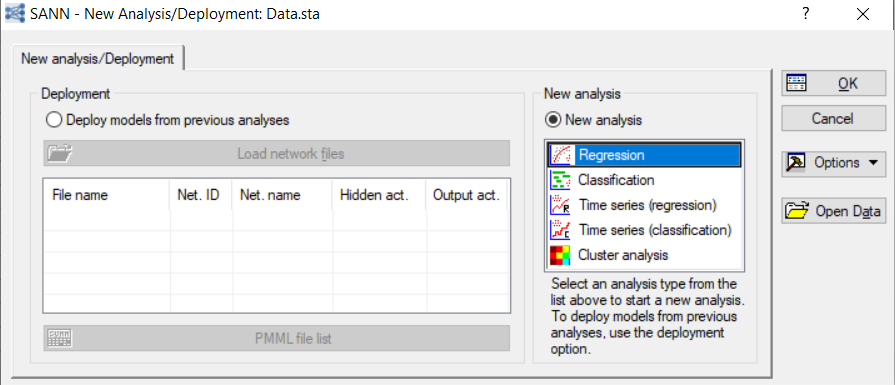


Рисунок 3.9 – Вікно вибору нейронних мереж

Далі, переходимо до задання змінних. Для цього у вікні, що відкрилося, натискаємо *Variables (Змінні)*.

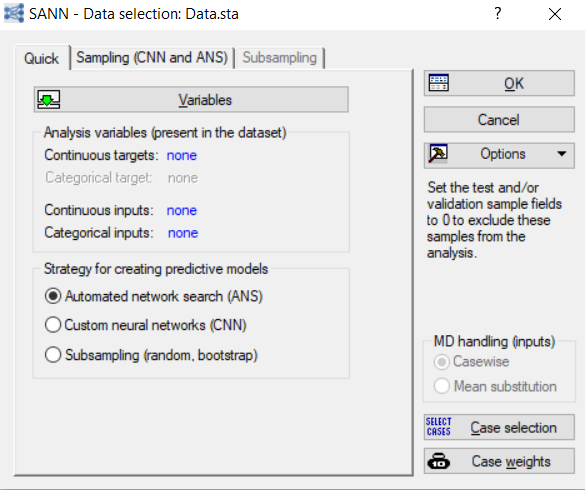


Рисунок 3.10 – Вікно налаштування нейронної мережі

У діалоговому вікні вибираємо змінні, в даному прикладі є вихідна (залежна) змінна і 6 вхідних змінних серед яких 4 — безперервні, а 2 — категоріальні.

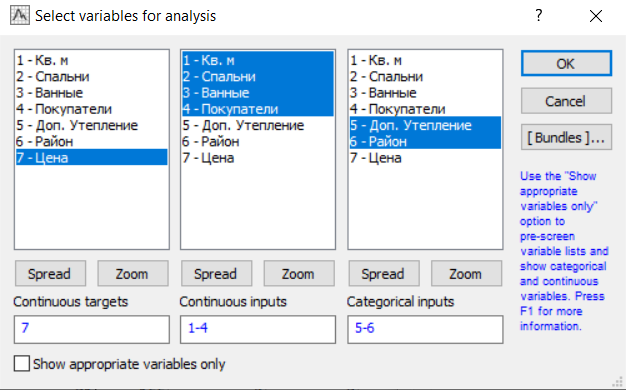


Рисунок 3.11 – Вікно задання змінних

Далі у наступній кладці обираємотривалість навчання мережі: *N мереж* – 100, *Зберегти мереж* – 7.

