ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ СИСТЕМ ТА ТЕХНОЛОГІЙ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «ІНТЕЛЕКТУАЛЬНА СИСТЕМА ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ІНФОРМАЦІЙНИХ ПРОЦЕСІВ ПІДПРИЄМСТВА»

на здобуття освітнього ступеня бакалавра (магістра)
зі спеціальності *124 Системний аналіз*
(код, найменування спеціальності)
освітньо-професійної програми Системний аналіз
(назва)

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

_____ Володимир ЛЕЙБЮК
(підпис) *Ім'я, ПРІЗВИЩЕ здобувача*

Виконав: здобувач вищої освіти гр. САД-41
ВОЛОДИМИР Лейбюк
Ім'я, ПРІЗВИЩЕ
Керівник: Ровіл НАФЄЄВ
к.ф.-м.н., доцент.. Ім'я, ПРІЗВИЩЕ
Рецензент: _____
науковий ступінь, Ім'я, ПРІЗВИЩЕ
вчене звання

Київ 2026

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Навчально-науковий інститут інформаційних технологій

Кафедра інформаційних систем та технологій

Ступінь вищої освіти Бакалавр

Спеціальність 124 Системний аналіз

Освітньо-професійна спеціальність Системний аналіз

ЗАТВЕРДЖУЮ

Завідувач кафедрою

іформаційних ситем та технологій

_____Каміла СТРОЧАК

«___»_____2026р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

Лейбюк Володимир Володимирович

Прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи: «Інтелектуальна система оцінювання ефективності інформаційних процесів підприємства»
керівник кваліфікаційної роботи:Нафєєв Ровіл Касимович к.ф.-м.н., доцент..
(ім'я ПРІЗВИЩЕ, наукова ступінь вчене звання)
затверджені наказом Державного університету інформаційно-телекомунікаційних технологій від «20» 2026р. № 51
2. Строк подання кваліфікаційної роботи: «01» червня 2026 р.
3. Вихідні дані до кваліфікаційної роботи**:** методи багатокритеріального аналізу (AHP, TOPSIS), теорія нечітких множин та нечітке виведення (система Мамдані), методи машинного навчання (LightGBM, LSTM), стандарти управління якістю інформаційних технологій (ITIL v4, ISO/IEC 20000), корпоративні дані підприємства за 24 місяці спостережень
4. Зміст розрахунково-пояснювальної записки:
1) Аналіз інформаційних процесів підприємства та методів їх оцінювання.
2)Розробка математичної моделі інтегрального показника ефективності (ІЕР).
3) Програмна реалізація інтелектуальної системи оцінювання (ІСО ЕІП).
4) Апробація системи та аналіз отриманих результатів.
5. Перелік ілюстративного матеріалу: *презентація.*
6. Дата видачі завдання: «20» лютого 2026 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів розробки	Примітка
1	Збір даних	15.03.2026 - 20.03.2026	Виконано
2	Оброботкa матеріалу	21.03.2026 - 31.03.2026	Виконано
3	Написання першого розділу розробки	01.04.2026 - 13.04.2026	Виконано
4	Написання другого розділу розробки	14.04.2026- 30.042026	Виконано
5	Написання третього розділу розробки	01.05.2026 - 09.05.2026	Виконано
6	висновки за результатами роботи	10.05.2026 - 13.05.2026	Виконано
7	Розробка призентцації	14.05.2026 - 19.05.2026	Виконано
8	ПІдготовка доповіді до захисту	20.05.2026 28.05.2026	Виконано

Здобувач вищої освіти _____ _____
(підпис) *(ім'я, ПРІЗВИЩЕ)*

Керівник кваліфікаційної роботи _____ _____
(підпис) *(ім'я ПРІЗВИЩЕ)*

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня бакалавра: 87 стор., 12 табл., 35 джерел.

Мета роботи - розробка та апробація інтелектуальної системи оцінювання ефективності інформаційних процесів підприємства, що забезпечує автоматизований розрахунок інтегрального показника ефективності на основі методів нечіткої логіки, багатокритеріального аналізу та машинного навчання.

Об'єкт дослідження: інформаційні процеси підприємства як сукупність операцій зі збирання, зберігання, опрацювання, передачі та використання інформації в межах організаційної структури для забезпечення управлінської діяльності.

Предмет дослідження: методи, моделі та програмні засоби інтелектуального оцінювання ефективності інформаційних процесів підприємства в умовах невизначеності та багатокритеріальності.

Короткий зміст роботи: У кваліфікаційній роботі проведено аналіз сучасних підходів до оцінювання ефективності інформаційних процесів підприємства, досліджено структуру інформаційних потоків та ключові показники ефективності (KPI), що характеризують якість, оперативність, вартість і результативність інформаційних процесів.

У другому розділі розроблено математичну модель інтегрального показника ефективності інформаційних процесів (IEP), сформовано систему показників оцінювання на основі методу аналізу ієрархій (AHP), реалізовано механізми нечіткого виведення за алгоритмом Мамдані та метод TOPSIS для ранжування підрозділів підприємства за рівнем ефективності.

У третьому розділі спроєктовано та реалізовано програмний прототип інтелектуальної системи оцінювання ефективності інформаційних процесів із використанням технологій Python, FastAPI, Apache Kafka, TimescaleDB та React. Проведено апробацію системи на тестових даних підприємства, виконано аналіз отриманих результатів та сформовано рекомендації щодо підвищення ефективності інформаційних процесів.

Результатом роботи є функціональний прототип інтелектуальної системи оцінювання ефективності інформаційних процесів підприємства, який забезпечує автоматизований моніторинг показників діяльності, підтримку прийняття управлінських рішень та підвищення ефективності використання інформаційних ресурсів. Розроблене рішення є масштабованим, адаптивним і може бути впроваджене на підприємствах різних галузей економіки.

КЛЮЧОВІ СЛОВА: ІНФОРМАЦІЙНІ ПРОЦЕСИ, ПІДПРИЄМСТВО, ІНТЕЛЕКТУАЛЬНА СИСТЕМА, KPI, ІНТЕГРАЛЬНИЙ ПОКАЗНИК ЕФЕКТИВНОСТІ, НЕЧІТКА ЛОГІКА, AHP, TOPSIS, МАШИННЕ НАВЧАННЯ, МІКРОСЕРВІСНА АРХІТЕКТУРА, АНАЛІЗ ДАНИХ, ЦИФРОВА ТРАНСФОРМАЦІЯ.

ЗМІСТ

ВСТУП

В умовах стрімкої цифровізації економіки інформаційні процеси перетворились зі допоміжної функції на стратегічний ресурс. Якість, швидкість та надійність інформаційних потоків безпосередньо визначають здатність підприємства реагувати на зміни ринкового середовища, приймати обґрунтовані управлінські рішення та підтримувати ефективність операційної діяльності. За оцінками міжнародних аналітичних агенцій, зокрема Gartner та IDC, підприємства, що цілеспрямовано управляють якістю своїх інформаційних процесів, досягають на 20–30 % вищої операційної продуктивності порівняно з тими, що не мають формалізованих систем такого управління.

Водночас переважна більшість вітчизняних підприємств середнього та великого бізнесу досі немає систематизованого підходу до оцінювання ефективності власних інформаційних процесів. Там, де оцінювання все-таки здійснюється, воно є фрагментарним, суб'єктивним та нерегулярним: IT-підрозділи відстежують окремі технічні метрики (час відповіді системи, кількість інцидентів), натомість бізнес-виміри якості даних - точність, повнота, своєчасність, узгодженість між системами - залишаються поза системним контролем. Відсутність єдиного погляду унеможливлює цілеспрямоване планування заходів із цифрової трансформації та не дозволяє об'єктивно оцінити їхній результат.

Вирішення цієї проблеми потребує розробки інтелектуальної системи, що поєднує методи багатокритеріального аналізу, нечіткої логіки та машинного навчання для автоматизованого, прозорого та регулярного оцінювання ефективності інформаційних процесів підприємства. Саме цим обумовлена актуальність теми кваліфікаційної роботи.

Актуальність теми. Цифрова трансформація підприємств є одним із пріоритетних напрямів розвитку вітчизняної економіки. Проте ефективність цієї трансформації неможливо оцінити без формалізованих метрик якості інформаційних процесів, що забезпечують цифрові ініціативи необхідними

даними. Нездатність вимірювати ефективність інформаційних процесів призводить до «цифрового туману» - ситуації, коли підприємство інвестує в IT-системи, не маючи об'єктивних даних про їхній вплив на бізнес-результати. Саме тому тема є актуальною як з наукової, так і з практичної точки зору.

Зв'язок роботи з науковими програмами, планами, темами. Кваліфікаційна робота виконана відповідно до навчального плану підготовки здобувачів вищої освіти ступеня бакалавра в Державному університеті інформаційно-комунікаційних технологій та відповідає науковому напряму кафедри інформаційних систем і технологій, пов'язаному з дослідженням методів інтелектуального аналізу даних та їхнього застосування в задачах управління підприємством.

Метою кваліфікаційної роботи є розробка та апробація інтелектуальної системи оцінювання ефективності інформаційних процесів підприємства (ІСО ЕІП), що забезпечує автоматизований розрахунок інтегрального показника ефективності (IEP) на основі методів нечіткої логіки, багатокритеріального аналізу та машинного навчання.

Для досягнення поставленої мети визначено такі завдання дослідження:

1. проаналізувати структуру інформаційних потоків підприємства та систематизувати підходи до оцінювання їхньої ефективності;

2. сформувати збалансовану систему ключових показників ефективності (KPI) та визначити їхні вагові коефіцієнти методом аналізу ієрархій (AHP);

3. розробити математичну модель інтегрального показника ефективності інформаційних процесів (IEP) на основі адитивно-мультиплікативної агрегації;

4. побудувати апарат нечіткого виведення (система Мамдані) для лінгвістичної класифікації IEP та модель TOPSIS для порівняльного ранжування підрозділів;

5. спроектувати мікросервісну архітектуру ІСО ЕІП та реалізувати програмний прототип з використанням Python, FastAPI, Apache Kafka, TimescaleDB та React;

6. провести апробацію системи та сформулювати рекомендації щодо підвищення ефективності інформаційних процесів підприємства.

Об'єктом дослідження є інформаційні процеси підприємства як сукупність операцій зі збирання, зберігання, опрацювання, передачі та використання інформації в межах організаційної структури з метою забезпечення управлінської діяльності.

Предметом дослідження є методи, моделі та програмні засоби інтелектуального оцінювання ефективності інформаційних процесів підприємства в умовах невизначеності та багатокритеріальності.

Методи дослідження. У процесі дослідження використано: структурно-функціональний аналіз - для вивчення архітектури інформаційних потоків та виявлення вузьких місць; метод аналізу ієрархій (AHP) - для визначення вагових коефіцієнтів критеріїв оцінювання; теорію нечітких множин та нечітке виведення за алгоритмом Мамдані - для лінгвістичної класифікації рівня ефективності; метод TOPSIS - для ранжування підрозділів за ефективністю інформаційних процесів; ансамблеві методи машинного навчання (LightGBM) та рекурентні нейронні мережі LSTM - для класифікації станів та прогнозування ІЕР; методи системного аналізу та об'єктно-орієнтованого проектування - для розробки архітектури програмного забезпечення.

Наукова новизна одержаних результатів:

Вперше запропоновано адитивно-мультиплікативну дворівневу модель інтегрального показника ефективності інформаційних процесів підприємства (ІЕР), що поєднує внутрішньокластерну зважену суму показників із міжкластерним зваженим добутком субіндексів - це усуває недолік повної компенсаторності адитивних моделей без надмірної жорсткості мультиплікативних.

Удосконалено методику лінгвістичної класифікації стану інформаційних процесів шляхом застосування п'ятирівневої нечіткої шкали на основі трикутних нечітких чисел (TFN), параметри якої визначено консенсусом дванадцяти експертів підприємства.

Дістало подальшого розвитку комплексне застосування методів AHP, TOPSIS та нечіткої логіки в єдиній інтегрованій системі оцінювання, що забезпечує одночасно кількісне порівняння альтернатив, якісну лінгвістичну інтерпретацію результатів та стійкість до невизначеності вхідних даних.

Практичне значення одержаних результатів. Розроблена система ІСО ЕІП надає підприємствам інструмент для переходу від суб'єктивного оцінювання до автоматизованого моніторингу 12 KPI у режимі реального часу. Апробація на виробничому підприємстві «АльфаМаш» підтвердила ефективність: загальний IEP зріс на 14,0 % за рік, IEP відділу закупівель - на 63,1 % завдяки цільовим заходам автоматизації, ідентифікованим системою. Термін окупності інвестицій - 20–24 місяці.

Апробація результатів дослідження. Основні положення та результати кваліфікаційної роботи апробовано в ході річного пілотного впровадження ІСО ЕІП на підприємстві «АльфаМаш» у режимі тіньового моніторингу. Результати можуть бути представлені на науково-практичних конференціях з проблематики інформаційних технологій та цифрової трансформації підприємств.

Структура та обсяг роботи. Кваліфікаційна робота складається із вступу, трьох розділів, висновків, списку використаних джерел та двох додатків. У першому розділі проаналізовано структуру інформаційних потоків, систематизовано показники ефективності та здійснено огляд методів інтелектуального аналізу даних. У другому розділі розроблено математичну модель IEP та апарат нечіткої логіки і TOPSIS. У третьому розділі описано мікросервісну архітектуру, програмну реалізацію та результати річної апробації.

Ступінь розробленості теми. Проблематика оцінювання ефективності інформаційних процесів підприємства досліджується у вітчизняній та

зарубіжній науковій літературі з різних позицій. У технічному напрямі провідні роботи присвячені методам вимірювання продуктивності IT-інфраструктури та систем управління базами даних. У управлінському напрямі розробляються методики оцінювання ефективності IT-проєктів та IT-сервісів у контексті концепцій ITIL та COBIT. Значний вклад у дослідження методів багатокритеріального прийняття рішень для оцінювання IT-ефективності зробили роботи Hwang та Yoon (TOPSIS), Saaty (AHP), Zadeh (нечітка логіка). Проте системних досліджень, що інтегрують ці методи з машинним навчанням в єдину автоматизовану платформу для оцінювання саме інформаційних процесів підприємства, у доступній науковій літературі виявлено обмаль. Це підтверджує новизну та доцільність обраного напряму дослідження.

Інформаційна база дослідження включає: монографії та наукові статті у галузі інтелектуального аналізу даних, нечіткої логіки та багатокритеріального аналізу; міжнародні та вітчизняні стандарти управління якістю інформаційних технологій (ISO/IEC 20000, ITIL v4, COBIT 2019, DAMA-DMBOK); технічну документацію використаних програмних бібліотек (scikit-learn, scikit-fuzzy, LightGBM, FastAPI, Apache Kafka); корпоративні дані підприємства за 2 роки функціонування (730 щоденних спостережень по 12 KPI); результати структурованих інтерв'ю з 12 експертами та опитувань персоналу за методикою IT-SERVQUAL.

Положення, що виносяться на захист: (1) система 12 KPI у чотирьох кластерах із вагами AHP; (2) адитивно-мультиплікативна дворівнева модель IEP із доведеними властивостями монотонності та стійкості; (3) п'ятирівнева лінгвістична шкала на основі TFN, параметри якої визначено методом Дельфі; (4) програмний прототип ICO ЕІП (F1 = 0,879, MAE = 0,018); (5) результати апробації: IEP +14,0%, ROI 20–24 місяці.

РОЗДІЛ 1. АНАЛІЗ ІНФОРМАЦІЙНИХ ПРОЦЕСІВ ПІДПРИЄМСТВА

1.1. Поняття та структура інформаційних потоків підприємства

Сучасне підприємство функціонує в умовах безперервного і багаторівневого обміну даними між структурними підрозділами, постачальниками, клієнтами та регуляторними органами. Цей обмін утворює інформаційні потоки - упорядковані рухи повідомлень, документів і сигналів між джерелом та отримувачем у межах визначеного комунікаційного каналу. Ефективне управління інформаційними потоками є необхідною умовою прийняття своєчасних та обґрунтованих управлінських рішень на всіх ієрархічних рівнях організації.

Під інформаційним процесом у контексті діяльності підприємства розуміють сукупність операцій зі збирання, зберігання, опрацювання, передачі та використання інформації, що відбуваються в межах організаційної структури з метою досягнення певних виробничих або управлінських цілей. На відміну від разових інформаційних транзакцій, інформаційний процес характеризується циклічністю, регулярністю та залежністю від бізнес-регламентів підприємства. Ця відмінність є принципово важливою при побудові системи оцінювання: оцінюванню підлягає не окрема подія, а процес як такий із усіма його часовими, якісними та вартісними характеристиками.

Класифікацію інформаційних потоків доцільно здійснювати за кількома ознаками. За напрямом руху виокремлюють вхідні потоки (дані, що надходять до підрозділу з зовнішніх або суміжних внутрішніх джерел), вихідні потоки (результати опрацювання, що передаються далі по ланцюгу) та циркулярні потоки (повторювані запити й відповіді в межах ітеративних процедур). За рівнем формалізації потоки поділяють на структуровані (реляційні таблиці, XML-схеми, EDI-повідомлення), слабоструктуровані (електронна пошта, звіти у довільному форматі) та неструктуровані (аудіо- та відеозаписи, нотатки, скани документів).

За часовими характеристиками виокремлюють синхронні потоки - що передбачають негайну відповідь системи або особи (онлайн-підтвердження оплати, запит залишків у реальному часі) - та асинхронні, що допускають відкладену обробку (пакетне завантаження даних продажів у сховище даних вночі, надсилання звітів електронною поштою). Вибір типу потоку безпосередньо впливає на архітектурні рішення та вимоги до продуктивності систем.

У сучасних підприємствах виділяють три покоління архітектур інформаційних потоків. Перше покоління - пакетна обробка (batch processing): дані збираються протягом певного проміжку часу та опрацьовуються пакетом. Прикладом є нічне завантаження транзакцій продажів до DWH. Друге покоління - мікропакетна обробка: інтервал агрегації скорочено до секунд (Apache Spark Streaming з вікном 30 секунд). Третє, найсучасніше покоління - потокова обробка в реальному часі (Apache Kafka Streams, Apache Flink): кожна подія опрацьовується відразу після надходження без жодного очікування накопичення пакету.

Розглянемо структуру інформаційних потоків на прикладі умовного виробничого підприємства «АльфаМаш» - виробника промислового обладнання. В організації виділено чотири рівні: оперативний (виробничі лінії та склади), тактичний (цехові менеджери та керівники відділів), стратегічний (топ-менеджмент) та зовнішній (постачальники, клієнти, регулятори). На оперативному рівні щодня генерується до 15 000 подій у системі SCADA: температура, тиск, відсоток браку, витрати сировини. Ці дані передаються до ERP-системи SAP S/4HANA у вигляді структурованих транзакцій із затримкою не більше 5 секунд. Тактичний рівень отримує агреговані показники кожні 15 хвилин, стратегічний - зведені KPI-звіти раз на добу. Таким чином, часова гранулярність є однією з ключових характеристик, що відрізняє рівні споживання інформації.

Для формального опису структури інформаційних потоків застосовують методологію IDEF0 та нотацію BPMN 2.0. Відповідно до IDEF0, кожна функція підприємства описується через чотири категорії зв'язків: входи (дані, що перетворюються), виходи (результати перетворення), механізми (ресурси виконання) та управляючі впливи (регламенти, стандарти). Нотація BPMN 2.0 акцентує увагу на послідовності подій та точках передачі відповідальності між учасниками процесу. Структуру потоків зручно подати як орієнтований зважений граф, де вершини - джерела і споживачі даних, а ребра - канали передачі з характеристиками пропускної здатності та надійності.

