

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»**

**Факультет прикладної математики  
Кафедра системного програмування  
і спеціалізованих комп'ютерних систем**

«На правах рукопису»  
УДК 004.89

До захисту допущено:  
Завідувач кафедри СПСКС  
Віталій РОМАНКЕВИЧ

\_\_\_\_\_

\_\_\_\_\_

“ \_\_\_ ” \_\_\_\_\_ 2024 р.

**Магістерська дисертація  
на здобуття ступеня магістра  
за освітньо-професійною програмою «Системне програмування та  
спеціалізовані комп'ютерні системи»  
зі спеціальності 123 «Комп'ютерна інженерія»**

**на тему: «Система інтелектуальної обробки даних про  
державні закупівлі з використанням нейромереж»**

Виконав:

студент II курсу, групи КВ-21 мп  
Бойчук Владислав Андрійович

\_\_\_\_\_

Науковий керівник:

доцент, с.н.с.,к.т.н.  
Боярінова Ю. Є.

\_\_\_\_\_

Рецензент:

ст.викладач кафедри ПЗКС  
к.т.н.  
Хіцко Я.В.

\_\_\_\_\_

Засвідчую, що у цій магістерській  
дисертації немає запозичень з праць  
інших авторів без відповідних посилань  
Студент \_\_\_\_\_

(підпис)

Київ - 2024 року

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»**

**Факультет прикладної математики**

**Кафедра системного програмування**

**і спеціалізованих комп'ютерних систем**

Рівень вищої освіти – другий (магістерський)

Спеціальність – 123 «Комп'ютерна інженерія»

Освітньо-професійна програма «Системне програмування та спеціалізовані комп'ютерні системи»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Віталій РОМАНКЕВИЧ

«\_\_» \_\_\_\_\_ 2022р.

**ЗАВДАННЯ**

**на магістерську дисертацію студенту**

**Бойчуку Владиславу Андрійовичу**

1. Тема дисертації «Система інтелектуальної обробки даних про державні закупівлі з використанням нейромереж», науковий керівник дисертації доцент кафедри СПіСКС, с.н.с.,к.т.н, Боярінова Ю.Є, затверджені наказом по університету від «25» листопада 2023 №4317-С .
2. Термін подання студентом дисертації 10 січня 2024
3. Об'єктом дослідження є система інтелектуальної обробки даних про державні закупівлі на основі нейромереж.
4. Предметом дослідження методи і алгоритми інтелектуальної обробки даних з використанням нейромереж для аналізу і оптимізації державних закупівель
5. Вихідні дані: мова програмування Javascript, інтеграція середовища розробки Visual Studio Code, бібліотека для використання chatgpt4 OpenAi, фреймворк NestJS, бібліотка React, Mobx
6. Перелік завдань:
  - a. провести аналітичний огляд наявних методів і алгоритмів інтелектуальної обробки даних з використанням нейромереж для аналізу і оптимізації державних закупівель;
  - b. проаналізувати існуючі методи і алгоритми обробки даних для аналізу і оптимізації державних закупівель;

- с. запропонувати спосіб оптимізації державних закупівель з використанням нейромереж;
- д. розробити програму інтелектуальної обробки даних з використанням нейромереж для аналізу і оптимізації державних закупівель
- е. проаналізувати результати аналізу і оптимізації державних закупівель

7. Перелік ілюстративного матеріалу – презентація

Дата видачі завдання: 1 листопада 2022 р.

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1.	Формулювання об'єкта, предмета, мети, завдань. Огляд існуючих методів	01.11.2022- 01.04.2023	
2.	Перший розділ, огляд літературно-інформаційних джерел	01.04.2023 – 1.07.2023	
3.	Другий розділ. Дослідження існуючих підходів до інтелектуальної обробки даних	01.09.2023 – 28.09.2023	
4.	Третій розділ. Застосування дослідженого матеріалу на практиці. Аналіз результатів програмного забезпечення для інтелектуального аналізу державних закупівель	29.09.2023 – 01.11.2023	
5.	Четвертий розділ. . Аналіз результатів програмного забезпечення для інтелектуального аналізу державних закупівель	02.11.2023 – 30.11.2023	
6.	Оформлення магістерської дисертації	01.12.2023 – 04.01.2024	

Студент

Владислав БОЙЧУК

Науковий керівник

Юлія БОЯРІНОВА

## РЕФЕРАТ

**Актуальність теми.** У сучасному світі штучний інтелект дедалі більше впливає на різні аспекти нашого життя, від спрощених завдань, таких як рекомендаційні системи, до складних проблем, таких як медична діагностика. Кластеризація, як один із методів машинного навчання, демонструє високий потенціал у вирішенні завдань групування даних на основі їхньої схожості. У контексті штучного інтелекту кластеризація може значно підвищити ефективність аналізу даних та прийняття рішень. Україна, як країна з розвиваються ІТ-індустрією, відчуває необхідність в глибокому дослідженні та розробці методів кластеризації. Тому дослідження цієї теми в контексті застосування штучного інтелекту є надзвичайно актуальним та необхідним для наукового та технічного прогресу нашої країни.

**Об'єктом дослідження** є система інтелектуальної обробки даних про державні закупівлі на основі нейромереж.

**Предметом дослідження** методи і алгоритми інтелектуальної обробки даних з використанням нейромереж для аналізу і оптимізації державних закупівель.

**Мета роботи:** покращення обробки даних про державні закупівлі завдяки використанню новітніх технологій – нейромереж та штучного інтелекту, оптимізація запропонованої системи для підвищення її ефективності та надійності.

**Наукова новизна** полягає в наступному:

1. запропоновано структуру вхідних даних для збереження інформації про державні закупівлі, яка відмінна від існуючих, для побудови системи інтелектуальної обробки даних
2. запропоновано архітектуру системи інтелектуальної обробки даних на основі великих мовних моделей для вирішення задачі кластеризації.

**Практична цінність** отриманих результатів дослідження полягає в можливості створення системи інтелектуальної обробки даних про державні закупівлі з використанням нейромереж і кластеризації. Ця система може ефективно аналізувати та групувати великий обсяг даних про закупівлі, що дозволить урядовим та комерційним структурам отримувати цінну інформацію для прийняття

рішень. Наприклад, вона може допомогти виявляти закономірності та тенденції в закупівлях, ідентифікувати аномалії та потенційні ризики корупції. Така система має великий потенціал для підвищення прозорості та ефективності державних закупівель, а також зниження витрат та покращення контролю над ними.

**Апробація роботи.** Основні положення і результати роботи були представлені та обговорювались на Міжнародній науково-практичній конференції "Сучасний стан та пріоритети модернізації науки, освіти та технологій (Біла Церква, 10 січня 2024 р.).

**Структура та обсяг роботи.** Магістерська дисертація складається з вступу, чотирьох розділів та висновків.

У *вступі* подано загальну характеристику роботи, зроблено оцінку сучасного стану проблеми, обґрунтовано актуальність напрямку досліджень, сформульовано мету і задачі досліджень, показано наукову новизну отриманих результатів і практичну цінність роботи, наведено відомості про апробацію результатів і їхнє впровадження.

У *першому розділі* надано теоретичні основи інтелектуальної обробки даних про державні закупівлі, включаючи визначення та класифікацію систем інтелектуальної обробки даних. Також розглянуто специфіку даних про державні закупівлі та їх особливості, а також обговорено необхідність та переваги використання нейромереж у проведенні аналізу та кластеризації цих даних. Аналіз, який дає змогу визначити основні переваги та недоліки цих навчальних посібників.

У *другому розділі* розглянуто основні поняття та класифікація систем інтелектуальної обробки даних, різні методи класифікації.

У *третьому розділі* описано розробку системи на основі нейромереж для інтелектуального аналізу та кластеризації даних про державні закупівлі, включаючи архітектуру та дизайн нейромережі, підготовку та передобробку даних для навчання, а також процес тренування, валідації та тестування на реальних даних.

У четвертому розділі роботи детально розглядається аналіз функціонування розробленої системи інтелектуальної обробки даних про державні закупівлі з використанням методів кластеризації.

У висновках представлені результати проведеної роботи. Робота представлена на **86** аркушах, містить посилання на список використаних літературних джерел.

**Ключові слова:** методичні вказівки, неймережі, кластеризація, держ закупівлі

## Abstract

**Actuality of theme.** In today's world, artificial intelligence is increasingly affecting various aspects of our lives, from simple tasks such as recommender systems to complex problems such as medical diagnostics. Clustering, as one of the methods of machine learning, shows high potential in solving tasks of grouping data based on their similarity. In the context of artificial intelligence, clustering can significantly improve the efficiency of data analysis and decision-making. Ukraine, as a country with a developing IT industry, feels the need for in-depth research and development of clustering methods. Therefore, the study of this topic in the context of the application of artificial intelligence is extremely relevant and necessary for the scientific and technical progress of our country.

**The object** of the study is a system of intelligent processing of data on public procurement based on neural networks.

**The subject** of research is the methods and algorithms of intelligent data processing using neural networks for the analysis and optimization of public procurement.

**The purpose of the work:** improving the processing of data on public procurement thanks to the use of the latest technologies - neural networks and artificial intelligence, optimization of the proposed system to increase its efficiency and reliability.

**The scientific novelty** consists in the following:

1. a structure of input data is proposed for saving information about public procurement, which is different from the existing ones, for building a system of intelligent data processing
2. the architecture of the intelligent data processing system based on large language models for solving the clustering problem is proposed.

**The practical value** of the obtained research results lies in the possibility of creating a system of intelligent processing of data on public procurement using neural networks and clustering. This system can effectively analyze and group a large volume of procurement data, allowing government and commercial entities to obtain valuable

information for decision making. For example, it can help identify patterns and trends in procurement, identify anomalies and potential corruption risks. Such a system has great potential to increase the transparency and efficiency of public procurement, as well as to reduce costs and improve control over them.

**Approbation of work.** The main provisions and results of the work were presented and discussed at the International scientific and practical conference "Current state and priorities of modernization of science, education and technologies" (Bila Tserkva, January 10, 2024).

**Structure and scope of work.** The master's thesis consists of an introduction, four chapters and conclusions.

*The introduction* provides a general description of the work, assesses the current state of the problem, substantiates the relevance of the research direction, formulates the purpose and tasks of the research, shows the scientific novelty of the obtained results and the practical value of the work, provides information on the approbation of the results and their implementation.

*The first chapter* provides the theoretical foundations of the intelligent processing of public procurement data, including the definition and classification of intelligent data processing systems. The specifics of public procurement data and their features are also considered, as well as the necessity and advantages of using neural networks in the analysis and clustering of these data. an analysis that makes it possible to identify the main advantages and disadvantages of these training aids.

In the *second chapter*, the main concepts and classification of intelligent data processing systems, various methods of classification are considered.

*The third chapter* describes the development of a neural network-based system for the intelligent analysis and clustering of public procurement data, including the architecture and design of the neural network, the preparation and refinement of training data, and the process of training, validation, and testing on real data.

*The fourth chapter* of the work examines in detail the analysis of the functioning of the developed system of intelligent processing of data on public procurement using clustering methods.

*The results of the work* are presented in the conclusions. The work is presented on 85 sheets, contains links to the list of used literary sources.

**Keywords:** methodological guidelines, neural networks, clustering, state procurement

## ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ _____	13
ВСТУП _____	14
РОЗДІЛ 1. АНАЛІЗ СУЧАСНИХ ДОСЯГНЕНЬ В ДОСЛІДЖЕННІ НЕЙРОМЕРЕЖ ТА ІНТЕЛЕКТУАЛЬНОГО ПОШУКУ _____	15
1.1. Сутність дослідження нейромереж та інтелектуального пошуку _____	15
1.2. Сучасні здобутки в області нейромереж _____	17
1.3. Розгляд сучасних способів кластеризації даних з використанням штучного інтелекту _____	20
Висновки до розділу _____	22
РОЗДІЛ 2. ТЕОРЕТИЧНІ ОСНОВИ ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ ДАНИХ ПРО ДЕРЖАВНІ ЗАКУПІВЛІ _____	23
2.1 Основні поняття та класифікація систем інтелектуальної обробки даних _	23
2.1.1 Класифікація систем ІОД за типом даних _____	23
2.1.1.1 Класифікація систем ІОД за типом даних. Текстові дані _____	24
2.1.1.2 Класифікація систем ІОД за типом даних. Числові дані _____	26
2.1.1.3 Класифікація систем ІОД за типом даних. Аудіо _____	28
2.1.1.4 Класифікація систем ІОД за типом даних. Відео _____	30
2.1.2 Класифікація систем ІОД за специфікою завдань. _____	32
2.1.2.1 Класифікація ІОД за специфікою завдань. Класифікація _____	33
2.1.2.3 Класифікація ІОД за специфікою завдань. Прогнозування _____	35
2.1.2.4 Класифікація ІОД за специфікою завдань. Оптимізація _____	36
2.2. Необхідність та переваги використання нейромереж у процесі державних закупівель _____	37

2.2.1 Процес державних закупівель _____	39
2.2.2 Інструменти штучного інтелекту в процесі державних закупівель _____	40
2.2.3 Інструменти штучного інтелекту для ручних процесів державних закупівель _____	42
2.3 Специфіка даних про державні закупівлі та їх особливості _____	44
Висновки до розділу _____	46
<b>РОЗДІЛ 3. АНАЛІЗ ДАНИХ, ПОБУДОВА ТА АРХІТЕКТУРА СИСТЕМИ ДЛЯ АНАЛІЗУ ДАНИХ ПРО ДЕРЖАВНІ ЗАКУПІВЛІ _____</b>	
3.1 Вибір платформи та мови програмування _____	47
3.2 Аналіз платформи державних закупівель _____	48
3.3 Аналіз структури вхідних даних _____	52
3.4 Попередня обробка даних _____	57
3.5 Побудова моделі _____	59
3.6 Архітектура системи _____	60
Висновки до розділу _____	62
<b>РОЗДІЛ 4. РЕАЛІЗАЦІЯ СИСТЕМИ ІНТЕЛКТУАЛЬНОЇ ОБРОБКИ ДАНИХ ПРО ДЕРЖАВНІ ЗАКУПІВЛІ НА ОСНОВІ НЕЙРОМЕРЕЖ _____</b>	
4.1 Опис роботи системи _____	63
4.2 Розроблення системи _____	64
4.2.1 Отримання даних від системи Prozorro API _____	64
Підеднаємо TendersModule який відповідає за функціонал, зв'язаний з тендерами. _____	64
4.2.2 Розібрати та відфільтрувати дані від непотрібного “шуму” _____	69
4.2.3 Надати можливість дописувати додаткові критерії користувачем _____	74
4.2.3 Отримати доступ до моделі GPT-4 _____	77
4.2.4 Перевести дані в формат який потрібний для моделі _____	78

4.3 Результати роботи програми _____	80
Висновки до розділу _____	84
ВИСНОВКИ _____	85
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ _____	86
ДОДАТКИ	

## **ПЕРЕЛІК СКОРОЧЕНЬ**

ІОД - інтелектуальної обробки даних

ШІ - штучний інтелект

ШНМ - штучні нейромережі

NLP - Обробка природної мови

## ВСТУП

У сучасному динамічному світі, де дані швидко стають новою валютою, здатність ефективно обробляти та аналізувати величезні обсяги інформації є ключовою для розвитку будь-якої сфери діяльності, зокрема у державному секторі. З ростом обсягів даних, пов'язаних з державними закупівлями, виникає нагальна потреба в розробці продуктивних інструментів для їх аналізу та використання, що, у свою чергу, може значно покращити ефективність державного управління та підвищити прозорість процесів. Розробка системи інтелектуальної обробки даних[1] на основі нейромереж[2], яка зосереджена на державних закупівлях, стоїть на передньому краї інновацій та може служити важливим інструментом для оптимізації та автоматизації цього процесу.

Застосування нейромереж у цій області є перспективним через їхню здатність до виявлення складних шаблонів та залежностей у великих наборах даних, що є критично важливим для ідентифікації ефективних шляхів проведення закупівель, а також для боротьби з неефективністю та корупцією. Впровадження таких систем уже показало свою ефективність у Сполучених Штатах Америки, де вони стають стандартом для аналізу державних даних.

У цій магістерській дисертації проведено дослідження сучасних підходів до обробки даних з використанням нейромереж, розглянуто потенціал їх впровадження в український контекст, та визначено, як такі системи можуть вдосконалити аналіз та кластеризацію даних [3] у державних закупівлях. Основна увага буде приділена вивченню можливостей нейронних мереж для підвищення прозорості та оптимізації процесів державних закупівель, а також розробці рекомендацій щодо їх впровадження у вітчизняні умови. Це дослідження спрямоване на створення теоретичної та практичної бази для впровадження інноваційних технологій в області державних закупівель в Україні, що відіграє критичну роль у забезпеченні ефективного та прозорого використання державних ресурсів.

## РОЗДІЛ 1. АНАЛІЗ СУЧАСНИХ ДОСЯГНЕНЬ В ДОСЛІДЖЕННІ НЕЙРОМЕРЕЖ ТА ІНТЕЛЕКТУАЛЬНОГО ПОШУКУ

### 1.1. Сутність дослідження нейромереж та інтелектуального пошуку

Нейронні мережі, також відомі як штучні нейронні мережі (ШНМ) або симульовані нейронні мережі, відіграють ключову роль у глибокому навчанні та є підмножиною машинного навчання. Вони були названі та структуровані за аналогією до людського мозку, імітуючи спосіб, яким біологічні нейрони передають сигнали один одному.

Штучні нейронні мережі складаються з ряду шарів вузлів: вхідного шару, одного або декількох прихованих шарів та вихідного шару. Кожен вузол, або штучний нейрон, пов'язаний з іншим, має асоційовану вагу та порогове значення. Якщо вихід конкретного вузла перевищує вказане порогове значення, цей вузол активується, передаючи дані на наступний шар мережі. У протилежному випадку дані на наступний шар не передаються.

Для навчання та покращення своєї точності з часом, нейронні мережі використовують тренувальні дані. Проте, коли алгоритми навчання налагоджені для максимальної точності, вони стають потужними інструментами в області комп'ютерних наук та штучного інтелекту, дозволяючи нам класифікувати та групувати дані з великою швидкістю. Завдання, такі як розпізнавання мови чи зображень, можуть виконуватися за хвилини, а не години, порівняно з ручною ідентифікацією людськими експертами. Однією з найвідоміших нейронних мереж є пошуковий алгоритм Google [4].

Обсяг світового ринку штучного інтелекту (ШІ) у 2022 році оцінювався у 454,12 млрд доларів США, а до 2032 року очікується, що він досягне приблизно 2 575,16 млрд доларів США. При цьому сукупний річний темп зростання (CAGR) становитиме 19% з 2023 по 2032 рік. У Північній Америці ринок штучного інтелекту у 2022 році оцінювався у 167,30 млрд доларів США.

