

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ХАРЧОВИХ ТЕХНОЛОГІЙ**

**Інститут (факультет) Автоматизації і комп'ютерних систем  
Кафедра інформаційних технологій, штучного інтелекту і кібербезпеки**

**«До захисту в ЕК»**  
Директор інституту (декан факультету)  
\_\_\_\_\_ Андрій ФОРСЮК  
(підпис) (прізвище та ініціали)

«02» \_\_\_\_\_ червня \_\_\_\_\_ 2025 р.

**«До захисту допущено»**  
Завідувач кафедри  
\_\_\_\_\_ Сергій ГРИБКОВ  
(підпис) (прізвище та ініціали)

«02» \_\_\_\_\_ червня \_\_\_\_\_ 2025 р.

**КВАЛІФІКАЦІЙНА РОБОТА  
НА ЗДОБУТТЯ ОСВІТНЬОГО СТУПЕНЯ БАКАЛАВРА**

зі спеціальності 122 «Комп'ютерні науки»  
(код і назва спеціальності)

освітньо-професійної програми «Інформаційні системи та штучний інтелект»  
на тему: «Розроблення інтелектуальної системи розпізнавання бухгалтерських документів»

Виконав: здобувач 4 курсу, групи КН-4-4.

\_\_\_\_\_ Москалюк Назар Русланович \_\_\_\_\_  
(прізвище, ім'я та по батькові повністю) (підпис)

Керівник \_\_\_\_\_ Чаплінський Юрій Петрович \_\_\_\_\_  
(прізвище, ім'я та по батькові повністю) (підпис)

Консультанти \_\_\_\_\_ (ім'я та прізвище) \_\_\_\_\_ (підпис)

\_\_\_\_\_ (ім'я та прізвище) \_\_\_\_\_ (підпис)

\_\_\_\_\_ (ім'я та прізвище) \_\_\_\_\_ (підпис)

Рецензент \_\_\_\_\_ (ім'я та прізвище) \_\_\_\_\_ (підпис)

*Я як здобувач(ка) Національного університету харчових технологій розумію і підтримую політику університету з академічної доброчесності. Я не надавав(-ла) і не одержував(-ла) незарядженої допомоги під час підготовки цієї роботи. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело*

Здобувач \_\_\_\_\_  
(підпис)

Київ — 2025р.

# НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ХАРЧОВИХ ТЕХНОЛОГІЙ

Інститут (факультет) Автоматизації і комп'ютерних систем

Кафедра інформаційних технологій, штучного інтелекту і кібербезпеки

Освітній ступінь бакалавр

Спеціальність 122 комп'ютерні науки

(код і назва)

Освітньо-професійна програма Інформаційні системи та штучний інтелект

(назва)

## ЗАТВЕРДЖУЮ

Завідувач кафедри інформаційних

технологій, штучного інтелекту і

кібербезпеки \_\_\_\_\_ **Сергій ГРИБКОВ**

«28» \_\_\_\_\_ квітня \_\_\_\_\_ 2025 року

## ЗАВДАННЯ

### НА КВАЛІФІКАЦІЙНУ РОБОТУ ЗДОБУВАЧА

Москалюка Назара Руслановича

(прізвище, ім'я, по батькові)

1. Тема роботи «Розроблення інтелектуальної системи розпізнавання бухгалтерських документів»

керівник роботи доцент, к.т.н, Чаплінський Юрій Петрович

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом закладу вищої освіти від «28» квітня 2025 р. № 254-кв

2. Строк подання здобувачем роботи: 30.05.2025 р.

3. Вихідні дані до роботи: Загальна інформація та історія ПрАТ «Білоцерківський консервний завод», документи відділу кадрів та бухгалтерії підприємства, схема організаційної структури підприємства, інформація про програмне забезпечення підприємства

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити):

1. Загальна характеристика ПрАТ «Білоцерківський консервний завод»

2. Структура підприємства

3. Аналіз комп'ютеризації підприємства

4. Виявлення проблем

5. Огляд існуючих рішень

6. Створення технічного завдання на проектування інтелектуальної системи розпізнавання бухгалтерських документів

7. Опис функцій інтелектуальної системи

8. Реалізація функцій інтелектуальної системи

9. Тестування розробки

5. Перелік графічного матеріалу:

Схеми структур, бізнес-процесів

6. Консультанти розділів роботи:

| Розділ | Прізвище, ініціали<br>та посада консультанта | Підпис, дата      |                     |
|--------|--|-------------------|---------------------|
|        |  | завдання<br>видав | завдання<br>прийняв |
| 1      | доцент, к.т.н, Чаплінський Ю. П.             | 28.04.2025        | 23.05.2025          |
| 2      | доцент, к.т.н, Чаплінський Ю. П.             | 28.04.2025        | 25.05.2025          |
| 3      | доцент, к.т.н, Чаплінський Ю. П.             | 28.04.2025        | 27.05.2025          |

7. Дата видачі завдання: \_\_\_\_\_ 28 квітня 2025 року \_\_\_\_\_

### КАЛЕНДАРНИЙ ПЛАН

| № | Назва етапів виконання<br>кваліфікаційної роботи | Строк виконання<br>етапів роботи | Примітка |
|---|--|----------------------------------|----------|
| 1 | Аналіз предметної області                        | 28.04.2025                       |          |
| 2 | Постановка задачі                                | 30.04.2025                       |          |
| 3 | Створення бізнес-моделі процесів                 | 01.05.2025                       |          |
| 4 | Огляд рішень-аналогів                            | 02.05.2025                       |          |
| 5 | Створення технічного завдання                    | 03.05.2025                       |          |
| 6 | Створення інтелектуальної системи                | 09.05.2025                       |          |
| 7 | Створення інструкції користувача                 | 25.05.2025                       |          |
| 8 | Оформлення пояснювальної записки                 | 26.05.2025                       |          |
| 9 | Створення презентації                            | 30.05.2025                       |          |

Здобувач

\_\_\_\_\_

(підпис)

Москалюк Н.Р.

\_\_\_\_\_

(прізвище та ініціали)

Керівник роботи

\_\_\_\_\_

(підпис)

Чаплінський Ю.П.

\_\_\_\_\_

(прізвище та ініціали)

## АНОТАЦІЯ

Метою кваліфікаційної роботи є розробка інтелектуальної системи розпізнавання бухгалтерських документів для мінімізації затрат людського ресурсу на виконання рутинних задач, пов'язаних з документообігом.

При виконанні кваліфікаційної роботи було досліджено особливості документообігу бухгалтерії на підприємстві ПрАТ «Білоцерківський консервний завод» та проаналізовано стан комп'ютеризації на підприємстві; досліджено методи та підходи оптичного розпізнавання тексту, інтелектуальної класифікації бухгалтерських документів та існуючі рішення для автоматизації документообігу; сформовано вимоги щодо розробки інтелектуальної системи розпізнавання бухгалтерських документів; створено технічне завдання та на його основі розроблено інтелектуальну систему розпізнавання документів

Для реалізації інтелектуальної системи розпізнавання бухгалтерських документів використовувалась мова програмування Python та бібліотека оптичного розпізнавання символів EasyOCR. Для створення графічного інтерфейсу користувача було використано бібліотеку PyQt.

Проведено тестування функціоналу розробленої інтелектуальної системи розпізнавання бухгалтерських документів на реальних документах ПрАТ «Білоцерківський консервний завод» та інших шаблонах первинних бухгалтерських документів.

Обсяг пояснювальної записки: 49 сторінок, 7 таблиць, 20 ілюстрацій, 25 використаних джерел.

**Ключові слова:** ІНТЕЛЕКТУАЛЬНА СИСТЕМА, ОПТИЧНЕ РОЗПІЗНАВАННЯ СИМВОЛІВ, АВТОМАТИЗАЦІЯ, БУХГАЛТЕРІЯ, ДОКУМЕНТООБІГ, Python.

## ABSTRACT

The goal of the thesis is to develop an intelligent accounting document-recognition system in a pursuit of minimization of the human resource costs that are wasted while performing routine tasks that deal with document flow and management.

The thesis investigates the features of accounting document flow at the PrJSC “Bilotserkivs’kyy konservnyy zavod” and analyses the state of computerization at the enterprise; methods and implementations of optical character recognition, intelligent document classification and existing solutions for document flow automation were studied; requirements for the development of an intelligent document-recognition system were actualized; a product requirements document was created and an intelligent document recognition system was developed on its basis.

To implement the intelligent accounting document-recognition system, the programming language Python and its libraries EasyOCR (for optical character recognition) and PyQt (for graphical user interface) were used.

The functionality of the developed intelligent accounting document-recognition system was tested with real documents of PrJSC “Bilotserkivs’kyy konservnyy zavod” and other templates of primary accounting documents.

Explanatory note consists of: 49 pages, 7 tables, 20 illustrations, 25 used sources.

**Keywords:** INTELLIGENT SYSTEM, OPTICAL CHARACTER RECOGNITION, AUTOMATIZATION, ACCOUNTING, DOCUMENT FLOW, Python.

## ЗМІСТ

|   |    |
|---|----|
| ВСТУП.....  | 8  |
| РОЗДІЛ 1. ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ.....  | 10 |
| 1.1. Загальна характеристика ПрАТ «Білоцерківський консервний завод» .....  | 10 |
| 1.2. Організаційна структура ПрАТ «Білоцерківський консервний завод», роль і взаємодія підрозділів .....                        | 12 |
| 1.3. Аналіз нинішнього стану комп'ютеризації ПрАТ «Білоцерківський консервний завод» .....                                      | 14 |
| 1.4. Функціональне моделювання та аналіз існуючих бізнес-процесів .....   | 16 |
| 1.5. Огляд існуючих рішень для розв'язання виявлених проблем .....  | 19 |
| 1.6. Обґрунтування доцільності проектування й розроблення інтелектуальної системи розпізнавання бухгалтерських документів ..... | 22 |
| РОЗДІЛ 2. ТЕХНІЧНЕ ЗАВДАННЯ НА ПРОЄКТУВАННЯ.....  | 24 |
| 2.1. Загальні положення .....   | 24 |
| 2.2. Призначення і цілі створення системи .....   | 24 |
| 2.3. Характеристики об'єкта автоматизації .....   | 24 |
| 2.4. Вимоги до системи .....  | 24 |
| 2.5. Склад і зміст робіт по створенню системи .....   | 29 |
| 2.6. Порядок контролю і приймання системи.....  | 30 |
| 2.7. Вимоги до складу і змісту робіт із підготовки до введення системи в дію .....  | 31 |
| 2.8. Вимоги до документації.....  | 31 |
| 2.9. Джерела розробки .....   | 32 |
| РОЗДІЛ 3. ПРОЄКТУВАННЯ, СТВОРЕННЯ ТА АПРОБАЦІЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ БУХГАЛТЕРСЬКИХ ДОКУМЕНТІВ .....            | 33 |

|  |    |
|--|----|
| 3.1. Опис та обґрунтування вибору програмно-технічних засобів розроблення програмного продукту ..... | 33 |
| 3.2. Проектування та створення бази даних-довідника .....  | 35 |
| 3.3. Реалізація функцій системи .....  | 40 |
| 3.4. Інструкція користувача .....  | 43 |
| 3.5. Тестування програмного продукту .....   | 45 |
| ВИСНОВКИ .....   | 47 |
| СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....   | 48 |
| ДОДАТКИ .....  | 50 |

## ВСТУП

У сучасному світі постійно росте рівень комп'ютеризації і вона охоплює все більше сфер діяльності людини — від побуту, торгівлі й до великих промислових підприємств. Усвідомлюючи потенційну користь оцифрування або автоматизації окремих процесів функціонування підприємства та рівень конкуренції у цих питаннях, компанії ставлять за мету інтегрування відповідних цифрових технологій у власні середовища виробництва.

Харчова промисловість не є винятком. Виробники продукції, окрім впровадження типових інформаційних систем, спираються на концепцію інтернету речей, штучний інтелект, машинне навчання та хмарні сервіси для підвищення ефективності і зниження витрат у виробничих і адміністративних ланцюгах [1, 2].

Доречними з точки зору автоматизації бізнес-процесів об'єктами є відділи або підрозділи, що на тому чи іншому рівні вже знайомі з перевагами комп'ютеризації. Окрім того, що необхідність у відведенні окремого приміщення і закупки обладнання й пристроїв частково відпадає, час та витрати на ознайомлення працівників із новими технологіями також значно зменшуються. Одним із подібних відділів й, у свою чергу, відділ, на якому сфокусована кваліфікаційна робота — бухгалтерія підприємства ПрАТ «Білоцерківський консервний завод».

В бухгалтерії та відділі кадрів підприємства вже пройдено перші кроки цифрової трансформації: принтери із мережевим підключенням дозволяють генерувати та архівувати документи, а сканери пришвидшують переведення паперових екземплярів існуючих документів у цифровий простір. Працівники відділу вже знайомі з електронними реєстрами та шаблонами MS Excel, Word для підготовки звітів, а також мають доступ до існуючих внутрішніх інформаційних систем підприємства, де зберігаються нормативні документи й постанови контролюючих органів.