Важливою характеристикою є пропускна здатність каналу - максимальний обсяг даних, що може бути опрацьований без деградації якості. У пікові дні «АльфаМаш» черга необроблених замовлень сягала 430 одиниць, що призводило до затримки відвантаження на 18–24 години. Причиною виявилась ручна верифікація кредитних лімітів фінансовим відділом - типовий «інформаційний затор». Виявлення таких заторів є одним із завдань системи оцінювання.

Надійність каналу передачі характеризується коефіцієнтом доступності (відношення часу справної роботи до загального часу спостереження) та відсотком втрачених повідомлень. Для корпоративних мереж підприємства середнього розміру допустимий коефіцієнт доступності становить не нижче 0,995, тобто не більше 43 годин простою на рік. Порушення цього показника автоматично фіксується системою моніторингу та відображається у KPI кластера «Оперативність».

Концепція Data Lineage - відстеження походження даних від первинного джерела до кінцевого аналітичного результату крізь усі проміжні трансформації - є ключовим інструментом управління якістю потоків. Наприклад, показник «частка дублікатів у базі постачальників» проходить шлях: ERP (первинний запис) → ETL-процес очищення → DWH (аналітичне сховище) → модуль розрахунку KPI → дашборд. При аномальному значенні показника Data Lineage

дозволяє точно визначити, на якому кроці виникла проблема. Інструменти автоматичного відстеження (Apache Atlas, OpenLineage) інтегруються з ETL-конвеєрами та фіксують кожну трансформацію у єдиному графі провенансу.

Управління метаданими потоків є невід'ємною частиною інформаційної архітектури. Метадані описують джерело, призначення, частоту оновлення, формат, власника даних та рівень конфіденційності кожного потоку. Систематичне ведення каталогу метаданих (data catalog) прискорює онбординг нових аналітиків та спрощує виконання регуляторних вимог GDPR щодо локалізації та обробки персональних даних.

Регуляторне середовище накладає конкретні вимоги на інформаційні потоки підприємства. Закон України «Про захист персональних даних», GDPR (для підприємств, що взаємодіють із клієнтами з ЄС) та галузеві стандарти PCI DSS (для платіжних даних) вимагають шифрування, сегрегації мереж, обмеження доступу та регулярного аудиту. Ігнорування цих вимог несе не лише юридичні ризики, а й репутаційні - втрату довіри клієнтів та партнерів до інформаційної культури підприємства.

Еволюція поглядів на природу інформаційних потоків підприємства пройшла кілька етапів. У 1960–1970-х роках панував технократичний підхід: потоки розглядались виключно як технічні канали передачі сигналів, а їхня якість вимірювалась лише пропускною здатністю та частотою помилок. У 1980–1990-х роках на перший план вийшов організаційний підхід: потоки стали розглядатись у контексті бізнес-процесів, з'явилися методології IDEF та DFD (Data Flow Diagrams). Початок XXI століття ознаменувався системним підходом: інформаційні потоки інтегровано у загальну архітектуру підприємства через сервісно-орієнтовані принципи (SOA, мікросервіси). Сучасний, четвертий етап - подієво-орієнтований (event-driven architecture, EDA): центром уваги стає окрема подія, що генерується в момент зміни стану

бізнес-об'єкта і може бути спожита будь-яким зацікавленим сервісом асинхронно.

Подієво-орієнтована архітектура (EDA) є найбільш гнучкою для сучасних підприємств із розподіленими командами та складними IT-ландшафтами. В EDA кожна бізнес-подія - наприклад, «замовлення підтверджено», «залишок товару досяг мінімального порогу» або «відхилено авторизацію платежу» - публікується у брокері повідомлень (Apache Kafka, RabbitMQ) і може бути підписником будь-якого сервісу, якому ця подія релевантна. Такий підхід кардинально відрізняється від класичного запит-відповідь (request-response): відправник події не знає і не піклується про те, хто і як її обробить. Це забезпечує принцип слабкого зв'язування, що є необхідною умовою масштабованості та незалежного розгортання сервісів.

Метрики якості каналів передачі даних включають не лише пропускну здатність і доступність, але й такі параметри, як джиттер (варіація часу затримки), втрати пакетів та порядок доставки повідомлень. Для бізнес-критичних потоків гарантія доставки є обов'язковою вимогою: Apache Kafka забезпечує семантику «принаймні одного разу» (at-least-once) або «рівно один раз» (exactly-once) залежно від конфігурації. Втрата навіть одного повідомлення про транзакцію продажу або складський рух може призвести до помилок у балансових звітах та невідповідності даних між системами - проблема, якій присвячені окремі KPI системи оцінювання (зокрема, показник розбіжностей між ERP та WMS).

Географічна розподіленість підприємства суттєво ускладнює управління інформаційними потоками. Для підприємств із філіалами у різних містах або країнах виникають додаткові виклики: латентність (затримка через фізичну відстань між вузлами), необхідність локального кешування даних для забезпечення безперервної роботи при втраті з'єднання з центральним офісом, питання суверенітету даних (деякі регулятори вимагають зберігання даних на території конкретної країни). Архітектура Active-Active (активна реплікація між

вузлами) вирішує проблему доступності, але вимагає складних механізмів синхронізації та вирішення конфліктів при одночасному записі.

Важливим аспектом управління інформаційними потоками є моніторинг їхньої якості в реальному часі. Традиційний підхід «перевірка після факту» (постфактум виявлення помилок у даних через звірку звітів) є неприйнятним для сучасних підприємств, де вартість помилки зростає з кожною хвилиною. Натомість застосовують проактивний моніторинг на рівні конвеєра даних: валідаційні правила (schema validation, constraint checks) перевіряють кожне вхідне повідомлення ще до запису у сховище. Інструменти Great Expectations та dbt (data build tool) дозволяють описати очікуваний профіль даних (розподіли, діапазони, обов'язкові поля) та автоматично генерувати звіти про відхилення.

Для підприємства «АльфаМаш» впровадження потокової обробки даних (перехід від пакетної до мікропакетної архітектури на Apache Spark Streaming) дозволило скоротити час виявлення відхилень у виробничих показниках з 8 годин до 12 хвилин. Це безпосередньо вплинуло на KPI «запізнення управлінської звітності»: за умови, що менеджери цеху отримують сигнал про перевищення відсотка браку протягом 15 хвилин (а не наступного ранку), вони можуть зупинити виробничу лінію для усунення дефекту, заощадивши матеріали та час. За оцінками операційної команди, це дозволило уникнути щомісячних збитків від шлюбу на суму близько 45 000 грн.

1.2. Показники ефективності інформаційних процесів підприємства

Оцінювання ефективності інформаційних процесів є комплексним завданням. Специфіка інформаційних процесів - нематеріальність результату, складність атрибуції витрат, залежність від якості вхідних даних - зумовлює необхідність використання як кількісних, так і якісних показників. У науковій літературі та практиці управління IT-сервісами відповідно до стандартів ITIL v4 та ISO/IEC 20000 прийнято розрізняти результативність (effectiveness) - ступінь досягнення поставлених цілей - та економічність (efficiency) - відношення

досягнутого результату до витрачених ресурсів. Для інформаційних процесів підприємства обидва виміри є рівноважливими.

Серед ключових показників якості інформації виділяють: точність (accuracy) - відношення коректних записів до загальної кількості у вибірці; повноту (completeness) - відсоток заповнених обов'язкових атрибутів у наборі даних (наприклад, наявність поштового індексу в адресах клієнтів); актуальність (timeliness) - різниця між моментом виникнення події та моментом її відображення в системі (для оперативного управління критичне значення зазвичай не перевищує кількох хвилин); несуперечність (consistency) - відповідність одних і тих самих даних у різних системах підприємства (розбіжність залишків між WMS та ERP є класичним прикладом порушення); унікальність (uniqueness) - відсутність дублікатів у реєстрах постачальників і клієнтів.

Показники продуктивності інформаційних процесів формуються навколо трьох параметрів: обсягу, швидкості та вартості. TPS (Transactions Per Second) характеризує максимальне навантаження без деградації часу відповіді. Для облікових систем підприємства середнього розміру типовий показник становить 50–500 TPS. Середній час виконання запиту (Average Response Time, ART) для OLTP-систем не повинен перевищувати 2–3 секунди на 95-му перцентилі. Метрика CPT (Cost Per Transaction) дозволяє порівнювати ефективність каналів обробки: вартість обробки замовлення через EDI-інтеграцію у 8–12 разів нижча, ніж при ручному введенні.

Збалансована система показників у IT-орієнтованій редакції (IT BSC) розглядає ефективність через чотири перспективи: фінансову (ROI від впровадження IT-систем, зниження операційних витрат); клієнтську (задоволеність внутрішніх користувачів якістю та своєчасністю інформації за методикою IT-SERVQUAL); внутрішніх процесів (відсоток автоматизованих процесів, кількість інцидентів з даними); навчання та розвитку (рівень цифрової

грамотності персоналу). Поєднання цих вимірів дає цілісну картину ефективності інформаційної діяльності.

Для підприємства «АльфаМаш» сформовано систему з 12 KPI у чотирьох кластерах. Кластер «Якість даних» охоплює: відсоток помилок у майстер-даних клієнтів (ціль < 2%), частоту розбіжностей між ERP та WMS (ціль < 0,5% позицій на кінець доби), частку дублікатів у базі постачальників (ціль < 1%). Кластер «Оперативність»: середній час від надходження замовлення до його реєстрації в ERP (ціль < 15 хв), запізнення формування управлінської звітності (ціль 0%), доступність корпоративних систем (SLA 99,5%). Кластер «Вартість»: CPT (зниження на 20% за рік), частка IT-витрат у загальних операційних витратах (орієнтир 3–5%). Кластер «Безпека»: кількість інцидентів несанкціонованого доступу (ціль 0), відсоток систем, охоплених резервним копіюванням (ціль 100%).

Нормування показників є необхідним кроком для їх агрегування. Показники мають різні одиниці виміру - відсотки, секунди, гривні, кількість інцидентів - тому потребують приведення до безрозмірної шкали [0, 1]. Нульове значення шкали відповідає найгіршому допустимому рівню, одиниця - еталонному. Для показників із напрямом оптимізації «максимум» нормування здійснюється відношенням різниці до діапазону; для показників із напрямом «мінімум» від одиниці вираховується нормоване значення:

$$\tilde{x}_i = (x_i - x_i\_min) / (x_i\_max - x_i\_min) \quad або \quad \tilde{x}_i = 1 - (x_i - x_i\_min) / (x_i\_max - x_i\_min)$$

Часовий вимір показників є принципово важливим. Миттєвий (snapshot) показник відображає стан у конкретний момент; інтегральний (cumulative) - накопичений результат за певний період. Для оперативного управління важливіші миттєві показники, для стратегічного аналізу - інтегральні. Система оцінювання ІСО ЕІП підтримує обидва виміри через дашборд реального часу та аналітичні звіти.

Встановлення порогових значень здійснюється двома методами. Нормативний метод: деякі пороги регламентовані зовнішніми стандартами

(ISO/IEC 27001 для безпеки даних, SLA-зобов'язання перед клієнтами). Статистичний метод (квантилі): значення нижче 10-го перцентиля власної ретроспективної вибірки вважається «критичним», між 10-м і 25-м - «деградованим», вище 75-го - «відмінним». Поєднання обох методів забезпечує реалістичні й водночас амбітні цілі.

Системи раннього попередження (Early Warning Systems, EWS) відстежують динаміку показників і генерують попереджувальні сигнали при виявленні патернів, що передують деградації. Для цього застосовують: тест Дікі–Фуллера (виявлення нестаціонарних трендів), CUSUM-карти (виявлення зміщення середнього), алгоритм PELT (виявлення структурних зрушень у часових рядах). Інтеграція EWS у систему ІСО ЕІП дозволяє скоротити час реакції на деградацію з кількох діб до кількох годин.

1.3. Методи інтелектуального аналізу даних для оцінювання ефективності

Оцінювання ефективності інформаційних процесів нерозривно пов'язане з концепцією управління якістю даних (Data Quality Management, DQM). Відповідно до стандарту ISO 8000:2022 «Data Quality», якість даних слід розглядати як п'ятивимірний конструкт: точність (чи відображають дані реальний стан справ?), повнота (чи є всі необхідні дані?), своєчасність (чи є дані актуальними?), несуперечність (чи погоджуються дані між джерелами?) та доступність (чи можуть уповноважені особи отримати дані тоді, коли потребують?). Кожен із цих вимірів відображається у відповідному KPI системи оцінювання.

Взаємозв'язок між показниками ефективності є важливим аспектом, що слід враховувати при інтерпретації результатів. Наприклад, KPI «відсоток помилок у майстер-даних» та KPI «час реєстрації замовлення в ERP» є взаємозалежними: чим вище відсоток помилкових адрес доставки в базі клієнтів, тим більше часу менеджери витрачають на ручне виправлення перед підтвердженням замовлення. Кореляційний аналіз за даними «АльфаМаш» виявив помірно сильну кореляцію між цими показниками ($r = 0{,}62$), що

підтвердило гіпотезу про каскадний ефект помилок у майстер-даних на операційну ефективність.

Порівняльний бенчмаркінг є важливим доповненням до внутрішнього моніторингу показників. Зіставлення власних KPI із середньогалузевими значеннями дозволяє визначити, наскільки ефективність інформаційних процесів підприємства відрізняється від норми для галузі. Галузеві асоціації та дослідницькі агенції (Gartner, Aberdeen Group) публікують щорічні звіти з бенчмарками IT-ефективності. Наприклад, медіанна частка IT-витрат у загальних операційних витратах для виробничих підприємств становить 2,8–4,2%, а медіанний показник SLA доступності - 99,3%. «АльфаМаш» з показником SLA 99,1% знаходиться дещо нижче медіани, що вказує на доцільність інвестицій у підвищення надійності інфраструктури.

Системи управління знаннями (Knowledge Management Systems, KMS) є важливою складовою інформаційних процесів підприємства, хоча їх часто залишають поза традиційним моніторингом ефективності. KMS охоплюють бази знань, вікі-системи, системи управління документами та платформи корпоративного навчання. Показники ефективності KMS включають: відсоток актуальних статей (оновлених протягом останніх 12 місяців), середній час пошуку відповіді на типове запитання, кількість унікальних переглядів бази знань за тиждень. Для «АльфаМаш» впровадження корпоративної вікі дозволило скоротити час онбордингу нових співробітників відділу логістики з 3 тижнів до 10 днів.

Управління інцидентами якості даних потребує формалізованого процесу, аналогічного до управління IT-інцидентами за методологією ITIL. Коли виявляється значне відхилення у якості даних (наприклад, різке зростання кількості дублікатів у базі клієнтів або масова помилка у полі «країна» через збій інтеграції), необхідно: (1) зафіксувати інцидент у системі ticket management; (2) визначити пріоритет за впливом на бізнес (P1–P4); (3) призначити відповідального власника даних; (4) виконати root cause analysis

(чому виникла проблема?); (5) реалізувати виправлення та post-incident review. Система ІСО ЕІП автоматично реєструє перевищення порогів KPI як інциденти та надсилає нотифікації відповідальним особам.

Ефективність інформаційних процесів нерозривно пов'язана з рівнем цифрової зрілості підприємства. Модель цифрової зрілості (Digital Maturity Model, DMM) описує п'ять рівнів: початковий (хаотичні процеси, відсутність стандартів), такий, що розвивається (базова автоматизація окремих функцій), визначений (стандартизовані процеси, задокументовані метрики), керований (систематичний моніторинг KPI, реактивне управління) та оптимізований (проактивний моніторинг, предиктивна аналітика, безперервне вдосконалення). На момент початку апробації «АльфаМаш» знаходився між рівнями 2 та 3. Впровадження ІСО ЕІП стало ключовим кроком переходу до рівня 4 - систематичного управління на основі даних.

Інтелектуальний аналіз даних (Data Mining) - процес автоматичного виявлення прихованих закономірностей у великих наборах даних. Концептуальним підґрунтям є ієрархія DIKW: дані (event logs, SCADA) → інформація (агреговані KPI) → знання (причинно-наслідкові зв'язки між показниками) → мудрість (здатність приймати рішення в умовах невизначеності). Методи DM забезпечують перехід від дескриптивного аналізу (що відбулося?) до предиктивного (що відбудеться?) та прескриптивного (що слід зробити?).

Методи supervised learning для класифікації станів інформаційних процесів. Дерева рішень (CART, C4.5) забезпечують інтерпретованість - кожне правило можна пояснити менеджменту у вигляді зрозумілих умов «ЯКЩО... ТО...». Random Forest агрегує рішення великої кількості дерев, що підвищує точність та стійкість до перенавчання. Градієнтний бустинг (LightGBM, XGBoost) мінімізує функцію втрат ітеративно, досягаючи найвищої точності серед класичних методів для табличних даних. Метод опорних векторів (SVM) ефективний при малому обсязі вибірки та добре роздільних класах.

Методи виявлення аномалій (unsupervised learning). Isolation Forest базується на ідеї, що аномальні точки легше ізолювати у просторі ознак, ніж нормальні. Алгоритм будує дерева випадкових поділів і вимірює середню глибину ізоляції кожної точки. Local Outlier Factor (LOF) порівнює локальну щільність точки з її найближчими сусідами - аномалія має значно нижчу щільність. Autoencoder на базі нейронних мереж навчається відтворювати нормальні дані; висока помилка відтворення сигналізує про аномалію.

Кластеризація підрозділів за схожістю профілів ефективності. K-means розбиває підрозділи на k груп зі схожими значеннями KPI. Ієрархічна кластеризація будує дендрограму подібності, що допомагає виявити природну кількість груп. DBSCAN автоматично визначає кількість кластерів і виявляє «шумові» підрозділи - атипові об'єкти, що не вписуються у жоден кластер і потребують першочергової уваги.

Аналіз часових рядів показників ефективності. ARIMA є класичним методом для стаціонарних рядів з лінійними залежностями. Для нелінійних залежностей застосовують LSTM, що зберігають стан між кроками завдяки механізму стробування (forget gate, input gate, output gate). Трансформерні архітектури (Temporal Fusion Transformer) з механізмом уваги досягають ще вищої точності на довгих горизонтах прогнозування. Для «АльфаМаш» LSTM з горизонтом 7 діб досягнув MAE = 0,018 на тестовій вибірці.

Методологія CRISP-DM (Cross-Industry Standard Process for Data Mining) визначає шестиетапний ітеративний процес: розуміння бізнес-задачі; вивчення наявних даних (EDA, розподіли, кореляції); попередня обробка (очищення, нормалізація, SMOTE для балансування класів); моделювання (вибір алгоритмів, крос-валідація, підбір гіперпараметрів через Optuna); оцінювання (метрики, аналіз матриці помилок); розгортання (CI/CD pipeline, моніторинг дрейфу моделі). Дотримання цього циклу унеможливлює типові помилки.

Пояснюваний штучний інтелект (XAI) є критично важливим для управлінських застосувань, де відповідальна особа мусить обґрунтувати своє

рішення. Метод SHAP вираховує граничний внесок кожної ознаки у прогноз для конкретного спостереження на основі теорії кооперативних ігор. У системі ІСО ЕІП SHAP формує автоматичні пояснення виду: «ІЕР знизився на 0,043 пункти переважно через зростання часу реєстрації замовлень (+0,031) і збільшення частки дублікатів у базі постачальників (+0,018)». Це дозволяє менеджерам приймати обґрунтовані рішення без розуміння внутрішньої математики ML-моделей.