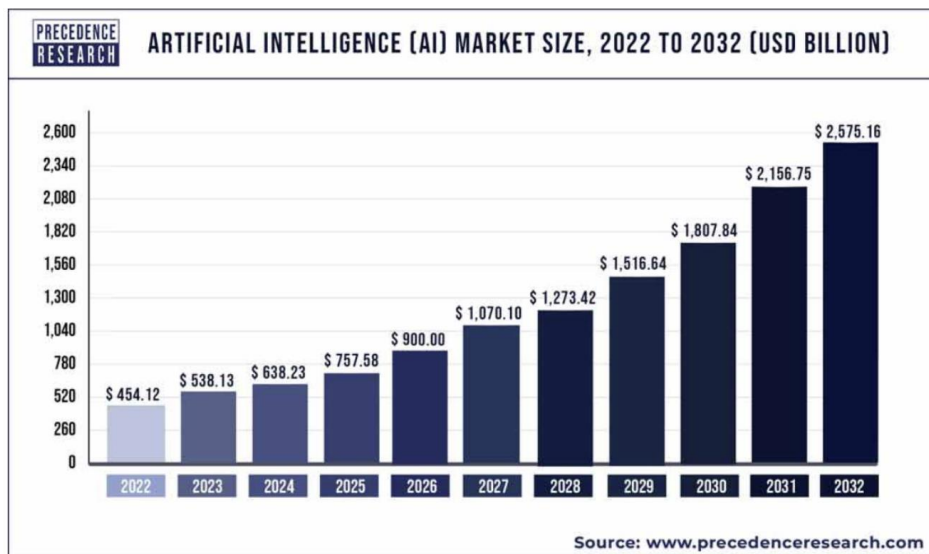


Рисунок 1.1 - Динаміка зростання світового ринку штучного інтелекту, 2022-2032

Штучний інтелект трансформує бізнес в усьому світі, надаючи можливість галузям, від фінансів до охорони здоров'я, оптимізувати операції, прогнозувати ринкові тренди та підвищувати продуктивність. У сфері освіти штучний інтелект дозволяє адаптувати навчальний матеріал під конкретного учня, використовуючи індивідуалізовані рекомендації та адаптивні підходи, що сприяє покращенню якості навчання. Ця адаптація веде до зміни на ринку праці, де деякі професії можуть стати менш затребуваними, але одночасно з'являються нові спеціалізовані робочі місця.

В контексті соціокультурного аспекту, штучний інтелект формує нові культурні і соціальні норми, впливаючи на спосіб, яким люди сприймають і взаємодіють з технологіями. Цей вплив привертає увагу до актуальних питань, таких як етика, безпека, приватність та права на дані, що стають ключовими при розгляді етичного використання технологій. Зокрема, в сферах, де раніше домінувало людське рішення, зараз алгоритми можуть аналізувати великі обсяги даних, щоб приймати обґрунтовані рішення. З погляду довкілля, штучний інтелект

відкриває можливості для оптимізації ресурсів, моніторингу змін клімату та прогнозування природних катастроф, що може призвести до більш стійкого майбутнього.

## **1.2. Сучасні здобутки в області нейромереж**

Штучний інтелект, безумовно, одна з найбільших революцій нашого часу, яка може похвалитися мільйонами дивовижних досягнень у різних галузях. Від автоматичного розпізнавання зображень до високоефективних систем рекомендацій, які керують сучасним онлайн-шопінгом, технології ШІ неодноразово демонстрували своє величезне потенційне вплив.

На медичному фронті штучний інтелект призвів до революційних проривів у діагностиці хвороб та плануванні лікування, дозволяючи лікарям працювати ефективніше та точніше [5]. У галузі автомобілебудування автономні транспортні засоби, засновані на ШІ, змінюють парадигму пересування людей і товарів. Тим часом в галузі розваг та медіа алгоритми, засновані на ШІ, створюють нові форми вмісту, від музики до візуального мистецтва. Ці мільйони досягнень лише підтверджують феноменальну здатність штучного інтелекту трансформувати світ навколо нас.

Швидке проникнення цифрових технологій та Інтернету в значущій мірі сприяло зростанню світового ринку штучного інтелекту за останні кілька років. Великі інвестиції технологічних гігантів у дослідження та розробки постійно сприяють технологічному прогресу в різних галузях. Очікується, що зростаючий попит на штучні технології серед різних секторів кінцевого використання, таких як автомобілебудування, охорона здоров'я, банківська справа та фінанси, виробництво, продукти харчування та напої, логістика та роздрібна торгівля, як очікується, значущо сприятиме зростанню світового ринку штучного інтелекту в найближчі роки. Технологічні інновації завжди були важливою частиною більшості галузей промисловості.

Зростаюча популярність різних медичних пристроїв, що рятують життя, та функцій автономного водіння в нових електромобілях істотно прискорює

зростання ринку штучного інтелекту по всьому світу. Перехід світу до цифровізації позитивно впливає на зростання ринку. Провідні світові технологічні гіганти, такі як Google, Microsoft, IBM, Amazon та Apple, збільшують свої інвестиції у модернізацію та розробку різних застосунків штучного інтелекту.

Очікується, що збільшені зусилля технологічних гігантів щодо поліпшення доступу до ШІ сприятиме зростанню світового ринку ШІ протягом передбачуваного періоду.

Очікується, що сприятливі ініціативи уряду матимуть позитивний вплив на рост промисловості. Створення підкомітетів з машинного навчання та штучного інтелекту в федеральному уряді сприяло розвитку галузі штучного інтелекту. У 2020 році уряд Індії збільшив витрати на "Цифрову Індію" до 477 мільйонів доларів на розвиток штучного інтелекту, Інтернету речей, великих даних, кібербезпеки, машинного навчання та робототехніки. Очікується, що на ринку штучного інтелекту відбудеться істотний ріст у секторі BFSI завдяки застосункам для інтелектуального аналізу даних, оскільки спостерігається ріст впровадження рішень штучного інтелекту в аналізі даних, виявленні шахрайства, кібербезпеці та системах баз даних.

Галузі кінцевого споживання почали інтегрувати штучний інтелект у свої бізнес-процеси, щоб оптимізувати свою діяльність. Технологія штучного інтелекту набирає обертів, оскільки вона допомагає організаціям кінцевих користувачів ставати більш ефективними та орієнтованими на результат. Зростаюче поширення штучного інтелекту спонукає нових учасників входити на ринок ШІ, пропонуючи нішеві продукти та рішення для конкретних застосунків. Додатково, компанії також здійснюють ряд стратегічних ініціатив з консолідації галузі для отримання конкурентних переваг.

#### *Технологічні тренди*

Запровадження систем виявлення захворювань на базі машинного навчання та глибокого навчання. Наприклад, технологія машинного навчання від IBM

допомагає у виявленні ранніх стадій діабетичних хвороб очей завдяки глибокому навчанню.

Використання глибоких нейронних мереж для розпізнавання мови привело до істотних покращень порівняно з існуючим станом розпізнавання мови та голосу. Прогрес в обробці природної мови (NLP) [6], алгоритмів машинного навчання [7], автоматизованого міркування та інтегрованої хмарної платформи відкрив нові можливості для розвитку штучного інтелекту.

#### *Тренди серед постачальників*

Деякі постачальники вступають у партнерські відносини з галузями кінцевих користувачів, щоб розширити свою присутність. Наприклад, інтелектуальна мережа Microsoft співпрацює з офтальмологічним інститутом LV Prasad у Гайдерабаді. Це співробітництво спрямоване на впровадження машинного навчання для надання послуг з офтальмології на даних в Індії.

Деякі стартапи у сфері штучного інтелекту пропонують новітні технології та інновації у сфері програмного та апаратного забезпечення глибокого навчання у галузях електронної комерції, кібербезпеки та роздрібною торгівлі.

Конечный пользователь	2022 год	2023 год	2027 год	2032 год
Здравоохранение	64,33	76,35	152,36	369,22
БФСИ	72,59	86,13	172,00	416,49
Закон	15,96	19,02	38,65	95,47
Розничная торговля	43,83	52,13	105,03	257,43
Рекламные СМИ	63,19	74,97	149,59	362,07
Автомобильная промышленность и транспорт	45,41	53,84	107,81	260,74
сельское хозяйство	29,26	34,78	70,02	171,16
Производство	43,44	51,58	103,75	252,81
Другие	76,11	89,34	170,89	389,77

Рисунок 1.2 - Доходи світового ринку штучного інтелекту (ШІ) за кінцевими користувачами, 2022–2032 рр. (млрд доларів США)

### *Тренди регулювання*

Запровадження систем виявлення захворювань на базі машинного та глибокого навчання. Наприклад, технологія машинного навчання IBM допомагає у виявленні ранніх стадій діабетичних хвороб очей.

Використання глибоких нейронних мереж для розпізнавання мови дало значущі покращення в порівнянні з поточним станом техніки. Авансування в обробці природної мови (NLP), алгоритмах машинного навчання, автоматичному міркуванні та інтегрованій хмарній платформі забезпечило ринкові можливості для розвитку ШІ.

### **1.3. Розгляд сучасних способів кластеризації даних з використанням штучного інтелекту**

Вчені та математики розробляли різні алгоритми для виявлення різних типів кластерів. Завданням є вибір найкращого рішення для конкретної проблеми. Алгоритми можуть не завжди бути точними. Науковці можуть використовувати методи, які входять до лише однієї категорії, або ж вони можуть застосовувати гібридні алгоритми, комбінуючи методи з декількох категорій. Ось категорії алгоритмів кластеризації:

1. Віднизу догори: ці алгоритми, відомі як агломеративні або ієрархічні [8], починають з того, що кожен об'єкт даних об'єднується в пари з найближчим до нього. Потім ці пари об'єднуються між собою. Кластери зростають до моменту, коли досягається певний поріг кількості кластерів або відстані між ними.
2. Роздільний: ці алгоритми працюють за принципом агломеративних, але вони стартують із всіх точок у одному кластері і намагаються розділити їх на два менших кластери. Зазвичай це передбачає пошук лінії чи іншої функції для чіткого розділення кластера.
3. К-середніх: цей метод шукає  $k$  різних кластерів, спочатку випадково розподіляючи точки між  $k$  групами. Розраховується середнє значення кожного кластера, після чого аналізується кожна точка для визначення, до якого кластера

вона найбільше підходить. Якщо точка не відповідає своєму кластеру, вона пересувається до іншого.

4. K-medoids: схожий на k-середніх, але центр кластера визначається за допомогою медіанного алгоритму.

5. Нечіткий: кожна точка може належати до декількох кластерів, які визначаються за допомогою різних методів. Це корисно, коли точки рівнодалекі від кожного центру кластера.

6. Сітчастий: алгоритми використовують заздалегідь визначену сітку для розділення даних. Точки призначаються кластерам відповідно до того, до якої частини сітки вони належать.

7. Хвильовий: дані спочатку стискаються або перетворюються за допомогою вейвлет-функції. Потім на стислі дані застосовується алгоритм кластеризації.

Алгоритми кластеризації застосовуються в широкому діапазоні технологій. Вчені, які працюють з даними, використовують ці алгоритми для класифікації та впорядкування.

Наприклад, численні додатки для взаємодії з людьми можуть бути ефективнішими, якщо вони використовують вдосконалені алгоритми кластеризації. Школи можуть бажати розподіляти учнів по класах відповідно до їхніх здібностей та можливостей. Алгоритми кластеризації групують учнів зі схожими інтересами та потребами.

Деякі компанії бажать розділити своїх потенційних клієнтів на категорії, щоб надавати їм більш цілеспрямоване обслуговування. Для нових покупців може бути запропонована докладна допомога, щоб вони могли ознайомитися з продукцією та варіантами. Досвідчені клієнти можуть одразу переглядати пропозиції та, можливо, отримувати спеціальні пропозиції, які відповідатимуть схожим покупцям.

Існує багато інших прикладів з різних галузей, таких як виробництво, банківська справа та мореплавство. Вони всі базуються на алгоритмах, які дозволяють розподіляти робочий обсяг на менші підмножини, які можна обробляти однаковою чином. Всі ці підходи у значній мірі залежать від збору даних.

### **Висновки до розділу**

В контексті сучасних науково-технічних викликів, ШІ виявляється як ключовий інструмент для вирішення складних задач. Особливо актуальним є застосування ШІ у сфері кластеризації, яка є основоположною для аналізу великих об'ємів даних. Кластеризація завдяки можливостям штучного інтелекту не лише підвищує ефективність роботи з даними, але і дозволяє видобувати з них нові, раніше недоступні для людини знання.

Враховуючи швидке зростання об'ємів інформації та її складності, інтеграція технік штучного інтелекту в процес кластеризації стає не тільки бажаним, але й необхідним кроком. Сучасний ШІ може адаптуватися, вчитися на власних помилках та постійно вдосконалювати свої методи, що робить його ідеальним інструментом для роботи з динамічними даними.

Таким чином, використання штучного інтелекту в кластеризації відкриває нові горизонти для наукових досліджень та комерційних застосувань, стаючи ключовим фактором успіху в сучасному цифровому світі.

## РОЗДІЛ 2. ТЕОРЕТИЧНІ ОСНОВИ ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ ДАНИХ ПРО ДЕРЖАВНІ ЗАКУПІВЛІ

### 2.1 Основні поняття та класифікація систем інтелектуальної обробки даних.

Основні поняття, що лежать в основі систем інтелектуальної обробки даних (ІОД), включають різноманіття методів та технологій, які використовуються для збору, аналізу, інтерпретації та представлення даних. Ці системи розроблені для того, щоб сприяти кращому розумінню великих та часто складних наборів даних, використовуючи алгоритми машинного навчання, статистичний аналіз та моделювання, штучний інтелект (AI) та інші обчислювальні методи.

Системи інтелектуальної обробки даних (ІОД) - це складна і багатогранна область, яка включає широкий спектр методів, технік та підходів. Ці системи розроблені для обробки і аналізу великих обсягів даних, використовуючи техніки штучного інтелекту та машинного навчання. Вони можуть варіюватися від простих алгоритмів до складних моделей, що імітують людське мислення.

#### 2.1.1 Класифікація систем ІОД за типом даних.

Класифікація систем інтелектуальної обробки даних за типом даних охоплює різні області та підходи. Вони поділяються на категорії:

1. Текстові дані - Обробка природної мови (NLP) та текстовий аналіз: NLP (Natural Language Processing) є галуззю штучного інтелекту, яка фокусується на взаємодії між комп'ютерами та людською (природною) мовою. Текстовий аналіз включає методи для обробки, класифікації, та вилучення значущої інформації з текстових даних.
2. Числові дані - Статистичний аналіз, прогнозування часових рядів: Статистичний аналіз включає в себе збір, аналіз, тлумачення та

представлення даних. Прогнозування часових рядів - це техніка, яка аналізує часові дані для виявлення патернів і передбачення майбутніх тенденцій.

3. Аудіо - Розпізнавання мови, аналіз аудіо-сигналів: Розпізнавання мови полягає в перетворенні усної мови в текст. Аналіз аудіо-сигналів зосереджений на обробці та інтерпретації звукових файлів.
4. Відео - Обробка та аналіз відеоданих, розпізнавання образів: Обробка відеоданих включає техніки для розуміння, інтерпретації та маніпуляції відеоконтенту. Розпізнавання образів орієнтоване на ідентифікацію об'єктів, людей, написів та інших елементів у відео.
5. Неструктуровані дані - Аналіз соціальних медіа, великі невизначені датасети: Аналіз соціальних медіа включає в себе збір та аналіз даних з соціальних мереж для вилучення інсайтів та патернів. Великі невизначені датасети - це набори даних, які не підлягають легкій класифікації або структуризації, і часто включають комбінації тексту, зображень, аудіо та інших типів даних

#### **2.1.1.1 Класифікація систем ЮД за типом даних. Текстові дані**

Обробка природної мови (NLP) є фундаментальною галуззю, яка поєднує області комп'ютерних наук, штучного інтелекту та лінгвістики. Її основна мета - створення систем, які можуть розуміти, інтерпретувати та реагувати на людську мову таким чином, як це робить людина. NLP охоплює різні аспекти обробки мови, від базового розуміння значення слів до складного розуміння контексту і нюансів людської мови. Одним з основних викликів у NLP є розуміння контексту та нюансів мови.

Людська мова може бути неоднозначною та складною, з численними винятками та ідіоматичними виразами. Системам NLP потрібно вчитися визначати смислові та емоційні відтінки, щоб точно відтворювати людське сприйняття тексту. Це досягається завдяки застосуванню складних алгоритмів машинного навчання, які аналізують великі обсяги текстових даних для вивчення мовних патернів [9]. NLP використовується в широкому спектрі застосувань, включаючи автоматичний переклад, аналіз настрою, витягування інформації, чат-боти та

голосові інтерфейси. Наприклад, у сфері обслуговування клієнтів чат-боти на основі NLP можуть відповідати на запитання та надавати інформацію без безпосередньої участі людини. Це дозволяє підвищити ефективність обслуговування та знизити часові витрати.

Останнім часом NLP досягла значного прогресу завдяки розвитку глибокого навчання та нейронних мереж. Сучасні моделі NLP, такі як BERT (Bidirectional Encoder Representations from Transformers)[10] та GPT (Generative Pre-trained Transformer)[11], здатні обробляти мову на неймовірно складному рівні, відтворюючи більш точні та змістовні відповіді. Ці моделі використовуються для створення більш розумних та інтуїтивно зрозумілих систем, які можуть адаптуватися та реагувати на змінні потреби користувачів.

1. Токенізація: Розбиває текст на менші одиниці, такі як слова, символи, числа, що є будівельними блоками для розуміння контексту в NLP моделях.
2. Стемінг та Лематизація: Обробляє слова для отримання їх кореневих форм, спрощуючи аналіз та зменшуючи обсяг оброблюваних даних.
3. Видалення Стоп-слів: Виключає з тексту незначущі слова (наприклад, сполучники та прийменники), щоб зосередитися на більш важливих словах.
4. TF-IDF (Term Frequency-Inverse Document Frequency): Статистична техніка для оцінки важливості слова в документі, яка ґрунтується на частоті його вживання та рідкості в інших документах.
5. Класифікація Тексту: Організація та структурування великих обсягів неструктурованого тексту для аналізу і отримання корисних даних.
6. Видобування Ключових Слів: Техніка аналізу тексту для вилучення значущих слів та фраз, що може допомогти у виявленні основних тем або питань у документі.
7. Векторне Представлення Слів (Word Embeddings): Перетворення слів на числові вектори для можливості їх обробки машинним навчанням та глибоким навчанням.