Завдяки такому рівню комп'ютеризації, облік руху коштів, формування податкових декларацій, формування звітів, тощо виконується швидше та з меншим рівнем ручної праці і у зручнішому форматі зберігання та передання інформації.

Тим не менш, введення даних із різномірних документів — рахунків-фактур, накладних і платіжних відомостей, тощо — у системи підприємства залишається процесом, затратним по часу через вимушеність ручного копіювання, перенесення або коригування даних з фізичних документів або їх сканувань.

Саме на цьому етапі дослідження підприємства з'явилася ідея розширення уже наявної інфраструктури підприємства інтелектуальним модулем для автоматичного розпізнавання змісту документів і валідації відповідних даних. Такий модуль дозволить не лише пришвидшити оцифрування інформації вхідних документів, а й інтегрувати отримані дані безпосередньо в базу даних підприємства, надаючи при цьому можливість користувачеві перевіряти й коригувати розпізнаний штучним інтелектом текст перед остаточним занесенням у базу [3, 4].

Подібне програмне рішення органічно доповнить існуючі цифрові інструменти підприємства та стане важливим кроком у підвищенні продуктивності його відділів. Відповідно, метою даної кваліфікаційної роботи є проєктування та розробка такої інтелектуальної системи для ПрАТ «Білоцерківський консервний завод».

## РОЗДІЛ 1. ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА

### ЗАДАЧІ

#### **1.1. Загальна характеристика ПрАТ «Білоцерківський консервний завод»**

У середині ХІХ століття Біла Церква стає осередком промислового розвитку. Одним із перших підприємств був пивоварний завод, який почав працювати у 1858 році і здобув популярність завдяки високій якості пива, що цінувалося не лише в Київській губернії. Проте, через фінансову нерентабельність, завод було закрито у 1863 році. Через 11 років виробництво було відновлено з випуском 86 500 відер пива за участю 30 осіб.

Після Жовтневої революції 1917 року завод зупинив своє виробництво пива, перетворившись на засолювальний пункт для огірків, помідорів та капусти. У 1920 році завод було реорганізовано в державне підприємство, і почався випуск солоду та оцту. У 1928 році, після зміни назви на «Укрглавкплодоовоч», завод отримує нову спеціалізацію.

Після закінчення Великої Вітчизняної війни завод було відновлено. Вже в 1949 році відкрився консервно-овочевий цех, де почали виробляти фруктові компоти та маринований перець, що стало поштовхом до переходу підприємства від сезонної до цілорічної роботи. У 1950 році було розпочато виробництво консервованих овочів, а в 1958 році — будівництво томатного цеху. З 1959 року комбінат працює цілорічно. У 1965 році завод впровадив обладнання для виробництва консервованого зеленого горошку, який здобув широку популярність, а в 1967 році — запроваджено лінію з виробництва соків.

У 1994 році, у контексті економічних реформ та приватизації, підприємство було реорганізовано в Акціонерне товариство відкритого типу «Білоцерківський консервний завод» на чолі з Крятюм Анатолієм Федоровичем.

Засновником Приватного акціонерного товариства «Білоцерківський консервний завод» є держава у особі Фонду комунального майна Білоцерківського району та організації орендарів підприємства «Білоцерківський консервний завод».

Відкрите акціонерне товариство «Білоцерківський консервний завод» зареєстроване 23 червня 1994 року Білоцерківської районною державною адміністрацією Київської області.

Згідно нової редакції Статуту затвердженого загальними зборами акціонерів (протокол № 1 від 29.04.2011 року), Відкрите акціонерне товариство «Білоцерківський консервний завод» перейменоване в Публічне акціонерне товариство «Білоцерківський консервний завод».

25.04.2017 загальними зборами акціонерів змінено тип акціонерного товариства з публічного акціонерного товариства на приватне акціонерне товариство.

Метою діяльності Товариства є здійснення підприємницької діяльності для одержання прибутку в інтересах акціонерів товариства, максимізації добробуту акціонерів у вигляді зростання ринкової вартості акцій Товариства.

Основні види діяльності:

- 10.39 — Інші види перероблення та консервування фруктів і овочів;
- 10.32 — Виробництво фруктових та овочевих соків.

Асортимент продукції, що випускає завод, разом з переліком відповідної нормативно-технічної документації на кожен продукт подано у Додатку А (Рис. А.1).

Нині виробнича потужність заводу складає 25 млн банок на рік. Продукція компанії з успіхом продається не тільки в Україні, але й експортується в наступні країни: Болгарія, Канада, Кіпр, Німеччина, Ізраїль, Латвія, Сполучені Штати, Греція, Вірменія, Прибалтика [5-7].

## **1.2. Організаційна структура ПрАТ «Білоцерківський консервний завод», роль і взаємодія підрозділів**

### **1.2.1. Загальна схема організаційної структури**

Організаційну структуру заводу у повному обсязі подано у Додатку А (Рис. А.2).

Голова правління установи безпосередньо керує підприємством і несе відповідальність за реалізацію затвердженої стратегії. Координує роботу заступників та підпорядкованих відділів, приймає ключові оперативні рішення.

Заступник голови правління з комерційних питань відповідає за комерційний блок підприємства, що включає в себе:

- відділ збуту – реалізація готової продукції на внутрішніх та зовнішніх ринках, формування збутової політики;
- відділ постачання – закупівля сировини, пакувальних матеріалів та інших ресурсів, необхідних для виробничих процесів.

Заступник голови правління по виробництву опікується виробничими процесами від підготовки сировини до випуску готової продукції. До його підпорядкування входять:

- виробничі цехи (зокрема консервний, томатний, фабрикатний) – безпосередня організація виробничих процесів;
- технологічна служба та лабораторія – контроль якості сировини та готової продукції, розробка й впровадження нових рецептур та технологій;
- механічна (технічна) служба – обслуговування, ремонт і модернізація обладнання для забезпечення безперервної роботи цехів.

Головний бухгалтер несе відповідальність за роботу бухгалтерії; приймає участь у керівництві підприємством щодо укладання договорів, контрактів.

Заступник головного бухгалтера здійснює зведений облік, оформляє та подає звітність.

Відділ кадрів відповідає за відбір, навчання та розвиток працівників, формування кадрової політики, ведення документації з питань праці та заробітної плати.

Інженер з охорони праці стежить за дотриманням техніки безпеки, проводить інструктажі для персоналу, контролює виконання норм і правил охорони праці, а також взаємодіє з відповідними державними органами.

### **1.2.2. Структура бухгалтерії ПрАТ «Білоцерківський консервний завод»**

Організаційна структура бухгалтерії консервного заводу будується за принципом функціонального розподілу облікових задач із урахуванням обсягів і характеру діяльності підприємства.

Застосовується схема з головним бухгалтером на чолі, заступником головного бухгалтера та декількома групами (облік реалізації, облік поставок, облік основних засобів, розрахунки з персоналом, звітність). Така ієрархія дозволяє централізовано контролювати ведення обліку і оперативно координувати взаємодію між внутрішніми обліковими та виробничими процесами.

Особливістю роботи бухгалтерії харчового виробництва є необхідність узгодження даних обліку із технологічними показниками заводу (обсяги випуску продукції, собівартість партій, списання сировини), оскільки саме через бухгалтерський облік відбувається фінансове оцінювання виробничих процесів. У буденному режимі бухгалтерія забезпечує прийом первинних документів, їх опрацювання, проведення господарських операцій в інформаційній системі та формування внутрішньої (управлінської) і зовнішньої фінансової звітності.

Бухгалтерія підприємства тісно співпрацює з виробничими та адміністративними службами, забезпечуючи інформаційні потоки в обидва боки (Таблиця 1.1).

*Таблиця 1.1 Потоки інформації між бухгалтерією та іншими структурними підрозділами консервного заводу*

| <b>Підрозділ</b>                | <b>Вхідна інформація до бухгалтерії</b>                                 | <b>Вихідна інформація з бухгалтерії</b>  |
|---------------------------------|---|--|
| Відділ постачань                | Рахунки-фактури, товарні накладні, договори постачання                  | Заяви на оплату, аналітика платежів, звіти по рахункам постачальників            |
| Виробничий відділ               | Накладні на сировину, журнали виробничих змін, обсяги випуску продукції | Звіт про собівартість продукції, калькуляції, довідки для планування             |
| Відділ збуту                    | Замовлення клієнтів, рахунки до сплати, відвантажувальні документи      | Податкові накладні, акти виконаних робіт, аналітика дебіторської заборгованості  |
| Відділ кадрів                   | Табелі обліку робочого часу, штатні розписи                             | Розрахунки зарплат, відомості по податках та соціальних внесках                  |
| Фінансовий відділ (керівництво) | Бюджетні заявки, планові показники, інформація по кредитних лініях      | Фінансова звітність, платіжні документи, аналітичні довідки для прийняття рішень |
| Юридичний відділ                | Договори та додатки, листування з контрагентами                         | Звіти про фінансові ризики, висновки по платіжних зобов'язаннях                  |

### **1.3. Аналіз нинішнього стану комп'ютеризації ПрАТ «Білоцерківський консервний завод»**

Ведення бухгалтерського та податкового обліку на підприємстві здійснюється за допомогою програмного комплексу М.Е.Дос, який включає базові модулі «Звітність» для формування та подання декларацій у контролюючі органи й «ЕДО» для обміну електронними документами з контрагентами. Додатково у

М.Е.Дос можуть використовуватися модулі «Зарплата» та «Акциз і ТТН», проте на заводі задіяні лише базові функції подання звітності та реєстрації податкових накладних [8].

Паралельно використовуються інструменти MS Excel та Word для підготовки внутрішніх табличних реєстрів і довідкових матеріалів. Все ж, різні потенційні інформаційні системи (управління виробництвом, логістика, контроль якості тощо) не впроваджено або впроваджено не в повному обсязі, що призводить до того, що на паперову та ручну обробку документів все одно витрачається значна частка людського ресурсу, що в свою чергу збільшує ризик розбіжностей між даними, ускладнює оперативне прийняття рішень та аналіз виробничих або бізнес процесів.

Зберігання та облік значної частини документів бухгалтерії і суміжних підрозділів ведеться через фізичне отримування, архівування і дублювання у відповідних журналах інформації. Процеси сканування та зберігання паперових первинних документів (рахунків-фактур, накладних, актів тощо) виконуються за допомогою багатофункціональних пристроїв, але інформація з отриманих документів далі не обробляється автоматично. Відсутність спеціалізованого рішення для вилучення й перевірки даних з первинних документів змушує операторів вручну вводити та коригувати інформацію у системі і базі даних підприємства.

Нинішній стан комп'ютеризації та автоматизації документообігу підприємства, зокрема у бухгалтерії, є незадовільним у сучасних умовах і потребує поступової модернізації і впровадження ряду комп'ютерних систем, адже покриває лише базові потреби в обліку й звітності, залишаючи поза увагою автоматизоване вилучення даних із різномірних первинних документів, контроль їх валідації та інтеграцію отриманих даних. Цей розрив призводить до дублювання ручної праці, затримок у обробці документів та підвищеного ризику помилок, що в свою чергу негативно сприяє росту прибутку компанії [9].

## 1.4. Функціональне моделювання та аналіз існуючих бізнес-процесів

### 1.4.1. Функціональна модель бухгалтерії ПрАТ «Білоцерківський консервний завод»

Діаграму поточної моделі бізнес-процесів, протікаючих між бухгалтерією та відділом постачання консервного заводу наведено у додатку Б (Рис. Б.1).

Загальна діаграма IDEF0 верхнього рівня, створена засобами AllFusion Process Modeler представлена на рисунку 1.1.