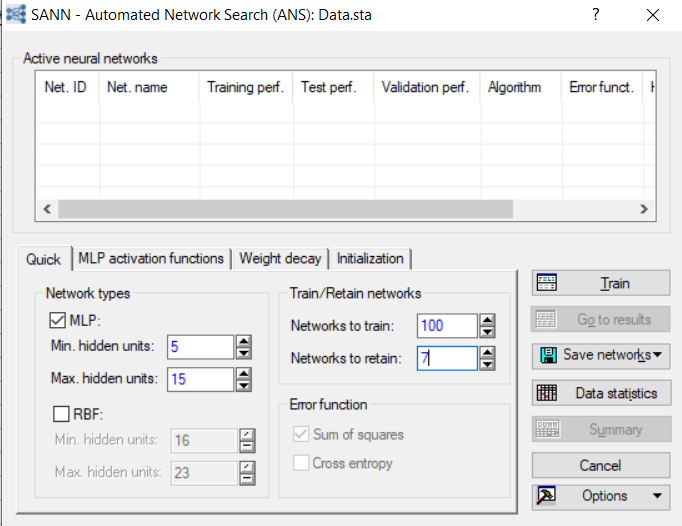


Рисунок 3.12 – Вікно налаштувань навчання

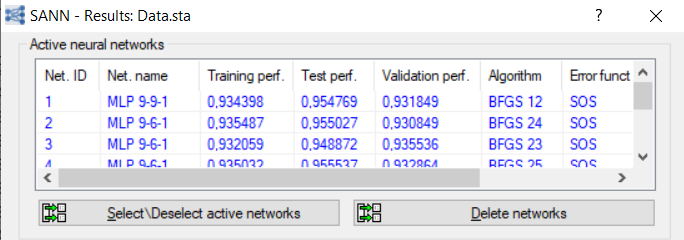
Далі натискаємо на *Train (Навчати)*.

Рисунок 3.13 – Результати навчання мережі

Отже, отримали 7 найкращих мереж. У вікні, що відкрилося натискаємо *Summary (Результати)* для відображення детальних результатів моделей.

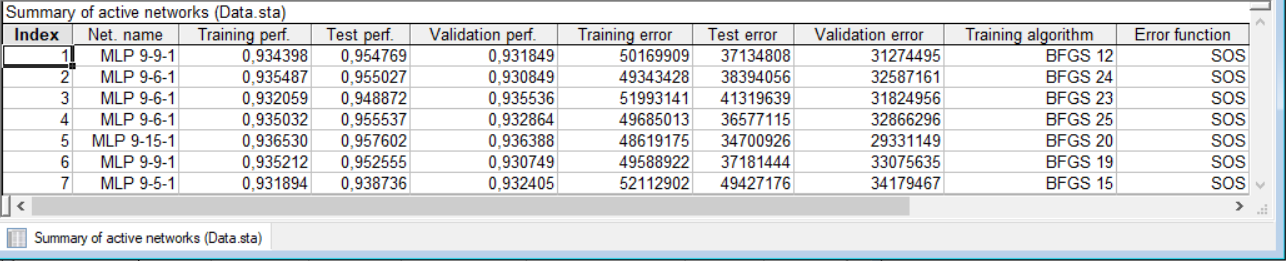


Рисунок 3.14 – Докладні результати моделей

Отже, мережа *MLP 9-15-1* дала найбільш оптимальний результат, а саме:

* Продуктивність – 0,936540

Виділяємо найбільш оптимальну модель

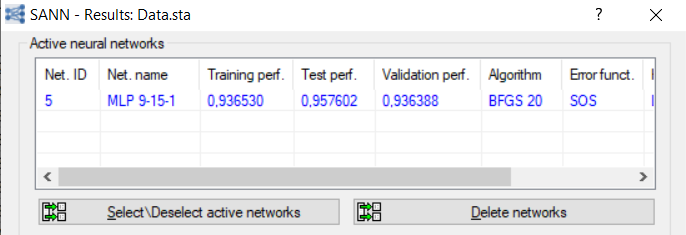
.

Рисунок 3.15 – Збереження моделі

Наступним кроком буде тестування навченої моделі на тестових даних. Отже, у вікні *Результати* вкладка *Custom predictions (Користувацькі передбачення)* натискаємо *Custom inputs (Користувацькі вхідні змінні)*.