8. Аналіз Сентименту (Sentiment Analysis): Оцінка емоційного забарвлення тексту, розділяючи його на позитивний, негативний або нейтральний.
9. Видобування Аспектів (Aspect Mining): Ідентифікація різних аспектів тексту, часто використовується разом з аналізом сентименту для отримання повного розуміння наміру тексту.
10. Моделювання Тем (Topic Modelling): Аналіз текстових документів для виявлення тем, які в них обговорюються, використовуючи непідконтрольні алгоритми машинного навчання.
11. Автоматичне Підсумовування Тексту (Text Summarization): Створення коротких, змістовних підсумків великих текстових документів.

#### **2.1.1.2 Класифікація систем ЮД за типом даних. Числові дані**

Аналіз часових рядів - це особливий спосіб аналізу послідовності точок даних, зібраних за певний інтервал часу. При аналізі часових рядів аналітики записують точки даних через постійні проміжки часу протягом заданого періоду часу, а чи не просто записують точки даних періодично чи випадковим чином. Однак цей тип аналізу - це не просто збір даних з часом.

Що відрізняє дані часових рядів від інших даних, то це те, що аналіз може показати, як змінні змінюються з часом. Іншими словами, час є найважливішою змінною, оскільки він показує, як дані коригуються з плином точок даних, а також остаточних результатів. Він забезпечує додаткове джерело інформації та встановлений порядок залежностей між даними.

Аналіз часових рядів зазвичай вимагає великої кількості точок даних для забезпечення узгодженості та надійності. Великий набір даних гарантує, що у вас буде репрезентативний розмір вибірки і аналіз зможе виключити зашумлені дані[12]. Це також гарантує, що будь-які виявлені тенденції чи закономірності не є викидами та можуть враховувати сезонні коливання. Крім того, дані часових рядів

можна використовувати для прогнозування – прогнозування майбутніх даних на основі історичних даних.

Аналіз часових рядів допомагає організаціям зрозуміти основні причини тенденцій чи системних закономірностей з часом. Використовуючи візуалізацію даних, бізнес-користувачі можуть побачити сезонні тенденції та глибше зрозуміти, чому ці тенденції виникають. Завдяки сучасним аналітичним платформам ці візуалізації можуть виходити далеко за межі лінійних графіків.

Коли організації аналізують дані за послідовні інтервали, вони можуть використовувати прогнозування часових рядів для прогнозування ймовірності майбутніх подій. Прогнозування часових рядів є частиною прогнозової аналітики. Він може показувати ймовірні зміни даних, такі як сезонність чи циклічне поведінка, що забезпечує краще розуміння змінних даних і допомагає краще прогнозувати.

Наприклад, державні школи Де-Мойна проаналізували дані про успішність учнів за п'ять років, щоб виявити учнів із групи ризику та відстежувати прогрес із плином часу. Сьогоднішні технології дозволяють нам збирати величезні обсяги даних щодня, і тепер простіше, ніж будь-коли, зібрати достатньо узгоджених даних для всебічного аналізу.

Аналіз часових рядів, як описано на Tableau [13], тісно пов'язаний з інтелектуальною обробкою даних та штучним інтелектом, оскільки цей метод використовує передові аналітичні та алгоритмічні підходи для інтерпретації даних, зібраних протягом певного часового періоду

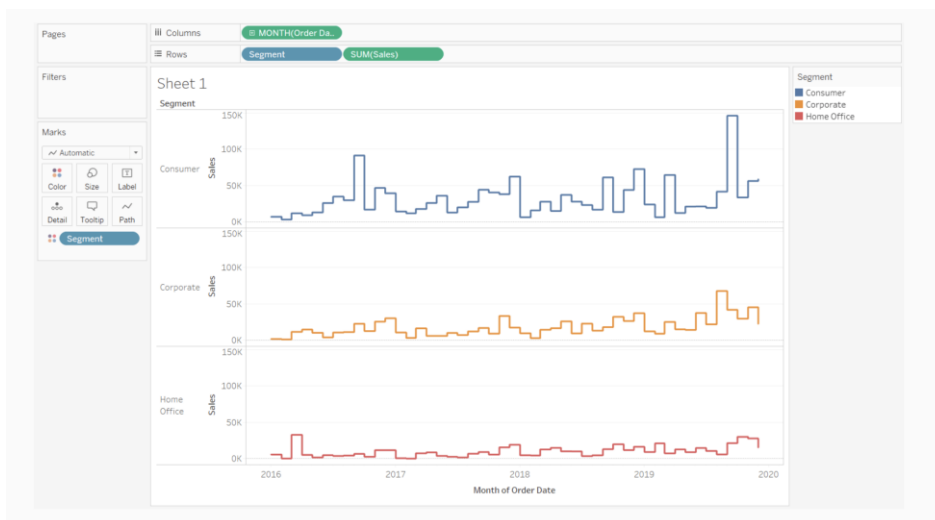


Рисунок 2.2.1.2 - приклад аналізу часових рядів для прогнозування продажів

### 2.1.1.3 Класифікація систем ІОД за типом даних. Аудіо

Аналіз аудіо-сигналів - це процес отримання інформації з аудіозаписів або акустичних сигналів для розуміння їхнього змісту або характеристик. Основні завдання в цій області включають:

1. **Визначення акустичних характеристик:** Для аналізу аудіо-сигналів важливо визначити різні акустичні характеристики, такі як частота, амплітуда, довжина, спектральний склад тощо. Для цього використовуються методи сигнальної обробки, такі як фур'є-аналіз [14] і вейвлет-аналіз [15].
2. **Виявлення шумів і фільтрація:** Однією з основних задач є виявлення шумів в аудіо-сигналах та їх подальша фільтрація. Це допомагає покращити якість аналізу та розпізнавання.
3. **Емоційний аналіз голосу:** Важливою областю є визначення емоцій, які виражені у голосі. Моделі глибокого навчання допомагають розпізнати емоційний стан говорящої особи, що може бути використано в різних

застосуваннях, таких як оцінка настрою в соціальних мережах або аналіз емоцій у клінічних дослідженнях.

4. Аналіз звукових паттернів: Це включає в себе виявлення і аналіз паттернів у звукових сигналах, таких як розпізнавання ритму, інтонації, мелодійних структур і гармоній. Це може бути корисним для аналізу музики, ритму мови та інших аспектів аудіо-сигналів.
5. Пошук аудіозаписів зі специфічними особливостями: Деякі дослідження включають в себе пошук аудіозаписів, які мають певні акустичні або музичні особливості. Наприклад, в медичинських дослідженнях можуть виявлятися аудіозаписи зі специфічними звуковими сигнатурами для діагностики певних станів.

Для досягнення успішних результатів у сфері аналізу аудіо-сигналів використовуються різні методи, включаючи машинне навчання, глибоке навчання, статистичні методи і обробку сигналів. Також важливим є збір і анотування великих обсягів аудіо-даних для навчання та валідації моделей.

Методами аудіо аналізу являються:

1. Перетворення Фур'є: Ця математична операція дозволяє розкласти аудіосигнал на різні частоти, виражаючи його як суму синусоїдальних і косинусоїдальних хвиль різних частот. Це дозволяє аналізувати частотний склад сигналу.
2. Спектрограма: Спектрограма - це візуальне відображення частотного змісту сигналу з плином часу. Вона дозволяє бачити, як частоти змінюються з часом.
3. Виявлення висоти звуку: Цей процес полягає в оцінці основної частоти музичної ноти або голосу. Це важливо для аналізу мови і музики.
4. Мел-частотні кепстральні коефіцієнти (MFCC): Цей метод витягує спектральні характеристики сигналу, які схожі на сприйняття людиною. Він широко використовується для аналізу мови та аудіо.

5. Аналіз форми хвилі: Досліджує амплітудні та часові характеристики сигналу, показуючи, як сигнал змінюється з часом.
6. Частотно-часовий аналіз: Дозволяє аналізувати, як частотний склад сигналу змінюється з часом, використовуючи методи, такі як БПФ [16] .
7. Спектральний аналіз: Досліджує частотний зміст сигналу за допомогою БПФ і оцінки спектральної щільності потужності.
8. Витягнення аудіохарактеристик: Витягує характеристики з сигналу, такі як спектральний центроїд, спектральний спад та інші, для використання в задачах аналізу звуку.
9. Машинне навчання і розпізнавання образів: Застосовує алгоритми машинного навчання для класифікації, розпізнавання мови та інших завдань аналізу аудіосигналів.
10. Обробка аудіосигналу: Маніпулює сигналами для покращення їх якості або зміни їх для конкретних цілей, включаючи фільтрацію, зменшення шуму, прибирання еха та зміну тривалості сигналу.

#### **2.1.1.4 Класифікація систем ІОД за типом даних. Відео**

Обробка та аналіз відеоданих - це галузь, яка вивчає методи та алгоритми для обробки, аналізу і розпізнавання відеосигналів. Ця область має ряд застосувань та включає в себе такі ключові аспекти:

1. Відеоаналітика: Відеоаналітика використовується для автоматичного аналізу відеоданих з метою виявлення об'єктів, подій або поведінки. Це може бути важливим для систем безпеки, відеоспостереження, аналізу трафіку і багатьох інших сфер.
2. Розпізнавання облич: Ця задача включає в себе виявлення та ідентифікацію облич людей на відео. Вона може використовуватися в системах розпізнавання облич для автентифікації або для виявлення осіб у масових заходах.

3. Відстеження об'єктів: Відстеження об'єктів відео дозволяє відслідковувати рух та зміну положення об'єктів на відеозапису. Це корисно в автономних транспортних засобах, роботах та системах відеоспостереження.
4. Розпізнавання жестів і міміки: Розпізнавання жестів і міміки обличчя може бути використано для аналізу комунікації людей на відео, що має застосування в розпізнаванні емоцій, інтерфейсах керування жестами та в інших областях.
5. Аналіз руху: Ця область займається аналізом руху об'єктів на відео для виявлення шаблонів, трендів та ризикових ситуацій. Це може бути використано в системах виявлення інцидентів та аналізу руху в містах.
6. Відеоіндексація і пошук: Щоб забезпечити ефективний пошук та організацію великих обсягів відеоданих, використовуються методи відеоіндексації, які дозволяють створювати ключові слова, описи і індекси для відеозаписів.
7. Розпізнавання об'єктів і сцен: Алгоритми розпізнавання об'єктів дозволяють ідентифікувати та класифікувати об'єкти на відео, такі як автомобілі, тварини, будівлі тощо. Це важливо в автоматизованих системах та роботах.
8. Синтез відео: Синтез відео використовується для створення комп'ютерної графіки і спеціальних ефектів у відеороликах та фільмах.
9. Розпізнавання дій і активностей: Дозволяє аналізувати дії та активності осіб або об'єктів на відео для визначення їхнього стану або намірів.
10. Відеоаналітика в реальному часі: Системи відеоаналітики можуть працювати в реальному часі, що важливо для нагляду та прийняття рішень в реальному часі, таких як системи безпеки та моніторингу.

Ця галузь включає в себе численні методи і технології, включаючи машинне навчання, обробку зображень та відео, комп'ютерне бачення та багато інших, і вона має широкий спектр застосувань у великій кількості галузей, включаючи технології розпізнавання образів та інтелектуального аналізу відеоданих.

### **2.1.1.5 Класифікація систем ІОД за типом даних. Неструктуровані дані**

Неструктуровані дані - це тип інформації, що не має чіткої організації або структури, і часто представлений у різних форматах, таких як текст, аудіо, відео, зображення тощо. Ця інформація не піддається традиційному табличному або реляційному способу збереження та обробки. Для обробки неструктурованих даних використовуються різні методи, включаючи числові дані для аналізу, часові ряди для виявлення трендів, аудіо- та відеоаналіз для обробки звукових і візуальних даних, а також природну мову обробки (NLP) для аналізу текстової інформації та витягнення семантики. Використання цих методів дозволяє отримувати цінну інформацію і приймати обґрунтовані рішення на основі неструктурованих даних у різних галузях.

### **2.1.2 Класифікація систем ІОД за специфікою завдань.**

Класифікація систем інтелектуальної обробки даних за специфікою завдань охоплює різні області та підходи. Вони поділяються на категорії:

Класифікація за специфікою завдань:

1. Класифікація: Розділення даних на певні категорії.
2. Кластеризація: Групування об'єктів на основі їх подібності.
3. Прогнозування: Використання історичних даних для передбачення майбутніх тенденцій.
4. Оптимізація: Пошук найбільш ефективних рішень на основі доступних даних.

### 2.1.2.1 Класифікація ІОД за специфікою завдань. Класифікація

Класифікація - це один із основних видів завдань у машинному навчанні та інтелектуальній обробці даних, яке полягає в розділенні вхідних даних на певні категорії або класи на основі їх характеристик і властивостей. Цей процес включає в себе навчання моделі на основі наявних даних і використання цієї моделі для призначення нових даних до відповідних категорій або класів. Класифікація має широкий спектр застосувань у різних галузях, включаючи розпізнавання образів, фільтрацію спаму в електронній пошті, медичну діагностику, прогнозування ризиків і багато інших областей. Основні етапи процесу класифікації включають наступне:

1. Збір та підготовка даних: Перший крок - це збір і підготовка набору даних, який включає в себе вхідні признаки та відповідні класи або мітки для кожного прикладу в наборі.
2. Вибір алгоритму класифікації: Обирається підходящий алгоритм або модель машинного навчання для завдання класифікації. Популярні алгоритми включають в себе логістичну регресію, метод опорних векторів, рішачі дерева, випадковий ліс, нейронні мережі тощо [17].
3. Навчання моделі: Модель навчається на тренувальних даних, де вона навчається визначати зв'язки між признаками і класами. Валідація та налаштування моделі: Модель перевіряється на валідаційному наборі даних для оцінки її точності і, за необхідності, налаштування параметрів моделі.
4. Тестування моделі: Завершальний етап включає в себе тестування моделі на незалежному тестовому наборі даних для оцінки її продуктивності та точності на нових прикладах.
5. Використання моделі для прогнозування: Після успішного навчання і валідації модель може бути використана для класифікації нових даних, які не були включені в тренувальний набір.
6. Оцінка результатів: Результати класифікації оцінюються за допомогою метрик, таких як точність, відсоток правильних класифікацій, матриця

плутанини і багато інших. Важливо відзначити, що ефективність класифікації значно залежить від якості даних, вибору правильного алгоритму та належної настройки моделі. Класифікація є потужним інструментом для автоматизованого визначення та сортування даних в різних галузях і застосовується для вирішення різних завдань, від розпізнавання образів до прийняття рішень на основі даних.

Класифікація є потужним інструментом для автоматизованого визначення та сортування даних в різних галузях і застосовується для вирішення різних завдань, від розпізнавання образів до прийняття рішень на основі даних.

#### **2.1.2.2 Класифікація ІОД за специфікою завдань. Кластеризація**

Кластеризація - це один із методів аналізу даних, який використовується для групування об'єктів (точок даних, спостережень, векторів) на основі їх подібності або схожості в одній групі, яку називають "кластером". Основна мета кластеризації полягає в розділенні набору даних на кілька груп або кластерів так, щоб об'єкти в одному кластері були схожі між собою, а об'єкти в різних кластерах були відмінні або менше схожі. Основні характеристики кластеризації включають наступне:

1. Подібність або відстань між об'єктами: Визначення того, наскільки два об'єкти схожі або відмінні між собою, грає ключову роль у кластеризації. Зазвичай використовуються метрики відстані, такі як евклідова відстань, косинусна схожість, Манхеттенська відстань і багато інших [18].
2. Кількість кластерів: Визначення кількості кластерів або груп, на які розділити дані, може бути завданням, яке вимагає попереднього аналізу або експертної оцінки. Існують алгоритми, які можуть визначити оптимальну кількість кластерів автоматично.
3. Алгоритми кластеризації: Існує багато алгоритмів для виконання кластеризації, таких як алгоритми K-середніх, ієрархічна кластеризація, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), [19]

агломеративна кластеризація і багато інших. Кожен з них має свої переваги і недоліки та підходить для різних типів даних та завдань.

4. Візуалізація результатів: Після завершення процесу кластеризації важливо візуалізувати отримані кластери, щоб зрозуміти їхню структуру та взаємодію об'єктів.
5. Застосування: Кластеризація знаходить застосування в багатьох галузях, включаючи аналіз соціальних мереж, рекомендаційні системи, генетику, обробку зображень, аналіз тексту, виявлення аномалій та багато інших областей.

Кластеризація є важливим інструментом для групування та виявлення патернів у наборах даних, допомагає в різних завданнях аналізу та прийняття рішень і є одним з методів навчання без учителя у машинному навчанні.

### **2.1.2.3 Класифікація ІОД за специфікою завдань. Прогнозування**

Одним із важливих аспектів є типи прогнозів. Існують різні типи прогнозів, які використовуються в різних областях:

1. Класичні часові ряди: Цей тип прогнозування використовується для прогнозування значень величини з часом, враховуючи їхню історію. Він широко використовується в економіці, метеорології, фінансах та інших галузях. Класифікаційні прогнози: Вони використовуються для класифікації об'єктів або подій у певні категорії. Наприклад, класифікація спаму в електронній пошті або класифікація пацієнтів на основі їх медичних даних.
2. Регресійні прогнози: Ці прогнози визначають числові значення або величини, наприклад, прогноз ціни на нерухомість, на основі історичних даних і різних факторів.
3. Прогнозування за допомогою машинного навчання: Використовується широкий спектр алгоритмів машинного навчання, таких як дерева рішень,

метод опорних векторів (SVM), нейронні мережі та інші для прогнозування подій або значень на основі великих обсягів даних.

4. Екстраполяція та інтерполяція: Ці методи використовуються для прогнозування значень поза наявними даними (екстраполяція) або відновлення проміжних значень між існуючими даними (інтерполяція).
5. Байєсівський прогноз: Використовується теорія ймовірностей для оцінки ймовірностей різних подій та їхніх можливих наслідків.
6. Прогнозування на основі тексту (NLP): Використовується обробка природної мови для аналізу текстових даних та прогнозування подій, настроїв або трендів на основі текстової інформації. Кожен з цих типів прогнозувань використовується у відповідних сферах та завданнях, і вони мають свої особливості та методи аналізу даних. Кожен з цих типів прогнозувань використовується у відповідних сферах та завданнях, і вони мають свої особливості та методи аналізу даних

#### **2.1.2.4 Класифікація ІОД за специфікою завдань. Оптимізація**

Оптимізація в контексті систем інтелектуальної обробки даних представляє собою процес пошуку найбільш ефективних рішень або конфігурацій на основі наявних даних. Вона важлива для вирішення різних завдань і включає в себе різні аспекти та підходи. Ось детальніші відомості:

1. Цільова функція: Оптимізація передбачає наявність цільової функції, яка визначає, як оцінюється якість рішення. Ця функція може бути максимізацією або мінімізацією певного параметра або показника ефективності.
2. Обмеження: У багатьох задачах оптимізації існують обмеження, які повинні бути враховані при пошуку оптимального рішення. Ці обмеження можуть включати фізичні обмеження, бюджетні обмеження чи інші обмеження вирішуваної задачі.