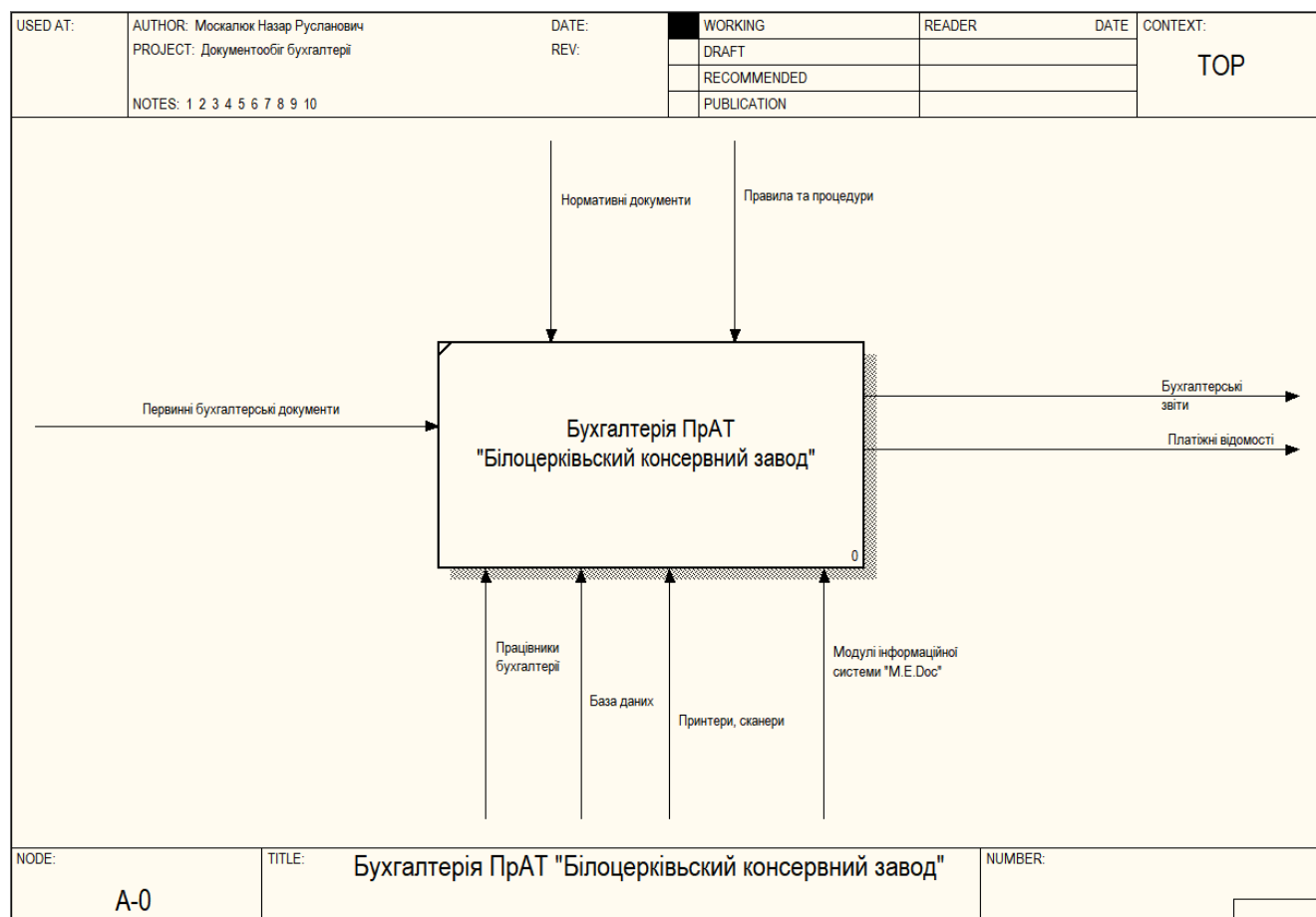


Рисунок 1.1 — Діаграма IDEF0 верхнього рівня

Опис функціональної моделі документообігу у бухгалтерії підприємства:

- вхідні дані: документи у вигляді рахунків-фактур, товарно-транспортних накладних, актів, звітів, тощо, які поступають від відділу постачання, виробничого відділу, відділу збуту, контролюючих органів, тощо, або формуються безпосередньо у бухгалтерії;
- обробка даних: прийом, друк та збереження первинних документів; перевірка та валідація їх змісту; сканування документів для цифрового

архіву; підготовка до імпорту перевірених документів у спеціально віднесені для них цифрові сховища (база даних, таблиці MS Excel); дублювання записів у паперові журнали;

- вихідні дані: звіти для керівництва; верифіковані та/або скориговані копії первинних документів; оновлені записи в електронних журналах обліку; відомості для архівів.

Часткова функціональна декомпозиція бухгалтерії проілюстрована на рисунку 1.2.

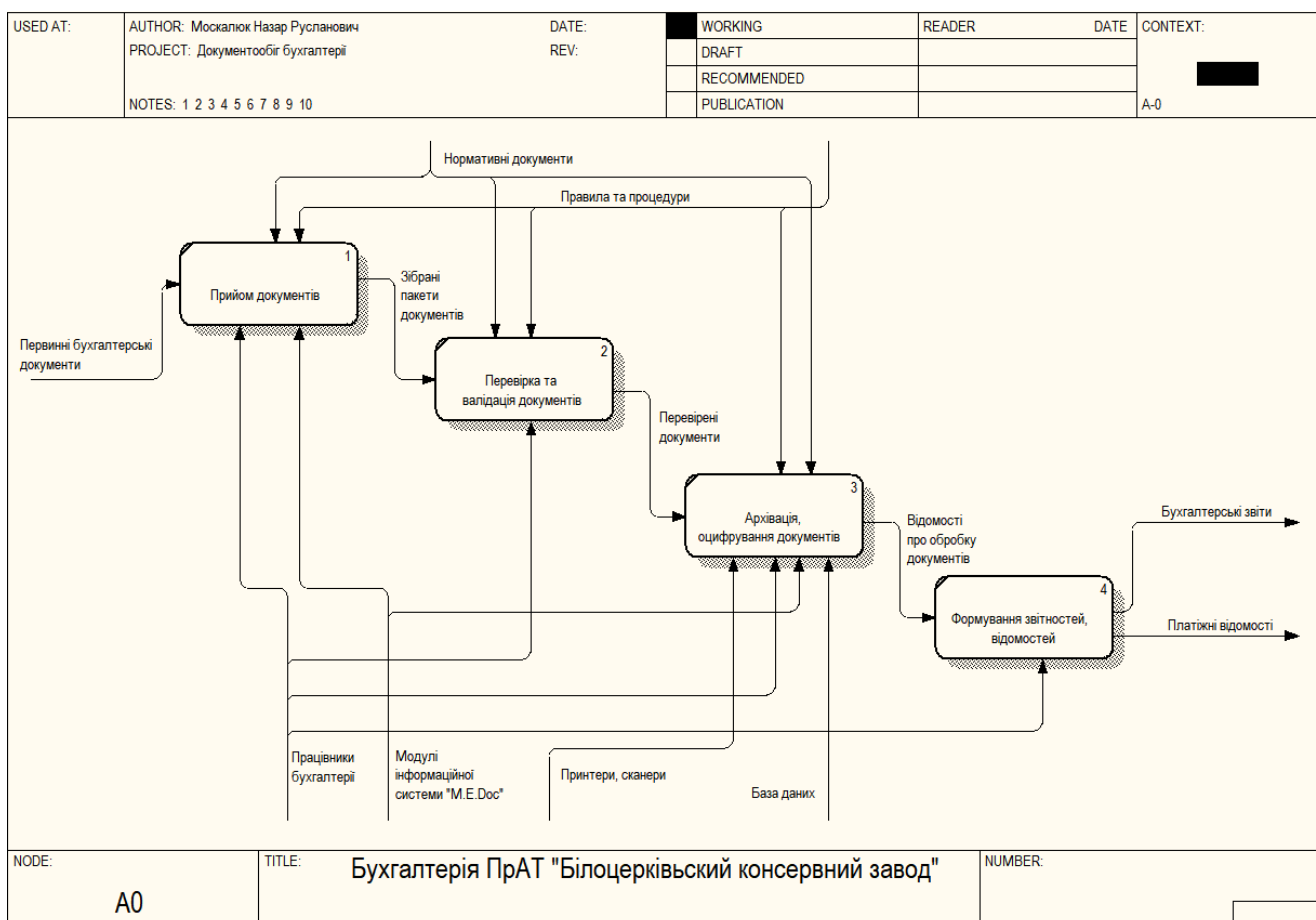


Рисунок 1.2 — Діаграма IDEF0 нижнього рівня

### 1.4.2. Виявлені проблеми та пропозиції щодо усунення

Низький рівень комп'ютеризація та використання паперових носіїв у процесах документообігу бухгалтерії супроводжуються низкою проблем.

Насамперед, паперовий документообіг є повільним і неефективним. Більшість життєвого циклу паперового документа витрачається на його передачу між співробітниками і менше часу проводиться за роботою з цим документом. Це виливається у затримки прийняття рішень та зниження загальної продуктивності підприємства.

Утримання архівів у паперовому форматі потребує значних витрат; на папір, на чорнила для друку, на простір для зберігання, на транспортування та утилізацію. Паперові документи легко пошкодити, втратити або знищити у разі надзвичайних ситуацій або просто необережності працівників. У разі відсутності резервних копій втрачається можливість відновлення.

У великих архівах паперових документів складно швидко знаходити необхідну інформацію, що ускладнює своєчасне прийняття рішень та може привести до дублювання роботи працівників.

Аналіз роботи бухгалтерії підприємства показав наступні проблеми:

- дублювання операцій – дані спочатку вводяться в Excel шаблони, а потім копіюються в інформаційну систему М.Е.Дос, що збільшує навантаження на персонал, підвищує ймовірність помилок та демонструє відсутність централізації документообігу;
- неможливість автоматичної валідації розпізнаних даних у відповідності з нормативами, що змушує працівників перевіряти кожен запис вручну;
- процес коригування й повторної перевірки документів займає значний час через відсутність звітності виявлених помилок та історії змін;
- не масштабованість існуючої архітектури систем, через що зростання обсягів документів призводить неприйнятних витрат часу і ресурсів на їх ручну обробку.

Пропозиції щодо усунення наявних проблем деяких аспектів бухгалтерії у форматі AS IS -> TO BE наведено у таблиці 1.2.

Таблиця 1.2 Нинішній стан документообігу в бухгалтерії проти запропонованих рішень

| Аспект                     | AS IS  | TO BE  |
|----------------------------|--|--|
| Введення даних             | Ручний запис<br>в/копіювання даних з<br>Excel                                      | Автоматичний експорт<br>даних у БД через<br>інтелектуальну систему |
| Валідація                  | Ручна перевірка<br>працівником   | Автоматична перевірка<br>словниками, нормативами<br>тощо           |
| Сховище документів         | Записи у паперових<br>журналах, скани<br>документів, форми в<br>Excel, записи у БД | Централізована база даних  |
| Затрати часу<br>працівника | Високі – рутинне<br>дублювання   | Низькі – переважно<br>контроль і корекція                          |
| Інтеграція                 | Мінімальна,<br>локальне/ручне  | Пряма, через з'єднання<br>інтелектуальної системи з<br>БД          |

Таким чином було вирішено розробити інтелектуальну систему розпізнавання бухгалтерських документів для ПрАТ «Білоцерківський консервний завод», адже інтеграція такого модулю посприє акселерації та якості виконання окремих процесів бухгалтерії.

## 1.5. Огляд існуючих рішень для розв'язання виявлених проблем

1.5.1. Система **Іннора RaccoonDoc** — сервіс від Української компанії Іннора, який забезпечує високоточне вилучення даних з комплексних форм бухгалтерських документів (рахунки-фактури, акти видаткові накладі, митні декларації, транспортні документи та ін.), а також пропонує вбудовані засоби машинного навчання для донавчання під специфіку підприємства. Документи для обробки імпортуються шляхом використання веб-додатку, запитів по API (прикладний

програмний інтерфейс) або надсилання файлів на відведену пошту. В функціонал веб-додатку також входить перегляд та коригування розпізнаних даних перед їх експортом [10].

1.5.2. Компанія «DMS Solutions» є українським розробником програмного забезпечення у сфері електронного документообігу. **eIDoc IDP** — інтегрована хмарна інтелектуальна платформа для автоматизації процесів документообігу та обробки документів. Платформа включає в себе модуль оптичного розпізнавання символів, машинне навчання для класифікації та автоматичні робочі процеси, які дозволяють повністю автоматизувати обробку вхідних документів підприємства. Інтеграція з існуючим програмним забезпеченням здійснюється через веб-додаток та готові API для роботи з популярними в Україні системами обліку [11].

1.5.3. **BitFaktura** — Сервіс для роботи з рахунками-фактурами, що у свою чергу підтримує модуль оптичного розпізнавання символів. Надають програмне забезпечення для бухгалтерських офісів [12].

1.5.4. **CleverForms** — це програмна платформа, що розроблена в Україні, з великою кількістю сервісів для автоматизації та побудови ефективних бізнес-процесів на підприємстві, що допомагає заощадити час і підвищити ефективність роботи, автоматизуючи цілком або частково різноманітні робочі процеси, що пов'язані з електронним документообігом. Її можна використовувати як в локальній мережі підприємства, так і через інтернет з будь-якого пристрою зі встановленим браузером, не встановлюючи ніякого додаткового програмного забезпечення.

OCR CleverForms — це інтегрований модуль, що базується на технології оптичного розпізнавання символів. Він автоматизує процес розпізнавання тексту на сканованих документах, миттєво перетворюючи його у зручний для редагування текстовий формат, зберігаючи структуру й зміст оригіналу [13].

