

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»  
НАВЧАЛЬНО-НАУКОВИЙ КОМПЛЕКС «ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ»  
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ**

# **Підвищення ефективності надання рекламної інформації шляхом масової розсилки**

**Виконав:** студент групи КА-21

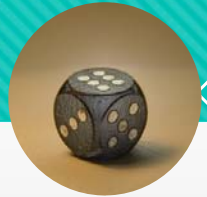
Логін В. В.

**Науковий керівник:**

д.т.н., професор Бідюк П. І.

Київ – 2016

# Структурні елементи дослідження



**Об'єкт дослідження:** масова рекламна розсилка.



**Предмет дослідження:** моделі відгуку клієнтів на розсилку рекламної інформації.



**Мета роботи:** підвищення ефективності масової рекламної розсилки за рахунок аналізу відгуків клієнтів, визначення потенційних клієнтів компанії та розробки стратегії мінімізації витрат.



**Методи дослідження:** методи інтелектуального аналізу даних, порівняльний аналіз ефективності побудованих моделей за передбачуваним прибутком.

# Актуальність роботи

**Підвищення  
ефективності  
рекламної  
розсилки  
необхідне для:**

- Підвищення конкурентоздатності;
- Збільшення доходів, зменшення витрат;
- Скорочення часу проведення розсилки.

**Існуючі  
проблеми:**

- Проведення рекламної розсилки без попереднього аналізу клієнтської бази може спричинити збитки для компанії замість отримання прибутку;
- Використання застарілих методів аналізу клієнтської бази.

# Змістовне формулювання задачі

Деяка торгова компанія, що здійснює продаж товарів, має інформацію щодо своїх клієнтів та їхніх покупок. При цьому компанія провела пробну рекламну розсилку усім своїм клієнтам та отримала певний відсоток відгуків.



За цими даними потрібно побудувати моделі відгуку та проаналізувати результати для того, щоб запропонувати способи мінімізації витрат на нові розсилки та розробити подальшу стратегію рекламної розсилки.

# Постановка задачі дослідження

Виконати загальний огляд методів аналізу даних та вибрати ті, які відповідають поставленій проблемі моделювання відгуків клієнтів на рекламну розсилку.

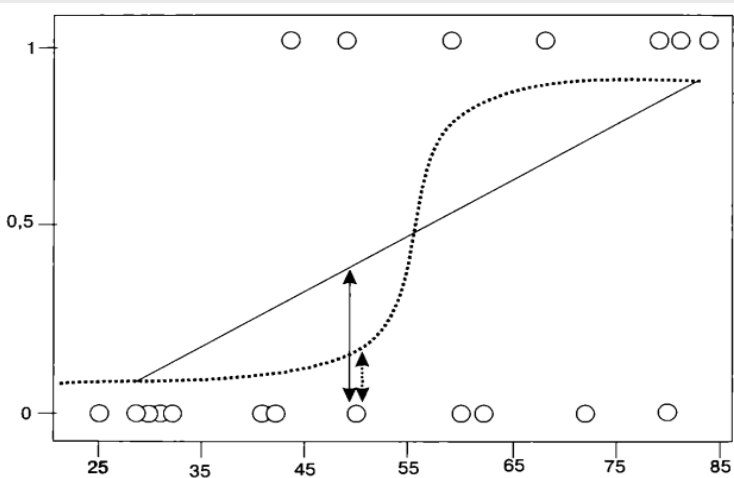
Побудувати моделі Data Mining за допомогою обраних підходів.

Порівняти отримані моделі та вибрати найкращу з точки зору ефективності.

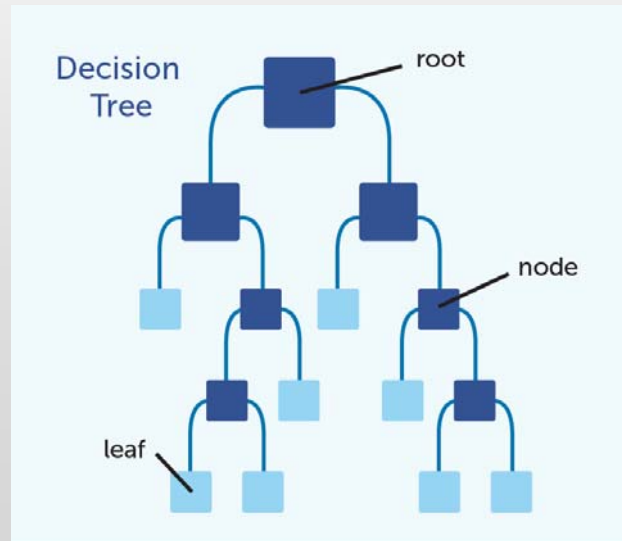
Виробити рекомендації щодо вдосконалення у майбутньому розроблених підходів до розв'язання даної задачі.