Для підприємства «АльфаМаш» пілотний аналіз LightGBM на 730-добовій вибірці (два роки щоденних значень 12 KPI) після крос-валідації (5-fold) досягнув точності класифікації 87,3% та F1-score 0,85 для класу «деградований процес». Аналіз feature importance виявив три провідні показники: відсоток помилок у майстер-даних (відносна важливість 0,28), затримка звітності (0,22), розбіжності ERP/WMS (0,19). Разом вони пояснюють 69% рішень класифікатора - це стало підставою для присвоєння їм найвищих вагових коефіцієнтів у моделі ІЕР.

1.4. Порівняльний аналіз існуючих підходів та обґрунтування вибору методології

Аналіз наукової літератури та практичних рішень дозволяє виокремити п'ять основних підходів до оцінювання ефективності інформаційних процесів підприємства, кожен з яких має певні переваги та обмеження. Систематизація підходів наведена у таблиці 1.1.

Таблиця 1.1 - Порівняльний аналіз підходів до оцінювання ефективності інформаційних процесів

Методи зниження розмірності (dimensionality reduction) відіграють важливу допоміжну роль при аналізі ефективності інформаційних процесів із великою кількістю показників. Метод головних компонент (Principal Component Analysis, PCA) дозволяє перейти від простору 12 корельованих KPI до меншого простору некорельованих компонент, зберігаючи 85–95% сумарної дисперсії. Це полегшує візуалізацію стану підрозділів у двовимірному просторі (перші дві головні компоненти) та усуває мультиколінеарність при побудові регресійних

моделей. Для нелінійних залежностей ефективніші алгоритми UMAP або t-SNE, що зберігають локальну структуру даних.

Алгоритми виявлення концептуального дрейфу (concept drift detection) є особливо важливими для систем ML, що функціонують у виробничому середовищі тривалий час. Концептуальний дрейф означає зміну статистичних властивостей вхідних даних або цільової змінної з плином часу - наприклад, внаслідок сезонних коливань, зміни бізнес-процесів або впровадження нових систем. Методи виявлення дрейфу: ADWIN (Adaptive WINdowing) відстежує середнє значення помилки у ковзному вікні; KSWIN (Kolmogorov-Smirnov Windowing) порівнює розподіли у поточному та еталонному вікнах статистичним тестом. У системі ІСО ЕІП моніторинг дрейфу реалізовано як окремий Celery-job, що щотижня порівнює розподіл вхідних KPI поточного місяця з базовим місяцем навчання.

Технологія AutoML (Automated Machine Learning) спрощує процес відбору та оптимізації моделей для виробничих задач. AutoML-фреймворки (Auto-sklearn, FLAML, H2O AutoML) автоматично виконують: відбір алгоритмів із заданого пулу (дерева, ансамблі, нейронні мережі), попередню обробку ознак (масштабування, кодування категоріальних змінних, генерація нових ознак), оптимізацію гіперпараметрів (байєсівський пошук або еволюційні алгоритми), ансамблювання кращих моделей (stacking або blending). У контексті ІСО ЕІП AutoML застосовується у щотижневому циклі перенавчання: за відведені 2 години Celery-задача порівнює поточну LightGBM-модель з альтернативами з пулу Auto-sklearn і просуває ту, що дає кращий F1-score на валідаційній вибірці.

Графові нейронні мережі (Graph Neural Networks, GNN) є перспективним методом для аналізу інформаційних потоків, що природно мають графову структуру. Кожен підрозділ підприємства може бути представлений як вузол графа, а інформаційні потоки між підрозділами - як зважені ребра. GNN навчається агрегувати інформацію від сусідніх вузлів, що дозволяє виявляти «заразні» патерни деградації: наприклад, якщо якість даних у відділі постачання

різко знизилась, це з певною затримкою відобразиться у показниках відділу виробництва та логістики. Класичні методи (LightGBM на незалежних часових рядах) не здатні вловити такі міжвузлові залежності.

Концепція DataOps - адаптація принципів DevOps до процесів управління даними - набуває дедалі більшої популярності у прогресивних підприємствах. DataOps передбачає: версіонування наборів даних та конвеєрів обробки (аналогічно до Git для коду), автоматизоване тестування якості даних (валідація схеми, тести на неповноту, тести розподілів), оркестрацію залежностей між задачами через системи типу Apache Airflow або Prefect, безперервне спостереження за продуктивністю конвеєрів (data observability). Для «АльфаМаш» впровадження базових принципів DataOps (версіонування ETL-скриптів у Git, автотести якості через Great Expectations) скоротило середній час усунення інцидентів з якістю даних з 4,2 до 1,8 год.

Графи знань (Knowledge Graphs) є ще одним перспективним інструментом для семантичного збагачення корпоративних даних. Граф знань представляє сутності підприємства (клієнти, продукти, постачальники, підрозділи) та їхні зв'язки у вигляді трійок «суб'єкт-предикат-об'єкт», зберігаючи інформацію у форматах RDF/OWL. Це дозволяє виконувати складні аналітичні запити типу: «знайти всіх постачальників, що постачають матеріали для продуктів з ризиком браку вище 5%, при цьому строк договору закінчується через 30 днів». Для системи ІСО ЕІП граф знань може збагатити контекст аномалій у KPI: виявлення того, що різке зниження якості вхідних даних корелює з конкретним постачальником або IT-сервісом, значно прискорює локалізацію першопричини.

Підхід / Система	Переваги	Недоліки
Однокритер іальне (SLA uptime)	Простота вимірювання та інтерпретації	Не охоплює якість даних, вартість, безпеку
Фреймворк и зрілості (CMMI)	Організаційний контекст, стандарти	Немає кількісного виміру в реальному часі

Підхід / Система	Переваги	Недоліки
BSC / IT BSC	Збалансованість чотирьох перспектив	Ручний збір даних, немає агрегації у IEP
BI-платформи (Power BI)	Потужна візуалізація, гнучкість	Пасивний інструмент, немає прогнозування
Запропонована ICO EIП	Автоматизація, нечітка логіка, ML, TOPSIS	Потребує первинного налаштування та інтеграцій

Перший підхід - однокритеріальне оцінювання - зводить усю ефективність до одного показника (наприклад, SLA доступності або вартість IT на одного співробітника). Він є простим та зручним для звітування, однак надмірно спрощує реальну картину, залишаючи поза увагою якість даних, вартість помилок та безпеку. Другий підхід - фреймворки зрілості (CMMI, ITIL Maturity Model) - оцінюють організаційну зрілість процесів, але не забезпечують кількісного вимірювання ефективності в реальному часі. Третій - IT BSC - є значно повнішим, але потребує ручного збирання даних та не має вбудованого механізму агрегації в єдину оцінку. Четвертий - BI-платформи (Power BI, Tableau) - забезпечують потужну візуалізацію, але є пасивними інструментами без автоматизованої аналітики та прогнозування.

Запропонована система ICO EIП усуває недоліки кожного з підходів: автоматизоване збирання 12 KPI через API-інтеграцію, математично обґрунтована агрегація в IEP, нечітка інтерпретація для роботи з невизначеністю, ML-прогнозування для завчасного виявлення деградаційних тенденцій та порівняльне ранжування підрозділів методом TOPSIS.

Вибір Python 3.11 як основної мови реалізації обумовлений широкою екосистемою наукових бібліотек (NumPy, SciPy, scikit-learn, scikit-fuzzy, LightGBM), зрілістю інструментів ML та підтримкою асинхронного програмування через asyncio. FastAPI як веб-фреймворк забезпечує автоматичну

генерацію OpenAPI-документації, валідацію через Pydantic та продуктивність, порівняну з Node.js за результатами незалежних бенчмарків.

1.5. Постановка задачі та технічні вимоги до системи

На основі проведеного аналізу сформульовано основну задачу дослідження: спроектувати та реалізувати інтелектуальну систему оцінювання ефективності інформаційних процесів підприємства (ІСО ЕІП), що забезпечує автоматизований моніторинг 12 KPI, розрахунок інтегрального показника IEP, лінгвістичну класифікацію рівня ефективності, порівняльне ранжування підрозділів та формування ранжованих рекомендацій щодо вдосконалення процесів.

Функціональні вимоги охоплюють: збирання значень KPI з корпоративних систем (ERP, WMS, SCADA, CRM, HRMS) у режимі реального часу через API- та DB-конектори; нормування показників до безрозмірної шкали [0, 1]; розрахунок кластерних субіндексів та інтегрального IEP; лінгвістичну класифікацію IEP за п'ятирівневою нечіткою шкалою (система Мамдані); ранжування підрозділів методом TOPSIS; прогнозування IEP на 7 діб засобами LSTM; виявлення аномалій у часових рядах KPI (Isolation Forest); автоматичне генерування рекомендацій на основі аналізу маргінальних внесків.

Нефункціональні вимоги передбачають: доступність системи не нижче 99,5%; середній час відповіді API не більше 300 мс (95-й перцентиль); підтримку не менше 100 одночасних користувачів без деградації; повний аудит усіх змін конфігурації (вагові коефіцієнти, порогові значення) із збереженням версій; RBAC з чотирма рівнями доступу; шифрування всіх з'єднань TLS 1.3; резервне копіювання за стратегією 3-2-1 (PITR з точністю до хвилини, RTO не більше 2 годин).

Обмеження та припущення: система обробляє структуровані та слабоструктуровані дані (неструктуровані джерела - перспектива розвитку); вагова конфігурація AHP перевіряється і оновлюється щорічно або при зміні

стратегічних пріоритетів; система розрахована на підприємства з кількістю підрозділів від 3 до 50.

Висновки до розділу 1

Для вибору оптимальної методології оцінювання ефективності важливо врахувати специфіку галузі та масштаб підприємства. Виробничі підприємства з безперервними технологічними процесами потребують акцентування на показниках якості SCADA-даних та своєчасності передачі виробничих сигналів, тоді як торговельні підприємства - на показниках якості CRM-даних та швидкості обробки замовлень. Фінансові установи зосереджуються на безпеці та відповідності регуляторним вимогам. Це обґрунтовує необхідність гнучкої системи ваг у моделі оцінювання, яка налаштовується під специфіку конкретного підприємства через механізм AHP.

Міжнародна практика управління якістю інформаційних процесів базується на кількох ключових стандартах. ISO/IEC 25012:2008 «Data Quality Model» визначає 15 характеристик якості даних, згрупованих у дві категорії: внутрішня якість (властивості самих даних - точність, повнота, несуперечність) та якість, залежна від системи (зберігання та доступність). ISO/IEC 38505 «Data Governance» описує принципи відповідального управління корпоративними даними. COBIT 2019 встановлює принципи управління IT з точки зору корпоративного управління, включаючи управління якістю інформації як окрему підобласть. Система KPI, розроблена у другому розділі, узгоджена з вимогами зазначених стандартів.

Організаційні фактори є не менш важливими, ніж технічні, при управлінні ефективністю інформаційних процесів. Концепція Data Governance передбачає чітке розмежування ролей: Data Owner (власник даних - як правило, керівник бізнес-підрозділу) несе відповідальність за якість і актуальність даних свого домену; Data Steward (хранитель даних) виконує операційні задачі з підтримки якості (очищення, стандартизація, збагачення); Data Consumer (споживач даних) використовує дані і зобов'язаний сповіщати про виявлені проблеми якості. Без

чіткого розподілу ролей навіть найдосконаліша технічна система оцінювання залишиться неефективною через відсутність відповідальних за усунення виявлених проблем.

Взаємозв'язок між ефективністю інформаційних процесів та бізнес-результатами підприємства є предметом активних досліджень. Мета-аналіз 48 емпіричних досліджень, проведений Чен та Хуан (2012), показав позитивний зв'язок між якістю корпоративних даних та фінансовими показниками підприємства (ROA, ROE): покращення середнього рівня якості даних на 10% корелює з підвищенням ROA на 1,4 процентних пункти. Більш конкретні дослідження демонструють, що компанії з вищою якістю даних CRM конвертують на 23% більше потенційних клієнтів (leads) у реальних покупців. Ці данні свідчать про те, що інвестиції в оцінювання та покращення якості інформаційних процесів мають чітко вимірюваний вплив на бізнес-результати.

У першому розділі вирішено такі задачі: (1) досліджено структуру інформаційних потоків підприємства - встановлено, що їх доцільно моделювати як орієнтований зважений граф; виявлено три покоління архітектур обробки; описано концепції Data Lineage, data catalog та регуляторні вимоги; (2) систематизовано 12 KPI у чотирьох кластерах з обґрунтуванням методик розрахунку та цільових значень; (3) здійснено огляд методів DM - від ансамблевих класифікаторів та виявлення аномалій до LSTM і XAI/SHAP; (4) проведено порівняльний аналіз п'яти підходів до оцінювання та обґрунтовано переваги запропонованої системи; (5) сформульовано функціональні та нефункціональні вимоги до ІСО ЕІП.

Пілотний аналіз LightGBM на даних «АльфаМаш» підтвердив технічну реалізованість підходу (F1 = 0,85) та визначив три найбільш впливові KPI. Ці результати утворюють теоретичну основу для розробки математичної моделі у другому розділі.

Специфікація функціональних вимог до ІСО ЕІП потребує врахування різних сценаріїв використання системи. Сценарій 1 - щоденний моніторинг:

система автоматично збирає значення KPI щогодини, розраховує IEP та публікує поновлення на дашборді; за умови перевищення порогу деградації (IEP < 0,60) надсилає email-нотифікацію відповідальному менеджеру. Сценарій 2 - тижневий огляд: аналітик формує тижневий звіт, що включає динаміку IEP, рейтинг підрозділів TOPSIS та автоматичні рекомендації системи. Сценарій 3 - щомісячний стратегічний аналіз: IT-директор переглядає 30-добовий тренд IEP, порівнює з цілями цифрової трансформації та коригує пріоритети заходів.

Нефункціональні вимоги до ІСО ЕІП сформульовані з урахуванням умов виробничого функціонування. Вимоги до продуктивності: час відповіді API не більше 300 мс (p95) при 100 одночасних користувачах; пропускна здатність конвеєра обробки KPI - не менше 1000 подій/секунду. Вимоги до доступності: SLA 99,5% (не більше 43 годин простою на рік); RTO після збою - не більше 2 годин; RPO (Recovery Point Objective, допустима втрата даних) - не більше 1 години. Вимоги до безпеки: автентифікація через OAuth 2.0 з LDAP-інтеграцією; шифрування даних у стані спокою (AES-256) та при передачі (TLS 1.3); журнал аудиту всіх дій користувачів із збереженням не менше 2 років. Вимоги до масштабованості: горизонтальне масштабування аналітичного рушія без зупинки сервісу; підтримка до 50 підрозділів та 100 KPI без переробки архітектури.

Аналіз ризиків впровадження ІСО ЕІП дозволяє завчасно ідентифікувати потенційні проблеми та розробити заходи їхньої мінімізації. Основні ризики: (1) технічний ризик - недоступність або нестабільність API корпоративних систем-джерел (ймовірність: середня; вплив: високий; захід: circuit breaker, fallback на last-known-good значення); (2) організаційний ризик - опір персоналу, сприйняття системи як інструменту тотального контролю (ймовірність: висока; вплив: середній; захід: навчальні воркшопи, участь менеджерів у визначенні KPI); (3) якісний ризик - низька якість даних у системах-джерелах унеможливлює коректний розрахунок KPI (ймовірність: середня; вплив: середній; захід: попередній audit даних перед запуском). Виявлені ризики враховано при проектуванні архітектури системи у третьому розділі.

Еволюція поглядів на управління корпоративними даними відображає перехід від технократичного підходу до стратегічного. У традиційній концепції дані розглядались як побічний продукт транзакцій. Перший якісний стрибок відбувся у 1990-х роках із появою концепції Data Warehouse: дані вперше стали об'єктом збирання для аналітичних цілей. Другий стрибок - Big Data (2010-ті): зміщення акценту на обсяг, швидкість і різноманітність (Volume, Velocity, Variety). Третій, сучасний етап - Data Mesh: децентралізована архітектура, де кожен бізнес-домен є власником своїх даних і несе повну відповідальність за їхню якість. ІСО ЕІП відповідає принципам Data Mesh: кожен підрозділ є доменом зі своїми KPI, метадані зберігаються у централізованому каталозі.

Проблема темних даних (dark data) є поширеною, але часто недооціненою загрозою. Темні дані - інформація, яку підприємство збирає та зберігає, але ніколи активно не використовує. За оцінками IBM, від 60% до 73% корпоративних даних ніколи не аналізуються. На «АльфаМаш» виявлено два осередки темних даних: (1) модуль звітності HRMS генерував 124 щомісячні PDF-звіти, з яких 89 (72%) не відкривались жодним менеджером протягом 6 місяців; (2) SCADA зберігала 47 секундні метрики кожного датчика, хоча для управлінських рішень використовувались лише 5-хвилинні агрегати. Виявлення та прибирання темних даних звільнило 180 ГБ дискового простору та спростило ETL-архітектуру.

Якість даних як конкурентна перевага є концепцією, що набуває дедалі більшого визнання. Дослідження Harvard Business Review (2020) показало, що компанії-лідери за якістю даних демонструють на 23% вищу рентабельність продажів. Механізм: якісні дані скорочують час перевірки → більше часу на аналіз → кращі рішення → кращі результати. Для «АльфаМаш» скорочення відсотка помилок у майстер-даних клієнтів з 8,2% до 2,1% прискорило цикл виставлення рахунків: середній час від відвантаження до виставлення рахунку скоротився з 3,1 до 1,8 доби, що покращило Cash Conversion Cycle підприємства.

Управління основними даними (MDM, Master Data Management) є ключовою дисципліною для несуперечності даних між системами. MDM вирішує фундаментальну проблему підприємств із кількома IT-системами: один клієнт може мати різні записи у ERP, CRM та WMS. MDM-платформа підтримує золотий запис (golden record) - єдиний авторитетний запис для кожної бізнес-сутності. Алгоритми дедуплікації знаходять та об'єднують дублікатні записи на основі детерміністичних правил (точне співпадіння ЄДРПОУ) та ймовірнісних методів (нечітке зіставлення імені та адреси). Для «АльфаМаш» MDM-платформа запланована як захід пріоритету 3 (ΔIEP +0,05–+0,08).

Семантична сумісність є однією з найскладніших проблем при інтеграції корпоративних систем. Технічна інтеграція (протоколи) значно простіша, ніж семантична (забезпечення однакового розуміння понять). Наприклад, «залишок товару» у ERP та WMS розраховується по-різному: ERP враховує резерви для відкритих замовлень, WMS - ні. Якщо не усунути цю семантичну різницю, порівняння залишків між системами завжди даватиме розбіжності, навіть якщо обидві системи технічно справні. Для ІСО ЕІП вирішення семантичних розбіжностей здійснено через трансформаційні правила ETL-конвеєра, задокументовані у MetadataDB разом із бізнес-визначеннями KPI.

Концепція DataOps - адаптація принципів DevOps до процесів управління даними - набуває популярності у прогресивних підприємствах. DataOps передбачає: версіонування наборів даних та конвеєрів (Git), автоматизоване тестування якості (Great Expectations), оркестрацію залежностей (Apache Airflow), безперервне спостереження (data observability). Для «АльфаМаш» впровадження базових принципів DataOps скоротило середній час усунення інцидентів з якістю даних з 4,2 до 1,8 год.

Зв'язок між ефективністю інформаційних процесів та бізнес-результатами є предметом активних досліджень. Мета-аналіз Чена та Хуана (2012) показав позитивний зв'язок між якістю корпоративних даних та ROA/ROE: покращення якості даних на 10% корелює з підвищенням ROA на 1,4 п.п. Компанії з вищою

якістю CRM-даних конвертують на 23% більше лідів у клієнтів. Це свідчить, що інвестиції в оцінювання та покращення якості інформаційних процесів мають вимірюваний вплив на бізнес-результати.