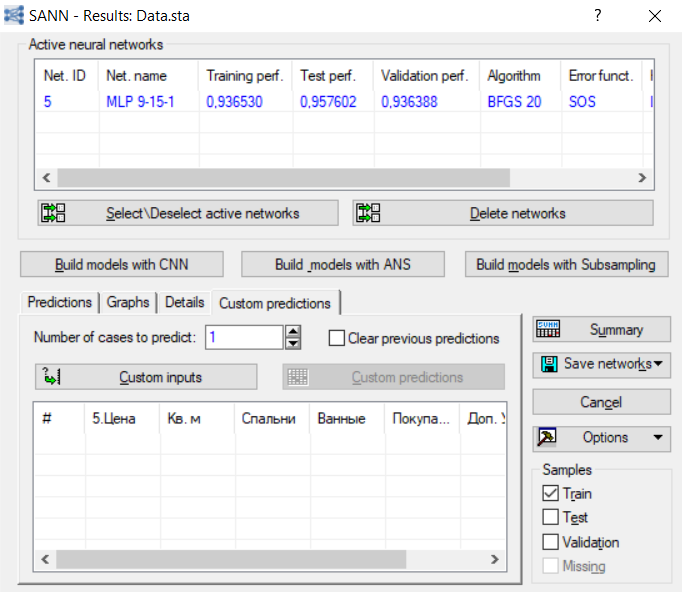


Рисунок 3.16 – Вікно *Результати*

Будемо тестувати одразу 5 прикладів. Вводимо вхідні дані до таблиці.

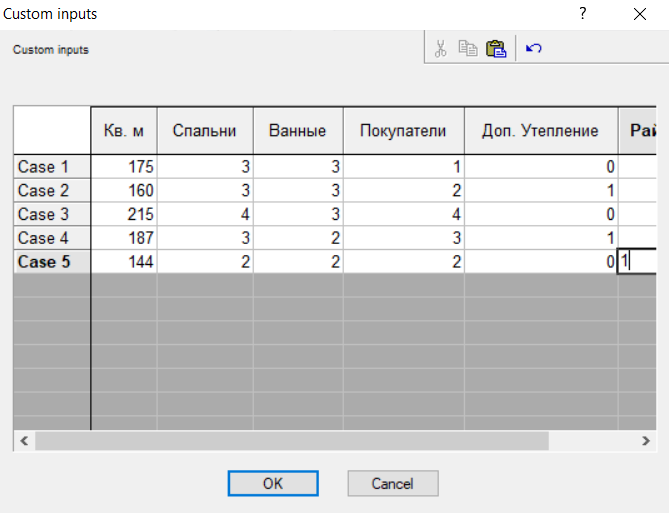


Рисунок 3.17 – Вікно *Користувацькі вхідні змінні*

Натиснувши *ОК*, повертаємося до вікна *Результати* та натискаємо *Custom predictions (Користувацькі передбачення)*. Отримали результат:

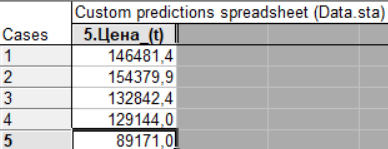


Рисунок 3.18 – Вікно *Користувацькі передбачення*

Отже, розглянувши декілька прикладів на нових даних, можна сказати, що система передбачає ціну загалом вірно з середньою помилкою 10000, а отже, модель дійсно гарно навчилася та може бути використана для автоматичного визначення ціни на нерухомість.

# ВИСНОВКИ

В процесі роботи була врахована система інтелектуального аналізу з практичним підтвердженням відповідно до заданої предметної області, що визначає ціну нерухомості за її знаком.