3. Пошуковий простір: Оптимізація включає аналіз пошукового простору, який представляє собою набір можливих рішень. Важливо визначити, яким чином досліджувати цей простір, чи то перебором всіх варіантів, чи застосовуючи різні алгоритми для знаходження оптимального рішення.
4. Алгоритми оптимізації: Оптимізація використовує різні алгоритми для пошуку оптимального рішення. До найпоширеніших алгоритмів оптимізації належать градієнтний спуск, алгоритми генетичного пошуку, алгоритми з частинками, методи опорних векторів (SVM), метаевристичні методи та інші.
5. Застосування: Оптимізація використовується в різних сферах, таких як фінанси (оптимізація інвестиційного портфеля), транспорт (маршрутизація), виробництво (планування виробництва), інженерія (проектування), логістика (оптимізація ланцюга постачання), енергетика (розподіл ресурсів) та багато інших галузей.

Оптимізація є потужним інструментом для прийняття рішень, покращення результатів і вирішення складних завдань. Вона дозволяє знайти найкращі рішення в умовах обмежень і оптимізувати різноманітні процеси в сферах діяльності

## **2.2. Необхідність та переваги використання нейромереж у процесі державних закупівель.**

Державні закупівлі охоплюють покупку товарів, послуг та виконання робіт урядами та державними підприємствами, становлячи близько 12% світового ВВП. Ці закупівлі є ключовими для надання державних послуг та досягнення політичних цілей, але вони можуть бути складними, вимагати багато часу та ресурсів, а також піддаватися ризику помилок, шахрайства та корупції.

Інструменти штучного інтелекту мають потенціал для поліпшення процесів прийняття рішень у сфері державних закупівель завдяки їх здатності до ефективної обробки даних і надання точної та своєчасної інформації для рішень. Недолік даних і систем, що сприяють відстеженню та вибору найкращих постачальників, може ускладнювати рішення щодо майбутніх закупівель.

Штучний інтелект (AI) має величезний потенціал для трансформації процесів державних закупівель, які становлять значну частину урядових витрат. Аспекти, як AI може бути використаний у цьому контексті:

1. Рівні можливості та зниження витрат: AI може зрівняти можливості для різних постачальників, прискорюючи процес написання пропозицій та дозволяючи більшій кількості фірм брати участь у торгах.
2. Економія ресурсів: Використання AI може привести до значних заощаджень як для уряду, так і для постачальників, особливо в країнах з високими витратами на управління закупівлями, як-от у Великобританії.
3. Збільшення конкуренції: AI сприяє залученню більшої кількості постачальників, що збільшує конкуренцію та може покращити умови для уряду.
4. Балансування інновацій та стандартів: Потрібно знайти баланс між використанням AI для автоматизації процесів та забезпеченням, що він не обмежує нових учасників чи інновації у сфері державних закупівель.
5. Реформування урядових процесів: Уряди починають розуміти необхідність адаптації своїх процесів для інтеграції з AI, що вимагає спрощення та переосмислення існуючих підходів.
6. Глобальний погляд: Досвід інших країн, таких як США, Естонія, Сінгапур, Дубай і Південна Корея, демонструє, як вони застосовують AI для внутрішньої оцінки пропозицій, підвищуючи прозорість і ефективність.
7. Обмеження AI: Є певні обмеження в тому, що може зробити AI, особливо у сферах, які вимагають більш тонкого людського розуміння, таких як культура та особливості роботи з людьми

### 2.2.1 Процес державних закупівель

Процес державних закупівель включає декілька ключових етапів, які можуть змінюватися залежно від типу та вартості контракту, способу закупівель та законодавчої основи. Огляд основних етапів:

1. Ідентифікація вимог: Закупівельна організація визначає потреби, обсяг контракту, очікувані результати та показники ефективності.
2. Вибір методу закупівель: Організація обирає найбільш підходящий метод на основі факторів, як-от складність, ризик, ринкові умови та конкуренція. Поширені методи включають відкриті тендери, обмежені торги, запит пропозицій, запит котирувань та прямі контракти.
3. Планування закупівель і розробка стратегії: Підготовка плану закупівель, який описує терміни, бюджет, ролі та відповідальності, критерії оцінки та стратегію управління ризиками.
4. Обробка заявок на закупівлі: Отримання необхідних схвалень та дозволів для розпочати процес закупівель.
5. Підготовка та публікація тендерної документації: Опис вимог умов та умов контракту, запрошення постачальників подати пропозиції.
6. Зустріч перед тендером/пропозицією та візит на об'єкт: Організація зустрічі або візиту для уточнення будь-яких питань або сумнівів.
7. Подача та відкриття тендерних пропозицій: Подача тендерних заявок постачальниками і їх прозоре відкриття організацією.
8. Оцінка пропозицій: Оцінювання пропозицій на основі встановлених критеріїв і методів.
9. Рекомендації щодо присудження контракту: Підготовка звіту про процес оцінки та рекомендації кращого постачальника.
10. Переговори щодо контракту: Обговорення з обраним постачальником нерозв'язаних питань або деталей контракту.
11. Присудження контракту (підписання): Підписання контракту з обраним постачальником, надаючи йому юридичну силу.

### 2.2.2 Інструменти штучного інтелекту в процесі державних закупівель

Інструменти штучного інтелекту можуть бути застосовані на різних етапах процесу державних закупівель для автоматизації конкретних завдань, оптимізації прийняття рішень, підвищення прозорості та відповідальності, а також зниження витрат та ризиків. Ось деякі приклади використання інструментів штучного інтелекту:

1. Ідентифікація вимог: Інструменти штучного інтелекту можуть допомогти проаналізувати історичні дані щодо минулих контрактів, поточних ринкових тенденцій та прогнозів майбутнього попиту для визначення оптимальних вимог і специфікацій для нових контрактів.
2. Визначення методу закупівлі: Інструменти штучного інтелекту можуть допомогти вибрати найбільш підходящий метод закупівлі з урахуванням таких факторів, як складність, ризик, ринкові умови та конкуренція.
3. Планування закупівель і розробка стратегії: Інструменти штучного інтелекту можуть допомогти створювати реалістичні плани закупівель, які враховують різні сценарії, обмеження та невизначеність.
4. Обробка заявок на закупівлю: Інструменти штучного інтелекту можуть допомогти спростити процес схвалення та авторизації за рахунок автоматизації робочих процесів, перевірки інформації та виявлення аномалій.
5. Підготовка та публікація тендерних документів: Інструменти штучного інтелекту можуть допомогти генерувати стандартизовані тендерні документи, які відповідають юридичним та етичним вимогам, а також публікувати їх на різних платформах для охоплення більшої аудиторії потенційних постачальників.
6. Спільні наради перед тендером/пропозицією та візити на об'єкт: Інструменти штучного інтелекту можуть полегшити комунікацію між закупівельними організаціями та потенційними постачальниками, надаючи чат-боти, які

можуть відповідати на поширені запитання та надавати пояснення або відгуки.

7. Подача та розкриття тендерних заявок/пропозицій Інструменти штучного інтелекту можуть бути використані для перевірки достовірності та повноти тендерних заявок або пропозицій, а також для забезпечення їх конфіденційності та безпеки. Також, вони можуть відкривати заявки або пропозиції у прозорий та перевірений спосіб. Наприклад, компанія IBM "Довіряйте своєму постачальнику" використовує технологію блокчейн[20] для підтвердження та обміну інформацією та обліковими даними постачальників.
8. Оцінка заявок/пропозицій Інструменти штучного інтелекту можуть допомогти оцінити заявки або пропозиції на основі заздалегідь визначених критеріїв і методів, а також виконувати складні розрахунки, порівняння та аналіз. Інструменти штучного інтелекту також можуть виявити будь-які помилки, упередженість або шахрайство під час оцінки. Наприклад, підрозділ SAS Procurement Integrity використовує передову аналітику та штучний інтелект для виявлення та запобігання шахрайству під час закупівель.
9. Рекомендації щодо присудження контракту Інструменти штучного інтелекту можуть допомогти підготувати комплексні та об'єктивні звіти, в яких узагальнюється процес оцінки та рекомендується найкращий постачальник для присудження контракту. Також вони можуть допомогти обґрунтувати рекомендації та надати докази для будь-яких суперечок або апеляцій. Наприклад, Power BI від Microsoft використовує візуалізацію даних та інформаційні панелі для представлення та обміну інформацією та рекомендаціями щодо закупівель.
10. Переговори щодо контракту Інструменти штучного інтелекту можуть допомогти проводити переговори з обраним постачальником щодо будь-яких невирішених питань або деталей контракту, таких як ціна, графік поставок або гарантії виконання. Вони також можуть допомогти контролювати процес

переговорів та забезпечувати його справедливість і дотримання вимог. Наприклад, Kira Systems використовує обробку природної мови та машинне навчання для видобування та аналізу даних та положень контракту.

11. Присудження контракту (підписання) Інструменти штучного інтелекту можуть допомогти завершити та підписати контракт, використовуючи цифрові підписи та технологію блокчейн[21] для забезпечення його дійсності та відстежування. Також вони можуть допомогти зберігати дані контрактів та керувати ними у безпечний та доступний спосіб. Наприклад, Icertis використовує штучний інтелект та технологію блокчейн[21] для автоматизації та оптимізації управління контрактами.

### **2.2.3 Інструменти штучного інтелекту для ручних процесів державних закупівель**

Навіть якщо процес державних закупів здійснюється вручну, інструменти ШІ все ж можна використовувати для покращення його результатів та впливу. Ось кілька прикладів використання інструментів ШІ для ручних процесів державних закупів: Аналіз даних. Інструменти штучного інтелекту можуть допомогти аналізувати обсяги даних з різних джерел, таких як контракти, рахунки-фактури, постачальники, ринки, нормативні акти і т. д., для створення ідей, тенденцій, закономірностей та прогнозів, які можуть сприяти прийняттю більш ефективних рішень щодо закупівель. Наприклад, SAP Ariba використовує штучний інтелект і машинне навчання для надання даних на основі рекомендацій щодо оптимізації закупівель.

1. Управління контрактами. Інструменти штучного інтелекту можуть допомогти управляти життєвим циклом контракту, від його укладення до припинення, автоматизуючи такі завдання, як відстеження продуктивності, платежів, виконання вимог, ризиків, змін, спорів і т. д., а також надавати сповіщення, рекомендації та звіти. Наприклад, Seal Software використовує

обробку природної мови та машинне навчання для автоматичного сканування та інтерпретації обсяжних та складних юридичних документів.

2. Вибір постачальника. Інструменти штучного інтелекту можуть допомогти вибирати найкращих постачальників на основі різних критеріїв, таких як якість, ціна, надійність, стійкість, інновації і т. д., за допомогою алгоритмів машинного навчання, які вчать на минулих даних і зворотньої зв'язку. Наприклад, Scoutbee використовує штучний інтелект і аналіз великих даних, щоб знаходити найкращих постачальників для кожного проекту і взаємодіяти з ними.
3. Управління відносинами з постачальниками. Інструменти штучного інтелекту можуть допомогти підтримувати і поліпшувати відносини з постачальниками, спрощуючи спілкування, співпрацю, зворотний зв'язок, визнання та інновації. Наприклад, Tradeshift використовує штучний інтелект і технологію блокчейну, щоб забезпечити безперебійні транзакції та взаємодію між покупцями та постачальниками.
4. Вимірювання ефективності закупівель. Інструменти штучного інтелекту можуть допомогти виміряти і оцінити ефективність закупівель на основі різних показників, таких як ефективність, результативність, етика, вплив і т. д., за допомогою візуалізації даних, інформаційних панелей і ключових показників. Наприклад, SpendHQ використовує штучний інтелект для надання аналітики витрат та інформації, яка допомагає вимірювати та покращувати ефективність закупівель.
5. Інновації у сфері закупівель. Інструменти ШІ можуть сприяти розвитку інновацій у сфері закупівель, виявляючи нові можливості, проблеми та рішення, а також підтримує експерименти, тестування і масштабування нових ідей. Наприклад, Amazon Web Services використовує штучний інтелект і хмарні обчислення для надання інноваційних рішень та послуг для державних закупівель.
6. Прозорість і відповідальність у закупівлях. Інструменти ШІ можуть допомогти підвищити прозорість і відповідальність у закупівлях, надаючи

доступ до відповідної інформації, даним і документам різним зацікавленим сторонам, таким як громадяни, засоби масової інформації, організації громадянського суспільства і т. д., а також забезпечуючи виконання законодавства та етичних стандартів. Наприклад, Open Contracting Partnership використовує відкриті дані та штучний інтелект для підтримки та моніторингу відкритих і справедливих державних контрактів.

7. Розвиток потенціалу у сфері закупівель. Інструменти ШІ можуть допомогти розвивати потенціал у сфері закупівель, надаючи навчання, рекомендації і підтримку посадовим особам, які займаються закупівлями, і постачальникам з різних аспектів процесів і практики закупівель. Наприклад, Coursera пропонує онлайн-курси і сертифікації з штучного інтелекту і закупівель.

### **2.3 Специфіка даних про державні закупівлі та їх особливості.**

Інструменти ШІ мають потенціал перетворити процес державних закупівель, роблячи його більш ефективним, результативним і етичним. Однак застосування цих інструментів супроводжується деякими викликами та ризиками, які потребують уважного вирішення:

1. Якість та доступність даних: Якість та доступність даних є критично важливими для функціонування інструментів ШІ. Необхідно забезпечити надійні джерела даних та системи, які збиратимуть, зберігатимуть та обмінюватимуться даними у стандартизований та сумісний спосіб.
2. Етичні принципи та цінності: Використання ШІ в державних закупівлях повинно відповідати етичним принципам та цінностям, які захищають людську гідність, права та інтереси. Однак інструменти ШІ можуть мати вбудовані упередження або обмеження, що може вплинути на їхні рішення або дії.
3. Правові рамки та регулювання: Використання ШІ у державних закупівлях має відповідати правовим рамкам та регулюванням, що забезпечують справедливість, прозорість та підзвітність. Однак інструменти ШІ можуть

породжувати нові правові виклики або невизначеності, які можуть не бути охоплені існуючими законами або правилами.

Щоб подолати ці виклики та ризики, організації закупівель повинні використовувати підхід, який залучає багато сторін та передбачає співпрацю та консультації з різними учасниками, включаючи політиків, регуляторів, постачальників, експертів, громадянське суспільство тощо. Вони також повинні використовувати людиноцентричний підхід, який гарантує, що людський нагляд та втручання завжди можливі та бажані при використанні інструментів ШІ у державних закупівлях.

### **Висновки до розділу**

У цьому розділі розглянуто різноманітні аспекти інтелектуальної обробки даних, які мають значний вплив на процес державних закупівель. Аналізуючи різні типи даних та специфічні завдання, включаючи класифікацію, прогнозування та оптимізацію, визначено необхідність та переваги застосування нейромереж у цій сфері. Особлива увага приділена специфіці даних про державні закупівлі, їхнім особливостям, а також впровадженню інструментів штучного інтелекту для поліпшення різних процесів державних закупівель. Використання ШІ не лише сприяє ефективності та прозорості закупівель, але й відкриває нові горизонти для їх оптимізації, для більш досконалого управління державними ресурсами.

## РОЗДІЛ 3. АНАЛІЗ ДАНИХ, ПОБУДОВА ТА АРХІТЕКТУРА СИСТЕМИ ДЛЯ АНАЛІЗУ ДАНИХ ПРО ДЕЖАВНІ ЗАКУПІВЛІ

### 3.1 Вибір платформи та мови програмування

З огляду на тему вашої роботи, "Система інтелектуальної обробки даних про державні закупівлі на основі нейромереж", інструменти та технології можуть бути спрямовані на обробку великих об'ємів даних, аналіз та прогнозування з використанням нейронних мереж. Ось як може виглядати адаптований опис розділу: Для розробки системи інтелектуальної обробки даних про державні закупівлі, яка використовує нейромережі для прогнозування та аналізу, було обрано ряд технологій, що підходять для задач штучного інтелекту та веб-розробки. React[21] використовується для створення інтерактивного користувацького інтерфейсу, дозволяючи ефективно представляти результати аналізу та візуалізації даних. NestJS[22] забезпечує надійну структуру для серверної частини додатку, що вимагає обробки великих об'ємів даних і взаємодії з нейронними мережами.

OpenAI[23] інтегровано для підвищення можливостей нейронних мереж, зокрема для розширеного аналізу текстових даних, пов'язаних з державними закупівлями. Bootstrap та Victory Charts використовуються для створення зрозумілих та доступних інтерфейсів з візуалізацією даних, що полегшує інтерпретацію результатів аналізу.

Структура проекту передбачає розробку модульного додатку, де фронтенд забезпечує відображення інтерактивних діаграм, таблиць та інших елементів для візуалізації аналізу державних закупівель. Бекенд, створений на основі NestJS, займається обробкою запитів, взаємодією з базами даних та виконанням розрахунків з використанням нейронних мереж. Завдяки використанню сучасних інструментів і технологій, система не тільки забезпечує точний аналіз даних, але й є масштабованою та легкою у підтримці додаткових вище зазначених бібліотеки.

### 3.2 Аналіз платформи державних закупівель

Для аналізу державних закупівель було обрано сайт Prozorro, де публікуються тендери з державних закупівель. Використано дані про тендери, їх учасників, обсяги та результати закупівель. Основа дослідження — інформація з сайту [prozorro.gov.ua](http://prozorro.gov.ua), яка охоплює широкий спектр тендерних процедур та їх характеристик.