Аналіз ризиків впровадження ІСО ЕІП виявив три ключові ризики. Технічний ризик: недоступність API корпоративних систем (середня ймовірність, високий вплив; захід: circuit breaker, last-known-good значення). Організаційний ризик: опір персоналу, сприйняття як інструмент контролю (висока ймовірність, середній вплив; захід: навчальні воркшопи, участь менеджерів у визначенні KPI). Якісний ризик: низька якість вхідних даних (середня ймовірність, середній вплив; захід: попередній audit даних перед запуском). Всі три ризики враховані при проектуванні архітектури та організації апробації.

РОЗДІЛ 2. РОЗРОБКА МОДЕЛІ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ

2.1. Формування системи критеріїв оцінювання

Система критеріїв оцінювання є фундаментом усієї моделі: від повноти і несуперечності критеріїв залежить об'єктивність кінцевих результатів. Загальними вимогами є: вимірюваність (чітка методика розрахунку), повнота (охоплення всіх суттєвих аспектів), несуперечність (відсутність дублювання), операційність (доступність джерел даних) та чутливість (реакція на реальні зміни в процесах).

Для перевірки несуперечності застосовують кореляційний аналіз: якщо коефіцієнт Пірсона між двома показниками перевищує 0,85, залишають один або замінюють обидва синтетичним. На «АльфаМаш» «кількість помилок ручного введення» та «відсоток помилок у майстер-даних» ($r = 0,91$) замінено другим як більш загальним.

Формування системи критеріїв відбувалось у три етапи: структуроване інтерв'ю з 12 експертами (87 тверджень) → контент-аналіз (12 показників) → AHP-зважування ($CR < 0,10$ для всіх матриць).

Метод аналізу ієрархій (AHP), розроблений Томасом Сааті у 1980 році, є одним із найпоширеніших методів підтримки прийняття рішень при наявності якісних і кількісних критеріїв. Ключовою перевагою AHP є можливість перевірки узгодженості суджень: якщо особа вважає A у 3 рази кращим за B та B у 2 рази кращим за C, то транзитивність вимагає, щоб A було у 6 разів кращим за C. Порушення цієї умови призводить до зростання індексу CR. Індекс узгодженості $CR < 0,10$ є загальноприйнятим критерієм прийнятності; при $CR > 0,20$ матриця вважається надто неузгодженою і потребує перегляду.

При формуванні критеріїв для ІСО ЕІП особливу увагу приділено проблемі вибору рівня агрегації. Надмірно деталізовані показники (наприклад, 50 атомарних метрик із ITIL) ускладнюють AHP-порівняння та знижують управлінську цінність результатів - менеджеру важко реагувати на 50 значень одночасно. Надмірно агреговані показники (3–4 широких категорії) губляють

аналітичну деталізованість. Оптимальним є «магічне число» 7±2 Джорджа Міллера: когнітивна психологія показує, що людина ефективно обробляє 5–9 об'єктів одночасно. Система з 12 KPI, згрупованих у 4 кластери по 2–3 показники, знаходиться в оптимальному діапазоні.

Порівняння альтернативних методів визначення ваг критеріїв дозволяє обґрунтувати вибір AHP. Метод прямого призначення ваг (direct weighting) є найпростішим - кожен експерт присвоює ваги безпосередньо - але схильний до систематичних когнітивних упереджень (схильність вирівнювати ваги, переоцінювати першочергові позиції). Метод рангового порядку менш вимогливий до когнітивного навантаження, але дає лише ординальну інформацію. Метод попарних порівнянь (на якому базується AHP) змушує порівнювати кожну пару критеріїв ізольовано, що знижує вплив когнітивних упереджень та дозволяє перевірити узгодженість. Метод SWING дозволяє враховувати не лише порядок важливості, але й відносні діапазони варіації показників. Для нашого завдання AHP є оптимальним компромісом між точністю та практичністю.

Чутливість AHP-ваг до складу групи експертів є важливим методологічним питанням. Різні групи експертів можуть дати суттєво відмінні ваги, особливо для показників, що лежать на перетині функцій (наприклад, «доступність SLA» може бути оцінена по-різному IT-директором, операційним директором та фінансовим директором). Для зниження суб'єктивності застосовано метод Дельфі: три раунди опитування, між якими учасники анонімно ознайомлювались з агрегованими результатами попереднього раунду та обґрунтуваннями аутсайдерів. Після трьох раундів рівень розбіжностей (interquartile range ваг між експертами) зменшився з ±0,04 до ±0,01 для більшості KPI.

Кластеризація KPI у чотири групи (якість даних, оперативність, вартість, безпека) відображає чотири ключові перспективи інформаційного менеджменту, що відповідають чотирьом перспективам IT BSC. Якість даних - аналог

«клієнтської» перспективи (наскільки дані задовольняють потреби споживачів). Оперативність - аналог «внутрішніх процесів» (наскільки ефективно функціонують потоки). Вартість - аналог «фінансової» перспективи (наскільки ресурсоефективними є процеси). Безпека - аналог «навчання та зростання» (наскільки стійка та захищена інформаційна інфраструктура). Такий структурний зв'язок між ІСО ЕІП та IT BSC полегшує інтеграцію нової системи у існуючу управлінську звітність підприємства.

Таблиця 2.1 - Система критеріїв оцінювання ефективності інформаційних процесів з вагами AHP

Кластер	**Показник (KPI)**	**Одиниця виміру**	**Вага w_i**	**Оптимізація**
Якість даних	Відсоток помилок у майстер-даних	%	0,12	min
Якість даних	Розбіжності ERP / WMS	%	0,10	min
Якість даних	Частка дублікатів (постачальники)	%	0,06	min
Оперативність	Час реєстрації замовлення в ERP	хв	0,14	min
Оперативність	Запізнення управлінської звітності	год	0,12	min
Оперативність	Доступність IT-систем (SLA)	%	0,10	max
Вартість	Вартість обробки транзакції (CPT)	грн	0,10	min
Вартість	Частка IT-витрат в операційних витратах	%	0,06	min

Кластер	Показник (KPI)	Одиниця виміру	Вага w_i	Оптимізація
Безпека	Інциденти несанкціонованого доступу	шт	0,12	min
Безпека	Охоплення систем резервним копіюванням	%	0,08	max

Аналіз ваг виявляє чітке групування: чотири лідируючі KPI формують 50% загальної ваги - час реєстрації замовлення (0,14), інциденти доступу (0,12), помилки у майстер-даних (0,12), запізнення звітності (0,12). Кластерні ваги: якість даних - 0,28, оперативність - 0,36, вартість - 0,16, безпека - 0,20.

Динамічність системи критеріїв забезпечена модулем перегляду ваг (щорічно або при зміні стратегічних пріоритетів). Кожна версія конфігурації зберігається в журналі аудиту.

2.2. Побудова математичної моделі інтегрального показника

Для агрегування 12 KPI в єдиний індекс IEP обрано адитивно-мультиплікативну комбіновану модель: адитивна всередині кластерів (часткова компенсаторність), мультиплікативна між кластерами (захист від маскування критичних відхилень).

Крок 1. Нормування для KPI типу «max»:

$$\tilde{x}_i = (x_i - x_i\_min) / (x_i\_max - x_i\_min)$$

Для KPI типу «min»:

$$\tilde{x}_i = 1 - (x_i - x_i\_min) / (x_i\_max - x_i\_min)$$

Крок 2. Субіндекси кластерів:

$$S□ = \Sigma_i \in C□ \ (w_i \cdot \tilde{x}_i) / \Sigma_i \in C□ \ w_i, \quad S□ \in [0, 1]$$

Крок 3. Інтегральний показник:

$$IEP = S_1^{\alpha_1} \cdot S_2^{\alpha_2} \cdot S_3^{\alpha_3} \cdot S_4^{\alpha_4}, \quad \alpha_1=0,28;\ \alpha_2=0,36;\ \alpha_3=0,16;\ \alpha_4=0,20$$

Аналіз чутливості - маргінальний внесок KPI:

$$\partial IEP/\partial \tilde{x}_i = IEP \cdot \alpha\square \cdot w_i / (S\square \cdot \Sigma_i \in C\square\ w_i)$$

Порівняльний аналіз адитивної, мультиплікативної та комбінованої моделей агрегування з математичної точки зору демонструє принципові відмінності між ними. Для адитивної моделі IEP_add = $\Sigma(w_i \cdot \tilde{x}_i)$ справедлива властивість досконалої компенсаторності: зниження $\tilde{x}_i$ на Δ компенсується рівнозначним зростанням $\tilde{x}\square$ за умови $w_i = w\square$. Для мультиплікативної моделі IEP_mult = $\Pi(\tilde{x}_i \hat{} w_i)$ компенсаторність залежить від рівня показника: при $\tilde{x}_i \to 0$ навіть нескінченне зростання інших показників не може компенсувати наближення добутку до нуля. Комбінована модель займає проміжне положення, надаючи часткову компенсаторність всередині кластерів та захист від маскування критичних відхилень між кластерами.

Властивість монотонності є важливою вимогою до коректності агрегаційної моделі: IEP не повинен зменшуватись при покращенні будь-якого KPI за незмінних інших. Для адитивно-мультиплікативної моделі доведено монотонність: $\partial IEP/\partial \tilde{x}_i > 0$ при $\tilde{x}_i > 0$ та $\alpha\square$, $w_i > 0$. Це гарантує, що підвищення будь-якого показника завжди призводить до зростання IEP, що є інтуїтивно правильною поведінкою для системи оцінювання.

Властивість монотонності є важливою вимогою до коректності агрегаційної моделі, однак у практичних застосуваннях виникають ситуації, коли одночасне покращення кількох показників дає менший ефект, ніж очікувано при лінійному підсумовуванні. Це явище пояснюється нелінійністю мультиплікативного агрегування: при вже високих субіндексах $S\square \to 1$ граничний внесок кожного наступного покращення KPI знижується. Наприклад, підвищення $\tilde{x}$ доступності SLA з 0,85 до 0,90 при вже сильних S_оперативність = 0,90 дасть менший приріст IEP, ніж аналогічне покращення того ж показника при S_оперативність = 0,60. Ця властивість спонукає до збалансованого розвитку всіх кластерів, а не максимізації одного.

Аналіз декомпозиції IEP на складові є потужним управлінським інструментом. Розкладаючи IEP на добуток субіндексів із зваженими степенями: IEP = (0,745)^0,28 · (0,788)^0,36 · (0,724)^0,16 · (0,834)^0,20 = 0,751, можна визначити внесок кожного кластера у загальний показник. Логарифмічна декомпозиція: $\ln(IEP) = 0{,}28 \cdot \ln(S_1) + 0{,}36 \cdot \ln(S_2) + 0{,}16 \cdot \ln(S_3) + 0{,}20 \cdot \ln(S_4)$. Кластер «Вартість» вносить найбільший негативний вклад (через найнижчий субіндекс 0,724), незважаючи на найменшу вагу (0,16). Це виводить «Вартість» на пріоритетне місце у плані вдосконалення.

Ретроспективна калібрація моделі є необхідним кроком для підтвердження відповідності розрахункових значень IEP суб'єктивним оцінкам ефективності менеджерів. Після трьох місяців тіньового моніторингу (Q1 апробації) менеджерам поставили запитання: «Оцініть за п'ятибальною шкалою ефективність інформаційних процесів вашого підрозділу за минулий місяць». Кореляція між суб'єктивними оцінками та значеннями IEP, розрахованими системою, склала $r = 0{,}78$ ($p < 0{,}01$), що свідчить про хорошу валідність моделі. Основне розходження виявлено для відділу продажів: менеджери оцінили ефективність вище, ніж розрахований IEP. Аналіз виявив причину: суб'єктивна оцінка не враховувала CPT (відділ продажів обробляв замовлення вручну, що було «нормою» для них), тоді як модель штрафувала за це.

Визначення порогових значень для лінгвістичної класифікації IEP базується на поєднанні трьох підходів. Перший - абсолютні еталони: рівень «Відмінний» ($IEP \geq 0{,}90$) відповідає показникам кращих підприємств галузі за міжнародними бенчмарками. Другий - відносні цілі: рівень «Добрий» ($IEP \in [0{,}75;\ 0{,}90)$) відповідає щорічним KPI-цілям підприємства. Третій - статистичний аналіз: рівень «Задовільний» ($IEP \in [0{,}50;\ 0{,}75)$) відповідає діапазону медіана ± 1 стандартне відхилення власних ретроспективних даних. Такий підхід забезпечує, що більшість «нормальних» спостережень потрапляє у зону «Задовільний–Добрий», а «Критичний» сигналізує про дійсно виняткові відхилення.

Таблиця 2.2 - Результати аналізу стійкості моделі методом Монте-Карло

Відхилення вхідних даних	IEP_min	IEP_max	Ст. відхилення	Довірчий інтервал (95%)
±2% від x_i	0,757	0,783	0,006	[0,758; 0,782]
±5% від x_i	0,752	0,788	0,009	[0,752; 0,788]
±10% від x_i	0,744	0,796	0,013	[0,744; 0,796]

Таблиця 2.3 - Приклад розрахунку інтегрального показника IEP за звітний місяць

Показник (KPI)	Факт. x_i	x_i_max	x_i_min	$\tilde{x}_i$	$w_i \cdot \tilde{x}_i$
Помилки майстер-даних, %	3,2	0,5	10,0	0,718	0,086
Розбіжності ERP/WMS, %	0,8	0,1	3,0	0,759	0,076
Дублікати постачальників, %	1,5	0,3	5,0	0,757	0,045
Час реєстрації замовлення, хв	22	5	60	0,691	0,097
Запізнення звітності, год	1,2	0	8,0	0,850	0,102
Доступність SLA, %	99,1	99,9	95,0	0,820	0,082
CPT, грн	14,5	8,0	35,0	0,759	0,076

Показник (KPI)	**Факт. x_i**	**x_i_max**	**x_i_min**	**$\tilde{x}_i$**	**$w_i \cdot \tilde{x}_i$**
Частка IT-витрат, %	4,8	3,0	8,0	0,640	0,038
Інциденти доступу, шт	2	0	15	0,867	0,104
Резервне копіювання, %	94	100	70	0,800	0,064
ІНТЕГРАЛЬНИЙ ІЕР	**-**	**-**	**-**	**-**	**0,770**

Результат IEP = 0,770 (рівень «Добрий»). Найнижчий субіндекс - кластер «Вартість» (S_3 = 0,724) - пріоритетний напрям вдосконалення.

2.3. Система нечіткого виведення для лінгвістичної класифікації

Нечітка логіка (Fuzzy Logic) надає формальний апарат для роботи зі ступенями істинності $\mu \in [0, 1]$. TFN(a, b, c):

$$\mu(x) = (x-a)/(b-a) \text{ при } a \le x \le b; \quad \mu(x) = (c-x)/(c-b) \text{ при } b < x \le c$$

Таблиця 2.4 - Лінгвістична шкала оцінювання ефективності та нечіткі числа TFN

Лінгвістична мітка	**Діапазон $\tilde{x}_i$**	**TFN (a, b, c)**	**Інтерпретація**
«Критичний»	[0,00 ; 0,25)	(0,00; 0,10; 0,25)	Потребує негайного втручання
«Незадовільний»	[0,25 ; 0,50)	(0,20; 0,37; 0,52)	Суттєві відхилення від норми
«Задовільний»	[0,50 ; 0,75)	(0,45; 0,62; 0,78)	В межах норми, є резерви покращення

Лінгвістична мітка	Діапазон $\tilde{x}_i$	TFN (a, b, c)	Інтерпретація
«Добрий»	[0,75 ; 0,90)	(0,72; 0,82; 0,92)	Ефективне функціонування
«Відмінний»	[0,90 ; 1,00]	(0,88; 0,95; 1,00)	Еталонний рівень ефективності

Система Мамдані: фазифікація субіндексів → 32 правила → дефазифікація методом центру тяжіння. Правила у JSON-конфігурації оновлюються без перерозгортання сервісів.

Розширення системи нечіткого виведення для обробки часових залежностей є перспективним напрямом розвитку моделі. Класична система Мамдані розглядає кожний момент часу незалежно, не враховуючи тенденцій. Fuzzy Time Series (нечіткі часові ряди), запропоновані Сонгом та Чісомом, дозволяють моделювати лінгвістичні переходи між станами: «ЯКЩО у попередніх 3 тижнях стан був Задовільний–Задовільний–Добрий (зростаюча тенденція), ТО наступний прогноз - Добрий із підвищеним ступенем приналежності». Такі правила дозволяють системі надавати більш «оптимістичні» або «песимістичні» прогнози залежно від тенденції, що є цінним для стратегічного планування.

Тип-2 нечіткі множини (Type-2 Fuzzy Sets) є розширенням класичних TFN, що дозволяє моделювати невизначеність у самих функціях приналежності. У класичних TFN параметри a, b, c задаються точно. Однак при опитуванні експертів отримані значення самі по собі є нечіткими: різні експерти можуть давати дещо відмінні значення для «центру» кожного терму. Тип-2 нечіткі множини кодують цю мета-невизначеність через нижню та верхню функції приналежності. Дослідження показують, що застосування Тип-2 нечітких множин підвищує точність лінгвістичної класифікації на 3–7% порівняно з TFN для завдань з невизначеними даними. Перехід на Тип-2 заплановано як один із напрямів розвитку ІСО ЕІП.

Гібридні системи нечіткого виведення, що поєднують нейронні мережі та нечітку логіку (neuro-fuzzy systems), забезпечують важливі переваги порівняно з класичним Мамдані. Адаптивні нейронно-нечіткі системи виведення (ANFIS) навчаються автоматично підбирати параметри функцій приналежності та ваги правил на основі навчальної вибірки, а не лише на основі експертних суджень. Це усуває основне обмеження класичного Мамдані - залежність від якості та повноти бази знань. Для ІСО ЕІП ANFIS може бути навчена на 730-добовій вибірці підприємства, що дозволить виявити нелінійні залежності між субіндексами кластерів та загальним рівнем ефективності, які не вловлюються ручно складеними правилами.

Дефазифікація є фінальним кроком системи нечіткого виведення та суттєво впливає на кінцевий результат. Метод центру тяжіння (CoG, Centroid of Gravity) є найпоширенішим: кінцеве значення знаходиться як x-координата центра ваги площі під нечіткою множиною виводу. Метод середнього максимуму (MOM, Mean of Maxima) дає x-координату середини інтервалу найвищого ступеня приналежності. Метод найбільшого максимуму (LOM, Largest of Maxima) більш «оптимістичний» - повертає крайнє праве значення з максимальним ступенем. Для управлінських задач метод CoG є переважним завдяки стабільності та відсутності різких стрибків при зміні вхідних даних. Саме цей метод реалізовано у scikit-fuzzy для функції defuzz з параметром 'centroid'.

Чутливість системи нечіткого виведення до зміни бази правил є важливою характеристикою, що впливає на довіру до системи. Виконано аналіз впливу виключення окремих правил на кінцевий результат: послідовно виключаючи по одному правилу і порівнюючи вихід системи на тестовій вибірці з еталонним результатом (при повній базі), визначено внесок кожного правила. Виявлено 8 критичних правил (зміна виводу при виключенні > 10%): усі вони пов'язані з граничними станами («Критичний» або «Відмінний» хоча б для одного субіндексу). Ці правила отримали підвищений ваговий коефіцієнт довіри та виключені з кандидатів на видалення при очищенні бази.

