

ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА»
Факультет автоматизації виробництва та цифрових технологій
Кафедра автоматизації, електро- та робототехнічних систем

«Допущено до захисту»

Гарант ОПП

Павло САГАЙДА

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня магістра

за підсумками виконання
освітньо-професійної програми
«Комп'ютерні науки та цифровий інтелект»
за спеціальністю 122 Комп'ютерні науки

на тему «ДОСЛІДЖЕННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ
АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЗАКУПІВЕЛЬНИХ ДАНИХ»

Керівник роботи

Валентина МОСКАЛЕНКО

Консультант від бази
практики

Валерій ДРОБОТ

*Кваліфікаційна робота містить результати власних досліджень.
Використання ідей, результатів і текстів інших авторів мають
посилання на відповідне джерело*

Здобувач

Сергій КРІПАК

Підсумкова оцінка за атестацію			
--------------------------------	--	--	--

Голова ЕК

Олена ПАВЛЕНКО

Кривий Ріг 2024

ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА»

Факультет	автоматизації виробництва та цифрових технологій
Кафедра	цифрових технологій та проектно-аналітичних рішень
Ступінь вищої освіти	магістр
Спеціальність	122 Комп'ютерні науки
ОПП	Комп'ютерні науки та цифровий інтелект

ЗАТВЕРДЖУЮ

Гарант ОПП

_____ Павло САГАЙДА
«06» листопада 2023 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРА

Кріпака Сергія Андрійовича

(прізвище, ім'я, по батькові здобувача)

1. Тема роботи Дослідження методів машинного навчання для аналізу та прогнозування закупівельних даних

керівник роботи Москаленко Валентина, професор, докт. техн. наук,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом Університету від 29.08. 2023 р. №137.1/29.08.2023

2. Термін подання роботи 10.01.2024 р.

3. Вихідні дані до роботи Навчальна література, державні стандарти, методична література з спеціальних дисциплін та дипломування, науково-дослідницькі роботи з тематики автоматизації обробки й аналізу даних та методів цифрового інтелекту, літературні джерела, результати власних експериментів та досліджень, технологічні інструкції тощо

4. Зміст пояснювальної записки (перелік питань) Реферат. Зміст. Вступ. 1. Дослідження методів машинного навчання для аналізу та прогнозування закупівельних даних. 2. Дослідження використання моделей машинного навчання для аналізу та прогнозування закупівельних даних у середовищі RapidMiner та MATLAB. 3. Дослідження моделей класифікації та прогнозування. 4. Розробка технічного завдання на створення засобів моделювання предметної області. 5. Економічне обґрунтування запропонованих технічних рішень. Висновки. Перелік використаних джерел. Додатки.

5. Перелік графічного (демонстраційного) матеріалу (з точним зазначенням обов'язкових креслень): Актуальність, мета, об'єкт, предмет та

завдання дослідження; розроблені або удосконалені процеси, методика дослідження; діаграми проекту програмно-методичного комплексу в нотації UML (діаграми прецедентів, класів, послідовностей, діяльності); результати розробки та експериментальних досліджень; результати економічних розрахунків; висновки до роботи; публікація результатів дослідження.

6. Консультанти по роботі, із зазначенням розділів роботи, що стосуються їх.

Розділ	Прізвище, ініціали та посада консультанта
1	Москаленко В.В., проф. каф. ЦТПАР
2	Москаленко В.В., проф. каф. ЦТПАР
3	Москаленко В.В., проф. каф. ЦТПАР
4	Москаленко В.В., проф. каф. ЦТПАР
5	Гетьман І.А., доцент каф. ЦТПАР

7. Дата видачі завдання 06.11.2023

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи
1	Розділ 1. Дослідження методів машинного навчання для аналізу та прогнозування закупівельних даних	25.12.2023 - 30.12.2023
2	Розділ 2. Дослідження використання моделей машинного навчання для аналізу та прогнозування закупівельних даних у середовищі RapidMiner та MATLAB	25.12.2023 - 30.12.2023
3	Розділ 3. Дослідження моделей класифікації та прогнозування	25.12.2023 – 02.01.2024
4	Розділ 4. Розробка технічного завдання на створення засобів моделювання предметної області	03.01.2024 - 07.01.2024
5	Розділ 5. Економічне обґрунтування запропонованих технічних рішень.	03.01.2024 - 07.01.2024
6	Висновки, перелік посилань, вступ, зміст, реферат	07.01.2024 – 08.01.2024
7	Подання завершеної роботи. Перевірка на академічний плагіат	10.01.2024 – 16.01.2024
8	Остаточне оформлення роботи, презентаційного матеріалу, автореферату	17.01.2024 – 19.01.2024
9	Рецензування завершеної роботи. Захист	19.01.2024 – 24.01.2024

Здобувач
Керівник роботи

Сергій Кріпак
Валентина Москаленко

РЕФЕРАТ

Дипломна робота: робота складається зі вступу, 5 розділів, висновків, списку використаних джерел, 7 додатків. Загальний обсяг роботи становить 118 сторінок, робота містить 31 рисуноків, 44 таблиць. Список використаних джерел складається з 35 джерел.

Мета дослідження. Вдосконалення процесів управління закупівлями та постачанням через вдосконалення стратегій закупівель на основі використання ПЗ, яке реалізує алгоритми машинного навчання.

Об'єкт дослідження. Процес підготовки та обробки даних для прийняття ефективних управлінських рішень в сфері закупівель.

Предмет дослідження. Проектування ПЗ для прискорення обробки даних обраним методом машинного навчання для підготовки та обробки даних для прийняття ефективних управлінських рішень в сфері закупівель.

Задачі дослідження: Задачами дослідження є формування:

- рекомендацій щодо відбору та підготовки закупівельних даних для подальшої їх обробки;
- оцінки переваг та недоліків різних методів машинного навчання та напрямки їх використання для аналізу закупівельних даних;
- критеріїв вибору найбільш відповідного методу машинного навчання для розв'язання задачі;
- рекомендацій щодо виявлення ключових закономірностей для здійснення точних прогнозів щодо закупівельних операцій.

В умовах стрімко змінюючогося світу та стану світової економіки аналіз закупівельних даних та прогнозування поведінки ринку стають ключовими для своєчасної реакції на зміни, умови та прийняття ефективних рішень.

Ця робота присвячена огляду методів машинного навчання та їх застосуванню для аналізу та прогнозування закупівельних даних. Практичне значення дослідження полягає в підтримці розуміння потенційних переваг, які надають методи машинного навчання для оптимізації процесів управління закупівлями. Дослідження акцентується на вивченні аспектів використання машинного навчання, таких як прогнозування обсягів, а також планування та оптимізація закупівель. Увага спрямована на розгляд переваг та недоліків методів машинного навчання, з більш детальним дослідженням методів прогнозування та виявлення аномалій та ризиків у закупівельних процесах. Також розглядається практичне застосування методів машинного навчання та архітектури нейронних мереж для кожного розглянутого методу. В результаті дослідження сформульовано технічне завдання на розробку програмного продукту "Система управління закупівлями".

Ключові слова: машинне навчання, кластерний аналіз, класифікація, регресійний аналіз, нейронна мережа, аналіз даних, прогнозування, моделювання, система управління закупівлями, бізнес-вимоги, специфікація вимог.

SUMMARY

The thesis comprises an introduction, 5 chapters, conclusions, a list of references, and 7 appendices. The total volume of the work is 170 pages, including 31 figures and 44 tables. The list of references encompasses 35 sources.

Research Objective: The aim of this study is to enhance procurement and supply management processes by improving procurement strategies through the use of software implementing machine learning algorithms.

Research Focus: The object of the research is the data preparation and processing processes to make effective managerial decisions in the procurement field.

Research Subject: The subject of the research is the design of software to expedite data processing using a chosen machine learning method for preparing and processing data to make effective managerial decisions in the procurement domain.

Research Tasks. The research tasks include:

Formulating recommendations for selecting and preparing procurement data for further processing.

Evaluating the advantages and disadvantages of various machine learning methods and their application directions for analyzing procurement data.

Defining criteria for selecting the most suitable machine learning method to address the research problem.

Providing recommendations for identifying key patterns to make accurate forecasts regarding procurement operations.

In a rapidly changing world and the state of the global economy, analyzing procurement data and forecasting market behavior are crucial for timely responses to changes, conditions, and making effective decisions.

This work is dedicated to reviewing machine learning methods and their application for analyzing and forecasting procurement data. The practical significance of the research lies in supporting an understanding of the potential advantages offered by machine learning methods to optimize procurement management processes. The research focuses on aspects of machine learning application such as volume forecasting, procurement planning, and optimization. Emphasis is placed on examining the pros and cons of machine learning methods, with a more detailed exploration of forecasting methods and the detection of anomalies and risks in procurement processes. The practical application of machine learning methods and neural network architecture for each discussed method is also considered. As a result of the research, a technical task for the development of the product "Procurement Management System" has been formulated.

Keywords: machine learning, cluster analysis, classification, regression analysis, neural network, data analysis, forecasting, modeling, procurement management system, business requirements, requirements specification.

ЗМІСТ

ВСТУП	8
РОЗДІЛ 1. ДОСЛІДЖЕННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЗАКУПІВЕЛЬНИХ ДАНИХ	11
1.1 Аналіз використання моделей та алгоритмів машинного навчання у сфері постачання та закупівель	11
1.2 Аналіз сучасних інформаційних технологій для аналізу та прогнозування закупівельних даних.....	19
1.2.1 Аналіз сучасних програмних інструментів аналізу та обробки даних	19
1.2.2 Дослідження сучасних методів машинного навчання для аналізу закупівельних даних	25
1.3 Постановка задачі дослідження	39
1.4 Висновки за розділом.....	41
РОЗДІЛ 2. ДОСЛІДЖЕННЯ ВИКОРИСТАННЯ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЗАКУПІВЕЛЬНИХ ДАНИХ У СЕРЕДОВИЩІ RAPIDMINER ТА MATLAB.....	42
2.1 Відбір моделей для проведення дослідження.....	42
2.1.1 Аналіз моделей класифікації.....	42
2.1.2 Аналіз моделей прогнозування.....	47
2.2 Огляд ключових елементів аналізу, підготовки та обробки даних у сфері закупівель.....	50
2.2.1 Підготовка та очищення даних є ключовим етапом у процесі аналізу.....	51
2.2.2 Визначення відповідного методу машинного навчання.....	52
2.2.3 Перевірка та оцінка достовірності моделей - важливий етап у процесі їх використання.....	54
2.2.4 Інтерпретування результатів	55
2.3 Аналіз та підготовка закупівельних даних для розв'язання управлінських задач.....	57
2.4 Висновки за розділом.....	65
РОЗДІЛ 3. ДОСЛІДЖЕННЯ МОДЕЛЕЙ КЛАСИФІКАЦІЇ ТА ПРОГНОЗУВАННЯ.....	66
3.1 Дослідження моделей класифікації.....	66
3.1.1 Дослідження моделей класифікації у середовищі RapidMiner .	66
3.1.2 Дослідження моделей класифікації у середовищі MATLAB	69
3.2 Дослідження моделей прогнозування.....	74
3.2.1 Дослідження моделей прогнозування у середовищі RapidMiner	74
3.2.2 Дослідження моделей прогнозування у середовищі MATLAB .	84
3.3 Аналіз результатів дослідження моделей класифікації	91
3.4 Висновки за розділом.....	92

РОЗДІЛ 4. РОЗРОБКА ТЕХНІЧНОГО ЗАВДАННЯ НА СТВОРЕННЯ СИСТЕМИ УПРАВЛІННЯ ЗАКУПІВЛЯМИ	94
4.1 Розробка бізнес-процесу управління закупівлями	94
4.2 Бізнес вимоги до системи управління закупівлями	95
4.3 Специфікація вимог до системи управління закупівлями	100
4.3.1 Функції системи управління закупівлями	100
4.3.2 Прогнозування потреб Матеріалів та Послуг у розрізі Замовників (FR01).....	102
4.3.3 Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів (FR02)	103
4.3.4 Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників (FR03).....	105
4.3.5 Створення заявок на закупівлю (FR04).....	106
4.3.6 Пропозиції щодо переліку Постачальників минулих періодів ..	107
4.3.7 Прогноз по виконанню поставок Постачальниками в строк....	108
4.4 Висновки	109
РОЗДІЛ 5. ЕКОНОМІЧНЕ ОБҐРУНТУВАННЯ ЗАПРОПОНОВАНИХ ПРОГРАМНИХ РІШЕНЬ.....	110
5.1 Визначення рангу значущості роботи	110
5.2 Визначення економічного ефекту від впровадження розробки	112
5.2.1 Оцінка вартості розробки програмного продукту	113
5.2.2 Оцінка очікуваного економічного ефекту	117
5.3 Висновки	119
ЗАГАЛЬНІ ВИСНОВКИ.....	120
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	122
ДОДАТОК А. ВІДОМІСТЬ РОБОТИ.....	126
ДОДАТОК Б. ПРИКЛАД КОДУ МОВИ MATLAB ДЛЯ МОДЕЛІ ENSEMBLE BAGGED TREE.....	127
ДОДАТОК В. ПРИКЛАД КОДУ МОВИ MATLAB ДЛЯ МОДЕЛІ DECISION TREE.....	130
ДОДАТОК Г. ПРИКЛАД КОДУ МОВИ MATLAB ДЛЯ МОДЕЛІ LINEAR REGRESSION.....	133
ДОДАТОК Д. ПРИКЛАД КОДУ МОВИ MATLAB ДЛЯ МОДЕЛІ NEURAL NETWORK.....	136
ДОДАТОК Е. BUSINESS REQUIREMENTS SPECIFICATION.....	139
ДОДАТОК Є. SOFTWARE REQUIREMENTS SPECIFICATION	146

ВСТУП

В сучасному конкурентному середовищі важливо швидко отримувати та аналізувати релевантну інформацію для своєчасного прийняття бізнесових рішень. Це особливо стосується закупівельних процесів, які включають великі об'єми даних, як наприклад назви товарів і послуг, дат закупівель і потреби, ціна, постачальники тощо.

Дані про закупівлі постійно збираються і можуть стати величезними, що уповільнює їх опрацювання. Аналітики отримують ці дані з різноманітних джерел, як наприклад облікові звіти організації, дані у електронних таблицях (наприклад Excel, LibreOffice Calc, Google Таблиці) тощо. Взаємодія з розосередженою інформацією вимагає потужних зусиль і часу від аналітика, що впливає на якість прийнятих управлінських рішень.

Актуальність вирішення цього питання проявляється в тому, що якість стратегічних управлінських рішень, які приймає ТОП-менеджмент компанії, залежить від якості аналітичних даних, які залежить від використовуваних ІТ-технологій. Ростуть вимоги до ІТ-обладнання та ПЗ, що призводить до суттєвого збільшення об'ємів даних закупівель.

У сучасних умовах, більшість компаній відмовляються від власних ІТ-підрозділів на користь спеціалізованих ІТ-компаній, які продають послуги з інтеграції та обслуговування інформаційних систем. Закупівельні процеси для ІТ-підприємств стають ключовим елементом забезпечення потрібними ресурсами та виконання бізнес-цілей. Збільшення об'ємів закупівельних даних вимагає особливих підходів до їх підготовки та аналізу.

Більша частина часу витрачається на підготовку даних через їх можливі помилки або неповноту. Висновки та рекомендації аналізу даних суттєво залежать від навичок аналітика та його вправності з системами обліку.

Отже, актуальним завданням є прискорення обробки даних, а ключовими інструментами для цього є data mining та машинне навчання. Застосування

методів машинного навчання спрямоване на підвищення ефективності бізнесових рішень у сфері закупівель. Розробка програмних систем з використанням моделей машинного навчання може значно поліпшити процеси управління закупівлями та прийняття рішень.

Метою дослідження є вдосконалення процесів управління закупівлями та постачанням в ІТ-компанії на основі використання ПЗ, яке реалізує алгоритми машинного навчання. Задачами дослідження є формування:

- рекомендацій по відбору та підготовці закупівельних даних з метою подальшої їх обробки;
- оцінка переваг використання моделей машинного навчання, їх недоліки та їх способи використання для обробки закупівельних даних;
- критеріїв вибору найбільш відповідного моделі машинного навчання для вирішення задач;
- рекомендацій з виявлення головних закономірностей для здійснення оціночних закупівельних прогнозів;
- рекомендацій з виявлення ризиків для своєчасного їх попередження.

Об'єктом дослідження є діяльність з підготовки, обробки даних для прийняття ефективних бізнесових рішень в сфері закупівель.

Предметом дослідження є проектування ПЗ для прискорення опрацювання даних завдяки обраної моделі машинного навчання для обробки підготовлених даних та прийняття ефективних бізнесових рішень у сфері закупівель.

Отже, отримання достовірних з закупівельних даних та її якісна обробка для бізнесових рішень є надважливим завданням. Застосування цих моделей дозволяє ефективно розв'язувати ці завдання, підвищуючи якість процесів управління закупівлями.

Для якісного опрацювання та аналізу даних потрібно використовувати моделі машинного навчання (Machine Learning – ML) [1]. Машинне навчання – це піднапрямок штучного інтелекту, який надає можливість:

- добувати нові знання з галузі знань (в цій роботі – закупівельні процеси) застосуванням алгоритмів знаходження взаємозв'язків та правил у закупівельних даних;

- вирішувати завдання керування закупівельними процесами із застосуванням побудованих моделей навчання на накопичених даних та інше.

Тому, постає завдання вибору моделей машинного навчання для підготовки та аналізу закупівельних даних з метою пошуку головних правил, зв'язків та для здійснення достовірних прогнозів щодо закупівельних операцій з метою своєчасного та достовірного формування планів закупівельної діяльності.

Практична цінність отриманих результатів відображається у розроблених моделях і реалізованих алгоритмах які рекомендується застосувати :

- при підтримці прийняття рішень менеджерами служб закупівель стосовно класифікації типів закупівель та

- у випадках прогнозування об'ємів заявок і навантаження на підрозділи.

За матеріалами дослідження опубліковано тези та стаття:

- Кріпак С.А., Москаленко В.В. Постановка задачі дослідження методів машинного навчання для аналізу та прогнозування закупівельних даних ІТ-підприємства / С.А. Кріпак, В.В. Москаленко // MININGMETALTECH 2023–Гірничо-металургійний комплекс: інтеграція бізнесу, технологій та освіти. Матеріали Міжнародної науково-технічної конференції. Частина І. Одеса, 29-30 листопада 2023 р. - Запоріжжя, Видавництво Baltija Publishing, 2023. - С. 241-242.

- Кріпак С.А., Москаленко В.В. Дослідження методів машинного навчання для аналізу та прогнозування закупівельних даних / С.А. Кріпак, В.В. Москаленко // Таврійський науковий вісник. Технічні науки. Випуск 4. Херсонський державний аграрно-економічний університет. 2023 р. - Херсон, Видавничий дім «Гельветика», 2023. - С. 61 - 68.

РОЗДІЛ 1. ДОСЛІДЖЕННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЗАКУПІВЕЛЬНИХ ДАНИХ

1.1 Аналіз використання моделей та алгоритмів машинного навчання у сфері постачання та закупівель

На сьогодні, області використання методів та алгоритмів машинного навчання в області закупівель присвячено праці іноземних та вітчизняних науковців, таких як В. Терещенко, А. Бугайов [1], А. Жиркова, О. Ігнатенко [2], Р. Резниченко, Л. Тимашова [3], К. Кононова [4], А. Мальцев [5], А. Чжан, З. Ліптон, М. Лі, А. Смола [6] та ін.

Наприклад, у дослідженні [1] акцентується увага на сучасних методах машинного навчання в сфері обробки великих даних. В роботі представлені головні питання та вигоди використання моделей машинного навчання в системі великих даних. Авторами у статті [3] досліджено використання моделей, таких як генетичний алгоритм, для розв'язання завдань у сфері закупівельних процесів, методів управління ланцюгами поставок та використання мови запитів DSQL для оптимізації рішень у сфері закупівель матеріалів. Робота [7] проводить аналіз застосування штучного інтелекту в сфері закупівель, визначаючи переваги різних підходів. У статті [8] розглядається загальний вплив штучного інтелекту на сферу закупівель, надаються приклади використання моделей штучного інтелекту в закупівлях та підтримці переговорів.

Отже, аналіз наукових робіт та досліджень підтверджує, що використання штучного інтелекту для розв'язання бізнес-задач у сфері закупівельної логістики викликає значний інтерес. Однак дослідження застосування моделей машинного навчання та різних математичних методів для управління закупівельною діяльністю великих компаній з географічно

розподіленою структурою не отримало належної уваги від дослідників і науковців. Таким чином, актуальним є проведення досліджень, спрямованих на вирішення завдань обробки та аналізу обсягів закупівельних даних за допомогою моделей машинного навчання, а також використання отриманих результатів для прогнозування обсягів у сфері закупівель.

Прогнозування міток об'єктів, відоме як класифікація, представляє лише одну з багатьох завдань у галузі машинного навчання. Ця галузь містить в собі великий набір варіантів практичних застосувань, розглянемо основні.

Класифікація тексту чи документа. Аналіз тексту або документа, що включає в себе завдання, такі як визначення теми тексту або документа, виявлення автоматичної непридатності чи надто розгорнутого контенту на веб-сторінці, а також виявлення спаму.

Обробка природної мови (NLP) охоплює різноманітні завдання, такі як визначення частин мови, іменованих сутностей, парсинг без контексту чи парсинг залежностей, що розглядаються як завдання навчання. В цих випадках прогнозування володіє конкретною структурою. Наприклад, у визначенні частин мови прогноз для речення представляє собою послідовність тегів частин мови, що відзначають кожне слово. У парсингу без контексту прогноз відображається у вигляді дерева. Це тільки декілька прикладів складніших завдань навчання, відомих як задачі структурованого прогнозування.

Застосування обробки мовлення. Сюди входять розпізнавання мови, синтез мовлення, верифікація та ідентифікація говорячої особи, а також пов'язані завдання, такі як моделювання мовлення та акустичне моделювання.

Застосування комп'ютерного зору. Сюди входить визначення об'єктів, ідентифікація об'єктів, розпізнавання обличчя, оптичне розпізнавання символів (OCR), а також пошук зображень на основі вмісту чи оцінка їх розташування.

Застосування обчислювальної біології. Це враховує передбачення функцій білка, ідентифікацію ключових сайтів та аналіз генетичних та білкових мереж.

Багато інших проблем. Виявлення шахрайства в кредитних картках, телефонних або страхових компаніях, вторгнення в мережі, вивчення вправності в грі у шахи, нарди чи Го, самостійне управління транспортними засобами, такими як роботи чи автомобілі, медична діагностика, розробка систем рекомендацій, пошукові системи чи системи вилучення інформації — ці завдання вирішуються з використанням методів машинного навчання. [9].

Для аналізу та прогнозування закупівельних даних можна використовувати основні методи машинного навчання, а саме:

- Кластерний аналіз (Clustering): Цей метод сприяє об'єднанню схожих об'єктів без передвизначення категорій. В закупівлях застосування кластерного аналізу дозволяє визначати категорії постачальників, типи закупок з подібними характеристиками і т.д.

- Класифікація (Classification): Цей метод визначає приналежність об'єкта до встановленої категорії. Наприклад, класифікація закупівельних даних на "проста закупівля", "стандартна закупівля" та "нестандартна закупівля". Категорії зазвичай встановлюються відповідно до чинної системи та політики закупівель організації. Результати кластерного аналізу також можна використовувати для визначення категорій.

- Регресійний аналіз (Regression): Ця модель використовується для отримання прогнозних числових значень на основі наявних даних. Наприклад, регресійний аналіз може використовуватися для прогнозування вартості, обсягів, трудомісткості закупівель тощо.

- Генетичний алгоритм (Genetic Algorithm): Це евристична модель пошуку, яка застосовується при розв'язанні оптимізаційних задач на основі імітації природного відбору. Він використовує механізми, аналогічні біологічній еволюції. Генетичний алгоритм ефективний у розв'язанні комбінаторних та

оптимізаційних завдань у закупівельній системі, враховуючи, що система є результатом еволюційного розвитку [3].

У кожній моделі машинного навчання є свої корисні та негативні аспекти, які необхідно врахувати при виборі моделі для окремої задачі аналізу чи прогнозування закупівельних даних (таблиця 1.1).

Таблиця 1.1 – Порівняння методів машинного навчання [1]

Метод	Переваги	Недоліки
Кластерний аналіз (Clustering)	- дозволяє виявити приховані залежності, аномалії та структури в наборах даних; - сегментація та групування даних за спільними ознаками	- вибір оптимальної кількості кластерів може бути складним завданням, залежить від вибору алгоритму та початкових параметрів; - чутливість до шуму та викидів в даних
Класифікація (Classification)	- ефективність в розв'язанні задач з поділом на визначені типи; - простота в реалізації та інтерпретації результатів	- низька ефективність у розв'язанні завдань, коли декілька класів частково перекриваються; - при роботі з великими наборами даних може виникнути перенавчання
Регресійний аналіз (Regression)	- простота та легка інтерпретація результатів; - використання для прогнозування числових значень	- передбачення може бути обмеженим простими лінійними залежностями; - недостатня гнучкість для деяких складних даних
Генетичні алгоритми (Genetic Algorithms)	можливість знаходження оптимальних рішень у складних задачах оптимізації	- часова та обчислювальна складність; - можливість потрапляння в локальні мінімуми
Глибоке навчання (Deep Learning)	- здатність вирішувати складні завдання аналізу даних; - висока точність прогнозів; - можливість відбирати ознаки з різнотипних даних і ігнорувати непоказові приклади у даних [1]	- потреба у великій кількості даних та обчислювальних ресурсів; - можливість перенавчання

Отже, можна визначити наступні напрямки використання моделей машинного навчання при обробці закупівельних даних:

1. Кластерний аналіз – для групування постачальників за спільними ознаками та групування даних використовується кластерний аналіз. Ця модель машинного навчання без вчителя дозволяє виділяти групи постачальників за подібними ознаками з метою якісного управління відносинами в закупівлях. Критерії сегментації включають характеристики різних етапів закупівельного процесу:

- для поділу постачальників на групи застосовують такі ознаки як форма власності, статутний фонд, належність до афілійованих компаній та партнерів, види діяльності, адреса, контактні дані та інші;

- для поділу матеріалів для закупівлі застосовують такі ознаки, як ціна, вага, розміри, тип, група матеріалів і інші характеристики;

- для виділення видів закупівельних операцій застосовують такі ознаки, як вартість закупівлі, категорія постачальника, приналежність постачальника до афілійованих компаній та партнерів, вид товару і інші параметри.

Основним завданням кластерного аналізу є групування подібних об'єктів у закупівельних даних у кластери так, щоб об'єкти, які належать до одного кластера, були більш схожі між собою, ніж з об'єктами з інших кластерів. Застосування кластерного аналізу в контексті закупівельних процесів дозволяє розподілити постачальників та інші дані про закупівлі. Це сприяє глибшому розумінню закупівельних процесів і виявленню можливостей для їх подальшого вдосконалення. Визначення кластерів постачальників з аналогічними характеристиками дозволяє компанії підвищити ефективність взаємодії з ними, використовуючи розроблені стратегії співпраці для кожної визначеної категорії.

2. Класифікаційні моделі – для розподілу закупівельних даних відповідно до встановлених категорій товарів, послуг або постачальників використовуються моделі класифікації, що відносяться до групи методів

навчання з вчителем. Такі моделі дозволяють відносити об'єкти до визначених раніше категорій за приналежними їм характеристикам та ознакам [17]. Щодо закупівельних даних, необхідно застосувати класифікацію для ідентифікації груп постачальників, товарів, послуг та інших параметрів закупівель. Класифікаційні алгоритми дозволяють ранжувати постачальників за різними критеріями, такими як тип діяльності, реєстраційні дані та досвід співпраці. Це полегшує визначення ключових постачальників для партнерства та різних видів закупівель, збільшуючи ефективність взаємодії. Класифікація товарів спрощує присвоєння їм визначених категорій, полегшуючи організацію та вибір необхідних товарів за типом, брендом чи характеристиками. Крім того, класифікаційні моделі дозволяють оцінити ризики, пов'язані з постачальниками на основі різних ознак та історії взаємовідносин. Це дозволяє підприємствам виявляти потенційно ненадійних постачальників і приймати рішення з урахуванням можливих наслідків. Класифікація закупівельних даних допомагає систематизувати їх за встановленими категоріями, такими як об'єми, види товарів чи послуг, строки поставки і т.д., що великою мірою покращує ефективність управління закупками через виділення різних видів операцій. Результати класифікації можуть бути використані при розробці або оновленні стратегій закупівель.

3. Регресійні моделі – прогнозування об'ємів та цін товарів для закупівлі на базі наявних накопичених даних та оптимізації закупівель застосовуються регресійні моделі. Враховуючи різноманітні фактори, такі як інвестиційні проекти, сезонність, маркетингові заходи та економічні коливання, ці моделі дозволяють прогнозувати попит на наступні періоди, що сприяє підвищенню достовірності планування закупівель та управлінню запасами.

Регресійні моделі надають можливість передбачати обсяги закупівель на визначений період, використовуючи дані про минулі закупівельні операції, попит, тенденції ринку та інші фактори. Період прогнозування встановлюється відповідно до періоду планування підприємства (місяць, квартал, рік), і

важливо враховувати циклічні коливання ринку та закупівельні цикли для конкретних товарів.

Ще одним застосуванням регресійних моделей є прогнозування коливань рівня цін на товари та послуги з урахуванням різних чинників, таких як вартість сировини, індекс інфляції, коливання валютного курсу, попит, конкуренція та економічні умови. Прогнозовані дані дозволяють приймати обґрунтовані рішення щодо бюджетів закупівель та встановлення оптимальних цінових пропозицій на продукцію.

Регресійні моделі визначають фактори, які найбільше впливають на обсяги закупівель та ціни, що дозволяє оптимізувати параметри закупівель для підвищення ефективності та зменшення витрат на закупівельні процеси. Ці моделі також можуть бути застосовані для визначення оптимальних обсягів запасів товарів, враховуючи вимоги споживачів та тимчасові зміни у попиті, з метою уникнення зайвих витрат і запасів та задоволення потреб споживачів.

Використання регресійних моделей для точного прогнозування необхідних фінансових ресурсів та ефективного управління бюджетом дозволяє розраховувати показники прибутку та витрат на основі запланованих обсягів закупівель та актуальних цін. Впровадження регресійних моделей може включати використання розширюваних нейронних мереж, що підвищує точність прогнозів. Загалом, використання регресійних моделей в області закупівель дозволяє приймати обґрунтовані рішення, знижувати витрати на закупівлі та оптимізувати стратегії закупівель.

4. Методи глибокого навчання – Застосування глибокого навчання для виявлення ризиків у закупівлях обумовлено його можливістю виявлення невідповідностей та шахрайства. Моделі глибокого машинного навчання визначають аномалії в виконаних угодах та ідентифікують подібні ризики у поточних даних. Ці методи виявляють сумнівні та неоднозначні закупівельні дані, виявляючи можливі випадки шахрайства або корупції. Використання моделей глибокого навчання дозволяє виявляти недостовірні дані, відсутність

підписів і печаток, а також сумнівні транзакції в інформації про оплати, з метою запобігання можливих втрат.

Для оперативного реагування на проблеми та уникнення катастрофічних наслідків необхідно використовувати алгоритми машинного навчання для виявлення порушень стандартів чи правил у закупівельних процесах шляхом постійного моніторингу закупівельних даних. Виявлення ризиків у ланцюгах постачань також є суттєвим аспектом, і методи глибокого навчання мають бути використані на всіх етапах ланцюга постачань для оцінки ризиків. З метою забезпечення стабільності поставок та розробки стратегій зменшення ризиків глибоке навчання виявить потенційні загрози.

Алгоритми машинного навчання можуть автоматично перевіряти контракти, відстежувати їх проходження в системі та своєчасно виявляти відхилення від умов договору. Такі невідповідності, як перевищення бюджету чи несвоєчасне виконання, будуть вчасно виявлені та своєчасно попереджені для уникнення втрат.

Ще одним напрямком використання методів глибокого навчання є вдосконалення закупівельних процесів. Оптимізаційні моделі включають різні параметри, такі як ціни постачальників, терміни поставок та обсяги попиту, сприяючи максимізації прибутку та зменшенню витрат. Машинне навчання дозволяє враховувати різноманітні змінні, такі як зміни у споживанні, економічні умови та соціальні фактори, для аналізу та ідентифікації тенденцій у закупівельних процесах. [4].

Застосування багатосарових нейронних мереж складної архітектури у моделях глибокого навчання має важливе значення і дозволяє аналізувати та виділяти шаблони в ієрархічній структурі [5].

1.2 Аналіз сучасних інформаційних технологій для аналізу та прогнозування закупівельних даних

1.2.1 Аналіз сучасних програмних інструментів аналізу та обробки даних

Сучасна бізнесова атмосфера визначається стрімким розвитком технологій обробки інформації, які надають розмаїття інструментів для аналізу та оптимізації управлінських процесів, зокрема в області закупівельних даних. Зокрема, важливо визначити ключові інформаційні технології, які широко використовуються з метою аналізу та обробки даних у сфері закупівель.

Однією з найпопулярніших інформаційних технологій є інтегровані системи аналізу та обробки даних, які забезпечують можливість оперативно реагувати на зміни в ринковому середовищі та проводити аналіз необхідних об'ємів даних. Серед таких систем варто відзначити табличні процесори, такі як Microsoft Excel, а також MatLab, RapidMiner Studio та Neural Designer.

Microsoft Excel є частиною пакету офісних програм Microsoft Office і представляє собою електронний аркуш де розміщуються дані у вигляді таблиць. Перелічимо основні функції Microsoft Excel.

1. Створення та редагування таблиць: Excel дозволяє створювати таблиці, де дані розташовуються у комірках, які формуються перетином рядків і стовпців. Кожна комірка може містити текст, числа, формули та інші дані.

2. Формули та функції: Excel має значний набір вбудованих формул і функцій, які дозволяють виконувати різноманітні обчислення. Тобто, можемо описувати формули для автоматичного розрахунку статистичних показників.

3. Графіки та Діаграми: Excel надає можливість зручно створювати графіки та діаграми, використовуючи дані, що містяться в таблицях. Це сприяє визначенню візуальної ясності та зрозумілості даних.

4. Сортування та фільтрація: Можливість впорядкованого розташування та відфільтрування даних надає можливість швидко структурувати та аналізувати обширні обсяги інформації.

5. Автоматизація завдань: Використовуючи макроси та мову програмування VBA (Visual Basic for Applications), можна автоматизувати повторювані завдання та створювати складні аналітичні інструменти. Вбудовані пакети аналізу та пошуку рішень стають особливо цінними для аналізу та обробки закупівельних даних.

6. Імпорт та експорт даних: Ексел дозволяє виконувати імпорт та експорт даних з різних джерел, таких як бази даних, текстові файли, Інтернет і інші.

7. Спільна робота та обмін даними: Ексел взаємодіє з іншими програмами Microsoft Office, сприяючи зручній спільній роботі та обміну даними між різними документами.

Microsoft Excel є популярним інструментом у бізнесі, фінансах, науці, освіті та інших сферах для обробки та аналізу даних.

MATLAB (MATrix LABoratory) – MATLAB представляє собою високорівневу мову програмування та інтерактивне середовище для виконання чисельних обчислень, візуального представлення даних, створення алгоритмів, моделювання та інших технічних областей. У своєму складі MATLAB має обширний набір вбудованих функцій, які дозволяють виконувати математичні обчислення, обробку сигналів, аналіз даних, створення графіків та вирішення інших завдань. Давайте розглянемо основні характеристики MATLAB.

1. Мова програмування: MATLAB використовується як мова програмування, обладнана власним синтаксисом та структурами даних. Її основна мета полягає в ефективній обробці матриць та виконанні чисельних операцій.

2. Чисельні обчислення: MATLAB є ефективним засобом для проведення чисельних обчислень і має вбудовані функції для вирішення різноманітних математичних завдань, таких як лінійна алгебра, диференціальні рівняння, статистика та інші.

3. Зображення та графіки: MATLAB надає доступні інструменти для візуалізації даних і побудови графіків. Користувачі можуть легко створювати

двовимірні та тривимірні графіки, зображення, анімації та інші візуальні елементи.

4. Робота з файлами та зовнішніми даними: MATLAB може легко взаємодіяти з різноманітними форматами файлів, такими як Excel, текстові файли, дані з приладів та інші. Це робить його важливим інструментом для обробки реальних даних.

5. Розробка алгоритмів: MATLAB використовується для створення та випробування алгоритмів, включаючи ті, які відносяться до штучного інтелекту та обробки сигналів.

6. Спеціалізовані інструменти: MATLAB володіє різноманітними спеціалізованими інструментами для виконання завдань у конкретних галузях, таких як обробка зображень, обробка сигналів, контроль та автоматизація.

MATLAB використовується в різних сферах, таких як інженерія, фізика, фінанси, біологія, комп'ютерні науки та інші, де необхідні чисельні обчислення та аналіз даних.

RapidMiner Studio – це комплексне середовище для обробки даних, використання методів машинного навчання та проведення аналітичних досліджень. Який дозволяє фахівцям з аналітики, дослідникам та бізнес-користувачам отримувати нові знання з даних за допомогою різноманітних алгоритмів машинного навчання та статистичних методів. Розглянемо основні риси RapidMiner Studio.

1. Візуальне програмування: RapidMiner використовує візуальний інтерфейс для створення проектів з аналізу даних та моделювання, що полегшує перетягування та розміщення вузлів (блоків) для різних етапів аналізу та обробки даних, моделювання.

2. Широкий набір алгоритмів: RapidMiner надає різноманітні вбудовані алгоритми класифікації, регресії, кластеризації, асоціативного аналізу та інших алгоритмів обробки даних.

3. Інтеграція з іншими інструментами: Забезпечення взаємодії з різними джерелами даних, базами даних та іншими інструментами для проведення аналізу даних.

4. Автоматизація процесів: Можливість автоматизації трудомістких процесів обробки даних та створення власних моделей машинного навчання.

5. Інтегроване управління даними: Засоби для витягування, трансформації та завантаження даних та очищення даних.

6. Візуалізація результатів: Засоби для побудови графіків та формування візуалізацій результатів аналізу для кращого розуміння отриманих знань.

RapidMiner широко використовується в галузі дослідження даних, бізнес-аналітики, прогнозування, аналізу текстів та інших областях для вирішення завдань, пов'язаних з аналізом та обробкою даних.

Neural Designer - це програмний продукт для розробки та тестування алгоритмів машинного навчання на основі нейронних мереж. Його використання дозволяє інженерам, дослідникам та аналітикам використовувати методи глибокого навчання без глибоких знань в області машинного навчання чи програмування. Надамо опис основним характеристикам Neural Designer.

1. Візуальне програмування: Neural Designer пропонує інтерфейс візуального програмування, який дозволяє користувачам конфігурувати та навчати і тестувати сформовані нейронні мережі, просто перетягуючи та розміщуючи вузли та з'єднуючи їх візуально.

2. Багатий набір алгоритмів: Neural Designer містить основні алгоритми машинного навчання для вирішення завдань класифікації, регресії, кластеризації та широкий набір прикладів таких задач.

3. Автоматична оптимізація параметрів: Neural Designer автоматично оптимізує гіперпараметри моделей для досягнення оптимальних результатів.

4. Підтримка глибокого навчання: Інструмент орієнтований на використання глибокого навчання, використовуючи багатoshарові нейронні мережі для ефективного вирішення складних завдань.

5. Візуалізація результатів: Neural Designer містить потужні інструменти для аналізу результатів навчання моделей та їх візуалізації.

6. Інтеграція з іншими інструментами: Можливість легко інтегрувати інструмент з іншими засобами для обробки інформації, аналізу даних.

Neural Designer спрощує процес створення та навчання і тестування нейронних мереж, дозволяючи фахівцям із різних галузей використовувати технології глибокого навчання, навіть не маючи глибокого технічного досвіду.

Також слід відзначити інструменти обробки та візуального представлення даних, такі як Power BI та Tableau, які завдяки технологіям штучного інтелекту, можуть забезпечити швидкий аналіз і візуалізацію даних.

Power BI - це сервіс бізнес-аналітики від компанії Microsoft, який дозволяє створювати візуалізації даних та формувати звіти та панелі для прийняття бізнес-рішень. Розглянемо основні характеристики Power BI.

1. З'єднання з джерелами даних: Power BI дозволяє використовувати різноманітні джерела інформації, такі як бази даних, електронні таблиці Excel, хмарові сервіси і т. д. Це може бути як внутрішня база даних компанії, так і зовнішні джерела інформації.

2. Візуалізація даних: Сервіс дозволяє створювати інтерактивні графіки, діаграми, карти та інші візуальні компоненти для ефективного подання даних. Користувачі можуть швидко аналізувати та сприймати великі обсяги інформації.

3. Формування звітів: Power BI дозволяє створювати звіти, які можуть включати кілька візуальних компонентів та графіків. Звіти можуть бути настроєні для автоматичного оновлення інформації з джерел даних.

4. Інтеграція з Excel: Power BI легко взаємодіє з Excel, що дозволяє використовувати дані та звіти з Excel у Power BI та навпаки.

5. Зберігання даних у хмарі: Інформацію та звіти можна зберігати у хмарному сервісі Microsoft Power BI, що спрощує обмін даними та спільну роботу в команді.

6. Поєднання з іншими сервісами: Power BI може бути легко інтегровано з іншими сервісами Microsoft, такими як Azure, SharePoint, різноманітні джерела даних і т. д.

Power BI надає можливість не лише візуалізації даних, але і отримання інсайтів, які сприяють ухваленню обґрунтованих бізнесових рішень. Цей інструмент виявляється важливим для аналізу даних у сфері бізнесу та організацій.

Tableau - це інноваційна платформа що здійснює візуалізацію та аналізу даних, яка забезпечує зручне об'єднання, візуалізацію та спільну роботу з інформацією. Tableau дозволяє створювати інтерактивні та професійно виглядаючі графіки, діаграми, табличні звіти та інші візуальні елементи. Основні риси платформи Tableau:

1. З'єднання з різними джерелами даних: Tableau здатний під'єднуватися до різних джерел інформації, таких як бази даних, електронні таблиці, веб-сервіси та інші. Це дозволяє користувачам об'єднувати дані з різних джерел для комплексного аналізу.

2. Створення візуалізацій за допомогою перетягування та викидання: Завдяки інтуїтивному інтерфейсу користувачі можуть легко створювати візуалізації, використовуючи перетягування та викидання полів даних.

3. Інтерактивність: Tableau дозволяє створювати інтерактивні візуалізації, що дозволяє користувачам прямо взаємодіяти з даними. Таким чином, користувач може вибирати елементи графіку, фільтрувати дані та досліджувати різні аспекти інформації.

4. Спільна робота та публікація: Tableau надає можливість користувачам ділитися своїми візуалізаціями та аналітичними звітами в Інтернеті, а також вбудовувати візуалізації в веб-сторінки чи додатки.

5. Автоматизація завдань та розкладу: Tableau дозволяє автоматизувати різні рутинні завдання та планувати оновлення даних з їх джерел.

6. Підтримка географічних даних: Tableau ефективно взаємодіє з географічними даними, що відкриває можливості створення карт та візуалізацій на географічній основі.

Tableau має широке застосування у різних напрямках, де велике значення має візуалізація даних, такі як бізнес-аналітика, фінанси, маркетинг, наука про дані та інші.

1.2.2 Дослідження сучасних методів машинного навчання для аналізу закупівельних даних

Машинне навчання широко визначається як обчислювальні методи, які використовують досвід для поліпшення продуктивності або точності прогнозів. Тут поняття "досвід" вказує на минулу інформацію, доступну для навчання, яка, зазвичай, представлена у формі електронних даних, зібраних і наданих для аналізу. Вони можуть мати вигляд цифрованих тренувальних наборів, які позначені людьми, або інших типів інформації, отриманих взаємодією з оточенням. У всіх випадках якість та розмір грають важливу роль у успішних прогнозах, здійснених системою навчання. [9]

Однією з задач навчання є використання обмеженого вибіркового набору випадково обраних документів, що мають позначену тему, для точного прогнозування теми невидимих документів. Зрозуміло, що чим більший вибіркового набір, тим простіше завдання. Проте складність завдання також залежить від якості міток, призначених документам у вибіркового наборі, оскільки мітки можуть не завжди бути правильними, а також від кількості можливих тем. [9]

Машинне навчання передбачає розробку дієвих та точних алгоритмів прогнозування. Як і в інших галузях комп'ютерних наук, деякі критичні показники якості цих алгоритмів - це їх складність за часом та простором. Але в машинному навчанні нам також потрібно поняття обсягу вибірки, щоб оцінити розмір вибірки, необхідний для того, щоб алгоритм вивчив родину концепцій. Загалом теоретичні гарантії для навчального алгоритму залежать від складності розглянутих класів концепцій та розміру навчальної вибірки. [9]

Оскільки успіх навчального алгоритму залежить від даних що використовувались, машинне навчання неодмінно пов'язане із аналізом даних та статистикою. Загалом техніки навчання – це дані-орієнтовані методи, які поєднують фундаментальні концепції інформатики з ідеями зі статистики, ймовірності та оптимізації. [9]

Методи машинного навчання для обробки і аналізу даних закупівель, зазвичай ділять за наступними напрямками:

1. Кластерний аналіз для розподілу постачальників і даних по закупівлям:
 - використання навчання без вчителя для ідентифікації постачальників за групами що мають подібні характеристики;
 - мета - ефективне керування закупівельними взаєминами та оптимізація співпраці.
2. Класифікаційні моделі при розпізнаванні видів товарів, послуг та постачальників:
 - використання навчання з учителем для присвоєння категорій у наявних закупівельних даних;
 - допомагає систематизувати та класифікувати продукти, послуги або постачальників.
3. Регресійні моделі при прогнозуванні об'ємів та вартості закупівель:
 - використання історичних даних для прогнозування об'ємів та вартості закупівель;
 - прогнозування попиту на товари та оптимізація стратегій закупівель.