В дипломній роботі було проведено аналіз для навчальної вибірки та за допомогою нейронних мереж знайдено найбільш оптимальну модель:

* Продуктивність – 0,960526

Після збереження тренованої моделі перевіряли якість функції лінійної регресії. Ви можете зробити це дуже зручно за допомогою набору програм STATISTICA. Для цього за допомогою навчання нейронних мереж вибирається найкраща модель і на основі неї заповнюються дані користувачів. Значення, отримані мережею, виправдовують очікування. Коротше кажучи, модель добре навчена і чудово підходить для вирішення проблем.

Отже, в ході виконання роботи було створено систему інтелектуального аналізу для автоматичного визначення ціни на нерухомість за певними ознаками.

# СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Навчальні матеріали онлайн [Електронний ресурс] – Режим доступу: http://pidruchniki.com/10811007/informatika/ekspertni\_sistemi
2. Салютін Ш.М. Штучний інтелект. — М.: Думка, 1995. — 200 с.
3. Вінер Н. Кібернетика. — М.: Наука. — 1983. — 356 с.
4. Сетлак Г. Интеллектуальные системы поддержки принятия решений. — К.: Логос, 2004. — 179 с.
5. Моїсеєв В.Б. Подання знань в інтелектуальних системах. Інформатика і освіта,.№ 2, 2003 р. с. 84-91
6. Використання нейронних мереж в моделюванні фінансових результатів бізнес-процесів [Електронний ресурс]. – Режим доступу: http://sf.wunu.edu.ua/index.php/sf/article/view/1249
7. Нейромережева методологія розпізнавання інтернет-орієнтованого шкідливого програмного забезпечення [Електронний ресурс]. – Режим доступу: http://jrnl.nau.edu.ua/index.php/Infosecurity/article/view/4688
8. Нейронні мережі: основні відомості [Електронний ресурс]. – Режим доступу: http://www.biometrica.tomsk.ru/statbook/modules/stneunet.html
9. Штучна нейронна мережа [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/Штучна\_нейронна\_мережа
10. Нейронні мережі з радіальними базисними функціями: лекція

[Електронний ресурс]. – Режим доступу:

<http://knowledge.allbest.ru/programming/3c0a65635b2bd68a4c43a89421306d37_0.html>

1. Методи класифікації і прогнозування. Нейронні мережі [Електронний ресурс]. – Режим доступу:

<http://www.intuit.ru/studies/professional_skill_improvements/1210/courses/6/lecture/178?page=5>

1. Data mining [Електронний ресурс]. – Режим доступу:

https://ru.wikipedia.org/wiki/Data\_mining

1. Нейронні мережі: основні відомості [Електронний ресурс]. – Режим доступу: <http://www.statsoft.ru/home/textbook/modules/stneunet.html#linear>
2. Імовірнісна нейронна мережа: основні відомості [Електронний ресурс]. – Режим доступу: <https://studopedia.ru/7_130500_veroyatnostnaya-neyronnaya-set.html>
3. Нейронні мережі. STATISTICA Neural Networks: Методологія і технології сучасного аналізу даних / за редакцією В. П. Боровикова. - 2-е вид. , перероб. і дод. - М.: Гаряча лінія-Телеком, 2008. - 392 с., мул.

# Додаток А

**Программный код проекта**

Dim newanalysis As Analysis

Set newanalysis = Analysis (scSANN, ActiveInputDataSet)

Dim oStaDocs As StaDocuments

Dim oAD1 As STASANN.SANNAnalysisType

Set oAD1 = newanalysis.Dialog

oAD1.NewAnalysis = True

oAD1.TypeOfAnalysis = scSANNRegression

newanalysis.Run

Dim oAD2 As STASANN.SANNDataSelection

Set oAD2 = newanalysis.Dialog

oAD2.Variables = "1 | 2-9"

