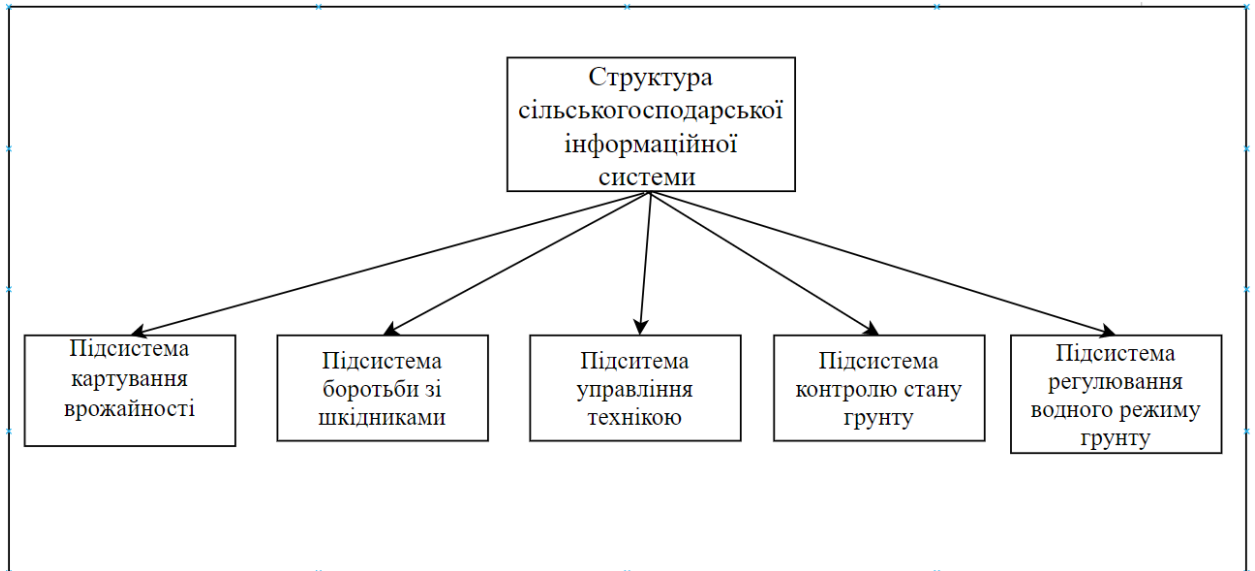


ДОДАТОК А
Графічний матеріал

Цілі та задачі досліджень

Об'єкт дослідження	Процес попередньої обробки даних щодо аграрних робіт у сільськогосподарських інформаційних системах
Предмет дослідження	Методи інтелектуального аналізу процесів
Мета роботи	Дослідження методів попередньої обробки даних у сільськогосподарських інформаційних системах для побудови моделей процесів аграрних робіт
Задачі дослідження	1. Аналізу особливостей сільськогосподарських інформаційних систем 2. Дослідження методів Process mining 3. Дослідження методів структуризації вхідних даних у задачах process mining 4. Удосконалення методу попередньої обробки вхідних даних для вирішення задачі інтелектуального аналізу при переході до процесного управління в сільськогосподарських системах 5. Експериментальна перевірка удосконаленого методу обробки вхідних даних
Наукова новизна	Удосконалений метод попередньої обробки даних на основі упорядкування подій за артибутами для побудови моделей процесу аграрних робіт методами process mining
Практичні результати	Підвищення точності побудови процесних моделей на неструктурованих даних