4. Виявлення ризиків та аномальних операцій у закупівельних процесах:

- використання методів глибокого навчання для виявлення ризиків та аномальних операцій у закупівельних угодах;
- максимізація ефективності системи виявлення небезпек та запобігання можливим проблемам.

Використання нейронних мереж для виявлення базових та прихованих взаємозв'язків у закупівельних даних є суттєвим аспектом машинного навчання. Ці потужні інструменти можуть ефективно розпізнавати та аналізувати неочевидні залежності та взаємодії в даних про закупівлі. Архітектура цих мереж подібна до структури нейронної системи людини, дозволяючи їм автоматично виявляти приховані закономірності в даних, які можуть залишатися непоміченими традиційними методами аналізу [4].

Нейронні мережі (НМ) представляють собою систему взаємопов'язаних штучних нейронів, які сполучені синоптичними вагами. Ці мережі взаємодіють із вхідною інформацією, обробляють її та формують сукупність вихідних сигналів [4]. Мережа нейронів включає кілька шарів: вхідний, внутрішні (приховані) і вихідний. Вхідний шар взаємодіє з вхідними даними, в той час як вихідний шар відповідає за вихідні сигнали (рис. 1.1).

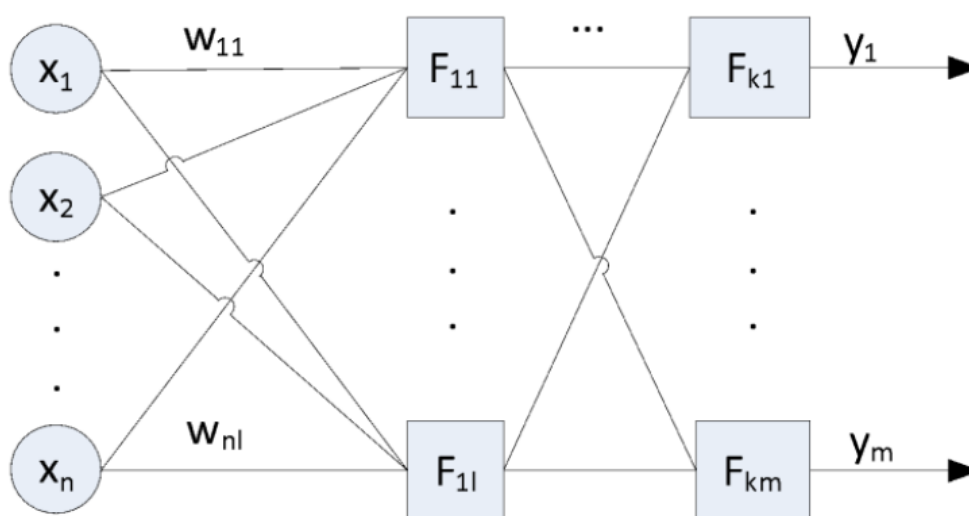


Рисунок 1.1 – Загальна архітектура штучної нейронної мережі [4]

Кожен шар включає кілька нейронів, які між собою взаємодіють через зв'язки з власними вагами. Навіть при поверхневій схожості штучні нейронні мережі виявляють властивості, що спільні з мозковою діяльністю [4]:

- вчитися;
- узагальнювати;
- абстрагувати.

Нейронні мережі виявляються надзвичайно потужним інструментом для аналізу складних залежностей у закупівельних даних, сприяючи прийняттю обґрунтованих рішень, прогнозуванню тенденцій та виявленню аномалій. Це допомагає підвищенню ефективності та оптимізації закупівельних процесів. Основні напрямки використання нейронних мереж включають кластерний аналіз, класифікацію, вирішення задач регресії та виявлення ризиків.

Кластерний аналіз є ефективним способом для виявлення нових залежностей і груп із схожими властивостями в закупівельних даних. Основна ціль цього методу це групування подібних ознак закупівельних даних в класи або кластери так, щоб згруповані об'єкти в кластері були схожими більше між собою, ніж з об'єктами розподіленими до інших кластерів. Використання кластерного аналізу в сфері закупівель дає можливість здійснити перерозподіл постачальників та даних закупівель, що сприяє глибшому розумінню та оптимізації процесів закупівель.

Ієрархічна кластеризація, спрямована на створення ієрархії кластерів "згори донизу" або "знизу догори", дозволяє розміщувати спостереження в одному кластері та рекурсивно розщеплювати їх вниз чи вгору по ієрархії [4].

Метод k-середніх – це процедура оптимального розбиття множини даних на певну кількість кластерів, спрямована на групування даних так, щоб досягти екстремуму цільової функції алгоритму розбиття [4].

Таблиця 1.2 – Порівняння моделей кластеризації [4]

Тип моделі	Переваги	Недоліки	Використання
Ієрархічна кластеризація	Отримання оптимальної кількості кластерів безпосередньо з моделі, легка інтерпретація дендрограми кластерів	Не ефективна для великих вибірок	Групування постачальників, Класифікація за географічним розташуванням
Метод k-середніх	Простий для зрозуміння, легко адаптується, ефективний як для великих, так і для малих вибірок	Вимагає задання кількості кластерів	Виявлення надійних постачальників, Виявлення ризиків

Кластерний аналіз дозволяє виконувати завдання, які наведені далі.

Групування постачальників за різноманітними факторами, такими як географічне розташування, обсяги постачань та якість матеріалів, для ефективного вибору учасників у закупівельних процесах. Це сприяє оптимальному вибору постачальників для різних категорій товарів чи послуг, розширює конкурентне середовище та полегшує співпрацю з різними типами постачальників.

Визначення груп постачальників за характеристиками надійності, якості та успішності у співпраці з організацією. Це дозволяє виділити надійних та довгострокових партнерів, зменшуючи ризики та підвищуючи ефективність закупівель.

Розділення постачальників за видом продукції або послуг, що дозволяє підприємствам спрямовувати зусилля на співпрацю з певними групами постачальників та встановлювати спеціальні умови для кожної групи.

Визначення груп постачальників за регіональним принципом, спрямоване на зосередження на локальних або міжнародних постачальниках для ефективного контролю над ланцюжками постачань.

Розуміння подібностей та потреб між постачальниками для розробки індивідуальних стратегій співпраці з кожною групою.

Виявлення ризикових груп постачальників, таких як ті з низькою надійністю чи обмеженим досвідом, для приділення їм спеціальної уваги та управління ризиками.

Виділення типів закупівельних операцій з різною трудомісткістю та різним життєвим циклом для ефективного розподілу ресурсів та аналізу закупівельних процесів.

Визначення ознак груп закупівельних операцій, що виходять за рамки звичайних та мають аномальні характеристики, для виявлення можливих ризиків та шахрайських дій.

Кластерний аналіз стає важливим інструментом для підприємств, що націлені на ефективне керування закупівельними процесами та встановлення успішної співпраці з різними постачальниками. Застосування кластерного аналізу для сегментації постачальників та закупівельних даних допомагає виявити та зрозуміти глибокі тенденції, взаємозв'язки та структуру даних. Це також дозволяє сформулювати критерії, на основі яких можна створювати класи постачальників та видів закупівель, сприяючи формуванню ефективних рішень та оптимізації процесів у сфері закупівель.

У вирішенні завдань кластеризації важлива здатність збереження топологічної подібності, тому застосовуються архітектури нейронних мереж. У цих архітектурах вхідний та вихідний шари складаються з радіальних нейронів, впорядкованих у структуру, а нейрони вихідного шару розташовані у вузлах двовимірної (об'єкт-нейрон) карти [4].

Класифікаційні методи використовуються для розпізнавання приналежності постачальників, товарів та послуг до встановлених груп у закупівельних даних. Класифікація, одне з пріоритетних завдань машинного навчання, використовується в різних галузях, включаючи аналіз закупівельної інформації. Ці методи сприяють віднесенню об'єктів до визначених груп або класів на підставі їх характеристик.

Логістична регресія, також відома як модель логітів, представляє собою статистичну модель для прогнозування ймовірності виникнення певної події (приналежність до конкретного класу) на основі значень набору ознак [4].

Метод опорних векторів — це модель, в якій побудова роздільної поверхні здійснюється за допомогою невеликої підмножини точок, які знаходяться в критичній зоні для розділення, в той час як інші спостереження в навчальній вибірці ігноруються [4].

Машина опорних векторів (Support Vector Machine, SVM) представляє собою нелінійне розширення лінійного класифікатора, що ґрунтується на розширенні розмірності вихідного простору предикторів за допомогою спеціальних ядерних функцій [4].

Метод k-найближчих сусідів (K-Nearest Neighbors, K-NN) передбачає, що тестований об'єкт отримає таку ж мітку класу, як і навчальні об'єкти в локальній області його найближчого оточення [4].

Наївний байєсовський класифікатор (Naive Bayes Classifier) — ймовірнісний класифікатор, що базується на теоремі Байєса з нежорсткими припущеннями про незалежність подій, надаючи формальний метод для ухвалення рішень та врахування нової інформації [4].

Таблиця 1.3 – Порівняння моделей класифікації [4]

Модель	Переваги	Недоліки	Застосування
Логістична регресія	Ймовірнісний підхід для оцінки значущості факторів	Припущення про лінійну розподіленість	Виявлення ризикових постачальників
Метод опорних векторів	Не чутливий до перенавчання	Потреба розрізняти лінійні та нелінійні випадки	Розподіл товарів за категоріями
Метод K-найближчих сусідів	Простий, швидкий і ефективний	Вручну обирається кількість сусідів	Розподіл товарів за категоріями

Продовження таблиці 1.3.

Байєсівський класифікатор	Стійкий до викидів ймовірнісний підхід, працює в лінійно нерозподільних просторах	Припущення про рівну статистичну значущість факторів	Визначення категорій закупівельних операцій
Дерева рішень	Добре інтерпретуються, дозволяють моделювати лінійні та нелінійні залежності	Погано працюють на малих вибірках, існує ризик перенавчання	Класифікація постачальників
Випадковий ліс	Висока точність, особливо для моделювання нелінійних залежностей	Неінтерпретовані, схильні до перенавчання, потрібно вручну обирати кількість дерев	Класифікація постачальників за типом товарів та рівнем послуг

Дерево рішень (Decision tree) – об'єднання логічних правил типу "якщо ... то ..." в структуру "дерева". Перевагою використання дерев рішень є їхня здатність виявляти нетипові випадки, що відрізняє їх, наприклад, від логістичної регресії [4].

Випадковий ліс (Random forest) – це процедура послідовної композиції алгоритмів на основі вирішальних дерев [19]. Для кожного розбиття ознаки, при навчанні, вибираються з деякої випадкової підмножини ознак [4].

Алгоритми класифікації дозволяють:

- класифікувати постачальників за різними групами відповідно до їх виду діяльності, використовуючи реєстраційні дані та інформацію, накопичену з попередніх взаємодій. Алгоритми класифікації проводять аналіз статусу підприємства, розмір його статутного фонду, кількість працівників, види діяльності, обсяги поставок товарів, строки надання послуг та інші атрибути постачальників. Це сприяє визначенню партнерів-постачальників для різних видів закупівель і полегшує взаємодію з ними;

- автоматично присвоювати відповідні ознаки групам даних в режимі реального часу, спрощуючи та прискорюючи процес організації поставок та

пошуку необхідних товарів. Наприклад, потрібні товари можуть бути класифіковані за типом, брендом, властивостями тощо;

- виявляти ризикових постачальників. Використання методів машинного навчання дозволяє оцінювати ризиковість постачальників на основі їх документації, досвіду співпраці та інших даних. Це сприяє ідентифікації потенційно ризикових постачальників та узгодженню взаємодії з ними, враховуючи можливі наслідки;

- групувати їх за чітко визначеними категоріями відповідно до стандартів організації, такими як обсяг, типи товарів чи послуг, терміни поставки і інші характеристики. Класифікація закупівельних операцій дозволяє ефективно управляти процесом закупівель, відділяючи на початкових етапах прості від складних, специфічні від стандартних, термінові від звичайних і т. д. Це спрощує вибір стратегій закупівель для конкретних об'єктів закупівель і дозволяє відразу направляти закупівлі відповідним спеціалізованим підрозділам для виконання.

Класифікаційні методи машинного навчання представляють собою потужний інструмент для ідентифікації та групування об'єктів за визначеними категоріями. Це дозволяє класифікувати закупівельні об'єкти в групи, до яких застосовують алгоритми аналізу та оцінки, спрямовані на підвищення ефективності та точності аналізу закупівельних процесів. У завданнях класифікації важлива узагальнююча здатність мережі, тому використовуються архітектури нейронних мереж, які звужуються, з меншою кількістю нейронів на першому шарі порівняно з вхідним [4].

Моделі регресії для прогнозування об'єму та вартості закупівель є ключовим елементом стратегічного управління організаціями. Прогнозування цих параметрів має критичне значення для забезпечення чіткого планування, оптимізації запасів та забезпечення необхідних ресурсів. Використання машинного навчання, зокрема регресійних моделей, сприяє отриманню точних та надійних прогнозів, що суттєво впливає на стратегічне прийняття рішень та раціоналізацію витрат.

Регресійний аналіз, який є ключовим методом машинного навчання, застосовується у сфері аналізу закупівельних даних для прогнозування об'єму та вартості товарів та послуг. Моделі регресії дозволяють виявити залежність між однією чи декількома незалежними змінними (наприклад, характеристиками товарів, термінами поставки, попитом) та залежною змінною (ціною або обсягом закупівель) [21].

Лінійна регресія (Linear regression) є прикладом моделі встановлення лінійної залежності між змінними залежною та незалежними, що дозволяє прогнозувати значення однієї змінної на основі інших факторів. [4].

Поліноміальна регресія (polynomial regression) – це модель, що визначає нелінійну залежність між незалежною та залежною змінними і є поліномом.

Дерево рішень (Decision Tree) – це об'єднання логічних правил у форматі "ЯКЩО ... ТО ..." в структуру "дерева", що створює ієрархічну послідовність правил. Воно складається з вузлів, де проводиться перевірка умови, та листя (вузлів рішення), які визначають клас [4].

Випадковий ліс (Random forest) – це процедура послідовної композиції алгоритмів машинного навчання, яка базується на вирішальних деревах. Під час навчання для кожного розбиття ознаки обираються з випадкової підмножини ознак [4].

Регресійні моделі можуть служити для таких цілей(таблицю 1.4):

- прогнозування попиту на різноманітні товари та послуги, базуючись на історичних даних про закупівлі;
- прогнозування обсягів закупівель протягом конкретного періоду за допомогою зібраних закупівельних даних, попиту, ринкових тенденцій та інших чинників попередніх періодів.

Таблиця 1.4 – Порівняння моделей регресії [4]

Модель	Переваги	Недоліки	Застосування
Лінійна регресія	Працює з вибірками будь-якого обсягу, дозволяє оцінити значущість факторів	Припущення лінійності	Прогнозування цін на товари та послуги, оптимізація запасів
Поліноміальна регресія	Працює з вибірками будь-якого обсягу, враховує нелінійність	Потрібно вручну підбирати ступінь полінома	Прогнозування попиту на товари, прогнозування обсягів закупівель, аналіз впливу факторів на обсяги та ціни
Дерева рішень	Добре інтерпретуються, дозволяють моделювати лінійні та нелінійні залежності	Погано працюють на малих вибірках, велика ймовірність перенавчання	Прогнозування обсягів закупівель
Випадковий ліс	Висока точність, особливо у разі моделювання нелінійних залежностей	Не інтерпретуються, легко перенавчаються, потрібно вручну підбирати кількість дерев	Прогнозування доходів та витрат, аналіз впливу факторів на обсяги та ціни

Враховуючи різні аспекти, такі як запуснені проекти, сезонні впливи, маркетингові заходи та економічні тренди, регресійні моделі допомагають здійснити максимально точний прогноз попиту на наступні періоди. Це підвищує ефективність у плануванні закупівель та управлінні запасами. Зазвичай період прогнозу відповідає установленим циклам життєдіяльності підприємства (місяць, квартал, рік). Однак, важливо також враховувати стан ринку та особливості циклів закупівель для конкретних товарних категорій, які можуть значно відрізнятися від встановлених циклів. Це дозволяє ефективно планувати та управляти запасами та ресурсами організації.

Прогнозування цін на послуги і товари є ключовим елементом стратегічного планування підприємств. Використання регресійних моделей дозволяє передбачати зміни цін на основі різноманітних факторів, таких як вартість сировини, інфляція, валютні коливання, сезонний попит, конкуренція та економічний стан. Це допомагає підприємствам раціонально виділяти бюджети на закупівлі та встановлювати оптимальні ціни для своїх продуктів.

Аналіз впливу цих факторів на об'єми та вартість закупівель дозволяє ідентифікувати ключові параметри, які впливають на ефективність закупівельних процесів. Регресійні моделі виявляють, які чинники мають більший вплив, допомагаючи оптимізувати стратегії для покращення ефективності.

Оптимізація рівнів запасів стає більш точною завдяки використанню регресійних моделей при прогнозуванні очікувань споживачів та змін у попиті. Це дозволяє уникнути зайвих запасів, що звільняє капітал та забезпечує належний рівень запасів продукції для задоволення попиту споживачів.

Прогнозування доходів та витрат, базуючись на регресійних моделях, сприяє точнішому плануванню фінансів та ефективному керуванню коштами організації.

Загалом, регресійні моделі виявляються потужним інструментом для прогнозування об'ємів та вартості закупівель, допомагаючи підприємствам ухвалювати обґрунтовані рішення, ефективно управляти запасами та ресурсами, а також оптимізувати ціноутворення для своїх продуктів і послуг. У завданнях прогнозування важлива точність, тому доцільно використовувати нейронні мережі з архітектурою, що розширюються, де кількість нейронів на першому шарі перевищує кількість на вхідному [4].

Застосування глибокого навчання для виявлення ризиків та аномальних операцій стає вирішальним завданням для ефективного керування процесами закупівель та запобігання можливим проблемам. Використовування методів глибокого машинного навчання дозволяє автоматизувати цей процес та забезпечити раннє виявлення небажаних аномалій та ризиків (таблиця 1.5).

Таблиця 1.5 – Порівняння моделей глибокого навчання [6]

Модель	Переваги	Недоліки	Використання
Згорткові нейронні мережі	Ефективно виявляють шаблони та об'єкти на великих обсягах візуальних даних	Можуть ускладнюватися під час навчання при обмежених даних, потребують значних обчислень	Виявлення шахрайства
Рекурентні нейронні мережі	Враховують послідовний контекст, здатні робити передбачення на основі попередніх даних	Мають проблеми з вивченням довгих залежностей (зниклі градієнти), не завжди підходять для довгих послідовностей	Виявлення шахрайства та корупції
Довга коротко-часна пам'ять	Розв'язують проблеми зниклих градієнтів в Рекурентних нейронних мережах, можуть враховувати довгі залежності в даних	Можуть ускладнюватися під час навчання та вимагати більше обчислень	Виявлення невідповідностей до контрактів, Оптимізація закупівель
Сіамські мережі	Зменшують вплив зміни властивостей даних, дозволяючи розв'язувати завдання порівняння та класифікації пар даних	Можуть бути складні для налаштування та вимагати більше даних для навчання	Моніторинг закупівельних процесів
Мережі засновані на увазі	Здатні акцентувати важливі частини вхідних даних, покращуючи точність	Вимагають більше ресурсів для обчислень та навчання	Оцінка ризиків, Персоналізовані рекомендації
Трансформери	Добре працюють з послідовними даними будь-якої довжини, можуть виявляти складні зв'язки	Вимагають більше обчислень, особливо при великих даних	Оптимізація закупівель, Аналіз тенденцій та ризиків

Згорткові нейронні мережі (Convolutional Neural Networks, CNN) використовуються для обробки зображень та візуальних даних, проявляючи здатність виявляти шаблони та об'єкти на зображеннях [6].

Рекурентні нейронні мережі (Recurrent Neural Networks, RNN) є відмінним вибором при діях з послідовними даними, такими як мовний текст чи часові ряди, завдяки їхній здатності враховувати попередній контекст [6].

Довга короткочасна пам'ять (Long Short-Term Memory, LSTM) представляє собою розширення Рекурентної нейронної мережі та відзначається високою здатністю враховувати довгострокові взаємозв'язки в послідовних даних допомагає уникнути проблеми втрати градієнтів [6].

Сіамські мережі (Siamese Neural Networks) використовуються для вирішення завдань схожості та класифікації пар даних, наприклад, для визначення, чи належать дві фрази одній темі.

Мережі, що базуються на увазі (Attention-based Neural Networks), використовують механізм зосередження уваги для акцентування на найбільш важливих частинах вхідних даних, що сприяє покращенню точності передбачень [6].

Трансформери (Transformers) представляють собою структуру моделей, яка використовує механізм зосередження уваги для обробки послідовних даних, демонструючи здатність працювати з послідовностями різної довжини [6].

Розглядаючи методи оптимізації для аналізу закупівельних даних, важливо зосередитися на використанні машинного навчання.

Алгоритми машинного навчання дають можливість ефективно прогнозувати тенденції в закупівельних процесах, враховуючи різноманітні параметри, що сприяє більш точному та оперативному управлінню закупівлями. Також важливо розглядати конкретні моделі та їх переваги у контексті аналізу закупівельних даних.

Сучасні технології мають потенціал покращити взаємодію між аналітичними інструментами та системами управління підприємством.

Важливо досліджувати можливості інтеграції з вже наявними ентерпрайз-системами, такими як ERP та CRM, для автоматизації обміну даними та створення єдиної електронної екосистеми для керування закупівельними процесами.

Поза аналізом сучасних технологій, важливо визначити виклики, які виникають при їхньому впровадженні, і потенційні перспективи для майбутнього розвитку.

Питання безпеки, етичного використання даних та постійного оновлення технічних знань персоналу є ключовими аспектами для розгляду.

Отже, на основі глибокого аналізу інформаційних технологій та методів оптимізації на сьогоднішній день, використовуваних для аналізу та прогнозування закупівельних даних, можемо здійснити постановку задачі нашого дослідження.

1.3 Постановка задачі дослідження

Типовий для більшості підприємств процес закупівель починається з розробки бізнес плану виробничих підрозділів на наступний рік. На цьому етапі, Замовник розробляє план реалізації проектів та формує перелік та об'єм матеріалів та послуг, необхідних для реалізації запланованих проектів.

Після перевірки сформованого списку із залишками матеріалів на складах, формується бюджет закупівель на рік і після його захисту та затвердження у системі обліку формуються заявки на закупівлю матеріалів та послуг.

Отримавши затвержені заявки в системі, служба закупівель аналізує заявки та формує план тендерів на рік та розраховує необхідні ресурси для здійснення закупівель і при потребі подає заявку на пошук персоналу, для забезпечення своєчасності виконання заявок. Після чого здійснюються

тендерні процедури та вибір постачальників, з якими служба постачання укладає договори та на фінальному етапі організовує поставку на склади підприємства.

Загальний процес закупівлі зображено на рисунку 1.2.

У вузькому розумінні, процес закупівлі обмежується етапами «Аналізу створених заявок на закупівлю» та «Поставкою». Що відноситься до компетенцій Служби закупівель.

До компетенцій Замовника відноситься визначення переліку матеріалів та послуг, що необхідні для закупівлі, їх об'єм, строки поставки (коли матеріали мають бути доступними на складі Підприємства, або послуги наданими), забезпечити наявність необхідного бюджету для здійснення їх закупівлі та сформулювати і подати заявки на закупівлю до служби закупівель.

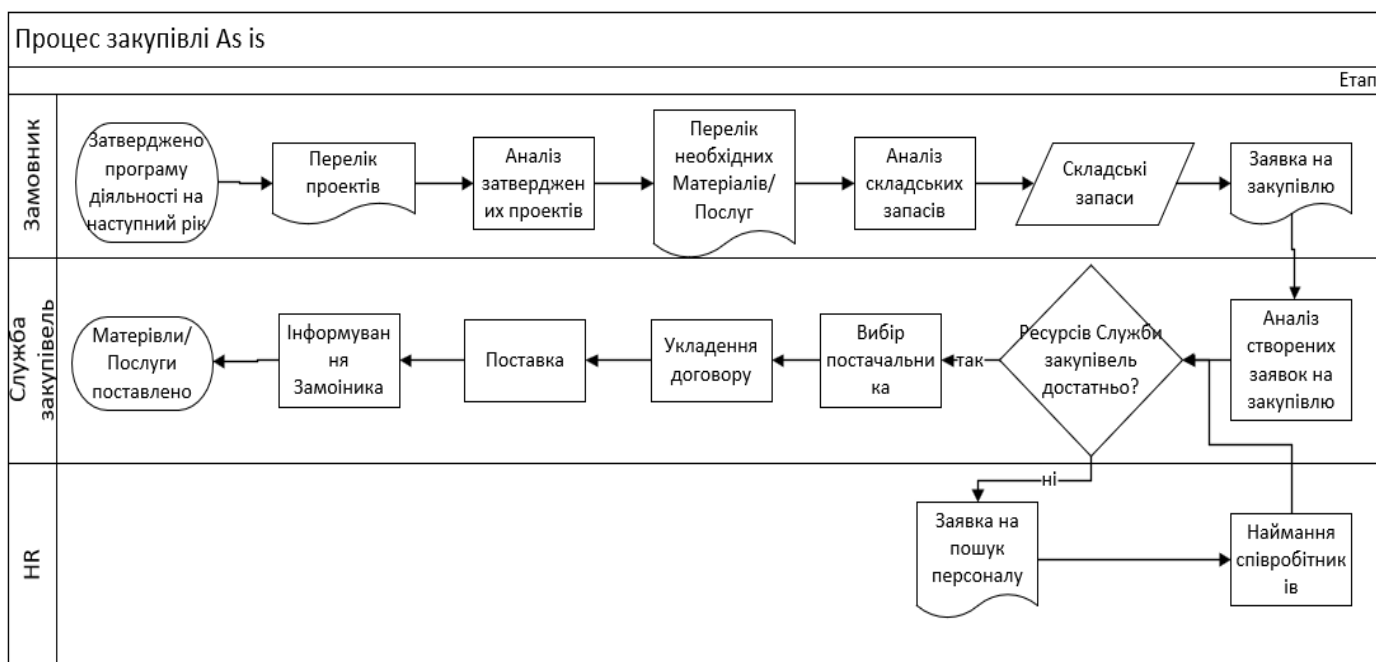


Рисунок 1.2 – Процес «as is»

Для Служби закупівель важливе розуміння загальної кількості заявок на період (рік), класифікація заявок за групами та типом тендера, строки

поставки, необхідні трудовитрати на одну заявку та тривалість процесу закупівлі.

1.4 Висновки за розділом

Слабким місцем у даному бізнес процесі є етап створення заявок та забезпечення необхідних ресурсів служб закупівель та постачання. Заявки створюються у не повному обсязі, в наслідок обмежень бюджетного процесу, крім того протягом року можуть виникати аварійні ситуації та створюватися нові проекти, що призводить до росту кількості заявок протягом року від 15% до 50%, що в свою чергу створює значні складнощі для забезпечення своєчасності виконання заявок на закупівлю. Такі значні коливання кількості заявок та розтягнутість процесу створення заявок у часі не дозволяє службам закупівель і постачання своєчасно забезпечити необхідну кількість персоналу для виконання заявок, оскільки середній період підбору та навчання нового персоналу становить від 3 до 5 місяців.

Вирішення описаної проблеми можуть забезпечити організаційні заходи та введення засобів обробки, прогнозування та аналізу даних із закупівель.

Отже головними завданнями дослідження є:

- аналіз та вибір моделей машинного навчання для обробки та аналізу закупівельних даних для пошуку головних закономірностей з метою здійснення точних прогнозів закупівельних операцій;
- розробка удосконаленого бізнес-процесу управління закупівлями;
- розробити технічне завдання на розробку Системи управління закупівлями, з використанням моделей машинного навчання для аналізу оброблення закупівельних даних.

РОЗДІЛ 2. ДОСЛІДЖЕННЯ ВИКОРИСТАННЯ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЗАКУПІВЕЛЬНИХ ДАНИХ У СЕРЕДОВИЩІ RAPIDMINER ТА MATLAB

2.1 Відбір моделей для проведення дослідження

2.1.1 Аналіз моделей класифікації

Враховуючи конфіденційний характер закупівельних даних, у цьому дослідженні будемо розглядати лише моделі класифікації та прогнозування, які використовують закодовані та анонімізовані дані, що не включають комерційну інформацію, таку як дані про постачальників, ціни, виявлені аномалії та ризики тощо.

Метод Наївного Баєса – це сімейство класифікаторів, що базується на ймовірнісній класифікації з використанням теореми Баєса і високого рівня незалежності між векторами ознак. Це ефективний метод для категоризації текстових даних, наприклад, класифікації електронних листів як спаму чи не спаму, використовуючи частоту слів як вектори ознак. Незважаючи на їх простий дизайн, ці класифікатори мають широке використання в різних реальних сценаріях [13].

Теорема Байєса є умовною ймовірнісною моделлю, яка призначає клас 'C_k' вектору ознак 'x' з ймовірністю, що задається:

$$Pr(C_k | x_1, x_2, \dots, x_n) \quad (2.1)$$

Цей тип моделі буде обчислювати умовну ймовірність для кожного класу 'k' на основі вектора ознак. Це означає, що якщо вектор ознак має велику розмірність, то цей процес є нереалізованим.

$$Pr(C_k | x) = \frac{Pr(C_k)Pr(x|C_k)}{Pr(x)} \quad (2.2)$$

Оскільки знаменник не залежить від 'C', і оскільки значення вектора ознак відомі, отже, вони є константами, ми можемо сфокусуватися на чисельнику. Чисельник можна розкласти на наступний вираз за допомогою правила ланцюга:

$$Pr(C_k, x_1, x_2, \dots, x_n) = Pr(x_1, x_2, \dots, x_n, C_k) \quad (2.3)$$

$$Pr(C_k, x_1, x_2, \dots, x_n) = Pr(x_1 | x_2, \dots, x_n, C_k) Pr(x_2 | x_3, \dots, x_n, C_k) \dots Pr(x_{n-1} | x_n | C_k) \quad (2.4)$$

Отже модель може бути виражена наступним чином:

$$Pr(C_k | x_1, x_2, \dots, x_n) = Pr(C_k) \prod_{i=1}^n Pr(x_i | C_k) \quad (2.5)$$

Цей метод базується на припущенні про незалежність між ознаками, що може вважатися недоліком, оскільки в багатьох випадках реальних даних це припущення може порушуватися, призводячи до неадекватних чи ненадійних результатів [18].

Незважаючи на це, цей класифікатор є дієвим та ефективним, особливо у завданнях багатокласової класифікації.

Загальна лінійна модель (Generalized Linear Model, GLM) є розширенням лінійної моделі, яке охоплює ширший спектр завдань, таких як лінійна регресія, логістична регресія та Пуассонівська регресія. Вона враховує лінійну комбінацію вхідних ознак, використовуючи функцію зв'язку та функцію розподілу, що робить її більш гнучкою та придатною для різноманітних завдань [13].

Математично GLM визначається так:

– лінійна складова:

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (2.6)$$

де $\beta_0, \beta_1, \dots, \beta_k$ - параметри моделі;

функція зв'язку (Link Function):

$$g(\mu) = \eta \quad (2.7)$$

Функція зв'язку $g(\mu)$ пов'язує математичне очікування μ вихідної змінної з лінійною складовою η .

Функція розподілу:

Модель GLM передбачає вибір певної функції розподілу залежно від типу задачі (регресія, класифікація тощо). Наприклад:

- для лінійної регресії: Гауссівський розподіл;
- для логістичної регресії: Біноміальний розподіл;
- для Пуассонівської регресії: Пуассонівський розподіл.

Логістична регресія – це метод бінарної (або мультікласової) класифікації, де для передбачення ймовірностей належності до певного класу використовується логістична функція. Математична модель у логістичній регресії описує ймовірність того, що змінна відгуку прийме одне з двох можливих значень, зазвичай 0 або 1. Для вираження цієї ймовірності можна використовувати логістичну функцію, також відому як сигмоїд-функція [18]. Функція виглядає наступним чином:

$$P(Y=1) = \frac{1}{1+e^{-(b_0+b_1X_1+\dots+b_nX_n)}} \quad (2.8)$$

де e - число Ейлера, приблизно 2.71828, $b_0, b_1, b_2, \dots, b_n$ - коефіцієнти моделі, X_1, X_2, \dots, X_n - значення вхідних ознак.

Функція сигмоїду перетворює значення з усього діапазону дійсних чисел в інтервал між 0 і 1, приймаючи величини від усього цього діапазону. Отримана ймовірність, що Y дорівнює 1, лежить між 0 і 1.

Дерево рішень. Математична модель дерева рішень ґрунтується на ідеї вибору рішень за допомогою вузлів і гілок, які представляють умови та результати відносно набору даних. Основна концепція полягає в рекурсивному розбитті даних на підгрупи, використовуючи найінформативніші ознаки для кожного поділу [13].

Математично цю модель можна виразити як послідовність умовних правил "якщо-то, то" для кожного вузла дерева. Умови ґрунтуються на

порогових значеннях ознак і розбивають набір даних на навчальну та тестову частини: одну, що задовольняє умові, і іншу, що не задовольняє. Цей процес відтворюється для кожної зновуствореної підгрупи рекурсивно до досягнення заданої умови зупинки або максимально визначеної глибини дерева.

Математично цю модель можна виразити як систему правил, де для кожного вузла обчислюється умова на основі ознаки, і дані розбиваються відповідно до цієї умови. Це дозволяє моделі приймати рішення з приводу класифікації або регресії для нових даних.

Метод опорних векторів (Support Vector Machine, SVM) представляє собою набір навчальних моделей, які навчаються на зазначеній навчальній вибірці даних та відносяться до типу наглядованого навчання. Ця модель тренується за допомогою набору точок навчальних даних, які входять в один з двох класів бінарного класифікатора. Засновані на цьому, Машини опорних векторів реалізують навчену модель на нових точках даних, які входять в тестовий зразок, і відносять їх до одного з двох класів. Таким чином, Машини опорних векторів є бінарним класифікатором. Ключовою ідеєю є використання гіперплощини $(n-1)$ -вимірному простору для лінійної класифікації n -вимірних векторів ознак на два окремі класи [18].

Очевидно, що існує нескінченна кількість гіперплощин, які можуть розділяти точки даних (при умові їх лінійної роздільності). Ми обираємо гіперплощину або лінійний класифікатор, який має максимальний розрив між класами. Цей розрив визначається як відстань від найближчих точок даних кожного класу до гіперплощини. Об'єктивна функція визначається цією відстанню, і метою оптимізації є максимізація цього розриву.

У Машин опорних векторів є параметр регуляризації, який допомагає уникнути перенавчання даних та може адаптуватися до нелінійних залежностей в даних. Недоліком є те, що модель вивчає параметри, що залежать від значення регуляризації та вибору ядра.

Лінійний дискримінантний аналіз (Linear Discriminant Analysis, LDA) - це статистичний метод для класифікації та зменшення розмірності вхідних даних.

Основна концепція полягає в тому, щоб знайти лінійну комбінацію вхідних ознак, яка найкраще розділяє класи [21].

Математично, модель LDA визначається наступним чином:

Математичне очікування класів:

$$\mu_k = \frac{1}{N_k} \sum_{i=1}^{N_k} X_{ik} , \quad (2.9)$$

де N_k - кількість спостережень в класі k , X_{ik} - вектор ознак для спостереження i в класі k .

Матриця коваріації для кожного класу:

$$S_k = \sum_{i=1}^{N_k} (X_{ik} - \mu_k)(X_{ik} - \mu_k)^T \quad (2.10)$$

Де T - оператор транспонування.

Математичне очікування загальне:

$$\mu = \frac{1}{N} \sum_{i=1}^K n_k \mu_k \quad (2.11)$$

Де N - загальна кількість спостережень.

Матриця коваріації загальна:

$$S = \sum_{i=1}^K S_k \quad (2.12)$$

Вагові коефіцієнти для лінійної комбінації:

$$W = S^{-1}(\mu_1 - \mu_2) \quad (2.13)$$

Функція дискримінанту:

$$g(x) = W^T x \quad (2.14)$$

Класифікація відбувається за допомогою порівняння значення $g(x)$ для кожного класу.

Метою Linear Discriminant Analysis (LDA) є максимізація відношення роздільної здатності між класами та дисперсії всередині класів. Це призводить

до отримання лінійної комбінації ознак, яка оптимально розділяє класи в просторі ознак [21].

KNN (к-найближчих сусідів) - це метод класифікації, також відомий як "Ліниве навчання". Алгоритм присвоює клас точці даних, базуючись на класах її k найближчих сусідів. Він вибирає клас, який є найбільш представленим серед k сусідів точки даних на зображенні [13].

Математично, принцип KNN описується наступним чином.

Визначення відстані: Метрика відстані між точками визначається у просторі ознак, зазвичай використовуються показники, такі як евклідова відстань або косинусна відстань.

Вибір параметра k: Визначається кількість сусідів, які будуть використовуватися для прийняття рішення. Вибір цього параметра важливий і може суттєво впливати на результати KNN.

Пошук k найближчих сусідів: Для нової точки обчислюється відстань до всіх інших точок у наборі даних, і вибираються k найближчих.

Вирішення завдання класифікації або регресії: Класифікація (KNN для класифікації) - визначається клас, який є найбільш представленим серед k найближчих сусідів, і нова точка призначається до цього класу. Регресія (KNN для регресії) - обчислюється середнє або зважене середнє значення вихідних змінних залежно від значень k найближчих сусідів [18].

2.1.2 Аналіз моделей прогнозування

Linear Regression. Математична модель у моделі Linear Regression виглядає наступним чином:

- для простої лінійної регресії (з однією залежною змінною) [13]:

$$y = \beta_0 + \beta_1 X + \varepsilon \quad (2.15)$$

– для багатовимірної лінійної регресії (з кількома залежними змінними) [13]:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (2.16)$$

де β_0 - це зсув (або інтерсепт), $\beta_1, \beta_2, \dots, \beta_n$ – це коефіцієнти регресії для кожної з незалежних змінних, x_1, x_2, \dots, x_n – це значення незалежних змінних, ε - є помилкою моделі, або терміном помилки.

Основний концепт лінійної регресії це виявлення лінійного зв'язку між залежною змінною та однією чи декількома незалежними змінними. Коефіцієнти $\beta_1, \beta_2, \dots, \beta_n$ обчислюються таким чином, щоб мінімізувати суму квадратів помилок (метод найменших квадратів) [21].

Polynomial Regression. Математична модель у моделі Polynomial Regression виглядає таким чином:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_n x^n + \varepsilon \quad (2.17)$$

де β_0 – це зсув (або інтерсепт), $\beta_1, \beta_2, \dots, \beta_n$ – це коефіцієнти регресії для кожної з незалежних змінних, x – це значення незалежних змінних, ε є помилкою моделі, або терміном помилки.

У моделі Polynomial Regression ми включаємо поліноміальні члени до лінійної моделі з метою врахування нелінійних взаємозв'язків між залежною та незалежними змінними. Порядок поліному (визначений параметром n , що вказує на кількість ступенів) зазвичай обирається на основі результату аналізу, і часто використовуються поліноми другого або третього порядку для моделювання криволінійних залежностей у даних [21].

Модель "Interactions Linear" або "Linear Regression with Interaction Terms" враховує взаємодію між різними незалежними змінними в контексті лінійної регресії. Математичне вираження для лінійної регресії з взаємодією має вигляд:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \varepsilon, \quad (2.18)$$

де β_0 – це зсув (або інтерсепт), β_1 , β_2 це коефіцієнти регресії для перших двох незалежних змінних, β_{12} – це коефіцієнт взаємодії між X_1 і X_2 , X_1 і X_2 – це значення незалежних змінних, ε – є помилкою моделі, або терміном помилки.

Взаємодія в даному контексті відображає вплив взаємодії двох змінних на залежну змінну, що дозволяє враховувати можливі взаємодії та підвищити точність прогнозування у випадках, коли взаємодія між змінними є суттєвою [21].

У моделі Neural Network (нейронна мережа) математична структура визначається конфігурацією нейронної мережі та ваговими коефіцієнтами, які піддаються налаштуванню при навчанні. Головна ідея полягає в моделюванні взаємозв'язків між даними вхідними та вихідними за допомогою штучних нейронів та їх вагових коефіцієнтів [14].

Для простої одношарової нейронної мережі з одним вихідним шаром математична модель може бути виражена наступним чином:

$$y = f(\sum_{i=1}^n w_i x_i + b) \quad (2.19)$$

де f – функція активації, яка застосовується до зваженої суми входів та вагових коефіцієнтів (наприклад, сигмоїдна, гіперболічний тангенс і т.д.);

w_i – ваговий коефіцієнт для виходу x_i ;

b - зсув (інтерсепт);

n - кількість входів.

У більш складних нейронних мережах, як багатшарові персептрони, математична модель стає складнішою за рахунок введення прихованих шарів та складніших функцій активації.

Навчання нейронної мережі містить налаштування вагових коефіцієнтів для максимізації точності прогнозування на тренувальних даних. Зазвичай цей процес реалізується за допомогою методів градієнтного спуску та зворотного поширення помилок [18].

2.2 Огляд ключових елементів аналізу, підготовки та обробки даних у сфері закупівель

Використання сучасних та високоякісних інструментів аналізу є необхідною передумовою для отримання достовірних результатів. Однак це лише одна з умов успішного аналізу. Ще однією ключовою умовою є використання якісних даних. Підприємства часто накопичують за попередні періоди значні обсяги даних про закупівлі, використовуючи системи обліку бізнес-процесів.

Сучасні системи обліку активно використовують різноманітні засоби для забезпечення повноти та цілісності даних. Вони використовують правила та контролі заповнення даних під час господарських операцій і можливість прикріплення сканкопій первинних документів до їх електронних еквівалентів у системі. Більшість систем мають широкі можливості настройки цих правил і контролів. Однак, щоб спростити процеси введення документів та прискорити їх обробку, часто обмежуються обов'язковістю введення багатьох атрибутів та послаблюють контролі, що може призводити до неповного введення або введення даних з помилками. Формальний підхід до навчання та збільшення інтенсивності праці фахівців, які вводять дані, також сприяє накопиченню помилок та пропущених даних [13].

Враховуючи вищезазначені фактори, можливий набір історичних даних, який може включати помилкові, пропущені, задубльовані та некоректно внесені дані. Присутність таких викривлень у даних може унеможливити їх використання в аналізі або призводити до значних похибок у результатах та, відповідно, до прийняття неточних рішень у закупівлях та можливих збитків.

Щоб уникнути можливих помилок та прийнятих на їх основі неточних рішень, необхідно перед обробкою введених даних для аналізу провести їхню додаткову обробку.

2.2.1 Підготовка та очищення даних є ключовим етапом у процесі аналізу

Перший етап у обробці закупівельних даних – це підготовка та очищення даних, що включає кілька ключових процесів. Спочатку проводиться збір та акумуляція необхідної інформації про закупівельні процеси з багатьох джерел, як бази даних, електронні таблиці, лог-файли та інші. Під час цього етапу важливо забезпечити повноту та репрезентативність даних, враховуючи їхню різноманітність та розподілення в існуючій системі обліку організації.

На другому етапі важливо провести перевірку та обробку пропущених даних, а також виявлення та обробку викидів. Це допомагає уникнути спотворень у результаті аналізу та забезпечує надійність отриманих результатів моделей машинного навчання [25].

Третій етап передбачає нормалізацію та стандартизацію даних для забезпечення їхньої порівнюваності та однорідності. Це сприяє підвищенню якості та консистентності даних, що є важливим для точних та достовірних результатів моделювання.