oAD2.CasewiseDeletionOfMD = True

oAD2.UseRandomSampling = True

oAD2.SizeOfTrainingSample = 70

oAD2.SizeOfTestSample = 15

oAD2.SizeOfValidationSample = 15

oAD2.SamplingSeed = 1000

oAD2.UseRandomSubsampling = True

oAD2.SizeOfTrainingSubsample = 70

oAD2.SizeOfTestSubsample = 15

oAD2.SizeOfValidationSubsample = 15

oAD2.NumberOfSubsamples = 2

oAD2.SubsamplingSeed = 1000

oAD2.CustomNetworkBuilding = True

oAD2.TrainSample = "off"

oAD2.TestSample = "off"

oAD2.ValidationSample = "off"

newanalysis.Run

Dim oAD3 As STASANN.SANNCustomNeuralNetwork

Set oAD3 = newanalysis.Dialog

oAD3.UseMultilayerPerceptron = True

oAD3.NumberOfNetworksToTrainCustom = 1

oAD3.HiddenLayerActivationFunction = "Гиперболическая"

oAD3.OutputActivationFunction = "Тождественная"

oAD3.NumberOfNeuronsForTheHiddenLayer = 7

oAD3.UseHidWeightDecay = False

oAD3.UseOutWeightDecay = False

oAD3.UseSumOfSquaresErrorFunction = True

oAD3.TrainingAlgorithm = 1

oAD3.NumberOfTrainingCycles = 200

oAD3.GaussRandomizeTheWeights = True

oAD3.MeanForGaussRandomization = 0

oAD3.VarianceForGaussRandomization = 0.1

oAD3.ApplyStoppingConditions = True

oAD3.ErrorImprovement = 1e-007

oAD3.NumberOfCyclesForErrorImprovement = 20

oAD3.UseFixedSeedForNetworkInitialization = False

oAD3.DisplayRealTimeTrainingGraph = False

newanalysis.Run

Dim oAD4 As STASANN.SANNResults

Set oAD4 = newanalysis.Dialog

oAD4.ActivateModels = "1 "

oAD4.ModelListView = "1 "

oAD4.UseTrainingSample = True

oAD4.UseTestSample = False

oAD4.UseValidationSample = False

oAD4.StandaloneNetworks = True

oAD4.IncludeInputs = False

oAD4.IncludeTargets = True

oAD4.IncludePredictions = True

oAD4.IncludeResiduals = False

oAD4.IncludeStandardResiduals = False

oAD4.IncludeAbsoluteResiduals = False

oAD4.IncludeSquareResiduals = False

oAD4.IncludeOtherVariables = False

oAD4.SelectedItemX = "1"

oAD4.SelectedItemY = "2"

oAD4.SelectedItemZ = "none"

oAD4.IncludeCaseNames = False

oAD4.NumberOfCustomPredictions = 1

oAD4.ClearCustomPredictions = False

Dim newanalysis As Analysis

Set newanalysis = Analysis (scSANN, ActiveInputDataSet)

Dim oStaDocs As StaDocuments

Dim oAD1 As STASANN.SANNAnalysisType

Set oAD1 = newanalysis.Dialog

oAD1.NewAnalysis = True

oAD1.TypeOfAnalysis = scSANNTimeseriesClassification

newanalysis.Run

Dim oAD2 As STASANN.SANNDataSelection

Set oAD2 = newanalysis.Dialog

oAD2.Variables = "11 | 1-10"

oAD2.UseRandomSampling = True

oAD2.SizeOfTrainingSample = 70

oAD2.SizeOfTestSample = 15

oAD2.SizeOfValidationSample = 15

oAD2.SamplingSeed = 1000

oAD2.UseRandomSubsampling = True

oAD2.SizeOfTrainingSubsample = 70

oAD2.SizeOfTestSubsample = 15

oAD2.SizeOfValidationSubsample = 15

oAD2.NumberOfSubsamples = 2

oAD2.SubsamplingSeed = 1000

oAD2.AutomaticNetworkSearch = True

oAD2.TrainSample = "off"