Структура інформаційної системи сільського господарства

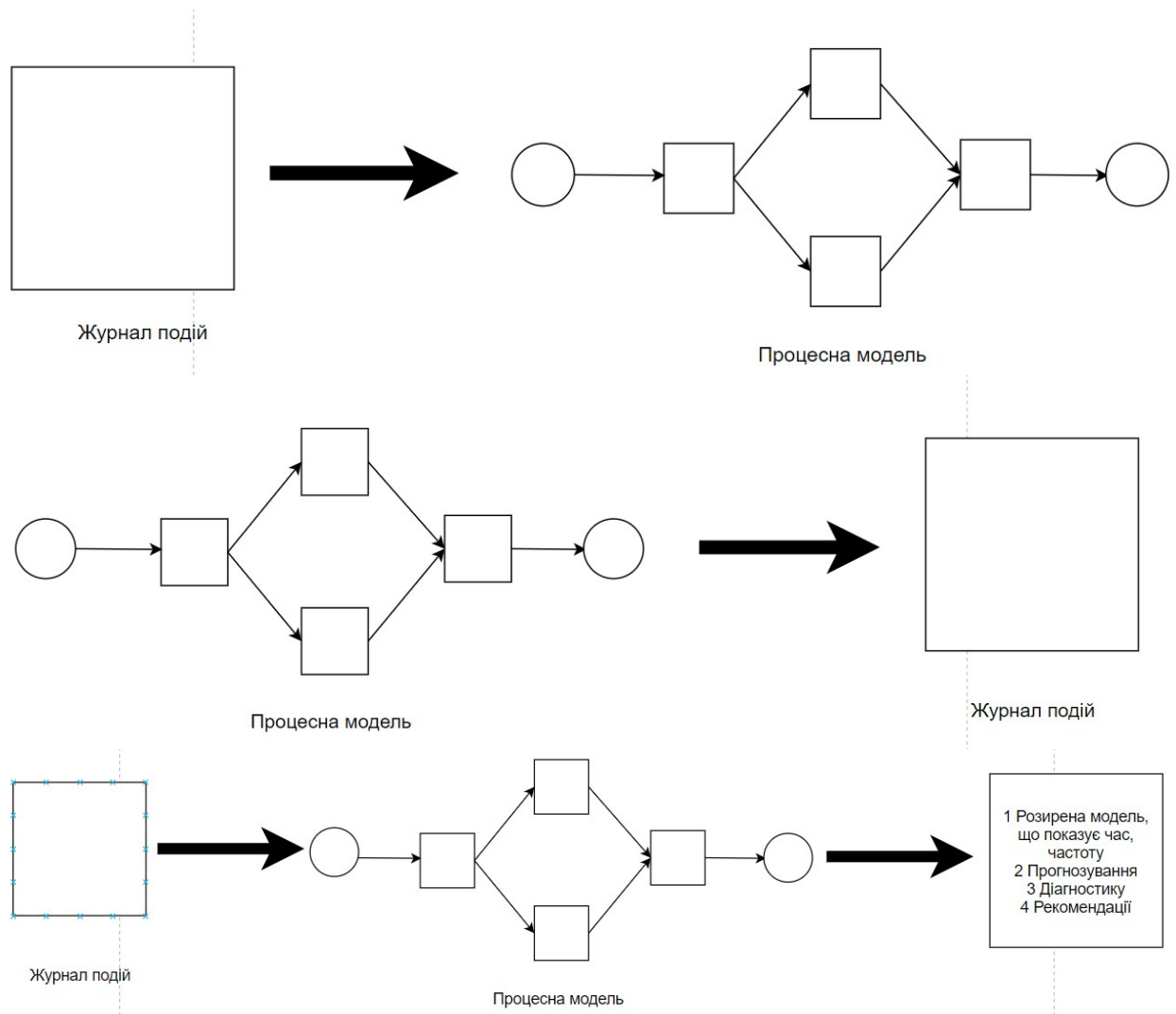


Порівняння процесного та функціонального управління

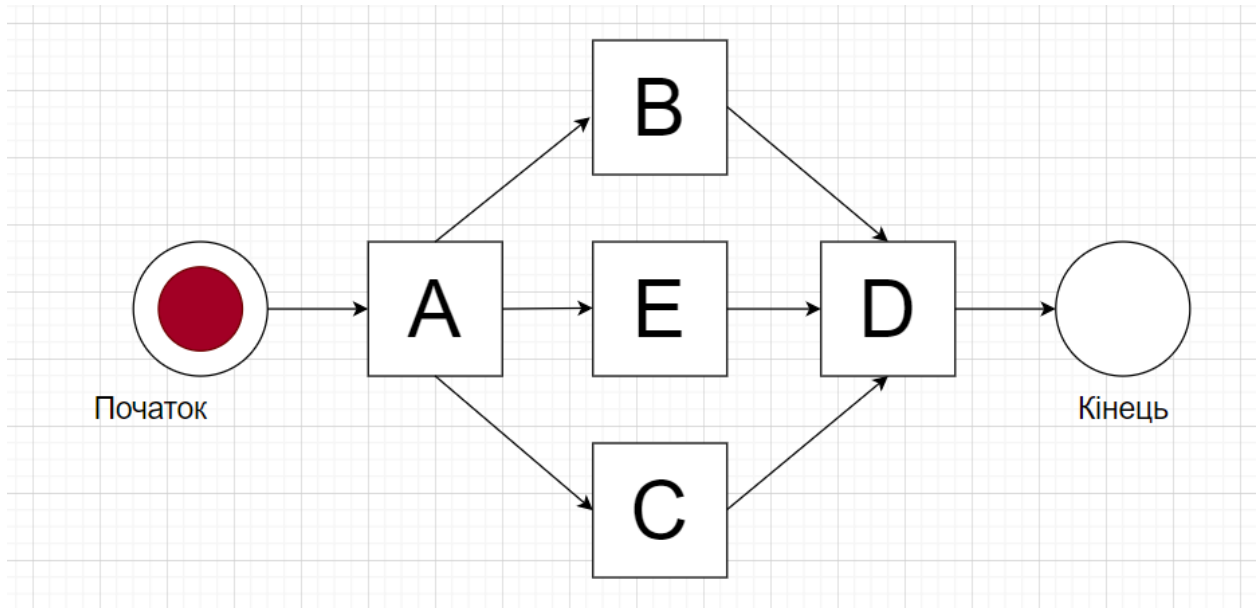
	Суть	Об'єкт	Недоліки	Переваги
Функціональний	Підприємство розглядається як механізм, що володіє набором функцій, які розподіляються серед підрозділів	Відділ, організаційна структура	1) незацікавлені працівники; 2) ускладнений обмін інформацією між підрозділами; 3) потрібно багато часу для реалізації управлінського впливу на виробничий процес; 4) авторитарне управління.	1) Легкий процес прийняття рішень
Процесний	Підприємство представляється набором процесів, що проходять через підрозділи і задіюють служби підприємства	Бізнес процес	1) взаємозалежність осіб, які приймають рішення	1) орієнтація на результат; 2) висока ефективність; 3) гнучкість реагування на зовнішні і внутрішні зміни; 4) контроль якості процесу, а не кінцевого продукту; 5) мотивація; 6) оперативність; 7) демократична система управління

Задачі інтелектуального аналізу процесів

Основною задачею інтелектуального аналізу процесів є побудова моделей бізнес-процесів, тобто перехід від функціонального до процесного управління.