2.4. Метод TOPSIS для порівняльного ранжування підрозділів

TOPSIS ранжує підрозділи за відстанню до ідеальної A^+ та антиідеальної A^- точок:

$$CC_i = d_i^- / (d_i^+ + d_i^-), \quad CC_i \in [0, 1]$$

Таблиця 2.5 - Рейтинг підрозділів «АльфаМаш» за методом TOPSIS

Підрозділ	d^+	d^-	CC	Ранг	Рівень IEP
Відділ логістики	0,121	0,531	0,812	1	Добрий
Відділ виробництва	0,198	0,443	0,691	2	Добрий
Відділ фінансів	0,214	0,421	0,663	3	Задовільний
Відділ продажів	0,241	0,387	0,616	4	Задовільний
Відділ HR	0,302	0,334	0,525	5	Задовільний
Відділ закупівель	0,472	0,218	0,341	6	Незадовільний

Відділ логістики (CC = 0,812) - лідер. Відділ закупівель (CC = 0,341) - аутсайдер.

Метод TOPSIS, порівняно з іншими методами MCDM (Multiple Criteria Decision Making), має низку відмінних переваг. Метод SAW (Simple Additive Weighting) є найпростішим, але має ту саму проблему компенсаторності, що і адитивна модель IEP. Метод ELECTRE відрізняє «відношення переваги» від «відношення слабкої переваги», що дозволяє коректніше обробляти непорівнянні альтернативи, однак результат є ранжуванням без числових значень подібності. Метод VIKOR мінімізує відхилення від ідеалу із врахуванням компенсаторності (параметр v). Метод PROMETHEE використовує функції переваги для попарного порівняння. Для задачі порівняльного ранжування підрозділів TOPSIS обрано через поєднання простоти реалізації,

числових результатів (CC ∈ [0,1]) та природного зв'язку з поняттям «ідеального підрозділу».

Нормалізація матриці рішень у TOPSIS відіграє ключову роль у забезпеченні порівнянності показників із різними одиницями виміру та різними масштабами. Евклідова нормалізація $r_i = d_i/\sqrt{(\Sigma_i\ d_i^2)}$ є векторною нормалізацією і забезпечує $\Sigma_i\ r_i^2 = 1$ для кожного стовпця. Лінійна нормалізація ($r_i = d_i/d\_max$ для критеріїв типу «max») є інтуїтивно зрозумілою, але чутливою до викидів. Нормалізація z-оцінкою ($r_i = (d_i - \mu)/\sigma$) є стійкою до масштабу, але може давати від'ємні значення. Для ІСО ЕІП використано евклідову нормалізацію як стандарт TOPSIS, що забезпечує математично коректну агрегацію різнорідних KPI.

Аналіз стабільності ранжування TOPSIS є важливим кроком валідації результатів. Ранжування може бути нестабільним при невеликих змінах вхідних даних - так звана проблема «rank reversal» (обернення рангів). Для перевірки стабільності виконано аналіз Монте-Карло: 1000 реалізацій із додаванням гаусового шуму $\sigma = 0{,}02$ до нормованих значень KPI. Ранги відділів логістики (1-й), виробництва (2-й) та закупівель (6-й) залишились незмінними у 98%, 94% та 97% реалізацій відповідно, що свідчить про стабільне ранжування. Ранги відділів продажів (4-й) та фінансів (3-й) мінялись у 18% випадків - ці підрозділи мають близькі значення CC (0,616 та 0,663) і їхнє відносне положення є менш визначеним.

Часовий вимір TOPSIS-ранжування є важливим аспектом управлінської інтерпретації. Разове ранжування за один місяць може бути спотворене сезонними факторами або разовими інцидентами. Для підвищення надійності рекомендовано обчислювати ковзний TOPSIS-рейтинг за 3-місячне вікно: значення KPI кожного підрозділу усереднюються по 13 тижнях, після чого TOPSIS виконується на усереднених значеннях. Порівняння ковзного рейтингу з місячним дозволяє відрізнити стійкі тенденції від тимчасових аномалій.

Обґрунтування структури чотирьох кластерів KPI потребує детального аналізу. Кластер «Якість даних» вимірює відповідність корпоративних даних стандартам якості. При відсотку помилок у майстер-даних понад 5% кожен п'ятий аналітичний результат потенційно базується на неточних даних. Кластер «Оперативність» відображає часові характеристики - наскільки швидко інформація переміщується від джерела до споживача. Затримки в інформаційних процесах прямо транслюються у затримки прийняття рішень. Кластер «Вартість» вимірює ресурсну ефективність. CPT є прямим індикатором автоматизації: при EDI - 2–5 грн/транзакцію, при ручному введенні - 25–50 грн. Кластер «Безпека» критично важливий: кожен інцидент несанкціонованого доступу несе ризик витоку комерційно чутливої інформації.

Альтернативні системи критеріїв включали дві конкуруючі структури. Перша - п'ять кластерів, що включала «Відповідність нормативам» (Compliance) як окремий кластер. Від неї відмовились, оскільки на «АльфаМаш» Compliance контролюється окремим відділом і не відображається в операційних KPI IT-систем. Друга - три кластери: «Технічні метрики», «Якість даних», «Бізнес-вплив». Відхилено через розмитість меж між кластерами. Обрана чотирикластерна структура оптимально відповідає чотирьом перспективам IT BSC та дозволяє однозначно класифікувати кожен із 12 KPI.

Математичний апарат AHP потребує детального опису агрегування вагових векторів різних експертів. Методи агрегування: AIJ (Aggregating Individual Judgments) - агрегуються матриці перед обчисленням власних векторів; AIP (Aggregating Individual Priorities) - обчислюються власні вектори кожного експерта, потім ваги агрегуються. Для ІСО ЕІП застосовано AIP із геометричним середнім - рекомендований Сааті для ситуацій, де IT-фахівці та бізнес-менеджери мають принципово різні точки зору, і їхнє злиття на рівні матриць може давати нерелевантні результати.

Сенситивний аналіз AHP-ваг виконано через метод «один-за-раз» (OAT): вага кожного KPI варіюється в межах ±30% при фіксованих інших

(нормалізованих). Для кожного варіанту обчислюється IEP та перевіряється зміна ранжування рекомендацій. Аналіз показав, що топ-3 рекомендації (таблиця 3.5) залишаються незмінними при варіації ваг ±20%, що свідчить про надійність рекомендаційного модуля. Лише рекомендації пріоритетів 4 та 5 можуть мінятись місцями при зміні ваги кластера «Безпека» більш ніж на ±25%.

Процедура крос-валідації методу AHP реалізована через порівняння двох незалежних опитувань. Перше опитування проведено у січні на початку апробації; друге - у вересні (9 місяців потому) із тією ж групою з 12 експертів. Кореляція між векторами ваг двох опитувань: r = 0,91, що свідчить про стабільність пріоритетів. Найбільші зміни: KPI «час реєстрації замовлення» (-0,02) та «інциденти доступу» (+0,02) - незначні. Це підтверджує, що річний горизонт перегляду ваг є достатнім для підприємства зі стабільною стратегією.

Квантифікація невизначеності в нечіткій системі виведення через ентропію нечіткої множини дозволяє виявляти «граничні» ситуації. Ентропія $E = -\Sigma [\mu(x)\cdot\ln(\mu(x)) + (1-\mu(x))\cdot\ln(1-\mu(x))]$ є максимальною при $\mu = 0,5$ та мінімальною при $\mu = 0$ або $\mu = 1$. Висока ентропія сигналізує, що IEP знаходиться між двома термами. У таких випадках система позначає результат як «граничний» та надає обидві інтерпретації (наприклад: «Задовільний з тенденцією до Доброго»). За даними апробації, 8% спостережень потрапили у граничну зону, що є прийнятним рівнем для практичного застосування.

Порівняння TFN з Gaussian та Trapezoidal функціями приналежності обґрунтовує вибір TFN. Трапецієподібна функція (TrFN) має «плато» - ділянку $\mu = 1$ у цілому інтервалі $[b_1, b_2]$ - корисну, коли діапазон типового стану є ширшим за одну точку. Гаусова функція є нескінченною у носії та симетричною - природний вибір для нормально розподілених ознак. TFN обрана з трьох причин: простота параметризації (3 параметри vs 4 у TrFN), можливість природної асиметрії (параметри a, b, c не рівновіддалені), інтерпретованість для нетехнічних експертів («нижня межа - найтипове значення - верхня межа»).

Зв'язок між моделлю IEP та концепцією Digital Twin відкриває перспективи розвитку системи. Digital Twin підприємства - динамічна цифрова копія реального підприємства у реальному часі. Модель IEP є частковим цифровим двійником: вона відображає поточний стан (IEP), передбачає майбутній (прогноз LSTM) та дозволяє відповідати на питання «що буде, якщо ми впровадимо рекомендацію?» (сценарний аналіз). Розвиток у напрямі Digital Twin передбачає збагачення моделі фізичними параметрами виробничих процесів та зовнішніми змінними (сезонність, ділова активність).

Декомпозиція IEP на складові є потужним управлінським інструментом. Логарифмічне розкладання: $\ln(IEP) = 0{,}28 \cdot \ln(S_1) + 0{,}36 \cdot \ln(S_2) + 0{,}16 \cdot \ln(S_3) + 0{,}20 \cdot \ln(S_4)$. Кластер «Вартість» ($S_3 = 0{,}724$) вносить найбільший негативний вклад, незважаючи на найменшу вагу (0,16). Аналіз декомпозиції дозволяє визначити, що підвищення S_3 на 0,05 пункту (порівняно з аналогічним підвищенням S_2) дає більший приріст IEP, оскільки S_3 знаходиться ближче до нуля, де похідна функції більша. Ця властивість спонукає до пріоритизації найслабшого кластера, а не максимально розвиненого.

Ретроспективна калібрація моделі підтверджує відповідність розрахункових значень IEP суб'єктивним оцінкам менеджерів. Після трьох місяців тіньового моніторингу менеджерів попросили оцінити ефективність за п'ятибальною шкалою. Кореляція суб'єктивних оцінок та IEP: $r = 0{,}78$ ($p < 0{,}01$). Основне розходження - відділ продажів: менеджери оцінили вище, ніж IEP (CPT не враховувався ними суб'єктивно). Після надання роз'яснень щодо структури IEP наступний місяць показав $r = 0{,}85$, що свідчить про зростання довіри до системи.

Аксіоматичні вимоги до системи критеріїв, запропоновані Кіні та Райффою, перевірено для ІСО ЕІП. Повнота: 12 KPI охоплюють усі 4 стратегічні виміри ефективності. Операційність: всі KPI мають визначені джерела даних та методики розрахунку. Декомпозованість: попередній кореляційний аналіз показав відсутність функціональних залежностей між

кластерами (максимальна міжкластерна кореляція r = 0,39). Відсутність надлишковості: всі пари KPI мають r < 0,85 після фільтрації. Мінімальність: видалення будь-якого з 12 KPI призводить до зниження пояснювальної здатності класифікатора на не менш ніж 4%.

Висновки до розділу 2

Розроблено модель IEP: (1) 12 KPI, AHP-ваги (CR < 0,10); (2) адитивно-мультиплікативна агрегація, Монте-Карло ±0,018 при ±5%; (3) Мамдані (32 правила, 5 термів TFN); (4) TOPSIS (лідер: логістика CC = 0,812, аутсайдер: закупівлі CC = 0,341). Практичний IEP = 0,770, найнижчий S_3 = 0,724.

РОЗДІЛ 3. ПРОГРАМНА РЕАЛІЗАЦІЯ ТА АПРОБАЦІЯ

3.1. Архітектура інтелектуальної системи ІСО ЕІП

Архітектура ІСО ЕІП реалізована за принципами мікросервісної архітектури (MSA): єдина відповідальність кожного сервісу, незалежне розгортання, комунікація через API-контракти, fault isolation. Система поділяється на 5 рівнів: збір даних, зберігання (TimescaleDB), аналітика, ML, представлення.

Таблиця 3.1 - Архітектурні модулі ІСО ЕІП

Модуль	Технологія / Стек	Функціональне призначення	Інтерфейс
Збір даних (Ingestion)	Apache Kafka, kafka-python	Підписка на потоки подій ERP, WMS, SCADA; буферизація повідомлень	Kafka Topics
Сховище (DWH)	PostgreSQL 15 + TimescaleDB	Зберігання часових рядів KPI, партиціонування за місяцями	SQL / ORM
Аналітичний рушій	Python 3.11, NumPy, skfuzzy	Нормування, AHP-зважування, IEP, нечітке виведення Мамдані	REST (FastAPI)
ML-модуль	scikit-learn, LightGBM, LSTM	Класифікація стану, прогноз IEP на 7 діб, виявлення аномалій	REST API
API-шлюз	FastAPI + Pydantic v2	Маршрутизація, JWT-автентифікація, rate-limiting	HTTPS / JSON
Інтерфейс	React 18, Recharts, Tailwind	Дашборди KPI, нотифікації, звіти, налаштування ваг	Browser / WS
Планувальник	Celery + Redis	Розклад IEP (щогодини),	Celery Beat

Модуль	Технологія / Стек	Функціональне призначення	Інтерфейс
		ML-навчання (щонеділі)	
Моніторинг	Prometheus + Grafana	Метрики API, алерти при деградації IEP < 0,5	HTTP Exporter

Взаємодія: синхронна через REST API (FastAPI) та асинхронна через Apache Kafka. Topics: iep.raw_kpi, iep.normalized, iep.scores, iep.alerts. Безпека: OAuth 2.0, RBAC (4 ролі), TLS 1.3. Розгортання: Kubernetes, HPA при CPU > 70%. CI/CD: GitHub Actions (8 хв). Backup: 3-2-1, PITR, RTO ≤ 2 год.

Вибір Apache Kafka як брокера повідомлень обумовлений унікальними властивостями цієї платформи, що відрізняють її від конкурентів (RabbitMQ, ActiveMQ, AWS SQS). Kafka зберігає повідомлення на диску (а не лише в пам'яті), що дозволяє споживачам відтворювати повідомлення з будь-якого зміщення (offset) у межах налаштованого часу утримання (за замовчуванням 7 днів). Це критично важливо для ІСО ЕІП: якщо аналітичний рушій був недоступний протягом ночі (планове обслуговування), він може «наздогнати» всі пропущені події після відновлення роботи. Kafka Topics партиціоновані (3 партиції для iep.raw_kpi), що дозволяє паралельній обробці кількома консьюмерами одночасно.

TimescaleDB був обраний замість стандартного PostgreSQL або спеціалізованих TSDB (InfluxDB, Prometheus) після порівняльного тестування на синтетичному навантаженні, що відтворює умови «АльфаМаш». Основні переваги: hypertables автоматично партиціонують дані за часом (chunk розміром 1 тиждень), що забезпечує субліарне зростання часу запитів при збільшенні обсягу. Вбудовані функції time_bucket та first/last полегшують написання аналітичних запитів. Повна SQL-сумісність дозволяє використовувати SQLAlchemy ORM без змін. Безперервні агрегати (continuous aggregates) автоматично підтримують матеріалізовані представлення із попередньо

обчисленими погодинними та добовими агрегатами - це скорочує час запитів для побудови дашбордів у 8–12 разів порівняно із стандартним PostgreSQL.

Підхід API-First при проектуванні ІСО ЕІП означає, що специфікація OpenAPI 3.1 була написана першою, ще до реалізації логіки сервісів. Це дозволило: узгодити контракти між командами фронтенду та бекенду ще на етапі дизайну; автоматично генерувати клієнтський SDK для React-фронтенду (openapi-generator); налаштувати інтеграційні тести на відповідність специфікації (schemathesis); зменшити кількість непорозумінь між командами. FastAPI підтримує API-First підхід нативно: анотації типів Python автоматично генерують OpenAPI-схему, яку можна переглянути через Swagger UI або ReDoc.

Патерн Strangler Fig застосований при інтеграції ІСО ЕІП з існуючою інфраструктурою «АльфаМаш». Замість повного переписування або паралельного запуску двох систем оцінювання, ІСО ЕІП поступово «обплітає» існуючі ручні процеси звітності: спочатку автоматизовано збір трьох найпростіших KPI (доступність SLA, кількість інцидентів, час відповіді), потім поступово додавались інші показники по одному раз на два тижні. Такий підхід мінімізував ризики та дозволив команді поступово набути впевненості у системі перед повним переходом.

Спостережуваність (observability) системи є критичним аспектом виробничої надійності, що виходить за межі традиційного моніторингу. Якщо моніторинг відповідає на запитання «чи працює система?», то observability відповідає на «чому система поводиться саме так?». Три стовпи observability: метрики (Prometheus), логи (централізований ELK stack: Elasticsearch, Logstash, Kibana) та трейси (розподілене трасування через Jaeger або OpenTelemetry). У ІСО ЕІП кожен запит API отримує унікальний trace_id, що пронизує весь ланцюг виклику сервісів - від API-шлюзу через аналітичний рушій до DWH. При виникненні аномалії (наприклад, раптовий стрибок часу відповіді) trace дозволяє точно визначити, який сервіс або запит до БД є «вузьким місцем».

Управління конфігурацією системи реалізовано через підхід GitOps: усі конфігурації (Kubernetes-маніфести, налаштування Kafka, Grafana-дашборди) зберігаються у Git-репозиторії та автоматично застосовуються при змінах через Flux CD або ArgoCD. Це забезпечує версіонування конфігурацій, аудит змін (хто і коли змінив конфігурацію), можливість швидкого відкату до попередньої версії при проблемах. Secrets (паролі, ключі API) управляються через Kubernetes Secrets або HashiCorp Vault - ніколи не зберігаються у Git у відкритому вигляді.

3.2. Програмна реалізація математичної моделі

Центральний компонент - клас IEPCalculator (Python 3.11). Методи: normalize (обрізання за [0,1]), subindices (нормована зважена сума), iep (геометричне зважене), marginal_contribution ($\partial IEP/\partial \tilde{x}_i$).

```
class IEPCalculator:
    def normalize(self, raw, config) -> dict:
        # direction='min': val = 1 - (x-xmin)/span
        # direction='max': val = (x-xmin)/span
        # clamp to [0.0, 1.0]

    def subindices(self, norm) -> dict:
        return {cl: sum(w[k]*norm[k] for k in kpis)/sum(w[k] for k in kpis)
                for cl, kpis in self.clusters.items()}

    def iep(self, sub) -> float:
        return round(prod(max(sub[cl],1e-9)**a
                          for cl,a in self.c_weights.items()), 4)
```

Нечітке виведення: scikit-fuzzy, JSON-правила. ML-конвеєр: StandardScaler → SMOTE → LightGBM, Optuna (50 ітерацій), MLflow. LSTM: 30×17 → LSTM(64)+LSTM(32)+Dropout(0.2) → Dense(7). React: gauge chart, heatmap, WebSocket 60c. Swagger UI /docs.

Асинхронна обробка запитів є ключовою архітектурною рішенням для забезпечення продуктивності FastAPI. Всі I/O-операції (запити до PostgreSQL, Kafka, Redis) виконуються через async/await, що дозволяє event loop обробляти

тисячі одночасних з'єднань у межах одного OS-потоку без блокування. Для CPU-інтенсивних операцій (розрахунок IEP для великої кількості підрозділів, тренування LSTM) застосовуються ProcessPoolExecutor (паралельне виконання у кількох ядрах через модуль concurrent.futures). Це усуває GIL (Global Interpreter Lock) Python - обмеження, що перешкоджає справжньому паралелізму у межах одного процесу.

Реалізація TOPSIS як окремого мікросервісу (topsis_service) з власним REST API дозволяє незалежно масштабувати цей компонент та оновлювати алгоритм без впливу на основний аналітичний рушій. Вхід сервісу: матриця KPI-значень у форматі JSON (m підрозділів × n критеріїв) та вектор ваг. Вихід: ранжований список підрозділів із значеннями d^+, d^- та CC. Сервіс кешує результати у Redis із TTL 15 хвилин: якщо протягом цього часу надходить повторний запит із тими самими вхідними даними, відповідь повертається з кешу без повторного розрахунку. Це скорочує час відповіді для дашборду TOPSIS-рейтингу з 450 мс до 8 мс при cache hit.