Додано примітку [1]: посилання на прозоро


Замовник: управління житлово-комунального господарства виконавчого комітету полтавської міської ради - 03365854	
Знайдено 4,524 тендерів	
Сортування: за замовчуванням	
<b>Поточний ремонт (негайна ліквідація пошкоджень) вулично-дорожньої мережі по просп. Миру в м. Полтаві</b> Завершена	Очікувана вартість <b>181 451,76 UAH</b>
Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради • м. Полтава ID: UA-2024-01-01-001424-a	
<b>Поточний ремонт (негайна ліквідація пошкоджень) вулично-дорожньої мережі по вул. Європейській в м. Полтаві</b> Завершена	Очікувана вартість <b>18 137,71 UAH</b>
Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради • м. Полтава ID: UA-2024-01-01-001201-a	
<b>Поточний ремонт (негайна ліквідація пошкоджень) вулично-дорожньої мережі по вул. Комарова ( вул. Польська) в м. Полтаві</b> Завершена	Очікувана вартість <b>85 129,74 UAH</b>
Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради • м. Полтава ID: UA-2023-12-29-002569-a	
<b>Додаткові послуги№ 4: Поточний ремонт покрівель, настінних жолобів, карни-</b>	Очікувана вартість

Рисунок 3.1 - результат фільтрації даних по замовнику з сайту prozorro

Євроконтейнери оцинковані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ, Контейнери для порталного сміттєвозу (контейнери для великогабаритних відходів)

UA-2023-10-04-009570-a • b42a52c5435f423781f255dc74f26950

Відкриті торги з особливостями **Торги відмінено**  
Електронний підпис накладено. [Перевірити](#)

Всі закупівлі замовника 

📌 На що звернути увагу потенційному постачальнику

оцінювана вартість  
**6 580 000,00** ЦАН

[Протокол розкриття](#)  
**Контакти**  
Ігор Ланюк  
+380532561938  
igorlts190@gmail.com

**ДЕРЖАВНА АУДИТОРСЬКА СЛУЖБА УКРАЇНИ**  
Знайшли порушення законодавства у сфері закупівель?  
[Інформувати про порушення](#)

📌 У яких випадках інформувати Держаудитслужбу

**DOZORRO**  
Відук про закупівлю на Dozorro

**Оголошення про проведення**  
Друквати оголошення [PDF](#) • [HTML](#)  
Друквати звіт про результати проведення процедури [PDF](#) • [HTML](#)  
[Перейти до обґрунтування та плану закупівлі](#)

**Інформація про замовників**  
📌 Як обрати замовника потенційним постачальникам

**Найменування:** **Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради**

Рисунок 3.2 - Основна детальна інформація про тендер

Прозоро надає таку інформацію про закупівлю:

1. Суму закупівлі
2. Назву закупівлі
3. Контакти замовника

## Інформація про процедуру

📍 [Гід по строкам проведення торгів](#)

<b>Дата оприлюднення:</b>	04 жовтня 2023 · 15:21
<b>Звернення за роз'ясненнями:</b>	до 09 жовтня 2023 · 00:00
<b>Оскарження умов закупівлі:</b>	до 09 жовтня 2023 · 00:00
<b>Кінцевий строк подання тендерних пропозицій:</b>	12 жовтня 2023 · 00:00
<b>Очікувана вартість:</b>	6 580 000,00 UAH з ПДВ
<b>Розмір мінімального кроку пониження ціни:</b>	65 800,00 UAH
<b>Розмір мінімального кроку пониження ціни, %:</b>	1,00%
<b>Мова (мови), якою (якими) повинні готуватися тендерні пропозиції:</b>	Мова тендерної пропозиції українська

## Інформація про предмет закупівлі

**Вид предмету закупівлі:** Товари

**Класифікатор та його відповідний код:** ДК 021:2015:44610000-9: Цистерни, резервуари, контейнери та посудини високого тиску

**Опис окремої частини або частин предмета закупівлі**

200 штуки

Євроконтейнери оцинковані (сміткові контейнери) 1,1 куб.м для збирання змішаних ТПВ

Місце поставки товарів або місце виконання робіт чи надання послуг: 36000, Україна, Полтавська область, м. Полтава, вул. Кагамлика, 84

Строк поставки товарів, виконання робіт чи надання послуг: 31 грудня 2023

ДК 021:2015: 44610000-9 — Цистерни, резервуари, контейнери та посудини високого тиску

20 штуки

Контейнери для портального смітвозову (контейнери для великобаритних відходів)

Місце поставки товарів або місце виконання робіт чи надання послуг: 36000, Україна, Полтавська область, м. Полтава, вул. Кагамлика, 84

Строк поставки товарів, виконання робіт чи надання послуг: 31 грудня 2023

ДК 021:2015: 44610000-9 — Цистерни, резервуари, контейнери та посудини високого тиску

Умови оплати договору (порядок здійснення розрахунків)

Рисунок 3.3 - Детальна інформація про процедуру та предмет закупівлі

Прозоро надає таку інформацію про закупівлю:

1. Дата закупівлі
2. Строки звернення за роз'ясненнями
3. Строки оскарження умов закупівлі
4. Кінцеві строки подання тендерних пропозицій
5. Очікувана вартість
6. Опис предмета закупівлі
7. Строки поставки товарів
8. Строки виконання робіт

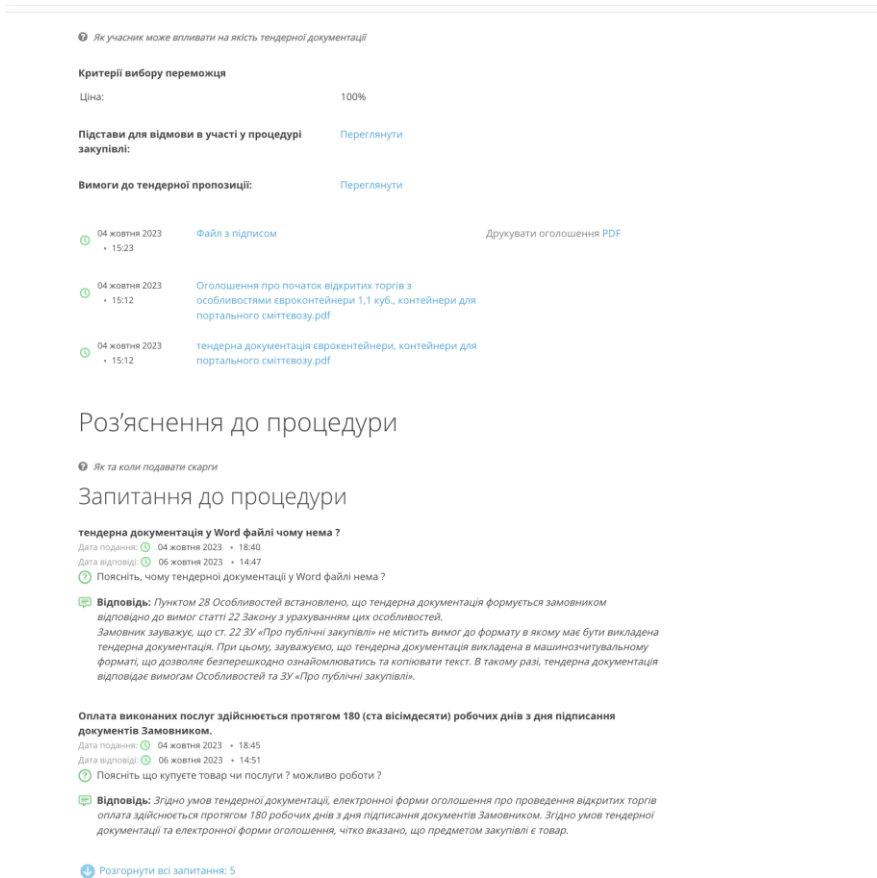


Рисунок 3.4 - Детальна інформація про критерії закупівлі

Прозоро надає таку інформацію про закупівлю:

1. Критерії вибору переможця
2. Підстави для відмову
3. Запитання та відповіді до процедури
4. Строки оплати тендера

Як бачимо на сайті про будь який тендер висвітлено дуже багато не зрозумілої інформації для пересічних людей. Щоб розібратись в ньому на це піде не одна година оскільки в ньому багато “підводних камнів”. З першого погляду цей тендер не відрізняється від сотні тисяч інших тендерів, але насправді він повністю корупційний. Постачальник і замовник хотіли “наваритись” на ньому на 2-3 мільони гривень що являється половиною суми очікуваної на закупівлю.

Коли йдеться про суть корупції у сфері закупівель, важливо вийти за межі спрощеного розуміння, яке обмежується лише завищенням очікуваних цін порівняно з ринковими. Це є поверхневим поглядом, який не в змозі охопити багатомірну та системну природу корупції в цій сфері. Насправді, корупція в закупівлях не тільки про платіжні перевищення над ринковими цінами; це, перш за все, про маніпуляцію процесами для вигоди обраних осіб. Корупція в закупівлях часто включає динаміку "свого постачальника", де певним продавцям надаються неправомірні переваги завдяки їхнім зв'язкам або угодам з тими, хто у владі.

### **3.3 Аналіз структури вхідних даних**

Використовуємо REST API, надане платформою Prozorro. REST API, що розшифровується як Representational State Transfer Application Programming Interface, представляє собою стандартний набір протоколів і інструментів для взаємодії з веб-ресурсами. Воно дозволяє здійснювати запити та отримувати відповіді через HTTP, що є універсальною мовою спілкування в інтернеті. Цей інструмент є невід'ємною частиною моєї роботи, оскільки він забезпечує доступ до великого масиву даних, який містить всю необхідну інформацію про державні тендери, їх учасників, а також результати торгів. Аналіз цих даних є ключовим для розуміння динаміки та ефективності системи державних закупівель в Україні. Через REST API Prozorro я маю змогу здійснювати фільтрацію, сортування та агрегацію даних, що відкриває широкі можливості для проведення глибокого структурованого аналізу.

Ось структура даних яка повертає платформою prozorro, з якою працює система штучного інтелекту для аналізу даних.



Рисунок 3.5 - Перший рівень вкладеності відповіді /tenders

Результат запиту /tenders/id - віддає нам об'єкт в якому є поля:

1. plans - масив в якому містяться id за допомогою яких можна отримати детальну інформацію про планінг тендера
2. documents - масив в якому містяться об'єкти в з інформацію про документ

3. value - поле в якому містяться дані про суму грошей в тендері
4. guarantee - поле в якому містяться дані про гарантію тендеру
5. milestones - поле в якому міститься масив з об'єктами даних про важливі періоди часу в цьому тендері
6. procuringEntity - поле в якому міститься дані про закупівельника
7. tenderPeriod - данні в якому містяться про початок та кінець тендера
8. items - масив в якому містять об'єкти з даними про предмети закупівлі
9. questions - масив в якому містяться питання від користувачів щодо тендеру

Детальніше розглянемо декілька важливих полів які сильно впливають на результат а саме - (documents,milestones,items,questions).

```
documentType: biddingDocum..
description: SHA1:97393bfe..
relatedItem: b42a52c5435f4..
id: 16fcc60b8a2e4d56856c1e..
datePublished: 2023-10-04T..
hash: md5:adac6b1e56338643..
title: тендерна документац..
format: application/pdf
url: https://public.api.op..
documentOf: tender
dateModified: 2023-10-04T1..
author: tender_owner
language: uk
```

Рисунок 3.6 - Огляд поля documents

Поле documents має декілька важливих полів а саме:

1. documentType - якщо значення biddingDocuments - це означає що цей файл містить детальну інформацію про критерії до об'єкта закупки
2. title - назва файлу, вона допоможе візуально розібратись про що йдеться в файлі

3. `format` - тип файлу, часто використовуюється для аналізу тендерів оскільки багато форматів є незручними для користування
4. `url` - посилання для скачування файлу



Рисунок 3.7 - Огляд поля milestones

Поле milestones має декілька важливих полів а саме:

1. `code` - поле яке несе в собі інформацію про тип оплати
2. `percentage` - поле яке вказує скільки оплатять в цьому майлстоні
3. `duration` - поле яке містить інформацію про період оплати



Рисунок 3.8 - Огляд поля items

Поле items має декілька важливих полів а саме:

1. unit - відповідає за тип послуги, який запрошується в тендері
2. deliveryDate - крайній срок до якого обект тендеру повинен бути доставлений
3. description - описує що за послугу потрібно надати
4. quantity - кількість яку потрібно надати

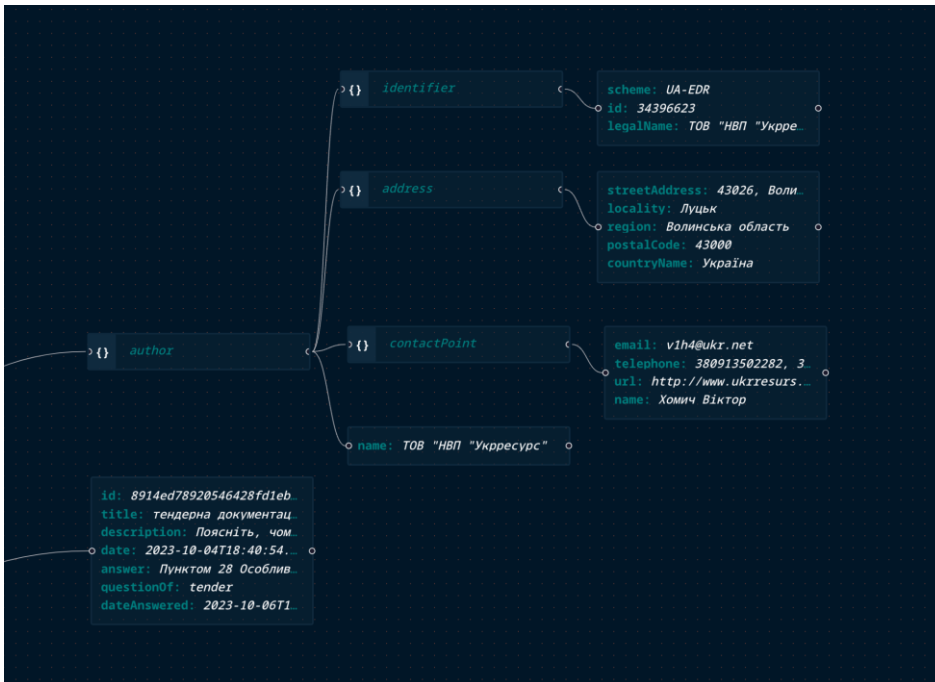


Рисунок 3.9 - Огляд поля questions

Поле questions має декілька важлих полів а саме:

1. description - містить текст питання користувача
2. answer - відповідь від замовника
3. legalName - назва користувача що задав питання

### 3.4 Попередня обробка даних

Для подальшої роботи нам потрібно обробити данні які було отримано використовуючи API та структуру даних яка повертається з Prozorro api. Відобразимо їх в читабельному форматі використовуючи React, Javascript.

**Євроконтейнери оцинковані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ, Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів)**

**UAH 6,580,000.00** включно з податками

**Предмет закупівлі**

200 штук: Євроконтейнери оцинковані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ до 31.12.2023, 00:00:00 (87 днів )  
20 штук: Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів) до 31.12.2023, 00:00:00 (87 днів )

**Строки подачі заявок**

Оголошення тендера: 04.10.2023, 15:21:33  
Кінець подачі заявок: 12.10.2023, 00:00:00  
Кількість днів на подачі заявок: 7

**Оплата послуг**

Післяплата протягом 180 днів

**Документи**

[Оголошення про початок відривних тогів з особливостями євроконтейнерів 1,1 куб. м контейнери для портального сміттєвозу.pdf](#)  
[тендерна документація євроконтейнерів\\_контейнери для портального сміттєвозу.pdf](#)  
[gpo-073](#)

Рисунок 3.10 - Огляд відфільтрованої інформації

**Додаткові критерії**

Євроконтейнери оцинковані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ    Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів)

**Питання від користувачів**

ТОВ "НВП "Укрресурс":Посиліть, чому тендерної документації у Word файлі нема?  
ТОВ "НВП "Укрресурс":Посиліть що купуєте товар чи послуги? можливо роботи?  
ТОВ "НВП "Укрресурс":Посиліть розкласти предмет закупівлі на лоти: лот 1- Євроконтейнери оцинковані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ, лот 2 - Контейнери для портального сміттєвозу з метою економії бюджетних коштів, оскільки коштує укласти велику кількість послуги тільки Євроконтейнери оцинковані (сміттєві контейнери) 1,1 куб.м по меншій ціні та меншій вазі, проте не займається поставкою Контейнери для портального сміттєвозу. Чи у замовника вже є парк машин?  
ТОВ "НВП "Укрресурс":Посиліть причину вимоги - Корпус контейнера складається з 3 елементів з'єднаних суцільним заварювальним швом: Дно контейнера, Передня панель з бочними стінками? стандарт B40-3 не передбачає саме, що контейнер буде складатися з 3-х елементів.  
ТОВ ВСМ ЕКОЛОДЖІ:шануємо запитання, а де би бачили дно плоске (на U подібне) з висотом борту по периметру не менше 60 мм., скрин ЄДИНОГО Сербського виробника з постачальниками з Білорусії під назвою ТОВ Рактоз Груп? ТОВ Юа Про (це одне і теж)?

**Procuring Entity**

Name: Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради

**Address**

вул. Спрієвська, 19  
м. Полтава, Полтавська область 36020  
Україна

**Identifier**

ID: 03365854  
Legal Name: Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради  
Scheme: UA-EDR

**Contact Point**

Name: Тарз Ланос  
Email: tarzlanos@gmail.com  
Telephone: +380532561938

Рисунок 3.11 - Огляд відфільтрованої інформації

Використовуючи поля, отримані від API, було відфільтровано непотрібні данні та чітко визначено данні які використовуються в аналізі тендера.

Після цього потрібно згенерувати опис тендеру в читебальному форматі який використовується для подальшого аналізу.

Назва Тендера: Сервісоконтейнери оцінювані (сміттєві контейнери) 1,1 куб.м для збирання відходів ТПВ, Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів), був створений 04.10.2023, 16:08:52, Строки подані заванж: до 12.10.2023, 00:00:00 Служба - 6580000 Пивень включно з податки. Тип оплати: Післяплата протягом 180 днів. Прямий закупівлі: 200 штук Сервісоконтейнери оцінювані (сміттєві контейнери) 1,1 куб.м для збирання відходів ТПВ, Строки постави товарів, виконана робота чи надана послуга: 31.12.2023, 00:00:00, 20 штук Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів), Строки постави товарів, виконана робота чи надана послуга: 31.12.2023, 00:00:00, Питання від користувачів: Поясність, чому тендерні документи у Word файлі нема ?Поясність що купити товар чи послугу ?Поясність розділити предмет закупівлі на лоти - лот 1 - Сервісоконтейнери оцінювані (сміттєві контейнери) 1,1 куб.м для збирання відходів ТПВ, лот 2 - Контейнери для портального сміттєвозу з метою економії бюджетних коштів, зокрема зупинити умови як можуть поставити умови Сервісоконтейнери оцінювані (сміттєві контейнери) 1,1 куб.м по низькій ціні та високій якості, проте не займається поставкою Контейнери для портального сміттєвозу. Чи у Замовника вже є парковки ?Поясність причому жевони - корути с контейнери складається з 3 елементів з'єднаних судильним зварювальним швом. Діаг контейнери. Передня панель з б'єними стійками. Задня панель із б'єними стійками ? стандарт B40-3 не передбачає саме , що контейнер буде складатися з 3-х елементів.Шановний замовник, а де Ви бачили дно плоске (не U подібне) з висотою борту по периметру не менше 60 мм., окрім ЄДИНОГО Сербського виробника з постачальниками з білоусой під назвою ТОВ Ренток Групу ТОВ Юа Про (за однієї ціною)?