Вимоги до вхідних даних аналізу процесів

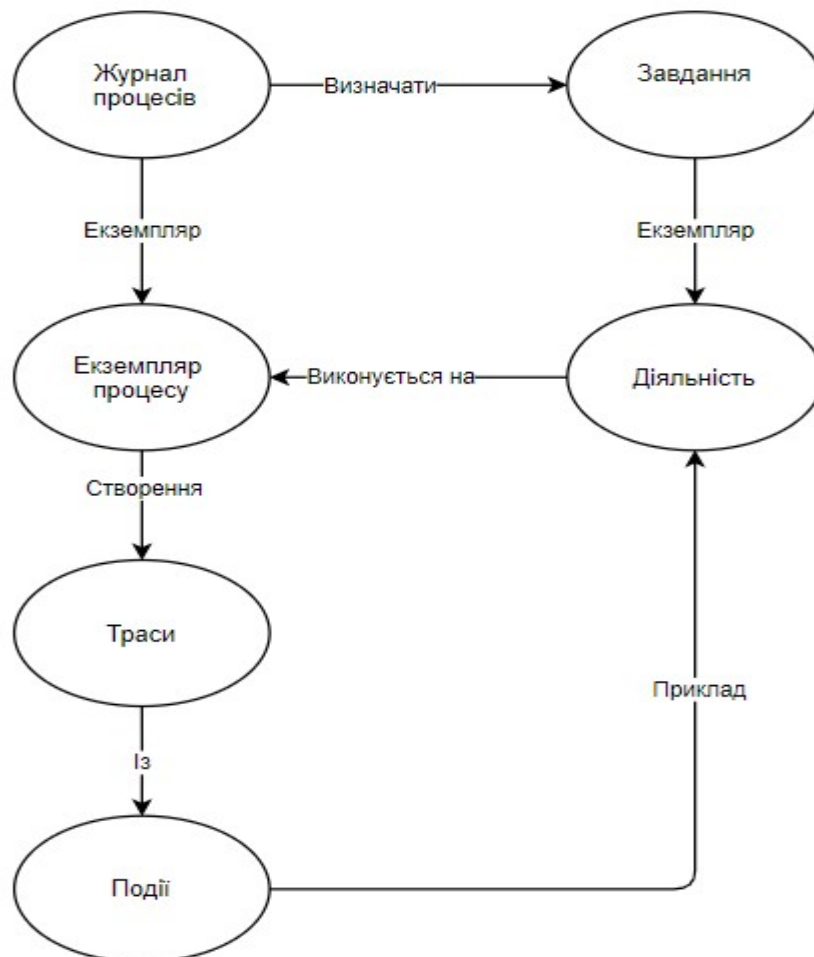


Для побудови процесної моделі використовуються методи process mining. Методи потребують вхідних даних у спеціалізованих форматах. Це викликає труднощі при побудові моделей, бо в сільському господарстві більшість процесів не надають інформацію у однакових форматах.

Вимоги до вхідних даних:

1. Дані необхідно представити у форматі логів журналу.
2. Дані повинні містити траси одного процесу.
3. Дані однієї моделі мають знаходитися у одному журналі подій;
4. Дані мають містити невелику кількість помилок;

Структуризація вхідних даних у задачах process mining



У процесі інтелектуального аналізу вхідні дані формуються у реляційну структуру наступного вигляду:

– події- найменша структурна одиниця;

- процес відображається множиною подій, що формують трасу;
- траса це послідовність подій, що відображає одну реалізацію процесу;
- журнал подій- набір трас;
- атрибути: мета-дані, що можуть характеризувати подію, або процес.

Постановка задачі

Сільськогосподарські роботи мають процесний характер у відповідності до сезонів року. Сільськогосподарські системи використовують функціональний підхід до управління, тому у сільськогосподарському виробництві відбувається перехід від традиційного функціонального до процесного управління. Функціонування сільськогосподарських систем процесного управління потребує побудови процесних моделей аграрного виробництва. Такі моделі формуються з використання методів process mining, з використанням вхідних даних у спеціальному процесно-орієнтованому форматі. Однак існуючі дані про сільськогосподарське виробництво не відповідають цьому формату, що потребує реалізації попередньої обробки вхідних даних та їх трансформації в форму придатну для використання методів process mining.

Послідовність перетворення вхідних даних для process mining

Схема перетворення даних із вхідних джерел до журналів подій XES:



Удосконалений метод підготовки вхідних даних для побудови процесної моделі

Етап 1. Формування набору обмежень по об'єктам процесу.

Етап 2. Відбір подій журналу, що відображають вирішення операції із відібраними об'єктами процесу.

Етап 3. Побудова підмножин подій логу для інтервалів часу .

Етап 4. Упорядкування підмножин подій за часом та формування трас:

$$T = \{t_1, t_2, \dots, t_n\}, \quad (1)$$

де t_i – подія, T – траса, M – множина трас.

Етап 5. Уточнення множини трас за частотою подій.

Крок 5.1 Обчислення вірогідності появи події у процесі відповідає відношенню кількості трас, що містять подію, до кількості усіх процесів:

$$P_i = \frac{N_i}{N}, \quad (2)$$

Крок 5.2 Відбір підмножини трас для яких виконується умова:

$$P_i \geq \theta, \quad (3)$$

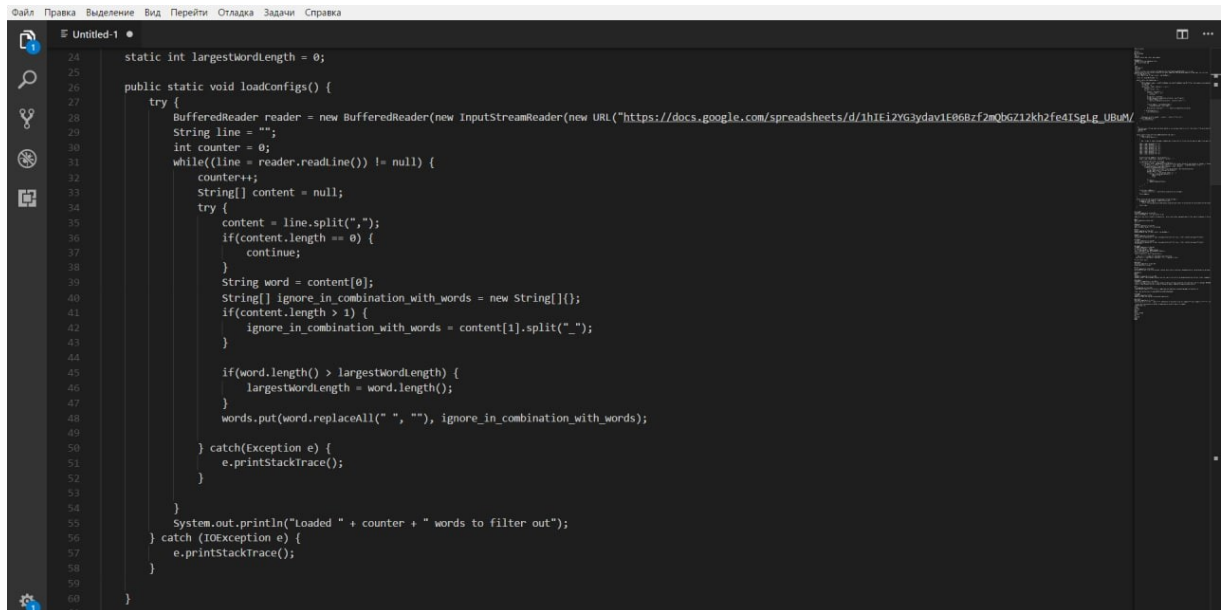
де θ – порогове значення частоти подій.

Алгоритм роботи доданого етапу



Реалізація удосконаленого методу

Формування набору підмножин подій журналу за класифікацією частоти виникання у процесах. Основна ідея запропонованого етапу – класифікація подій за частотою вживання у процесі.



```

24 static int largestWordLength = 0;
25
26 public static void loadConfigs() {
27     try {
28         BufferedReader reader = new BufferedReader(new InputStreamReader(new URL("https://docs.google.com/spreadsheets/d/1h1Ei2YG3yday1E868zf2mQbGZ12kh2fe4I5gt_g_UBu4/
29         String line = "";
30         int counter = 0;
31         while((line = reader.readLine()) != null) {
32             counter++;
33             String[] content = null;
34             try {
35                 content = line.split(",");
36                 if(content.length == 0) {
37                     continue;
38                 }
39                 String word = content[0];
40                 String[] ignore_in_combination_with_words = new String[]{};
41                 if(content.length > 1) {
42                     ignore_in_combination_with_words = content[1].split("_");
43                 }
44
45                 if(word.length() > largestWordLength) {
46                     largestWordLength = word.length();
47                 }
48                 words.put(word.replaceAll(" ", ""), ignore_in_combination_with_words);
49             } catch(Exception e) {
50                 e.printStackTrace();
51             }
52         }
53         System.out.println("Loaded " + counter + " words to filter out");
54     } catch (IOException e) {
55         e.printStackTrace();
56     }
57 }
58
59 }
60

```

Формування процесної моделі

repaExample.xes

Log Summary

Total number of events: 11066

MXML Legacy Classifier

Event classes defined by MXML Legacy Classifier

All events

Total number of classes: 12

Class	Occurrences (absolute)	Occurrences (relative)
Test.Repair.complete	1618	12.72%
Test.Repair.start	1518	12.72%
Register.complete	1114	9.313%
Analyze.Defect.complete	1114	9.313%
Analyze.Defect.start	1114	9.313%
Inform.User.complete	1112	9.296%
Archive.Repair.complete	1010	8.435%
Repair (Simple).complete	781	6.922%
Repair (Simple).start	781	6.922%
Repair (Complex).start	725	6.116%
Repair (Complex).complete	724	6.107%
Restart.Repair.complete	401	3.420%