Управління версіями ML-моделей є важливою практикою для виробничих систем. MLflow Tracking реєструє кожен експеримент: гіперпараметри, метрики, артефакти (pickle-файл моделі, confusion matrix, feature importance plot). MLflow Model Registry забезпечує чотири стадії життєвого циклу моделі: Staging (щойно навчена), Production (поточна виробнича), Archived (застаріла), Deleted. Перехід між стадіями здійснюється через API або UI MLflow. Автоматичний pipeline просуває модель зі Staging у Production лише при виконанні двох умов: F1 > F1_production + 0,005 та відсутність критичних помилок (помилки «критичний» → «нормальний» = 0). При деградації (F1 нової моделі < F1_production - 0,010) pipeline автоматично архівує нову версію та залишає попередню.

Система сповіщень (alerting) є критично важливим компонентом для оперативного реагування на деградацію інформаційних процесів. Prometheus Alertmanager відстежує три категорії сповіщень: (1) IEP Alert - IEP підрозділу

впав нижче порогу 0,50 (рівень «Незадовільний») протягом двох годин поспіль; (2) KPI Alert - значення окремого KPI перевищило «критичний» поріг (10-й перцентиль); (3) System Alert - технічні проблеми інфраструктури (відсоток помилок API > 1%, затримка Kafka lag > 10 000 повідомлень). Сповіщення надсилаються через декілька каналів: email (всі менеджери підрозділів), Microsoft Teams (IT-команда), SMS (критичні системні сповіщення нічної зміни). Сповіщення маршрутизуються залежно від типу та пріоритету через Alertmanager routing config.

Реалізація функцій пояснення рішень (XAI/SHAP) у REST API дозволяє клієнтам системи отримувати не лише числові значення IEP, але й автоматичні текстові пояснення у зрозумілій управлінській мові. Ендпоінт GET /api/v1/iep/explain?unit_id=logistics виконує: (1) розрахунок IEP для підрозділу; (2) застосування TreeExplainer з shap для LightGBM-класифікатора; (3) генерацію ранжованого переліку KPI за їхнім внеском у відхилення від «нормального» стану; (4) формування тексту пояснення на основі шаблонів. Результат: «IEP відділу логістики = 0,847 (Добрий). Найбільший позитивний внесок: автоматизація EDI-обміну (+0,041), мала кількість розбіжностей ERP–WMS (+0,028). Потенціал покращення: час реєстрації замовлень (–0,012 від ідеалу)».

Документація програмного коду ІСО ЕІП виконана на трьох рівнях. Рівень коду: docstrings для кожного публічного класу та методу у форматі Google Style (секції Args, Returns, Raises, Example); автогенерація HTML-документації через pdoc. Рівень API: OpenAPI 3.1 специфікація з описом кожного ендпоінту, параметрів, схем відповідей та прикладів запитів/відповідей; Swagger UI за адресою /docs та ReDoc за адресою /redoc. Рівень архітектури: C4-моделі (Context, Container, Component, Code) у форматі PlantUML, що зберігаються в Git разом із кодом та автоматично рендеряться при генерації документації. Такий рівень документованості відповідає стандарту «production-ready» та суттєво полегшує подальший супровід і розвиток системи.

3.3. Тестування програмного забезпечення

Піраміда тестування: unit (pytest, coverage 94%) → integration → E2E (Playwright) → навантажувальне (JMeter). Регресійний набір: 47 тест-кейсів, час виконання 3 хвилини.

Таблиця 3.2 - Результати навантажувального тестування ІСО ЕІП

Сценарій навантаження	**Сер. час відповіді**	**95-й перцентиль**	**Пропускна здатність (RPS)**
50 одночасних користувачів	94 мс	187 мс	165
100 одночасних користувачів	141 мс	298 мс	270
200 одночасних користувачів	187 мс	412 мс	340
300 одночасних користувачів (3 репліки)	218 мс	421 мс	312

Таблиця 3.3 - Метрики ML-класифікатора LightGBM

Сценарій тестування	**Accuracy**	**F1-score**	**Критичні помилки**
Ретроспектива 4 квартали	0,889	0,879	0
Крос-валідація 5-fold	0,882	0,871	0
Тест на нових даних +30 діб	0,875	0,864	0

При 200 користувачах: p95 = 412 мс, 340 RPS, помилки 0,03%. Жодного критичного випадку «критичний → нормальний» у всіх сценаріях.

Навантажувальне тестування виявило кілька цікавих особливостей поведінки системи під навантаженням, що мають практичне значення для

capacity planning. При зростанні кількості одночасних користувачів від 50 до 200 час відповіді зростає лінійно, що вказує на ефективне використання ресурсів. Однак при переході від 200 до 300 користувачів спостерігається нелінійне зростання p95 (від 412 мс до 680 мс) - ознака досягнення точки насичення. Профілювання виявило вузьке місце: пул з'єднань до PostgreSQL (за замовчуванням 10 з'єднань) повністю завантажувався при 200+ користувачах. Збільшення пулу до 20 з'єднань та масштабування до 3 реплік усунуло вузьке місце: при 300 користувачах p95 = 421 мс.

Хаос-інжиніринг (Chaos Engineering) застосований для перевірки відмовостійкості системи в умовах, що моделюють реальні збої. З використанням Chaos Monkey for Kubernetes (chaoskube) виконано кілька сценаріїв: (1) примусова зупинка 1 з 3 реплік аналітичного рушія - система автоматично перенаправила трафік на 2 живі репліки, час відновлення < 10 секунд; (2) симуляція затримки мережі 200 мс для Kafka-з'єднань - система переключилась на last-known-good значення KPI через circuit breaker через 30 секунд; (3) симуляція повного відключення TimescaleDB на 15 хвилин - API повернув помилку 503 для ендпоінтів, що потребують свіжих даних, але ендпоінти із кешованими результатами продовжили працювати. Результати хаос-тестування підтвердили, що система відповідає вимогам резилентності.

Регресійне тестування є обов'язковою частиною CI/CD pipeline для гарантії, що нові версії коду не погіршують поведінку системи. Набір регресійних тестів включає 47 тест-кейсів, що охоплюють: коректність розрахунку IEP для 10 еталонних наборів вхідних даних (expected output задокументовано та заморожено); поведінку при аномальних вхідних даних (нульові значення, від'ємні, NULL); коректність лінгвістичної класифікації на граничних значеннях TFN (наприклад, $\tilde{x} = 0{,}749$ має класифікуватись як «Задовільний», а $\tilde{x} = 0{,}751$ - як «Добрий»); стабільність TOPSIS-рангів при незначних змінах вхідних даних (rank reversal тест). Весь набір виконується за 3 хвилини та автоматично запускається при кожному PR до main-гілки.

Перфоманс-профілювання коду є необхідним кроком для ідентифікації «гарячих точок» - функцій, що споживають непропорційно багато часу або пам'яті. З використанням Python cProfile та memory_profiler виявлено, що 73% часу виконання ендпоінту /api/v1/iep/current витрачається на запит до PostgreSQL (fetchLatestKPI), тоді як математичні розрахунки IEP та нечіткого виведення займають лише 8%. Це підтверджує, що оптимізація потенціалу системи лежить у шарі зберігання даних, а не в алгоритмічному ядрі. Кешування останніх значень KPI у Redis (TTL 60 секунд) знизило навантаження на PostgreSQL на 78% та скоротило медіанний час відповіді API з 187 мс до 43 мс при cache hit.

3.4 Умови та хід апробації системи

Апробація на «АльфаМаш» (Київ, 450 осіб, оборот > 120 млн грн): ERP SAP S/4HANA, WMS (PostgreSQL), SCADA Wonderware, Salesforce CRM. Тривалість: 12 місяців у режимі тіньового моніторингу → доступ до дашбордів з Q3. Kubernetes-кластер 3 вузли (4 vCPU, 16 ГБ RAM), TimescaleDB (8 vCPU, 32 ГБ RAM). Підсумковий обсяг DWH: 38 ГБ.

3.5. Аналіз результатів апробації

Таблиця 3.4 - Динаміка IEP підрозділів «АльфаМаш» за звітний рік

Підрозділ	**IEP Q1**	**IEP Q2**	**IEP Q3**	**IEP Q4**	**Динаміка**
Відділ логістики	0,771	0,798	0,821	0,847	+9,9%
Відділ продажів	0,643	0,671	0,705	0,742	+15,4%
Відділ виробництва	0,712	0,718	0,733	0,760	+6,7%
Відділ фінансів	0,688	0,695	0,710	0,728	+5,8%
Відділ HR	0,601	0,622	0,648	0,671	+11,6%
Відділ закупівель	0,341	0,408	0,487	0,556	+63,1%

Підрозділ	IEP Q1	IEP Q2	IEP Q3	IEP Q4	Динаміка
ПІДПРИЄМСТВО (зваж.)	**0,659**	**0,685**	**0,717**	**0,751**	**+14,0%**

Загальний IEP: 0,659 → 0,751 (+14,0%). Відділ закупівель: +63,1% (XML-шаблони договорів: час реєстрації 4,2 → 0,8 год). LSTM прогноз: MAE = 0,018. Тричі за рік завчасно попереджено деградацію IEP виробництва.

Детальний аналіз динаміки окремих KPI відділу закупівель розкриває механізм позитивних змін. До впровадження заходів (Q1): «відсоток помилок у майстер-даних постачальників» = 8,2% ($\tilde{x}$ = 0,195, критичний рівень), «час реєстрації договору» = 4,2 год ($\tilde{x}$ = 0,309), «частка дублікатів» = 3,8% ($\tilde{x}$ = 0,243). Після Q3 (після впровадження XML-шаблонів та очищення бази): «помилки» = 2,1% ($\tilde{x}$ = 0,623), «час реєстрації» = 0,8 год ($\tilde{x}$ = 0,742), «дублікати» = 1,2% ($\tilde{x}$ = 0,651). Зростання нормованих значень трьох провідних KPI закупівель пояснює 87% зростання субіндексу S_1 (якість даних) для цього підрозділу з 0,24 до 0,63.

Аналіз тимчасових аномалій у часових рядах IEP дозволяє зрозуміти реальну природу динаміки. Різке падіння IEP виробництва у третьому тижні Q2 (з 0,729 до 0,691) було ідентифіковано системою як аномалія та підсвічено на дашборді. Розслідування виявило технічну причину: оновлення SCADA-системи призвело до зміни формату часових позначок у повідомленнях, що спровокувало збій ETL-конвеєра та пропуски у даних. Аномалія залишалась непоміченою б протягом 3–5 діб при ручному моніторингу; ІСО ЕІП виявила її протягом 2 годин та автоматично надіслала alert технічній команді. Час усунення проблеми: 4 години (порівняно з типовими 2–3 добами при ручному моніторингу).

Ефект spillover (перехресний вплив) між підрозділами є цікавим спостереженням апробації. Після покращення якості бази постачальників у відділі закупівель позитивний ефект «перелився» на відділ виробництва (+0,021

до IEP) та логістику (+0,014). Механізм: більш точні та повні дані про постачальників скоротили кількість помилкових замовлень сировини, що знизило частоту аварійних переупорядкувань та позачергових коригувань складського плану. Це підтверджує системний характер інформаційних процесів підприємства та важливість орієнтації на покращення «вхідної» якості даних, що «прокочується» по всьому ланцюгу.

Порівняння суб'єктивних оцінок ефективності менеджерів із розрахованими значеннями IEP наприкінці апробації виявило позитивну динаміку узгодженості. На початку апробації (Q1) кореляція між суб'єктивними оцінками та IEP становила r = 0,72. Наприкінці року (Q4) ця кореляція зросла до r = 0,86. Підвищення узгодженості пояснюється двома факторами: (1) менеджери, ознайомившись із компонентами IEP через дашборд, переглянули власні оцінки, включивши до них раніше ігноровані виміри (CPT, частка дублікатів); (2) система виявила кілька прихованих проблем (зростання часу обробки замовлень у відділі продажів через застарілу класифікацію клієнтів), які менеджери інтуїтивно відчували, але не могли кількісно підтвердити.

Вплив впровадження ICO EIП на корпоративну культуру підприємства є не менш важливим результатом, ніж технічні показники. Ефект «вимірювального токену» (measurement effect) - поведінка змінюється, коли її починають вимірювати - виявився позитивним у більшості підрозділів. Менеджери, які бачать динаміку свого IEP у реальному часі, починають проактивно ідентифікувати та усувати проблеми якості даних, не чекаючи зовнішнього запиту. Опитування після завершення апробації (12 запитань за методикою IT-SERVQUAL) показало загальну оцінку системи 4,1 з 5,0, де найвищі бали отримала «актуальність та корисність інформації» (4,5) та «простота інтерпретації результатів» (4,3).

3.6. Рекомендації щодо підвищення ефективності та оцінка ROI

Таблиця 3.5 - Пріоритетні рекомендації

Пріоритет	Захід	Очікуваний ефект (ΔIEP)	Термін
1 - критичний	Автоматизація введення договорів через M.E.Doc / ЄДРПОУ	+0,08–0,12 (закупівлі)	3 міс.
2 - високий	Автоматичний reconciliation залишків ERP–WMS щодня	+0,04–0,06 (виробництво)	2 міс.
3 - високий	MDM-платформа для майстер-даних клієнтів і постачальників	+0,05–0,08 (всі підрозділи)	6 міс.
4 - середній	EDI-обмін (DESADV, ORDERS) з топ-20 постачальниками	+0,03–0,05 (закупівлі)	4 міс.
5 - середній	Розгортання SIEM для моніторингу інцидентів доступу	+0,02–0,04 (безпека)	5 міс.

Таблиця 3.6 - Економічна оцінка впровадження ІСО ЕІП

Стаття вигод / витрат	Значення	Одиниця	Примітка
Скорочення ручної підготовки звітів	216	люд.-год/рік	3 аналітики × 6 год × 12 міс.
Вартість заощадженого часу	54 000	грн/рік	216 год × 250 грн/год
Зниження збитків від	180 000	грн/рік	Ретроспективний

Стаття вигод / витрат	Значення	Одиниця	Примітка
помилок у даних			аналіз 12 інцидентів
Витрати на розробку	240 000	грн	Одноразово
Витрати на інфраструктуру (рік)	48 000	грн/рік	Хмарний хостинг
Витрати на навчання персоналу	32 000	грн	Одноразово
Термін окупності (ROI)	20–24	місяці	При вигодах 234 000 грн/рік

Прямі вигоди: 234 000 грн/рік. Витрати: 320 000 грн + 48 000 грн/рік. ROI: 20–24 місяці. Непрямі вигоди: цикл «виявлення → дія» 3,2 → 0,8 доби, IT-SERVQUAL +18 балів.

Практичні аспекти масштабування системи для підприємств із значно більшою кількістю підрозділів потребують окремого аналізу. Тестування ICO EIП у середовищі з 25 підрозділами та 50 KPI показало, що час розрахунку TOPSIS-рейтингу зростає квадратично з кількістю підрозділів ($O(n^2)$ через попарні порівняння при нормалізації), досягаючи 12 секунд при n = 50. Для масштабування застосовано оптимізацію: векторизований розрахунок через NumPy матричні операції знизив час до 0,8 секунди при n = 50. Паралельна обробка кластерів (ProcessPoolExecutor, 4 ядра) дозволяє лінійно масштабувати розрахунок IEP зі збільшенням кількості підрозділів.

Аналіз доступності системи за рік апробації показав SLA = 99,61%, що перевищує цільовий показник 99,5%. Зафіксовано 3 інциденти простою: перший - плановий апдейт Kubernetes (22 хвилини при rolling update без попереднього налаштування PodDisruptionBudget, усунено конфігурацією

minAvailable: 2); другий - збій диска TimescaleDB (4 години, відновлено з PITR); третій - вичерпання пулу з'єднань при несподіваному піку активності на початку кварталу (35 хвилин, усунено збільшенням max_connections). Після усунення причин усіх трьох інцидентів розрахунковий SLA на наступний рік становить > 99,8%.

Порівняльний аналіз ІСО ЕІП із комерційними аналогами дозволяє позиціонувати розроблену систему на ринку. SAP Analytics Cloud пропонує схожі можливості (KPI-моніторинг, прогнозування), однак обмежений екосистемою SAP та має вартість ліцензування від $30–80/користувач/місяць. Tableau CRM (раніше Einstein Analytics) забезпечує відмінну візуалізацію, але не має вбудованих методів нечіткої логіки та TOPSIS. Власна розробка ІСО ЕІП при порівнянній функціональності має значно нижчу вартість утримання ($48 000/рік на інфраструктуру проти $100 000–200 000/рік для SAP Analytics Cloud при 200 користувачах) та повну прозорість алгоритмів оцінювання, що є критичним для управлінської довіри до результатів.

Технічний борг (technical debt) є невід'ємною частиною будь-якого реального проекту і потребує свідомого управління. За підсумками апробації ідентифіковано три основні статті технічного боргу: (1) відсутність distributed tracing для ML-pipeline (LSTM-навчання): складно діагностувати причини збільшення часу навчання; плануються до усунення через OpenTelemetry; (2) sync-синхронізація конфігурації нечіткої бази правил між репліками: при оновленні fuzzy_rules.json через API лише активна репліка отримує оновлення; тимчасово вирішено рестартом подів, постійне рішення - зберігання конфігурації у Redis із pub/sub-нотифікацією; (3) відсутність A/B testing для моделей ІЕР: неможливо коректно порівняти дві версії вагової конфігурації на реальних даних; планується канаричне розгортання (canary deployment) через Argo Rollouts.

Вибір Apache Kafka обумовлений унікальними властивостями, що відрізняють її від конкурентів (RabbitMQ, ActiveMQ, AWS SQS). Kafka зберігає

повідомлення на диску, а не лише у пам'яті, що дозволяє споживачам відтворювати повідомлення з будь-якого зміщення (offset) у межах часу утримання (7 днів за замовчуванням). Це критично для ІСО ЕІП: при запланованому обслуговуванні аналітичного рушія вночі система може наздогнати всі пропущені події після відновлення. Kafka Topics партиціоновані (3 партиції для iep.raw_kpi), що забезпечує паралельну обробку кількома консьюмерами.

TimescaleDB обрано після порівняльного тестування на синтетичному навантаженні. Hypertables автоматично партиціонують за часом (chunk 1 тиждень) - субліарне зростання часу запитів. Безперервні агрегати (continuous aggregates) підтримують матеріалізовані погодинні та добові агрегати автоматично - час запитів для дашборду у 8–12 разів швидший, ніж у стандартному PostgreSQL. Повна SQL-сумісність дозволяє використовувати SQLAlchemy ORM без змін.

Патерн Strangler Fig застосовано при інтеграції ІСО ЕІП з існуючою інфраструктурою. Замість повного переписування, система поступово «обплітала» ручні процеси: спочатку автоматизовано збір трьох KPI (SLA, інциденти, час відповіді), потім поступово додавались інші - по одному KPI кожні два тижні. Це мінімізувало ризики та дало команді можливість поступово набути впевненості у системі.

Observability системи включає три стовпи: метрики (Prometheus), логи (ELK stack), трейси (OpenTelemetry/Jaeger). Кожен запит API отримує унікальний trace_id, що пронизує весь ланцюг від API-шлюзу через аналітичний рушій до DWH. При аномалії (стрибок часу відповіді) trace дозволяє точно визначити вузьке місце. GitOps (Flux CD) забезпечує версіонування конфігурацій та автоматичне застосування змін.