## 1.5.5. Порівняння систем-аналогів та висновки.

Таблиця 1.3 Порівняння представлених рішень-систем

| <b>Характеристика</b>   | <b>Іннора<br/>RaccoonDoc</b>             | <b>elDoc IDP</b>                        | <b>BitFaktura</b>                 | <b>CleverForms</b>                                |
|---|--|---|-----------------------------------|---|
| Вартість впровадження або підписки                              | Не опублікована                          | 3000,00 ₴ на рік                        | 3600,00 ₴ на рік                  | 576 000,00 ₴                                      |
| Підтримка усіх необхідних форм документів, інтеграція у M.E.Doc | Так, Так                                 | Так, Ні                                 | Ні, Так                           | Так, Так  |
| Кількість користувачів одночасно                                | Необмежено                               | 3                                       | 10                                | Необмежено  |
| Користування  | Веб-додаток, API, через електронну пошту | API, ERP-конектори                      | Веб-додаток, CSV/XML експорт, API | За бажанням під час та після впровадження системи |
| Підтримка розпізнавання мов документів                          | Українська, англійська та понад 50 ін.   | Українська, англійська та понад 100 ін. | Українська, англійська            | Українська, англійська та ін.                     |

Як правило, модуль оптичного розпізнавання символів є лише частиною більшої системи, направленої на автоматизацію документообігу загалом. Окрім функції зчитати та експортувати текст з документу, впровадження одного з таких рішень у відділ бухгалтерії принесе з собою ряд функцій, необхідності в яких немає, використання їх неможливе у контексті підприємства або витрати на навчання персоналу кардинально змінити свій типовий алгоритм роботи не виправдають себе.

Окрім того, більшість рішень працюють у хмарі, по підписці, що потребує постійного підключення до мережі інтернет, та не гарантує безперебійності системи. А рішення, що надає змогу розгорнути усю систему індивідуально для підприємства є надзвичай дорогим, хоч і якісним. Необхідності у таких просунутих багатофункціональних систем у консервного заводу немає.

Виділити як перспективне варто рішення RaccoonDoc від Іннора. Сервіс направлений виключно на роботу з розпізнаванням та експортуванням даних з документів, має зручний інтерфейс, нативну підтримку з'єднання з інформаційною системою М.Е.Doc та можливість використання через API.

### **1.6. Обґрунтування доцільності проєктування й розроблення інтелектуальної системи розпізнавання бухгалтерських документів**

У результаті проведеного аналізу бізнес-процесів бухгалтерії і існуючих рішень інтелектуальної роботи з документами виявлено, що поточна автоматизація бізнес-процесів бухгалтерії ПрАТ «Білоцерківський консервний завод» не знаходиться на достатньо високому рівні, щоб її стан можна було ігнорувати.

Оптимізація завдань з вилучення, перевірки та інтеграції даних із первинних призведе до економії часу працівників бухгалтерії, що дозволить витратити його на продуктивніші задачі підрозділу [14]. Ринкові інтелектуальні системи демонструють значні переваги в підвищенні продуктивності, точності даних, потенційному зниженні операційних витрат і ризиків невідповідності нормативним вимогам.

Проте, їхнє впровадження у пропонованому вигляді й з урахуванням нинішнього стану комп'ютеризації консервного заводу або високо вартісне, або є складним у інтегруванні з усіма процесами відділу і не виправдовує витрат.

Пропонована розробка інтелектуальної системи націлена на вирішення однієї з наявних проблем автоматизації бухгалтерії та, в силу індивідуальності проєкту, врахує недоліки існуючих рішень у контексті функціонування відділу бухгалтерії саме ПрАТ «Білоцерківський консервний завод» та його документів.

Розробка рішення на базі безкоштовних бібліотек з відкритим кодом Python дозволить дешево та швидко розробити програмний продукт, який можна буде розгорнути для використання підприємством на локальних серверах. Пряме з'єднання зі сховищами даних підприємства, модульність коду розробки та вузький напрям підприємства сприятимуть простоті підтримки та масштабуванню програмного продукту.

Хоч кінцевий продукт кваліфікаційної роботи першочергово орієнтований на експлуатацію бухгалтерією об'єкта дослідження, модульність такої розробки і її гнучкість дозволить розширення функціоналу і галузей застосування шляхом доповнення змісту відповідної БД-довідника.

Все це і визначило наступні основні завдання кваліфікаційної роботи:

- дослідити особливості документообігу бухгалтерії ПрАТ «Білоцерківський консервний завод» та проаналізувати стан комп'ютеризації на підприємстві;
- дослідити сучасні методи оптичного розпізнавання тексту, підходи до інтелектуальної класифікації бухгалтерських документів та існуючі рішення для автоматизації документообігу на вітчизняному ринку;
- визначити проблеми та причини неефективності ручного введення даних у процесах документообігу;
- сформувати вимоги щодо розробки інтелектуальної системи розпізнавання бухгалтерських документів;
- створити технічне завдання на розробку інтелектуальної системи розпізнавання бухгалтерських документів;
- розробити інтелектуальну систему розпізнавання бухгалтерських документів;
- провести тестування функціоналу розробленої інтелектуальної системи розпізнавання бухгалтерських документів.

## **РОЗДІЛ 2. ТЕХНІЧНЕ ЗАВДАННЯ НА ПРОЄКТУВАННЯ**

### **2.1. Загальні положення**

2.1.1. Найменування системи: «Інтелектуальна система розпізнавання бухгалтерських документів».

2.1.2. Результати кожного етапу розроблення оформлюються згідно з вимогами ДСТУ. Порядок передачі результатів спирається на затверджений календарний план виконання розробки.

2.1.3. У випадку необхідності на наступних стадіях робіт по створенню системи окремі положення можуть уточнюватися і розвиватися.

### **2.2. Призначення і цілі створення системи**

2.2.1. Система призначена для автоматизації та централізації окремих процесів документообігу бухгалтерських документів на підприємстві. А саме процесів зчитування, оцифрування, валідації й занесення даних із первинних бухгалтерських документів у базу даних (БД) підприємства.

2.2.2. Основною метою створення інтелектуальної системи розпізнавання бухгалтерських документів є інформаційне та програмне забезпечення для мінімізації затрат часу на оцифрування інформації з бухгалтерських документів та їх імпорт в базу даних шляхом зменшення об'єму обробки документів ручними методами.

### **2.3. Характеристики об'єкта автоматизації**

2.3.1. Об'єктом автоматизації є документообіг бухгалтерії ПрАТ «Білоцерківський консервний завод».

### **2.4. Вимоги до системи**

2.4.1. Вимоги до системи в цілому.

2.4.1.1. Вимоги до структури і функціонування системи:

- система повинна мати клієнт-серверну архітектуру та передбачати під'єднання та виконання запитів до БД підприємства;
- система повинна бути пов'язана в мережі з персональними комп'ютерами (ПК), розміщеними в бухгалтерії;

- система має використовувати відкриту для доповнення БД-довідник для класифікації бухгалтерських документів та визначення ключових полів документів;
- діагностування функціонування системи має передбачати виявлення відхилень від нормального процесу експорту зчитаних даних до БД і порушень у роботі комп'ютерно-технічних засобів, а також програмних помилок, забезпечуючи користувачів відповідними діагностичними повідомленнями;
- розвиток і модернізація системи повинні проводитися шляхом уточнення чи нарощення виконуваних функцій, модернізації технічних і програмних засобів по мірі розроблення і впровадження нових поколінь комп'ютерів;
- структура і технологія програмного забезпечення системи повинні забезпечити простоту їх модернізації та розвитку, з можливістю збільшення розмірності задач і масивів інформації, а також можливості реалізації їх на нових ПК;
- модуль оптичного розпізнавання символів (ОРС) повинен утилізувати графічну карту за її наявності для забезпечення максимальної швидкодії.

#### 2.4.1.2. Вимоги до кваліфікації персоналу:

- повинен пройти навчання і отримати навички роботи на ПК;
- повинен дотримуватись технологічних інструкцій при роботі з системою;
- повинен дотримуватись умов експлуатації ПК у відповідності з інструкціями по експлуатації;
- повинен дотримуватись правил зберігання інформації і організації резервних копій БД-довідника і оцифрованих даних бухгалтерських документів;
- повинен дотримуватись правил техніки безпеки при роботі на ПК.

#### 2.4.1.3. Вимоги до надійності системи:

- система є багатофункціональною і призначена для використання протягом робочого дня;

- всі функції системи виконуються дискретно й у відповідності з ДСТУ 2226-93 оцінка надійності проводиться по кожній функції окремо;
- передбачене використання модульного, структурного і об'єктно-орієнтованого програмування;
- інформація про вхідні файли повинна видавати користувачеві повідомлення про виявлені помилки;
- система повинна бути захищена від збоїв, несанкціонованого доступу, помилкових дій персоналу тощо.

#### 2.4.1.4. Вимоги до безпеки:

- для забезпечення безпеки при експлуатації, налагодженні, монтажі, обслуговуванні і ремонті технічних засобів системи потрібно дотримуватись вимог ДСТУ: ДСТУ 2293-99, ДСТУ ISO 6309:2007, ДСТУ 12.0.230:2008, ДСТУ 7237:2011, ДСТУ 7238:2011, ДСТУ 7239:2011;
- по доступним рівням освітленості, вібраційних і шумових навантажень слід дотримуватися вимог відповідно ДСТУ Б А.3.2-15:2011, ДСТУ EN 14253:2018, ДСТУ 2867-94.

#### 2.4.1.5. Вимоги з ергономіки та технічної естетики.

Загальні ергономічні і естетичні вимоги до системи повинні відповідати державним стандартам ДСТУ 8604:2015, ДСТУ 7298:2013. Освітленість робочого місця повинна відповідати ДСТУ EN 12464-1:2016, ДБН В.2.5-28-2006.

Засоби відображення повинні розміщуватися таким чином, щоб кут спостереження екрану складав не більше, ніж 45 градусів, мінімальна відстань спостереження екрану — 0,3 м, рекомендована — 0,5 м.

При розробленні програмного забезпечення слід створити зручний інтерфейс для запобігання втомлюваності користувача.

#### 2.4.1.6. Вимоги по експлуатації, технічного обслуговування, ремонту і зберігання компонентів системи:

- види обслуговування системи визначаються у відповідності з ДСТУ EN 13306:2019. Загальні вимоги по експлуатації, технічному обслуговуванню і ремонту повинні відповідати ДСТУ 3576-97;

- для розміщення технічних засобів системи необхідні площі, визначені в ДБН В.2.2-9-2009. При цьому слід дотримуватися вимог, зазначених в експлуатаційній документації. Напряга живлення технічних засобів системи 220/380 В змінного струму, частотою  $(50 \pm 1)$  Гц. Допустиме відхилення напруги від +10 до -15%, тривалість перерв у живленні не повинна перевищувати 0,001 с;
- склад, розміщення і умови зберігання компонентів технічних засобів системи визначається рекомендаціями, зазначеними в експлуатаційній документації на ці елементи;
- регламент обслуговування повинен відповідати їхньому рівню і умовам роботи, щоб у випадку відмови системи забезпечити роботу в аварійному режимі.

#### 2.4.2. Вимоги до функцій.

2.4.2.1. Перелік функцій із зазначенням вхідної та вихідної інформації наведено в таблиці 2.1.

Функції мають забезпечити раціональну організацію роботи користувача на основі безперервної технології: класифікація документа, розпізнавання тексту документа, перевірка правопису і виконання інших функцій, визначених чинним документом. При цьому пріоритетом є зручність передачі у систему зображень документів та експорту розпізнаних даних у базу даних за рахунок підказок і меню на екрані монітора.

Таблиця 2.1. Перелік функцій вхідної та вихідної інформації

| № п/п | Найменування функції                                   | Вхідна інформація   | Вихідна інформація                                    |
|-------|--|---|---|
| 1     | Завантаження документів                                | Файли сканів документів (по одному або пакетно)                               | Готові до обробки файли                               |
| 2     | Перегляд зображення та результатів попередньої обробки | Зображення із черги на обробку  | Результат застосування фільтрів попередньої обробки   |
| 3     | Класифікація документа                                 | Оброблене зображення документа; ключові слова з БД-довідника                  | Найвірогідніший тип документа                         |
| 4     | Зчитування даних з таблиці документа                   | Оброблене зображення документа та його тип                                    | Текст з клітинок таблиці                              |
| 5     | Зчитування даних з полів документа                     | Оброблене зображення документа та його тип; допоміжні значення з БД-довідника | Текст з ключових полів документа                      |
| 6     | Перевірка правопису слів з помилками                   | Зчитаний текст документа  | Пропозиції по виправленню слів                        |
| 7     | Ручне виправлення даних                                | Зчитаний текст документа  | Виправлений текст документа                           |
| 8     | Підтвердження успішності оцифрування даних             | Виправлений текст документа   | Структурування значень для імпорту в БД               |
| 9     | Експорт результатів до БД підприємства                 | Структуровані дані документа  | Сформовані та виконанні SQL-запити; внесені у БД дані |

### 2.4.3. Вимоги до видів забезпечення.

#### 2.4.3.1. Вимоги до лінгвістичного забезпечення (ЛЗ):

- використання мов високого рівня, які забезпечують створення структурних програм;
- використання фреймворків з відкритим кодом;
- використання мови СУБД підприємства для здійснення доступу та імпорту даних;
- організація діалогу користувача до системи має будуватися на наборах меню і підказок, орієнтованих на виконання користувачем функцій.