Завершальний етап передбачає відбір ознак, що є ключовим для побудови надійних моделей машинного навчання. Цей процес допомагає скоротити кількість ознак до найбільш інформативних, що сприяє ефективному аналізу та прогнозуванню закупівельних даних.

Перевірка та обробка пропущених даних становлять важливий етап в підготовці закупівельних даних для подальшого аналізу. Оскільки ці дані формуються співробітниками, які беруть участь у закупівельних процесах, можливість пропущених значень виникає через те, що системи обліку не завжди вимагають заповнення всіх ознак. Один із простих методів обробки помилок та відсутніх значень - це їх видалення. Однак для більш доречного підходу варто використовувати ймовірні значення ознаки або використовувати сусідні значення (наступне або попереднє), особливо для упорядкованих даних [25].

Обробка викидів є іншим важливим кроком, оскільки вони можуть вплинути на точність моделей. Викиди - це аномальні значення, які відрізняються від інших даних. Видалення викидів, що виходять за зазначений інтервал, може поліпшити якість моделі. Це важливо враховувати, оскільки викиди часто є окремими випадками в закупівельних процесах і мають епізодичний характер.

Нормалізація та стандартизація даних стають важливими для забезпечення однакового масштабу всіх ознак, оскільки різний масштаб може негативно позначитися на результативності методів машинного навчання.

Останній етап – відбір ознак, що дозволяє знизити кількість ознак із збільшеної кількості в закупівельних даних. Це важливо для уникнення перенавчання моделі та забезпечення ефективної та компактною моделі. Налаштовані статистичні інструменти можуть використовуватись для видалення змінних, дисперсія яких нижча за певний поріг, що робить модель більш ефективною.

Обробка та очищення даних визнаються ключовими етапами перед використанням методів машинного навчання для обробки та аналізу закупівельних даних. Забезпечення якості та чистоти даних гарантує достовірність і точність отриманих результатів, а також забезпечує ефективне функціонування інтелектуальної системи керування закупівлями [13].

2.2.2 Визначення відповідного методу машинного навчання

В залежності від мети аналізу та оцінки закупівельних операцій, вибір правильного методу машинного навчання визначається характером даних та очікуваними результатами. Це є ключовим етапом у розробці системи інтелектуального управління закупівлями, оскільки коректний вибір методу

визначає надійність та точність отриманих результатів, а помилки можуть призвести до критичних наслідків при ухваленні рішень [9].

Методи машинного навчання мають свої унікальні особливості та придатність для різних видів даних та завдань. При виборі методу важливо враховувати наступні аспекти:

Тип даних: Окремі методи машинного навчання добре пристосовані до певних типів даних. Наприклад, структурованим даним можуть бути ефективні класифікаційні методи, наприклад метод опорних векторів (SVM) або навчання дерев рішень. Для даних без жорсткої структури, таких як текст чи аудіо, можуть бути більш підходящі методи глибокого навчання [9].

Кількість даних: Більше дані зазвичай сприяють кращому навчанню, але з великим об'ємом даних збільшується час обробки, відбір репрезентативних ознак та необхідність у високій обчислювальній потужності. Для завдань з обмеженим об'ємом даних будуть доцільними методи навчання з вчителем, тоді як для великих обсягів даних можуть підходити методи глибокого навчання.

Врахування цих аспектів при виборі методу машинного навчання допомагає досягти оптимальних результатів у вирішенні конкретних завдань управління закупівлями.

Складність завдання також впливає на підбірметоду машинного навчання. Так, у випадку лінійно роздільної задачі ефективними можуть бути класифікаційні методи, такі як логістична регресія. Проте, якщо завдання має багатовимірні та нелінійні залежності, регресійні моделі та глибоке навчання можуть забезпечити кращі результати [14].

Інтерпретованість результатів також важлива. Деякі методи, такі як дерева рішень чи логістична регресія, надають змогу легко інтерпретувати результати. Це дозволяє зрозуміти вплив різних ознак на прийняття рішення моделі. З іншого боку, глибоке навчання може мати труднощі з інтерпретацією, але здатним виявляти складні залежності у даних.

Обчислювальні вимоги різних методів машинного навчання також розглядаються. Наприклад, нейронні мережі можуть вимагати значних обчислювальних потужностей та графічних процесорів. Складнощі викликані тим, що традиційні алгоритми машинного навчання не завжди враховують можливість роботи в умовах Великих Даних, тому алгоритми навчання часто розраховані на дані, які повністю поміщаються в пам'ять. Це впливає на обчислювальні вимоги системи. Розвиток технологій полегшив доступ до потужних систем, але зростання різноманітності даних підвищує вимоги до швидкості реагування на зміни та прийняття рішень [14].

При проведенні прикладних досліджень у сфері аналізу та обробки закупівельних даних важливо здійснювати відбір методу машинного навчання на основі аналізу особливостей даних, поставленої задачі та наявних ресурсів. Це сприяє досягненню високої точності та ефективності в системі інтелектуального управління закупівлями. Залежно від характеру задачі, такої як кластерний аналіз, класифікація, прогнозування чи виявлення аномалій, може бути обрано різні методи машинного навчання [9].

2.2.3 Перевірка та оцінка достовірності моделей - важливий етап у процесі їх використання

Оцінка точності моделей машинного навчання та їх адекватність є важливою складовою аналізу та обробки закупівельних даних. Після створення моделей важливо провести їх перевірку на точність та ефективність. Для цього використовується метод валідації, який дозволяє оцінити роботу моделі на різних частинах даних. Уникнути перенавчання та забезпечити генералізацію моделей на нових даних дає можливість валідація, визначаючи, наскільки добре вони взаємодіють з існуючими даними та якість їх прогнозів [11].

Основні етапи валідації включають розділення даних на навчальну та тестову вибірку, навчання моделі на навчальній вибірці та оцінку її точності на тестовій вибірці. Валідація використовує різні метрики, такі як точність чи середньоквадратична помилка, для визначення ефективності моделі.

Також важливо враховувати різні види навчання в машинному навчанні, такі як навчання з вчителем, без вчителя та з підкріпленням, які використовуються в залежності від типу задачі та наявності позначених даних.

Оцінка точності включає в себе розгляд різних метрик та аспектів, таких як чутливість, специфічність та F-оцінка. Повторення процесу використовується для зниження дисперсії та отримання більш об'єктивних оцінок.

Завершивши валідацію, можна зробити висновки щодо того, яка модель найкраще відповідає конкретній задачі аналізу закупівельних даних. Валідація має свої переваги, але також може вимагати значних обчислювальних ресурсів та виявляти обмеження в якості наявних даних [18].

2.2.4 Інтерпретування результатів

Розшифрування отриманих результатів є ключовим етапом аналізу. Після проведення валідації та оцінки точності моделей машинного навчання важливо розуміти, що вони намагаються повідомити про закупівельні дані.

На цьому етапі важливо враховувати контекст завдання та природу використовуваних даних. Важливість інтерпретації полягає в тому, щоб визначити, які особливості чи параметри внесли найбільший внесок у формування результатів. Додатково, розглядання аспектів, таких як чутливість моделі до різних видів даних чи невизначеності, може забезпечити глибше розуміння, як модель взаємодіє з реальними ситуаціями.

Важливо також визначити, як вибір конкретних методів машинного навчання вплинув на трактування результатів і на які аспекти аналізу закупівельних даних вони найбільше впливають. Це може допомогти вдосконалити моделі та забезпечити їхню кращу адаптацію до конкретних вимог та особливостей аналітичної задачі.

Інтерпретація результатів аналізу закупівельних даних виступає важливим етапом, оскільки вона допомагає зрозуміти вплив факторів на закупівельні процеси та процес формування прогнозів моделей. Визначення суттєвих залежностей між змінними та їх взаємозв'язки допомагає зробити обґрунтовані висновки та прийняти ефективні рішення для оптимізації закупівельних процесів [25].

Інтерпретація включає кілька ключових аспектів

Вагові коефіцієнти: Використовуючи моделі машинного навчання, які базуються на лінійній або логістичній регресії, отримані вагові коефіцієнти для кожної змінної. Це дозволяє з'ясувати вплив кожної змінної на вихідні результати моделі.

Важливість змінних: У моделях, таких як дерева рішень та випадковий ліс, можна оцінити важливість кожної змінної для процесу прийняття рішення. Це допомагає виділити ключові фактори, що впливають на закупівельні процеси.

Візуалізація результатів: Використання графіків, діаграм та дашбордів для візуального представлення результатів. Візуалізація допомагає виявити тенденції та зв'язки в даних, що полегшує їх розуміння.

Висновки та рекомендації: В результаті інтерпретації слід формулювати висновки про ефективність методів, точність прогнозів та глибину аналізу закупівельних даних. Отримані результати можуть бути використані для розробки практичних рекомендацій з управління закупівлями, оптимізації процесів та зменшення ризиків.

Інтерпретація результатів машинного навчання є важливим кроком у дослідженні та аналізі закупівельних даних. Вона робить складні математичні

моделі зрозумілими та застосовними в реальних ситуаціях, що сприяє поліпшенню управління закупівлями та розробці успішних стратегій.

2.3 Аналіз та підготовка закупівельних даних для розв'язання управлінських задач

Зазвичай підприємства накопичують значні обсяги закупівельних даних у сховищах даних, які часто містять значну кількість надлишкових даних що частково дублюють, розширюють або є похідними від базового набору даних.

Дані, які найчастіше можуть міститися у початковому наборі закупівельних даних, надані у таблиці 2.1. Після завершення закупівлі, набір даних значно розширюється, оскільки до початкового набору додаються дані про результати закупівлі та проміжні етапи закупівлі.

Таблиця 2.1 – Закупівельні дані.

Атрибут	Назва атрибуту	Визначення	Пояснення
BU	Bussiness Unit	Структурна одиниця	Підрозділ який ініціює закупівлю
ReqDate	Request Date	Дата потреби	Дата коли потрібна наявність предмету закупівель
PRDate	Purchasing Request Date	Дата заявки на закупівлю	Дата подання заявки на закупівлю
ReqDateW	Request Date Week	Тиждень Дати потреби	Похідна від Дати потреби
PRDateW	Purchasing Request Date Week	Тиждень Дати заявки на закупівлю	Похідна від Дати заявки на закупівлю
ReqDateM	Request Date Mounth	Місяць Дати потреби	Похідна від Дати потреби
PRDateM	Purchasing Request Date Mounth	Місяць Дати заявки на закупівлю	Похідна від Дати заявки на закупівлю

Продовження таблиці 2.1.

Атрибут	Назва атрибуту	Визначення	Пояснення
ReqDateY	Request Date Year	Рік Дати потреби	Похідна від Дати потреби
PRDateY	Purchasing Request Date Year	Рік Дати заявки на закупівлю	Похідна від Дати заявки на закупівлю
Urgency	Urgency	Терміновість закупівлі	Ідентифікує важливість закупівлі
Material	Material	Назва матеріалу/послуг закупівля	Ідентифікує що саме треба закупити
MatGroup	Material Group	Група матеріалів	Визначає клас предмету закупівель, похідна від матеріалу
ProcGroup	Procurement Group	Група закупівель	Визначає специфіку закупівель
SPP	Service Parts Planning	Планування складових частин проекту	Визначає проект до якого відноситься закупівля
BudProg	Budget Programm	Програма бюджету	Визначає джерело фінансування
FinPosition	Financial Position	Фінансова позиція	Визначає категорію фінансових ресурсів
Qneeds	Quantity needed	Необхідна кількість	Визначає яка кількість повинна бути закуплена
Unit	Unit	Одиниця виміру	Ідентифікує важливість виміру
Price	Price	Ціна	Ціна одиниці предмету закупівлі
PriceUnit	Price Unit	Одиниця ціни	Визначає за який об'єм предмету закупівель встановлено ціну
SumPR	Summa Purchasing Request	Вартість заявки	Загальна вартість повного об'єму закупівлі, похідна від Ціни та кількості
Currency	Currency	Одиниця валюти	Визначає валюту розрахунку
Duration	Duration	Тривалість закупівлі	Нормативний строк закупівлі від дати заявки до дати поставки

Продовження таблиці 2.1.

Атрибут	Назва атрибуту	Визначення	Пояснення
FTEproc	FTE of process	FTE процесу	Необхідна кількість співробітників для здійснення закупівлі
FTEsupl	FTE of supplying	FTE постачання	Необхідна кількість співробітників для організації поставки
Type	Type	Тип закупівель	Визначає особливості обліку предмету закупівлі
Direction	Direction	Напрямок закупівель	Визначає специфіку процесу закупівлі
TnderType	TnderType	Тип тендеру	Визначає процедуру закупівлі

Дані Замовника як правило відображають конкретні потреби та наявні ресурси і як правило фіксуються в системі, проте дані Служби закупівель, не завжди однозначні і часто потребують використання низки припущень, що значною мірою впливає на процеси прийняття рішень керівництвом Служби закупівель та на самі результати закупівель.

Використовуючи Rapid miner, Excel та MATLAB, побудуємо кореляційну матрицю. Для цього необхідне додаткове опрацювання н'аявних історичних даних та кодування текстових полів у числові значення, оскільки коефіцієнти кореляції Пірсона і Фехнера розраховуються лише для кількісних значень. Саме для кількісних ознак найчастіше застосовують ці коефіцієнти [10].

Результати наведені у таблицях 2.2-2.4 відповідно для Rapid miner [28], Excel та MATLAB [25].

Усі інструменти показали умовно ідентичні результати, різниця складає мільйонні долі, тому можемо їх не враховувати і вважати результати ідентичними.

Аналіз кореляційних матриць показав, що частина полів мають дуже високий позитивний зв'язок і можуть бути виключені із подальшої обробки з метою зменшення обсягів оброблюваних даних.

Таблиця 2.5 – Закупівельні дані з коефіцієнтом Пірсона більше 0,7

	ReqDate	PRDate	ReqDateW	PRDateW	ReqDateM	PRDateM	ReqDateY	PRDateY	MatGroup	ProcGroup	SPP	BudProg	Qneeds	SumPR
ReqDate	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86						
PRDate	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93						
ReqDateW	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86						
PRDateW	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93						
ReqDateM	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86						
PRDateM	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93						
ReqDateY	0,93	0,91	0,93	0,91	0,93	0,91	1,00	0,84						
PRDateY	0,86	0,93	0,86	0,93	0,86	0,93	0,84	1,00						
MatGroup									1,00	0,71				
ProcGroup									0,71	1,00				
SPP											1,00	0,76		
BudProg											0,76	1,00		
Qneeds													1,00	0,99
SumPR													0,99	1,00

Можемо виключити поля PRDate, ReqDateW, PRDateW, ReqDateM, PRDateM, ReqDateY, PRDateY, ProcGroup, BudProg, FinPosition, SumPR з подальшої обробки оскільки вони є похідними від полів: ReqDate, MatGroup, SPP, Qneeds.

Отже, для подальшого аналізу достатньо використовувати наступні дані (10 з 28 початкових даних, таблиця 2.6).

Таблиця 2.2 – Кореляційна матриця RapidMiner.

	BU	ReqDate	PRDate	ReqDateW	PRDateW	ReqDateM	PRDateM	ReqDateY	PRDateY	Urgency	Material	MatGroup	ProcGroup	SPP	BudProg	FinPosition	Qneeds	Unit	Price	PriceUnit	SumPR	Currency	Duration	FTEproc	FTEsupl	Type	Direction	TnderType
BU	1,00	0,05	0,00	0,05	0,00	0,07	0,00	0,09	0,02	0,54	0,13	0,01	0,05	0,19	0,24	0,42	0,01	0,07	0,04	0,02	0,01	0,02	0,22	0,04	0,11	0,18	0,11	0,00
ReqDate	0,05	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86	0,02	0,24	0,06	0,01	0,05	0,03	0,02	0,01	0,00	0,04	0,01	0,01	0,03	0,11	0,07	0,12	0,19	0,07	0,21
PRDate	0,00	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93	0,14	0,19	0,03	0,05	0,00	0,03	0,05	0,01	0,04	0,05	0,01	0,01	0,04	0,19	0,05	0,19	0,27	0,01	0,25
ReqDateW	0,05	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86	0,02	0,24	0,06	0,01	0,05	0,03	0,02	0,01	0,00	0,04	0,01	0,01	0,03	0,11	0,07	0,12	0,19	0,07	0,21
PRDateW	0,00	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93	0,14	0,19	0,03	0,05	0,00	0,03	0,05	0,01	0,04	0,05	0,01	0,01	0,04	0,20	0,05	0,19	0,27	0,01	0,25
ReqDateM	0,07	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86	0,00	0,26	0,07	0,01	0,06	0,03	0,02	0,01	0,01	0,04	0,01	0,01	0,03	0,11	0,07	0,11	0,17	0,08	0,20
PRDateM	0,00	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93	0,14	0,19	0,03	0,05	0,00	0,03	0,04	0,01	0,04	0,05	0,01	0,01	0,04	0,19	0,05	0,18	0,27	0,01	0,25
ReqDateY	0,09	0,93	0,91	0,93	0,91	0,93	0,91	1,00	0,84	0,00	0,27	0,08	0,01	0,05	0,02	0,03	0,00	0,01	0,06	0,01	0,01	0,02	0,13	0,08	0,11	0,17	0,09	0,18
PRDateY	0,02	0,86	0,93	0,86	0,93	0,86	0,93	0,84	1,00	0,14	0,18	0,03	0,05	0,01	0,04	0,03	0,01	0,03	0,02	0,01	0,01	0,03	0,18	0,04	0,18	0,25	0,02	0,26
Urgency	0,54	0,02	0,14	0,02	0,14	0,00	0,14	0,00	0,14	1,00	0,33	0,17	0,29	0,19	0,12	0,07	0,01	0,28	0,12	0,01	0,01	0,00	0,01	0,09	0,34	0,63	0,25	0,06
Material	0,13	0,24	0,19	0,24	0,19	0,26	0,19	0,27	0,18	0,33	1,00	0,65	0,60	0,27	0,28	0,41	0,00	0,39	0,09	0,01	0,01	0,01	0,03	0,09	0,31	0,48	0,25	0,08
MatGroup	0,01	0,06	0,03	0,06	0,03	0,07	0,03	0,08	0,03	0,17	0,65	1,00	0,71	0,16	0,18	0,56	0,00	0,30	0,07	0,00	0,00	0,00	0,01	0,03	0,16	0,32	0,24	0,13
ProcGroup	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,29	0,60	0,71	1,00	0,23	0,24	0,52	0,00	0,33	0,08	0,01	0,01	0,00	0,03	0,07	0,23	0,52	0,25	0,13
SPP	0,19	0,05	0,00	0,05	0,00	0,06	0,00	0,05	0,01	0,19	0,27	0,16	0,23	1,00	0,76	0,40	0,00	0,16	0,03	0,00	0,00	0,04	0,07	0,00	0,10	0,27	0,16	0,06
BudProg	0,24	0,03	0,03	0,03	0,03	0,03	0,03	0,02	0,04	0,12	0,28	0,18	0,24	0,76	1,00	0,46	0,00	0,15	0,00	0,00	0,00	0,02	0,04	0,03	0,10	0,35	0,03	0,03
FinPosition	0,42	0,02	0,05	0,02	0,05	0,02	0,04	0,03	0,03	0,07	0,41	0,56	0,52	0,40	0,46	1,00	0,00	0,23	0,05	0,00	0,00	0,04	0,20	0,02	0,16	0,30	0,21	0,06
Qneeds	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,99	0,00	0,01	0,00	0,00	0,01	0,02	0,00
Unit	0,07	0,00	0,04	0,00	0,04	0,01	0,04	0,01	0,03	0,28	0,39	0,30	0,33	0,16	0,15	0,23	0,00	1,00	0,04	0,00	0,01	0,01	0,02	0,19	0,21	0,39	0,03	0,05
Price	0,04	0,04	0,05	0,04	0,05	0,04	0,05	0,06	0,02	0,12	0,09	0,07	0,08	0,03	0,00	0,05	0,00	0,04	1,00	0,00	0,00	0,04	0,04	0,02	0,03	0,10	0,15	0,05
PriceUnit	0,02	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,01	0,00	0,00	0,01	0,01
SumPR	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00	0,01	0,00	0,00	0,00	0,99	0,01	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,01	0,01	0,01
Currency	0,02	0,03	0,04	0,03	0,04	0,03	0,04	0,02	0,03	0,00	0,01	0,00	0,00	0,04	0,02	0,04	0,00	0,01	0,04	0,00	0,00	1,00	0,05	0,06	0,05	0,01	0,04	0,05
Duration	0,22	0,11	0,19	0,11	0,20	0,11	0,19	0,13	0,18	0,01	0,03	0,01	0,03	0,07	0,04	0,20	0,01	0,02	0,04	0,00	0,00	0,05	1,00	0,00	0,07	0,09	0,08	0,07
FTEproc	0,04	0,07	0,05	0,07	0,05	0,07	0,05	0,08	0,04	0,09	0,09	0,03	0,07	0,00	0,03	0,02	0,00	0,19	0,02	0,01	0,00	0,06	0,00	1,00	0,24	0,06	0,56	0,06
FTEsupl	0,11	0,12	0,19	0,12	0,19	0,11	0,18	0,11	0,18	0,34	0,31	0,16	0,23	0,10	0,10	0,16	0,00	0,21	0,03	0,00	0,00	0,05	0,07	0,24	1,00	0,46	0,03	0,59
Type	0,18	0,19	0,27	0,19	0,27	0,17	0,27	0,17	0,25	0,63	0,48	0,32	0,52	0,27	0,35	0,30	0,01	0,39	0,10	0,00	0,01	0,01	0,09	0,06	0,46	1,00	0,18	0,11
Direction	0,11	0,07	0,01	0,07	0,01	0,08	0,01	0,09	0,02	0,25	0,25	0,24	0,25	0,16	0,03	0,21	0,02	0,03	0,15	0,01	0,01	0,04	0,08	0,56	0,03	0,18	1,00	0,00
TnderType	0,00	0,21	0,25	0,21	0,25	0,20	0,25	0,18	0,26	0,06	0,08	0,13	0,13	0,06	0,03	0,06	0,00	0,05	0,05	0,01	0,01	0,05	0,07	0,06	0,59	0,11	0,00	1,00

Таблиця 2.3 – Кореляційна матриця Excel

	BU	ReqDate	PRDate	ReqDateW	PRDateW	ReqDateM	PRDateM	ReqDateY	PRDateY	Urgency	Material	MatGroup	ProcGroup	SPP	BudProg	FinPosition	Qneeds	Unit	Price	PriceUnit	SumPR	Currency	Duration	FTEproc	FTEsupl	Type	Direction	TnderType
BU	1,00																											
ReqDate	0,05	1,00																										
PRDate	0,00	0,94	1,00																									
ReqDateW	0,05	1,00	0,94	1,00																								
PRDateW	0,00	0,94	1,00	0,94	1,00																							
ReqDateM	0,07	1,00	0,94	1,00	0,94	1,00																						
PRDateM	0,00	0,94	1,00	0,94	1,00	0,94	1,00																					
ReqDateY	0,09	0,93	0,91	0,93	0,91	0,93	0,91	1,00																				
PRDateY	-0,02	0,86	0,93	0,86	0,93	0,86	0,93	0,84	1,00																			
Urgency	0,54	-0,02	-0,14	-0,02	-0,14	0,00	-0,14	0,00	-0,14	1,00																		
Material	0,13	0,24	0,19	0,24	0,19	0,26	0,19	0,27	0,18	0,33	1,00																	
MatGroup	0,01	0,06	0,03	0,06	0,03	0,07	0,03	0,08	0,03	0,17	0,65	1,00																
ProcGroup	0,05	-0,01	-0,05	-0,01	-0,05	0,01	-0,05	0,01	-0,05	0,29	0,60	0,71	1,00															
SPP	-0,19	0,05	0,00	0,05	0,00	0,06	0,00	0,05	-0,01	0,19	0,27	0,16	0,23	1,00														
BudProg	-0,24	0,03	-0,03	0,03	-0,03	0,03	-0,03	0,02	-0,04	0,12	0,28	0,18	0,24	0,76	1,00													
FinPosition	-0,42	-0,02	-0,05	-0,02	-0,05	-0,02	-0,04	-0,03	-0,03	0,07	0,41	0,56	0,52	0,40	0,46	1,00												
Qneeds	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,00											
Unit	0,07	0,00	-0,04	0,00	-0,04	0,01	-0,04	-0,01	-0,03	0,28	0,39	0,30	0,33	0,16	0,15	0,23	0,00	1,00										
Price	-0,04	-0,04	-0,05	-0,04	-0,05	-0,04	-0,05	-0,06	-0,02	-0,12	-0,09	-0,07	-0,08	-0,03	0,00	-0,05	0,00	-0,04	1,00									
PriceUnit	0,02	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	-0,01	0,00	-0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,00								
SumPR	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	-0,01	0,00	-0,01	0,00	0,00	0,00	0,99	-0,01	0,00	0,00	1,00							
Currency	-0,02	-0,03	-0,04	-0,03	-0,04	-0,03	-0,04	-0,02	-0,03	0,00	0,01	0,00	0,00	0,04	0,02	0,04	0,00	-0,01	0,04	0,00	0,00	1,00						
Duration	-0,22	-0,11	-0,19	-0,11	-0,20	-0,11	-0,19	-0,13	-0,18	0,01	-0,03	-0,01	0,03	0,07	0,04	0,20	-0,01	0,02	0,04	0,00	0,00	0,05	1,00					
FTEproc	-0,04	-0,07	-0,05	-0,07	-0,05	-0,07	-0,05	-0,08	-0,04	-0,09	0,09	0,03	0,07	0,00	0,03	-0,02	0,00	0,19	0,02	0,01	0,00	0,06	0,00	1,00				
FTEsupl	0,11	-0,12	-0,19	-0,12	-0,19	-0,11	-0,18	-0,11	-0,18	0,34	0,31	0,16	0,23	0,10	0,10	0,16	0,00	0,21	-0,03	0,00	0,00	0,05	0,07	0,24	1,00			
Type	0,18	-0,19	-0,27	-0,19	-0,27	-0,17	-0,27	-0,17	-0,25	0,63	0,48	0,32	0,52	0,27	0,35	0,30	0,01	0,39	-0,10	0,00	-0,01	0,01	0,09	0,06	0,46	1,00		
Direction	0,11	0,07	0,01	0,07	0,01	0,08	0,01	0,09	-0,02	0,25	0,25	0,24	0,25	0,16	0,03	0,21	-0,02	0,03	-0,15	-0,01	-0,01	-0,04	0,08	-0,56	0,03	0,18	1,00	
TnderType	0,00	0,21	0,25	0,21	0,25	0,20	0,25	0,18	0,26	-0,06	0,08	0,13	0,13	0,06	0,03	0,06	0,00	0,05	0,05	-0,01	0,01	-0,05	-0,07	-0,06	-0,59	-0,11	0,00	1,00

Таблиця 2.4 – Кореляційна матриця MATLAB

	BU	ReqDate	PRDate	ReqDateW	PRDateW	ReqDateM	PRDateM	ReqDateY	PRDateY	Urgency	Material	MatGroup	ProcGroup	SPP	BudProg	FinPosition	Qneeds	Unit	Price	PriceUnit	SumPR	Currency	Duration	FTEproc	FTEsupl	Type	Direction	TnderType
BU	1,00	0,05	0,00	0,05	0,00	0,07	0,00	0,09	-0,02	0,54	0,13	0,01	0,05	-0,19	-0,24	-0,42	0,01	0,07	-0,04	0,02	0,01	-0,02	-0,22	-0,04	0,11	0,18	0,11	0,00
ReqDate	0,05	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86	-0,02	0,24	0,06	-0,01	0,05	0,03	-0,02	0,01	0,00	-0,04	0,01	0,01	-0,03	-0,11	-0,07	-0,12	-0,19	0,07	0,21
PRDate	0,00	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93	-0,14	0,19	0,03	-0,05	0,00	-0,03	-0,05	0,01	-0,04	-0,05	0,01	0,01	-0,04	-0,19	-0,05	-0,19	-0,27	0,01	0,25
ReqDateW	0,05	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86	-0,02	0,24	0,06	-0,01	0,05	0,03	-0,02	0,01	0,00	-0,04	0,01	0,01	-0,03	-0,11	-0,07	-0,12	-0,19	0,07	0,21
PRDateW	0,00	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93	-0,14	0,19	0,03	-0,05	0,00	-0,03	-0,05	0,01	-0,04	-0,05	0,01	0,01	-0,04	-0,20	-0,05	-0,19	-0,27	0,01	0,25
ReqDateM	0,07	1,00	0,94	1,00	0,94	1,00	0,94	0,93	0,86	0,00	0,26	0,07	0,01	0,06	0,03	-0,02	0,01	0,01	-0,04	0,01	0,01	-0,03	-0,11	-0,07	-0,11	-0,17	0,08	0,20
PRDateM	0,00	0,94	1,00	0,94	1,00	0,94	1,00	0,91	0,93	-0,14	0,19	0,03	-0,05	0,00	-0,03	-0,04	0,01	-0,04	-0,05	0,01	0,01	-0,04	-0,19	-0,05	-0,18	-0,27	0,01	0,25
ReqDateY	0,09	0,93	0,91	0,93	0,91	0,93	0,91	1,00	0,84	0,00	0,27	0,08	0,01	0,05	0,02	-0,03	0,00	-0,01	-0,06	0,01	0,01	-0,02	-0,13	-0,08	-0,11	-0,17	0,09	0,18
PRDateY	-0,02	0,86	0,93	0,86	0,93	0,86	0,93	0,84	1,00	-0,14	0,18	0,03	-0,05	-0,01	-0,04	-0,03	0,01	-0,03	-0,02	0,01	0,01	-0,03	-0,18	-0,04	-0,18	-0,25	-0,02	0,26
Urgency	0,54	-0,02	-0,14	-0,02	-0,14	0,00	-0,14	0,00	-0,14	1,00	0,33	0,17	0,29	0,19	0,12	0,07	0,01	0,28	-0,12	0,01	0,01	0,00	0,01	-0,09	0,34	0,63	0,25	-0,06
Material	0,13	0,24	0,19	0,24	0,19	0,26	0,19	0,27	0,18	0,33	1,00	0,65	0,60	0,27	0,28	0,41	0,00	0,39	-0,09	-0,01	-0,01	0,01	-0,03	0,09	0,31	0,48	0,25	0,08
MatGroup	0,01	0,06	0,03	0,06	0,03	0,07	0,03	0,08	0,03	0,17	0,65	1,00	0,71	0,16	0,18	0,56	0,00	0,30	-0,07	0,00	0,00	0,00	-0,01	0,03	0,16	0,32	0,24	0,13
ProcGroup	0,05	-0,01	-0,05	-0,01	-0,05	0,01	-0,05	0,01	-0,05	0,29	0,60	0,71	1,00	0,23	0,24	0,52	0,00	0,33	-0,08	-0,01	-0,01	0,00	0,03	0,07	0,23	0,52	0,25	0,13
SPP	-0,19	0,05	0,00	0,05	0,00	0,06	0,00	0,05	-0,01	0,19	0,27	0,16	0,23	1,00	0,76	0,40	0,00	0,16	-0,03	0,00	0,00	0,04	0,07	0,00	0,10	0,27	0,16	0,06
BudProg	-0,24	0,03	-0,03	0,03	-0,03	0,03	-0,03	0,02	-0,04	0,12	0,28	0,18	0,24	0,76	1,00	0,46	0,00	0,15	0,00	0,00	0,00	0,02	0,04	0,03	0,10	0,35	0,03	0,03
FinPosition	-0,42	-0,02	-0,05	-0,02	-0,05	-0,02	-0,04	-0,03	-0,03	0,07	0,41	0,56	0,52	0,40	0,46	1,00	0,00	0,23	-0,05	0,00	0,00	0,04	0,20	-0,02	0,16	0,30	0,21	0,06
Qneeds	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,99	0,00	-0,01	0,00	0,00	0,01	-0,02	0,00
Unit	0,07	0,00	-0,04	0,00	-0,04	0,01	-0,04	-0,01	-0,03	0,28	0,39	0,30	0,33	0,16	0,15	0,23	0,00	1,00	-0,04	0,00	-0,01	-0,01	0,02	0,19	0,21	0,39	0,03	0,05
Price	-0,04	-0,04	-0,05	-0,04	-0,05	-0,04	-0,05	-0,06	-0,02	-0,12	-0,09	-0,07	-0,08	-0,03	0,00	-0,05	0,00	-0,04	1,00	0,00	0,00	0,04	0,04	0,02	-0,03	-0,10	-0,15	0,05
PriceUnit	0,02	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	-0,01	0,00	-0,01	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	-0,01	-0,01
SumPR	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	-0,01	0,00	-0,01	0,00	0,00	0,00	0,99	-0,01	0,00	0,00	1,00	0,00	0,00	0,00	0,00	-0,01	-0,01	0,01
Currency	-0,02	-0,03	-0,04	-0,03	-0,04	-0,03	-0,04	-0,02	-0,03	0,00	0,01	0,00	0,00	0,04	0,02	0,04	0,00	-0,01	0,04	0,00	0,00	1,00	0,05	0,06	0,05	0,01	-0,04	-0,05
Duration	-0,22	-0,11	-0,19	-0,11	-0,20	-0,11	-0,19	-0,13	-0,18	0,01	-0,03	-0,01	0,03	0,07	0,04	0,20	-0,01	0,02	0,04	0,00	0,00	0,05	1,00	0,00	0,07	0,09	0,08	-0,07
FTEproc	-0,04	-0,07	-0,05	-0,07	-0,05	-0,07	-0,05	-0,08	-0,04	-0,09	0,09	0,03	0,07	0,00	0,03	-0,02	0,00	0,19	0,02	0,01	0,00	0,06	0,00	1,00	0,24	0,06	-0,56	-0,06
FTEsupl	0,11	-0,12	-0,19	-0,12	-0,19	-0,11	-0,18	-0,11	-0,18	0,34	0,31	0,16	0,23	0,10	0,10	0,16	0,00	0,21	-0,03	0,00	0,00	0,05	0,07	0,24	1,00	0,46	0,03	-0,59
Type	0,18	-0,19	-0,27	-0,19	-0,27	-0,17	-0,27	-0,17	-0,25	0,63	0,48	0,32	0,52	0,27	0,35	0,30	0,01	0,39	-0,10	0,00	-0,01	0,01	0,09	0,06	0,46	1,00	0,18	-0,11
Direction	0,11	0,07	0,01	0,07	0,01	0,08	0,01	0,09	-0,02	0,25	0,25	0,24	0,25	0,16	0,03	0,21	-0,02	0,03	-0,15	-0,01	-0,01	-0,04	0,08	-0,56	0,03	0,18	1,00	0,00
TnderType	0,00	0,21	0,25	0,21	0,25	0,20	0,25	0,18	0,26	-0,06	0,08	0,13	0,13	0,06	0,03	0,06	0,00	0,05	0,05	-0,01	0,01	-0,05	-0,07	-0,06	-0,59	-0,11	0,00	1,00

Таблиця 2.6 – Закупівельні дані для подальшої обробки та аналізу

Атрибут	Позначення	Назва атрибуту	Визначення	Пояснення
BU	X ₁	Bussiness Unit	Структурна одиниця	Підрозділ який ініціює закупівлю
ReqDate	X ₂	Request Date	Дата потреби	Дата коли потрібна наявність предмету закупівель
Urgency	X ₃	Urgency	Терміновість закупівлі	Ідентифікує важливість закупівлі
Material	X ₄	Material	Назва матеріалу/послуг закупівля	Ідентифікує що саме треба закупити
MatGroup	X ₅	Material Group	Група матеріалів	Визначає клас предмету закупівель, похідна від матеріалу
SPP	X ₆	Service Parts Planning	Планування складових частин проекту	Визначає проект до якого відноситься закупівля
FinPosition	X ₇	Financial Position	Фінансова позиція	Визначає категорію фінансових ресурсів
Qneeds	X ₈	Quantity needed	Необхідна кількість	Визначає яка кількість повинна бути закуплена
Unit	X ₉	Unit	Одиниця виміру	Ідентифікує важливість виміру
Price	X ₁₀	Price	Ціна	Ціна одиниці предмету закупівлі
PriceUnit	X ₁₁	Price Unit	Одиниця ціни	Визначає за який об'єм предмету закупівель встановлено ціну
Currency	X ₁₂	Currency	Одиниця валюти	Визначає валюту розрахунку
Duration	X ₁₃	Duration	Тривалість закупівлі	Нормативний строк закупівлі від дати заявки до дати поставки

Продовження таблиці 2.6

Атрибут	Позначення	Назва атрибуту	Визначення	Пояснення
FTEproc	X ₁₄	FTE of process	FTE процесу	Необхідна кількість співробітників для здійснення закупівлі
FTEsupl	X ₁₅	FTE of supplying	FTE постачання	Необхідна кількість співробітників для організації поставки
Type	X ₁₆	Type	Тип закупівель	Визначає особливості обліку предмету закупівлі
Direction	X ₁₇	Direction	Напрямок закупівель	Визначає специфіку процесу закупівлі
TnderType	I ₁ , I ₂ , ..., I ₆	TnderType	Тип тендеру	Визначає процедуру закупівлі

2.4 Висновки за розділом

Для проведення дослідження побудовані кореляційні матриці для закупівельних даних, які було розраховано за допомогою інструментів Excel, RapidMiner та MATLAB які показали ідентичний результат. Аналіз кореляційних матриць показав, що частина даних мають високий позитивний зв'язок і можуть бути виключені із подальшої обробки з метою зменшення обсягів оброблюваних даних. Таким чином з 28 полів базових закупівельних даних, для подальшої обробки відібрано 18 полів.

РОЗДІЛ 3. ДОСЛІДЖЕННЯ МОДЕЛЕЙ КЛАСИФІКАЦІЇ ТА ПРОГНОЗУВАННЯ

3.1. Дослідження моделей класифікації

3.1.1 Дослідження моделей класифікації у середовищі RapidMiner

На початковому етапі аналізу необхідно виконати завдання Класифікації записів даних та віднесення кожного запису до конкретної категорії типу тендеру, для чого використовуються вже упорядковані, очищені та підготовлені дані відповідно до п. 2.1.1. Правильний розподіл забезпечить можливість:

- оптимального розподілу навантаження між співробітниками підрозділів закупівель та постачання;
- планування необхідних ресурсів для організації та здійснення закупівель;
- формування графіку проведення закупівель;
- адаптувати потреби бізнесу до ринкового середовища.

Для кожного типу тендеру встановлено конкретні умови, процедури, часові обмеження та правила здійснення закупівель, які можуть відрізнятися у різних підприємств і не розглядаються в рамках даної роботи.

У загальному випадку, спочатку необхідно розв'язати задачу кластеризації для об'єднання об'єктів в групи (кластери) на основі схожості ознак для об'єктів однієї групи і відмінностей між об'єктами з різних груп [11]. Але, оскільки системні ІТ компанії, як правило, мають чітко визначені вимоги до конкретних типів закупівель, то задача кластерного аналізу замінюється аналізом нормативних документів компанії, та виділенням визначених ними типів з подальшим накладенням цих типів на історичні дані із системи обліку підприємства для формування навчальної бази, яка буде виконувати роль вчителя для машинного навчання, тобто класифікації. Задача класифікації

полягає у віднесенні об'єкта з полями (x_1, x_2, \dots, x_{17}) до одного класу з множини (l_1, l_2, \dots, l_6) [11].

Для визначення найбільш ефективної моделі класифікації, проведемо навчання на основі підготовленої бази історичних даних для декількох основних методів класифікації за допомогою інструментів RapidMiner [28] та MATLAB [25].

Для дослідження будемо використовувати інструменти автомоделей для прискорення процесу.

Обробка навчальних даних у середовищі RapidMiner надали результати, які наведені у таблиці 3.1.

Таблиця 3.1 – Результати обробки у середовищі RapidMiner

Model	Accuracy	Standard Deviation	Gains	Total Time, s	Training Time (1,000 Rows)	Scoring Time (1,000 Rows)
Naive Bayes	0,674	0,0	2 914,0	401 145,0	5,1	367,5
Generalized Linear Model	0,642	0,0	2 450,0	654 876,0	62,2	244,2
Logistic Regression	0,672	0,0	2 886,0	754 973,0	34,5	724,3
Fast Large Margin	0,583	0,0	1 730,0	1 214 840,0	816,2	524,2
Deep Learning	0,693	0,0	3 146,0	642 404,0	335,1	261,4
Decision Tree	0,714	0,0	3 414,0	571 064,0	12,3	203,4
Random Forest	0,706	0,0	3 320,0	1 772 464,0	51,2	1 805,8
Gradient Boosted Trees	0,697	0,0	3 172,0	1 642 577,0	176,0	1 203,3
Support Vector Machine	0,478	0,0	230,0	2 412 118,0	1 293,3	2 952,7

За результатами обробки даних у середовищі RapidMiner най точнішою виявилася модель Decision Tree, яка досягла точності у 71,4% і прийнятний

час, 571 064,0 секунди, на навчання бази з 20 000 записів про закупівлі за останні 3 роки, що в результаті дає найбільший рівень оптимального відсічення 3 414,0. Найшвидшою є модель Naive Bayes, але її точність 67,4% та рівень оптимального відсічення 2 914,0 суттєво нижчі за модель Decision Tree (рисунки 3.1 -3.3).

Accuracy

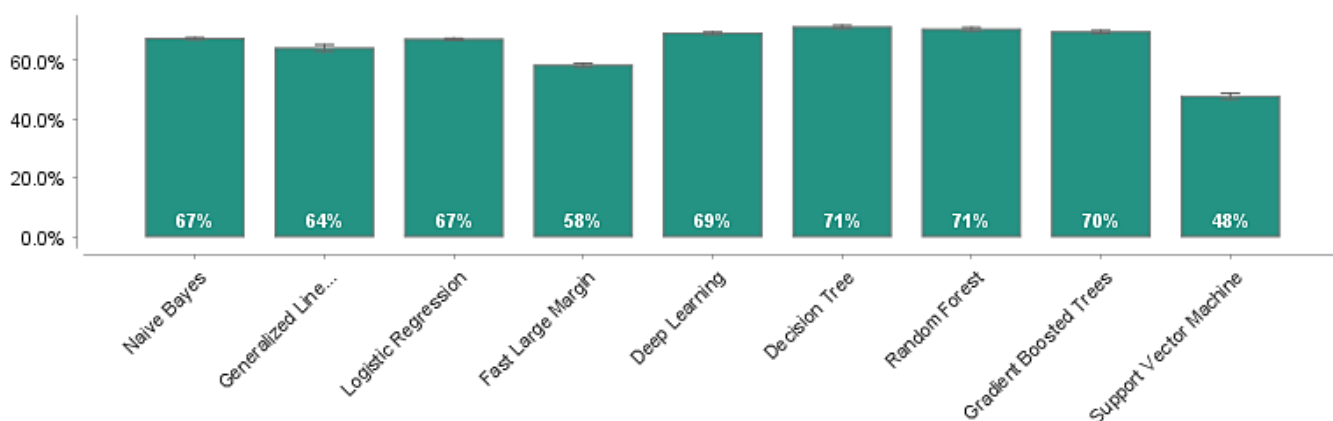


Рисунок 3.1 – Точність моделей класифікації у середовищі RapidMiner

Runtimes (ms)

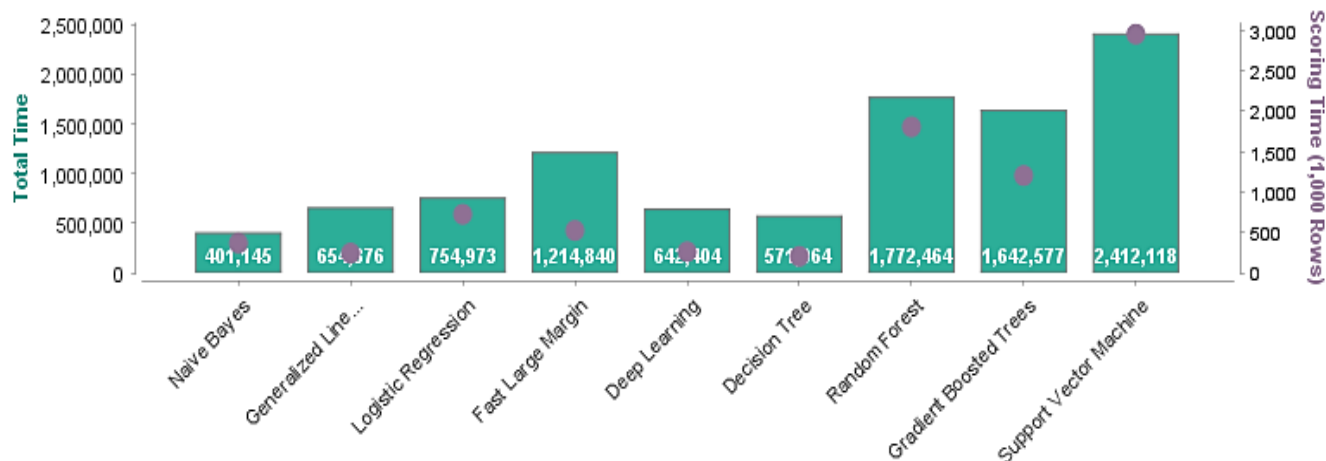


Рисунок 3.2 – Загальний час навчання моделей класифікації у середовищі RapidMiner

Model	Classification Error	Standard Deviation	Gains	Total Time	Training Time (1,000 Ro...	Scoring Time (1,000 Ro...
Naive Bayes	32.6%	± 0.3%	2,914	6 min 41 s	5 ms	367 ms
Generalized Linear Model	35.8%	± 1.1%	2,450	10 min 54 s	62 ms	244 ms
Logistic Regression	32.8%	± 0.3%	2,886	12 min 34 s	35 ms	724 ms
Fast Large Margin	41.7%	± 0.6%	1,730	20 min 14 s	816 ms	524 ms
Deep Learning	30.7%	± 0.4%	3,146	10 min 42 s	335 ms	261 ms
Decision Tree	28.6%	± 0.6%	3,414	9 min 31 s	12 ms	203 ms
Random Forest	29.4%	± 0.6%	3,320	29 min 32 s	51 ms	2 s
Gradient Boosted Trees	30.3%	± 0.5%	3,172	27 min 22 s	176 ms	1 s
Support Vector Machine	52.2%	± 1.0%	230	40 min 12 s	1 s	3 s

Рисунок 3.3 – Помилки класифікації моделей у середовищі RapidMiner

Виходячи з результатів отриманих у середовищі RapidMiner, можемо рекомендувати для виконання задач класифікації закупівельних даних використання моделі Decision Tree.