Start events

Total number of classes: 1

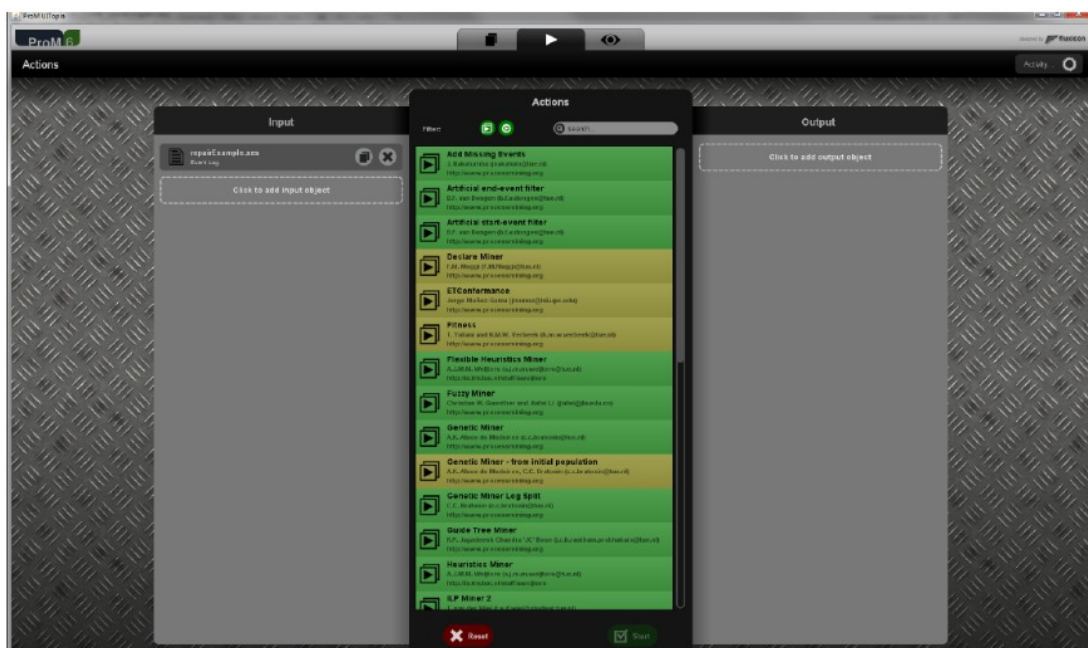
Class	Occurrences (absolute)	Occurrences (relative)
Register.complete	1114	100.0%

End events

Total number of classes: 5

Class	Occurrences (absolute)	Occurrences (relative)
Archive.Repair.complete	1010	90.58%
Test.Repair.complete	75	6.793%
Inform.User.complete	27	2.440%
Repair (Complex).start	1	0.001%
Repair (Complex).complete	1	0.001%