#### 2.4.3.2. Вимоги до програмного забезпечення (ПЗ):

- операційна система (ОС) сервера — Linux;

- ОС користувача — Windows 10;
- Python версії 3.9.x;
- програмний засіб для управління БД.
- повне задоволення потреб функціональних завдань системи;
- виведення необхідних даних на екран у вигляді відповідних відеограм;
- керований комп'ютером діалог при обробці документа;

2.4.3.3. При розробленні системи слід виконати наступні вимоги:

- використані програми мають бути сумісні між собою та із загальносистемним ПЗ;
- ПЗ має розроблятися засобами об'єктно-орієнтованого програмування;
- забезпечити відповідність інтерфейсу користувача стандартам Windows;
- структура програм представлена модулями;
- передбачена можливість розширення складу задач у відповідності з новими функціональними потребами (нові типи документів);
- діалог із користувачем повинен проводитись за допомогою клавіатури або миші з поясненням виконання дій і можливістю отримання підказки.

2.4.3.4. Технічні засоби системи (таблиця 2.2.) повинні забезпечувати виконання функцій, перерахованих в таблиці 2.2.

*Таблиця 2.2. Вимоги до технічного забезпечення системи*

| № п/п  | Основні характеристики комп'ютера   |
|--|---|
| <b>Технічне забезпечення для сервера</b>     |   |
| 1  | (CPU)Intel Xeon Silver\ (GPU)NVIDIA Tesla T4\ (RAM)32 GB\ (Накопичувачі)NVMe SSD 512GB+HDD 1TB\ (Network)LAN 1 Gbit |
| <b>Технічне забезпечення для користувача</b> |   |
| 1  | (CPU)Intel Core i5/AMD Ryzen 5, 3.0 GHz\ (RAM)8 GB\ (Накопичувач)SSD 256GB  |
| 2  | Монітор >13"  |
| 3  | Миша  |
| 4  | Клавіатура  |
| 5  | Сканер  |

## 2.5. Склад і зміст робіт по створенню системи

2.5.1. Стадії створення системи і терміни виконання робіт наведені в таблиці

2.3.

Таблиця 2.3. Найменування робіт при створенні системи

| № п/п | Найменування робіт                      | Строки виконання робіт |
|-------|---|------------------------|
| 1.1   | Дослідження документообігу підприємства | 30.04.2025             |
| 1.2   | Документування процесів                 | 02.05.2025             |
| 1.3   | Аналіз та звітування                    | 03.05.2025             |
| 2.1   | Створення прототипу технічного завдання | 05.05.2025             |
| 2.2   | Перегляд ТЗ та внесення правок          | 07.05.2025             |
| 2.3   | Затвердження ТЗ                         | 08.05.2025             |
| 3.1   | Дизайн архітектури системи              | 12.05.2025             |
| 3.2   | Специфікація модулів                    | 13.05.2025             |
| 3.3   | Моделювання БД-довідника                | 14.05.2025             |
| 3.4   | Програмування інтерфейсу користувача    | 17.05.2025             |
| 4.1   | Тестування розробки                     | 23.05.2025             |
| 5.1   | Створення користувацького мануала       | 30.05.2025             |
| 5.2   | Створення технічної документації        | 30.05.2025             |

2.5.2. Діаграму Ганта із відповідними підзадачами стадій створення системи зображено на рисунку 2.1.

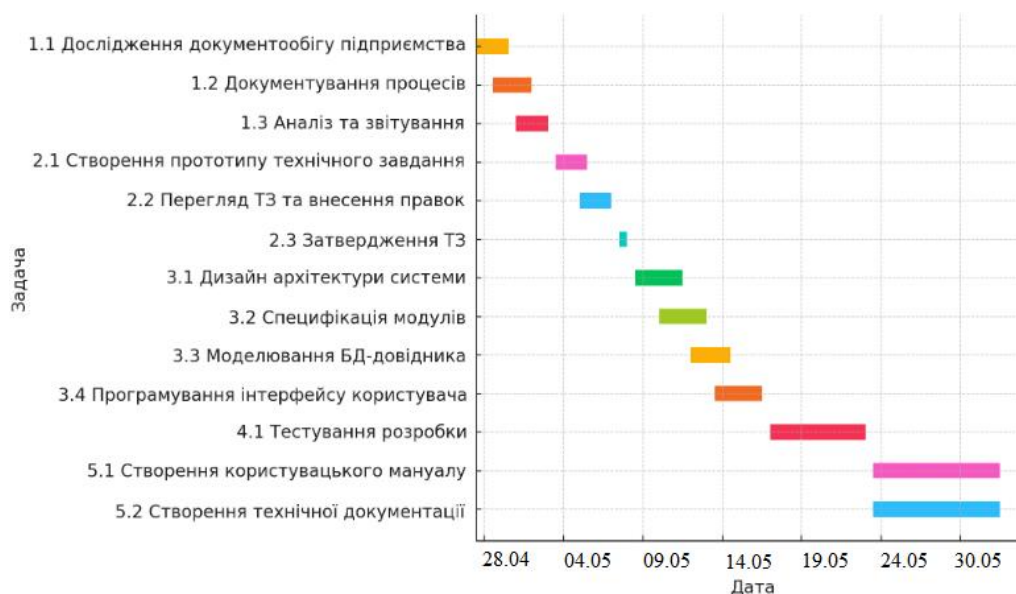


Рисунок 2.1 — Діаграма Ганта проекту

## 2.6. Порядок контролю і приймання системи

2.6.1 Користувацький модуль системи вводиться на комп'ютерах працівників бухгалтерії. Модуль, що виконуватиме обчислення під час обробки зображень документів, класифікації та розпізнавання символів повинен знаходитись на спеціально відведеній системі-сервері з наявною дискретною графічною картою.

При введенні в дію система повинна пройти приймальні випробування згідно з ДСТУ 3974-2000.

2.6.2. Випробування для визначення працездатності і рішення про можливість приймання системи в дослідну експлуатацію проводять розробники разом із замовником. Програму випробувань складає розробник і затверджує замовник.

2.6.3. Здача в дослідну експлуатацію здійснюється на основі технічного завдання та інструкції користувача. За результатами дослідної експлуатації формується перелік доробок і рекомендовані строки їх виконання.

2.6.4. Введення в дію системи оформлюється актом здачі-прийому.

## **2.7. Вимоги до складу і змісту робіт із підготовки до введення системи в дію**

Для введення системи в дію замовник виконує ряд робіт із підготовки об'єкта:

- проводить укомплектування технічних засобів;
- організовує навчання користувачів системи роботі на персональному комп'ютері і вивчення інструкції з її експлуатації;
- організовує навчання користувачів системи роботі з сканером і вивчення інструкції з його експлуатації;
- ознайомлює користувачів із моделлю бази даних підприємства;
- проводить дослідну експлуатацію і вводить систему в дію.

## **2.8. Вимоги до документації**

2.8.1. На систему розробляється комплекс документації у складі: технічне завдання та технічний проєкт.

2.8.2. Документація на систему розробляється у відповідності з вимогами Державних стандартів серії 19 «Єдина система програмної документації» та серії 24 «Єдина система стандартів автоматизованих систем управління».

## 2.9. Джерела розробки

2.9.1 При розробці технічного завдання на систему використано наступні документи:

- ДСТУ 3008-2015. Документація. Звіти у сфері науки і техніки. Структура та правила оформлювання;
- ДСТУ 3973–2000 Система розроблення та поставлення продукції на виробництво;
- ДСТУ Б В.2.5–82:2016 Електробезпека в будівлях і спорудах. Вимоги до захисних заходів від ураження електричним струмом.

## **РОЗДІЛ 3. ПРОЄКТУВАННЯ, СТВОРЕННЯ ТА АПРОБАЦІЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ БУХГАЛТЕРСЬКИХ ДОКУМЕНТІВ**

### **3.1. Опис та обґрунтування вибору програмно-технічних засобів розроблення програмного продукту**

Для реалізації системи використано об'єктно-орієнтовану мову програмування високого рівня Python версії 3.9.13 та ряд безкоштовних бібліотек [15].

#### **3.1.1. Python.**

Python є оптимальним вибором для реалізації проєкту з оптичним розпізнаванням символів та класифікації документів завдяки своїй гнучкості, швидкості та простоті у написанні, багатій екосистемі бібліотек і активній спільноті розробників.

#### **3.1.2. EasyOCR.**

Легка та проста у використанні бібліотека Python для оптичного розпізнавання символів (optical character recognition – OCR). Вибір впав саме на неї через підтримку української та англійської мов, доступний для швидкого освоєння функціонал, зручну документацію. Також вона підтримує навчання та інтеграцію власних моделей розпізнавання. В планах розширення функціоналу проєкту EasyOCR зазначено про майбутню імплементацію розпізнавання людської каліграфії, що в майбутньому може знайти застосування в інтелектуальній системі кваліфікаційної роботи [16, 17].

#### **3.1.3. OpenCV.**

Це потужна бібліотека, призначена для обробки зображень, комп'ютерного зору та машинного навчання. Вона надає широкий спектр інструментів для аналізу та обробки зображень, включаючи розпізнавання об'єктів, обличь, тексту, відстеження руху, калібрування камер, а також підтримку глибоких нейронних мереж.

Завдяки своїй ефективності та гнучкості, OpenCV широко використовується в наукових дослідженнях, промислових застосуваннях, робототехніці, системах безпеки та мобільних додатках [18, 19].

Ця бібліотека допоможе проводити попередню обробку зображення шляхом вирівнювання зображення, застосування фільтрів, пошуку таблиць та маніпуляції пікселями зображення, щоб спростити процес оптичного розпізнавання символів модулем EasyOCR.

#### 3.1.4. Hunspell.

Проста бібліотека для перевірки правопису окремих слів, створена на базі порту на Python бібліотеки Hunspell з підтримкою українського словника. Hunspell — це програма перевірки орфографії, якою користуються LibreOffice, OpenOffice.org, Mozilla Firefox та Thunderbird, Google Chrome, а також такі власницькі програмні пакети, як macOS, InDesign, memoQ, Opera та SDL Trados [20, 21].

Простий у використанні і зрозумілий функціонал бібліотеки та можливість використовувати її офлайн задовольняє усі базові потреби системи стосовно перевірки правопису.

#### 3.1.5. PyQt.

Бібліотека для реалізації графічного інтерфейсу користувача на Python на базі інструментарію поширеного крос-платформеного фреймворка Qt. Великий об'єм класів віджетів для будь яких потреб та наявність нативного додатку Qt Designer для конструювання графічного інтерфейсу шляхом перетягування віджетів на поле робить цю бібліотеку дуже зручним інструментом для реалізації проекту [22, 23].

#### 3.1.6. SQLAlchemy.

SQLAlchemy — це потужний інструмент для роботи з реляційними базами даних у Python, який поєднує два рівні взаємодії:

1. Дає змогу будувати та виконувати SQL-запити програмно, використовуючи знайомі Python-конструкції. Це повний контроль над

генерованими запитами, але при цьому нема необхідності писати SQL рядок за рядком.

2. Дозволяє відображати таблиці бази даних у вигляді Python-класів, а рядки — як об'єкти цих класів. Це спрощує роботу з даними: достатньо створити екземпляр класу та додати його в сесію і потім робити потрібні запити через функції сесії.

Бібліотека підтримує різні СУБД: проєкт зможе працювати як з SQLite (для швидкого прототипування), так і з PostgreSQL чи MySQL, міняючи лише рядок підключення [24].

Така бібліотека ідеально вписується в архітектуру розробки системи, де необхідні лише функції перегляду БД-довідки та вставки у БД підприємства.

### 3.1.7. SQLite.

Для створення та управління БД-довідкою використано програмний застосунок DB Browser for SQLite. Мінімальна необхідна кількість функцій у системі управління і збереження усієї БД у одному файлі дозволяє легко та швидко розробити БД-довідку для класифікації документів. За потреби, базу даних завжди можна буде мігрувати до іншого, більш захищеного середовища [25].