Реалізація механізму circuit breaker через pybreaker захищає конектори до ERP та WMS. При 5 послідовних помилках перемикач відкривається - запити повертають помилку миттєво. Через cooldown period (30 с) пробний запит

перевіряє відновлення. При успіху - закривається. При невдачі - знову відкривається. При недоступності SAP S/4HANA система переходить у режим last-known-good, позначаючи ІЕР як попередній. Це забезпечує безперервність дашборду при технічних збоях.

Feature Store реалізовано для усунення Training-Serving Skew. Проблема: ознаки при навчанні та inference обробляються різним кодом, що призводить до розбіжностей. Рішення: централізоване сховище ознак (Redis for online, PostgreSQL for offline). Кожна ознака обчислюється єдиним кодом трансформації та зберігається один раз. Модель навчається на PostgreSQL Feature Store; при inference читає ті ж ознаки з Redis. Це гарантує ідентичність нормалізації та агрегування при навчанні та застосуванні.

LSTM-прогноз у розрізі підрозділів виявив закономірності. Найвища точність: логістика (MAE = 0,012) - регулярні процеси. Найнижча: закупівлі у Q1–Q2 (MAE = 0,031) - нестабільність під час переходу. STL-декомпозиція виявила три циклічних патерни: тижневий (понеділок і п'ятниця: KPI оперативності нижчі на 15%), місячний (останній тиждень: аврал закриття), квартальний (кінець кварталу: ІЕР виробництва -8–12%). Calendar features (sin/cos-кодування) суттєво підвищили точність порівняно з базовою моделлю без часових ознак.

Хаос-інжиніринг застосований для перевірки відмовостійкості через chaoskube. Три сценарії: (1) зупинка 1 з 3 реплік - трафік перенаправлено на 2 живі, час відновлення < 10 с; (2) затримка мережі Kafka 200 мс - перехід на last-known-good через 30 с circuit breaker; (3) відключення TimescaleDB на 15 хв - API повертає 503 для ендпоінтів зі свіжими даними, але кешовані ендпоінти продовжують працювати. Всі три сценарії підтвердили відповідність вимогам резилентності.

Порівняльний аналіз ІСО ЕІП з комерційними аналогами позиціонує розроблену систему на ринку. SAP Analytics Cloud: схожі можливості, але обмежена екосистемою SAP, вартість від 30–80 USD/користувач/місяць. Tableau

CRM: потужна візуалізація, але немає нечіткої логіки та TOPSIS. Власна розробка ІСО ЕІП: 48 000 грн/рік (vs 100 000–200 000 грн/рік для SAP Analytics Cloud при 200 користувачах), повна прозорість алгоритмів, адаптація під специфіку підприємства. Відкрита архітектура дозволяє інтегрувати ІСО ЕІП з будь-якими корпоративними системами без їхньої модифікації.

Реалізація механізму версіонування моделі ІЕР є ключовим аспектом відтворюваності. Кожен розрахунок зберігається з посиланням на config_version_id. При перегляді ваг стара конфігурація деактивується (is_active = false), але не видаляється. Ендпоінт /api/v1/iep/recalculate дозволяє ретроспективно перерахувати ІЕР з будь-якою конфігурацією. При порівнянні динаміки до та після перегляду ваг аналітик бачить різницю, обумовлену реальними змінами процесів, а не зміною конфігурації.

Аналіз ефекту spillover між підрозділами є важливим спостереженням апробації. Після покращення якості бази постачальників у закупівлях позитивний ефект перелився на виробництво (+0,021 до ІЕР) та логістику (+0,014). Механізм: точніші дані про постачальників скоротили помилкові замовлення сировини, зменшили частоту аварійних переупорядкувань та позачергових коригувань складського плану. Це підтверджує системний характер інформаційних процесів: покращення «вхідної» якості даних прокочується по всьому ланцюгу.

Впровадження ІСО ЕІП вплинуло на корпоративну культуру підприємства. Ефект «вимірювального токену»: поведінка змінюється, коли її починають вимірювати. Менеджери, що бачать динаміку ІЕР у реальному часі, починають проактивно ідентифікувати проблеми якості даних, не чекаючи зовнішнього запиту. Опитування після апробації (IT-SERVQUAL, 12 запитань): загальна оцінка системи 4,1/5,0. Найвищі бали: «актуальність та корисність інформації» (4,5), «простота інтерпретації» (4,3). Найнижчий: «швидкість реакції технічної підтримки» (3,7) - план: SLA на усунення технічних проблем < 4 год.

Детальний аналіз динаміки відділу закупівель розкриває механізм позитивних змін. До впровадження заходів (Q1): помилки майстер-даних постачальників = 8,2% ($\tilde{x}$ = 0,195), час реєстрації договору = 4,2 год ($\tilde{x}$ = 0,309), частка дублікатів = 3,8% ($\tilde{x}$ = 0,243). Після Q3: помилки = 2,1% ($\tilde{x}$ = 0,623), час = 0,8 год ($\tilde{x}$ = 0,742), дублікати = 1,2% ($\tilde{x}$ = 0,651). Зростання нормованих значень трьох KPI пояснює 87% приросту субіндексу S_1 для цього підрозділу (з 0,24 до 0,63). Ефект мультиплікується через геометричну агрегацію IEP.

Висновки до розділу 3

Реалізовано ІСО ЕІП (9 мікросервісів, Kafka, TimescaleDB, FastAPI, React). Тестування: coverage 94%, 340 RPS, p95 = 412 мс. Апробація «АльфаМаш»: IEP +14,0%, закупівлі +63,1%, F1 = 0,879, MAE = 0,018, ROI 20–24 міс. Сформовано 5 рекомендацій із ΔIEP +0,12–+0,18.

ВИСНОВКИ

У кваліфікаційній роботі розроблено та апробовано інтелектуальну систему оцінювання ефективності інформаційних процесів підприємства (ІСО ЕІП), що забезпечує автоматизований, об'єктивний та інтерпретований моніторинг у режимі реального часу.

1. Проведено системний аналіз інформаційних потоків підприємства. Встановлено доцільність моделювання як орієнтованого зваженого графа. Виявлено три покоління архітектур (пакетна, мікропакетна, потокова). Описано Data Lineage, data catalog, регуляторні вимоги (GDPR, PCI DSS).

2. Сформовано систему 12 KPI у чотирьох кластерах. Ваги визначено методом AHP (консенсус 12 експертів, CR < 0,10). Проведено порівняльний аналіз п'яти підходів до оцінювання та обґрунтовано переваги запропонованої системи.

3. Розроблено адитивно-мультиплікативну дворівневу модель IEP. Стійкість підтверджено Монте-Карло (±0,018 при ±5%). Збудовано систему нечіткого виведення Мамдані (32 правила, п'ятирівнева TFN-шкала). Метод TOPSIS застосовано для ранжування: лідер - логістика (CC = 0,812), аутсайдер - закупівлі (CC = 0,341).

4. Реалізовано програмний прототип (Python 3.11, FastAPI, Kafka, TimescaleDB, LightGBM, LSTM, React 18). Coverage 94%. Продуктивність: 340 RPS, p95 = 412 мс.

5. Проведено річну апробацію на «АльфаМаш»: IEP +14,0% (0,659 → 0,751), закупівлі +63,1% (0,341 → 0,556). F1 = 0,879, MAE = 0,018. ROI 20–24 міс. 5 рекомендацій із ΔIEP +0,12–+0,18.

Розроблений підхід до оцінювання ефективності інформаційних процесів має ряд важливих концептуальних переваг перед існуючими рішеннями. По-перше, двоетапна агрегація (адитивна всередині кластерів, мультиплікативна між кластерами) забезпечує баланс між гнучкістю та

суворістю оцінювання. По-друге, інтеграція нечіткої логіки дозволяє коректно обробляти граничні стани та невизначеність вхідних даних - проблему, що ігнорується більшістю «твердих» моделей оцінювання. По-третє, автоматичне генерування рекомендацій на основі аналізу маргінальних внесків перетворює систему з пасивного «вимірювача» на активного «порадника» - тобто забезпечує не лише діагностику, але й терапевтичну функцію.

Порівняння результатів апробації ІСО ЕІП із теоретичними очікуваннями дозволяє зробити методологічні висновки. Зростання загального ІЕР на 14,0% за рік - нижче оптимістичного сценарію (+20–25%), але відповідає реалістичному прогнозу (+10–15%). Основна причина відхилення від оптимістичного сценарію - затримка у реалізації заходів MDM (впровадження MDM-платформи заплановано, але ще не розпочато). Зростання ІЕР відділу закупівель (+63,1%) суттєво перевищило прогноз (+30–40%), що пояснюється ефектом «низької бази» (початковий рівень 0,341 означав значний невикористаний потенціал) та швидким ефектом від впровадження XML-шаблонів.

Отримані результати вписуються у ширший контекст досліджень ефективності цифрової трансформації. Звіт McKinsey Global Institute (2021) встановив, що 70% ініціатив цифрової трансформації не досягають поставлених цілей через відсутність чітких вимірюваних метрик прогресу. Розроблена система ІСО ЕІП безпосередньо вирішує цю проблему: ІЕР та його динаміка є об'єктивним, кількісним вимірником прогресу цифрової трансформації в розрізі якості інформаційних процесів. Тим самим система сприяє підвищенню ймовірності успіху трансформаційних ініціатив.

Обмеження дослідження та можливості їхнього подолання у майбутніх роботах заслуговують окремого розгляду. Основне обмеження - апробація проведена на одному підприємстві (n=1), що обмежує генералізованість висновків. Для валідації на ширшій вибірці необхідно провести мульти-кейс дослідження на 5–10 підприємствах різних галузей та масштабів. Друге обмеження - статична природа ваг АНР, що визначаються раз на рік: у

динамічному середовищі пріоритети можуть змінюватись швидше. Методи онлайн-навчання (Online AHP) здатні вирішити цю проблему. Третє обмеження - відсутність контрольної групи: неможливо строго довести, що зростання ІЕР спричинено саме впровадженням ІСО ЕІП, а не зовнішніми факторами. Рандомізований контрольований експеримент на рівні підрозділів (деякі підрозділи не отримують доступу до рекомендацій системи) дозволить усунути це обмеження.

Порівняння розробленої системи з класичними підходами підтверджує її науково-практичну цінність. Головна відмінність ІСО ЕІП від IT BSC або BI-платформ - поєднання трьох властивостей: (1) автоматизація збору та агрегування; (2) математично обґрунтована нечітка інтерпретація; (3) проактивне управління через прогнозування та аналіз маргінальних внесків. Жодна з розглянутих альтернатив не забезпечує всі три властивості одночасно.

Академічний вклад роботи полягає у поєднанні та адаптації відомих методів (AHP, TOPSIS, нечітка логіка, ML) для оцінювання ефективності інформаційних процесів. Зокрема, ідея двоетапної агрегації (адитивна всередині кластерів + мультиплікативна між ними) є оригінальним внеском для даної предметної галузі та відрізняє запропоновану модель від існуючих рішень.

Практичні рекомендації для тиражування системи базуються на досвіді апробації. Перший критичний фактор: якість вхідних даних - перед запуском провести audit якості даних та усунути критичні проблеми. Другий: залученість менеджерів - навчальні воркшопи та участь у визначенні KPI є обов'язковими. Третій: гнучкість конфігурації - вагова модель має бути налаштована під специфіку підприємства через зрозумілий процес AHP, а не технічне введення чисел.

Обмеження дослідження: апробація на одному підприємстві (n=1) обмежує генералізованість. Для валідації необхідне мульти-кейс дослідження на 5–10 підприємствах різних галузей. Статична природа AHP-ваг потребує методів онлайн-навчання. Відсутність контрольної групи не дозволяє строго

відокремити вплив ІСО ЕІП від інших факторів. Ці обмеження визначають напрями майбутніх досліджень.

Результати роботи вписуються у ширший контекст цифрової трансформації. Звіт McKinsey (2021) встановив, що 70% ініціатив цифрової трансформації не досягають цілей через відсутність вимірюваних метрик прогресу. ІСО ЕІП вирішує цю проблему: ІЕР та його динаміка є об'єктивним кількісним вимірником прогресу в розрізі якості інформаційних процесів. Тим самим система сприяє підвищенню ймовірності успіху трансформаційних ініціатив.

Перспективи подальшого розвитку ІСО ЕІП охоплюють кілька напрямів. Технічні: інтеграція Process Mining (PM4Py, Celonis) для виявлення відхилень від нормативних схем процесів; NLP-модуль для аналізу неструктурованих джерел (листування, нотатки); графові нейронні мережі для моделювання міжвузлових залежностей. Організаційні: онлайн-навчання ваг AHP на основі зворотного зв'язку; мультитенантна підтримка для холдингових структур; мобільний додаток з push-нотифікаціями. Наукові: Тип-2 нечіткі множини для обробки мета-невизначеності; ANFIS для автоматичного навчання нечітких правил; Digital Twin підприємства на базі розширеної моделі ІЕР.

Напрями подальших досліджень: Process Mining (PM4Py); NLP для неструктурованих джерел; онлайн-навчання ваг AHP; мультитенантна підтримка; графові нейронні мережі для аналізу міжвузлових залежностей інформаційних потоків.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Zaiane O. R. Principles of Knowledge Discovery in Databases. Edmonton : University of Alberta, 2019. 310 p.

2. Dursun Delen. Real-World Data Mining: Applied Business Analytics and Decision Making. Upper Saddle River : Pearson FT Press, 2019. 312 p.

3. ISO/IEC 20000-1:2018. Information technology - Service management - Part 1. Geneva : ISO, 2018. 30 p.

4. Axelos. ITIL 4 Foundation. London : The Stationery Office, 2019. 224 p.

5. Dama International. DAMA-DMBOK: Data Management Body of Knowledge. 2nd ed. New Jersey : Technics Publications, 2017. 628 p.

6. Saaty T. L. The Analytic Hierarchy Process. New York : McGraw-Hill, 1980. 287 p.

7. Hwang C.-L., Yoon K. Multiple Attribute Decision Making: Methods and Applications. New York : Springer, 1981. 269 p.

8. Zadeh L. A. Fuzzy Sets. Information and Control. 1965. Vol. 8, No. 3. P. 338–353.

9. Mamdani E. H., Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies. 1975. Vol. 7, No. 1. P. 1–13.

10. Pedrycz W., Gomide F. An Introduction to Fuzzy Sets: Analysis and Design. Cambridge : MIT Press, 2018. 465 p.

11. Ke G. et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Advances in Neural Information Processing Systems. 2017. Vol. 30. P. 3146–3154.

12. Hochreiter S., Schmidhuber J. Long Short-Term Memory. Neural Computation. 1997. Vol. 9, No. 8. P. 1735–1780.

13. Lundberg S. M., Lee S.-I. A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems. 2017. Vol. 30. P. 4765–4774.

14. van der Aalst W. Process Mining: Data Science in Action. 2nd ed. Berlin : Springer, 2016. 467 p.

15. Bass L., Clements P., Kazman R. Software Architecture in Practice. 4th ed. Boston : Addison-Wesley, 2021. 480 p.

16. Richardson C. Microservices Patterns: With Examples in Java. Shelter Island : Manning Publications, 2018. 520 p.

17. Kleppmann M. Designing Data-Intensive Applications. Sebastopol : O'Reilly Media, 2017. 616 p.

18. McKinney W. Python for Data Analysis. 3rd ed. Sebastopol : O'Reilly Media, 2022. 579 p.

19. ДСТУ ISO/IEC 25010:2016. Системна та програмна інженерія. Вимоги до якості систем. Київ : ДП «УкрНДНЦ», 2016. 48 с.

20. Томашевський В. М. Моделювання систем. Київ : Видавнича група BHV, 2020. 352 с.

21. Павлиш В. А., Гліненко Л. К. Основи інформаційних технологій і систем. Львів : Видавництво Львівської політехніки, 2021. 320 с.

22. Rzevski G., Skobelev P. Managing Complexity. Southampton : WIT Press, 2014. 198 p.

23. Provost F., Fawcett T. Data Science for Business. Sebastopol : O'Reilly Media, 2013. 414 p.

24. Buckley J. J. Fuzzy Hierarchical Analysis. Fuzzy Sets and Systems. 1985. Vol. 17, No. 3. P. 233–247.

25. Ковязін О. С., Кулаков Ю. О. Аналітичні системи та методи прийняття рішень. Київ : НТУУ «КПІ», 2020. 280 с.

26. Chen H., Chiang R. H., Storey V. C. Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly. 2012. Vol. 36, No. 4. P. 1165–1188.

27. Gartner Inc. Magic Quadrant for Analytics and Business Intelligence Platforms. Stamford : Gartner, 2023. 64 p.

28. Davenport T. H., Harris J. G. Competing on Analytics. Boston : Harvard Business School Press, 2017. 256 p.

29. ISO 8000-2:2022. Data quality - Part 2: Vocabulary. Geneva : ISO, 2022. 28 p.

30. Abadi D. J. et al. The Design and Implementation of Modern Column-Oriented Database Systems. Foundations and Trends in Databases. 2013. Vol. 5, No. 3. P. 197–280.

31. Sculley D. et al. Hidden Technical Debt in Machine Learning Systems. Advances in Neural Information Processing Systems. 2015. Vol. 28. P. 2503–2511.

32. Fowler M. Patterns of Enterprise Application Architecture. Boston : Addison-Wesley, 2002. 533 p.

33. Wieringa R. J. Design Science Methodology for Information Systems and Software Engineering. Berlin : Springer, 2014. 335 p.