Зведений вигляд на прикладі журналу подій

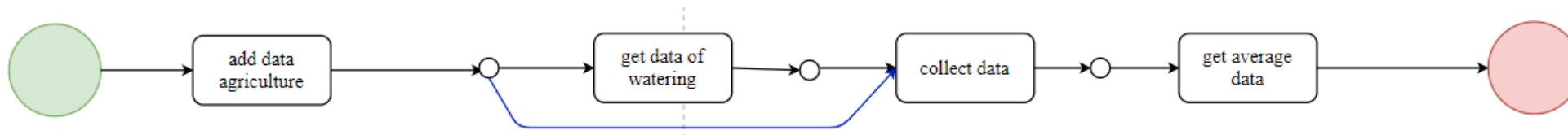


Експериментальна перевірка отриманих результатів

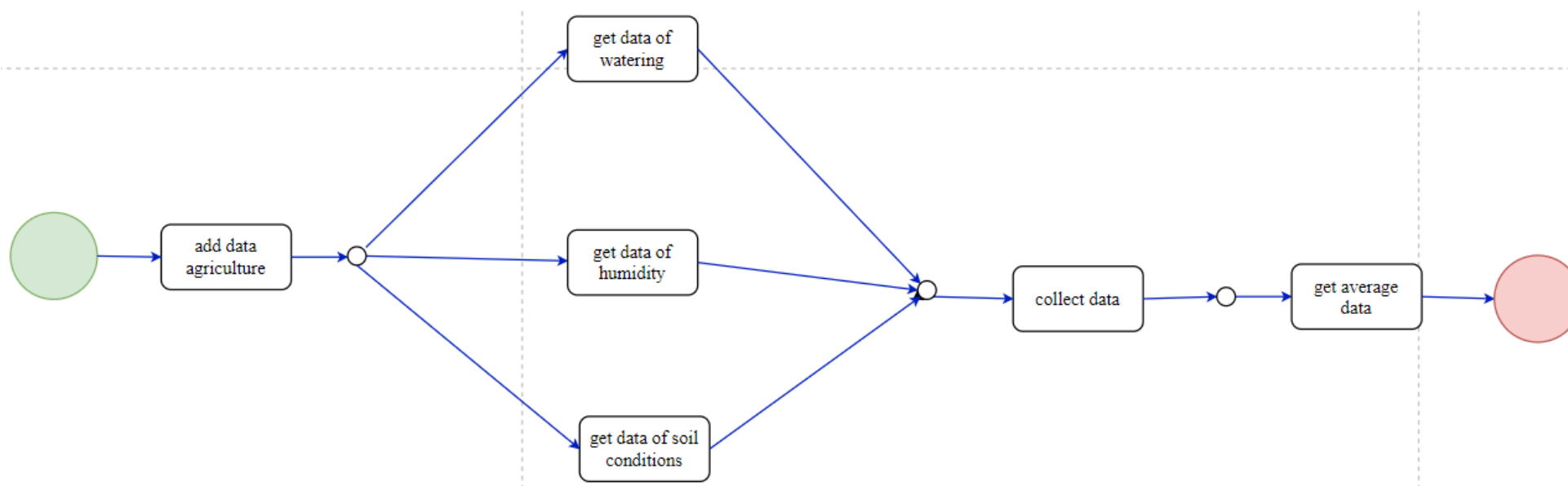
	Без використання модифікованого методу, відсоток подій що включений до складу моделі	Із використанням модифікованого методу, відсоток подій що включений до складу моделі
лог 1	70%	97%
лог 2	35%	90%

Приклад побудови процесної моделі

Без використання попередньої обробки вхідних даних



З використанням попередньої обробки вхідних даних



Експериментальна перевірка отриманих результатів

	Без використання модифікованого методу, відсоток подій що включений до складу моделі	Із використанням модифікованого методу, відсоток подій що включений до складу моделі
лог 1	70%	97%
лог 2	35%	90%

Висновки

В магістерській атестаційній роботі було досліджено проблему реалізації процесного управління в сільськогосподарських інформаційних системах. При переході до процесного управління використовується моделі процесів аграрних робіт. Побудова моделей таких процесів потребує виконання попередньої обробки даних.

В результаті виконання магістерської роботи був виконаний огляд методів попередньої обробки даних у сільськогосподарських інформаційних системах.

Був проведений аналіз особливостей сільськогосподарських інформаційних систем та досліджено методи process mining, які в них застосовуються.

Було досліджено методи структуризації вхідних даних у задачах process mining.

Було проведено аналіз вимог до вхідних даних, методи структуризації вхідних даних, процес конвертації вхідних даних до формату XES, послідовність перетворення вхідних даних.

Було удосконалено метод попередньої обробки неструктурованих даних у сільськогосподарських інформаційних системах .

Метод реалізовано у вигляді програмного додатку, що виконує попереднє упорядкування вхідних даних для побудови процесної моделі сільськогосподарського виробництва засобами process mining. Виконано експериментальну перевірку методу з використанням платформи ProM. Результати показали, що із використанням модифікованого методу, відсоток подій, що включений до складу моделі значно виріс.

Практичним результатом стало підвищення точності побудови процесних моделей на неструктурованих даних.