Рисунок 3.12 - Приклад згенерованого тексту для подальшого аналізу

### 3.5 Побудова моделі

У рамках магістерської роботи, для аналізу тендерів на вміст корупційної складової, розроблена модель з наступною архітектурою: fine-tuned версія GPT-4 від OpenAI, адаптована до контексту тендерних документів. Модель використовує глибоку архітектуру з багатьма шарами трансформерів, які на вхід приймають текстові дані з документів тендерів, проаналізовані та розмічені на основі корупційних індикаторів.

Трансформери в моделі дозволяють здійснювати глибоке семантичне аналізування тексту, вивчаючи контекст та зв'язки між словами та фразами, що є критично важливим для ідентифікації складних індикаторів корупції. Кожен шар трансформера складається з механізмів уваги, які аналізують та враховують взаємозв'язки між усіма словами в тексті, незалежно від їхньої позиції. На виході модель генерує вектори ембеддінгів для тендеру, що представляють семантичне узагальнення інформації, містять оцінки ризику та важливі корупційні індикатори. Ці ембеддінги подаються на вхід у класифікаційний шар, зазвичай з декількома прихованими шарами, що завершується вихідним шаром з функцією активації, яка прогнозує ймовірність корупційної складової.

Для тренування та налаштування моделі використовується набір даних з анотованими тендерами, де кожен документ містить позначення про наявність або відсутність корупційних ознак. Використовуючи методики машинного навчання, зокрема, стохастичний градієнтний спуск, модель ітеративно оновлює свої ваги для мінімізації функції втрат, зазвичай MSE (середньоквадратична помилка) або cross-

ентропу, для покращення точності прогнозування. Завдяки цьому підходу модель здатна виявляти складні та непрямі індикатори корупції.

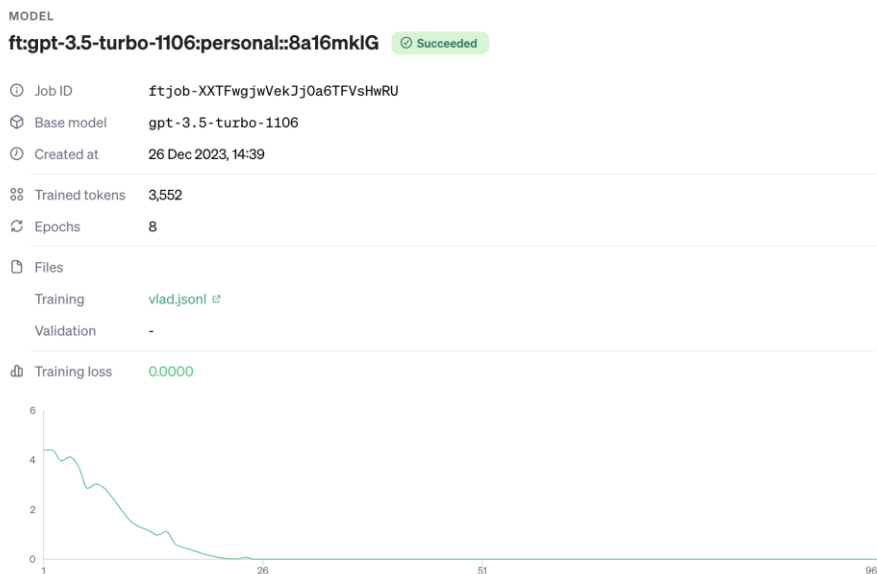


Рисунок 3.13 - Функція втрат під час навчання моделі

### 3.6 Архітектура системи

Побудована модель є основною для прогнозування того, що тендер корупційний (рис 3.14), оскільки в ній міститься вся логіка. Всі інші учасники архітектури відповідають за надання даних для подальшого прогнозу

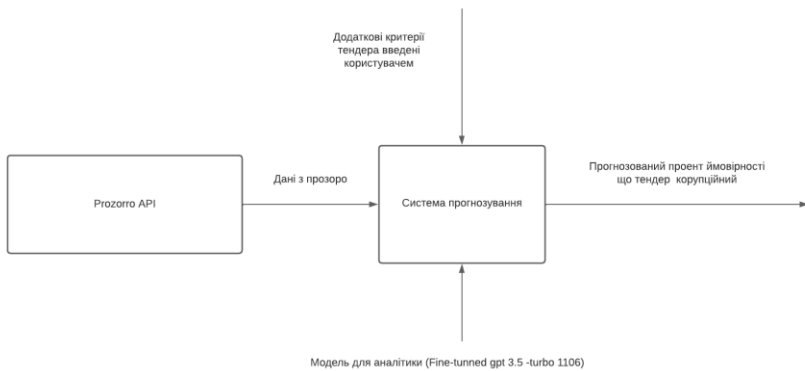


Рисунок 3.14 - Структура системи прогнозування корупційної складової

Для роботи системи прогнозування необхідні дані на вхід:

- інформація про тендер від prozorro API
- інформація від користувача про додаткові критерії
- модель GPT

Дані від системи Prozorro фільтруються щоб користувач міг з ними легко та зрозуміло працювати. Після чого користувач може надати додаткові вимоги, які легко знайти, оскільки файли з технічними критеріями висвітлюється в видному місці. Дані трансформуються в зручний для моделі формат для прогнозування, в результаті чого отримується прогноз

### **Висновки до розділу**

У цьому розділі було проведено детальний аналіз структури даних з системи ProzoGo, включаючи опис полів, їхню попередню обробку та фільтрацію для оптимізації подальшого аналізу. Використання налаштованої моделі штучного інтелекту дозволило значно підвищити точність виявлення ризиків корупції та нераціонального використання державних коштів. Архітектура системи була ретельно розроблена з метою забезпечення гнучкості, масштабованості та високої продуктивності, що робить її ефективним інструментом для моніторингу та аналізу державних закупівель.

## **РОЗДІЛ 4. РЕАЛІЗАЦІЯ СИСТЕМИ ІНТЕЛКТУАЛЬНОЇ ОБРОБКИ ДАНИХ ПРО ДЕРЖАВНІ ЗАКУПІВЛІ НА ОСНОВІ НЕЙРОМЕРЕЖ**

### **4.1 Опис роботи системи**

Для того, щоб створити систему інтелектуальної обробки даних про державні закупівлі на основі нейромереж потрібно виконати наступні дії:

1. Отримати дані від Prozorro API
2. Розібрати та відфільтрувати непотрібні дані від інформаційного “шуму”
3. Надати можливість дописувати додаткові критерії користувачем, якщо якісь дані були втрачені
4. Отримати доступ до моделі gpt3.5-turbo-1106 або gpt 4
5. Перевести дані про тендер в формат який підходить для текстової моделі
6. Вивести результат роботи моделі

Вибираємо id тендера з Prozorro. Програма повністю обробить всі дані про цей тендер та виведе найбільш потрібні, розпарсивши предмети закупівлі. Користувач надає додаткові критерії до кожної з предмета закупівлі, якщо такі є. Згенерувати опис тендера та відправити моделі на обробку. Модель, яка має дуже широкі знання у всіх сферах, максимально підходить для цієї роботи, а додаткові налаштування виводять результативність на максимум. Отримавши згенерований опис тендера, модель видає свій висновок про можливу ймовірність корупційної складової в тендері, яку користувач по бачить на сайті

## 4.2 Розроблення системи

### 4.2.1 Отримання даних від системи Prozorro API

Перш за все потрібно мати сервер на порті 3000. Для цього я обрав популярний framework NestJS, оскільки він надає модульність, масштабованість та його використовують тисячі відомих продуктів.

```
import { NestFactory } from '@nestjs/core';
import { AppModule } from './app.module';
import OpenAI from 'openai';

async function bootstrap() {
  const app = await NestFactory.create(AppModule);
  app.enableCors();
  await app.listen(3000);
}
bootstrap();
```

Тепер нам потрібно забезпечити оброблення запитів.

```
1 import { Module } from '@nestjs/common';
2 import { AppController } from './app.controller';
3 import { AppService } from './app.service';
4 import { TendersModule } from './tenders/tenders.module';
5
6 export const PROZORO_API_URL =
7   'https://public.api.openprocurement.org/api/2.5/';
8
9 @Module({
10   imports: [TendersModule],
11   controllers: [AppController],
12   providers: [AppService],
13 })
14 export class AppModule {}
15
```

Підєднаємо TendersModule який відповідає за функціонал, зв'язаний з тендерами.

```

1 import { Module } from '@nestjs/common';
2 import { TendersService } from './tenders.service';
3 import { TendersController } from './tenders.controller';
4 import { OpenAIService } from 'src/services/open-ai.service';
5
6 @Module({
7   controllers: [TendersController],
8   providers: [TendersService, OpenAIService],
9 })
0 export class TendersModule {}
1

```

TendersModule, має в собі контролер TendersController - саме він відповідає за оброблення запитів які надходять з клієнта. TendersService - містить в собі набірв функцій які оброблять дані отримані з клієнта.

```

@Controller('tenders')
export class TendersController {
  constructor(
    private readonly tendersService: TendersService,
    private readonly openAIService: OpenAIService,
  ) {}

  @Get('all')
  async getAll(@Res() res: Response) {
    const shortTendersInfo = await this.tendersService.getAll();

    if (shortTendersInfo) {
      res.status(HttpStatus.OK).json({ data: shortTendersInfo });
    } else {
      res
        .status(HttpStatus.BAD_REQUEST)
        .json({ error: 'Something went wrong' });
    }
  }
}

```

Отримавши від клієнта запит /all на отримання даних про останні 10 тендерів, сервер викликає функцію getAll tendersService

```

export class TendersService {
  async getAll(): Promise<TenderShortInfo[] | null> {
    try {
      const response = await fetch(PROZORO_API_URL + `tenders`);
      const json = await response.json();
      if (json.data) {
        return json.data as TenderShortInfo[];
      } else return null;
    } catch (err) {
      console.error(err);
    }
  }
}

```

getAll робить запит на prozorro\_api\_url обробляє отримані дані та повертає тільки потрібну коротку інформацію.

Для відображення даних на клієнті я використовую React - популярна бібліотека для маніпуляції з ДОМ-деревом. Mobx - для роботи з станом програми.

```

You, 3 weeks ago | 1 author (You)
import React, { memo } from "react";
import "./App.css";
import ioc from "./ioc";
import { AppBLoC } from "./App.bloc";
import { BrowserRouter, Navigate, Route, Routes } from "react-router-dom";
import { Observer } from "mobx-react-lite";
import { Main } from "./pages/Main";
import { DetailTenderInfoPage } from "./pages/TenderInfo/TenderInfo.page";

export const App = memo(() => {
  const bloc = ioc.useBLoC2(AppBLoC, {});

  return (
    <Observer>
      {() => (
        <BrowserRouter>
          <Routes>
            <Route path="/main" element={<Main />} />
            { /* </Route> */ }
            <Route path="/tenders/:uid" element={<DetailTenderInfoPage />} />
            <Route path="*" element={<Navigate to={"/main"} replace />} />
          </Routes>
        </BrowserRouter>
      )}
    </Observer>
  );
});
export default App;

```

Потрібно вказати шлях url, щоб знати які сторінки відображати на якому посиланні. Дефолтним посиланням з цього коду, видно, що буде /main який на якому відрендериться компонент <Main/>

```
8 import { useCallback, useMemo } from 'react';
9
10 export const Main = memo((props) => {
11   const navigate = useNavigate();
12
13   const bloc = ioc.useBLoC2(MainBLoC, props);
14   const handleClick = (uid: string) => {
15     bloc.onDetailTenderClick(() => navigate(`/${uid}`));
16   };
17   return (
18     <Observer>
19       {() => (
20         <ListContainer items={bloc.shortTenderInfo} onClick={handleClick} />
21       )}
22     </Observer>
23   );
24 };
25 const ListItem = ({
26   item,
27   onClick,
28 }): {
29   onClick: () => void;
30   item: DetailTenderInfo;
31 } => (
32   <div className="list-item" onClick={onClick}>
33     <div className="list-item-header">
34       <span className="list-item-id">{item.id}</span>
35       <span className="list-item-date">
36         {parseISO(item.dateModified).toLocaleDateString()}
37       </span>
38     </div>
39     <div className="list-item-header">
40       <span className="list-item-id" style={{ fontWeight: 700 }}>
41         {item.title}
42       </span>
43       <span
44         className="list-item-date"
45         style={{ color: "green", fontWeight: 700 }}
46       >
47         {item.value.amount.toLocaleString("en-US", {
48           style: "currency",
49           currency: "UAH",
50         })}
51       </span>
52     </div>
53   </div>
54 );
55
56 export type ListContainerProps = {
57   items: DetailTenderInfo[];
58   onClick: (uid: string) => void;
59 };
60 const ListContainer = ({ items, onClick }: ListContainerProps) => (
61   <div className="list-container">
62     {items.map((item) => (
63       <ListItem key={item.id} item={item} onClick={() => onClick(item.id)} />
64     ))}
65   </div>
66 );
67
```

Рисунок 4.2 - реалізація компоненту <Main/>

Як видно з коду, компонент <Main/> рендерить listContainer який ітерується через дані отримані з нашого сервера, та рендерить для кожного відповідний ListItem.

За менеджмент стейту відповідає MainBloc.

```
@injectable()
export class MainBloc implements IBLoC {
  @inject(NavigationParamsSymbol) params!: any;
  @inject(InitialPropsSymbol) props!: any;
  @inject(TenderServiceSymbol) tenderService!: ITenderService;
  @inject(TenderStoreSymbol) tenderStore!: ITenderStore;
  constructor() {
    makeAutoObservable(this);
    setImmediate(() => this.initialize());
  }
  @action initialize = async () => {
    if (this.shortTenderInfo?.length == 0) {
      await this.tenderService.syncShortTenderInfo();
    }
    const shortTenderInfo = this.tenderStore.getShortTenderInfo();
    await Promise.all(
      shortTenderInfo.map((el) =>
        this.tenderService.syncDetailTenderInfo(el.id)
      )
    );
  };

  onDetailTenderClick = async (navigate: () => void) => {
    navigate();
  };
  @computed get shortTenderInfo(): DetailTenderInfo[] {
    return this.tenderStore.getAllDetailTenderInfo() ?? [];
  }

  onParamsChange(props: any, params: any) {
    this.params = params;
    this.props = props;
  }
}

/*"1e4a279d7aa24b3a9ba3d0f9ffdb3906"*/
```

Рисунок 4.3 - реалізація MainBloc

Як видно з реалізації, в конструкторі класа MainBloc виконується запит syncShortTenderInfo, який робить запит на наш сервер та отримує данні для показу.

0072c6eb8a544b308b0f48d8073db94e	06/12/2023
<b>Додаткові послуги№3: Поточний ремонт покрівель, настінних жолобів, карнизних звисів та водостічних труб житлових будинків та нежитлових будівлях (приміщення) у місті Полтава</b>	
	<b>UAH 142,803.49</b>
19ccfb76bf92439288b2ef151285c2cf	06/12/2023
<b>Поточний ремонт зупинки громадського транспорту с. Івренці (Клуб) (на с. Івонченці), а саме: павільйонів, лав для сидіння, інших елементів та улаштування посадкового майданчика у м. Полтава, що здійснюється у 2023 році</b>	
	<b>UAH 133,154.35</b>
30c61ea5c30948e98aa7b354a5acf254	04/01/2024
<b>Провід самоутримний з ізоляцією</b>	
	<b>UAH 3,727,000.00</b>
4bf9561893244c09915d044dfb5c38c1	03/01/2024
<b>«ДК 021:2015 – 15850000-1 Макаронні вироби»</b>	
	<b>UAH 998,400.00</b>
5a8203b8ba1c40c1924866566bd0d076	11/12/2023
<b>Поточний ремонт зупинки громадського транспорту вул. Адміністративна (на с. Рибці), а саме: павільйонів, лав для сидіння, інших елементів та улаштування посадкового майданчика у м. Полтава, що здійснюється в 2023 році</b>	
	<b>UAH 90,987.11</b>
8c0b2445c9154d50ad94123853c7604d	15/12/2023
<b>Контейнери для портального сміттевозу (контейнер для великогабаритних відходів)</b>	
	<b>UAH 1,800,000.00</b>

Рисунок 4.4 - результат роботи клієнта з сервером по отриманні коротких даних пиро тендера

#### 4.2.2 Розібрати та відфільтрувати дані від непотрібного “шуму”

При натисканні на елемент з списку тендерів користувач буде перенаправлений на сторінку `tenders/:uid`. Де буде відображено детальна, оброблена та відфільтрована інформація про тендер.

```

export const App = memo(() => {
  const bloc = ioc.useBLoC2(AppBLoC, {});

  return (
    <Observer>
      {() => (
        <BrowserRouter>
          <Routes>
            <Route path="/main" element={<Main />} />
            { /* </Route> */ }
            <Route path="/tenders/:uid" element={<DetailTenderInfoPage />} />
            <Route path="*" element={<Navigate to={"/main"} replace />} />
          </Routes>
        </BrowserRouter>
      )}
    </Observer>
  );
});

```

За це відповідає роутинг в корені програми. При переході повинен відрендеритись компонент `DetailTenderInfoPage`. Який відображає відфільтровану інформацію отриману з сервера за запитом `/tenders/:id`.

```

async getDetailTenderInfo(id: string): Promise<DetailTenderInfo> {
  try {
    const response = await fetch(PROZORO_API_URL + `tenders/${id}`);
    const json = await response.json();
    if (json.data) {
      return json.data as DetailTenderInfo;
    } else return null;
  } catch (err) {
    console.error(err);
  }
}