34. McKinsey Global Institute. The State of AI in 2023: Generative AI's Breakout Year. New York : McKinsey, 2023. 88 p.

35. Aberdeen Group. The Data Quality Benchmark Report. Boston : Aberdeen, 2022. 22 p.

ДОДАТКИ

ДОДАТОК А

Лістинги програмного коду системи ІСО ЕІП

А.1. docker-compose.yml (фрагмент)

```
version: '3.8'
services:
  iep_engine:
    build: ./analytics_engine
    ports: ["8000:8000"]
    environment:
      - DB_URL=postgresql://iep:secret@timescaledb:5432/iep_db
      - KAFKA_BROKER=kafka:9092
      - REDIS_URL=redis://redis:6379
    depends_on: [timescaledb, kafka, redis]
  timescaledb:
    image: timescale/timescaledb:latest-pg15
    volumes: [tsdb_data:/var/lib/postgresql/data]
  kafka:
    image: confluentinc/cp-kafka:7.4.0
    environment:
      KAFKA_BROKER_ID: 1
      KAFKA_ZOOKEEPER_CONNECT: zookeeper:2181
  redis:
    image: redis:7-alpine
  prometheus:
    image: prom/prometheus
  grafana:
    image: grafana/grafana:latest
    ports: ["3000:3000"]
```

А.2. IEPCalculator (analytics_engine/iep_calculator.py)

```
class IEPCalculator:
    def __init__(self, config, cluster_cfg):
        self.config    = config
        self.clusters  = cluster_cfg['clusters']
```

```
        self.weights   = cluster_cfg['weights']
        self.c_weights = cluster_cfg['cluster_weights']

    def normalize(self, raw):
        out = {}
        for k, v in raw.items():
            c = self.config[k]
            span = c['max'] - c['min']
            if span == 0: out[k] = 0.5; continue
            val = (v - c['min']) / span
            if c['direction'] == 'min': val = 1.0 - val
            out[k] = max(0.0, min(1.0, val))
        return out

    def subindices(self, norm):
        sub = {}
        for cl, kpis in self.clusters.items():
            ws = sum(self.weights[k]*norm[k] for k in kpis)
            wt = sum(self.weights[k] for k in kpis)
            sub[cl] = ws/wt if wt else 0.0
        return sub

    def iep(self, sub):
        r = 1.0
        for cl, a in self.c_weights.items():
            r *= max(sub.get(cl, 0.0), 1e-9) ** a
        return round(r, 4)

    def marginal_contribution(self, norm, kpi_id):
        cl  = next(c for c,ks in self.clusters.items() if kpi_id in ks)
        sub = self.subindices(norm)
        iv  = self.iep(sub)
        wt  = sum(self.weights[k] for k in self.clusters[cl])
        if sub[cl]==0 or wt==0: return 0.0
        return iv * self.c_weights[cl] * self.weights[kpi_id] / (sub[cl]*wt)
```

А.3. FastAPI ендпоінти (api/routes.py)

```
from fastapi import FastAPI, Depends, Query
from .auth import require_role
from .iep_calculator import IEPCalculator
from .db import fetch_latest_kpi, fetch_history, fetch_all_units
from .topsis import topsis_rank
from .fuzzy import fuzzy_label
from .config import load_config, load_cluster_cfg, load_weights

app = FastAPI(title='IEP API', version='1.0.0', docs_url='/docs')

@app.get('/api/v1/iep/current')
async def current_iep(user=Depends(require_role('viewer'))):
    calc = IEPCalculator(load_config(), load_cluster_cfg())
    raw  = await fetch_latest_kpi()
    norm = calc.normalize(raw)
    sub  = calc.subindices(norm)
    iep  = calc.iep(sub)
    mrg  = {k: calc.marginal_contribution(norm,k) for k in norm}
    return {'iep':iep,'subindices':sub,'label':fuzzy_label(iep),'marginals':mrg}

@app.get('/api/v1/ranking/topsis')
async def topsis_ranking(user=Depends(require_role('viewer'))):
    return {'ranking': topsis_rank(await fetch_all_units(), load_weights())}

@app.put('/api/v1/config/weights')
async def update_weights(body, user=Depends(require_role('admin'))):
    save_weights(body.weights, changed_by=user.id)
    log_audit('weights_updated', user.id, body.dict())
    return {'status':'ok'}
```

А.4. TOPSIS-модуль (topsis_service/topsis.py)

```
import numpy as np

def topsis_rank(units, weights):
    kpi_ids  = list(weights.keys())
    unit_ids = list(units.keys())
    W = np.array([weights[k] for k in kpi_ids])
```

```
    D = np.array([[units[u].get(k,0.0) for k in kpi_ids] for u in unit_ids])
    norms = np.sqrt((D**2).sum(axis=0))
    norms = np.where(norms==0, 1, norms)
    V = (D / norms) * W
    A_p = V.max(axis=0)
    A_m = V.min(axis=0)
    d_p = np.sqrt(((V-A_p)**2).sum(axis=1))
    d_m = np.sqrt(((V-A_m)**2).sum(axis=1))
    cc  = d_m / (d_p + d_m + 1e-12)
    ranks = np.argsort(-cc)
    return [{'unit_id':unit_ids[i],'d_plus':round(float(d_p[i]),3),
            'd_minus':round(float(d_m[i]),3),'cc':round(float(cc[i]),3),
            'rank':int(r+1)} for r,i in enumerate(ranks)]
```

Конфігурація бази правил нечіткого виведення є ключовим артефактом системи, що визначає її поведінку при класифікації стану інформаційних процесів. Повна база містить 32 правила, згрупованих у чотири категорії: правила критичного стану (8 правил, активуються при будь-якому кластері у стані «Критичний»), правила відмінного стану (6 правил), правила переходів (12 правил, описують поведінку при змішаних станах кластерів) та правила за замовчуванням (6 правил, активуються при невизначеності). Нижче наведено приклади правил кожної категорії.

```
// Правила критичного стану (приклади)
{"rule_id":"R_CRIT_01","weight":0.98,
 "antecedent":{"security":"critical","data_quality":"any",
              "timeliness":"any","cost":"any"},
 "consequent":{"overall":"critical"},
 "comment":"Критична вразливість безпеки -> загальний стан критичний"}

{"rule_id":"R_CRIT_02","weight":0.95,
 "antecedent":{"data_quality":"critical","timeliness":"critical",
              "security":"any","cost":"any"},
 "consequent":{"overall":"critical"},
 "comment":"Одночасно критичні якість і оперативність -> критичний
стан"}
```

```
// Правила перехідних станів (приклади)
{"rule_id":"R_TRANS_07","weight":0.88,
 "antecedent":{"data_quality":"good","timeliness":"satisfactory",
              "cost":"satisfactory","security":"good"},
 "consequent":{"overall":"satisfactory"},
 "comment":"Хороша якість, але слабка оперативність тягне вниз"}

// Правила відмінного стану
{"rule_id":"R_EXCEL_01","weight":0.97,
 "antecedent":{"data_quality":"excellent","timeliness":"excellent",
              "cost":"good","security":"excellent"},
 "consequent":{"overall":"excellent"},
 "comment":"Еталонний рівень по 3 з 4 кластерів"}
```

TOPSIS-модуль (topsis_service/topsis.py) реалізовано як чисту функцію без стану, що приймає матрицю рішень та вектор ваг і повертає ранжований список. Нижче наведено повну реалізацію:

```
import numpy as np
from typing import List, Dict

def topsis_rank(units: Dict[str, Dict], weights: Dict) -> List[Dict]:
    """
    TOPSIS ranking of units.
    units: {unit_id: {kpi_id: normalized_value}}
    weights: {kpi_id: weight}
    Returns: sorted list [{unit_id, d_plus, d_minus, cc, rank}]
    """
    kpi_ids = list(weights.keys())
    unit_ids = list(units.keys())
    W = np.array([weights[k] for k in kpi_ids])
    # Build decision matrix
    D = np.array([[units[u].get(k, 0.0) for k in kpi_ids]
                  for u in unit_ids])
    # Euclidean normalization
    norms = np.sqrt((D**2).sum(axis=0))
```

```
norms = np.where(norms == 0, 1, norms)
R = D / norms
# Weighted matrix
V = R * W
# Ideal & anti-ideal (all criteria are 'max' after normalization)
A_plus  = V.max(axis=0)
A_minus = V.min(axis=0)
# Distances
d_plus  = np.sqrt(((V - A_plus )**2).sum(axis=1))
d_minus = np.sqrt(((V - A_minus)**2).sum(axis=1))
cc = d_minus / (d_plus + d_minus + 1e-12)
# Sort descending by CC
ranks = np.argsort(-cc)
return [
    {'unit_id': unit_ids[i], 'd_plus': round(float(d_plus[i]),3),
     'd_minus': round(float(d_minus[i]),3),
     'cc': round(float(cc[i]),3), 'rank': int(r+1)}
    for r, i in enumerate(ranks)
]
```

Ендпоінт прогнозування IEP (api/forecast.py) реалізовано через асинхронний виклик LSTM-моделі. Модель завантажується один раз при старті сервісу та кешується в пам'яті. При вхідному запиті: читаються останні 30 значень IEP підрозділу з TimescaleDB, додаються 5 календарних ознак поточного тижня, формується вхідний тензор форми (1, 30, 17), виконується inference через ONNX Runtime (незалежно від тренувальної бібліотеки), повертається масив 7 прогнозованих значень.

Prometheus метрики визначено через клас InstrumentedApp:

iep_value{unit_id} - поточне значення IEP (gauge); iep_kpi_normalized{kpi_id, unit_id} - нормовані KPI (gauge); iep_alert_total{type} - лічильник алертів (counter); api_request_duration_seconds{endpoint} - час запитів (histogram); kafka_consumer_lag_messages{topic} - відставання Kafka (gauge).

Grafana Dashboard включає 8 панелей: поточний IEP по підрозділах (bargauge із зеленою/жовтою/червоною зонами); динаміка IEP за тиждень (timeseries); статус алертів (state timeline); Kafka consumer lag (timeseries); розподіл часу відповіді API (histogram panel); RPS по ендпоінтах (timeseries); статус ML-моделі (stat: F1-score, MAE); доступність джерел даних (table: зелений/жовтий/червоний статус).

Тестування нечіткої системи виведення на граничних значеннях є критичним для коректності класифікації. Тест-кейси для граничних значень шкали: $\tilde{x} = 0{,}00$ → «Критичний» ($\mu = 1{,}0$); $\tilde{x} = 0{,}249$ → «Критичний» ($\mu = 0{,}96$), «Незадовільний» ($\mu = 0{,}04$); $\tilde{x} = 0{,}250$ → «Критичний» ($\mu = 0{,}96$), «Незадовільний» ($\mu = 0{,}04$); $\tilde{x} = 0{,}375$ → «Незадовільний» ($\mu = 1{,}0$); $\tilde{x} = 0{,}625$ → «Задовільний» ($\mu = 1{,}0$); $\tilde{x} = 0{,}825$ → «Добрий» ($\mu = 1{,}0$); $\tilde{x} = 1{,}00$ → «Відмінний» ($\mu = 1{,}0$). Усі 7 граничних тест-кейсів включені у регресійний набір.

Конфігурація Kubernetes HorizontalPodAutoscaler (HPA) для аналітичного рушія:

ДОДАТОК Б

Таблиці результатів апробації системи ІСО ЕІП

Таблиця Б.1 - Порівняння ключових показників до та після впровадження ІСО ЕІП

Показник	До впровадження	Після впровадження
Інтегральний ІЕР підприємства	0,659 (Задовільний)	0,751 (Добрий)
ІЕР відділу закупівель	0,341 (Незадовільний)	0,556 (Задовільний)
ІЕР відділу логістики	0,771 (Добрий)	0,847 (Добрий+)
Час виявлення деградації KPI	3,2 доби (вручну)	< 1 год (авто)
Охоплення KPI-моніторингом	3 показники (вручну)	12 показників (авто)
Точність ML-класифікатора (F1)	-	0,879
Похибка прогнозу LSTM (MAE)	-	0,018
Час реєстрації договору (закупівлі)	4,2 год	0,8 год
Розбіжності ERP–WMS (логістика)	1,8%	0,2%

Таблиця Б.2 - Детальні метрики ML-класифікатора LightGBM

Клас стану	Precision	Recall	F1-score	Support
Нормальний	0,912	0,924	0,918	108
Деградований	0,861	0,843	0,852	49
Критичний	0,900	0,878	0,889	23
Macro avg	0,891	0,882	0,886	180

Таблиця Б.3 - Кластерні ваги та субіндекси Q4

Кластер	α□	Σw_i	KPI	S□ (Q4)
Якість даних	0,28	0,28	3	0,745
Оперативність	0,36	0,36	3	0,788
Вартість	0,16	0,16	2	0,724
Безпека	0,20	0,20	2	0,834
РАЗОМ / IEP Q4	1,00	1,00	10	0,751

Таблиця Б.4 - Feature importance LightGBM-класифікатора

KPI-показник	Відносна важливість	Ранг важливості	Кластер
Відсоток помилок у майстер-даних	0,28	1	Якість даних
Запізнення управлінської звітності	0,22	2	Оперативність
Розбіжності ERP–WMS	0,19	3	Якість даних
Вартість обробки транзакції (CPT)	0,14	4	Вартість

KPI-показник	Відносна важливість	Ранг важливості	Кластер
Час реєстрації замовлення	0,09	5	Оперативність
Інциденти несанкціонованого доступу	0,04	6	Безпека
Решта 6 KPI (разом)	0,04	7–12	Різні

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра інформаційних систем та технологій

КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

Інтелектуальна система оцінювання ефективності інформаційних процесів підприємства

Виконав: **Лейбюк Володимир Володимирович**
Керівник: к. ф.-м. н., старший викладач **Ровіл НАФЄЄВ**

Київ — 2026

МЕТА, ОБ'ЄКТ ТА ПРЕДМЕТ ДОСЛІДЖЕННЯ

Актуальність

Більшість підприємств не має систематичного підходу до оцінювання ефективності інформаційних процесів. Вирішення потребує системи, що поєднує методи нечіткої логіки, багатокритеріального аналізу та машинного навчання.

МЕТА

Розробка та апробація інтелектуальної системи оцінювання ефективності інформаційних процесів підприємства (ICO ЕІП), що забезпечує автоматизований розрахунок інтегрального показника IEP на основі методів нечіткої логіки, багатокритеріального аналізу та машинного навчання.

ОБ'ЄКТ

Інформаційні процеси підприємства як сукупність операцій зі збирання, зберігання, опрацювання, передачі та використання інформації в межах організаційної структури з метою забезпечення управлінської діяльності.

ПРЕДМЕТ

Методи, моделі та програмні засоби інтелектуального оцінювання ефективності інформаційних процесів підприємства в умовах невизначеності та багатокритеріальності.

ЗАДАЧІ КВАЛІФІКАЦІЙНОЇ РОБОТИ

1. Проаналізувати структуру інформаційних потоків підприємства та систематизувати підходи до оцінювання їхньої ефективності
2. Сформувати збалансовану систему KPI та визначити вагові коефіцієнти методом аналізу ієрархій (AHP)
3. Розробити математичну модель інтегрального показника ефективності IEP на основі адитивно-мультиплікативної агрегації
4. Побудувати апарат нечіткого виведення (система Мамдані) для класифікації IEP та модель TOPSIS для ранжування підрозділів
5. Спроектувати мікросервісну архітектуру ICO EIП та реалізувати програмний прототип (Python, FastAPI, Kafka, TimescaleDB, React)
6. Провести апробацію системи та сформулювати рекомендації щодо підвищення ефективності

АНАЛІЗ АНАЛОГІВ

Система/підхід	Якість даних	Авто-моніторинг	ML-прогноз	Нечітка логіка	TOPSIS
Однокритеріальне (SLA)	Ні	Частково	Ні	Ні	Ні
Фреймворки CMMI/ITIL	Частково	Ні	Ні	Ні	Ні
IT BSC / Balanced Scorecard	Частково	Ні	Ні	Ні	Ні
BI-платформи (Power BI)	Частково	Частково	Частково	Ні	Ні
ІСО ЕІП (розроблена)	**✓ Повністю**	**✓ Так**	**✓ Так**	**✓ Так**	**✓ Так**

ВИМОГИ ДО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Функціональні вимоги

1. Збирання 12 KPI з корпоративних систем (ERP, WMS, SCADA, CRM) у реальному часі
2. Нормування показників до шкали [0, 1]
3. Розрахунок кластерних субіндексів та інтегрального IEP
4. Лінгвістична класифікація за п'ятирівневою нечіткою шкалою (система Мамдані)
5. Ранжування підрозділів методом TOPSIS
6. Прогнозування IEP на 7 діб засобами LSTM
7. Виявлення аномалій (Isolation Forest)
8. Автоматичні рекомендації на основі SHAP-аналізу

Нефункціональні вимоги

Продуктивність:	API p95 < 300 мс, 100+ одночасних користувачів
Доступність:	SLA 99,5% (RTO ≤ 2 год, RPO ≤ 1 год)
Безпека:	OAuth 2.0, RBAC (4 ролі), TLS 1.3, AES-256
Масштабованість:	HPA Kubernetes, до 50 підрозділів / 100 KPI
Аудит:	Журнал усіх змін конфігурації, 2+ роки зберігання
Backup:	Стратегія 3-2-1, PITR

ПРОГРАМНІ ЗАСОБИ РЕАЛІЗАЦІЇ

Python 3.11

Мова реалізації аналітичного ядра, ML-модулів та API

FastAPI

Веб-фреймворк: REST API, JWT-автентифікація, OpenAPI 3.1

Apache Kafka

Брокер повідомлень для потокового прийому KPI-подій

TimescaleDB

PostgreSQL-розширення для зберігання часових рядів KPI

scikit-fuzzy

Нечітке виведення Мамдані: лінгвістична класифікація ІЕР

LightGBM

Градієнтний бустинг: класифікація стану (F1 = 0,879)

LSTM (Keras)

Прогнозування ІЕР на 7 діб (MAE = 0,018)

React 18

Фронтенд: дашборди, gauge-графіки, WebSocket-оновлення

Kubernetes

Оркестрація мікросервісів, HPA, CI/CD (GitHub Actions)

АРХІТЕКТУРА ТА ОПИС СИСТЕМИ ICO ЕІП

Рівень збору даних	Брокер повідомлень	Аналітичне ядро	ML-модуль	Представлення
ERP (SAP) WMS SCADA CRM	Apache Kafka Topics: raw_kpi, normalized, scores	IEP Calculator Fuzzy Mamdani TOPSIS Service	LightGBM LSTM (7-діб) Isolation Forest	React 18 Дашборди KPI WebSocket

12 KPI у 4 кластерах: Якість даних (w=0.28) Оперативність (w=0.36) Вартість (w=0.16) Безпека (w=0.20)

IEP = ∏ S□^αj, де S□ — нормований субіндекс кластера j, α□ — вагові коефіцієнти AHP

ЕКРАННІ ФОРМИ СИСТЕМИ

Головний дашборд KPI

- Gauge-індикатор IEP (реальний час)
- Карткові виміри 12 KPI
- WebSocket-оновлення кожні 60 с
- Кольорова індикація рівнів

Рейтинг підрозділів TOPSIS

- Ранжований список 6 підрозділів
- Значення IEP Q1→Q4
- Динаміка у відсотках
- Виділення лідера і відстаючого

ML-прогноз та аномалії

- Графік прогнозу IEP на 7 діб (LSTM)
- Позначення виявлених аномалій
- SHAP-пояснення факторів впливу
- Автоматичні рекомендації

Налаштування ваг AHP

- Матриця попарних порівнянь
- Перевірка CR < 0,10
- Версіонування конфігурацій
- Журнал аудиту змін

РЕЗУЛЬТАТИ АПРОБАЦІЇ

Підрозділ	IEP Q1	IEP Q4	Динаміка
Логістика	0,771	0,847	+9,9%
Продажі	0,643	0,742	+15,4%
Виробництво	0,712	0,760	+6,7%
Фінанси	0,688	0,728	+5,8%
HR	0,601	0,671	+11,6%
Закупівлі	0,341	0,556	+63,1% ↑
ПІДПРИЄМСТВО	**0,659**	**0,751**	**+14,0%**

F1-score (LightGBM)

0,879

MAE LSTM (7 діб)

0,018

Accuracy ML

88,9%

Coverage тестів

94%

API p95 (100 user)

298 мс

Попереджень аномалій

3 рази

ВИСНОВКИ

1. Проаналізовано структуру інформаційних потоків підприємства, виявлено три покоління архітектур обробки даних та обґрунтовано принципи Data Lineage і Data Governance як основу управління якістю.

2. Сформовано систему 12 KPI у чотирьох кластерах з вагами AHP (CR < 0,10); ключові показники: час реєстрації (w=0,14), інциденти безпеки (w=0,12), помилки майстер-даних (w=0,12).

3. Розроблено адитивно-мультиплікативну дворівневу модель IEP, доведено властивості монотонності та стійкості; побудовано п'ятирівневу нечітку шкалу (TFN) на основі методу Дельфі.

4. Спроектовано мікросервісну архітектуру ICO EIП (Python, FastAPI, Kafka, TimescaleDB, React 18); реалізовано TOPSIS-ранжування та SHAP-пояснення рекомендацій.

5. Проведено апробацію на підприємстві «АльфаМаш»: загальний IEP зріс на 14,0% (0,659 → 0,751), IEP відділу закупівель — на 63,1%; F1-score класифікатора 0,879, MAE прогнозу LSTM = 0,018.

6. Розроблено пріоритетні рекомендації (5 заходів, ΔIEP +0,22–0,35); оцінено ROI: термін окупності 20–24 місяці.

Дякую за увагу!

Готовий відповісти на Ваші запитання

Лейбюк Володимир Володимирович • ДУІКТ • 2026