Найближчим за точністю до моделі Decision Tree є модель Random Forest, але витрати часу на її навчання та тестування більше ніж в три рази перевищує час моделі Decision Tree, що є суттєвим недоліком при прийнятті бізнес рішень.

3.1.2 Дослідження моделей класифікації у середовищі MATLAB

Середовище MATLAB надає значно ширший набір моделей класифікації, близько 30 моделей. Серед них базові моделі такі як Naive Bayes, Logistic Regression, Decision Tree, Neural Network, KNN, SVM та їх варіації [26].

Обробка навчальних даних у середовищі MATLAB надали такі результати (таблиця 3.2).

Таблиця 3.2 – Результати обробки у середовищі MATLAB

Model	Accuracy	Total cost	Prediction speed, obs/sec	Training Time	Model size (Compact), Mb
Naive Bayes	0,498	10 013,000	~110	741,3	~9
Linear Discriminant	0,679	6 404,000	~97 000	15,0	~11
Efficient Logistic Regression	0,463	10 833,000	~50 000	18,6	~0,196
Ensemble Bagged Tree	0,945	1 164,000	~15 000	1 625,9	~13
KNN	0,899	2 248,000	~6200	1 198,4	~3
Decision Tree	0,826	3 465,000	~69 000	15,1	~0,041
Neural Network	0,868	3 594,000	~140 000	2 162,9	~0,011
Boosted Trees	0,776	4 464,000	~25 000	1 570,5	~0,384
Support Vector Machine	0,698	6 030,000	~45 000	4 155,2	~0,095

Як бачимо, у середовищі MATLAB найбільшу точність 94,5% показала модель Ensemble Bagged Tree, яка створює кілька дерев рішень та об'єднує їх результати для отримання більш стабільного та точного прогнозу, та забезпечує прийнятний час на обробку даних 1 625,9 (рисунки 3.4, 3.5) [25].

Найшвидшою виявилась модель Linear Discriminant, оскільки використовує найпростіший алгоритм розрахунку, але забезпечує значно нижчу точність, лише 67,9% , проте забезпечує найменший час на обробку даних 15,0 (рисунки 3.6, 3.7) .

Також, порівняною по швидкості, але значно вищою за точністю є модель Decision Tree яка забезпечує значно вищу точність 82,6% , та забезпечує майже найменший час на обробку даних 15,1 (рисунки 3.8, 3.9).

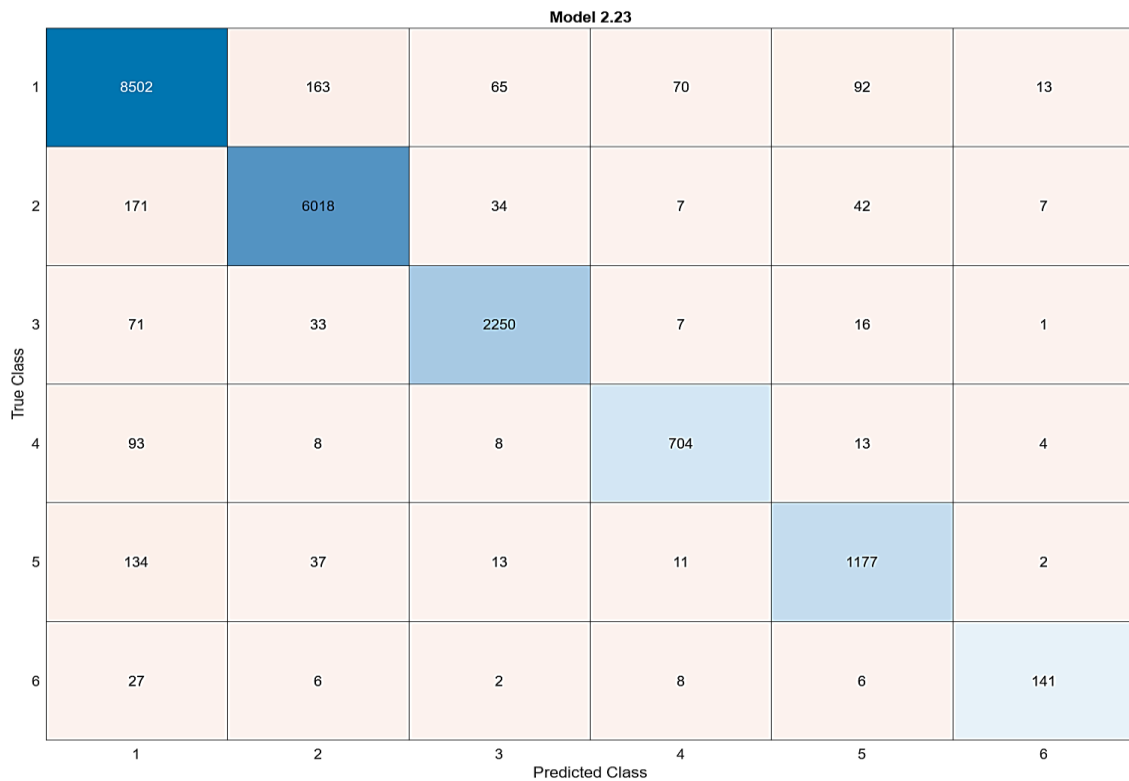


Рисунок 3.4 – Confusion Matrix (матриця плутанини) моделі Ensemble Bagged Tree у середовищі MATLAB

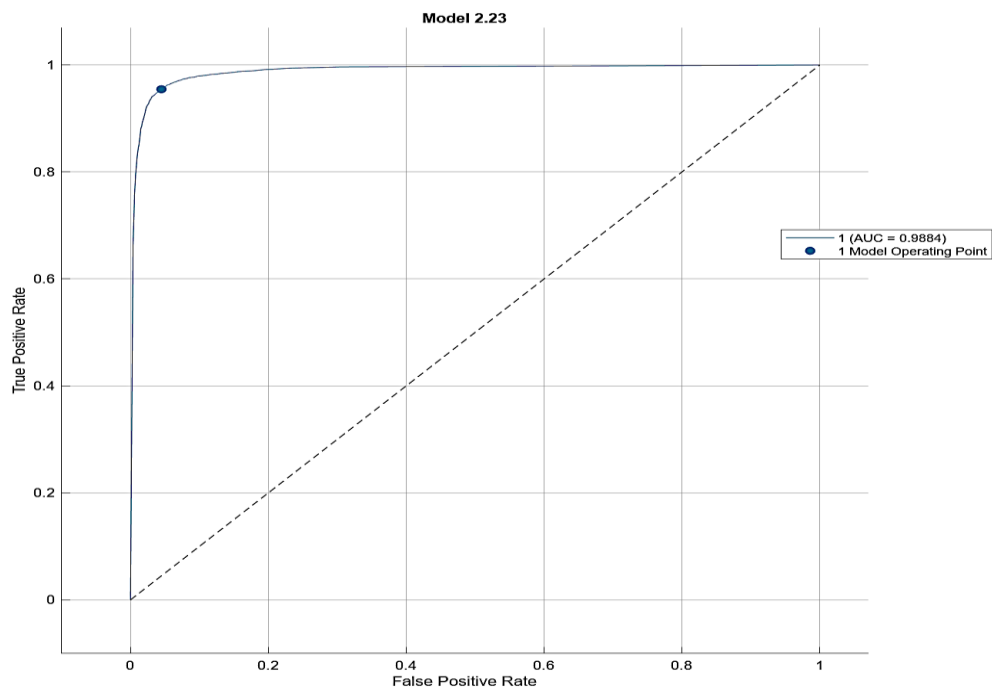


Рисунок 3.5 – Validation ROC Curve моделі Ensemble Bagged Tree у середовищі MATLAB

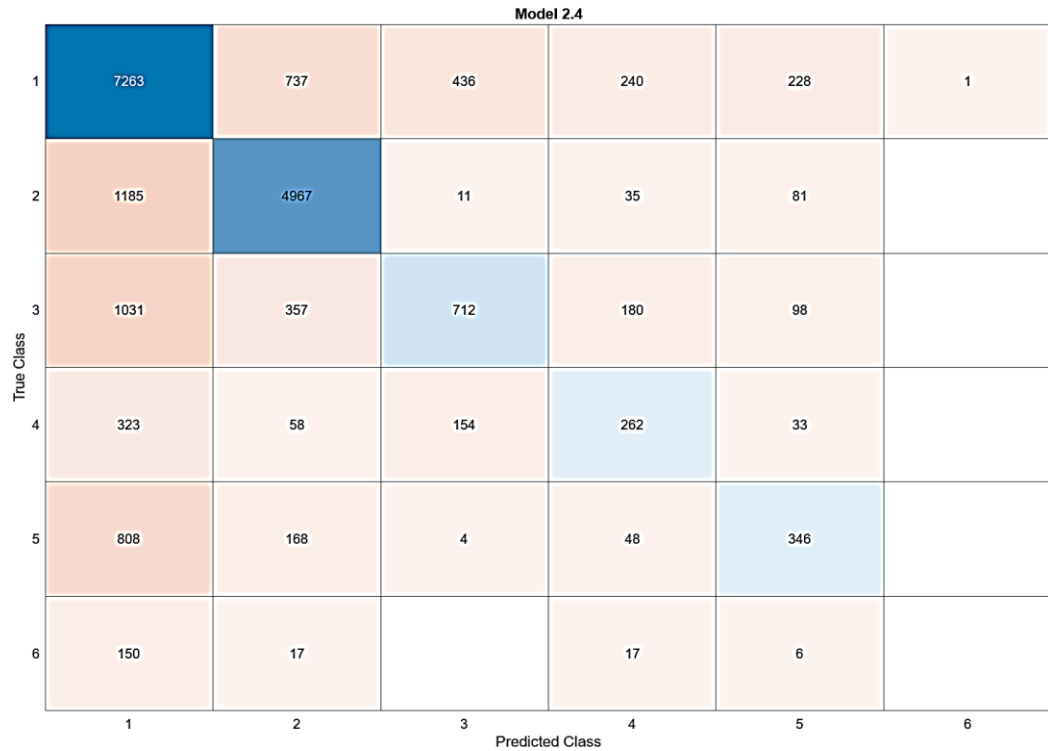


Рисунок 3.6 – Confusion Matrix (матриця плутанини) моделі Linear Discriminant у середовищі MATLAB

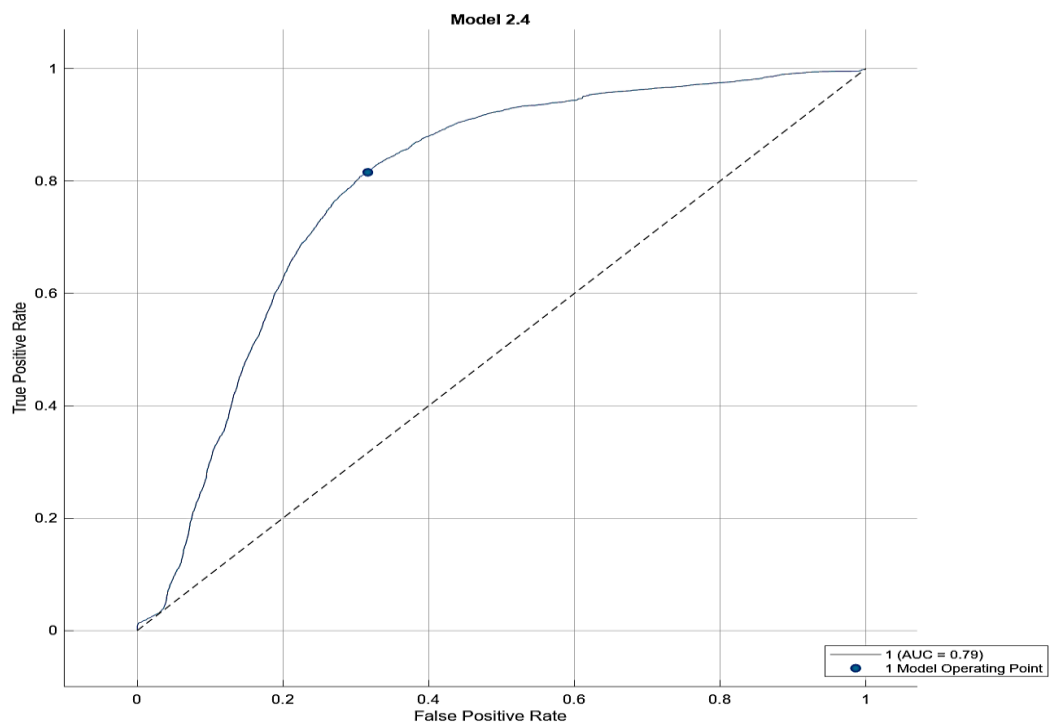


Рисунок 3.7. Validation ROC Curve моделі Linear Discriminant у середовищі MATLAB

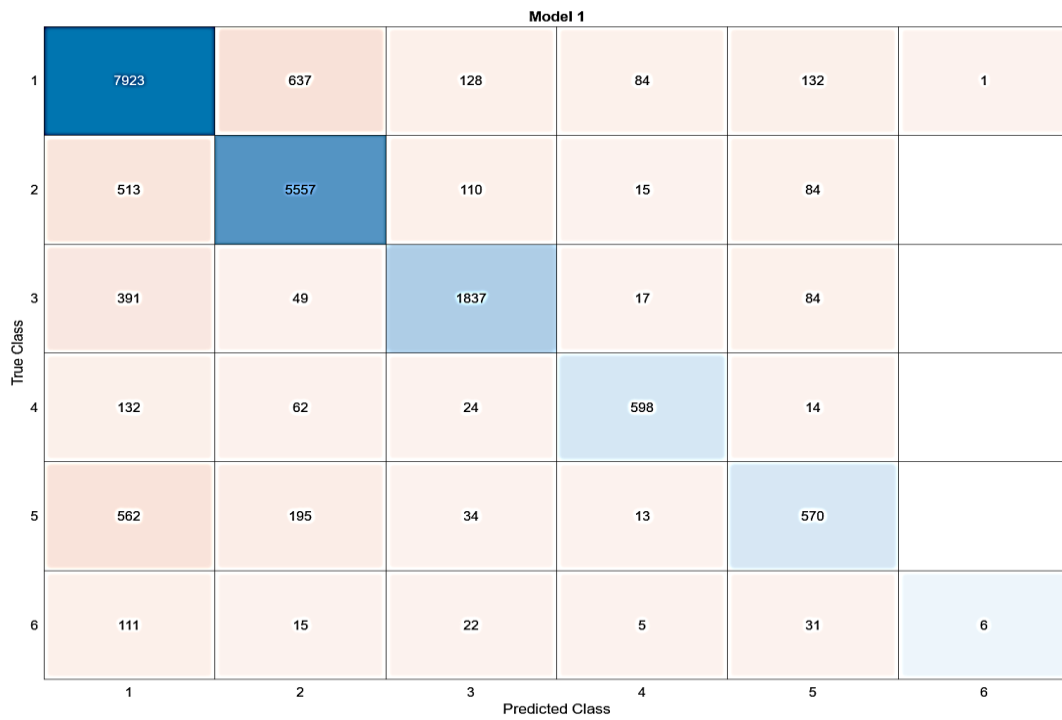


Рисунок 3.8 – Confusion Matrix (матриця плутанини) моделі Decision Tree у середовищі MATLAB

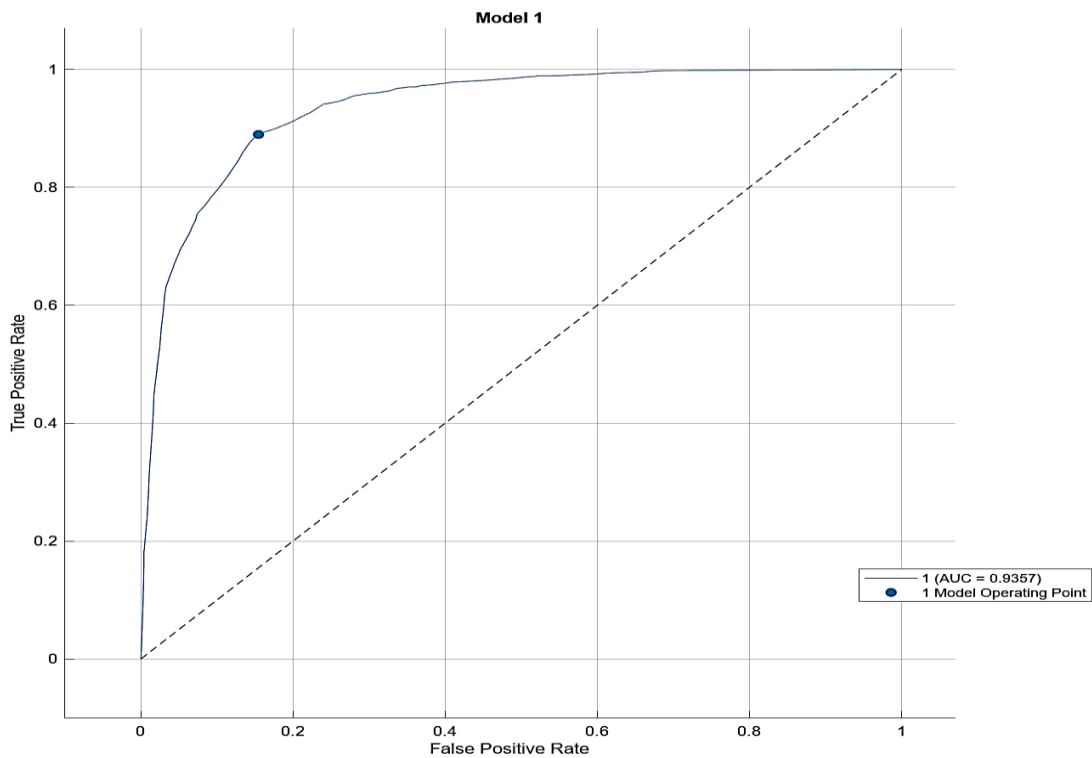


Рисунок 3.9 – Validation ROC Curve моделі Decision Tree у середовищі MATLAB

Отже, враховуючи отримані результати, можемо зробити висновки, що найефективнішою моделлю класифікації закупівельних даних є:

- для прийняття стратегічних рішень є модель Ensemble Bagged Tree;
- для прийняття операційних рішень є модель Decision Tree.
- середовище MATLAB дозволяє згенерувати код моделі

автоматично:

- Ensemble Bagged Tree (Додаток Б);
- Decision Tree (Додаток В).

3.2. Дослідження моделей прогнозування

3.2.1 Дослідження моделей прогнозування у середовищі RapidMiner

Регресія є поширеною задачею в машинному навчанні, суть якої в тому, щоб використовуючи наявні дані прогнозувати можливі значення у майбутньому, як найточніших значень у вигляді дійсних чисел для розглянутих даних.

Для прогнозування закупівельних даних, необхідна додаткова обробка наявних історичних даних. Зазвичай, прогнозування здійснюється на певні часові інтервали, тому наявні закупівельні дані попередніх періодів необхідно згрупувати за часовими проміжками. Якщо для задач класифікації ми відмовились від надлишкових даних, що виокремлюють часові проміжки, то для задач прогнозування нам необхідно повернутися до них та відмовитися від частини даних, які були необхідні для здійснення класифікації [9].

В нашому наборі мають залишитися дані з атрибутами за якими буде здійснено групування та ті, в розрізі яких нам потрібно прогнозування за періодами та агреговані дані атрибутів, значення яких необхідно спрогнозувати (таблиця 3.3).

Таблиця 3.3 – Закупівельні дані для прогнозування

Атрибут	Назва атрибуту	Визначення	Пояснення
BU	Bussiness Unit	Структурна одиниця	Підрозділ який ініціює закупівлю
ReqDateM	Request Date Mounth	Місяць Дати потреби	Похідна від Дати потреби
ReqDateY	Request Date Year	Рік Дати потреби	Похідна від Дати потреби
Urgency	Urgency	Терміновість закупівлі	Ідентифікує важливість закупівлі
MatGroup	Material Group	Група матеріалів	Визначає клас предмету закупівель, похідна від матеріалу
TnderType	TnderType	Тип тендеру	Визначає процедуру закупівлі
AvDuration	Average Duration	Середня тривалість закупівлі	Середній строк закупівлі від дати заявки до дати поставки
SumFTEproc	FTE of procurement summ	Потреба FTE закупівель	Необхідна кількість співробітників для здійснення закупівлі
SumFTEsupl	FTE of supplying summ	Потреба FTE постачання	Необхідна кількість співробітників для організації поставки
PRsQnt	Quantity of PRs	Кількість заявок на закупівлю	Визначає обсяги закупівель та постачання у часовий період

Розглянемо які засоби нам надає середовище RapidMiner. Насамперед RapidMiner надає зручний візуальний інструмент моделювання з достатньою кількістю операторів та змінюваних параметрів для побудови різноманітних моделей прогнозування та класифікації [28].

Побудуємо регресійну модель процесу Linear Regression та розглянемо її результати (рисунок 3.10) [28].

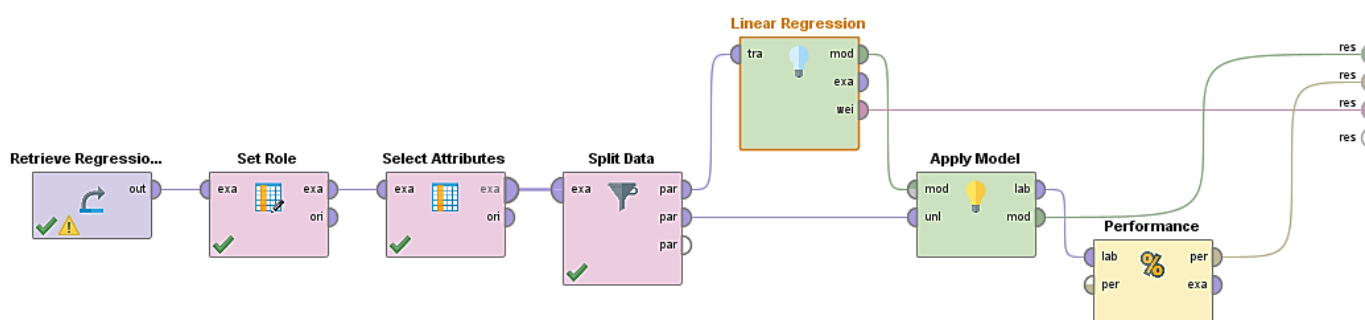


Рисунок 3.10 – Модель процесу Linear Regression у середовищі RapidMiner

У якості джерела даних для лінійної регресії використовуємо таблицю історичних закупівельних даних (таблиця 3.4).

Таблиця 3.4 – Закупівельні дані для дослідження лінійної регресії

Атрибут	Назва атрибуту	Визначення	Пояснення
ReqDateM	Request Date Mounth	Місяць Дати потреби	Похідна від Дати потреби
ReqDateY	Request Date Year	Рік Дати потреби	Похідна від Дати потреби
TnderType	TnderType	Тип тендеру	Визначає процедуру закупівлі
AvDuration	Average Duration	Середня тривалість закупівлі	Середній строк закупівлі від дати заявки до дати поставки
SumFTEproc	FTE of procurement summ	Потреба FTE закупівель	Необхідна кількість співробітників для здійснення закупівлі
SumFTEsupl	FTE of supplying summ	Потреба FTE постачання	Необхідна кількість співробітників для організації поставки
PRsQnt	Quantity of PRs	Кількість заявок на закупівлю	Визначає обсяги закупівель та постачання у часовий період

У таблиці 3.4 поля PRsQnt, AvDuration, SumFTEproc, SumFTEsupl будемо використовувати у якості міток для навчання моделі. Запустивши процес отримаємо дані, наведені у таблицях 3.5 та 3.6.

Таблиця 3.5 – Результати обробки закупівельних даних вектору продуктивності моделі Linear Regression

		root mean square d error (RMSE)	absolute error (MAE)	relative error	root relative square d error	square d error (MSE)	Correlation
Linear Regression	Criterion	25,8	17,2	701%	1,0	668,2	0,2
	+/-	0,0	19,3	683%	NaN	2627,2	NaN

Таблиця 3.6 – Результати обробки закупівельних даних моделі Linear Regression

Attribute	Weight	Coefficient	Std.Error	Std.Coefficient	Tolerance	p-Value
TnderType	-2,9	-2,9	0,5	-0,2	1,0	0,00
ReqDateM	-0,3	-0,3	0,1	-0,1	1,0	0,01
(Intercept)	NaN	34,4	4,5	NaN	NaN	0,00

У результаті обробки RapidMiner згенерував рівняння для Linear Regression:

$$PRsQnt = - 2.934 * TnderType - 0.253 * ReqDateM + 34.443 \quad (3.1)$$

Висновок: модель лінійної регресії має помірні значення RMSE та MAE. Високий відносний та кореневий відносний показники пояснюють велику різницю між спостережуваними та прогнозованими значеннями.

Побудуємо регресійну модель процесу Generalized Linear Model GLM (рисунок 3.11) та розглянемо її результати (таблиці 3.7, 3.8) [28].

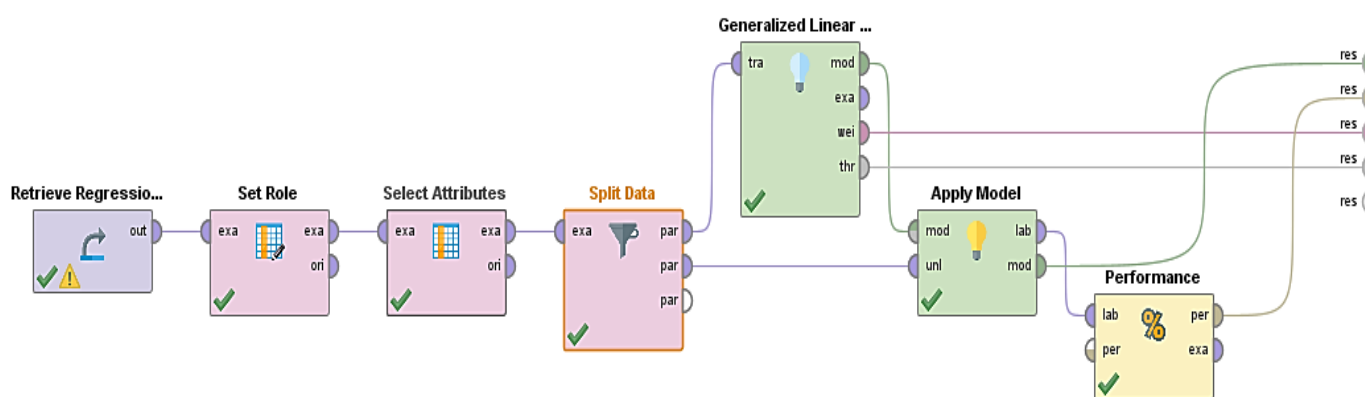


Рисунок 3.11 – Модель процесу Generalized Linear Model у середовищі RapidMiner

Таблиця 3.7 – Результати обробки закупівельних даних Вектору продуктивності Generalized Linear Model

		root mean square d error (RMSE)	absolut e error (MAE)	relativ e error	root relative square d error	square d error (MSE)	Correla -tion
Generalize d Linear Model	Criteria n	31,3	17,4	536%	1,0	976,8	0,2
	+/-	0,0	50,9	593%	n/a	4574,2	n/a

Таблиця 3.8 – Результати обробки закупівельних даних Generalized Linear Model

Attribute	Weight	Coefficient	Std.Error	Std.Coeficient	Tolerance	p-Value
TnderType	-4,8	-2,9	NaN	-4,8	NaN	NaN
ReqDateM	-2,2	-0,3	NaN	-2,2	NaN	NaN
(Intercept)	NaN	34,4	NaN	15,5	NaN	NaN

У результаті обробки RapidMiner згенерував рівняння для Generalized Linear Model:

$$PRsQnt = - 4.818 * TnderType - 0.225 * ReqDateM + 34.385 \quad (3.2)$$

Висновок: модель GLM показує більше помилок (вищі значення RMSE та MAE) порівняно з лінійною регресією та має більше високих значень відносних помилок.

Побудуємо регресійну модель процесу Polynomial Regression (рисунок 14) та розглянемо її результати (таблиця 3.9) [28].

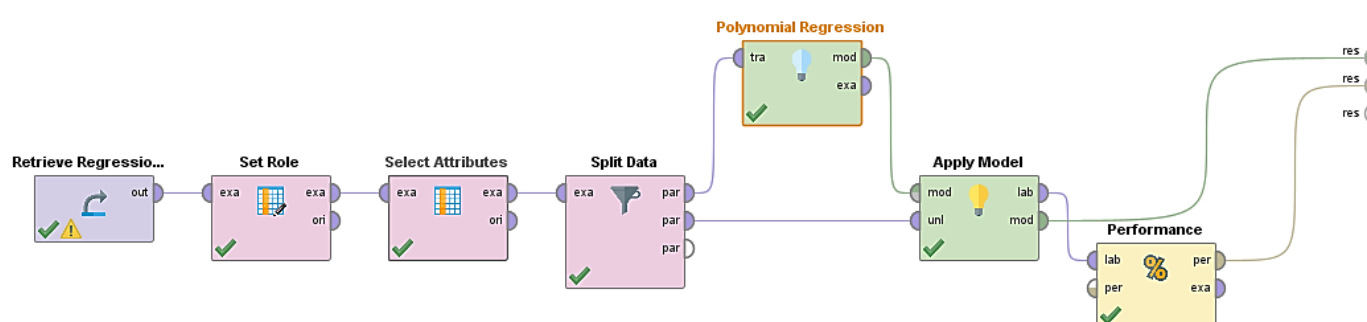


Рисунок 3.12 – Модель процесу Polynomial Regression у середовищі RapidMiner

Таблиця 3.9 – Результати обробки закупівельних даних Вектору продуктивності Polynomial Regression

		root mean squared error (RMSE)	absolute error (MAE)	relative error	root relative squared error	squared error (MSE)	Correlation
Polynomial Regression	Criterion	242,1	184,1	10598 %	9,2	58610,3	0,1
	+/-	0,0	157,2	14461 %	NaN	94521,5	NaN

У результаті обробки RapidMiner згенерував рівняння для Polynomial Regression:

$$PRsQnt = - 24.022 * TnderType ^ 2.000 + 0.002 * ReqDateM ^ 3.000 - 30.861 \quad (3.3)$$

Висновок: поліноміальна регресія має дуже великі значення помилок та відносних помилок, що свідчить про низьку точність моделі.

Побудуємо модель процесу Decision Tree (рисунок 3.13) та розглянемо її результати (таблиця 3.10) [28].

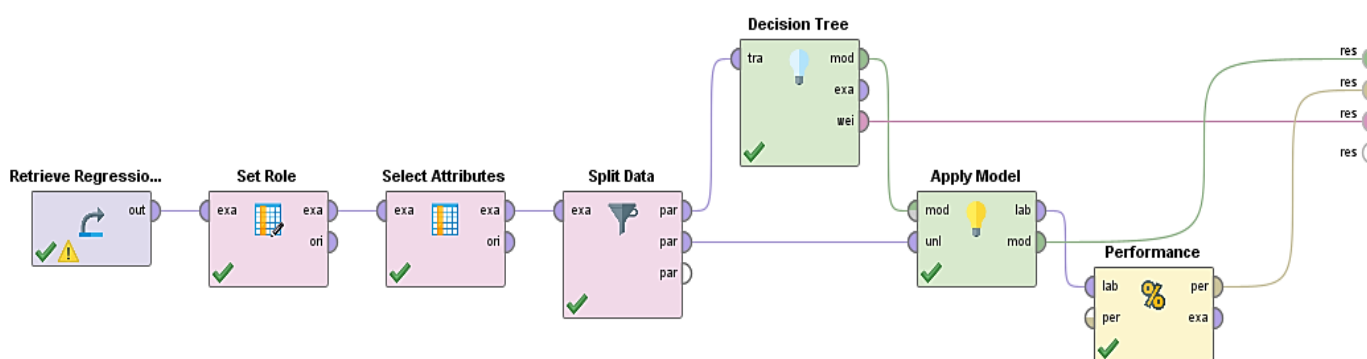


Рисунок 3.13 – Модель процесу Decision Tree у середовищі RapidMiner

Таблиця 3.10 – Результати обробки закупівельних даних Вектору продуктивності Decision Tree

		root mean square d error (RMSE)	absolut e error (MAE)	relativ e error	root relative square d error	square d error (MSE)	Correla -tion
Polynomial Regressio n	Criteria n	26,0	17,3	693%	1,0	676,4	1,0
	+/-	0,0	19,4	695%	NaN	2628,9	NaN

У результаті обробки RapidMiner згенерував дерево рішень для моделі RegressionTree:

```
TnderType > 3.500
| TnderType > 5.500
| | ReqDateM > 44.500
| | | ReqDateM > 45.500
```

```

| | | | ReqDateM > 47.500
| | | | | ReqDateM > 50.500
| | | | | | ReqDateM > 51.500
| | | | | | | ReqDateM > 56.500: 1.000 {count=2}
| | | | | | | ReqDateM ≤ 56.500
| | | | | | | | ReqDateM > 54.500: 3.000 {count=2}
| | | | | | | | ReqDateM ≤ 54.500: 1.750 {count=8}
| | | | | | | ReqDateM ≤ 51.500: 1.000 {count=2}
| | | | | | ReqDateM ≤ 50.500: 3.833 {count=12}
| | | | | ReqDateM ≤ 47.500
| | | | | | ReqDateM > 46.500: 1.250 {count=4}
| | | | | | ReqDateM ≤ 46.500: 1.000 {count=3}
| | | | | ReqDateM ≤ 45.500: 14.400 {count=5}
| | | | ReqDateM ≤ 44.500
| | | | ReqDateM > 37.500
| | | | | ReqDateM > 41.500
| | | | | | ReqDateM > 42.500
| | | | | | | ReqDateM > 43.500: 1.000 {count=2}
| | | | | | | ReqDateM ≤ 43.500: 1.667 {count=3}
| | | | | | | ReqDateM ≤ 42.500: 2.500 {count=2}
| | | | | | ReqDateM ≤ 41.500
| | | | | | | ReqDateM > 39.500
| | | | | | | | ReqDateM > 40.500: 1.250 {count=4}
| | | | | | | | ReqDateM ≤ 40.500: 1.000 {count=3}
| | | | | | | ReqDateM ≤ 39.500: 1.444 {count=9}
| | | | | | ReqDateM ≤ 37.500
| | | | | | | ReqDateM > 35.500: 2.500 {count=4}
| | | | | | | ReqDateM ≤ 35.500: 1.000 {count=2}
| | | | | | ReqDateM ≤ 33.500: 1.000 {count=2}
| | | | | ReqDateM ≤ 31.500: 1.000 {count=2}
| | | | | | ReqDateM ≤ 29.500: 1.000 {count=2}
| | | | | | | ReqDateM ≤ 27.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 25.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 23.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 21.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 19.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 17.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 15.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 13.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 11.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 9.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 7.500: 1.000 {count=2}
| | | | | | | | ReqDateM ≤ 5.500: 6.977 {count=344}
| | | | | | | | ReqDateM ≤ 3.500: 19.614 {count=904}

```

Висновок: модель дерева рішень має схожі показники з лінійною регресією, але високі значення відносних помилок також вказують на значну різницю між спостережуваними та прогнозованими значеннями.

Побудуємо модель процесу Neural Net (рисунок 3.14) та розглянемо її результати (таблиця 3.11) [28].

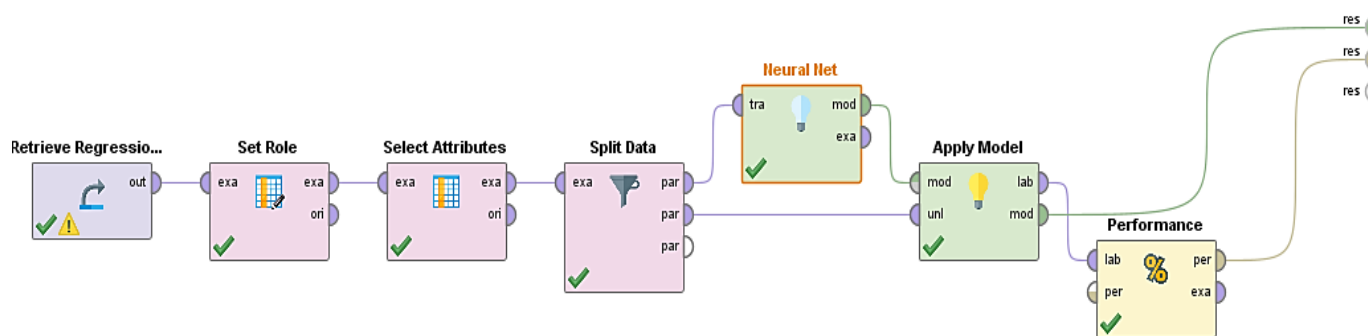


Рисунок 3.14 – Модель процесу Neural Net у середовищі RapidMiner

Таблиця 3.11 – Результати обробки Вектору продуктивності Neural Net

		root mean square d error (RMSE)	absolut e error (MAE)	relativ e error	root relative square d error	square d error (MSE)	Correla -tion
Polynomial Regression	Criterion	25,7	13,8	420%	1,0	662,9	0,2
	+/-	0,0	21,8	447%	NaN	2913,6	NaN

В результаті обробки RapidMiner згенерував нейронну мережу прогнозування для моделі Neural Net (рисунок 3.15, таблиця 3.12).

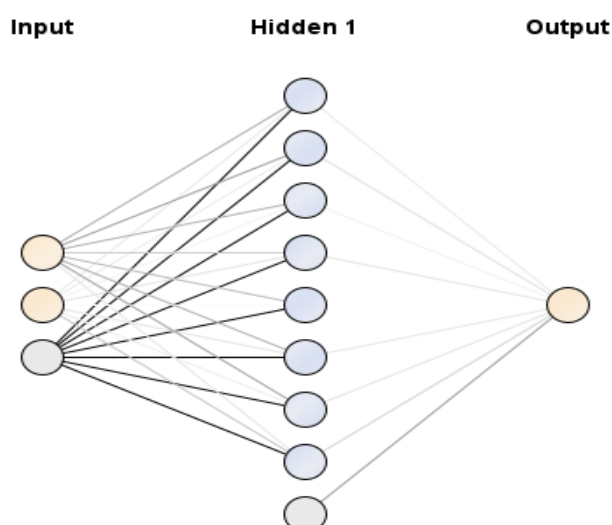


Рисунок 3.15 – Модель нейронної мережі Neural Net у середовищі RapidMiner

Таблиця 3.12 – Параметри нейронної мережі Neural Net

	Hidden 1 (Sigmoid)							
	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8
TnderType:	0,69	0,82	0,67	0,45	0,58	0,68	0,80	0,20
ReqDateM:	0,17	0,14	0,09	-0,23	0,03	0,22	0,15	0,48
Bias:	-2,12	-2,10	-2,12	-2,08	-2,13	-2,13	-2,10	-2,16
Output	-0,17	-0,25	-0,10	0,22	0,00	-0,20	-0,23	-0,36
Threshold:	-0,83							

Висновок: нейронна мережа має найменші значення RMSE та MAE серед усіх моделей, що свідчить про найкращу точність прогнозів. Однак відносна помилка є високою.

Зведемо отримані дані в одну таблицю та порівняємо отримані результати (таблиця 3.13).

Таблиця 3.13 – Зведені дані прогнозування в середовищі RapidMiner

	root mean squared error (RMSE)	absolute error (MAE)	relative error	root relative squared error	squared error (MSE)	Correlation
Linear Regression	25,8	17,2	701%	1,0	668,2	0,24
Generalized Linear Model	31,3	17,4	536%	1,0	976,8	0,18
Polynomial Regression	242,1	184,1	10598%	9,2	58610,3	0,11
Decision Tree	26,0	17,3	693%	1,0	676,4	0,21
Neural Net	25,7	13,8	420%	1,0	662,9	0,23

Враховуючи усі ці параметри, слід врахувати конкретні вимоги задачі аналізу закупівельних даних та важливість різних аспектів (наприклад, точності, відносних помилок тощо) при виборі оптимальної моделі.

Виходячи з отриманих даних, найефективнішою моделлю у середовищі RapidMiner для прогнозування закупівельних даних є нейронна мережа Neural Net або регресійна модель Linear Regression, показники яких дуже близькі.

3.2.2 Дослідження моделей прогнозування у середовищі MATLAB

Розглянемо які можливості надають нам середовище MATLAB. Для дослідження будемо використовувати ті самі дані, що були використані при дослідженнях у середовищі RapidMiner (таблиця 3.14).

Використаємо стандартну модель Linear Regression у середовищі MATLAB та розглянемо її результати (рисунки 3.16, 3.17) [26].