## 3.2. Проєктування та створення бази даних-довідника

Така база даних потрібна для динамічного підтримання ознак для класифікації і коректного зчитування даних документів. Модуль розпізнавання тексту очікуватиме будь які дані з цієї бази даних та не потребуватиме змін у коді у разі внесення нових ознак. Для реалізації БД-довідника була змодельована наступна база даних:

- таблиця з назвами типів документів, з якими використовуватиметься система;
- таблиця з назвами стовпчиків у таблицях документів (Рис. 3.1);
- таблиця з назвами полів документа (Рис. 3.2).

| № | Товар   | Кількість | Од.   | Ціна без ПДВ | Сума без ПДВ |
|---|---|-----------|-------|--------------|--------------|
| 1 | лак меблевий акрил-поліуретановий Trae Lux Moebel lak (0,25л) | 10        | банки | 150,00       | 1500,00      |
| 2 | ДВП СТ-40 (2.5мм×2440 мм ×1220 мм)                            | 5         | лист  | 95,00        | 475,00       |

Основні економічні показники господарської діяльності  
ПРАТ «Білоцерківський консервний завод»

| № п/п | Показники            | Один виміру | Факт з початку року | Відповідний період минулого року | Приріст, зниження (+, -) | %     |
|-------|----------------------|-------------|---------------------|----------------------------------|--------------------------|-------|
| 1.    | Обсяги виробництва   |             | 2127                | 2024                             |                          |       |
|       | у порівняльних цінах | тис.грн     |                     |                                  |                          |       |
|       | у діючих цінах       | тис.грн     | 56544,6             | 53750,6                          | +2824,0                  | 105,2 |

Рисунок 3.1 — Приклад ключових назв таблиць

**Постачальник:** ТОВ «Едельвейс»,  
14032, м. Чернігів, вул. 1-ї танкової бригади, 17, тел. +380(462)151554,  
р/р UA8130033500002600333333333 у банку АКБ «Аваль», м. Чернігів

**Покупець:** ТОВ «Орхідея»,  
14032, м. Чернігів, вул. Доценка, 35, тел. +380(46)181864,  
р/р UA8130033500002600333333333 у банку АКБ «Приватбанк», м. Чернігів,

**Договір:** купівлі-продажу № 1 від 03.01.2024 р.

| Автопідприємство   | Водій       | Бердочник І.        |
|--|-------------|---------------------|
| назва  | п. І. по б. | п. І. по б.         |
| <b>Замовник (платник)</b> Мартинюк Наталія Василівна, ФО-П                                   |             |                     |
| <b>Вантажовідправник</b> ЗАТ Консьюмерс Скло Зоря 35314 с. Зоря, Промислова 1 00380362692000 |             |                     |
| <b>Вантажоодержувач</b> Мартинюк Наталія Василівна, ФО-П                                     |             |                     |
| <b>Пункт вантаження</b> ЗАТ Консьюмерс Скло 35314 Зоря вул Промисл                           |             | Пункт розвантаження |
| <b>Переадресування</b> назва і адреса нового вантажоодержувача, підля відповідальної особи   |             | 1. П<br>2. П        |

Рисунок 3.2 — Приклад ключових полів на документі

Полями таблиці з назвами типів документів є (Рис. 3.3):

- назва документа;
- опис типу документа (опціонально);
- формат документа (якщо один і той же документ може бути представлено у різних формах, це можна вказати).

| doc_types   |         |
|-------------|---------|
| doc_type_id | INTEGER |
| doc_name    | TEXT    |
| description | TEXT    |
| format      | INTEGER |

Рисунок 3.3 — Представлення таблиці назв документів у DB Browser for SQLite

Полями таблиці з назвами ключових стовпчиків документів є (Рис. 3.4):

- зовнішній ключ, що посилається на назву документа із таблиці про типи документів;
- текст ключової назви;
- опис стовпчику документа (опціонально).

| columns     |         |
|-------------|---------|
| column_id   | INTEGER |
| doc_type    | INTEGER |
| column_text | TEXT    |
| description | TEXT    |

Рисунок 3.4 — Представлення таблиці ключових стовпчиків документа у DB Browser for SQLite

Полями таблиці з ключовими написами на документах є (Рис. 3.5):

- зовнішній ключ, що посилається на назву документа із таблиці про типи документів;
- текст ключової фрази;
- опис поля документа (опціонально);
- значення відношення висоти рядка документа до висоти регіону, де знаходиться значення цього поля;
- значення відношення висоти рядка документа до ширини регіону, де знаходиться значення цього поля.

| labels        |         |
|---------------|---------|
| label_id      | INTEGER |
| doc_type      | INTEGER |
| label_text    | TEXT    |
| description   | TEXT    |
| field_h_ratio | NUMERIC |
| field_w_ratio | NUMERIC |

Рисунок 3.5 — Представлення таблиці ключових написів документа у DB Browser for SQLite

Поточні записи у таблиці типів документів зображено на рисунку 3.6.

|   | <u>doc_type_id</u> | doc_name                                       | description | format |
|---|--------------------|--|-------------|--------|
|   | Filter             | Filter   | Filter      | Filter |
| 1 | 1                  | Товарно-транспортна накладна                   | NULL        | NULL   |
| 2 | 2                  | Звіт про рух грошових коштів                   | NULL        | NULL   |
| 3 | 3                  | Основні економічні показники господарської ... | NULL        | NULL   |
| 4 | 4                  | Видаткова накладна                             | NULL        | NULL   |

Рисунок 3.6 — Типи документів для майбутньої обробки

Частина поточних записів у таблиці назв стовпчиків документів зображено на рисунку 3.7.

|    | <u>column_id</u> | doc_type | column_text *                              | description |
|----|------------------|----------|--|-------------|
|    | Filter           | Filter   | Filter                                     | Filter      |
| 1  | 31               | 3        | %  | NULL        |
| 2  | 9                | 1        | Вид пакування                              | NULL        |
| 3  | 29               | 3        | Відповідний період минулого року           | NULL        |
| 4  | 8                | 1        | Документи з вантажем                       | NULL        |
| 5  | 24               | 2        | За аналогічний період попереднього року    | NULL        |
| 6  | 23               | 2        | За звітний період                          | NULL        |
| 7  | 13               | 1        | Клас вантажу                               | NULL        |
| 8  | 22               | 2        | Код  | NULL        |
| 9  | 12               | 1        | Код вантажу                                | NULL        |
| 10 | 10               | 1        | Кількість місць                            | NULL        |
| 11 | 14               | 1        | Маса брутто, т                             | NULL        |
| 12 | 3                | 1        | Назва продукції, товару (вантаж) або № ... | NULL        |
| 13 | 1                | 1        | Номенкл.№, код                             | NULL        |

Рисунок 3.7 — Назви стовпчиків різних деяких документів

Частина поточних записів у таблиці ключових написів документів зображено на рисунку 3.8.

|    | label_id | doc_type | label_text *                            | description | field_h_ratio | field_w_ratio |
|----|----------|----------|---|-------------|---------------|---------------|
|    | Filter   | Filter   | Filter                                  | Filter      | Filter        | Filter        |
| 1  | 17       | 1        | 1. Прицеп,прицеп,1прицеп,1 причеп       | NULL        | 1.7           | 20            |
| 2  | 18       | 1        | 2. Прицеп,прицеп,2прицеп,2 причеп       | NULL        | 1.7           | 20            |
| 3  | 7        | 1        | Автомобіль                              | NULL        | 1.7           | 20            |
| 4  | 8        | 1        | Автопідприємство                        | NULL        | 1.7           | 14            |
| 5  | 12       | 1        | Вантажовідправник                       | NULL        | 1.7           | 40            |
| 6  | 13       | 1        | Вантажоодержувач                        | NULL        | 1.7           | 40            |
| 7  | 10       | 1        | Вид перевезень,вид,перевезень           | NULL        | 1.7           | 12            |
| 8  | 9        | 1        | Водій                                   | NULL        | 1.7           | 19.5          |
| 9  | 19       | 1        | Всього відпущено на суму,відпущено,суму | NULL        | 1.7           | 16            |
| 10 | 6        | 4        | Всього з ПДВ                            | NULL        | 1.2           | 7             |
| 11 | 5        | 4        | Від покупця                             | NULL        | 1.2           | 11.5          |
| 12 | 4        | 4        | Від постачальника                       | NULL        | 1.2           | 13.3          |
| 13 | 24       | 2        | Головний бухгалтер,головний,бухгалтер   | NULL        | 2             | 27            |
| 14 | 3        | 4        | Договір                                 | NULL        | 1.2           | 27            |
| 15 | 11       | 1        | Замовник (платник),замовник,платник     | NULL        | 1.7           | 40            |
| 16 | 22       | 3        | Керівник                                | NULL        | 1.5           | 16            |
| 17 | 23       | 2        | Керівник                                | NULL        | 2             | 27            |
| 18 | 15       | 1        | Переадресування                         | NULL        | 5.5           | 30            |
| 19 | 2        | 4        | Покупець                                | NULL        | -3            | 27            |
| 20 | 1        | 4        | Постачальник                            | NULL        | -3            | 27            |

Рисунок 3.8 — Ключові поля деяких документів

Потенційне виявлення ключового поля навіть у випадку помилок під час розпізнавання тексту документа досягається шляхом переліку через кому окремих унікальних частин ключової фрази (наприклад, замість «Замовник (платник)» без помилок розпізнало лише слово «платник» - цього достатньо).

Значення відношення по висоті, наприклад «1.7» означає, що якщо висотою рядка поля документа виявиться, нехай, 20 пікселів, то висота регіону, де знаходитиметься значення цього поля дорівнює  $20 * 1.7 = 34$  пікселі. Відповідно, ширина пошуку того значення дорівнює  $20 * 20 = 400$  пікселів. У модуль оптичного розпізнавання символів буде передано червоний регіон, який відраховується від правого нижнього кута ключового напису (Рис. 3.9).

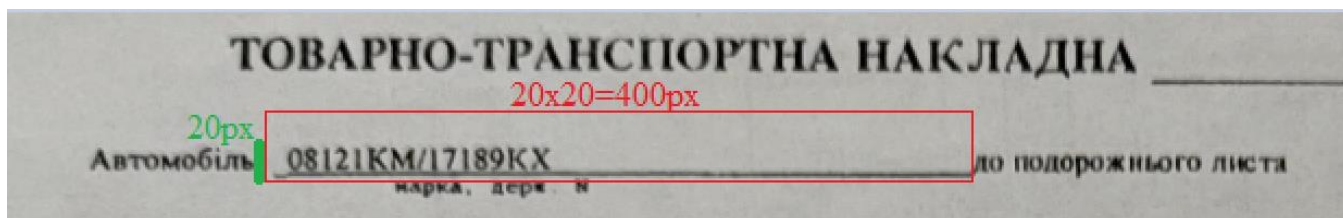


Рисунок 3.9 — Розрахунок розмірів і позиції регіону, де знаходиться значення поля

Значення відношення з мінусом (останні два рядка на рисунку 3.8) означають, що регіон значення заходить під рядок ключового слова (Рис. 3.10).

Постачальник: ТОВ «Едельвейс»,  
14032, м. Чернігів, вул. 1-ї танкової бригади, 17, тел. +380(462)151554,  
р/р UA81300335000026003333333333 у банку АКБ «Аваль», м. Чернігів

Покупець: ТОВ «Орхідея»,  
14032, м. Чернігів, вул. Доценка, 35, тел. +380(46)181864,  
р/р UA81300335000026003333333333 у банку АКБ «Приватбанк», м. Чернігів,

Рисунок 3.10 — Приклад особливого випадку місцезнаходження значення поля  
У такому випадку мінусове значення обертає направлення розгортання регіону і він вже відраховується від правого верхнього кута ключового напису.

Яким чином БД-довідник допомагає класифікувати документ буде описано в підрозділі 3.3.

### 3.3. Реалізація функцій системи

Інтелектуальну систему створено та всі її функції реалізовано з використанням Python та бібліотек Python.

Основними модулями розробки виступають:

- Модуль препроцесингу (попередньої обробки зображення), в основі якого лежить бібліотека Python OpenCV. Основні задачі, що виконує модуль препроцесингу, наступні:
  - метод `cv2.imread` завантажує надане йому зображення у пам'ять для майбутньої з ним роботи. Одночасно, цей метод переводить зображення на канал кольору відтінків сірого, бо присутність кольорів лише сповільнить майбутню обробку зображення та не грає ролі у розпізнаванні символів;

- метод `cv2.blur` застосовує фільтр розмивання за вказаними параметрами. Необхідність такого фільтру полягає у позбавленні зображення від шуму або в стовщенні занадто тонких або дрібних символів з ціллю покращення їх розпізнавання;
- метод `cv2.threshold` (або `cv2.adaptiveThreshold`) реалізує бінаризування зображення у відповідності з вказаним йому методом та параметрами. Бінаризація перетворює пікселі зображення у один з двох кольорів – чорний або білий в залежності в вказаного порогу яскравості пікселя або параметрів алгоритму (у випадку використання іншого методу бінаризації). У програмній роботі з документами, бінаризація допомагає розділити зображення на пікселі аркуша/фону та пікселі символів/тексту для максимально швидкого та точного розпізнавання тексту на наступних етапах обробки;
- метод `cv2.resize` дозволяє маніпулювати розміром зображення та стискати/збільшувати його у разі необхідності у відповідності з обраним алгоритмом. Зменшення розміру зображення корисне у разі занадто великого початкового значення, що робить його обробку повільним. Збільшення зображення допомагає у випадках, коли дрібним символам не достає пікселів, щоб алгоритм розпізнавання розпізнав його ознаки (різні алгоритми масштабування зображення генеруватимуть нові пікселі зображення за різними принципами);
- модуль оптичного розпізнавання символів, в основі якого лежить бібліотека Python EasyOCR, допомагає реалізувати зчитування тексту з документів:
  - клас `easyocr.Reader` ініціалізує модель штучного інтелекту для оптичного розпізнавання символів. Конструктор цього класу приймає параметри мов для розпізнавання та відомість про те, чи можна використовувати обчислювальні потужності графічної карти комп'ютера;

