

ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ  
ПОЛІТЕХНІКА»

Факультет автоматизації виробництва та цифрових технологій  
Кафедра цифрових технологій та проектно-аналітичних рішень

**АВТОРЕФЕРАТ**  
**кваліфікаційної роботи**

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»**

Здобувач



Сергій КРИПАК

Кривий Ріг 2024

Кваліфікаційною магістерською роботою є рукопис.

Робота виконана у Технічному університеті «МЕТІНВЕСТ ПОЛІТЕХНІКА» на кафедрі цифрових технологій та проектно-аналітичних рішень.

Керівник:



МОСКАЛЕНКО Валентина  
Володимирівна, доктор  
технічних наук, професор,  
професор кафедри цифрових  
технологій та проектно-  
аналітичних рішень

Захист відбудеться 23 січня 2024 р. о 09:00 год на засіданні  
екзаменаційної комісії .

Електронна версія автореферату розміщена в Інституційному  
репозитарії ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ  
ПОЛІТЕХНІКА» 20 січня 2024 р.

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА КВАЛІФІКАЦІЙНОЇ РОБОТИ

**Актуальність теми роботи.** Актуальність проблеми полягає в тому, що якість стратегічних управлінських рішень, які приймає ТОП-менеджмент компанії, залежить від якості аналітичних даних, що, в свою чергу, залежить від використовуваних ІТ-технологій. Ростуть вимоги до ІТ-обладнання та програмного забезпечення, що призводить до значного збільшення обсягів закупівельних даних.

У сучасних умовах більшість компаній відмовляються від власних ІТ-підрозділів на користь спеціалізованих ІТ-компаній, які надають послуги з інтеграції та підтримки інформаційних систем. Закупівельні процеси для ІТ-підприємств стають ключовим елементом забезпечення необхідними ресурсами та досягнення бізнес-цілей. Збільшення обсягів закупівельних даних вимагає особливих підходів до їх обробки та аналізу.

Значна частина часу витрачається на підготовку даних через їх можливі помилки або неповноту. Результати аналізу даних залежать від досвіду аналітика та його досвіду роботи з системами обліку.

Отже, актуальним завданням є прискорення обробки даних, а ключовими інструментами для цього є data mining та машинне навчання. Використання методів машинного навчання спрямоване на підвищення ефективності управлінських рішень у сфері закупівель. Розробка програмних систем з використанням методів машинного навчання може значно поліпшити процеси управління закупівлями та прийняття рішень.

**Постановка проблеми.** Для служби закупівель важливо знати загальну кількість заявок на плановий період, класифікацію заявок за групами та типом тендеру, строки поставки, необхідні трудовитрати на одну заявку та тривалість процесу закупівлі.

Слабким місцем у бізнес-процесі управління закупівлями є етап створення заявок та забезпечення необхідними ресурсами службою закупівлі та постачання. Заявки виконуються не повному обсязі в наслідок обмежень бюджету, крім того, протягом року можуть виникати аварійні ситуації та створюватися нові проекти, що призводить до збільшення кількості заявок протягом року від 15% до 50%. Все це створює значні складнощі для забезпечення своєчасності виконання заявок на закупівлю. Значні коливання кількості заявок та розтягнутість процесу створення заявок у часі не дозволяє службам закупівлі та постачання своєчасно забезпечити необхідну кількість персоналу для виконання заявок. На забезпеченість персоналу виконання заявок також впливає те, що середній період підбору та навчання нового персоналу становить від 3 до 5 місяців.

Вирішення цієї проблеми може забезпечити управлінські заходи та засоби обробки, аналізу та прогнозування закупівельних даних.

Головним завданням дослідження є аналіз та вибір алгоритмів машинного навчання, які застосовуються у процесі підготовки та обробки закупівельних даних, вдосконалення бізнес-процесу управління закупівлями та постачанням через формування ефективних стратегій у сфері закупівель, а також проектування програмного забезпечення (ПЗ), яке буде реалізувати обрані методи машинного навчання для аналізу закупівельних даних.