oAD2.TestSample = "off"

oAD2.ValidationSample = "off"

oAD2.StepsUsedToPredict = 1

oAD2.StepsAhead = 1

newanalysis.Run

Dim oAD3 As STASANN.SANNAutomaticNetworkSearch

Set oAD3 = newanalysis.Dialog

oAD3.UseMLPNetwork = True

oAD3.UseRBFNetwork = False

oAD3.UseSOSError = True

oAD3.UseEntropyError = True

oAD3.IdentityHiddenActivationFunction = True

oAD3.SigmoidHiddenActivationFunction = True

oAD3.TanhHiddenActivationFunction = True

oAD3.ExpHiddenActivationFunction = True

oAD3.SineHiddenActivationFunction = False

oAD3.IdentityOutputActivationFunction = True

oAD3.SigmoidOutputActivationFunction = True

oAD3.TanhOutputActivationFunction = True

oAD3.ExpOutputActivationFunction = True

oAD3.SineOutputActivationFunction = False

oAD3.MinimumNumberOfHiddenUnits = 4

oAD3.MaximumNumberOfHiddenUnits = 13

oAD3.UseHiddenWeightDecay = False

oAD3.UseOutputWeightDecay = False

oAD3.NumberOfNetworksToTrain = 20

oAD3.NumberOfNetworksToRetain = 5

oAD3.UseFixedSeedForNetworkInitialization = False

newanalysis.Run

Dim oAD4 As STASANN.SANNResults

Set oAD4 = newanalysis.Dialog

oAD4.ActivateModels = "1-5 "

oAD4.ModelListView = "1 2 3 4 5 "

oAD4.UseTrainingSample = True

oAD4.UseTestSample = False

oAD4.UseValidationSample = False

oAD4.StandaloneNetworks = True

oAD4.IncludeInputs = False

oAD4.IncludeTargets = True

oAD4.IncludePredictions = True

oAD4.IncludeErrorbars = False

oAD4.IncludeResiduals = False

oAD4.SelectedItemX = "1"

oAD4.SelectedItemY = "2"

oAD4.SelectedItemZ = "none"

oAD4.IncludeCaseNames = False

oAD4.GainLiftchart = True

oAD4.SelectCategory = "К"

oAD4.CumulativeLiftchart = False

oAD4.NumberOfCustomPredictions = 1

oAD4.ClearCustomPredictions = False

Sub Main

Dim newanalysis As Analysis

Set newanalysis = Analysis (scClusterAnalysis, ActiveInputDataSet)

Dim oStaDocs As StaDocuments

Dim oAD1 As STACluster.CluStartup

Set oAD1 = newanalysis.Dialog

oAD1.ClusterMethod = scCluKMeans

newanalysis.Run

Dim oAD2 As STACluster.CluKMeansSpecifications

Set oAD2 = newanalysis.Dialog

oAD2.CasesOrVariables = scCluVariables

oAD2.Variables = "1-10"

oAD2.NumberOfClusters = 2

oAD2.NumberOfIterations = 10

oAD2.CasewiseDeletionOfMD = True

oAD2.ConstantIntervals = True

oAD2.BatchProcessingAndPrinting = False

newanalysis.Run

Dim oAD3 As STACluster.CluKMeansResults

Set oAD3 = newanalysis.Dialog

Set oStaDocs = oAD3.ClusterMeansEuclideanDistances

newanalysis.RouteOutput(oStaDocs).Visible = True

Set oStaDocs = Nothing

Set oStaDocs = oAD3.AnalysisOfVariance

newanalysis.RouteOutput(oStaDocs).Visible = True

Set oStaDocs = Nothing

Set oStaDocs = oAD3.MembersOfEachClusterDistances

newanalysis.RouteOutput(oStaDocs).Visible = True

Set oStaDocs = Nothing

newanalysis.Dialog.ClassificationsAndDistances.Visible = True

End Sub