- метод `Reader.readtext` приймає зображення та опціональні додаткові параметри обробки і повертає результат розпізнавання тексту у вигляді списку з трьома об'єктами: координати знайденого слова, розпізнане слово, шкала впевненості розпізнавання. Завдяки координатам, що надає цей метод, та розпізаному тексту можлива реалізація більшості інших алгоритмів класифікації та пошуку потрібних даних на зображенні документа;
- модуль перевірки правопису, в основі якого лежить бібліотека Python `Phunspell`, допомагає значно покращити результати розпізнавання та використати їх як для класифікації так і для майбутнього експорту:
  - клас `Phunspell` отримує у свій конструктор мову, словник якої користувач бажає використовувати в якості перевірки слів;
  - метод `Phunspell.lookup` приймає на вхід слово та повертає логічне значення про те, чи слово знайдено у словнику. Якщо ні, то, скоріше всього, у слові помилка;
  - метод `Phunspell.suggest` приймає на вхід слово з помилкою та повертає список слів із свого словника у якості пропозиції для виправлення;
- модуль класифікації документа це набір алгоритмів для пошуку ключових слів у розпізаному на документі тексті та визначенні його типу. Бібліотека Python `SQLAlchemy` та ряд її класів і функцій дозволяють підключення до БД-довідника, описаного у підрозділі 3.2. Модуль отримує з бази даних потрібні таблиці з назвами типів документів, ключовими словами та параметрами для зчитування даних з документу. Основна ідея алгоритму класифікації полягає у підрахунку кількості знайдених унікальних ключових слів на зображенні документа. Тип документа з найбільшою кількістю знайдених ключових слів вважається типом документа на зображенні. Визначення правильного типу документа це вирішальний фактор для майбутнього успішного експорту даних з зображення документа;

- модуль графічного інтерфейсу користувача, в основі якого лежить бібліотека Python PyQt5. Створений у окремому середовищі-дизайнері макет графічного інтерфейсу експортується у окремий модуль з кодом Python. Його підключення до власних модулів дозволяє прописати логіку інтерфейсу.

### 3.4. Інструкція користувача

Експлуатація системи починається з вибору користувачем файлів зображень документів, які він бажає обробити (Рис. 3.11).

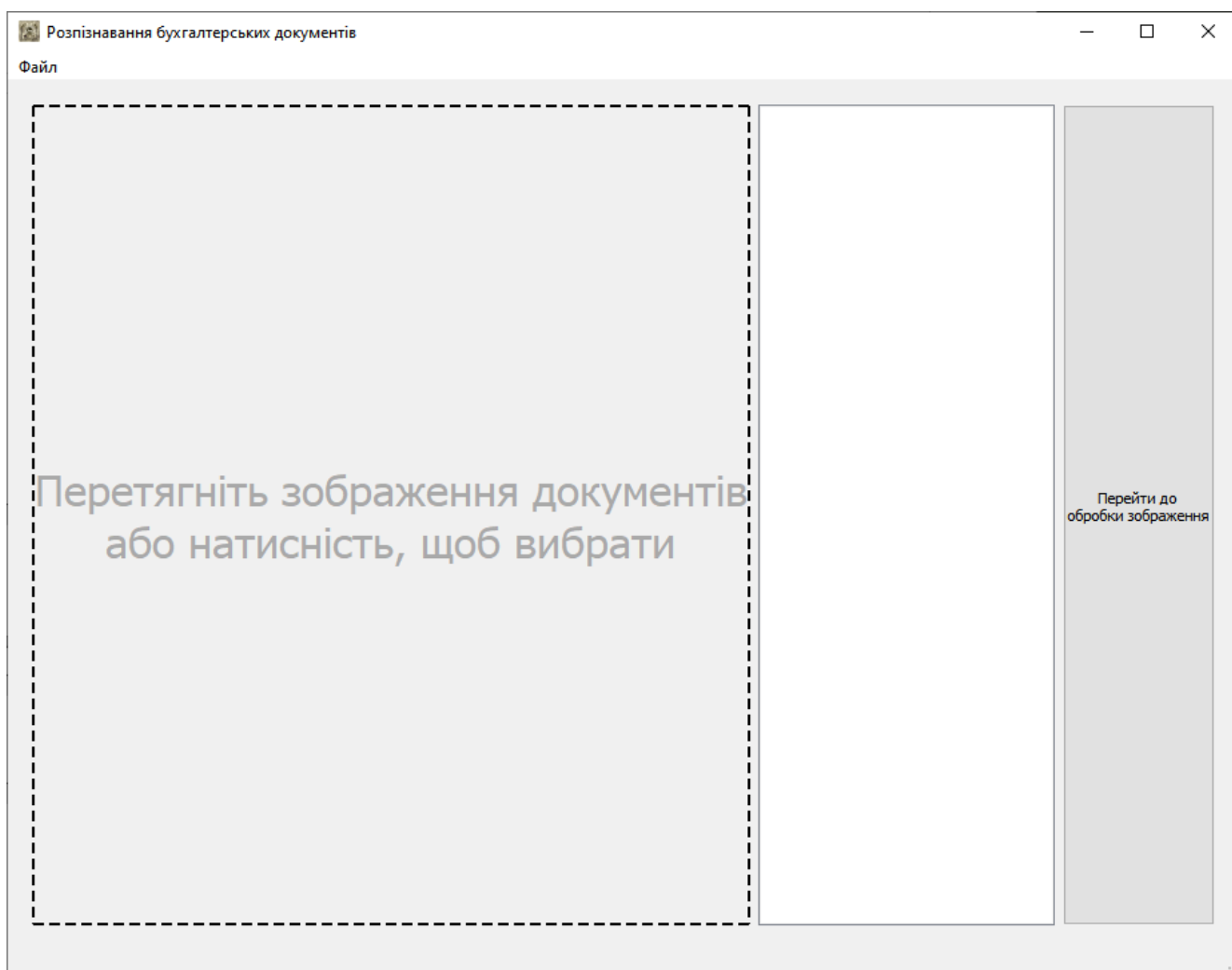


Рисунок 3.11 – Головне меню програмного застосунку

Після імпорту зображень документів, користувач може перейти до наступного етапу роботи з програмою, а саме – застосування фільтрів попередньої обробки (Рис. 3.12). Повзунки та поля з випадаючим списком дозволяють тонко налаштувати фільтри для зображення.

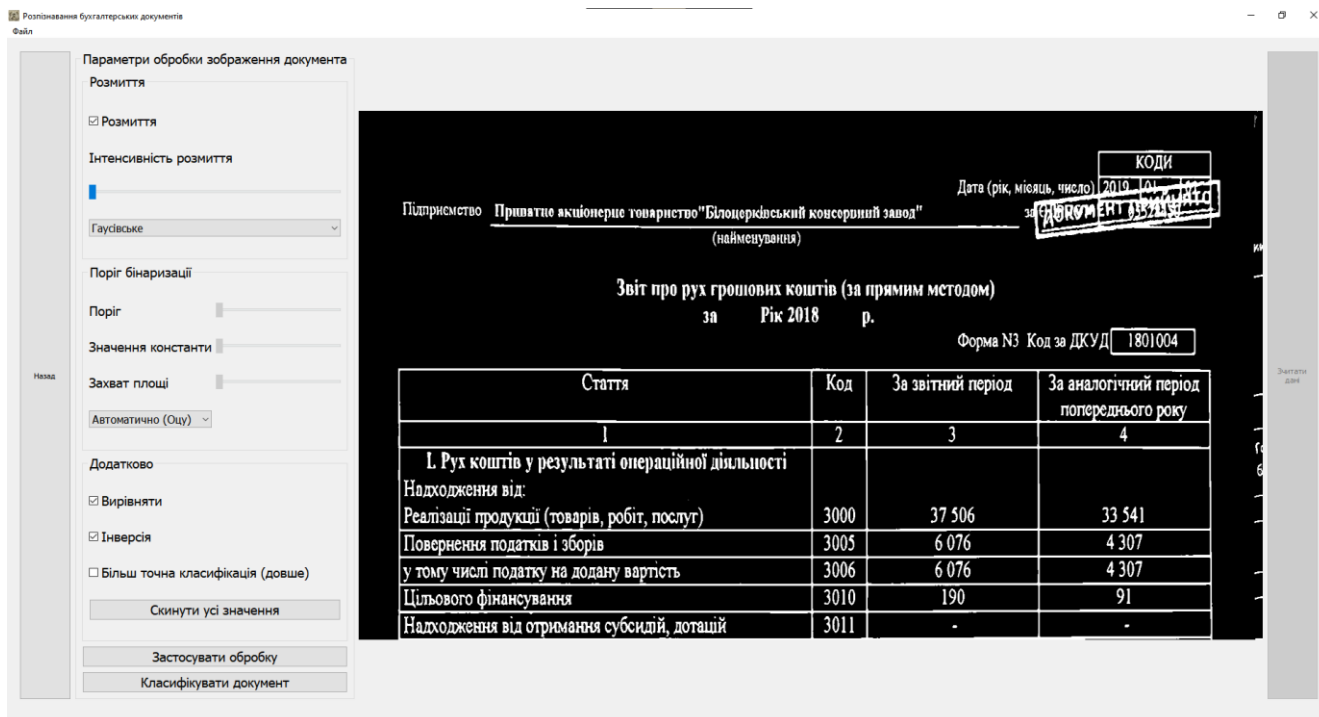


Рисунок 3.13 — Повідомлення про завершення алгоритму класифікації

У випадку коректного визначення типу документа, користувач може перейти до етапу зчитування та експорту даних з документа у відповідні таблиці бази даних підприємства.

У вікні результату зчитування даних з документа, користувач має право скорегувати потрібні значення у полях таблиці перед їх відправкою у базу даних.

Результати виконаних запитів передачі даних до бази даних підприємства користувач може переглянути у відповідній інформаційній системі підприємства, якщо у нього є доступ (Рис 3.14).

| document_id | document_name                                    | document_date | head        |
|-------------|--|---------------|-------------|
| Filter      | Filter   | Filter        | Filter      |
| 1           | Звіт про рух грошових коштів (за прямим методом) | 2018 рік      | Керівник: ? |

| funds_operations_id | funds_movement_id | article                                       | row_code | reported | prior_reported |
|---------------------|-------------------|---|----------|----------|----------------|
| Filter              | Filter            | Filter  | Filter   | Filter   | Filter         |
| 1                   | 1                 | Реалізація продукції (товарів, робіт, послуг) | 3000     | 37506    | 33541          |
| 2                   | 1                 | Повернення податків і зборів                  | 3005     | 6076     | 4307           |
| 3                   | 1                 | у тому числі податку на додану вартість       | 3006     | 6076     | 4307           |
| 4                   | 1                 | Цільового фінансування                        | 3010     | 190      | 91             |
| 5                   | 1                 | Надходження від отримання субсидій, дотацій   | 3011     | NULL     | NULL           |

Рисунок 3.14 — Результат імпорту записів у таблиці БД

### 3.5. Тестування програмного продукту

Тест-план охоплює чотири рівні: модульні, інтеграційні та системні. Для кожного рівня визначено набір тест-кейсів, вхідні дані, очікуваний результат та статус проходження (таблиця 3.1).

Таблиця 3.1. Проведені тести системи

| № | Рівень        | Опис тест-кейсу  | Вхідні дані                            | Очікуваний результат                            | Статус   |
|---|---------------|--|--|---|----------|
| 1 | Модульний     | Перевірка функції збільшення та попередньої обробки зображень  | Тестове зображення (300×200)           | Масив пікселів розміром 800×600, без артефактів | Пройдено |
| 2 | Модульний     | Перевірка OCR-модуля на простому текстовому шаблоні            | Зображення з чітким текстом "Test 123" | Рядок "Test 123"                                | Пройдено |
| 3 | Модульний     | Перевірка працездатності модуля перевірки правопису            | Слово "документт"                      | Рекомендація "документ"                         | Пройдено |
| 5 | Інтеграційний | Імпорт ключових слів з БД-довідника                            | З'єднання до бази даних                | Об'єкти-словники з повним набором записів БД    | Пройдено |
| 6 | Системний     | Повний сценарій користувача: від вибору файлу до експорту в БД | Фотографії, скани документів           | Автоматичний імпорт у систему підприємства      | Пройдено |

З існуючих проблем, пов'язаних з точністю розпізнавання та експлуатації варто виявити наступне:

1. Часткове або зовсім неточне розпізнавання прописних літер, цифр, слів під нахилом, значень поза полями таблиць документа.
2. Відсутність підказок в інтерфейсі користувача у повному обсязі. У повзунків не відображається явне числове значення.
3. Програма «зависає» під час виконання важкого алгоритму розпізнавання тексту або класифікації документа.
4. Якщо документ більший за вікно програми, то він або не вміщається цілком, або спотворюється під час зміни розміру вікна (лише візуальний ефект, що не впливає на розпізнавання тексту).
5. Програма автоматично застосовує виправлення слів з помилками замість пропозиції користувачу списку можливих слів. Це призводить до виправлення слів на існуючі, проте некоректні варіанти.