**Мета дослідження.** Вдосконалення процесів управління закупівлями та постачанням в ІТ-компанії на основі використання ПЗ, яке реалізує алгоритми машинного навчання.

**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 напрямків їх використання для аналізу закупівельних даних;
- критеріїв вибору ефективного методу машинного навчання для розв'язання задачі;
- рекомендацій щодо виявлення ключових закономірностей для здійснення прогнозів щодо закупівельних операцій;
- рекомендацій з виявлення ризиків здійснення закупівельних операцій для своєчасного їх попередження.

**Об'єкт дослідження.** діяльність з підготовки, обробки даних для прийняття ефективних бізнесових рішень у сфері закупівель.

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

**Результати та обґрунтування їх новизни / інноваційності.** Новизна роботи полягає в удосконаленні бізнес-процесу управління закупівлями та постачанням завдяки включенню етапу підготовки та обробки закупівельних даних на основі алгоритмів машинного навчання, це реалізується за допомогою запропонованого ПЗ.

Інноваційність роботи полягає у використанні методів та моделей машинного навчання для прогнозування та аналізу закупівельних даних, які на сьогоднішній день не застосовуються у сфері закупівель.

**Структура та обсяг роботи.** Робота складається зі вступу, 5 розділів, висновків, списку використаних джерел, 6 додатків. Загальний обсяг роботи становить 118 сторінок, робота містить 31 рисунок, 44 таблиці. Список використаних джерел складається з 35 джерел.

## ОСНОВНА ЧАСТИНА КВАЛІФІКАЦІЙНОЇ РОБОТИ

Дипломна робота присвячена дослідженню та аналізу методів машинного навчання з метою оптимізації процесу управління закупівельними операціями.

У вступі визначено проблеми використання закупівельних даних, обґрунтовано актуальність роботи, зокрема проведення досліджень методів аналізу та прогнозування закупівельних даних, сформульована мета та завдання дослідження, а також об'єкт та предмет дослідження. Вказано, що більша частина часу витрачається на підготовку даних і вміст у них помилок та неповноти інформації, як наслідок, висновки й рекомендації аналізу даних суттєво залежать від навичок аналітика та його вправності з системами обліку. Розглянуто тенденції розвитку ІТ-компаній.

У першому розділі зроблено дослідження публікацій та праць іноземних та вітчизняних науковців в області закупівель, яке підтверджує, що використання штучного інтелекту для розв'язання бізнес-задач у сфері закупівельної логістики викликає значний науково-практичний інтерес. Проаналізовано можливості методів та алгоритмів машинного навчання та розглянуті питання практичного використання моделей та алгоритмів машинного навчання у сфері закупівель та постачання, їх призначення та класифікації. Розглянуто сучасні програмні інструменти аналізу та обробки даних. Виділено чотири напрямки аналізу та обробки закупівельних даних: кластерний аналіз, методи класифікації, регресійні моделі та методи виявлення аномалій та ризиків у закупівельних процесах. Надано порівняльний аналіз методів машинного навчання, моделей кластерного аналізу, моделей класифікації, регресії та глибокого навчання. Зазначено переваги, недоліки та напрямки практичного використання вказаних моделей машинного навчання. Описано типовий бізнес-процес управління закупівлями підприємства та побудовано схему діючого бізнес-процесу закупівель в нотатії BPMN. Здійснено постановку задачі дослідження.

У другому розділі розглянуто теоретичні основи методів машинного навчання: класифікації та прогнозування. Надано короткий опис, математичне представлення моделей класифікації та напрямки їх застосування. Розглянуто метод наївного Баєса, загальна лінійна модель (Generalized Linear Model), дерева рішень (Decision Tree), метод опорних векторів (Support Vector Machine), лінійний дискримінантний аналіз (Linear Discriminant Analysis), KNN (k-найближчих сусідів). Зазначено короткий опис, напрямки використання і математичне представлення моделей для прогнозування, а саме: лінійна регресія (Linear Regression); поліноміальна регресія (Polynomial Regression); лінійна регресія з

взаємодією елементів (Linear Regression with Interaction Terms); нейронні мережі (Neural Network). Розглянуто ключові елементи аналізу, підготовки та обробки закупівельних даних: підготовка та очищення; визначення відповідного методу машинного навчання; перевірка та оцінка достовірності моделі; інтерпретування результатів. Проведено аналіз використання моделей машинного навчання для аналізу та прогнозування закупівельних даних у різних середовищах. Проведено обробку початкового набору закупівельних даних, побудовано кореляційну матрицю та сформовано оптимальний набір закупівельних даних для подальшого використання для розв'язання задач машинного навчання.