Таблиця 3.14 – Результати обробки закупівельних даних моделлю Linear Regression

Linear Regression	RMSE	MSE	R-Squared	MAE
Validation	31,028	962,77	0,028651	17,402
Test	28,678	822,44	0,06	16,14

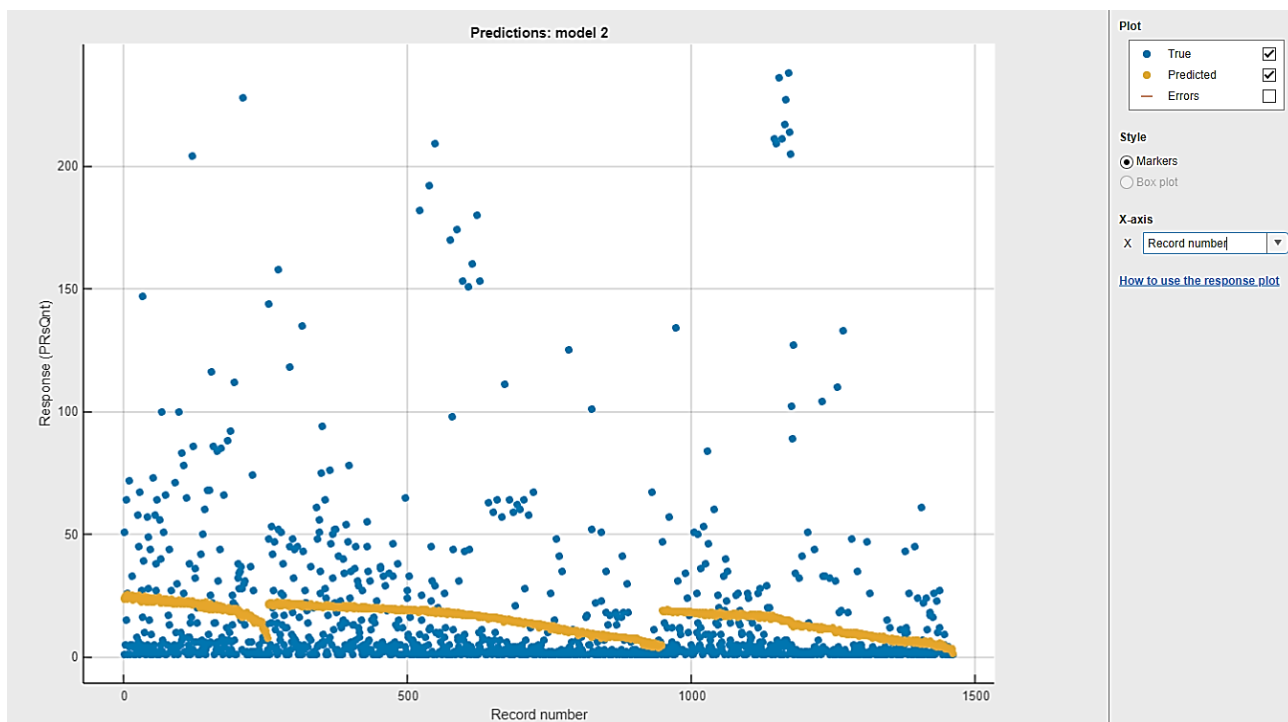


Рисунок 3.16 – Порівняння кількості фактичних і прогнозних даних моделі Linear Regression

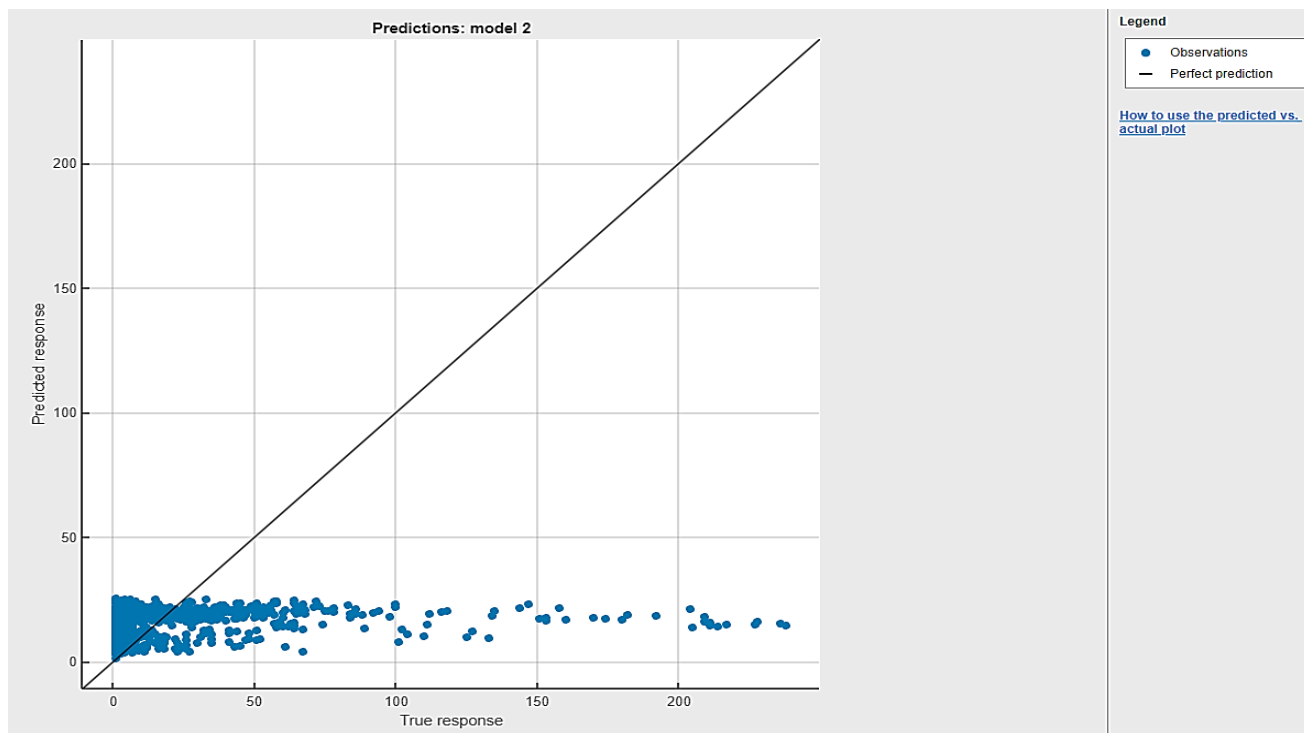


Рисунок 3.17 – Відповідність фактичних і прогнозних даних моделі Linear Regression

Висновок: модель лінійної регресії має помірно великі значення RMSE та MSE, а R-Squared показує низьку властивість пояснення варіації вихідної змінної. MAE також має помірне значення, що вказує на середню абсолютну точність прогнозів.

Використаємо стандартну модель Lnteractions Linear у середовищі MATLAB та розглянемо її результати (таблиця 3.15, рисунки 3.18, 3.19) [26].

Таблиця 3.15 – Результати обробки закупівельних даних моделлю Lnteractions Linear

Lnteractions Linear	RMSE	MSE	R-Squared	MAE
Validation	31,02	962,26	0,029162	17,351
Test	28,69	823,09	0,06	16,094

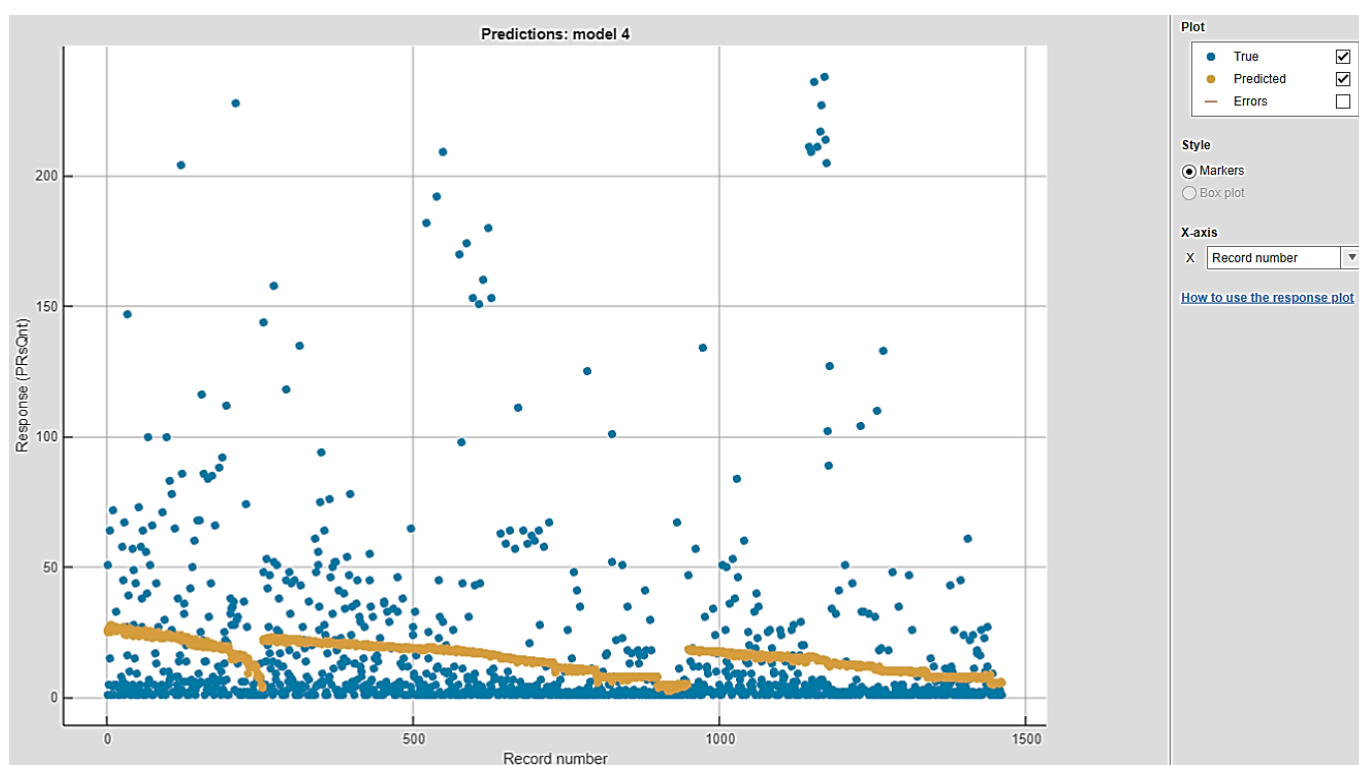


Рисунок 3.18 – Порівняння кількості фактичних і прогнозних даних моделі Lnteractions Linear

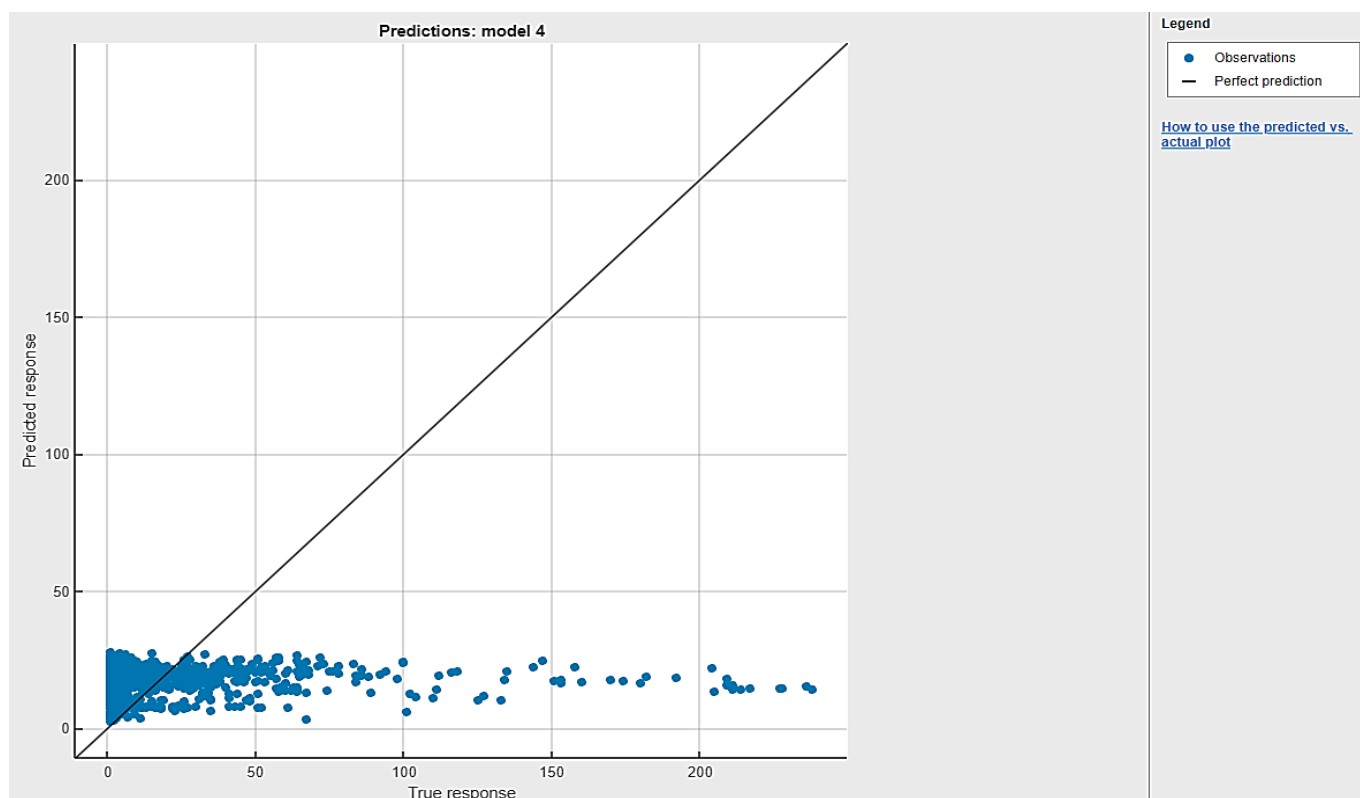


Рисунок 3.19 – Відповідність фактичних і прогнозних даних моделі
Lnteractions Linear

Висновок: модель лінійної регресії з інтеракціями має схожі показники з моделлю лінійної регресії, і вони обидві демонструють помірний рівень точності.

Використаємо стандартну модель Tree у середовищі MATLAB та розглянемо її результати (таблиця 3.16, рисунки 3.20, 3.21) [26].

Таблиця 3.16 – Результати обробки закупівельних даних моделлю Tree

Tree	RMSE	MSE	R-Squared	MAE
Validation	32,421	1051,1	-0,060504	18,555
Test	32,535	1058,5	-0,21	17,463

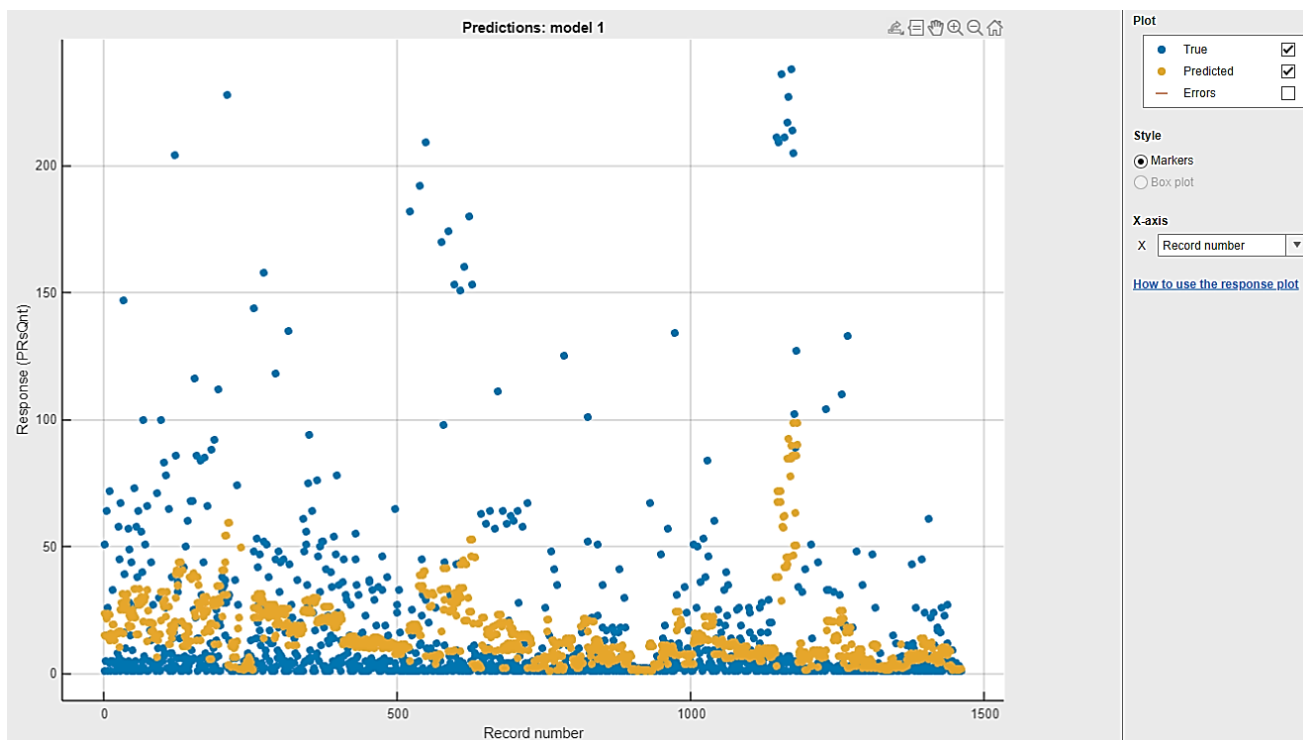


Рисунок 3.20 – Порівняння кількості фактичних і прогнозних даних моделі Tree

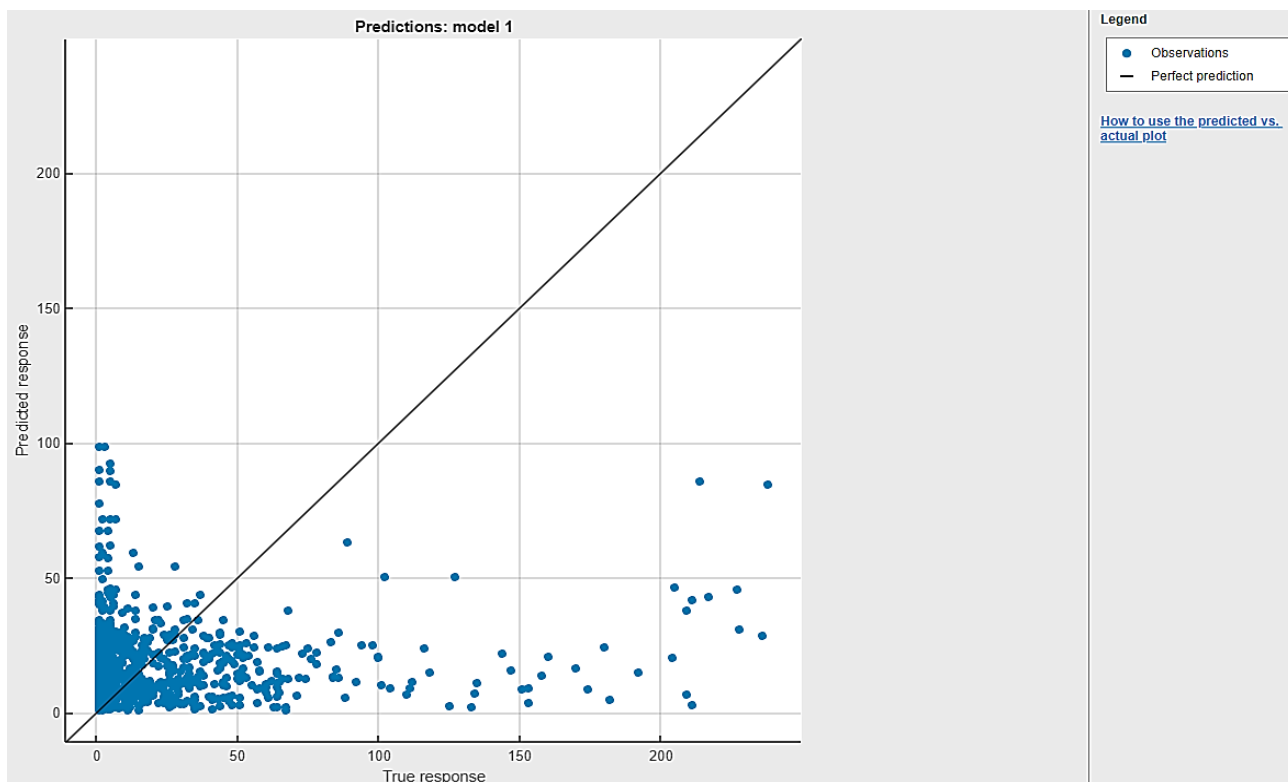


Рисунок 3.21 – Відповідність фактичних і прогнозних даних моделі Tree

Висновок: модель дерева рішень має великі значення RMSE та MSE, а від'ємне значення R-Squared вказує на те, що модель слабо пояснює варіацію вихідної змінної. Високе значення MAE свідчить про низьку точність прогнозів.

Використаємо стандартну модель Neural Network у середовищі MATLAB та розглянемо її результати (таблиця 3.17, рисунки 2.22, 2.23) [26].

Таблиця 3.17 – Результати обробки закупівельних даних моделлю Neural Network

Neural Network	RMSE	MSE	R-Squared	MAE
Validation	31,17	971,55	0,019793	17,488
Test	28,728	825,32	0,05	16,17

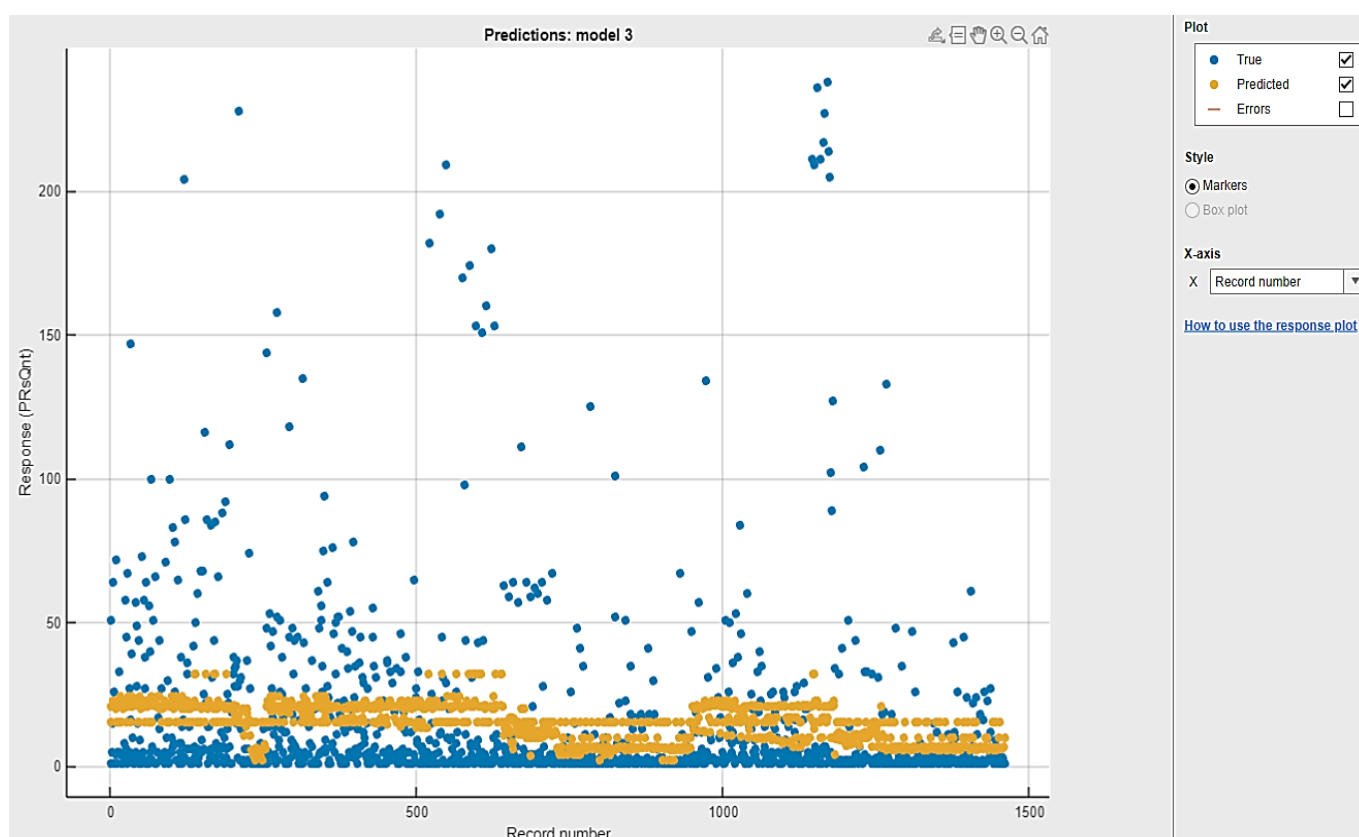


Рисунок 2.22 – Порівняння кількості фактичних і прогнозних даних моделі Neural Network

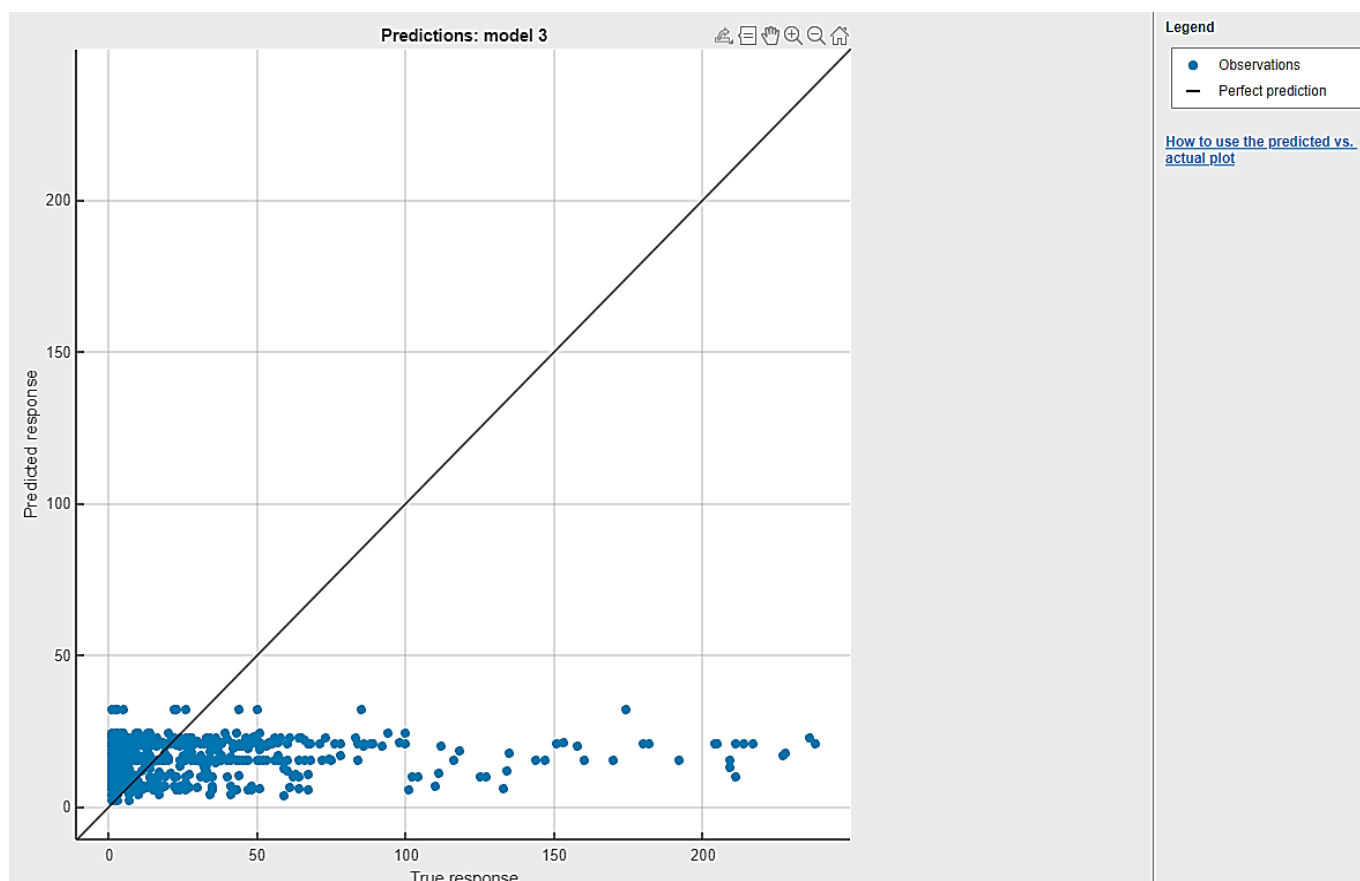


Рисунок 3.23 – Відповідність фактичних і прогнозних даних моделі Neural Network

Висновок: модель нейронної мережі має подібні показники до моделей лінійної регресії, але все ще показує низьку ефективність у поясненні варіації.

Зведемо отримані дані в одну таблицю та порівняємо отримані результати (таблиця 3.18).

Таблиця 3.18. Зведені дані прогнозування в середовищі MATLAB

	RMSE	MSE	R-Squared	MAE
Linear Regression	28,678	822,44	0,06	16,14
Lnteractions Linear	28,69	823,09	0,06	16,094
Tree	32,535	1058,5	-0,21	17,463
Neural Network	28,728	825,32	0,05	16,17

Узагальнюючи, можемо сказати, що всі чотири моделі мають свої переваги та недоліки.

Вибір оптимальної моделі може залежати від конкретних вимог задачі та характеристик даних. Великі значення RMSE та MSE можуть вказувати на те, що модель не дуже точно передбачає значення, а від'ємне значення R-Squared свідчить про те, що модель не може пояснити варіацію вихідної змінної.

Виходячи з отриманих даних, найефективнішою моделлю у середовищі MATLAB для прогнозування закупівельних даних є нейронна мережа Neural Network або регресійна модель Linear Regression, показники яких дуже близькі.

Середовище MATLAB дозволяє згенерувати код моделі автоматично:

- Linear Regression (Додаток Г);
- Neural Network (Додаток Д).

3.3. Аналіз результатів дослідження моделей класифікації

Порівнюючи результати досліджень середовищ RapidMiner та MATLAB (таблиці 3.17 та 3.18) можемо визначити деякі загальні тенденції та відмінності між моделями.

1. Точність прогнозів:

- модель Neural Network має найменші значення RMSE та MAE в обох таблицях, що говорить про високу точність прогнозів моделі;
- моделі Linear Regression у середовищах RapidMiner і MATLAB та Decision Tree у середовищі RapidMiner також мають подібні значення RMSE та MAE;
- модель Polynomial Regression у середовищі RapidMiner виявляється найменш точною серед розглянутих.

2. Відносні помилки: У обох таблицях відносні помилки (Relative Error) для всіх моделей є високими, але найменший відсоток помилки спостерігається у моделі Neural Network.

3. Інші показники:

- MSE: Модель Polynomial Regression в обох таблицях має найбільші значення MSE, що говорить про значну кількість квадратичних помилок;
- кореляція (Correlation): Модель Neural Network має найвищий показник кореляції серед розглянутих моделей в обох середовищах.

Загалом, модель Neural Network є найкращою при прогнозуванні та аналізу закупівельних даних з позиції точності прогнозів та кореляції. Однак, враховуючи контекст задачі та вимоги, в якості альтернативної моделі може використовуватися модель Linear Regression.

Сформовані рекомендації для вибору середовища моделювання, у якому краще здійснювати аналіз та прогнозування закупівельних даних, на основі набору їх інструментів та вартості ліцензій (таблиця 3.19).

3.4 Висновки за розділом

1. Neural Designer. Враховуючи високу вартість продукту, обмеження використання тільки для нейронних мереж, особливості приготування даних та можливість точного налаштування нейронної мережі і широкий вибір аналітичних звітів – може використовуватись великими комерційними компаніями та ІТ компаніями з високим рівнем фінансування аналітичних підрозділів.

2. RapidMiner. Враховуючи широкий вибір інструментів аналізу та обробки даних, зручний візуальний інтерфейс та наявність безкоштовних ліцензій для освітніх та особистих цілей – може використовуватись освітніми

зкладами для навчання та середніми та невеликими комерційними компаніями для аналізу та обробки закупівельних даних.

3. MATLAB. Враховуючи широкий вибір інструментів аналізу та обробки даних, невисоку вартість, наявність власної мови програмування, різноманітність інтегрованих моделей та інструментів візуалізації – може використовуватися освітніми закладами для навчання та дослідницької роботи.

Таблиця 3.19 – Вартість ліцензій у доларах США

Neural Designer	USD	Small year (per user)	Medium year (per user)	Large year (per user)	Research year (per user)	Teaching year (per module)
		2 495,00	4 995,00	7 495,00	2 495,00	4 995,00
RapidMiner	USD	Studio (Commercial)	Server (Commercial)	AI Hub (Commercial)	Studio (Community)	Server (Community)
		No information available			Free	Free
MATLAB	USD	Home (per user)	Standart year (per user)	Startups year (per user)	Student (per user)	Academic year (per user)
		95,00	940,00	3 850,00	55,00	275,00

Загальні висновки: модель Neural Network є найкращою при прогнозуванні та аналізу закупівельних даних з позиції точності прогнозів та кореляції. Однак, враховуючи контекст задачі та вимоги, в якості альтернативної моделі може використовуватися модель Linear Regression. Було доведено, що застосування сучасних інструментів data mining та машинного навчання дозволить досягти значного прискорення обробки даних.

РОЗДІЛ 4. РОЗРОБКА ТЕХНІЧНОГО ЗАВДАННЯ НА СТВОРЕННЯ СИСТЕМИ УПРАВЛІННЯ ЗАКУПІВЛЯМИ

4.1 Розробка бізнес-процесу управління закупівлями

З метою усунення слабких місць у процесі закупівель, необхідно розробити «Систему управління закупівлями», яка буде застосовувати розглянуті у роботі моделі машинного навчання при обробці, аналізі та прогнозуванні закупівельних даних [32]. Бізнес вимоги до «Системи управління закупівлями» (Business Requirements Specification) представлені у Додатку 5 [23].

Використання системи, потребує коригування діючого процесу закупівель. Удосконалений використанням системи управління закупівлями бізнес-процес, після формування потреб у матеріалах та послугах, дані про наявні на поточний момент потреби направляються в систему для обробки та аналізу (рисунок 4.1) [30]. Після підготовчої обробки даних із системи згідно рекомендацій описаних у п. 2.2.1 система прогнозує необхідні для прийняття стратегічних рішень дані та, отримавши із системи обліку наявні залишки матеріалів на складах, формує прогноз закупівель матеріалів та послуг на рік, з помісячним прогнозом. Після здійснюється розрахунок необхідних ресурсів системою управління закупівлями, автоматично формуються заявки на підбір персоналу (із зазначенням періоду коли новий персонал має приступити до виконання заявок) та заявки на закупівлю матеріалів та послуг. Отримавши затверджені заявки в системі, служба закупівель аналізує заявки та формує план тендерів на рік та розраховує необхідні ресурси для здійснення закупівель і при потребі подає заявку на пошук персоналу, для забезпечення своєчасності виконання заявок. Після чого здійснюються тендерні процедури

та вибір постачальників, з якими служба постачання укладає договори та на фінальному етапі організовує поставку на склади підприємства.

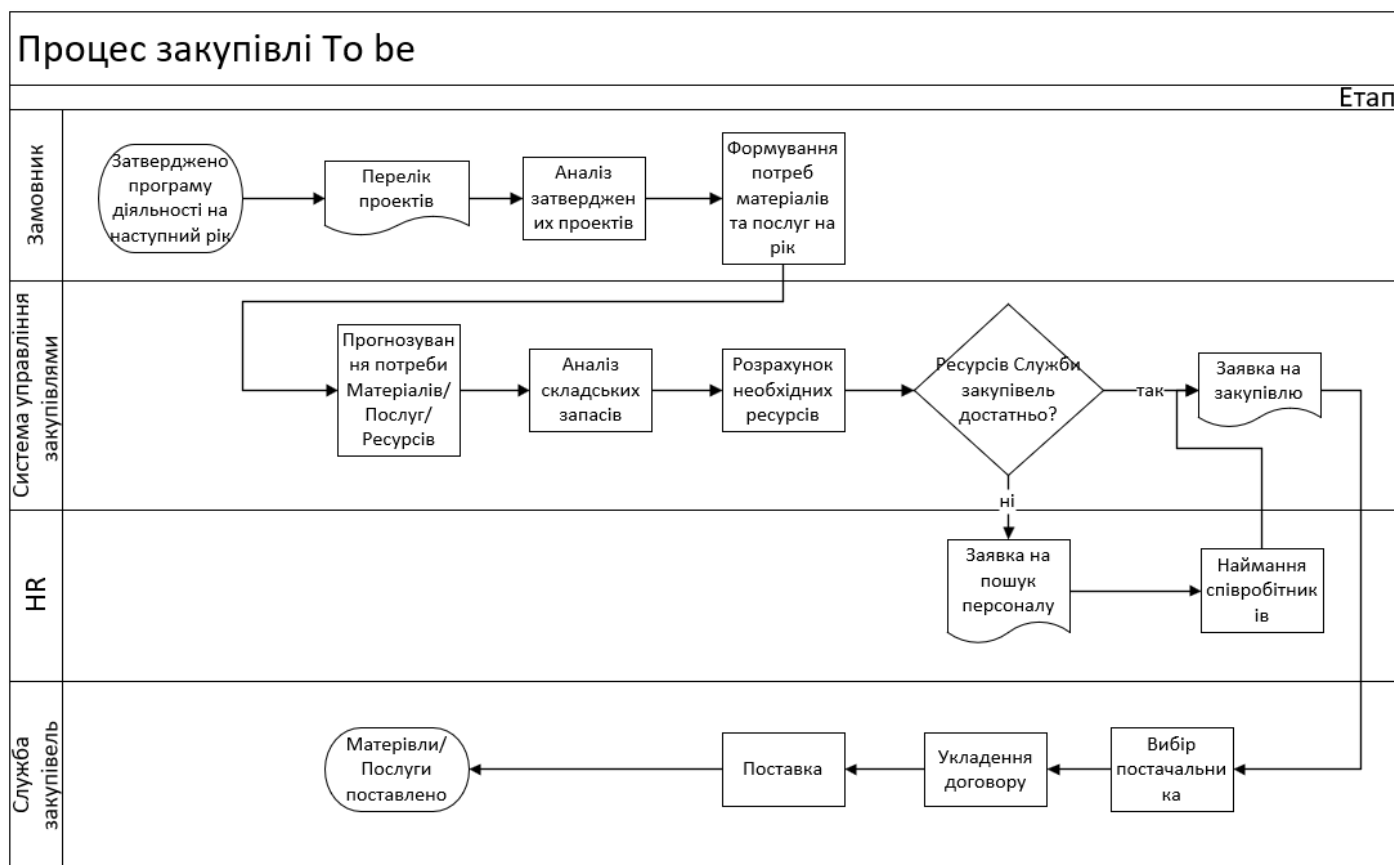


Рисунок 4.1 – Бізнес-процес управління закупівлями «to be»

4.2 Бізнес вимоги до системи управління закупівлями

Передбачається, що «Система управління закупівлями» буде використовуватись службою закупівель підприємства для прогнозування об'ємів закупівель майбутніх періодів (місяць, квартал, рік) та необхідних для цього ресурсів (бюджету, співробітників) та орієнтовні строки закупівель. Система повинна бути оптимізована для забезпечення ефективного функціонування бізнес-процесів.

Машинне навчання має бути використано для прогнозування попиту на матеріали та послуги, виявлення аномальних закупівельних операцій, автоматизованого вибору постачальників та оптимального розподілу ресурсів для закупівель.

Програмні компоненти повинні бути проєктовані та розроблені для забезпечення відповідності системи закупівель та постачання вимогам бізнесу та наукових стандартів. Недостатня ефективність в процесі закупівель, що пов'язані із складнощами у виборі постачальників, управлінні ризиками, контролі витрат, не відповідність поставки вимогам, не можливість точно передбачити попит на товари та послуги

Програмний продукт повинен вирішувати ці проблеми шляхом використання алгоритмів машинного навчання, які забезпечують більш точний аналіз та прогнозування даних про постачальників, допомагати встановлювати відповідність між параметрами поставки та вимогами, аналізувати дані та передбачати попит на основі різноманітних факторів.

Система взаємодіє [31] (рисунок 4.2):

- зі Системою обліку підприємства для отримання даних щодо закупівель попередніх періодів (Матеріали/Послуги, кількість, ціна, строки, умови тощо);
- з нормативними документами підприємства, щоб відповідати встановленим вимогам до процесу та строкам закупівель;
- з Замовником, інформуючи його про очікувані строки та вартість закупівель та прогнозування виконання замовлень ;
- зі Службою Закупівель, прогнозуючи очікувані об'єми та вартість закупівель та необхідні ресурси для їх своєчасного виконання.

Діаграма потоків даних «Системи управління закупівлями», яка демонструє обмін даними з зовнішнім середовищем розроблюваної Системи, показана на рисунку 4.3 [31].



Рисунок 4.2 – Контекстна діаграма Системи управління закупівлями

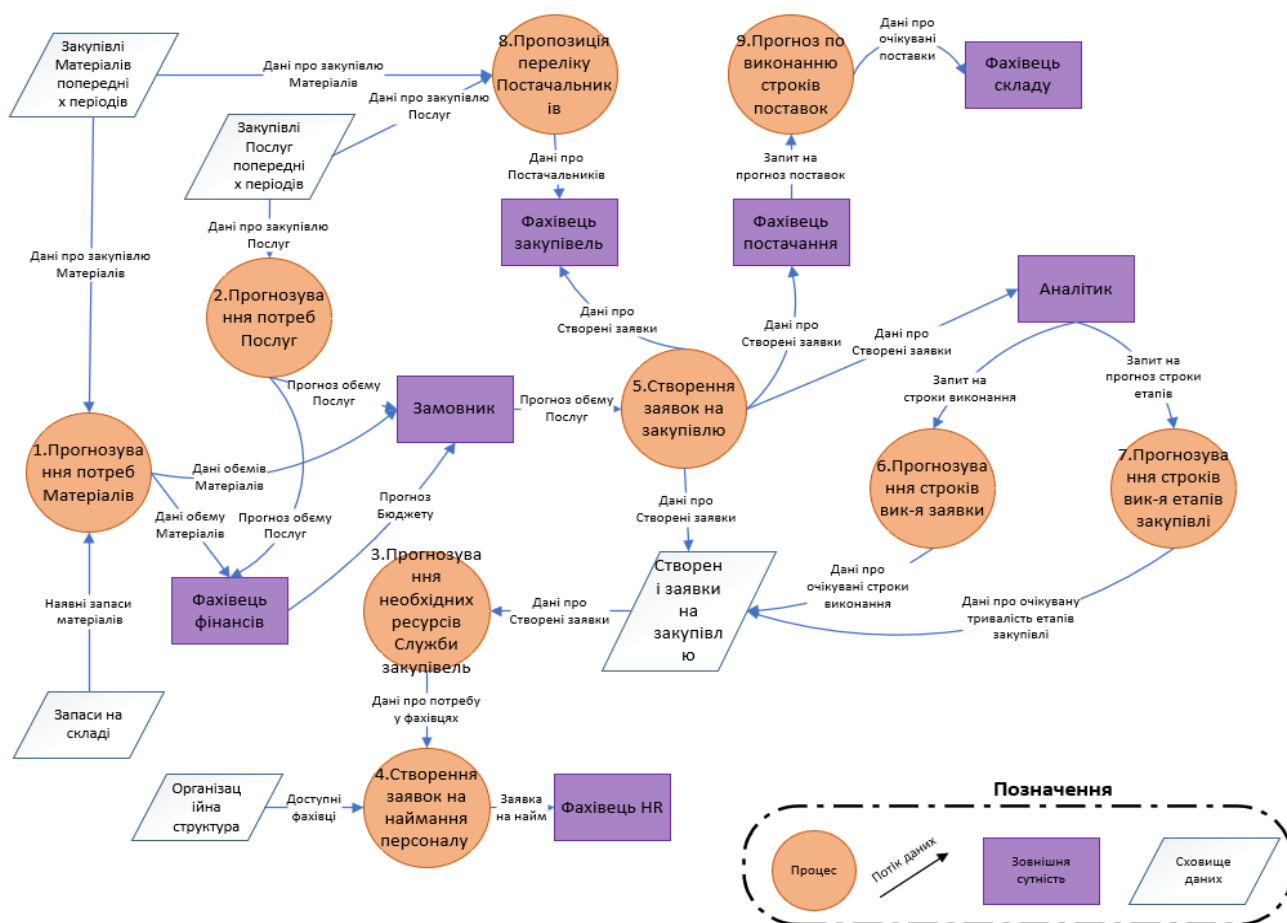


Рисунок 4.3 – Діаграма потоків даних Системи управління закупівлями

Сфера застосування. Для досягнення цілей проекту, поставлених бізнесом, та реалізації максимально можливих матеріальних та нематеріальних переваг (зисків) система має забезпечити:

- своєчасність та безперебійність поставок матеріалів та послуг, необхідних для виробництва;
- оптимізацію та прогнозування навантаження на співробітників Служби закупівель;
- автоматизацію процесу закупівлі матеріалів, обладнання та послуг;
- зниження витрат на закупівлю, що дозволить забезпечити оптимальне співвідношення між вартістю та якістю наданого товару або послуги;
- відслідковування всіх етапів процесу закупівлі, що покращить контроль та управління;
- скорочення строків поставок.

Для цього потрібно буде здійснити широкий спектр заходів, включаючи:

- збір інформації, формування навчальної вибірки;
- маркування даних навчальної вибірки;
- створення нейронної мережі;
- тренування НМ на основі навчальної вибірки;
- тестування моделі ШІ;
- підготовку необхідної звітності та документації.

На рисунку 4.4 представлено діаграму варіантів використання (Use Case) Системи [31].

За межами проекту. Система планується для використання Службою закупівель підприємства, але вона може бути використана і для потреб Служби реалізації матеріалів та послуг. Також вона може бути корисною в різних організаційних структурах, де необхідно здійснювати закупівлі. Наприклад, у великих корпораціях, державних органах, медичних установах, тощо.