```

Порвернення результату як `DetailTenderInfo` - фільтрує дані від інформаційного “шуму” для подальшої роботи моделі.

```

export type DetailTenderInfo = {
  id: string;

  tenderID: string;
  owner: string;
  title: string;
  mainProcurementCategory: string;
  procurementMethod: string;
  procurementMethodType: string;
  procuringEntity: {
    name: string;
    identifier: {
      id: string;
      legalName: string;
      scheme: string;
    };
    address: {
      countryName: string;
      locality: string;
      postalCode: string;
      region: string;
      streetAddress: string;
    };
    contactPoint: {
      email: string;
      name: string;
      telephone: string;
    };
    kind: string;
  };
  status: string;

  value: {
    amount: number;
    currency: string;
    valueAddedTaxIncluded: boolean;
  };

  items: {
    classification: { description: string };
    deliveryDate: { startDate: string; endDate: string };
    quantity: number;
    unit: { name: string };
    description: string;
  }[];

  plans: { id: string }[];
  dateModified: string;
  dateCreated: string;
  date: string;

  tenderPeriod?: { startDate: string; endDate: string };
  questions?: {
    description: string;
    answer?: string;
    author?: { name: string };
  }[];
  milestones?: {
    code: string;
    duration: { days: number; type: string };
    type: string;
  }[];
  documents: {
    documentType: string;
    format: string;
    url: string;
    title: string;
  }[];
};

```

Рисунок 4.5 Інтерфейс детальної інформації про тендер

Отримавши цю інформацію, ми можемо обробити її на клієнті та показати користувачу для подальшої роботи.

```

export const DetailTenderInfoComponent = memo(
  (props: DetailTenderInfoProps) => {
    const bloc = ioc.useBloc2(DetailTenderInfoBloc, props);
    return (
      <Observer>
        {() => (
          <div className="tender-details">
            <h2 className="title">{bloc.tender.title}</h2>
            <PriceBlock value={bloc.tender.value} />

            {bloc.tender.items && (
              <PurchaseSubject
                items={bloc.tender.items}
                startDate={
                  bloc.tender.tenderPeriod?.startDate ?? bloc.tender.date
                }
              />
            )}
            {bloc.tender.tenderPeriod && (
              <ApplicationDates tenderPeriod={bloc.tender.tenderPeriod} />
            )}

            {bloc.paymentInfo && <PaymentInfo paymentInfo={bloc.paymentInfo} />}
            {bloc.tender.documents && (
              <DocumentInfo docs={bloc.tender.documents} />
            )}
            {bloc.tender.items && (
              <AdditionalRequirementsComponent tender={bloc.tender} />
            )}
            {bloc.tender?.questions && (
              <Questions questions={bloc.tender.questions} />
            )}

            <div className="procuring-entity">
              <h3>Procuring Entity</h3>
              <span>Name: {bloc.tender.procuringEntity.name}</span>
              <AddressBlock address={bloc.tender.procuringEntity.address} />
              <IdentifierBlock
                identifier={bloc.tender.procuringEntity.identifier}
              />
              <ContactPointBlock
                contactPoint={bloc.tender.procuringEntity.contactPoint}
              />
            </div>

            <div className="is-flex is-flex-direction-row is-justify-content-center mb-3">
              <Button
                label="Згенерувати опис тендера"
                onClick={bloc.onGenerateDescriptionClick}
              />
            </div>
            {bloc.tenderDescription && (
              <div className="logic-block">
                <p>{bloc.tenderDescription}</p>
              </div>
            )}
          </div>
        )}
      </Observer>
    );
  }
);

```

Рисунок 4.6 - Компонент DetailTenderInfoComponent

Цей компонент репрезентує дані з попередньою обробкою для того щоб користувачу було легко працювати з інформацією

**Євроконтейнери оциквані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ, Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів)**

**UAN 6,580,000.00** включно з податками

**Предмет закупівлі**  
200 штук: Євроконтейнери оциквані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ до 31.12.2023, 00:00:00 (87 днів)  
20 штук: Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів) до 31.12.2023, 00:00:00 (87 днів)

**Строки подачі заявок**  
Оголошення тендера: 04.10.2023, 15:21:33  
Кінець подачі заявок: 12.10.2023, 00:00:00  
Кількість днів на подачу заявок: 7

**Оплата послуг**  
Послугата протягом 180 днів

**Документи**  
[Оголошення про продаж відеонаписів записів з особливостями євроконтейнерів 1,1 куб.м - контейнери для портального сміттєвозу.pdf](#)  
[Заявка на придбання євроконтейнерів, контейнерів для портального сміттєвозу.pdf](#)  
[irola.ua](#)

**Документи**  
[Оголошення про продаж відеонаписів записів з особливостями євроконтейнерів 1,1 куб.м - контейнери для портального сміттєвозу.pdf](#)  
[Заявка на придбання євроконтейнерів, контейнерів для портального сміттєвозу.pdf](#)  
[irola.ua](#)

**Додаткові критерії**

Євроконтейнери оциквані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ	Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів)
<input type="text"/>	<input type="text"/>
<a href="#">Додати критерію</a>	<a href="#">Додати критерію</a>

**Питання від користувачів**

ТОВ "НВП "Укрресурс" Поясніть, чому тендерної документації у Word файлі нема ?  
ТОВ "НВП "Укрресурс" Поясніть що купуєте товар чи послуги ? можливо роботи ?  
ТОВ "НВП "Укрресурс" Просимо розділити предмет закупівлі на лоти : лот 1 - Євроконтейнери оциквані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ, лот 2 - Контейнери для портального сміттєвозу з метою економії бюджетних коштів, оскільки існують учасники які можуть поставити тільки Євроконтейнери оциквані (сміттєві контейнери) 1,1 куб.м по низькій ціні та високій якості, проте не займаються поставанням Контейнери для портального сміттєвозу. Чи у Замовника вже є переможць ?  
ТОВ "НВП "Укрресурс" Поясніть причину вимоги - Корпус контейнера складається з 3 елементів з'єднаних суцільним заварювальним швом. Дно контейнера, Передня панель з бічними стінками. Задня панель із бічними стінками ? стандарт 840-3 не передбачає саме, що контейнер буде складатися з 3-х елементів.  
ТОВ ВСМ ЕКОЛОДЖІ Шановний замовник, а де Ви бачили дно плоске (не U подібне) з висотою борту по периметру не менше 60 мм., окрім ЄДИНОГО Сербського виробника з постачальниками з білорусії під назвою ТОВ Ректор Груп? ТОВ Юа Про (це одне і теж)?

**Procuring Entity**  
Name: Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради

**Address**  
вул. Стрітенська, 19  
м. Полтава, Полтавська область 36020  
Україна

**Питання від користувачів**

ТОВ "НВП "Укрресурс" Поясніть, чому тендерної документації у Word файлі нема ?  
ТОВ "НВП "Укрресурс" Поясніть що купуєте товар чи послуги ? можливо роботи ?  
ТОВ "НВП "Укрресурс" Просимо розділити предмет закупівлі на лоти : лот 1 - Євроконтейнери оциквані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ, лот 2 - Контейнери для портального сміттєвозу з метою економії бюджетних коштів, оскільки існують учасники які можуть поставити тільки Євроконтейнери оциквані (сміттєві контейнери) 1,1 куб.м по низькій ціні та високій якості, проте не займаються поставанням Контейнери для портального сміттєвозу. Чи у Замовника вже є переможць ?  
ТОВ "НВП "Укрресурс" Поясніть причину вимоги - Корпус контейнера складається з 3 елементів з'єднаних суцільним заварювальним швом. Дно контейнера, Передня панель з бічними стінками. Задня панель із бічними стінками ? стандарт 840-3 не передбачає саме, що контейнер буде складатися з 3-х елементів.  
ТОВ ВСМ ЕКОЛОДЖІ Шановний замовник, а де Ви бачили дно плоске (не U подібне) з висотою борту по периметру не менше 60 мм., окрім ЄДИНОГО Сербського виробника з постачальниками з білорусії під назвою ТОВ Ректор Груп? ТОВ Юа Про (це одне і теж)?

**Procuring Entity**  
Name: Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради

**Address**  
вул. Стрітенська, 19  
м. Полтава, Полтавська область 36020  
Україна

**Identifier**  
ID: 03365854  
Legal Name: Управління житлово-комунального господарства виконавчого комітету Полтавської міської ради  
Scheme: UA-EDR

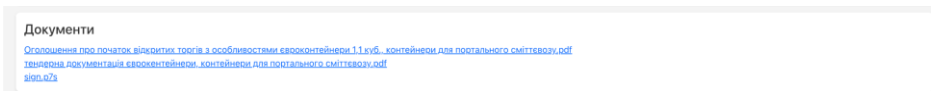
**Contact Point**  
Name: Ігор Лівко  
Email: igoruly190@gmail.com  
Telephone: +380532661938

[Згенерувати опис тендера](#)

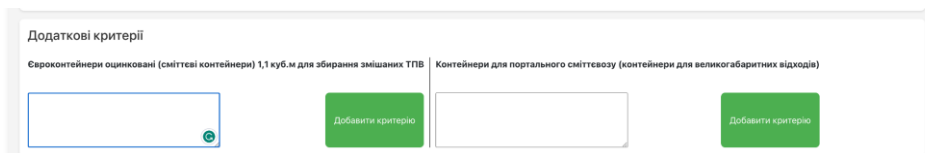
Рисунок 4.7 - Детальна інформація про тендер

### 4.2.3 Надати можливість дописувати додаткові критерії користувачем

Оскільки деякі тендери містять в собі документи, де вписані додаткові критерії для продукту, нема ніякої змоги отримати цю інформацію від прозоро. Тому найкраще що можна зробити це надати легкий доступ до документів через наш сервіс.



По кліку на документ, файл буде завантажений на ваш компютер. Найшовши додаткові критерії, користувач зможе надати цю інформацію до об'єктів закупівлі



За це відповідає компонент `AdditionalRequirement`

```
0 export const AdditionalRequirementsComponent = memo(  
1   (props: AdditionalRequirementsProps) => {  
2     const bloc = ioc.useBLoC2(AdditionalInfoBloc, props);  
3     return (  
4       <Observer>  
5         {() => (  
6           <div className="logic-block">  
7             <h4>Додаткові критерії</h4>  
8             <div  
9               className="is-flex is-flex-direction-row pt-3 is-flex-wrap-wrap"  
10              style={{ gap: "10px" }}>  
11               {bloc.PurchaseObjects.map((el, index) => (  
12                 <>  
13                   <PurchaseObjectRequirementsList  
14                     name={el.name}  
15                     index={index}  
16                     requirements={el.requirements}  
17                     onAddRequirement={bloc.onAddRequirement}  
18                     onRemoveRequirement={bloc.onRemoveRequirement}  
19                   />  
20                   {index !== bloc.PurchaseObjects.length - 1 && <Vr />}  
21                 </>  
22               )>  
23             </div>  
24           </div>  
25         )>  
26       </Observer>  
27     )>  
28   );  
29 }  
30 );
```

Рисунок 4.8 - `AdditionalRequirementComponent`

Як видно з реалізації він рендерить список об'єктів закупівлі до яких можна додати та видалити додаткові критерії.

```
export const PurchaseObjectRequirementsList = ({
  name,
  index,
  requirements,
  onAddRequirement,
  onRemoveRequirement,
}) => {
  name: string;
  index: number;
  requirements: string[];
  onAddRequirement: (purchaseObjectIndex: number, text: string) => void;
  onRemoveRequirement: (
    purchaseObjectIndex: number,
    requirementIndex: number
  ) => void;
}) => {
  const [input, setTextInput] = useState("");

  return (
    <div>
      <div style={{ fontWeight: "bold" }}>{name}</div>
      <ol>
        {requirements?.map((requirement, requirementIndex) => (
          <RequirementItem
            name={requirement}
            index={requirementIndex}
            onRequirementRemove={() =>
              onRemoveRequirement(index, requirementIndex)
            }
          </RequirementItem>
        ))}
      </ol>
      <div className="is-flex is-flex-direction-row pt-4 is-justify-content-space-between">
        <textarea
          style={{ width: "350px", height: "100px" }}
          value={input}
          onChange={(e) => {
            setTextInput(e.currentTarget.value);
          }}
        ></textarea>
        <button
          className="button"
          style={{ height: "100px" }}
          onClick={() => {
            onAddRequirement(index, input);
            setTextInput("");
          }}
        >
          Додати критерію
        </button>
      </div>
    </div>
  );
};

export const RequirementItem = ({
  name,
  index,
  onRequirementRemove,
}) => {
  name: string;
  index: number;
  onRequirementRemove: () => void;
}) => {
  return (
    <div style={{ position: "relative", display: "flex" }}>
      <div style={{ flex: 1 }}>
        <div>{index + 1}. {name}</div>
      </div>
      <div style={{ flex: 1, borderLeft: "1px solid black" }}></div>
      <CloseIcon onClick={() => onRequirementRemove()} />
    </div>
  );
};

export const Vr = () => {
  return (
    <div style={{ width: "1px", borderLeft: "1px solid black" }}></div>
  );
};
```

Данні зберігають в стореджі RequirementsStore в форматі : ключ - id тендера, значення: об'єкт, ключами якого є індекси а значенням - масиви строк.

```

@Injectable()
export class RequirementsStore implements IRequirementsStore {
  @observable requirements = observable.map<
    string,
    { [index: number]: string[] }
  >();

  constructor() {
    makeAutoObservable(this);
  }

  addRequirement(
    tenderId: string,
    purchaseObjectId: number,
    requirement: string
  ): void {
    const tenderObjectsRequirements = this.requirements.get(tenderId) ?? [];
    const objectRequirements =
      tenderObjectsRequirements[purchaseObjectId] ?? [];
    objectRequirements.push(requirement);
    this.requirements.set(tenderId, {
      ...tenderObjectsRequirements,
      [purchaseObjectId]: objectRequirements,
    });
  }

  removeRequirement(
    tenderId: string,
    purchaseObjectId: number,
    index: number
  ): void {
    const tenderRequirements = this.requirements.get(tenderId);
    if (!tenderRequirements || !tenderRequirements[purchaseObjectId]) {
      console.warn("No such purchase object or tender found");
      return;
    }
    if (index >= 0 && index < tenderRequirements[purchaseObjectId].length) {
      const updatedPurchaseObjectRequirements = (
        tenderRequirements[purchaseObjectId] ?? []
      ).filter((_, i) => i !== index);

      this.requirements.set(tenderId, {
        ...tenderRequirements,
        [purchaseObjectId]: updatedPurchaseObjectRequirements,
      });
    } else {
      console.warn("Invalid index for requirement removal");
    }
  }

  getRequirements(tenderId: string): { [index: number]: string[] } {
    return this.requirements.get(tenderId) ?? {};
  }
}

```

Рисунок 4.9 - Реалізація стореджа RequirementsStore.

**Додаткові критерії**

**Середоконтейнери оцинковані (сміттєві контейнери) 1,1 куб. м для збирання змішаних ТПВ**

1. 1. Середоконтейнери оцинковані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ 2. Об'єм 1100 літра. EN 840-3. Гарнеоцинкований. 3. Корпус контейнера складається з 3 елементів з'єднаних суцільним зварювальним швом. Дно контейнера. Передня панель з б'ичними стійками. Задня панель із б'ичними стійками. 4. Дно контейнера: Штамповане із цинкового листа металу без зварних з'єднань. Зі штампованими ребрами жорсткості на дні контейнера. Плоске (не U подібне) з висотою борту по периметру не менше 80 мм. Контейнер повинен стійко стояти на землі без коліс. Товщина металу не менше 1,5 мм. Зварний отвір. 5. Передня панель з б'ичними стійками: Штампована із цинкового листа металу без зварних з'єднань з привареними напрямними для кришки. З тривірними вертикальними штампованими ребрами жорсткості. 6. Штампованими горизонтальними ребрами жорсткості на кутах. 7. Штампована на передній стороні прямокутним заглубленим розміром не менше 100 (висота) на 250 (ширина) мм для розміщення наклейки з інформацією. Товщина металу не менше 1,25 мм. 8. Задня панель із б'ичними стійками: Штампована із цинкового листа металу без зварних з'єднань з привареними напрямними для кришки. 9. Штампованими горизонтальними ребрами жорсткості на кутах. 10. З тривірними вертикальними штампованими ребрами жорсткості. Товщина металу не менше 1,25 мм. 11. Опори коліс: Зварне кріплення до дна контейнера. Товщина металу не менше 3 мм

**Контейнери для портового сміттєвозу (контейнери для великогабаритних відходів)**

1. Контейнери для побутових відходів, типу: RLT (KMT), RLR (KCS), KKH (K38), KKO-MUL (K3P) об'єм 7000 літрів або 7 м<sup>3</sup>. 2. Стінки та дно контейнера мають бути підсилені ребра жорсткості. 3. Кришка висотою 5. Колеса висотою 8. Загальна вантажопідйомність 4 т. 9. Матеріал контейнера: Листовий метал 4 мм, 10 мм, шовері товщиною 4 мм, уголки 40\*40, 100\*100 круг 8 мм, 32 мм, 45 мм. 10. Колір зелений П1. Стандарти контейнери мають відповідати ДСТУ 8476:2015 «Контейнери для побутових відходів. Загальні технічні вимоги». 12. Габаритні розміри, мм: довжина 3590 (+/-5%) висота 1240 (+/-5%) ширина 1900 (+/-5%) 13. Гарантійний строк не менше 10 років. 14. Особливості: Контейнер покривається антикорозійним епоксидним грунтом та фарбується поліуретановою надстійкою емаллю

Рисунок 4.10 - Приклад добавлених критеріїв

### 4.2.3 Отримати доступ до моделі GPT-4

Для роботи з моделлю gpt-4 я використав бібліотеку openai. Оскільки вона містить хорошу документацію для демонстрації роботи моделі, а також являється однією з найвідоміших та найбільш натренованою моделю в світі.

Для початку початку напишемо сервіс OpenAiService, який містить підключення до моделі

```

3
4  ✓ const openai = new OpenAI({
5     apiKey: process.env.OPENAI_API_KEY,
6   });
7

```

Цей кусок кода відповідає за отримання доступу до моделі openai.

```

@Injectable()
export class OpenAIService {
  async getCorruptionPercentage(text: string): Promise<string> {
    try {
      const response = await openai.chat.completions.create({
        messages: [
          {
            role: 'system',
            content: 'Ви помічник, здатний аналізувати дані на наявність можливих ознак корупції в
          },
          {
            role: 'user',
            content: `Ось тендер: ${text};
            Дай відповідь в форматі: {процент корупції : (процент від 0 до 100), пояснення чо
          },
        ],
        model: 'gpt-4',
        temperature: 1,
      });

      return response.choices[0].message.content;
    } catch (err) {
      console.log(err);
    }
  }
  async getCorruptionExplanation(

```

Рис 4.11 - OpenAIService

Цей сервіс містить функціонал який відповідає за створення промптів до моделі, а також, налаштування temperature - що відповідає за надання моделі більшої “свободи”, тобто, вона пише не чистим патерном, який в неї попросили, а може сама шукати додаткові дані в інтернеті, писати свої “думки”. Також ми даємо мінімальні настанови моделі, щоб вона розуміла що саме вона повина зробити. Відбувається імітація чату, та ставиться запитання на яке модель повина дати відповідь.