У третьому розділі відображено результати дослідження моделей класифікації та прогнозування за допомогою інструментів RapidMiner та MATLAB. У середовищі RapidMiner досліджено такі моделі класифікації: наївний Баєс; загальна лінійна модель (Generalized Linear Model); логістична регресія (Logistic Regression), глибокого навчання (Deep Learning); дерева рішень (Decision Tree); випадковий ліс (Random Forest ); метод опорних векторів (Support Vector Machine. Тут же досліджено моделі прогнозування лінійної регресії (Linear Regression), загальної лінійної моделі (Generalized Linear Model), поліноміальної регресії (Polynomial Regression), дерева рішень (Decision Tree) та нейронних мереж (Neural Network). У середовищі MATLAB досліджено базовий набір моделей класифікації: наївний Баєс, лінійний дискримінантний аналіз (Linear Discriminant), логістична регресія (Efficient Logistic Regression), Ensemble Bagged Tree, KNN (к-найближчих сусідів), дерева рішень (Decision Tree, Ensemble Bagged Tree, Boosted Trees), нейронні мережі (Neural Network), метод опорних векторів (Support Vector Machine). Також досліджено прогнозування за допомогою лінійної регресії (Linear Regression), дерева рішень (Tree), нейронних мереж (Neural Network) та лінійної регресії з взаємодією елементів (Interactions Linear). Сформовано рекомендації щодо використання моделей машинного навчання та інструментів для аналізу та прогнозування закупівельних даних.

У четвертому розділі розроблено та описано удосконалений бізнес-процес управління закупівлями підприємства з використанням моделей машинного навчання для класифікації та прогнозування закупівельних даних на основі наявних історичних даних. Розроблено удосконалений бізнес-процес закупівель з використанням моделей машинного навчання в нотації BPMN. Визначено проблемні питання, які мають бути вирішені для розробки Системи управління закупівлями. Проаналізовано потреби служб закупівель та її взаємодію з системою обліку підприємства. Розроблено контекстну діаграму Системи управління закупівлями, діаграму потоків даних,

діаграму використання у нотації UML. Сформовано бізнес-вимоги до програмного продукту «Система управління закупівлями». Визначені ділова мета та бізнес-цілі розробки системи. Детально описані функціональні вимоги. Розроблено технічне завдання на розробку програмного продукту «Система управління закупівлями».

У п'ятому розділі надано економічне обґрунтування проєкту щодо створення системи управління закупівлями. Визначено ранги пріоритетності, новизни, готовності та опрацьованості наукової теми. Побудовано графік рангу значущості роботи. З'ясовано, що дослідницький проєкт з розробки системи управління закупівлями має велику значущість, оскільки має високий ранг новизни та пріоритетності дослідження. При визначенні економічного ефекту оцінено вартість дослідження та розробки програмного продукту з використанням калькулятора СОСОМО. Оцінено очікуваний економічний ефект від впливу системи на вартість процесу обробки та аналізу закупівельних даних. Встановлено, що економічна вигода зростатиме пропорційною кількості підприємств до сум економічної вигоди від впровадження, витрати окупляться протягом першого року використання. Визначено напрямок отримання додаткового прибутку при продажі даного продукту третім особам.

У додатках надано приклади коду мови MATLAB для моделей класифікації Ensemble Bagged Tree і дерева рішень (Decision Tree), приклади коду мови MATLAB для моделей прогнозування лінійної регресії (Linear Regression) та нейронних мереж (Neural Network). Також додатки містять повний текст розроблених документів Business Requirements Specification та Software Requirements Specification.

## ВИСНОВКИ

У результаті виконання роботи було отримано такі результати.