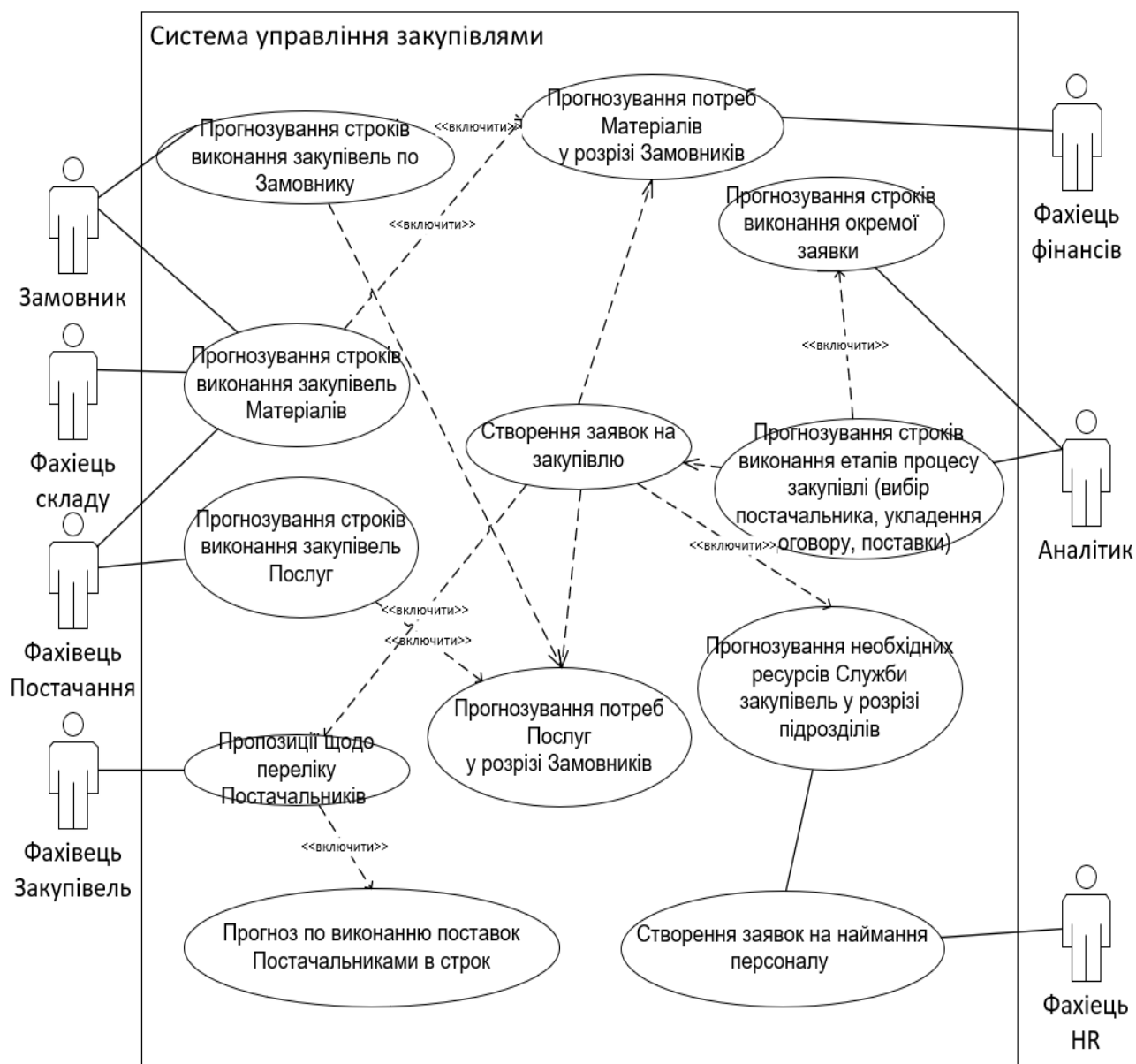


Рисунок 4.4 – Діаграма використання Системи управління закупівлями

Отже, сформовано мету та бізнес-цілі проекту щодо розробки та впровадження Системи [32].

Ділова мета проекту – розробити систему, яка в автоматичному режимі буде здійснювати обробку заявок на закупівлю в системі обліку компанії, та надавати прогноз закупівель у вигляді розгорнутого списку матеріалів та послуг, очікуваної кількості, необхідного бюджету та ресурсів для забезпечення своєчасної закупівлі, укладення договорів та поставок на склади Компанії необхідних для виробництва матеріалів.

Бізнес-цілі проекту:

- збільшити лояльність Замовників закупівель до Служби закупівель;
- усунення інформаційного вакууму між різними учасниками процесу закупівель;
- мінімізація впливу людського фактору на результати прогнозування, яке в поточний час проводиться вручну;
- збільшення операційної ефективності Служби закупівель.

4.3 Специфікація вимог до системи управління закупівлями

4.3.1 Функції системи управління закупівлями

Виходячи з аналізу діючого процесу закупівель та проведених інтерв'ю та опитувань фахівців та керівників служб закупівель і постачання, було розроблено технічне завдання на розробку Системи управління закупівлями. Система повинна забезпечити функціональність, яка представлена у таблиці 4.1 [23].

Таблиця 4.1 – Функції системи управління закупівлями

ID	Особливість системи	Пріоритет
FR01	Прогнозування потреб Матеріалів та Послуг у розрізі Замовників	M
FR02	Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів	M
FR03	Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників	M
FR04	Створення заявок на закупівлю	W

Продовження таблиці 4.1

FR05	Пропозиції щодо переліку Постачальників минулих періодів	C
FR06	Прогноз по виконанню поставок Постачальниками в строк	W

Пріорітизацію функціональності системи визначено за системою MoSCoW (таблиця 4.2) [23].

Таблиця 4.2 – Рейтинг пріоритетів

Рейтинг пріоритетів	Опис
M – Повинен мати	Опис вимог, що повинні бути задоволені у фінальному представленні рішення для досягнення успіху.
S – Варто було б мати	Представляє високо-пріоритетні деталі (пункти), що повинні бути добавлені у рішення, якщо це можливо.
C – Можливо мати	Описує вимоги котрі вважаються бажаними, але не обов'язковими. Вони будуть включені, якщо дозволять час і ресурси.
W – Хотілося б мати	Представляє вимоги, які були погоджені зацікавленими сторонами, що не будуть додаватися до анонсування, проте можуть бути розглянуті у подальшому.

Послідовності стимулів/відповідей (Stimulus/Response Sequences) для реалізації основних функцій Системи [23]:

- запит від користувача. Stimulus Користувач формує в системі запит, який включає основні параметри прогнозування: Базовий період, прогнозований період, напрямок та обмеження прогнозування (підрозділ);
- аналіз запиту. Response Система аналізує запит, та формує запити до бази даних;
- направлення запиту до бази даних. Stimulus Сформований запит направляється для обробки та формування даних у систему підприємства;
- отримання даних з бази даних. Response Дані з бази даних отримуються та зберігаються у тимчасовому файлі;

- підготовка даних. Stimulus Отримані дані приводяться до необхідного формату, виявляються та заповнюються пропуски, коригуються аномальні значення;
- обробка даних. Response Система обробляє дані і формує прогнози згідно запиту на основі генетичних алгоритмів машинного навчання;
- формування відповіді користувачеві. Response Система формує та надсилає відповідь користувачеві у вигляді таблиці або файлу;
- інформування HR Служби. Response Система, засобами електронної пошти та MS Teams розсилає інформацію фахівцям HR Служби, про необхідність пошуку персоналу для Служби закупівель за встановленою формою.

4.3.2 Прогнозування потреб Матеріалів та Послуг у розрізі Замовників (FR01).

Опис і пріоритет. Ця функція за допомогою генетичних алгоритмів машинного навчання на основі аналізу закупівель за попередні роки, моделює очікувані закупівлі та їх вартість і об'єми, поточного та майбутнього періодів. Прогнозування здійснюється в розрізі конкретних матеріалів, послуг, Замовників, Постачальників, груп матеріалів. Пріоритет: Високий

Основні дії Системи щодо реалізації цієї функції:

- прогнозування потреб Матеріалів в розрізі Замовників;
- прогнозування потреб Послуг в розрізі Замовників;
- прогнозування потреб Матеріалів в розрізі Постачальників;
- прогнозування потреб Послуг в розрізі Постачальників;
- прогнозування орієнтовного бюджету закупівель;
- формування звітів.

Опис вимоги «Прогнозування потреб Матеріалів та Послуг у розрізі Замовників» надано у таблиці 4.3.

Таблиця 4.3 – Функція з прогнозування потреб Матеріалів та Послуг

ID функціональної вимоги	FR01
Тип вимоги	Функціональна
Пріоритет	М
Опис вимоги	За допомогою генетичних алгоритмів машинного навчання на основі аналізу закупівель за попередні роки, моделювання очікуваних закупівель та їх вартість
Перехресне посилання на вимоги бізнесу	BR01, BR02
Перехресне посилання на використання	UC01, UC02
Бізнес правило	Вимоги до товару, що закупається описуються у Технічному зайданні Закупівля однотипних трварів здійснюється в рамках одного тендеру Закупівлі можуть здійснюватись в рамках затвердженого бюджету Бізнес-плану діяльності Виконання бюджету закупівель розраховується як % суми створених заявок до затвердженого бюджету діяльності
Джерело	Протокол інтерв'ю №1

4.3.3 Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів (FR02)

Опис і пріоритет. Ця функція на основі встановлених на підприємстві нормативів навантаження співробітників Служби закупівель, проаналізувавши навантаження за попередні роки, моделює очікувані ресурси поточного та

майбутнього періодів. Прогнозування здійснюється в розрізі підрозділів Служби закупівель підприємства. Пріоритет: Високий.

Основні дії щодо реалізації цієї функції:

- прогнозування потреб об'ємів (кількість заявок/договорів) закупівель на вказаний період;
- прогнозування потреб Служби закупівель у розрізі підрозділів;
- створення заявок на наймання персоналу для Служби закупівель;
- формування звітів.

Опис функції «Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів» надано у таблиці 4.4.

Таблиця 4.4 – Функції системи з прогнозування необхідних ресурсів

ID функціональної вимоги	FR02
Тип вимоги	Функціональна
Пріоритет	М
Опис вимоги	Моделювання необхідних ресурсів Служби закупівель поточного та майбутнього періодів.
Перехресне посилання на вимоги бізнесу	BR03
Перехресне посилання на використання	UC03
Бізнес правило	<p>Процес закупівлі поділяється на окремі підпроцеси: закупівлю та постачання</p> <p>Якщо жоден з ключових показників не виконано, робота служби закупівель вважається незадовільною.</p> <p>Якщо кількість отриманих заявок перевищує пропускну здатність закупівель більше ніж на 20%, фахівець HR повинен почати пошук кандидатів для найому у підрозділ закупівель.</p> <p>Якість роботи служби закупівель розраховується як середнє між % виконання поставок в строк та % виконання бюджету в строк.</p>
Джерело	Протокол інтерв'ю №1

4.3.4 Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників (FR03)

Опис і пріоритет. Ця функція на основі даних за попередні роки, моделює очікувані строки виконання закупівель поточного та майбутнього періодів. Прогнозування здійснюється в розрізі Матеріалів/Послуг/Замовників. Пріоритет: Високий.

Основні дії Системи щодо реалізації цієї функції:

- прогнозування строків виконання закупівель Матеріалів у розрізі Замовників;
- прогнозування строків виконання закупівель Послуг у розрізі Замовників;
- інформування Замовників та Служби закупівель про очікувані строки виконання заявок та необхідні ресурси;
- формування звітів.

Опис вимоги «Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників» надано у таблиці 4.5.

Таблиця 4.5 – Функції системи з прогнозування строків виконання закупівель

ID функціональної вимоги	FR03
Тип вимоги	Функціональна
Пріоритет	М
Опис вимоги	Моделювання строків виконання закупівель поточного та майбутнього періодів.
Перехресне посилання на вимоги бізнесу	BR04, BR05, BR06, BR09
Перехресне посилання на використання	UC06, UC07

Продовження таблиці 4.5

ID функціональної вимоги	FR03
Бізнес правило	<p>Кожна груп матеріалів , що закупається має свої визначені строки закупівлі</p> <p>Закупівлі мають здійснюватись в затверджені строки</p> <p>Якщо матеріал поставлено на склад пізніше дати потреби вказаної у заявці на закупівлю, заявка вважається простроченою</p> <p>Якщо жоден з ключових показників не виконано, робота служби закупівель вважається незадовільною</p> <p>Якість роботи служби закупівель розраховується як середнє між % виконання поставок в строк та % виконання бюджету в строк</p> <p>Строк виконання заявки розраховується як різниця дати поставки та дати створення заявки</p>
Джерело	Протокол інтерв'ю №1

4.3.5 Створення заявок на закупівлю (FR04)

Опис і пріоритет. Ця функція на основі даних за попередні роки , моделює очікувані об'єми закупівель поточного та майбутнього періодів. Прогнозування здійснюється в розрізі Матеріалів/Послуг/Замовників. На основі прогнозних даних формує перелік заявок, які необхідно створити. Користувач переглядаючи перелік запропонованих заявок підтверджує необхідність їх створення. Система створює заявки. Пріоритет: Низький.

Основні дії Системи щодо реалізації цієї функції:

- прогнозування об'ємів закупівель Матеріалів у розрізі Замовників;
- прогнозування об'ємів закупівель Послуг у розрізі Замовників;
- формування пропозицій по створенню заявок на закупівлю;
- підтвердження та створення заявок на закупівлю;
- формування звітів.

Опис функції «Створення заявок на закупівлю» надано у таблиці 4.6.

Таблиця 4.6 – Функція системи із створення заявок на закупівлю

ID функціональної вимоги	FR04
Тип вимоги	Функціональна
Пріоритет	W
Опис вимоги	Автоматизація процесу створення заявок.
Перехресне посилання на вимоги бізнесу	BR07
Перехресне посилання на використання	UC04
Бізнес правило	Закупівля здійснюється виключно при наявності заявки в системі обліку.
Джерело	Інтерв'ю

4.3.6 Пропозиції щодо переліку Постачальників минулих періодів

Опис і пріоритет. Ця функція на основі даних за попередні роки, моделює можливості постачальників по виконанню заявок на закупівлю. Прогнозування здійснюється в розрізі Матеріалів/Послуг. На основі прогнозних даних формує перелік Постачальників для участі у тендері. Пріоритет: Середній.

Основні дії в межах цієї функції:

- аналіз Постачальників попередніх періодів;
- формування пропозиції по Постачальникам в розрізі Матеріалів;
- формування пропозиції по Постачальникам в розрізі Послуг;
- формування звітів.

Опис функціональної вимоги «Пропозиції щодо переліку Постачальників минулих періодів» надано у таблиці 4.7.

Таблиця 4.7 – Функція з пропозиції щодо переліку Постачальників

ID функціональної вимоги	FR05
Тип вимоги	Функціональна
Пріоритет	C
Опис вимоги	Автоматизація процесу пошуку постачальників для участі у тендері на закупівлю.
Перехресне посилання на вимоги бізнесу	BR10
Перехресне посилання на використання	UC08
Бізнес правило	Закупівля однотипних товарів здійснюється в рамках одного тендеру Якщо поставлений матеріал Постачальника не прийнято складом, або поставка цим самим Постачальником регулярно здійснюється із запізненням, то Постачальник вважається не надійним
Джерело	Інтерв'ю

4.3.7 Прогноз по виконанню поставок Постачальниками в строк

Опис і пріоритет. Ця функція на основі даних за попередні роки, моделює можливості служби закупівель та постачання по спроможності виконання поставок в строк. Прогнозування здійснюється в розрізі Матеріалів. На основі прогнозних даних формує очікувані строки виконання заявок на закупівлю в розрізі Матеріалів. Пріоритет: Низький.

Основні дії в межах цієї функції:

- аналіз строків виконання заявок на закупівлю попередніх періодів;
- прогнозування строків виконання заявок на закупівлю по Замовникам в розрізі Матеріалів;
- формування звітів.

Опис функціональної вимоги «Прогноз щодо строків виконання заявок на закупівлю» надано у таблиці 4.8.

Таблиця 4.8 – Функція з прогнозу строків виконання заявок

ID функціональної вимоги	FR06
Тип вимоги	Функціональна
Пріоритет	W
Опис вимоги	Прогнозування строків виконання поставок Постачальниками
Перехресне посилання на вимоги бізнесу	BR11
Перехресне посилання на використання	UC09
Бізнес правило	Якщо поставлений матеріал Постачальника не прийнято складом, або поставка цим самим Постачальником регулярно здійснюється із запізненням, то Постачальник вважається не надійним
Джерело	Інтерв'ю

4.4 Висновки

У результаті виконання роботи проведена розробка технічного завдання на створення системи управління закупівлями. У результаті розроблено удосконалений бізнес-процес управління закупівлями, сформовані вимоги до програмного забезпечення для системи управління закупівлями. Впровадження такої системи на підприємстві дозволить значно поліпшити процеси управління закупівлями.

РОЗДІЛ 5. ЕКОНОМІЧНЕ ОБҐРУНТУВАННЯ ЗАПРОПОНОВАНИХ ПРОГРАМНИХ РІШЕНЬ

5.1 Визначення рангу значущості роботи

Написання цієї роботи призвело до створення технічного завдання для розробки програмного продукту. Оскільки магістерська робота є науково-дослідною, була проведена оцінка її значущості.

Визначено ранг пріоритетності роботи (таблиця 5.1) [35].

Таблиця 5.1 – Ранги пріоритетності НДРС

П1	4	НДРС, що виконується на замовлення держави відповідно до річних планів і програм
П2	3	НДРС, що виконуються на замовлення вишів
П3	2	НДРС, що виконуються на замовлення директорів підприємств
П4	1	НДРС, що не відносяться до вищевказаних видів, в т.ч, ініціативні НДР, не підтверджені замовником

Згідно таблиці 5.1 дана робота має ранг П2 (2 бали), оскільки тема роботи підтверджена керівництвом підприємства і виконувалась на його замовлення. Визначено ранг новизни НДРС (таблиця 5.2) [35].

Таблиця 5.2 – Ранги новизни НДРС

Н1	3	принципово нові дослідження (світова новизна)
Н2	2	удосконалення в рамках розвитку певного напрямку
Н3	1	удосконалення, розвиток частині відомого дослідження

Згідно таблиці 5.2 дана робота має ранг Н2 (2 бали), оскільки метою роботи є вдосконалення процесів управління закупівлями та постачанням підприємства через вдосконалення процесу підготовки та обробки

закупівельних даних для прийняття ефективних управлінських рішень на основі розробки ПЗ, яке реалізує алгоритми машинного навчання [35].

Визначено ступінь готовності НДРС (таблиця 5.3).

Таблиця 5.3 – Ранги готовності НДРС

Г1	2	підготовка до використання ПП
Г2	1	підготовка проекту ПП

Згідно таблиці 5.3 дана робота має ранг Г2 (1 бал), оскільки результатом роботи є розроблене технічне завдання на розробку ПП, для обробки закупівельних даних завдяки обраного в роботі методу машинного навчання для підготовки та обробки даних щодо прийняття ефективних управлінських рішень в сфері закупівель [35].

Визначимо ступінь опрацьованості наукової теми (таблиця5.4).

Таблиця 5.4 –Ранги опрацьованості наукової теми

О1	3	кількість опублікованих статей автора
О2	2	кількість вивчених монографій і дисертацій, що знайшли відображення в НДРС
О3	1	кількість вивчених статей за своєю тематикою (3х річної давності)

Згідно таблиці 5.4 дана робота має ранг О1 (3 бали *1 опубліковану статтю) та О3 (1 бал * 5 вивчених статей за тематикою 3х річної давності) .

Ранг розраховуємо за формулою [35]:

$$O = \left(\frac{3*O_1 + 2*O_2 + 1*O_3}{45} \right) * 4 = \left(\frac{3*1 + 2*0 + 1*5}{45} \right) * 4 = 0,71 \quad (5.1)$$

На підставі отриманих результатів побудовано графік рангу значущості роботи (рисунок 5.1).

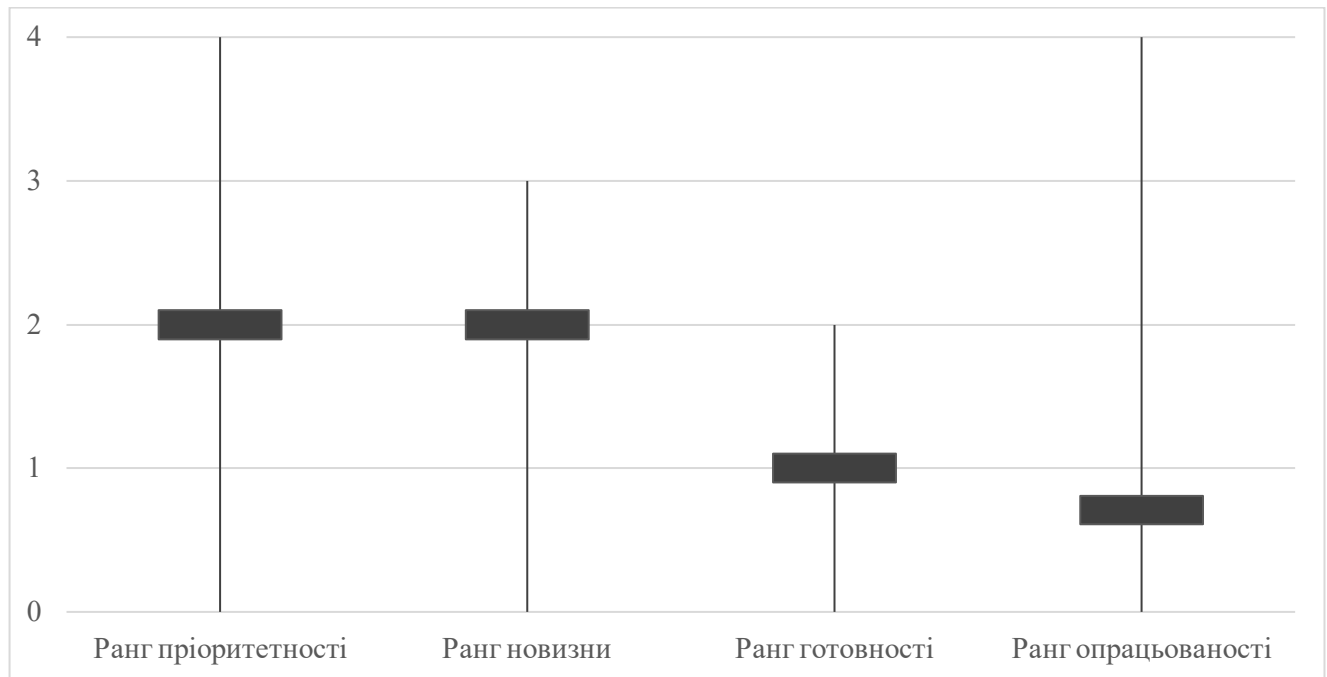


Рисунок 5.1 – Графік рангу значущості роботи

Зрозуміло, що заявлена робота відзначається високою важливістю та готовністю до використання. Запропоновані покращення представляють собою інноваційний підхід в рамках досліджуваного напрямку. Виділяється достатньо високий рівень дослідженості у цьому конкретному напрямку.

5.2 Визначення економічного ефекту від впровадження розробки

На другому етапі, коли вже є розроблений програмний продукт, рекомендується провести детальну оцінку якості цього продукту, отриманого в результаті науково-дослідної роботи, і порівняти його з аналогічним програмним забезпеченням. Оцінка якості програмного продукту включає в себе ряд операцій, таких як визначення критеріїв якості, обрання показників для оцінки, визначення конкретних значень цих показників та їх порівняння з базовими параметрами. Цей процес дозволяє отримати об'єктивну картину

ефективності та якості програми, враховуючи його конкурентні переваги та характеристики.

Зазвичай оцінку якості здійснює група експертів у складі п'яти-семи чоловік, які дають оцінку елементам факторів якості:

- надійність;
- супровідність;
- зручність застосування;
- ефективність;
- універсальність;
- коректність.

Оцінка здійснюється в сім етапів, відповідно до «ДСТУ ISO/IEC 9126-1:2013 (ISO/IEC 9126-1:2001, IDT) Національний стандарт України. Програмна інженерія. Якість продукту. Частина 1. Модель якості виконаємо оцінку показників якості програмного продукту», що описує номенклатуру показників якості програмних продуктів і властивості, які їх характеризують.

Однак, враховуючи, що програмний продукт ще не розроблено, а за результатами дослідження сформовано лише технічне завдання на його розробку, та відсутність аналогічних продуктів на ринку оцінку якості провести не можливо, тому розрахуємо очікувану економічну вигоду від впровадження запропонованого програмного продукту.

5.2.1 Оцінка вартості розробки програмного продукту

Вартість розробки програмного продукту складається з вартості дослідження та вартості розробки програмного забезпечення. Вартість дослідження включає в себе витрати на придбання програмних продуктів середовища дослідження (наприклад Neural Designer, RapidMiner Studio,

MSTLAB тощо), вартості ресурсів та часу використаного співробітниками для виявлення вимог користувачів, проведення дослідження, обробки та аналізу даних, навчання та тестування моделі, формування технічного завдання на розробку програмного забезпечення та комунікацію з розробником (таблиця 5.5). Для розрахунку вартості дослідження та економічного ефекту, будемо використовувати середню заробітну плату в Україні яка на січень 2022 року, становила 14 577 грн. [33].

Таблиця 5.5 – Вартість дослідження

No.	Етап	Потреба в ресурсах	Вартість, грн
1	Вивчення питання		72 885 ₴
1.1	Огляд процесу аналізу та прогнозування закупівельних даних		40 087 ₴
1.1.1	Дослідження процесу аналізу та прогнозування закупівельних даних	персонал + робоче місце	14 577 ₴
1.1.2	Дослідження сучасних методів аналізу закупівельних даних	персонал + робоче місце	10 933 ₴
1.1.3	Дослідження сучасних методів прогнозування закупівельних даних	персонал + робоче місце	14 577 ₴
1.2	Огляд можливостей ПЗ Neural Designer	ПЗ Neural Designer + персонал + робоче місце	10 933 ₴
1.3	Огляд можливостей ПЗ Rapidminer Studio	ПЗ Rapidminer Studio + персонал + робоче місце	10 933 ₴
1.4	Огляд можливостей ПЗ MATLAB	ПЗ MATLAB + персонал + робоче місце	10 933 ₴
2	Планування		76 529 ₴
2.1	Розробка Плану управління дослідженням		58 308 ₴
2.1.1	Розробка Робочої структури кваліфікаційної роботи (WBS)	персонал + робоче місце	10 933 ₴
2.1.2	Побудова діаграми Ганта	персонал + робоче місце	7 289 ₴
2.1.3	Розробка RACI-матриці	персонал + робоче місце	7 289 ₴
2.1.4	Розробка плану комунікацій	персонал + робоче місце	10 933 ₴

Продовження таблиці 5.5

No.	Етап	Потреба в ресурсах	Вартість, грн
2.1.5	Розробка Плану залучення стейкхолдерів	персонал + робоче місце	10 933 ₴
2.1.6	Розробка Плану управління ризиками	персонал + робоче місце	10 933 ₴
2.2	Погодження плану управління дослідженням	персонал + робоче місце	18 221 ₴
3	Дослідження		373 972 ₴
3.1	Дослідження методів навчання машинного навчання		297 443 ₴
3.1.1	Дослідження переваг та недоліків методів машинного навчання для класифікації закупівельних даних	ПЗ Rapidminer Studio + персонал + робоче місце	79 451 ₴
3.1.2	Дослідження переваг та недоліків методів машинного навчання для прогнозування закупівельних даних	ПЗ Neural Designer + персонал + робоче місце	149 474 ₴
3.1.3	Дослідження ключових закономірностей для здійснення точних прогнозів щодо закупівельних операцій	ПЗ MATLAB + персонал + робоче місце	68 518 ₴
3.2	Розробка моделі аналізу та прогнозування закупівельних даних	персонал + робоче місце	76 529 ₴
3.2.1	Розробка моделі та алгоритму класифікації закупівельних даних	персонал + робоче місце	32 798 ₴
3.2.2	Розробка моделі та алгоритму прогнозування закупівельних даних	персонал + робоче місце	43 731 ₴
4	Тестування		145 770 ₴
4.1	Перевірка розроблених моделей на історичних даних підприємства		69 241 ₴
4.1.1	Тестування моделі класифікації закупівельних даних	ПЗ Rapidminer Studio, MATLAB + персонал + робоче місце	32 798 ₴
4.1.2	Тестування моделі прогнозування закупівельних даних	ПЗ Neural Designer + персонал + робоче місце	36 443 ₴
4.2	Уточнення моделей за отриманими результатами	персонал + робоче місце	43 731 ₴
4.3	Тестування уточнених моделей	персонал + робоче місце	32 798 ₴
5	Оформлення результатів		54 664 ₴
5.1	Оформлення результатів дослідження		32 798 ₴
5.1.1	Формування рекомендацій щодо класифікації закупівельних даних	персонал + робоче місце	10 933 ₴

Продовження таблиці 5.5

No.	Етап	Потреба в ресурсах	Вартість, грн
5.1.2	Формування рекомендацій щодо прогнозування закупівельних даних	персонал + робоче місце	10 933 ₴
5.1.3	Формування рекомендацій щодо виявлення ризиків	персонал + робоче місце	10 933 ₴
5.2	Оформлення Технічного завдання на розробку	персонал + робоче місце	21 866 ₴
			723 820 ₴

Як бачимо, орієнтовна вартість дослідження становить 723 820 грн. За результатами дослідження сформовано технічне завдання на розробку, яке має свою вартість яка має бути також врахована у вартості продукту. Розрахунок виконаємо використовуючи «Калькулятор СОСОМО», що дозволяє виконати орієнтовні розрахунки трудомісткості та термінів розробки програмного продукту [34]. Для розрахунків встановлюємо вихідні дані, що відповідають розробці програмного продукту об'ємом 10 000 рядків коду та середньої складності розробки продукту (рисунок 5.2).

COCOMO II - Constructive Cost Model

Software Size Sizing Method:

[SLOC](#) % Design Modified % Code Modified % Integration Required Assessment and Assimilation (0% - 8%) Software Understanding (0% - 50%) Unfamiliarity (0-1)

New Reused Modified

Software Size Probability Distribution

Iterations

Software Equivalent Size (KSLOC)	# Iterations
5-6	28
6-8	168
8-10	331
10-11	340
11-13	101
13-14	31

Software Equivalent Size (KSLOC)

Software Scale Drivers

Precedentedness Architecture / Risk Resolution Process Maturity

Development Flexibility Team Cohesion

Software Cost Drivers

Product

Required Software Reliability Data Base Size Product Complexity Developed for Reusability Documentation Match to Lifecycle Needs

Personnel

Analyst Capability Programmer Capability Personnel Continuity Application Experience Platform Experience Language and Toolset Experience

Platform

Time Constraint Storage Constraint Platform Volatility

Project

Use of Software Tools Multisite Development Required Development Schedule

Maintenance

Software Labor Rates

Cost per Person-Month (Dollars)

Рисунок 5.2 – Параметри для розрахунку «Калькулятор СОСОМО II»

Після розрахунків отримуємо дані щодо орієнтовної вартості та строку розробки програмного продукту «Система управління закупівлями» (рис. 5.3).

Results

Software Development (Elaboration and Construction) Staffing Profile

Effort = 37.0 Person-months
 Schedule = 11.6 Months
 Cost = \$55480

Total Equivalent Size = 10000 SLOC
 Effort Adjustment Factor (EAF) = 1.00

Acquisition Phase Distribution

Phase	Effort (Person-months)	Schedule (Months)	Average Staff	Cost (Dollars)
Inception	2.2	1.4	1.5	\$3329
Elaboration	8.9	4.3	2.0	\$13315
Construction	28.1	7.2	3.9	\$42165
Transition	4.4	1.4	3.1	\$6658

Рисунок 5.3 – Результати розрахунку «Калькулятор COCOMO II»

Прийнявши курс долару рівним 38 грн за 1 долар США, розрахуємо, що вартість розробки програмного забезпечення складає 2 108 240 гривень. Таким чином орієнтовна вартість розробки програмного Системи управління закупівлями складає 2 832 060,25 грн. Також важливо пам'ятати, що річне використання ліцензій програмних засобів Neural Designer, RapidMiner Studio та MATLAB в подальшому становитиме орієнтовно 166 250 грн.

5.2.2 Оцінка очікуваного економічного ефекту

Для визначення економічного ефекту від впровадження Системи управління закупівлями після його розробки, необхідно порівняти як зміниться вартість виконання регулярних процесів аналізу закупівельних даних, після розробки та впровадження Системи управління закупівлями. Для розрахунку вартості процесів до та після впровадження впровадження Системи управління закупівлями, будемо використовувати середню заробітну плату в Україні яка на січень 2022 року, становила 14 577 грн. [33].

Проаналізовано діючий процес та вплив впровадження впровадження Системи управління закупівлями (СУЗ) (таблиця 5.6).

Таблиця 5.6 – Вплив СУЗ на вартість аналізу закупівельних даних

	As is			To be		
	Відповідальна особа	Кількість у рік	Час на виконання, годин	Відповідальна особа	Кількість у рік	Час на виконання, годин
Вивантажити заявки створені за попередній місяць	Аналітик	12	0,5	Аналітик	12	0,5
Проаналізувати заявки та розподілити їх за напрямками	Аналітик	12	0,1	СУЗ	12	0,1
Направити перелік заявок на фахівців закупівель	Аналітик	12	0,1	-	-	-
Проаналізувати заявки та визначити тип закупівлі	Закупівельник	84	2	СУЗ	12	0,2
Направити розподіл заявок аналітику закупівель	Закупівельник	84	0,1	-	-	-
Консолідувати отриману інформацію	Аналітик	12	1	-	-	-
Сформувати перелік заявок на рік	Аналітик	12	0,5	Аналітик	12	0,5
Розрахувати прогнозну кількість заявок на наступний місяць	Аналітик	12	1	СУЗ	12	0,2
Розрахувати кількість ресурсів на виконання заявок на наступний місяць	Аналітик	12	2	СУЗ	12	0,2
Проаналізувати достатність наявних ресурсів	Аналітик	12	0,5	СУЗ	12	0,1
Надати інформацію керівникам відділів про недостатність ресурсів	Аналітик	2	0,1	-	-	-
Сформувати заявку на підбір персоналу	Керівник відділу	3	1,5	СУЗ	3	0,1
Направити заявку в службу HR	Керівник відділу	3	0,1	СУЗ	3	0,1
Час на обробку даних			250			22
ЗП співробітників			305 567			27 156
Економічна вигода від впровадження СУЗ						-278 411

Порівнюючи орієнтовну вартість розробки програмного Системи управління закупівлями та економію від її впровадження бачимо, що суму витрат на дослідження та розробку Системи управління закупівлями перевищує розмір ефекту від впровадження на 2 553 650 гривень, тобто інвестиції в дослідження та розробку окупляться протягом 10 років використання Системи управління закупівлями на 1 підприємстві. При впровадженні цієї системи на всіх підприємствах групи, економічна вигода зростатиме пропорційно кількості підприємств до сум економічної вигоди від впровадження і витрати окупляться протягом першого року використання, оскільки впровадження на інших підприємствах не потребуватиме додаткових розробок.

Також може бути розглянута можливість продажу даного продукту третім особам, але це питання потребує додаткового дослідження та аналізу, оскільки з високою ймовірністю продукт потребуватиме доробки та адаптації до систем обліку Покупців продукту.

5.3 Висновки

Проведено економічне обґрунтування запропонованих програмних рішень для системи управління закупівлями з використанням методів машинного навчання, доведено економічну доцільність проведених досліджень та розробки програмних рішень.

Отже, враховуючи досить високий ранг значущості роботи та можливість повернути витрачені кошти на дослідження та розробку протягом першого року використання продукту при впровадженні на всіх підприємствах групи, можемо зробити висновок про доцільність та необхідність здійснення розробки та впровадження у відповідності з розробленим Технічним завданням (додаток Є).

ЗАГАЛЬНІ ВИСНОВКИ

У результаті виконання роботи було отримано такі результати.

Проведено аналіз процесів управління закупівлями, аналіз проблем обробки, аналізу та інтерпретації даних у системах управління закупівлями та поставками. Доведено актуальність вирішення проблеми прискорення обробки, аналізу та прогнозування закупівельних даних. За результатами аналізу сформовано постановку задачі дослідження з урахуванням особливостей процесу закупівлі на підприємстві.

Проведено аналіз середовищ, які можна використовувати для розв'язання задач обробки, аналізу та прогнозування закупівельних даних (Neural Designer, RapidMiner Studio та MatLab). Сформовано такі рекомендації щодо вибору середовища.

- Neural Designer: висока вартість продукту, але відмінно підходить для великих комерційних компаній та IT-підрозділів з високим рівнем фінансування.

- RapidMiner: зручний візуальний інтерфейс, широкий вибір інструментів та безкоштовні ліцензії для освіти роблять його ідеальним для освітніх організацій та невеликих комерційних компаній.

- MATLAB: рекомендується для освітніх закладів, проведення дослідницької роботи завдяки різноманітності інтегрованих моделей та інструментів візуалізації, а також доступної вартості.

Проведено дослідження методів машинного навчання для аналізу та прогнозування закупівельних даних. На основі проведених досліджень та порівняння методів аналізу та прогнозування даних закупівель, які реалізовані у середовищах RapidMiner та MATLAB, зроблено висновки щодо використання моделей машинного навчання на основі таких критеріїв:

- точність прогнозів: модель Neural Network виявилася найточнішою, маючи найменші значення RMSE та MAE в обох середовищах. Моделі Linear

Regression та Decision Tree також показали схожу точність між середовищами RapidMiner та MATLAB. Модель Polynomial Regression у середовищі RapidMiner виявилася менш точною порівняно із іншими моделями;

- відносні помилки: відносні помилки для всіх моделей виявилися високими, проте модель Neural Network мала найменший відсоток помилки;
- інші показники: модель Polynomial Regression має найвищі значення MSE, що вказує на значну кількість квадратичних помилок. Neural Network вирізняється найвищим показником кореляції серед усіх розглянутих моделей.

Загальні висновки: модель Neural Network є найбільш ефективною для аналізу та прогнозування закупівельних даних завдяки високій точності прогнозів та сильній кореляції. З урахуванням контексту задачі та вимог, модель Linear Regression може бути використана як альтернатива. Було доведено, що застосування сучасних інструментів data mining та машинного навчання дозволить досягти значного прискорення обробки даних.

У результаті виконання роботи проведена розробка технічного завдання на створення системи управління закупівлями. Розроблено удосконалений бізнес-процес управління закупівлями, сформовані вимоги до програмного забезпечення для системи управління закупівлями. Впровадження системи на підприємстві повинно значно поліпшити процеси управління закупівлями.

Проведено економічне обґрунтування запропонованих програмних рішень для системи управління закупівлями з використанням методів машинного навчання, доведено економічну доцільність проведених досліджень та розробки програмних рішень.

Отже, дослідження обґрунтувало необхідність змін в управлінський процес закупівлями. Доведено значущість використання алгоритмів машинного навчання для розв'язання задач управління закупівлями, визначено переваги формування рекомендацій на основі розв'язків цих задач. Запропоновано використовувати спроектовану систему на алгоритмах машинного навчання, для підвищення ефективності процесів та рішень.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Терещенко В.М., Бугайов А.Д. Алгоритми машинного навчання у контексті великих даних. *Штучний інтелект*, 3, 2018. С. 80-86. URL: <http://dspace.nbu.gov.ua/handle/123456789/162446>
2. Жиркова А.П., Ігнатенко О.П. Аналіз методів машинного навчання в задачі класифікації документів. *Проблеми програмування*, 4, 2020. С. 81-87. URL: <https://doi.org/10.15407/pp2020.04.081>
3. Резниченко Р.В., Тимашова Л.А. Оптимізація прийняття рішень для закупівель і поставок на віртуальному підприємстві. *Економіко-математичне моделювання соціально-економічних систем. Збірник наукових праць*, 18, 2013. С. 200-211. URL: <http://dspace.nbu.gov.ua/handle/123456789/83564>
4. Кононова К.Ю. Машинне навчання: методи та моделі: підручник для бакалаврів, магістрів та докторів філософії спеціальності 051 «Економіка». ХНУ імені В. Н. Каразіна, 2020. 301 с.
5. Мальцев А.Ю. Огляд принципів глибокого навчання як динамічної теорії штучного інтелекту. *Інформатика, обчислювальна техніка та автоматизація*, 6, 2021. С. 97-102. URL: <https://doi.org/10.32838/2663-5941/2021.6/16>
6. Aston Zhang, Zack C. Lipton, Mu Li, Alex J. Smola. Dive into Deep Learning. *Interactive deep learning book*. URL: <https://d2l.ai/index.html#jupyter>.
7. Arielle Kushner. Artificial Intelligence in Procurement. URL: <https://www.approve.com/blog/artificial-intelligence-in-procurement/>
8. Ines Schulze-Horn, Sabrina Hueren, Paul Scheffler & Holger Schiele. Artificial Intelligence in Purchasing: Facilitating Mechanism Design-based Negotiations. *Applied artificial intelligence*, 8(34), 2020. P. 618-642. URL: <https://www.tandfonline.com/doi/full/10.1080/08839514.2020.1749337>
9. Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar. Foundations of Machine Learning. *second edition*. The MIT Press, Cambridge, Massachusetts,

London, England. 2018. 486 с. URL: <https://www.dbooks.org/foundations-of-machine-learning-0262039400/read/>

10. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя : КПУ. 2011. 268 с. URL: https://moodle.znu.edu.ua/pluginfile.php/464907/mod_folder/intro/Посібник_1.1.pdf

11. Штовба С. Д., Козачко О. М. Machine Learning: стартовий курс: електронний навчальний посібник. Вінниця : ВНТУ. 2020. – 81 с. URL: https://www.researchgate.net/publication/338924246_Machine_Learning_startovij_kurs

12. Nils J. Nilsson. Introduction to machine learning. Stanford, CA 94305. Robotics Laboratory Department of Computer Science Stanford University. 1998. 188 p. URL: <https://ai.stanford.edu/~nilsson/MLBOOK.pdf>

13. Miroslav Kubat. An Introduction to Machine Learning. Second Edition. Springer. 2017. 348 p. URL: <https://www.dbooks.org/an-introduction-to-machine-learning-3319639137/read/>

14. Charu C. Aggarwal. Neural Networks and Deep Learning. Springer. 2018. 512 p. URL: <https://www.dbooks.org/neural-networks-and-deep-learning-3319944630/>

15. Richard S. Sutton, Andrew G. Barto. Reinforcement Learning. An Introduction. Second Edition. The MIT Press, Cambridge Massachusetts, London England. URL : <https://www.dbooks.org/reinforcement-learning-0262039249/>

16. Ankur Moitra. Algorithmic Aspects of Machine Learning. E-publishing. URL : <https://www.dbooks.org/algorithmic-aspects-of-machine-learning-5656491604/>

17. Ariruna Dasgupta, Asoke Nath. Classification of Machine Learning Algorithms. Article. International Journal of Innovative Research in Advanced Engineering. 2016. 7 p. URL: https://www.researchgate.net/publication/297775323_Classification_of_Machine_Learning_Algorithms

18. Vishal Sharma. Survey of Classification Algorithms and Various Model Selection Methods. Article. Journal of Machine Learning Research. 2000. 14 p. URL : https://www.academia.edu/37254443/Classification_Algorithms_in_Machine_Learning
19. Haoyuan Tan. Machine Learning Algorithm for Classification. Article. International Conference on Big Data and Intelligent Algorithms. 2021. 6 p. URL: https://www.researchgate.net/publication/353814583_Machine_Learning_Algorithm_for_Classification
20. Aized Amin, Arshad Awan. Classification Techniques in Machine Learning: Applications and Issues. Article. Journal of Basic & Applied Sciences. 2017. 7 p. URL : https://www.academia.edu/34405782/Classification_Techniques_in_Machine_Learning_Applications_and_Issues
21. Shai Shalev-Shwartz, Shai Ben-David. Understanding Machine Learning. From Theory to Algorithms. Cambridge University Press. 2014. 449 p. URL : <https://www.dbooks.org/understanding-machine-learning-1107057132/>
22. Ricardo Portilla, Brenner Heintz, Denny Lee. The Big Book of Machine Learning Use Cases. Your complete how-to guide to putting ML to work - plus code samples and notebooks. eBook. 2019. 78 p. URL : <https://www.dbooks.org/the-big-book-of-machine-learning-use-cases-5664978964/>
23. IIBA Global Business Analysis Core Standard. A Companion to A Guide to the Business Analysis Body of Knowledge® (BABOK® Guide). Version 3. 2015. 615 p. URL: [IIBA Global Business Analysis Core Standard](#)
24. PMBOK® Guide. Seventh Edition. 2021. 370 p. URL: [PMBOK Guide | Project Management Institute \(pmi.org\)](#)
25. Data Analysis. MATLAB® Documentation. 2023. 126 p. URL: [MATLAB Data Analysis \(mathworks.com\)](#)
26. App Building. MATLAB® Documentation. 2023. 514 p. URL: https://www.mathworks.com/help/pdf_doc/matlab/creating_guis.pdf

27. Data Import and Export. MATLAB® Documentation. 2023. 624 p. URL: https://www.mathworks.com/help/pdf_doc/matlab/import_export.pdf
28. RapidMiner Documentation. 2023. URL: [Getting started with RapidMiner Studio - RapidMiner Documentation](#)
29. Neural Designer. User's guide. 2023. URL: <https://www.neuraldesigner.com/learning/user-guide/user-guide/>
30. Business Process Model and Notation (BPMN). Version 2.0. Object Management Group, Business Process Model and Notation 2011. 538 p. URL: <https://www.omg.org/spec/BPMN/2.0/PDF>
31. Grady Booch, James Rumbaugh, Ivar Jacobson. The Unified Modeling Language. User Guide Second edition. Addison-Wesley. 2005. 493 p. URL: [\(PDF\) Unified Modeling Language User Guide, The \(2nd Edition\) \(Addison-Wesley Object Technology Series\) \(researchgate.net\)](#)
32. Karl Wieggers. Successful Business Analysis Consulting. Strategies and Tips for Going It Alone. J. Ross Publishing. 2019. 329 p. URL: [Successful Business Analysis Consulting Strategies and Tips For Going It Alone \(Karl Wieggers\) \(Z-Library\) | PDF | Business Analysis | Consultant \(scribd.com\)](#)
33. Середня зарплата в Україні. Мінфін. URL: [Середня зарплата в Україні \[2024\] ► Середня зарплата в Києві та регіонах \(minfin.com.ua\)](#)
34. COCOMO II - Constructive Cost Model. URL: <http://softwarecost.org/tools/COCOMO/>
35. В. В. Ровенська. Виконання економічної частини дипломних проектів. Методичні вказівки. Краматорськ. ДДМА. 2019. 75 с.