#### 4.2.4 Перевести дані в формат який потрібний для моделі

GPT-4 - мовна модель, що працює з текстом. Тому ми повині отримати дані про тендер в текстовому форматі. Для цього ми генеруємо опис тендера на фронтенді.

## Згенерувати опис тендера

Нажимаючи на кнопку “Згенерувати опис тендера”, користувач отримує опис, який формується функцією `generateTenderDescription`

```
export const generateTenderDescriptionIpd = (
  tender: DetailTenderInfo,
  paymentInfo: {
    code: string;
    duration: {
      days: number;
      type: string;
    };
    type: string;
  } | null,
  purchaseObjects: {
    name: string;
    deliveryDate: string;
    quantity: number;
    unit: string;
    requirements: string[];
  }[],
  questions: string[]
) => {
  return `Назва Тендера: ${tender.title}, був створений ${parseISO(
    tender.dateCreated
  )}.toLocalDateString("uk-UA", {
    hour: "numeric",
    minute: "numeric",
    second: "numeric",
  })`;

  ${
    tender.tenderPeriod?.endDate
    ? `Строки подачі заявок: до ${parseISO(
      tender.tenderPeriod?.endDate
    )}.toLocalDateString("uk-UA", {
      hour: "numeric",
      minute: "numeric",
      second: "numeric",
    })`
    : ""
  }

  Оплата - ${tender.value.amount} гривень ${
    tender.value.valueAddedTaxIncluded ? "включно з" : "не рахуючи"
  } податки.

  ${
    paymentInfo
    ? `Тип оплати: ${translateLocal(paymentInfo.code)} протягом ${
      paymentInfo.duration.days
    }`
    : ""
  }

  днів;

  `;

  `Предмет закупівлі: ${purchaseSubject(purchaseObjects)}`

  ${
    questions?.length > 0
    ? `Питання від користувачів: ${questions.join(";")}`
    : ""
  }

  `;
};
```

Рис 4.12 - Реалізація функції для генерації опису тендера

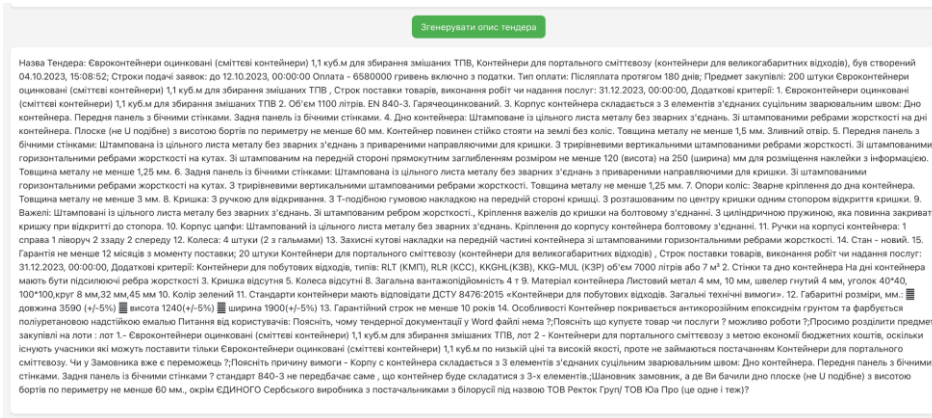


Рис 4.13 - Результат роботи функції для генерації опису тендера

### 4.3 Результати роботи програми

Для прикладу візьмемо скандальний тендер, який засвітився в інтернеті через свої корупційні признаки. Назва тендера - “Євроконтейнери оцинковані (сміттєві контейнери) 1,1 куб.м для збирання змішаних ТПВ, Контейнери для портального сміттєвозу (контейнери для великогабаритних відходів)”



Рисунок 4.14 - згенерований опис тендера

(процент корупції: 85, пояснення чому: Велика сума оплати вказує на потенційне завищення ціни. Зробивши додаткове дослідження, ціна за євроконтейнер становить близько 7000-11000 грн, залежно від виробника. За цими розрахунками, вартість 220 контейнерів повинна коливатися від 1,5 до 2,42 мільйонів грн. Тім часом, вартість тендера складає 6,58 млн грн, що в 2-4 рази вище основної вартості. Також є питання від користувачів, які вказують на специфічні вимоги до товару та пропозицію розділити тендер на два лоти, щоб підвищити конкурентність та знизити вартість. Ці коментарі можуть вказувати на те, що замовник має конкретного виробника на увазі, що є ознакою корупції.)

## Рисунок 4.15 - висновок моделі

Як бачимо, висновок справді дуже схожий на те що йдеться в статті.

<https://poltava.to/news/73459/>

Давайте візьмемо декілька тендерів про яких не знаємо ніякої інформації задалегідь.

«ДК 021:2015 – 15850000-1 Макаронні вироби»

УАН 998,400.00 включно з податками

Предмет закупівлі

12480 кілограм: «ДК 021:2015 – 15850000-1 Макаронні вироби» до 31.12.2024, 00:00:00 (362 днів)

Строки подачі заявок

Оголошення тендера: 03.01.2024, 14:53:50  
Кінець подачі заявок: 11.01.2024, 00:00:00  
Кількість днів на подачі заявок: 7

Оплата послуг

Післяплата протягом 30 днів

Документи

[\[1\]Д макарони.1178.2024.docx](#)  
[sign.p7s](#)

Додаткові критерії

«ДК 021:2015 – 15850000-1 Макаронні вироби»

Добавити критерію

Procuring Entity

Часте: Управління освіти Солом'янської районної в місті Києві державної адміністрації

Address

вулиця П'ятерська будинок 12

## Рисунок 4.16 - тендер “ДК 021:2015 – 15850000-1 Макаронні вироби”

Це один з найсвіжіших тендерів в прозоро на “кругленьку” суму, не читаючи нічого про нього, давайте отримаємо аналітику від моделі.

Аналітика тендеру корупції в цьому тендері

Для визначення потенціальних ознак корупції потрібно порівнювати ціни закупівлі в даному тендері з іншими тендерами або ринковими цінами. В даному випадку маємо замовлення на 12480 кг макаронних виробів за 998400 грн. Це означає, що ціна за кг складає приблизно 80 грн. Якщо порівняти з ринковими цінами, то вони в середньому становлять близько 20-40 грн за 1 кг макаронних виробів (дані можуть варіюватися і залежать від бренду, виду макаронних виробів та їх якості). Таким чином, можна сказати, що ціна завищена майже в 2 рази, що може свідчити про ознаки корупції. Також, якщо відсутні умови конкурентних торгів, і постачальник вже визначений за наперед, це теж дає істотні підозрювати корупцію. (процент корупції: 60, пояснення чому: завищення ціни майже вдвічі порівняно з ринковими показниками, можливість обрання "свого" постачальника без конкурсу)

По результату видно, чому тендер може бути корупційний, але так як немає повної інформації про тип макаронів для закупівлі, модель видає різні результати, бо не знає про який тип макаронів йдеться в тендері.

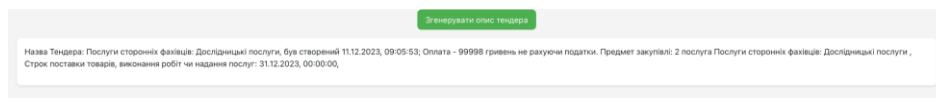


Рисунок 4.17 - Приклад 3 “Послуги сторонніх фахівців: Дослідницькі послуги”

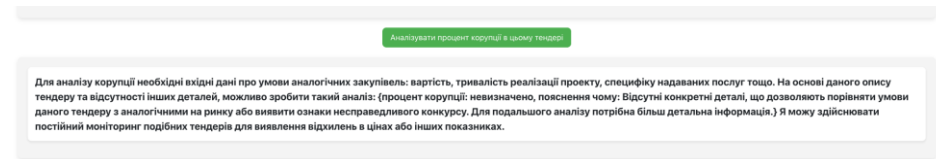


Рисунок 4.17 - Результат аналізу по тендеру “Послуги сторонніх фахівців: Дослідницькі послуги”

Видно, що було досягнуто не overfitting а саме перевивчення моделі, що призвело до того, що модель вказувала б на корупцію у всіх тендерах, що не являються елементами тренувальної вибірки.

Виберемо ще один будь який тендер для прикладу а саме “Хліб пшеничний, хліб житній, паски”

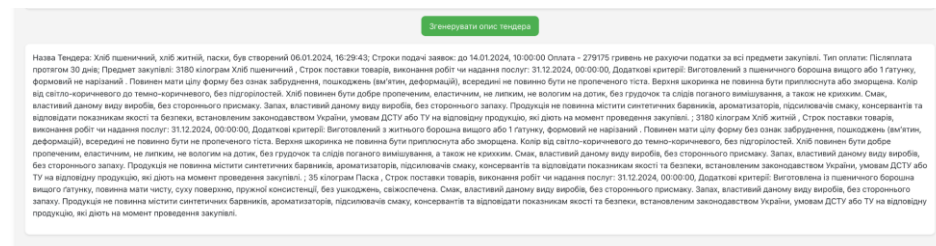


Рисунок 4.18 - Результат генерації даних про тендер

Аналізувати процент корупції в цьому тендері

Для аналізу корупційних ризиків у даних закупівлях, необхідно розглянути вартість кожного типу продукту окремо і порівняти її з ринковою вартістю. 1. Хліб пшеничний: 3180 кг, загальна вартість закупівлі не вказана. 2. Хліб житній: 3180 кг, загальна вартість закупівлі не вказана. 3. Паска: 35 кг, загальна вартість закупівлі не вказана. У відкритих джерелах в інтернеті вказують, що середня вартість 1 кг хліба в 2024 році коливається в межах 25-30 гривень. Якщо поррахувати загальну вартість хліба за цими показниками, ми отримаємо: 1. Хліб пшеничний: 3180 кг \* 25 грн (мінімальна ціна) = 79,500 грн 2. Хліб житній: 3180 кг \* 25 грн (мінімальна ціна) = 79,500 грн 3. Ціна за паску в роздріб може бути вище, але не критично, тому беремо її за 30 грн/кг: 35 кг \* 30 грн (приблизна ціна) = 1,050 грн У сумі отримуємо: 79,500 грн + 79,500 грн + 1,050 грн = 160,050 грн Загальна вартість закупівлі, вказана в тендері - 279,175 грн, що значно вище від нашого розрахунку. Таким чином, можливо, що ціни на хліб завищені на приблизно 70%. Однак, потрібно врахувати, що розрахунок був здійснений за мінімальною вартістю хліба на ринку, а також не були враховані можливі витрати на доставку, зберігання та інші пов'язані витрати. Тому, для точного визначення корупційних ризиків, необхідно більше інформації про постачальника, його минулу діяльність та деталі контракту. (процент корупції : 70, пояснення чому : Вартість закупівлі значно вище порівняно з посланнями на ціни на ринку.)

Рисунок 4.19 - Висновок моделі

Як бачимо, модель не спішить клеймувати “тендери” вона дає чіткий аналіз про отримані дані, також вказує які дані можна надати додатково. Що дає перспективи для подальшої розробки

### **Висновки до розділу**

У четвертому розділі розроблено та реалізовано систему інтелектуальної обробки даних про державні закупівлі на основі нейромереж. Детально описано архітектуру програми, методи розпізнавання та обробки даних, і представлено результати з їх візуалізацією. Система демонструє здатність ефективно аналізувати та виявляти потенційні ризики в державних закупівлях, пропонуючи важливий інструмент для підвищення прозорості та зниження корупції. Було показано результат роботи fine-tuned моделі. Також важливим моментом є те що результат показує різні дані на не тестових даних, це означає що модель не була “перенавчена”, що свідчить про правильне налаштування моделі

## **ВИСНОВКИ**

У магістерській дисертації розроблена аналітична система для виявлення корупційних ризиків у державних закупівлях за допомогою використання NestJS, React, OpenAI, Bootstrap, і NestJS. Проаналізовано багатошаровість і складність процесу державних закупівель, що включає не лише завищення цін, а й маніпулювання умовами тендерів для забезпечення вигоди обраним постачальникам. Здійснено вивчення українського сайту Prozorro, де виявлено типові шляхи зловживань та запропоновано методи їх виявлення через автоматизацію процесу контролю. Розроблена система використовує передові технології штучного інтелекту для аналізу та візуалізації даних, забезпечуючи прозорість та ефективність процесів аудиту. Програмний продукт протестовано та підтверджено його здатність ідентифікувати потенційно корупційні закупівлі. Це відкриває нові можливості для підвищення чесності та справедливості у сфері публічних фінансів та може слугувати ефективним інструментом у боротьбі з корупцією на національному рівні.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Інтелектуальна обробка даних. On-line: <https://kiis.knu.ua/intelektualnij-analiz-danih/>
2. Нейромережа. On-line: <https://termin.in.ua/neyromerezha/>
3. Кластеризація даних. On-line: [https://csc.knu.ua/media/study/asp/mod\\_probl\\_inf\\_tech\\_sys\\_analysis\\_ivohin/lecture/lec2.pdf](https://csc.knu.ua/media/study/asp/mod_probl_inf_tech_sys_analysis_ivohin/lecture/lec2.pdf)
4. Пошуковий алгоритм Google. On-line: <https://ahrefs.com/blog/google-search-algorithm/>
5. Штучний інтелект в медицині. On-line: <https://aiconference.com.ua/uk/news/primenenie-iskusstvennogo-intellekta-v-meditsine-effektivnaya-diagnostika-i-sozdanie-novih-lekarstv-92604>
6. NLP. On-line: <https://habr.com/ru/companies/Voximplant/articles/446738/>
7. Машинне навчання. On-line: <https://www.it.ua/knowledge-base/technology-innovation/machine-learning>
8. Алгоритми кластеризації. On-line: <https://ukrayinska.libretexts.org/>
9. Мовні патерни. On-line: <https://news.mit.edu/2022/ai-learn-patterns-language-0830>
10. BERT. On-line: [https://en.wikipedia.org/wiki/BERT\\_\(language\\_model\)](https://en.wikipedia.org/wiki/BERT_(language_model))
11. GPT. On-line: <https://aws.amazon.com/what-is/gpt/>
12. Зашумлені дані. On-line: <https://uk.wikipedia.org/wiki/%D0%9F%D0%B5%D1%80%D0%B5%D0%BD%D0%B0%D0%B2%D1%87%D0%B0%D0%BD%D0%BD%D1%8F>
13. Аналіз часових рядів на tablau. On-line: <https://www.tableau.com/solutions/time-series-analysis>
14. Аналіз Фур'є. On-line: [https://uk.wikipedia.org/wiki/%D0%90%D0%BD%D0%B0%D0%BB%D1%96%D0%B7\\_%D0%A4%D1%83%D1%80%27%D1%91](https://uk.wikipedia.org/wiki/%D0%90%D0%BD%D0%B0%D0%BB%D1%96%D0%B7_%D0%A4%D1%83%D1%80%27%D1%91)
15. Вейвлет аналіз. On-line: <https://habr.com/ru/articles/449646/>

- 16.БПФ. On-line: <https://habr.com/ru/companies/otus/articles/449996/>
- 17.Алгоритми класифікації даних. On-line:  
[https://ela.kpi.ua/bitstream/123456789/26128/1/Khanko\\_magistr.docx](https://ela.kpi.ua/bitstream/123456789/26128/1/Khanko_magistr.docx)
- 18.Манхетенська відстань. On-line:  
<https://uk.wikipedia.org/wiki/%D0%9C%D0%B0%D0%BD%D0%B3%D0%B5%D1%82%D1%82%D0%B5%D0%BD%D1%81%D1%8C%D0%BA%D0%B0%D0%BC%D0%B5%D1%82%D1%80%D0%B8%D0%BA%D0%B0>
- 19.DBSCAN. On-line: <https://ru.wikipedia.org/wiki/DBSCAN>
- 20.Блокчейн. On-line:  
<https://uk.wikipedia.org/wiki/%D0%91%D0%BB%D0%BE%D0%BA%D1%87%D0%B5%D0%B9%D0%BD>
- 21.React. On-line: <https://ru.legacy.reactjs.org/>
- 22.Nestjs - On-line: <https://docs.nestjs.com/>
- 23.Open AI API. On-line: <https://platform.openai.com/docs/api-reference>
- 24.How can AI transform government procurement ? On-line:  
<https://www.government-transformation.com/innovation/can-ai-transform-government-procurement>
- 25.What does artificial intelligence mean for government procurement? On-line:  
<https://www.instituteforgovernment.org.uk/event/artificial-intelligence-government-procurement>
- 26.How does AI works? On-line:  
<https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence>
- 27.Prozorro API. On-line: <https://prozorro-api-docs.readthedocs.io/uk/master/>
- 28.Open AI API. On-line: <https://platform.openai.com/docs/api-reference>
- 29.Deep learning на пальцях (урок 1). On-line:  
[https://www.youtube.com/watch?v=q46x0tq2FQ&t=3893s&ab\\_channel=sim0nsay](https://www.youtube.com/watch?v=q46x0tq2FQ&t=3893s&ab_channel=sim0nsay)

30. Deep learning на пальцах (урок 2) - елементи машинного навчання. On-line: [https://www.youtube.com/watch?v=1BUuB28FDOc&t=2567s&ab\\_channel=simOnsays](https://www.youtube.com/watch?v=1BUuB28FDOc&t=2567s&ab_channel=simOnsays)
31. Deep learning на пальцах (урок 3) - нейронні мережі. On-line: [https://www.youtube.com/watch?v=kWTC1NvL894&t=1883s&ab\\_channel=simOnsays](https://www.youtube.com/watch?v=kWTC1NvL894&t=1883s&ab_channel=simOnsays)
32. Фінальний Fine-tuning GPT Assistants з OpenAI. On-line: <https://dou.ua/forums/topic/45345/>
33. Machine Learning in Procurement. On-line: <https://planergy.com/blog/machine-learning-in-procurement/>
34. Корупційні ризики в публічних закупівлях під час воєнного стану: рекомендації НАЗК. On-line: <https://nazk.gov.ua/uk/novyny/koruptsijni-ryzyky-v-publichnyh-zakupivlyah-pid-chas-voyennogo-stanu-rekomendatsiyi-nazk/>
35. How to avoid overfitting? On-line: <https://aws.amazon.com/what-is/overfitting/>
36. Overfitting in Machine Learning: What It Is and How to Prevent It. On-line: <https://elitedatascience.com/overfitting-in-machine-learning>
37. Приклад корупційного тендера. On-line: <https://poltava.to/news/73459/>