Проведено аналіз процесів управління закупівлями і постачанням, 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 і 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, які реалізовані у середовищах RapidMiner та MATLAB, зроблено такі 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 найменше значення помилки;

- інші показники: модель Polynomial Regression має найвищі значення MSE, що вказує на значну кількість квадратичних помилок. Модель 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. 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. У результаті розроблено удосконалений бізнес-процес управління закупівлями, сформовані вимоги до програмного забезпечення для системи управління закупівлями. Впровадження такої 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 управлінського процесу закупівлями та постачання. Доведено значущість використання алгоритмів машинного навчання для розв'язання задач управління закупівельними процесами, визначено переваги формування рекомендацій на основі розв'язків цих задач. Запропоновано використовувати спроектовану систему управління закупівлями, яка заснована на алгоритмах машинного навчання, для підвищення ефективності процесів та прийняття рішень щодо закупівлі та постачання в ІТ-компанії.

У ході виконання дипломної роботи було повністю виконані поставлені завдання. Основна мета вдосконалення процесів управління закупівлями та постачанням в ІТ-компанії на основі використання ПЗ, яке реалізує алгоритми машинного навчання, була досягнута.

## ПЕРЕЛІК ПУБЛІКАЦІЙ ЗА РЕЗУЛЬТАТАМИ ВИКОНАННЯ КВАЛІФІКАЦІЙНОЇ РОБОТИ

1. 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.

2. 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 I. Одеса, 29-30 листопада 2023 р. - Запоріжжя, Видавництво Baltija Publishing, 2023. - С. 241-242.

### АНОТАЦІЯ

Кріпак С.А. Дослідження методів машинного навчання для аналізу та прогнозування закупівельних даних. Кваліфікаційна робота на здобуття ступеню вищої освіти – магістр за спеціальністю 122 Комп'ютерні науки, освітньо-професійна програма «Комп'ютерні науки та цифровий інтелект». – ТОВ ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА», Кривий Ріг, 2024.

Закупівельні процеси є важливою та невід'ємною частиною функціонування бізнес-структур та організацій, забезпечуючи можливість отримання ресурсів, матеріалів та послуг для підтримки їх операцій та досягнення стратегічних бізнес-цілей. Отже, в умовах стрімко мінливого ринкового середовища та стану світової економіки аналіз закупівельних даних і прогнозування поведінки ринку стають ключовими для своєчасної реакції на такі зміни з метою прийняття ефективних управлінських рішень.

Робота присвячена дослідженню методів машинного навчання та їх застосування для аналізу й прогнозування закупівельних даних. Практичне значення дослідження полягає у визначенні певних переваг використання методів машинного навчання для оптимізації процесів управління закупівлями. Дослідження направлене на аспекти використання машинного навчання для прогнозування обсягів планування та оптимізації закупівель. Розглянуто переваги та недоліки методів машинного навчання для прогнозування, виявлення аномалій і ризиків у закупівельних процесах. Увагу приділено практичному застосуванню різних методів машинного навчання,

зокрема окремих нейронних мереж для задачі класифікації в управлінні закупівлями. У 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.

## ABSTRACT

Kripak S.A. Research on machine learning methods for purchasing data analysis and forecasting. Theses for a Master's degree in the specialty 122 Computer Science, Educational Programme "Computer Science and Digital Intelligence". – TECHNICAL UNIVERSITY "METINVEST POLYTECHNICS", LLC, Kryvyi Rih, 2024.

Procurement processes are an essential and integral part of the functioning of business structures and organizations, providing the means to acquire resources, materials, and services to support their operations and achieve core business objectives. Therefore, in the rapidly changing world and the state of the global economy, the analysis of procurement data and forecasting market behavior become crucial for timely response to changing conditions and making effective decisions.

This work is dedicated to reviewing machine learning methods and their application for the analysis and prediction of 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 studying aspects of machine learning application, such as volume forecasting, as well as procurement planning and optimization. Attention is directed towards examining the advantages and disadvantages 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 method discussed is also considered. As a result of the research, a technical task for the development of the software 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.