ДОДАТОК А. ВІДОМІСТЬ РОБОТИ

п/п	Назва документу	Формат	Найменування об'єкту або виробу	Кількість сторінок
1	Пояснювальна записка	A4	КЦТПАР.122-22-2м.01.00.КР.ПЗ	170
2	Актуальність дослідження	Ел-на	КЦТПАР.122-22-2м.02.00.КР.ПЛ	1
3	Мета та об'єкт дослідження	Ел-на	КЦТПАР.122-22-2м.03.00.КР.ПЛ	1
4	Кореляційна матриця закупівельних даних	Ел-на	КЦТПАР.122-22-2м.04.00.КР.ПЛ	1
5	Закупівельні дані для класифікації	Ел-на	КЦТПАР.122-22-2м.05.00.КР.ПЛ	1
6	Результати використання моделей класифікації у середовищі RapidMiner	Ел-на	КЦТПАР.122-22-2м.06.00.КР.ПЛ	2
7	Результати використання моделей класифікації у MATLAB	Ел-на	КЦТПАР.122-22-2м.07.00.КР.ПЛ	1
8	Закупівельні дані для прогнозування	Ел-на	КЦТПАР.122-22-2м.08.00.КР.ПЛ	1
9	Результати прогнозування у середовищі RapidMiner	Ел-на	КЦТПАР.122-22-2м.09.00.КР.ПЛ	1
10	Результати прогнозування у MATLAB	Ел-на	КЦТПАР.122-22-2м.10.00.КР.ПЛ	1
11	Типовий процес закупівлі	Ел-на	КЦТПАР.122-22-2м.11.00.КР.ПЛ	1
12	Бізнес- процес управління закупівлями	Ел-на	КЦТПАР.122-22-2м.12.00.КР.ПЛ	1
13	Контекстна діаграма Системи управління закупівлями	Ел-на	КЦТПАР.122-22-2м.13.00.КР.ПЛ	1
14	Діаграма використання	Ел-на	КЦТПАР.122-22-2м.14.00.КР.ПЛ	1
15	Діаграма потоків даних для Системи управління закупівлями	Ел-на	КЦТПАР.122-22-2м.15.00.КР.ПЛ	1
16	Економічне обґрунтування проєкту	Ел-на	КЦТПАР.122-22-2м.16.00.КР.ПЛ	1
17	Висновки	Ел-на	КЦТПАР.122-22-2м.17.00.КР.ПЛ	3

ДОДАТОК Б. ПРИКЛАД КОДУ МОВИ MATLAB ДЛЯ МОДЕЛІ ENSEMBLE BAGGED TREE

```
function [trainedClassifier, validationAccuracy] = trainClassifier(trainingData)
% [trainedClassifier, validationAccuracy] = trainClassifier(trainingData)
% Returns a trained classifier and its accuracy. This code recreates the
% classification model trained in Classification Learner app. Use the
% generated code to automate training the same model with new data, or to
% learn how to programmatically train models.
%
% Input:
%   trainingData: A table containing the same predictor and response
%   columns as those imported into the app.
%
% Output:
%   trainedClassifier: A struct containing the trained classifier. The
%   struct contains various fields with information about the trained
%   classifier.
%
%   trainedClassifier.predictFcn: A function to make predictions on new
%   data.
%
%   validationAccuracy: A double representing the validation accuracy as
%   a percentage. In the app, the Models pane displays the validation
%   accuracy for each model.
%
% Use the code to train the model with new data. To retrain your
% classifier, call the function from the command line with your original
% data or new data as the input argument trainingData.
%
% For example, to retrain a classifier trained with the original data set
% T, enter:
% [trainedClassifier, validationAccuracy] = trainClassifier(T)
%
% To make predictions with the returned 'trainedClassifier' on new data T2,
% use
```

```

% [yfit,scores] = trainedClassifier.predictFcn(T2)
%
% T2 must be a table containing at least the same predictor columns as used
% during training. For details, enter:
% trainedClassifier.HowToPredict
% Auto-generated by MATLAB on 19-Dec-2023 23:19:50
% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'BU', 'ReqDate', 'Urgency', 'Material', 'MatGroup', 'SPP', 'FinPosition', 'Qneeds',
'Unit', 'Price', 'PriceUnit', 'Currency', 'Duration', 'Type', 'Direction'};
predictors = inputTable(:, predictorNames);
response = inputTable.TenderType;
isCategoricalPredictor = [false, false, false, false, false, false, false, false, false, false, false, false, false, false, false];
classNames = [1; 2; 3; 4; 5; 6];
% Train a classifier
% This code specifies all the classifier options and trains the classifier.
template = templateTree(...
    'MaxNumSplits', 19955, ...
    'NumVariablesToSample', 'all');
classificationEnsemble = fitcensemble(...
    predictors, ...
    response, ...
    'Method', 'Bag', ...
    'NumLearningCycles', 30, ...
    'Learners', template, ...
    'ClassNames', classNames);
% Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
ensemblePredictFcn = @(x) predict(classificationEnsemble, x);
trainedClassifier.predictFcn = @(x) ensemblePredictFcn(predictorExtractionFcn(x));
% Add additional fields to the result struct
trainedClassifier.RequiredVariables = {'BU', 'ReqDate', 'Urgency', 'Material', 'MatGroup', 'SPP',
'FinPosition', 'Qneeds', 'Unit', 'Price', 'PriceUnit', 'Currency', 'Duration', 'Type', 'Direction'};
trainedClassifier.ClassificationEnsemble = classificationEnsemble;

```

```
trainedClassifier.About = 'This struct is a trained model exported from Classification Learner R2023b.';
```

```
trainedClassifier.HowToPredict = sprintf('To make predictions on a new table, T, use: \n [yfit,scores] = c.predictFcn(T) \nreplacing "c" with the name of the variable that is this struct, e.g. "trainedModel". \n \nThe table, T, must contain the variables returned by: \n c.RequiredVariables \nVariable formats (e.g. matrix/vector, datatype) must match the original training data. \nAdditional variables are ignored. \n \nFor more information, see <a href="matlab:helpview(fullfile(docroot, "stats", "stats.map"), "appclassification_exportmodeltoworkspace")>How to predict using an exported model</a>');
```

```
    % Extract predictors and response
```

```
    % This code processes the data into the right shape for training the  
    % model.
```

```
    inputTable = trainingData;
```

```
    predictorNames = {'BU', 'ReqDate', 'Urgency', 'Material', 'MatGroup', 'SPP', 'FinPosition', 'Qneeds',  
'Unit', 'Price', 'PriceUnit', 'Currency', 'Duration', 'Type', 'Direction'};
```

```
    predictors = inputTable(:, predictorNames);
```

```
    response = inputTable.TenderType;
```

```
    isCategoricalPredictor = [false, false, false, false, false, false, false, false, false, false, false,  
false, false, false];
```

```
    classNames = [1; 2; 3; 4; 5; 6];
```

```
    % Perform cross-validation
```

```
    partitionedModel = crossval(trainedClassifier.ClassificationEnsemble, 'Kfold', 10);
```

```
    % Compute validation predictions
```

```
    [validationPredictions, validationScores] = kfoldPredict(partitionedModel);
```

```
    % Compute validation accuracy
```

```
    validationAccuracy = 1 - kfoldLoss(partitionedModel, 'LossFun', 'ClassifError');
```

ДОДАТОК В. ПРИКЛАД КОДУ МОВИ МАТЛАВ ДЛЯ МОДЕЛІ DECISION TREE

```
function [trainedClassifier, validationAccuracy] = trainClassifier(trainingData)
% [trainedClassifier, validationAccuracy] = trainClassifier(trainingData)
% Returns a trained classifier and its accuracy. This code recreates the
% classification model trained in Classification Learner app. Use the
% generated code to automate training the same model with new data, or to
% learn how to programmatically train models.
%
% Input:
%   trainingData: A table containing the same predictor and response
%   columns as those imported into the app.
%
% Output:
%   trainedClassifier: A struct containing the trained classifier. The
%   struct contains various fields with information about the trained
%   classifier.
%
%   trainedClassifier.predictFcn: A function to make predictions on new
%   data.
%
%   validationAccuracy: A double representing the validation accuracy as
%   a percentage. In the app, the Models pane displays the validation
%   accuracy for each model.
%
% Use the code to train the model with new data. To retrain your
% classifier, call the function from the command line with your original
% data or new data as the input argument trainingData.
%
% For example, to retrain a classifier trained with the original data set
% T, enter:
% [trainedClassifier, validationAccuracy] = trainClassifier(T)
%
% To make predictions with the returned 'trainedClassifier' on new data T2,
% use
% [yfit,scores] = trainedClassifier.predictFcn(T2)
```

```

%
% T2 must be a table containing at least the same predictor columns as used
% during training. For details, enter:
% trainedClassifier.HowToPredict
% Auto-generated by MATLAB on 19-Dec-2023 23:22:25
% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'BU', 'ReqDate', 'Urgency', 'Material', 'MatGroup', 'SPP', 'FinPosition', 'Qneeds',
'Unit', 'Price', 'PriceUnit', 'Currency', 'Duration', 'Type', 'Direction'};
predictors = inputTable(:, predictorNames);
response = inputTable.TenderType;
isCategoricalPredictor = [false, false, false, false, false, false, false, false, false, false, false, false, false, false, false];
classNames = [1; 2; 3; 4; 5; 6];
% Train a classifier
% This code specifies all the classifier options and trains the classifier.
classificationTree = fitctree(...
    predictors, ...
    response, ...
    'SplitCriterion', 'gdi', ...
    'MaxNumSplits', 100, ...
    'Surrogate', 'off', ...
    'ClassNames', classNames);
% Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
treePredictFcn = @(x) predict(classificationTree, x);
trainedClassifier.predictFcn = @(x) treePredictFcn(predictorExtractionFcn(x));
% Add additional fields to the result struct
trainedClassifier.RequiredVariables = {'BU', 'ReqDate', 'Urgency', 'Material', 'MatGroup', 'SPP',
'FinPosition', 'Qneeds', 'Unit', 'Price', 'PriceUnit', 'Currency', 'Duration', 'Type', 'Direction'};
trainedClassifier.ClassificationTree = classificationTree;
trainedClassifier.About = 'This struct is a trained model exported from Classification Learner
R2023b.';
trainedClassifier.HowToPredict = sprintf('To make predictions on a new table, T, use: \n
[yfit,scores] = c.predictFcn(T) \nreplacing "c" with the name of the variable that is this struct, e.g.
"trainedModel". \n \nThe table, T, must contain the variables returned by: \n c.RequiredVariables

```

\nVariable formats (e.g. matrix/vector, datatype) must match the original training data. \nAdditional variables are ignored. \n\nFor more information, see [How to predict using an exported model](matlab:helpview(fullfile(docroot,).);

```

    % Extract predictors and response
    % This code processes the data into the right shape for training the
    % model.
    inputTable = trainingData;
    predictorNames = {'BU', 'ReqDate', 'Urgency', 'Material', 'MatGroup', 'SPP', 'FinPosition', 'Qneeds',
'Unit', 'Price', 'PriceUnit', 'Currency', 'Duration', 'Type', 'Direction'};
    predictors = inputTable(:, predictorNames);
    response = inputTable.TnderType;
    isCategoricalPredictor = [false, false, false, false, false, false, false, false, false, false, false, false, false, false, false];
    classNames = [1; 2; 3; 4; 5; 6];
    % Perform cross-validation
    partitionedModel = crossval(trainedClassifier.ClassificationTree, 'KFold', 10);
    % Compute validation predictions
    [validationPredictions, validationScores] = kfoldPredict(partitionedModel);
    % Compute validation accuracy
    validationAccuracy = 1 - kfoldLoss(partitionedModel, 'LossFun', 'ClassifError');

```

ДОДАТОК Г. ПРИКЛАД КОДУ МОВИ MATLAB ДЛЯ МОДЕЛІ LINEAR REGRESSION

```

function [trainedModel, validationRMSE] = trainRegressionModel(trainingData)
% [trainedModel, validationRMSE] = trainRegressionModel(trainingData)
% Returns a trained regression model and its RMSE. This code recreates the
% model trained in Regression Learner app. Use the generated code to
% automate training the same model with new data, or to learn how to
% programmatically train models.
%
% Input:
%   trainingData: A table containing the same predictor and response
%   columns as those imported into the app.
%
% Output:
%   trainedModel: A struct containing the trained regression model. The
%   struct contains various fields with information about the trained
%   model.
%
%   trainedModel.predictFcn: A function to make predictions on new data.
%
%   validationRMSE: A double representing the validation RMSE. In the
%   app, the Models pane displays the validation RMSE for each model.
%
% Use the code to train the model with new data. To retrain your model,
% call the function from the command line with your original data or new
% data as the input argument trainingData.
%
% For example, to retrain a regression model trained with the original data
% set T, enter:
% [trainedModel, validationRMSE] = trainRegressionModel(T)
%
% To make predictions with the returned 'trainedModel' on new data T2, use
% yfit = trainedModel.predictFcn(T2)
%
% T2 must be a table containing at least the same predictor columns as used
% during training. For details, enter:
%   trainedModel.HowToPredict
% Auto-generated by MATLAB on 24-Dec-2023 00:08:55

```

```

% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'TnderType', 'ReqDateM'};
predictors = inputTable(:, predictorNames);
response = inputTable.PRsQnt;
isCategoricalPredictor = [false, false];
% Train a regression model
% This code specifies all the model options and trains the model.
concatenatedPredictorsAndResponse = predictors;
concatenatedPredictorsAndResponse.PRsQnt = response;
linearModel = fitlm(...
    concatenatedPredictorsAndResponse, ...
    'linear', ...
    'RobustOpts', 'off');
% Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
linearModelPredictFcn = @(x) predict(linearModel, x);
trainedModel.predictFcn = @(x) linearModelPredictFcn(predictorExtractionFcn(x));
% Add additional fields to the result struct
trainedModel.RequiredVariables = {'TnderType', 'ReqDateM'};
trainedModel.LinearModel = linearModel;
trainedModel.About = 'This struct is a trained model exported from Regression Learner R2023b.';
trainedModel.HowToPredict = sprintf('To make predictions on a new table, T, use: \n yfit = c.predictFcn(T)
\nreplacing "c" with the name of the variable that is this struct, e.g. "trainedModel". \n \nThe table, T, must contain
the variables returned by: \n c.RequiredVariables \nVariable formats (e.g. matrix/vector, datatype) must match the
original training data. \nAdditional variables are ignored. \n \nFor more information, see <a
href="matlab:helpview(fullfile(docroot, "stats", "stats.map"), "appgression_exportmodeltoworkspace")>How to
predict using an exported model</a>');
% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'TnderType', 'ReqDateM'};
predictors = inputTable(:, predictorNames);
response = inputTable.PRsQnt;
isCategoricalPredictor = [false, false];
% Perform cross-validation
KFolds = 10;
cvp = cvpartition(size(response, 1), 'KFold', KFolds);
% Initialize the predictions to the proper sizes

```

```

validationPredictions = response;
for fold = 1:KFolds
    trainingPredictors = predictors(cvp.training(fold), :);
    trainingResponse = response(cvp.training(fold), :);
    foldIsCategoricalPredictor = isCategoricalPredictor;
    % Train a regression model
    % This code specifies all the model options and trains the model.
    concatenatedPredictorsAndResponse = trainingPredictors;
    concatenatedPredictorsAndResponse.PRsQnt = trainingResponse;
    linearModel = fitlm(...
        concatenatedPredictorsAndResponse, ...
        'linear', ...
        'RobustOpts', 'off');
    % Create the result struct with predict function
    linearModelPredictFcn = @(x) predict(linearModel, x);
    validationPredictFcn = @(x) linearModelPredictFcn(x);
    % Add additional fields to the result struct
    % Compute validation predictions
    validationPredictors = predictors(cvp.test(fold), :);
    foldPredictions = validationPredictFcn(validationPredictors);
    % Store predictions in the original order
    validationPredictions(cvp.test(fold), :) = foldPredictions;
end
% Compute validation RMSE
isNotMissing = ~isnan(validationPredictions) & ~isnan(response);
validationRMSE = sqrt(nansum(( validationPredictions - response ).^2) / numel(response(isNotMissing) ));

```

ДОДАТОК Д. ПРИКЛАД КОДУ МОВИ МАТЛАВ ДЛЯ МОДЕЛІ NEURAL NETWORK

```
function [trainedModel, validationRMSE] = trainRegressionModel(trainingData)
% [trainedModel, validationRMSE] = trainRegressionModel(trainingData)
% Returns a trained regression model and its RMSE. This code recreates the
% model trained in Regression Learner app. Use the generated code to
% automate training the same model with new data, or to learn how to
% programmatically train models.
%
% Input:
%   trainingData: A table containing the same predictor and response
%   columns as those imported into the app.
%
% Output:
%   trainedModel: A struct containing the trained regression model. The
%   struct contains various fields with information about the trained
%   model.
%
%   trainedModel.predictFcn: A function to make predictions on new data.
%
%   validationRMSE: A double representing the validation RMSE. In the
%   app, the Models pane displays the validation RMSE for each model.
%
% Use the code to train the model with new data. To retrain your model,
% call the function from the command line with your original data or new
% data as the input argument trainingData.
%
% For example, to retrain a regression model trained with the original data
% set T, enter:
% [trainedModel, validationRMSE] = trainRegressionModel(T)
%
% To make predictions with the returned 'trainedModel' on new data T2, use
```

```

% yfit = trainedModel.predictFcn(T2)
%
% T2 must be a table containing at least the same predictor columns as used
% during training. For details, enter:
% trainedModel.HowToPredict
% Auto-generated by MATLAB on 24-Dec-2023 00:09:51
% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'TnderType', 'ReqDateM'};
predictors = inputTable(:, predictorNames);
response = inputTable.PRsQnt;
isCategoricalPredictor = [false, false];
% Train a regression model
% This code specifies all the model options and trains the model.
regressionNeuralNetwork = fitrnet(...
    predictors, ...
    response, ...
    'LayerSizes', [3 9 1], ...
    'Activations', 'sigmoid', ...
    'Lambda', 0, ...
    'IterationLimit', 1000, ...
    'Standardize', true);
% Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
neuralNetworkPredictFcn = @(x) predict(regressionNeuralNetwork, x);
trainedModel.predictFcn = @(x) neuralNetworkPredictFcn(predictorExtractionFcn(x));
% Add additional fields to the result struct
trainedModel.RequiredVariables = {'TnderType', 'ReqDateM'};
trainedModel.RegressionNeuralNetwork = regressionNeuralNetwork;
trainedModel.About = 'This struct is a trained model exported from Regression Learner
R2023b.';
trainedModel.HowToPredict = sprintf('To make predictions on a new table, T, use: \n yfit
= c.predictFcn(T) \nreplacing "c" with the name of the variable that is this struct, e.g.

```

"trainedModel". \n \n The table, T, must contain the variables returned by: \n c.RequiredVariables \n Variable formats (e.g. matrix/vector, datatype) must match the original training data. \n Additional variables are ignored. \n \n For more information, see How to predict using an exported model.);

```
% Extract predictors and response
```

```
% This code processes the data into the right shape for training the
% model.
```

```
inputTable = trainingData;
```

```
predictorNames = {'TnderType', 'ReqDateM'};
```

```
predictors = inputTable(:, predictorNames);
```

```
response = inputTable.PRsQnt;
```

```
isCategoricalPredictor = [false, false];
```

```
% Perform cross-validation
```

```
partitionedModel = crossval(trainedModel.RegressionNeuralNetwork, 'KFold', 10);
```

```
% Compute validation predictions
```

```
validationPredictions = kfoldPredict(partitionedModel);
```

```
% Compute validation RMSE
```

```
validationRMSE = sqrt(kfoldLoss(partitionedModel, 'LossFun', 'mse'));
```

ДОДАТОК Е. BUSINESS REQUIREMENTS SPECIFICATION

Затвердження

Власник бізнесу/Зацікавлена сторона	Затверджено / Не затверджено	Підпис	Посада	Дата
Дробот В.В.	затверджено		Директор із закупівель	
Кріпак С.А.	затверджено		Бізнес аналітик	

Список розповсюджувачів

Усі, хто перераховані нижче повинні допомогти розповсюдженню ресурсів та усвідомлюють що наступну інформацію так детально розглянуто з метою погодження.

Ім'я	Посада
Дробот В.В.	Директор із закупівель
Кріпак С.А.	Бізнес аналітик

Автор: Сергій Кріпак, Бізнес аналітик,

Версія: 01

Дата: 10 квітня 2026

Контроль версій

Версія	Ім'я	Посада	Контактна Інформація	Дата	Підсумок Змін
01	Сергій Кріпак	Бізнес аналітик	serhii.kripak@mipolytech.education	01.11.2023	

Деталі Документа

Ім'я Документа	Місцезнаходження Документа
Business Requirements Specification	Кріпак С.А.

Е.1. Вступ

Е.1.1. Призначення

Призначення цього документа уточнити вимоги до бізнесу, що було надано зацікавленим сторонам для проекту Система управління закупівлями на основі машинного навчання.

Е.1.2. Цільова аудиторія

Цей документ призначений для читання:

- Спонсорами проекту, Володарями бізнесу та Делегатами проекту та усіма іншими ключовими зацікавленими сторонами для забезпечення того, щоб усі ділові задачі були розглянуті та для підписання вимог цього проекту.
- Командою розробників для подальшого використання як бази з технічними специфікаціями, тестування дизайну та виконання тестового плану.

Е.1.3. Визначення, термінологія, скорочення та аббревіатури

Визначення, термінологія, скорочення та аббревіатури, що використовуються в межах цього проекту, містяться у цій таблиці:

Визначення, термінологія, скорочення та аббревіатури	Визначення/Опис
АС – автоматизована система	організаційно-технічна система, що забезпечує вироблення рішень на основі автоматизації інформаційних процесів
БД – база даних	сукупність даних, організованих відповідно до концепції, яка описує характеристики цих даних і взаємозв'язки між їх елементами, ця сукупність підтримує щонайменше одну з областей застосування
Замовник	
Компанія	ТОВ «МЕТІНВЕСТ ДІДЖИТАЛ»
Матеріали	Матеріальні товари, необхідні для виробництва товарів Компанії, що можуть бути виміряні фізичними величинами
НВ – навчальна вибірка	деяка підмножина досліджуваної загальної (генеральної) сукупності. На основі вивчення навчальної вибірки висновки про генеральну сукупність
Послуги	Не матеріальні товари, чи спеціалізовані дії необхідні для виробництва товарів Компанії, що можуть бути виміряні умовними одиницями
Університет	ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА»
Система управління закупівлями	Автоматизована система для обробки та аналізу накопичених даних та знань з метою прогнозування закупівель наступних періодів і прийняття рішень

	щодо забезпечення своєчасних закупівель та поставок
Служба закупівель	Підрозділ на який покладено обов'язки з управління та здійснення процесу закупівель
ШІ - Штучний інтелект	Система в основі якої лежить машинне навчання
Машинне навчання	Виявлення закономірностей в історичних даних для прогнозування очікуваних закупівель наступних періодів

Е.1.4. Припущення

Передбачається, що система управління закупівлями буде використовуватись службою закупівель підприємства для прогнозування об'ємів закупівель майбутніх періодів (рік, квартал, місяць, тиждень) та необхідних для цього ресурсів (бюджету, співробітників) та орієнтовні строки закупівель.

Е.1.5. Залежності

Система управління закупівлями повинна бути оптимізована для забезпечення ефективного функціонування бізнес-процесів.

Машинне навчання може бути застосовано для прогнозування попиту на матеріали та послуги, виявлення аномальних закупівельних операцій, автоматизованого вибору постачальників та оптимального розподілу ресурсів для закупівель.

Програмні компоненти повинні бути проєктовані та розроблені для забезпечення відповідності системи закупівель та постачання вимогам бізнесу та наукових стандартів.

Е.1.6. Обмеження

Система управління закупівлями повинна відповідати затвердженим вимогам та бізнес-процесам закупівель на підприємстві.

Е.1.7. Підхід до збору вимог

- Інтерв'ювання зацікавлених сторін
- Розроблено процеси "As Is"
- Розроблено шаблони "To Be" бізнес процесів
- Підсумок "To Be" процесів, що було обговорено на семінарах зацікавлених

сторін

Е.1.8. Контекст бізнесу. Контекстна діаграма

Недостатня ефективність в процесі закупівель, що пов'язані із складнощами у виборі постачальників, управлінні ризиками, контролі витрат, не відповідність поставки вимогам, не можливість точно передбачити попит на товари та послуги

Програмний продукт повинен вирішувати ці проблеми шляхом використання алгоритмів машинного навчання, які забезпечують більш точний аналіз та прогнозування даних про постачальників, допомагати встановлювати відповідність між параметрами поставки та вимогами, аналізувати дані та передбачати попит на основі різноманітних факторів.



Рис.Е.1 Система управління закупівлями (контекстна діаграма)

Система управління закупівлями взаємодіє:

- з Системою обліку підприємства для отримання даних щодо закупівель попередніх періодів (Матеріали/Послуги, кількість, ціна, строки, умови тощо);
- з нормативними документами підприємства, щоб відповідати встановленим вимогам до процесу та строкам закупівель;
- із Замовником, інформуючи його про очікувані строки та вартість закупівель та прогнозування виконання замовлень ;
- із Службою Закупівель прогнозуючи очікувані об'єми та вартість закупівель та необхідні ресурси для їх своєчасного виконання.

Е.2. Сфера застосування

Е.2.1. В межах проекту

Для досягнення цілей проекту, поставлених бізнесом, та реалізації максимально можливих матеріальних та нематеріальних переваг (зисків) система має забезпечити:

- забезпечення своєчасності та безперебійності поставок матеріалів та послуг необхідних для виробництва;
- оптимізувати та прогнозувати навантаження на співробітників Служби закупівель;
- автоматизації процесу закупівлі матеріалів, обладнання та послуг;
- знизити витрати на закупівлю, що дозволяють забезпечити оптимальне співвідношення між вартістю та якістю наданого товару або послуги;
- дозволить відслідковувати всі етапи процесу закупівлі, що покращить контроль та управління;
- скоротити строки поставок.

Для цього потрібно буде здійснити широкий спектр заходів, включаючи:

- збір інформації, формування навчальної вибірки;
- маркування даних навчальної вибірки;
- створення нейронної мережі;
- тренування НМ на основі навчальної вибірки;
- тестування моделі ШІ;
- підготовку необхідної звітності та документації.

Е.2.2. За межами проекту

Система планується для використання Службою закупівель підприємства, але вона може бути використана і для потреб Служби реалізації матеріалів та послуг. Також вона може бути корисною в різних організаційних структурах, де необхідно здійснювати закупівлі. Наприклад, у великих корпораціях, державних органах, медичних установах, тощо.

Е.3. Ділова мета, цілі/Завдання

Розробити систему, яка в автоматичному режимі буде здійснювати обробку заявок на закупівлю в системі обліку компанії, та надавати прогноз закупівель у вигляді розгорнутого списку матеріалів та послуг, очікуваної кількості, необхідного бюджету та ресурсів для забезпечення своєчасної закупівлі, укладення договорів та поставок на склади Компанії необхідних для виробництва матеріалів.

Збільшити лояльність Замовників закупівель до Служби закупівель.

Усунення інформаційного вакууму між різними учасниками процесу закупівель.

Мінімізація впливу людського фактору на результати прогнозування, яке в поточний час проводиться вручну.

Збільшення операційної ефективності Служби закупівель.

Е.4. Ключові зацікавлені сторони

Кріпак С.А. – розробник

Дробут В.В. – керівник проєкту зі сторони Компанії

Москаленко В.В. – керівник проєкту зі сторони Університету

Кінцевий користувач – бере участь у опитуваннях, необхідних для написання бізнес вимог.

Е.5. Моделі бізнес-процесів

Процес As is:

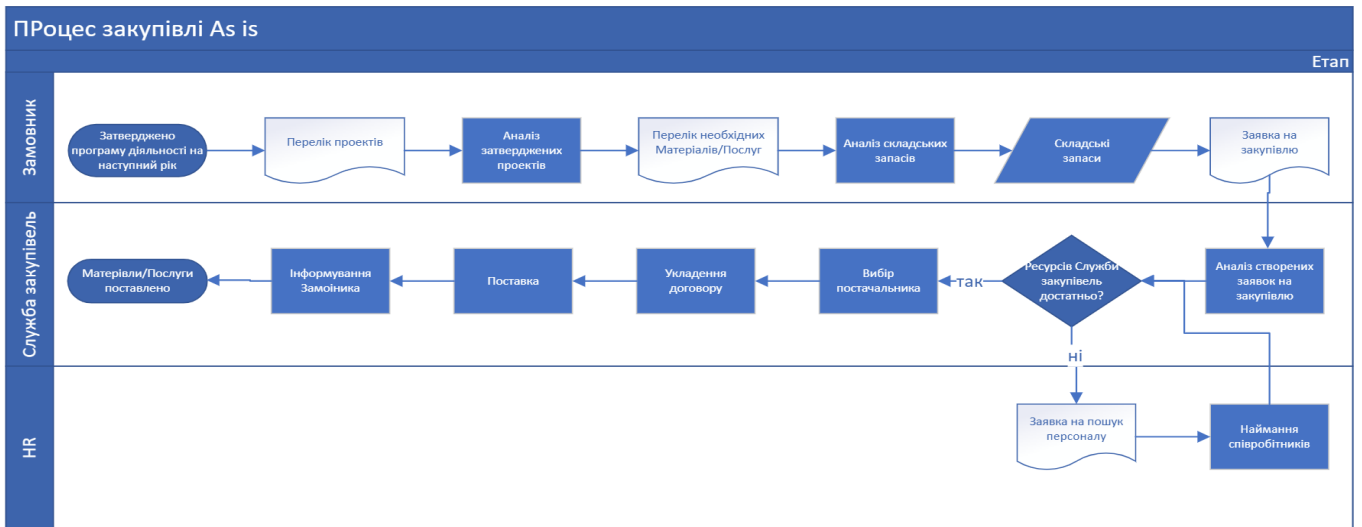


Рис.Е.2 Система управління закупівлями (As is)

Процес To Be:

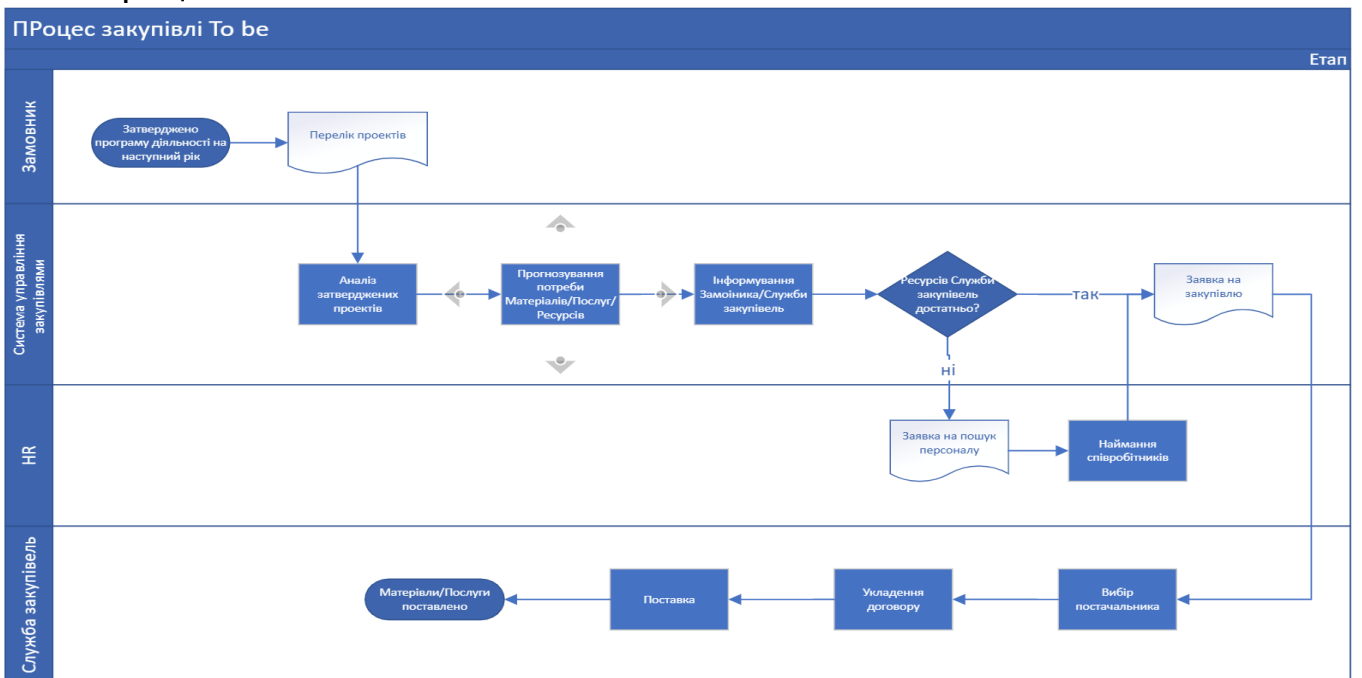


Рис.Е.3 Система управління закупівлями (To be)

Е.6. Вимоги розміщення пріоритетів

Вимоги будуть розміщені згідно пріоритетам, що базуються на техніці MoSCoW, котра розділяє вимоги на наступні категорії:

Рейтинг пріоритетів	Опис
П – Повинен мати (M – Must Have)	Описуються вимоги, що повинні бути задоволені у фінальному представленні рішення для досягнення успіху.
В – Варто було б мати (S – Should Have)	Представляє високо-пріоритетні деталі (пункти), що повинні бути добавлені у рішення, якщо це можливо. Дуже часто це вирішальні вимоги, проте кожен з них може бути задоволений іншим шляхом, якщо суворо необхідно.
М – Можливо мати (C – Could Have)	Описуються вимоги, які вважаються бажаними, але не обов'язковими. Вони будуть включені, якщо дозволять час і ресурси.
Х – Хотілося б мати (W – Won't Have)	Представляє вимоги, які були погоджені зацікавленими сторонами, що не будуть додаватися до анонсування, проте можуть бути розглянуті у майбутньому.

Е.7. Вимоги бізнесу (BR)

Посилання	Вимоги	Пріоритет	Джерело
BR01	Прогнозування потреб Матеріалів у розрізі Замовників	M	Протокол 1
BR02	Прогнозування потреб Послуг у розрізі Замовників	M	Протокол 1
BR03	Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів	M	Протокол 1
BR04	Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників	M	Протокол 1
BR05	Інформування Замовників та Служби закупівель про очікувані строки виконання заявок та необхідні ресурси	S	Інтерв'ю
BR06	Прогнозування строків виконання етапів процесу закупівлі (вибір постачальника, укладення договору, поставки)	C	Інтерв'ю
BR07	Створення заявок на закупівлю	W	Інтерв'ю
BR08	Створення заявок на наймання персоналу для Служби закупівель за результатами прогнозування	W	Інтерв'ю
BR09	Прогнозування строків виконання окремої заявки поетапно	S	Протокол 1
BR10	Пропозиції щодо переліку Постачальників минулих періодів	C	Інтерв'ю
BR11	Прогноз по виконанню поставок Постачальниками в строк	W	Інтерв'ю

ДОДАТОК Є. SOFTWARE REQUIREMENTS SPECIFICATION

Специфікація вимог до програмного продукту для Системи управління закупівлями

Затверджено версію 1.0
Підготовлено Кріпак С.А.
Метінвест Політехніка
01.01.2024

Журнал версій (Revision History)

Ім'я	Дата	Причина для змін	Версія
Кріпак С.А.	01.01.2024	створення	1.0

Є.1 Введення (Introduction)

Є.1.1 Мета (Purpose)

Даний документ описує вимоги до Системи інтелектуального управління закупівлями, розробленої з метою надання співробітникам Служби закупівель та Служб Замовника даних прогнозу закупівель матеріалів та послуг, виявлення аномальних закупівельних операцій, автоматизованого вибору постачальників та оптимального розподілу ресурсів для закупівель.

Є.1.2 Умовні позначення документів (Document Conventions)

У даному документі використовуються такі умовні позначення:

- Функціональні вимоги (Functional Requirements, FR): Вимоги до функціональності системи.
- Нефункціональні вимоги (Non-functional Requirements, NFR): Вимоги до якості системи, її робочих характеристик та інших аспектів, які не стосуються функціональності.

Є.1.3 Цільова аудиторія і пропозиції з читання (Intended Audience and Reading Suggestions)

Даний документ призначений для команди розробників, тестувальників, менеджерів проектів, бізнес-аналітиків, внутрішніх користувачів, вищого керівництва компанії та інших стейкхолдерів, які беруть участь у розробці, впровадженні та використанні системи. Рекомендується ознайомитися з усіма розділами документа для повного розуміння вимог та обмежень системи.

Є.1.4 Область застосування продукту (Product Scope)

Система буде використовуватися внутрішніми співробітниками Служби закупівель та Служб Замовника компанії для отримання даних прогнозу закупівель матеріалів та послуг, виявлення аномальних закупівельних операцій, автоматизованого вибору постачальників та оптимального розподілу ресурсів для закупівель. Система повинна бути інтегрована з існуючими корпоративними системами та базами даних.

Є.1.5 Посилання (References)

- IEEE стандарт 830-1998: Стандарт для опису вимог до програмного забезпечення.
- Документи, що стосуються корпоративних стандартів, політик та процедур компанії.

- Законодавчі та нормативні акти, які стосуються юридичної сфери та можуть мати вплив на вимоги до системи.

Є.2 Загальний опис (Overall Description)

Є.2.1 Перспектива продукту (Product Perspective)

Система інтелектуального управління закупівлями, призначена для автоматизації процесів прогнозування закупівель матеріалів та послуг, виявлення аномальних закупівельних операцій та оптимального розподілу ресурсів для закупівель, автоматичного створення заявок на закупівлю та на підбір персоналу закупівель на підприємстві. Система спрямована на підвищенні та зменшення часу на прогнозування потреб та об'ємів закупівель, необхідних фінансових та трудових ресурсів, а також підвищення ефективності роботи користувачів.



Рис.Є.1 Контекстна діаграма Системи інтелектуального управління закупівлями

Є.2.2 Особливості продукту (Product Features)

- Прогнозування потреб Матеріалів у розрізі Замовників
- Прогнозування потреб Послуг у розрізі Замовників
- Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів
- Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників
- Інформування Замовників та Служби закупівель про очікувані строки виконання заявок та необхідні ресурси
- Прогнозування строків виконання етапів процесу закупівлі (вибір постачальника, укладення договору, поставки)
- Створення заявок на закупівлю
- Створення заявок на наймання персоналу для Служби закупівель за результатами прогнозування
- Прогнозування строків виконання окремої заявки поетапно
- Пропозиції щодо переліку Постачальників минулих періодів
- Прогноз по виконанню поставок Постачальниками в строк

Є.2.3 Класи та характеристики користувачів (User Classes and Characteristics)

Користувач	Варіант використання
Замовник	Прогнозування строків виконання закупівель по Замовнику
	Прогнозування строків виконання закупівель Матеріалів
	Прогнозування потреб Матеріалів у розрізі Замовників
	Прогнозування потреб Послуг у розрізі Замовників
	Створення заявок на закупівлю
Аналітик	Створення заявок на наймання персоналу
	Прогнозування строків виконання етапів процесу закупівлі (вибір постачальника, укладення договору, поставки)
	Створення заявок на закупівлю
Фахівець закупівель	Створення заявок на закупівлю
	Пропозиції щодо переліку Постачальників
	Прогноз по виконанню поставок Постачальниками в строк
	Прогнозування потреб Послуг у розрізі Замовників
	Прогнозування потреб Матеріалів у розрізі Замовників
Фахівець складу Фахівець постачання	Прогнозування строків виконання закупівель Матеріалів
	Прогнозування строків виконання закупівель Послуг
Фахівець HR	Прогнозування строків виконання закупівель Матеріалів
	Прогнозування строків виконання закупівель Послуг
	Прогнозування потреб Послуг у розрізі Замовників
	Прогнозування потреб Матеріалів у розрізі Замовників
	Створення заявок на закупівлю
Фахівець фінансів	Прогноз по виконанню поставок Постачальниками в строк
	Створення заявок на наймання персоналу

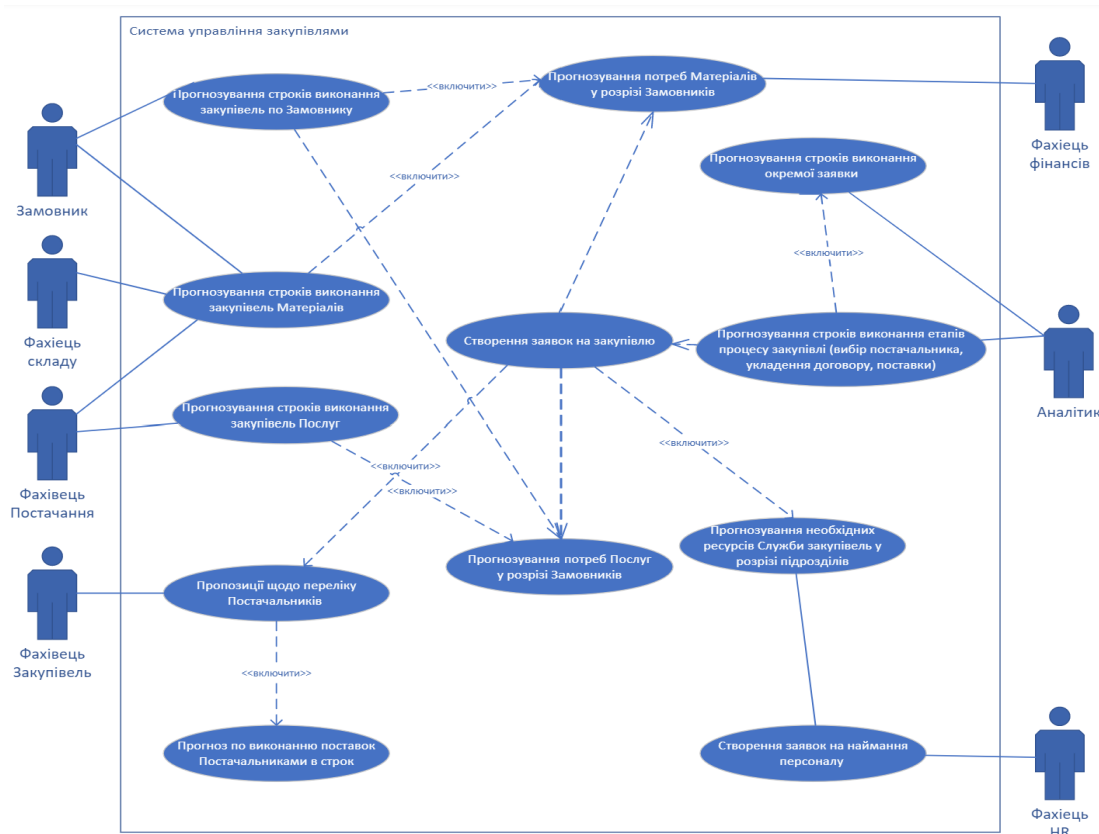


Рис.Є.2.1 Діаграма використання варіантів

Є.2.4 Операційне середовище (Operating Environment)

Система повинна бути сумісна з наступними операційними системами: Windows, macOS. Клієнтська частина повинна працювати на основних веб-браузерах (Google Chrome, Mozilla Firefox, Safari, Microsoft Edge. Також система повинна працювати з програмним продуктом SAP, Microsoft Teams, Microsoft Outlook, Power BI.