## ВИСНОВКИ

В результаті виконання кваліфікаційної роботи були виконані усі завдання що, стосуються розробки інтелектуальної системи розпізнавання бухгалтерських документів. А саме:

Досліджено документообіг бухгалтерії ПрАТ «Білоцерківський консервний завод» та проаналізовано стан комп'ютеризації на підприємстві. Результати досліджень дали змогу виявити проблеми та неефективність ручного введення даних для оцифрування та архівації первинних бухгалтерських документів.

Досліджено методи оптичного розпізнавання тексту, підходи до інтелектуальної класифікації бухгалтерських документів та існуючі рішення для автоматизації документообігу на вітчизняному ринку, та програмні інструменти для реалізації системи розпізнавання бухгалтерських документів.

Результати проведення аналізу дали змогу сформулювати вимоги щодо розробки інтелектуальної системи розпізнавання бухгалтерських документів.

Це в свою чергу дало змогу створити технічне завдання на розробку інтелектуальної системи розпізнавання бухгалтерських документів.

Базуючись на створеному технічному завданні, розроблено інтелектуальну систему розпізнавання бухгалтерських документів.

Розроблена інтелектуальна система розпізнавання бухгалтерських документів надає змогу розв'язання низки актуальних у контексті документообігу задач та проблем, включаючи автоматизацію процесу введення даних з паперових документів; підвищення ефективності та точності обробки бухгалтерських документів; економію часу працівників; безпеку документів.

Отримані результати розробки та тестування дозволяють розглядати цю систему, як складову частину системи управління, підприємствами, що бажають використовувати сучасні інтелектуальні технології та технології автоматизації діяльності підприємства.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Андросова О. Ф. Цифрові інструменти на промислових підприємствах України. *Актуальні проблеми економіки*. 2025. Т. 2. С. 6–19.
2. Бергер А. Д. Цифрова трансформація бізнесу в компаніях харчової індустрії. *Наукові праці НУХТ*. 2024. Т.30, вип. 1. С. 81-90. DOI: 10.24263/2225-2924-2024-30-1-9.
3. @klear-stack. *Medium*. *Key Benefits of OCR Technology in Accounting and Finance*. URL: <https://medium.com/%40klear-stack/key-benefits-of-ocr-technology-in-accounting-and-finance-27f0ad26de97> (дата звернення: 15.05.2025).
4. IBM. *What is optical character recognition (OCR)?*. URL: <https://www.ibm.com/think/topics/optical-character-recognition> (дата звернення: 15.05.2025).
5. Київський обласний центр зайнятості. ВАТ "Білоцерківський консервний завод" постійний партнер обласної служби зайнятості. URL: [https://web.archive.org/web/20150407010730/http://www.dcz.gov.ua/kir/control/uk/publish/article?art\\_id=7011](https://web.archive.org/web/20150407010730/http://www.dcz.gov.ua/kir/control/uk/publish/article?art_id=7011) (дата звернення: 15.05.2025).
6. "Білоцерківський консервний завод" – Приватне акціонерне товариство. ПРОВАХІВАННЯ. URL: <http://patbkz.in.ua> (дата звернення: 15.05.2025).
7. Бізнес-Гід. БІЛОЦЕРКІВСЬКИЙ КОНСЕРВНИЙ ЗАВОД, ПРОВАХІВАННЯ. URL: <https://krjat.business-guide.com.ua> (дата звернення: 15.05.2025).
8. M.E.Doc. Рішення. URL: <https://medoc.ua> (дата звернення: 15.05.2025)
9. INNORA. Інтелектуальне розпізнавання документів. URL: <https://innora.com.ua/rozpiznavannya-dokumentiv> (дата звернення: 15.05.2025).
10. eIDoc. інтелектуальна обробка та розпізнавання документів. URL: <https://dms-solutions.co/uk/eldoc/eldoc-intelligent-document-processing> (дата звернення: 15.05.2025).
11. BitFaktura. URL: <https://bitfaktura.com.ua> (дата звернення: 15.05.2025)
12. CleverForms. Модуль OCR CleverForms. URL: <https://clever-forms.com/module-ocr-cleverforms> (дата звернення: 15.05.2025).

13. Trivedi A. *Analytics Vidhya. Top 8 OCR Libraries in Python to Extract Text from Image*. URL: <https://www.analyticsvidhya.com/blog/2024/04/ocr-libraries-in-python> (дата звернення: 15.05.2025).
14. Осмятченко В., Матюха М. Електронне документування у бухгалтерському обліку. *Вісник КНТЕУ*. 2013. вип. 2. ISSN 1727-9313.
15. *Python. Python 3.9.22 documentation*. URL: <https://docs.python.org/3.9/> (дата звернення: 15.05.2025).
16. *GitHub. JaidedAI / EasyOCR*. URL: <https://github.com/JaidedAI/EasyOCR> (дата звернення: 15.05.2025).
17. *IBM. What is optical character recognition (OCR)?*. URL: <https://www.ibm.com/think/topics/optical-character-recognition> (дата звернення: 15.05.2025).
18. Kaehler A. & Bradski G. *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*. Sebastopol, CA : O'Reilly Media, Inc.
19. *OpenCV*. URL: <https://opencv.org/> (дата звернення: 15.05.2025).
20. *GitHub. dvwright / phunspell*. URL: <https://github.com/dvwright/phunspell> (дата звернення: 15.05.2025).
21. *Hunspell. About*. URL: <https://hunspell.github.io> (дата звернення: 15.05.2025)
22. *RiverBank Computing. What is PyQt?*. URL: <https://riverbankcomputing.com/software/pyqt/intro> (дата звернення: 15.05.2025).
23. *PythonGUIs. The complete PyQt5 tutorial*. URL: <https://www.pythonguis.com/pyqt5-tutorial/> (дата звернення: 15.05.2025).
24. *PyPi. SQLAlchemy*. URL: <https://pypi.org/project/SQLAlchemy> (дата звернення: 15.05.2025).
25. *SQLite. What Is SQLite?*. URL: <https://www.sqlite.org> (дата звернення: 15.05.2025).

## ДОДАТКИ

## Додаток А.

| № п/п | Назва консервів                                      | НТД на продукцію, що випускається |
|-------|--|-----------------------------------|
| 1.    | Салати овочеві                                       | ДСТУ 3749-98                      |
| 2.    | Овочі фаршировані в томатному соусі                  | ГОСТ 1016-90                      |
| 3.    | Овочі різані в томатному соусі                       | ГОСТ 18611-73                     |
| 4.    | Горошок зелений консервований                        | ГОСТ 15842-90                     |
| 5.    | Горошок суповий                                      | ТУ 46.72 України 1'093            |
| 6.    | Маринади овочеві                                     | ГОСТ 1633-73                      |
| 7.    | Томати консервовані                                  | ГОСТ 7231-90                      |
| 8.    | Огірки консервовані                                  | ГОСТ 20144-74                     |
| 9.    | Перець маринований цілий                             | ТУ 61 України 90-92               |
| 10.   | Соуси томатні  | ДСТ України 2118-93               |
| 11.   | Перші обідні страви                                  | ГОСТ 18316-73                     |
| 12.   | Солянки. Овочі з м'ясом                              | ГОСТ 18224-72                     |
| 13.   | Соки плодови і ягідні натуральні                     | ГОСТ 656-79                       |
| 14.   | Компоти із плодів і ягід                             | ГОСТ 816-91                       |
| 15.   | Джем плодово-ягідний                                 | ГОСТ 87009-88                     |
| 16.   | Плоди і ягоди протерті або дроблені, з цукром        | ГОСТ 22371-77                     |
| 17.   | Варення  | ГОСТ 7061-88                      |
| 18.   | Повидло  | ГОСТ 6929-88                      |
| 19.   | Ікра із кабачків "Літня"                             | ТУ У 46.72.064-95                 |
| 20.   | Яловичина тушкована                                  | ГОСТ 5284-84                      |
| 21.   | Свинина тушкована                                    | ГОСТ 697-84                       |
| 22.   | Паштет м'ясний                                       | ГОСТ 12318-91                     |
| 23.   | Каша з м'ясом  | ГОСТ 8286-90                      |
| 24.   | Гриби мариновані і відварені                         | ГОСТ 286649-90                    |
| 25.   | Сік березовий з цукром                               | ТУ У 46.72.101                    |
| 26.   | Кавуни консервовані                                  | ТУ 61 України 501-92              |
| 27.   | Огірки консервовані, столові н/ форми                | ТУ 61 УССР 196-91                 |
| 28.   | Огірки солоні  | ГОСТ 7180-73                      |
| 29.   | Напівфабрикати овочеві для промислового використання | ТУ 46.12 України 17-93            |
| 30.   | Квасоля у томатному соусі                            | ТУ У 46.72.163-99                 |

Рисунок А.1 — Асортимент продукції ПрАТ «Білоцерківський консервний завод»



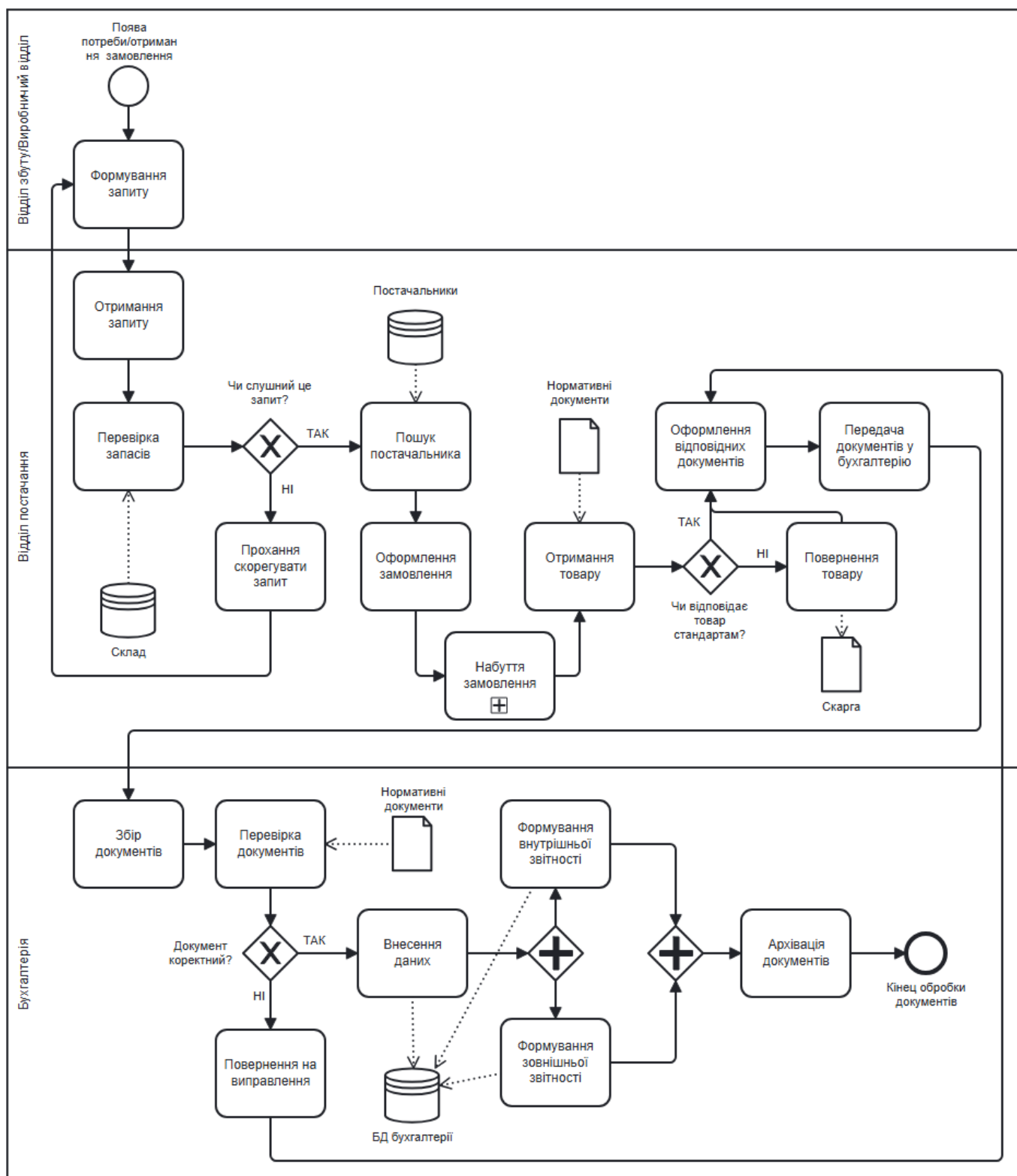


Рисунок Б.1 — Схема бізнес-процесів взаємодії бухгалтерії та відділу постачання на ПрАТ «Білоцерківський консервний завод»