Є.2.5 Обмеження щодо проектування та впровадження (Design and Implementation Constraints)

- Система повинна бути сумісна з існуючими інформаційними системами підприємства.
- Система має відповідати діючим нормативним документам підприємства.
- Система повинна відповідати вимогам щодо безпеки даних та конфіденційності інформації.
- Система повинна бути масштабована та легко адаптована до зростання організації та змін в її структурі.

Є.2.6 Документація користувача (User Documentation)

- Посібник користувача: опис функціональності та інструкції з використання системи для різних категорій користувачів.
- Довідкова система: пошукова система та відповіді на часті запитання користувачів.
- Технічна підтримка: контактні дані служби підтримки для вирішення проблем та надання консультацій.

Є.2.7 Припущення та залежності (Assumptions and Dependencies)

Передбачається, що система управління закупівлями буде використовуватись службою закупівель підприємства для прогнозування об'ємів закупівель майбутніх періодів (рік, квартал, місяць, тиждень) та необхідних для цього ресурсів (бюджету, співробітників) та орієнтовні строки закупівель.

Система управління закупівлями повинна бути оптимізована для забезпечення ефективного функціонування бізнес-процесів.

Машинне навчання може бути застосовано для прогнозування попиту на матеріали та послуги, виявлення аномальних закупівельних операцій, автоматизованого вибору постачальників та оптимального розподілу ресурсів для закупівель.

Програмні компоненти повинні бути проєктовані та розроблені для забезпечення відповідності системи закупівель та постачання вимогам бізнесу та наукових стандартів.

Є.3 Вимоги до зовнішнього інтерфейсу (External Interface Requirements)

Є.3.1 Інтерфейси користувача (User Interfaces) -(UINT)

Ідентифікатор вимог до інтерфейсів	Опис
UINT-01	Авторизація користувачів за корпоративним ідентифікатором
UINT-02	Головний екран з можливістю вибору функціоналу та рядком пошуку
UINT-03	Екран функції з фільтрами по напрямкам, строкам та видам закупівель
UINT-04	Формування звітів
UINT-05	Управління користувачами та їх ролями

Є.3.2 Апаратні інтерфейси (Hardware Interfaces) - (HINT))

Система планується використовуватись на наступних типах пристроїв:

- Персональні комп'ютери
- Ноутбуки

Дані між програмним продуктом та апаратними компонентами будуть передаватись через стандартні протоколи зв'язку, такі як HTTP, HTTPS та FTP.

Є.3.3 Інтерфейси програмного забезпечення (Software Interfaces) - (SwINT)

Ідентифікатор	Компонент	Тип інтерфейсу	Протокол	Опис
SwINT-02	Внутрішні системи компанії	REST/SOAP	API-ключі/OAuth 2.0	Підключення до внутрішніх систем компанії за допомогою REST або SOAP протоколів
SwINT-04	Бази даних	DB Driver/Library	Залежить від бази даних	Взаємодія з різними типами баз даних (реляційні та нереляційні) за допомогою драйверів та бібліотек
SwINT-06	Outlook 365	RESTful API	JSON, OAuth 2.0	Інтеграція з сервісом Outlook 365 для надсилання та отримання електронних листів

Є.3.4 Інтерфейси зв'язку (Communications Interfaces) -(CINT)

1. Протоколи зв'язку мережевих серверів: продукт буде використовувати TCP/IP для забезпечення зв'язку між різними компонентами системи, включаючи сервери баз даних, веб-сервери та сервери застосунків.

2. Електронна пошта: продукт буде інтегрований з Outlook 365 для надсилання та отримання електронних листів. Використовуватиметься протокол «Simple Mail Transfer Protocol» (SMTP) для відправки листів та «Internet Message Access Protocol» (IMAP) або «Post Office Protocol» (POP3) для отримання листів.

3. Форматування повідомлень: JSON буде використовуватись як основний формат для обміну даними між різними компонентами системи.

4. Стандарти зв'язку: продукт буде використовувати загальноприйняті стандарти зв'язку, такі як HTTP/HTTPS, SMTP, IMAP, POP3, та TCP/IP для підтримки різних аспектів взаємодії.

Є.4 Особливості системи (System Features)

Рейтинг пріоритетів	Опис
M – Повинен мати	Опис вимог, що повинні бути задоволені у фінальному представленні рішення для досягнення успіху.
S – Варто було б мати	Представляє високо-пріоритетні деталі (пункти), що повинні бути добавлені у рішення, якщо це можливо. Дуже часто це вирішальні вимоги, проте кожен з них може бути задоволений іншим шляхом, якщо суворо необхідно.
C – Можливо мати	Описує вимоги котрі вважаються бажаними, але не обов'язковими. Вони будуть включені, якщо дозволять час і ресурси.
W – Хотілося б мати	Представляє вимоги, які були погоджені зацікавленими сторонами, що не будуть додаватися до анонсування, проте можуть бути розглянуті у подальшому.

ID	Особливість системи	Пріоритет
FR01	Прогнозування потреб Матеріалів та Послуг у розрізі Замовників	M
FR02	Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів	M
FR03	Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників	M
FR04	Створення заявок на закупівлю	W
FR05	Пропозиції щодо переліку Постачальників минулих періодів	C
FR06	Прогноз по виконанню поставок Постачальниками в строк	W

Є.4.1 Функція системи 1 (System Feature 1) Прогнозування потреб Матеріалів та Послуг у розрізі Замовників.

Основні модулі:

1. Прогнозування потреб Матеріалів в розрізі Замовників
2. Прогнозування потреб Послуг в розрізі Замовників
3. Прогнозування потреб Матеріалів в розрізі Постачальників
4. Прогнозування потреб Послуг в розрізі Постачальників
5. Прогнозування орієнтовного бюджету закупівель
6. Формування звітів

Є.4.1.1 Опис і пріоритет (Description and Priority)

Ця функція за допомогою генетичних алгоритмів машинного навчання на основі аналізу закупівель за попередні роки, моделює очікувані закупівлі та їх вартість і об'єми, поточного та майбутнього періодів. Прогнозування здійснюється в розрізі конкретних матеріалів, послуг, Замовників, Постачальників, груп матеріалів.

Пріоритет: Високий

Є.4.1.2 Послідовності стимулів/відповідей (Stimulus/ Response Sequences)

- Запит від користувача. *Stimulus* Користувач формує в системі запит, який включає основні параметри прогнозування: Базовий період, прогнозований період, напрямок та обмеження прогнозування (по замовнику, по постачальнику) мета (визначення об'ємів/вартості).
- Аналіз запиту. *Response* Система аналізує запит, та формує запити до бази даних.
- Направлення запиту до бази даних. *Stimulus* Сформований запит направляється для обробки та формування даних у систему підприємства.
- Отримання даних з бази даних. *Response* Дані з бази даних отримуються та зберігаються у тимчасовому файлі.
- Підготовка даних. *Stimulus* Отримані дані приводяться до необхідного формату, виявляються та заповнюються пропуски, коригуються аномальні значення.
- Обробка даних. *Response* Система обробляє дані і формує прогнози згідно запиту на основі генетичних алгоритмів машинного навчання.
- Формування відповіді користувачеві. *Response* Система формує та надсилає відповідь користувачеві у вигляді таблиці або файлу.

Є.4.1.3 Функціональні вимоги (Functional Requirements)

Функція: Прогнозування потреб Матеріалів та Послуг у розрізі Замовників

ID функціональної вимоги	FR01
Тип вимоги	Функціональна
Пріоритет	М
Опис вимоги	За допомогою генетичних алгоритмів машинного навчання на основі аналізу закупівель за попередні роки, моделювання очікуваних закупівель та їх вартість
Перехресне посилання на вимоги бізнесу	BR01, BR02
Перехресне посилання на використання	UC01, UC02
Бізнес правило	<ul style="list-style-type: none"> - Вимоги до товару, що закупається описуються у Технічному заїданні - Закупівля однотипних трварів здійснюється в рамках одного тендеру - Закупівлі можуть здійснюватись в рамках затвердженого бюджету Бізнес-плану діяльності - Виконання бюджету закупівель розраховується як % суми створених заявок до затвердженого бюджету діяльності
Джерело	Протокол інтерв'ю №1

Є.4.2 Системна функція 2 (System Feature 2) Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів

Основні модулі:

1. Прогнозування потреб об'ємів (кількість заявок/договорів) закупівель на вказаний період
2. Прогнозування потреб Служби закупівель у розрізі підрозділів
3. Створення заявок на наймання персоналу для Служби закупівель
4. Формування звітів

Є.4.2.1 Опис і пріоритет (Description and Priority)

Ця функція на основі встановлених на підприємстві нормативів навантаження співробітників Служби закупівель, проаналізувавши навантаження за попередні роки, моделює очікувані ресурси поточного та майбутнього періодів. Прогнозування здійснюється в розрізі підрозділів Служби закупівель підприємства.

Пріоритет: Високий

Є.4.2.2 Послідовності стимулів/відповідей (Stimulus/Response Sequences)

- Запит від користувача. *Stimulus* Користувач формує в системі запит, який включає основні параметри прогнозування: Базовий період, прогнозований період, напрямок та обмеження прогнозування (підрозділ).
- Аналіз запиту. *Response* Система аналізує запит, та формує запити до бази даних.
- Направлення запиту до бази даних. *Stimulus* Сформований запит направляється для обробки та формування даних у систему підприємства.
- Отримання даних з бази даних. *Response* Дані з бази даних отримуються та зберігаються у тимчасовому файлі.

- Підготовка даних. *Stimulus* Отримані дані приводяться до необхідного формату, виявляються та заповнюються пропуски, коригуються аномальні значення.
- Обробка даних. *Response* Система обробляє дані і формує прогнози згідно запиту на основі генетичних алгоритмів машинного навчання.
- Формування відповіді користувачеві. *Response* Система формує та надсилає відповідь користувачеві у вигляді таблиці або файлу.
- Інформування HR Служби. *Response* Система, засобами електронної пошти та MS Teams розсилає інформацію фахівцям HR Служби, про необхідність пошуку персоналу для Служби закупівель за встановленою формою.

Є.4.2.3 Функціональні вимоги (Functional Requirements)

Функція: Прогнозування необхідних ресурсів Служби закупівель у розрізі підрозділів

ID функціональної вимоги	FR02
Тип вимоги	Функціональна
Пріоритет	M
Опис вимоги	Моделювання необхідних ресурсів Служби закупівель поточного та майбутнього періодів.
Перехресне посилання на вимоги бізнесу	BR03
Перехресне посилання на використання	UC03
Бізнес правило	<ul style="list-style-type: none"> - Процес закупівлі поділяється на окремі підпроцеси: закупівлю та постачання - Якщо жоден з ключових показників не виконано, робота служби закупівель вважається незадовільною - Якщо кількість отриманих заявок перевищує пропускну здатність закупівель більше ніж на 20%, фахівець HR повинен почати пошук кандидатів для найому у підрозділ закупівель - Якість роботи служби закупівель розраховується як середнє між % виконання поставок в строк та % виконання бюджету в строк
Джерело	Протокол інтерв'ю №1

Є.4.3 Системна функція 3 (System Feature 3) Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників

Основні модулі:

1. Прогнозування строків виконання закупівель Матеріалів у розрізі Замовників
2. Прогнозування строків виконання закупівель Послуг у розрізі Замовників
3. Інформування Замовників та Служби закупівель про очікувані строки виконання заявок та необхідні ресурси
4. Формування звітів

Є.4.3.1 Опис і пріоритет (Description and Priority)

Ця функція на основі даних за попередні роки, моделює очікувані строки виконання закупівель поточного та майбутнього періодів. Прогнозування здійснюється в розрізі Матеріалів/Послуг/Замовників.

Пріоритет: Високий

Є.4.3.2 Послідовності стимулів/відповідей (Stimulus/ Response Sequences)

- Запит від користувача. *Stimulus* Користувач формує в системі запит, який включає основні параметри прогнозування: Базовий період, прогнозований період, напрямок та обмеження прогнозування (Матеріали/Послуги/Замовники).
- Аналіз запиту. *Response* Система аналізує запит, та формує запити до бази даних.
- Направлення запиту до бази даних. *Stimulus* Сформований запит направляється для обробки та формування даних у систему підприємства.
- Отримання даних з бази даних. *Response* Дані з бази даних отримуються та зберігаються у тимчасовому файлі.
- Підготовка даних. *Stimulus* Отримані дані приводяться до необхідного формату, виявляються та заповнюються пропуски, коригуються аномальні значення.
- Обробка даних. *Response* Система обробляє дані і формує прогнози згідно запиту на основі генетичних алгоритмів машинного навчання.
- Формування відповіді користувачеві. *Response* Система формує та надсилає відповідь користувачеві у вигляді таблиці або файлу.
- Інформування Замовників та Служби закупівель. *Response* Система, засобами електронної пошти та MS Teams розсилає інформацію Замовникам та фахівцям Служби закупівель, про очікувані строки виконання заявок та необхідні ресурси.

Є.4.3.3 Функціональні вимоги (Functional Requirements)

Функція: Прогнозування строків виконання закупівель у розрізі Матеріалів, Послуг, Замовників

ID функціональної вимоги	FR03
Тип вимоги	Функціональна
Пріоритет	M
Опис вимоги	Моделювання строків виконання закупівель поточного та майбутнього періодів.
Перехресне посилання на вимоги бізнесу	BR04, BR05, BR06, BR09
Перехресне посилання на використання	UC06, UC07
Бізнес правило	Кожна групв матеріалів , що закупається має свої визначені строки закупівлі Закупівлі мають здійснюватись в затверджені строки Якщо матерівл поставлено на склад пізніше дати потреби вказаної у заявці на закупівлю, заявка вважається простроченою Якщо жоден з ключових показників не виконано, робота служби закупівель вважається незадовільною Якість роботи служби закупівель розраховується як середнє між % виконання поставок в строк та % виконання бюджету в строк Строк виконання заявки розраховується як різниця дати поставки та дати створення заявки
Джерело	Протокол інтерв'ю №1

Є.4.4 Системна функція 4 (System Feature 4) Створення заявок на закупівлю

Основні модулі:

1. Прогнозування об'ємів закупівель Матеріалів у розрізі Замовників
2. Прогнозування об'ємів закупівель Послуг у розрізі Замовників
3. Формування пропозицій по створенню заявок на закупівлю
4. Підтвердження та створення заявок на закупівлю
5. Формування звітів

Є.4.4.1 Опис і пріоритет (Description and Priority)

Ця функція на основі даних за попередні роки, моделює очікувані об'єми закупівель поточного та майбутнього періодів. Прогнозування здійснюється в розрізі Матеріалів/Послуг/Замовників. На основі прогнозних даних формує перелік заявок, які необхідно створити. Користувач переглядаючи перелік запропонованих заявок підтверджує необхідність їх створення. Система створює заявки.

Пріоритет: Низький

Є.4.4.2 Послідовності стимулів/відповідей (Stimulus/Response Sequences)

- Запит від користувача. *Stimulus* Користувач формує в системі запит, який включає основні параметри прогнозування: Базовий період, прогнозований період, напрямок та обмеження прогнозування (Матеріали/Послуги/Замовники).
- Аналіз запиту. *Response* Система аналізує запит, та формує запити до бази даних.
- Направлення запиту до бази даних. *Stimulus* Сформований запит направляється для обробки та формування даних у систему підприємства.
- Направлення запиту про складські залишки. *Stimulus* Сформований запит направляється для обробки та формування даних у систему підприємства.
- Отримання даних з бази даних. *Response* Дані з бази даних отримуються та зберігаються у тимчасовому файлі.
- Підготовка даних. *Stimulus* Отримані дані приводяться до необхідного формату, виявляються та заповнюються пропуски, коригуються аномальні значення.
- Обробка даних. *Response* Система обробляє дані і формує прогнози згідно запиту на основі генетичних алгоритмів машинного навчання.
- Урахування наявних складських запасів. *Response* Система зменшує прогнозні дані по об'ємам закупівель на об'єми наявних запасів і формує пропозиції, щодо створення заявок.
- Формування відповіді користувачеві. *Response* Система формує та надсилає відповідь користувачеві у вигляді таблиці або файлу.

Є.4.4.3 Функціональні вимоги (Functional Requirements)

Функція: Створення заявок на закупівлю

ID функціональної вимоги	FR04
Тип вимоги	Функціональна
Пріоритет	W
Опис вимоги	Автоматизація процесу створення заявок.
Перехресне посилання на вимоги бізнесу	BR07
Перехресне посилання на використання	UC04
Бізнес правило	Закупівля здійснюється виключно при наявності заявки в системі обліку.
Джерело	Інтерв'ю

Є.4.5 Системна функція 5 (System Feature 5) Пропозиції щодо переліку Постачальників минулих періодів

Основні модулі:

1. Аналіз Постачальників попередніх періодів
2. Формування пропозиції по Постачальникам в розрізі Матеріалів
3. Формування пропозиції по Постачальникам в розрізі Послуг
4. Формування звітів

Є.4.5.1 Опис і пріоритет (Description and Priority)

Ця функція на основі даних за попередні роки, моделює можливості постачальників по виконанню заявок на закупівлю. Прогнозування здійснюється в розрізі Матеріалів/Послуг. На основі прогнозних даних формує перелік Постачальників для участі у тендері.

Пріоритет: Середній

Є.4.5.2 Послідовності стимулів/відповідей (Stimulus/Response Sequences)

- Запит від користувача. *Stimulus* Користувач формує в системі запит, який включає основні параметри прогнозування: Базовий період, прогнозований період, напрямок та обмеження прогнозування (Матеріали/Послуги).
- Аналіз запиту. *Response* Система аналізує запит, та формує запити до бази даних.
- Направлення запиту до бази даних. *Stimulus* Сформований запит направляється для обробки та формування даних у систему підприємства.
- Отримання даних з бази даних. *Response* Дані з бази даних отримуються та зберігаються у тимчасовому файлі.
- Підготовка даних. *Stimulus* Отримані дані приводяться до необхідного формату, виявляються та заповнюються пропуски, коригуються аномальні значення.
- Обробка даних. *Response* Система обробляє дані і формує прогнози згідно запиту на основі генетичних алгоритмів машинного навчання.
- Формування відповіді користувачеві. *Response* Система формує та надсилає відповідь користувачеві у вигляді таблиці або файлу.

Є.4.5.3 Функціональні вимоги (Functional Requirements)

Функція: Пропозиції щодо переліку Постачальників минулих періодів

ID функціональної вимоги	FR05
Тип вимоги	Функціональна
Пріоритет	C
Опис вимоги	Автоматизація процесу пошуку постачальників для участі у тендері на закупівлю.
Перехресне посилання на вимоги бізнесу	BR10
Перехресне посилання на використання	UC08
Бізнес правило	Закупівля однотипних трварів здійснюється в рамках одного тендеру Якщо поставлений матеріал Постачальника не прийнято складом, або поставка цим самим Постачальником регулярно здійснюється із запізненням, то Постачальник вважається не надійним
Джерело	Інтерв'ю

Є.4.6 Системна функція 6 (System Feature 6) Прогноз по виконанню поставок Постачальниками в строк

Основні модулі:

1. Аналіз строків виконання заявок на закупівлю попередніх періодів
2. Прогнозування строків виконання заявок на закупівлю по Замовникам в розрізі

Матеріалів

3. Формування звітів

Є.4.6.1 Опис і пріоритет (Description and Priority)

Ця функція на основі даних за попередні роки, моделює можливості служби закупівель та постачання по спроможності виконання поставок в строк. Прогнозування здійснюється в розрізі Матеріалів. На основі прогнозних даних формує очікувані строки виконання заявок на закупівлю в розрізі Матеріалів.

Пріоритет: Низький

Є.4.6.2 Послідовності стимулів/відповідей (Stimulus/Response Sequences)

- Запит від користувача. *Stimulus* Користувач формує в системі запит, який включає основні параметри прогнозування: Базовий період, прогнозований період, напрямок та обмеження прогнозування (Матеріали/Послуги).
- Аналіз запиту. *Response* Система аналізує запит, та формує запити до бази даних.
- Направлення запиту до бази даних. *Stimulus* Сформований запит направляється для обробки та формування даних у систему підприємства.
- Отримання даних з бази даних. *Response* Дані з бази даних отримуються та зберігаються у тимчасовому файлі.
- Підготовка даних. *Stimulus* Отримані дані приводяться до необхідного формату, виявляються та заповнюються пропуски, коригуються аномальні значення.
- Обробка даних. *Response* Система обробляє дані і формує прогнози згідно запиту на основі генетичних алгоритмів машинного навчання.
- Формування відповіді користувачеві. *Response* Система формує та надсилає відповідь користувачеві у вигляді таблиці або файлу.

Є.4.6.3 Функціональні вимоги (Functional Requirements)

Функція: Пропозиції щодо переліку Постачальників минулих періодів

ID функціональної вимоги	FR06
Тип вимоги	Функціональна
Пріоритет	W
Опис вимоги	Прогнозування строків виконання поставок Постачальниками
Перехресне посилання на вимоги бізнесу	BR11
Перехресне посилання на використання	UC09
Бізнес правило	- Якщо поставлений матеріал Постачальника не прийнято складом, або поставка цим самим Постачальником регулярно здійснюється із запізненням, то Постачальник вважається не надійним
Джерело	Інтерв'ю

Є.5 Нефункціональні вимоги (Other Nonfunctional Requirements)

Функція: Забезпечення безпеки та конфіденційності даних користувачів

ID нефункціональної вимоги	NFR01
Тип вимоги	Безпека
Пріоритет	M
Опис вимоги	Забезпечити захист даних користувачів, відповідність GDPR та виконання інших стандартів безпеки. Недопущення витоку Комерційної інформації
Бізнес правило	Компютери та мережа компанії мають відповідати встановленим вимогам компанії Процес закупівлі поділяється на окремі підпроцеси: закупівлю та постачання Система має відповідати всім затвердженим нормативним документам підприємства
Джерело	Нормативні документи

Функція: Швидкий відгук системи на запит користувача

ID нефункціональної вимоги	NFR02
Тип вимоги	Продуктивність
Пріоритет	M
Опис вимоги	Забезпечити час відгуку системи на запит користувача менше 5 секунд.
Бізнес правило	-
Джерело	Інтерв'ю

Функція: Можливість розширення функціоналу

ID нефункціональної вимоги	NFR03
Тип вимоги	Масштабованість
Пріоритет	S
Опис вимоги	Система повинна мати можливість масштабуватися згідно з потребами (горизонтально та вертикально)
Бізнес правило	Компютери та мережа компанії мають відповідати встановленим вимогам компанії Система має відповідати всім затвердженим нормативним документам підприємства
Джерело	Інтерв'ю

Функція: Забезпечення високої відмовостійкості системи

ID нефункціональної вимоги	NFR04
Тип вимоги	Надійність
Пріоритет	M
Опис вимоги	Забезпечення високої відмовостійкості системи на рівні не менш ніж 98%

Бізнес правило	- Кількість відмов на 1000 випадків використання не більше 10-17
Джерело	Інтерв'ю

Функція: Інтеграція з існуючими системами та базами даних компанії

ID нефункціональної вимоги	NFR05
Тип вимоги	Сумісність
Пріоритет	М
Опис вимоги	Забезпечити сумісність з SAP, Microsoft Teams, Outlook та іншими сервісами, що використовуються
Бізнес правило	Компютери та мережа компанії мають відповідати встановленим вимогам компанії Система має відповідати всім затвердженим нормативним документам підприємства
Джерело	Нормативні документи

Функція: Підтримка різних мов інтерфейсу для зручності користувачів

ID нефункціональної вимоги	NFR06
Тип вимоги	Підтримка мов
Пріоритет	С
Опис вимоги	Система повинна підтримувати державну та міжнародну мови інтерфейсу
Бізнес правило	Система має відповідати всім затвердженим нормативним документам підприємства
Джерело	Нормативні документи

Є.5.1 Вимоги до продуктивності (Performance Requirements) - (PER)

Ідентифікатор вимог до продуктивності	Опис вимоги
PER01	Система повинна підтримувати обробку та аналіз великих обсягів даних з бази даних облікової системи підприємства.
PER02	Система має обробляти не менше 20 одночасних запитів користувачів без втрати продуктивності.
PER03	Система повинна забезпечувати плавне масштабування ресурсів відповідно до навантаження.
PER04	Система повинна забезпечувати швидкий відгук на запит користувача, з відгуком не більше 5 секунд для 80% запитів.
PER05	Час відновлення системи після збою не повинен перевищувати 0,5 години, з гарантією збереження даних користувачів.

Є.5.2 Вимоги щодо неушкодженості (техніки безпеки) (Safety Requirements)

Ідентифікатор вимог щодо неушкодженості	Опис вимоги
SAF01	Регулярне оновлення антивірусного програмного забезпечення (щонайменше щодня) для захисту від вірусів та інших шкідливих програм.
SAF02	Використання захищеного протоколу передачі даних TLS 1.2 або вище з сильним шифруванням (наприклад, AES-256) для забезпечення конфіденційності і цілісності обміну даними через HTTPS.
SAF03	Розгортання системи резервного копіювання та відновлення даних з використанням віддаленого сховища Azure Blob Storage, з можливістю автоматичного відновлення даних у випадку збою або атаки, при кожному використанні системи.
SAF04	Реалізація рівнів доступу (адміністратор, користувач).
SAF05	Шифрування даних користувачів з використанням симетричного шифрування AES-256, зберіганих в базі даних, для запобігання несанкціонованому доступу або злому. Ключі шифрування повинні зберігатися окремо від даних користувачів.
SAF06	Встановлення системи моніторингу та аудиту дій користувачів з метою виявлення ненормальної або підозрілої активності, що може свідчити про атаку або порушення безпеки. Зберігання журналів аудиту протягом щонайменше 12 місяців.
SAF07	Використання регулярного патч-менеджменту для забезпечення оновлення операційних систем, серверів, баз даних та застосунків до останніх версій та застосування патчів безпеки для запобігання вразливостей та атак.
SAF08	Розгортання системи захисту від DDoS-атак з використанням хмарних сервісів, таких як Azure DdoS Protection, для забезпечення стабільності, доступності та продуктивності системи незалежно від зовнішніх загроз.

Є.5.3 Вимоги безпеки(Security Requirements)- (SEC)

Ідентифікатор вимог безпеки	Опис вимоги
SEC01	Використання протоколу HTTPS з TLS 1.2 або вище для забезпечення безпечного та зашифрованого передавання даних між клієнтом та сервером.
SEC02	Реалізація аутентифікації та авторизації на основі ролей з використанням OAuth 2.0 та OpenID Connect для забезпечення безпечного доступу користувачів до даних і функцій системи згідно з їхніми дозволами.
SEC03	Регулярне сканування та аналіз коду застосунку на наявність вразливостей і відповідність передовим стандартам безпеки, таким як OWASP Top Ten.
SEC04	Застосування принципу найменшого привілею для обмеження доступу користувачів та систем до ресурсів, які є необхідними для виконання їх обов'язків.

SEC05	Захист від атак типу SQL Injection, Cross-Site Scripting (XSS) та інших загальних веб-загроз шляхом використання параметризованих запитів, екранування виводу даних та валідації вводу користувачів.
SEC06	Використання сучасних методів хешування та соління паролів, таких як bcrypt або Argon2, для забезпечення безпечного зберігання паролів користувачів.
SEC07	Реалізація політики паролів, яка вимагає від користувачів встановлювати сильні та унікальні паролі з використанням мінімальної довжини паролів, спеціальних символів, цифр та великих літер, а також регулярного оновлення паролів.

Є.5.4 Атрибути якості програмного забезпечення (Software Quality Attributes)

1. Надійність (Reliability):
 - Відсоток доступності системи: 99,5%
 - Mean Time Between Failures (MTBF): 3000 годин
2. Продуктивність (Performance):
 - Час відгуку для ключових операцій: менше 300 мс
 - Пропускна здатність: 1000 операцій на секунду
3. Масштабованість (Scalability):
 - Час додавання нових ресурсів до системи: до 0,5 години
 - Здатність системи працювати зі збільшенням навантаження на 200% без зниження продуктивності
4. Безпека (Security):
 - Відсоток забезпечення безпечної передачі даних: 100%
 - Кількість виявлених та усунених вразливостей: 98% від загальної кількості
5. Сумісність (Compatibility):
 - Відсоток сумісності зі стандартами та протоколами: 100%
6. Зручність використання (Usability):
 - Середній час виконання основних завдань у системі: до 10 хвилин
 - Відсоток задоволення користувачів: >80%
7. Модифікованість (Maintainability):
 - Час виявлення та виправлення помилок: до 24 годин
 - Відсоток коду, покритого юніт-тестами та інтеграційними тестами: 80%

Є.6 Інші вимоги (Other Requirements)

1. Дотримання нормативно-правових вимог: Система має дотримуватися місцевих та міжнародних нормативних актів з охорони праці, захисту даних (наприклад, GDPR) та інтелектуальної власності.
2. Система має інтегруватися з існуючими корпоративними системами, такими як системи обліку, CRM та системи звітності (PowerBI), забезпечуючи мінімальні зміни в робочих процесах. Система має підтримувати багатомовність, забезпечуючи зручність використання для користувачів з різних країн та культурних середовищ.
3. Доступність та інклюзивність: Система має бути доступною для користувачів з різними рівнями знань та досвіду, а також для осіб з інвалідністю, забезпечуючи рівні можливості для всіх користувачів.

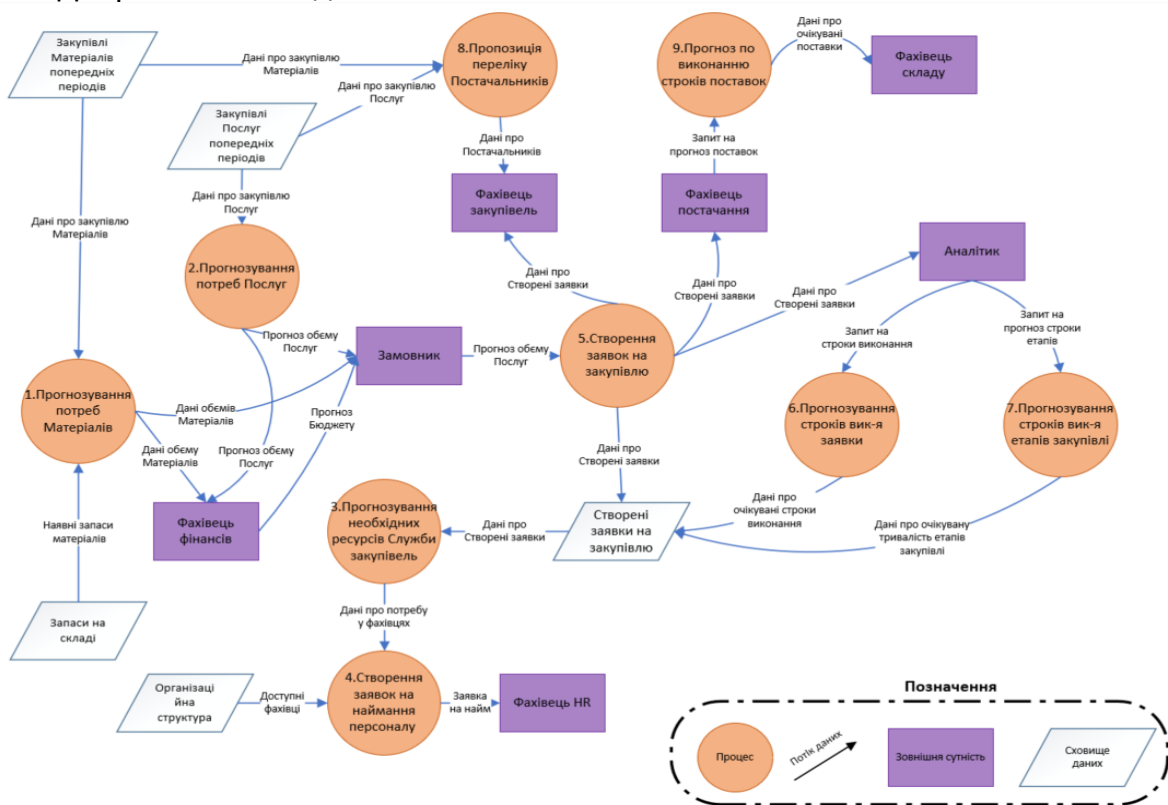
Є.Додаток А: Глосарій

Визначення, термінологія, скорочення та аббревіатури, що використовуються в межах цього проекту, містяться у цій таблиці:

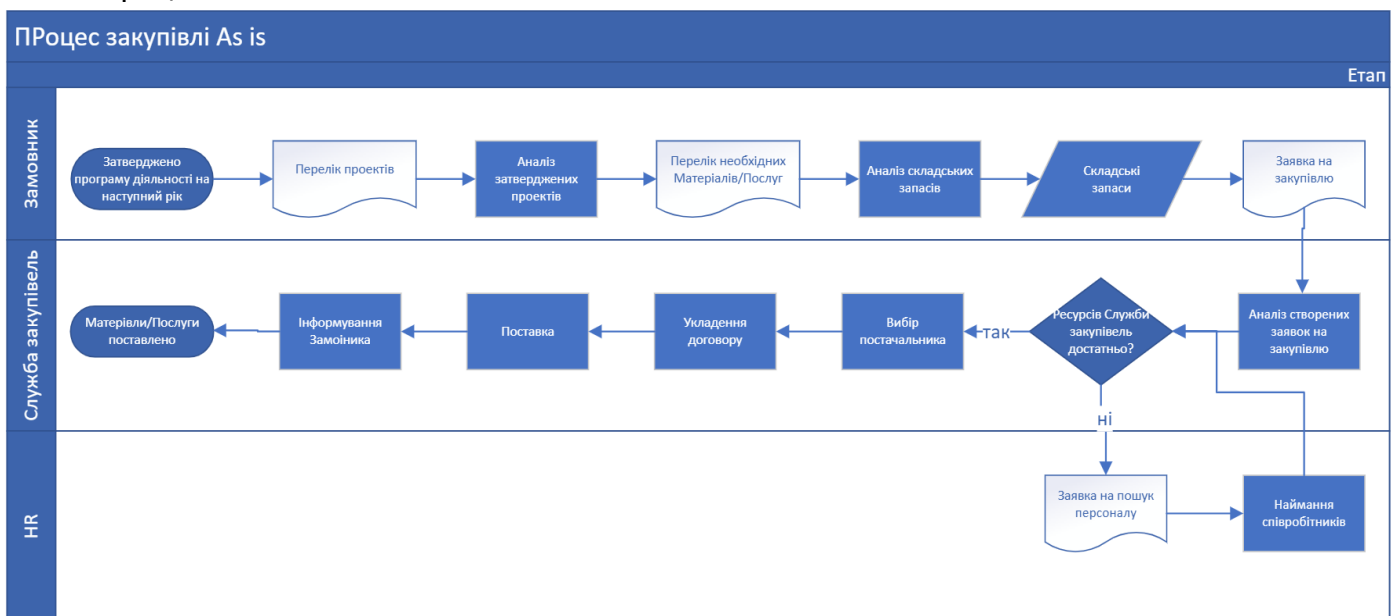
Визначення, термінологія, скорочення та аббревіатури	Визначення/Опис
АС – автоматизована система	організаційно-технічна система, що забезпечує вироблення рішень на основі автоматизації інформаційних процесів
БД – база даних	сукупність даних, організованих відповідно до концепції, яка описує характеристику цих даних і взаємозв'язки між їх елементами, ця сукупність підтримує щонайменше одну з областей застосування
Замовник	
Компанія, підприємство	ТОВ «МЕТІНВЕСТ ДІДЖИТАЛ »
Матеріали	Матеріальні товари, необхідні для виробництва товарів Компанії, що можуть бути виміряні фізичними величинами
НВ – навчальна вибірка	деяка підмножина досліджуваної загальної (генеральної) сукупності. На основі вивчення навчальної вибірки висновки про генеральну сукупність
Послуги	Не матеріальні товари, чи спеціалізовані дії необхідні для виробництва товарів Компанії, що можуть бути виміряні умовними одиницями
Університет	ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА»
Система управління закупівлями	Автоматизована система для обробки та аналізу накопичених даних та знань з метою прогнозування закупівель наступних періодів і прийняття рішень щодо забезпечення своєчасних закупівель та поставок
Служба закупівель	Підрозділ на який покладено обов'язки з управління та здійснення процесу закупівель
ШІ – Штучний інтелект	Система в основі якої лежить машинне навчання
Машинне навчання	Виявлення закономірностей в історичних даних для прогнозування очікуваних закупівель наступних періодів

Є.Додаток В: Моделі аналізу (Analysis Models)

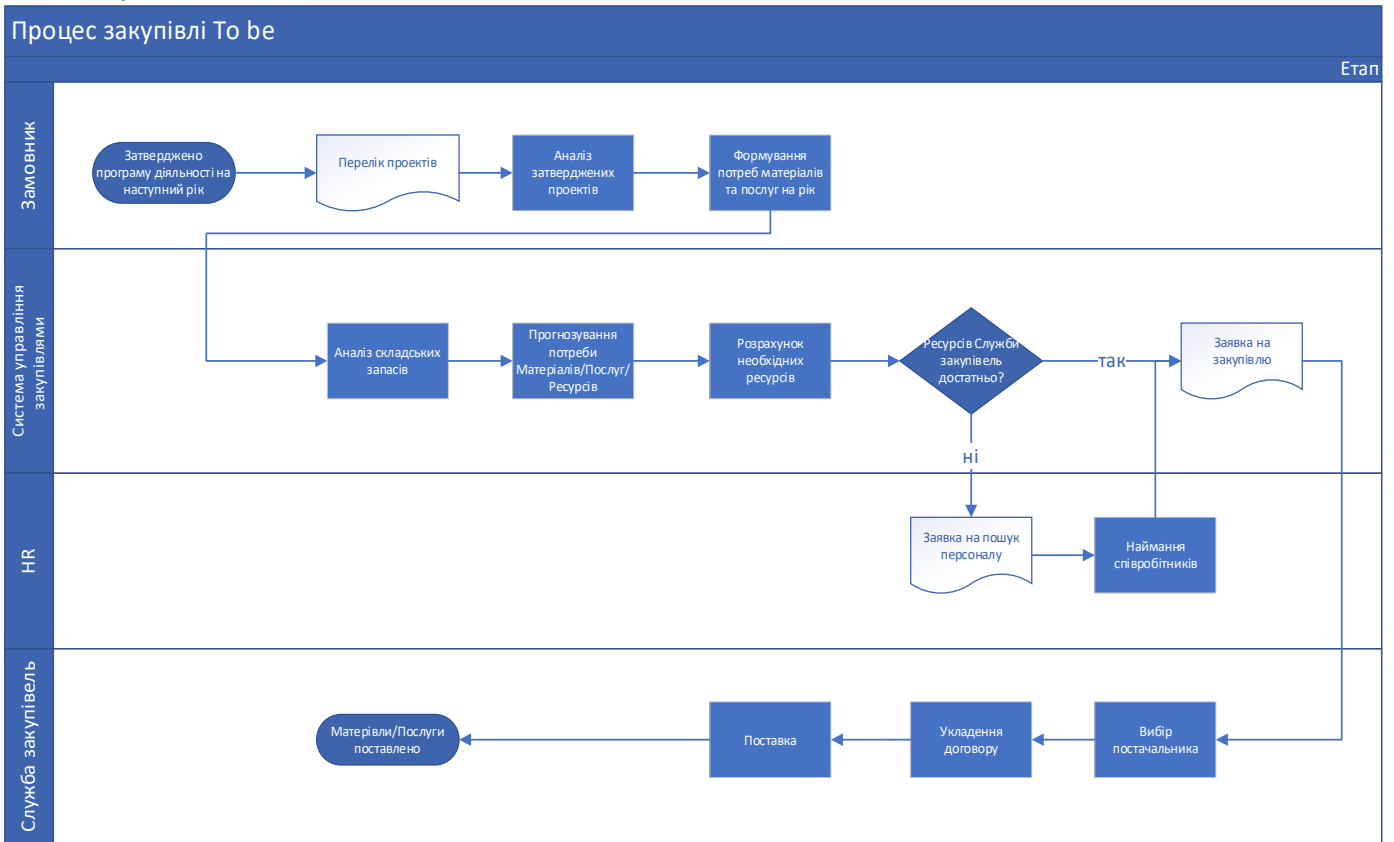
Діаграма потоків даних



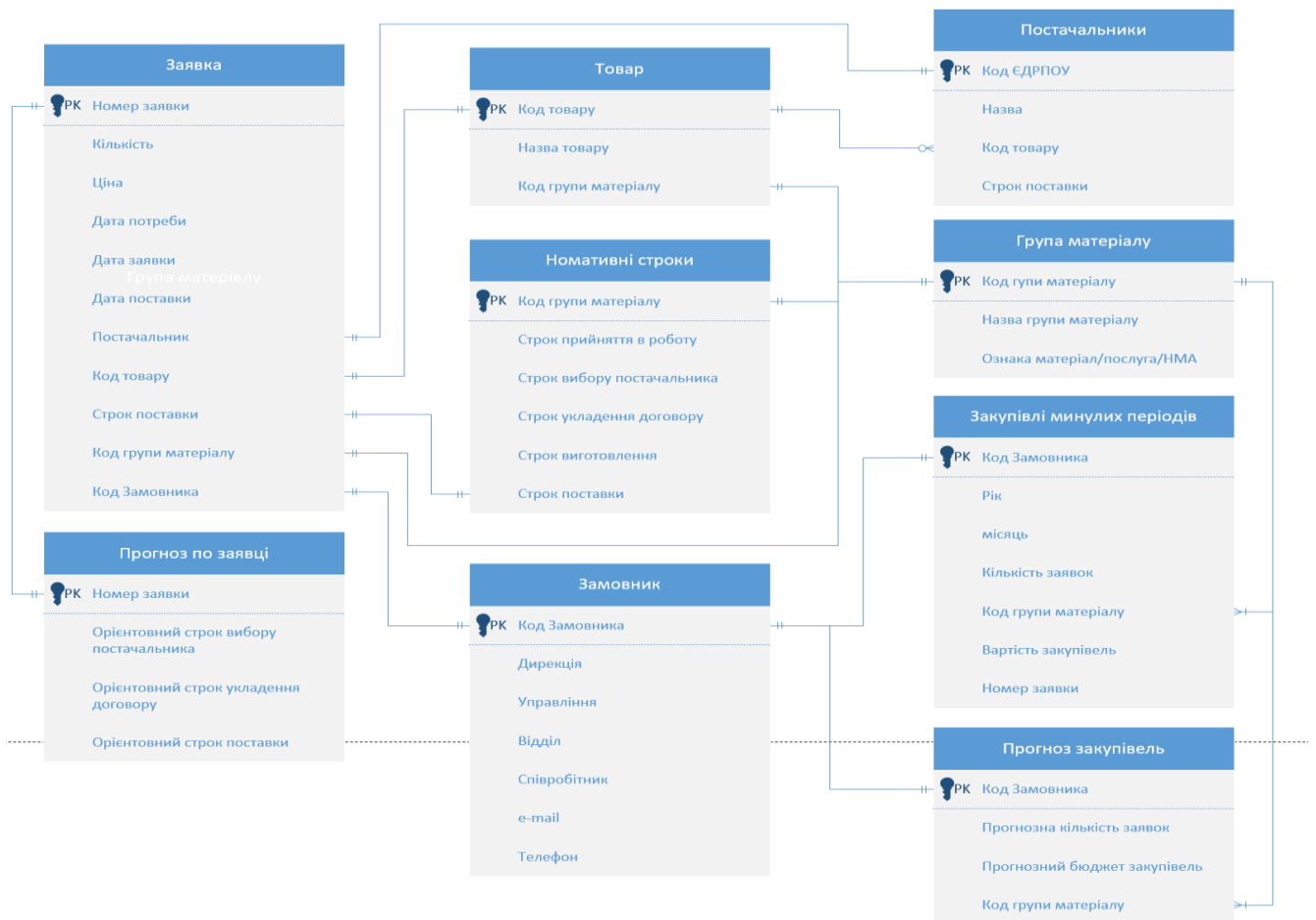
Моделі бізнес-процесів Процес As is:



Процес To Be:



Діаграма «сутність-зв'язок».



Є.Додаток С: Список проблем (Issues List)

Нааявні не вирішені проблеми, а саме:

- Прогнозування об'ємів та кількості заявок закупівель на квартал, рік виконується вручну із застосуванням Excel.
- Прогнозування необхідного бюджету закупівель на квартал, рік виконується вручну із застосуванням Excel.
- Прогнозування строків виконання етапів закупівель та строків виконання заявок виконується вручну із застосуванням Excel.
- Прогнозування необхідної чисельності служби закупівель для забезпечення виконання заявок у строк виконується вручну із застосуванням Excel.

З якими стикаються:

- Технічні, виробничі та адміністративні підрозділи підприємства, що здійснюють закупівлі.
- Служба закупівель підприємства, що виконує заявки на закупівлю.
- Фінансова служба.
- Служба HR.
- Склад.