# Моделі, що використані в роботі

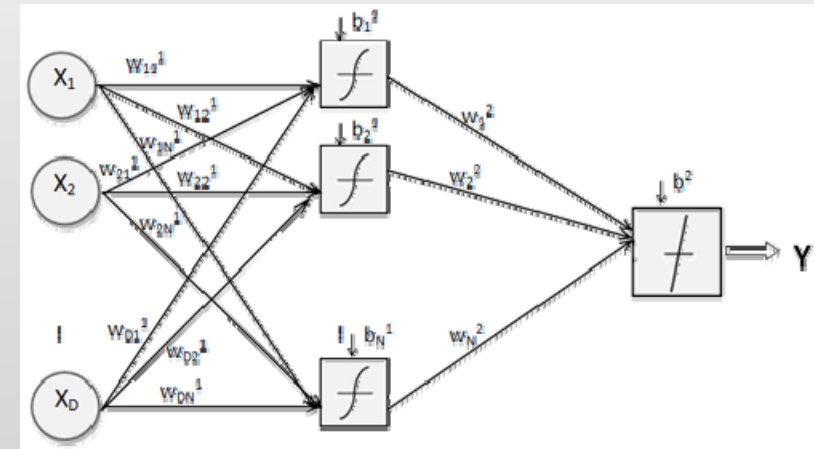
## Логістична регресія



## Дерево рішень



## Нейронна мережа



# Програмне забезпечення, використане в роботі

## Deductor

«Deductor» – аналітична платформа для створення завершених аналітичних рішень.



### Переваги платформи:



Інтеграція різноманітних даних



Зручна візуалізація



Можливість очищення даних



Наявні засоби Data Mining



Реалізовано сценарний підхід

# Загальний вигляд платформи Deductor

The screenshot displays the Deductor Studio Academic interface. The main window is titled "Deductor Studio Academic (D:\Vadimuch\Універ\Диплом (бакалавр)\Робота\DW.ded) - [Нейромережа [4 x 3 x 1]]". The interface is divided into several panes:

- Сценарии (Scenarios):** A tree view on the left showing the workflow: "Вхідні дані" (Input data) -> "Навчальна множина" (Training set) -> "Заміна значень: Стать (чоловіча -> 1, жіноча -> 0)" (Value replacement: Gender) -> "Логістична регресія" (Logistic regression) -> "Процедура Undersampling" (Undersampling procedure) -> "Фільтр ([Відгук] = 0)" (Filter [Review] = 0) -> "Фільтр ([Відгук] = 1)" (Filter [Review] = 1) -> "Об'єднання (Фільтр ([Cut\_coef] > 0,8235294118))" (Join (Filter [Cut\_coef] > 0,8235294118)) -> "Дерево рішень" (Decision tree) -> "Нейромережа [4 x 3 x 1]" (Neural network [4 x 3 x 1]).
- Граф нейросети (Neural network graph):** A central diagram showing a neural network with four input nodes: "Скільки років є клієнтом" (How many years is the client), "Кількість позицій товарів" (Number of product positions), "Дохід з клієнта, тис. од." (Revenue from the client, thousand units), and "Загальна кількість покупок" (Total number of purchases). These are connected to three hidden nodes, which are then connected to a single output node labeled "Відгук" (Review).
- Таблиця сопряженности (Correlation matrix table):** A table on the right showing the relationship between the input variables and the output variable. The columns are "Фактически" (Actual), "Классифицировано" (Classified), and "Итого" (Total). The rows are "0" and "1" for the "Відгук" variable.
- Что-если (What-if):** A table at the bottom showing the results of a "what-if" analysis for a specific input set.

Фактически	Классифицировано		
	0	1	Итого
0	1116	110	1226
1	367	801	1168
Итого	1483	911	2394

Поле	Значение
<b>Входные</b>	
9.0 Скільки років є к...	4
9.0 Кількість позицій...	4
9.0 Дохід з клієнта, т...	17,77
9.0 Загальна кількіс...	3
<b>Выходные</b>	
12 Відгук	0







# Блок-схема розв'язання задачі



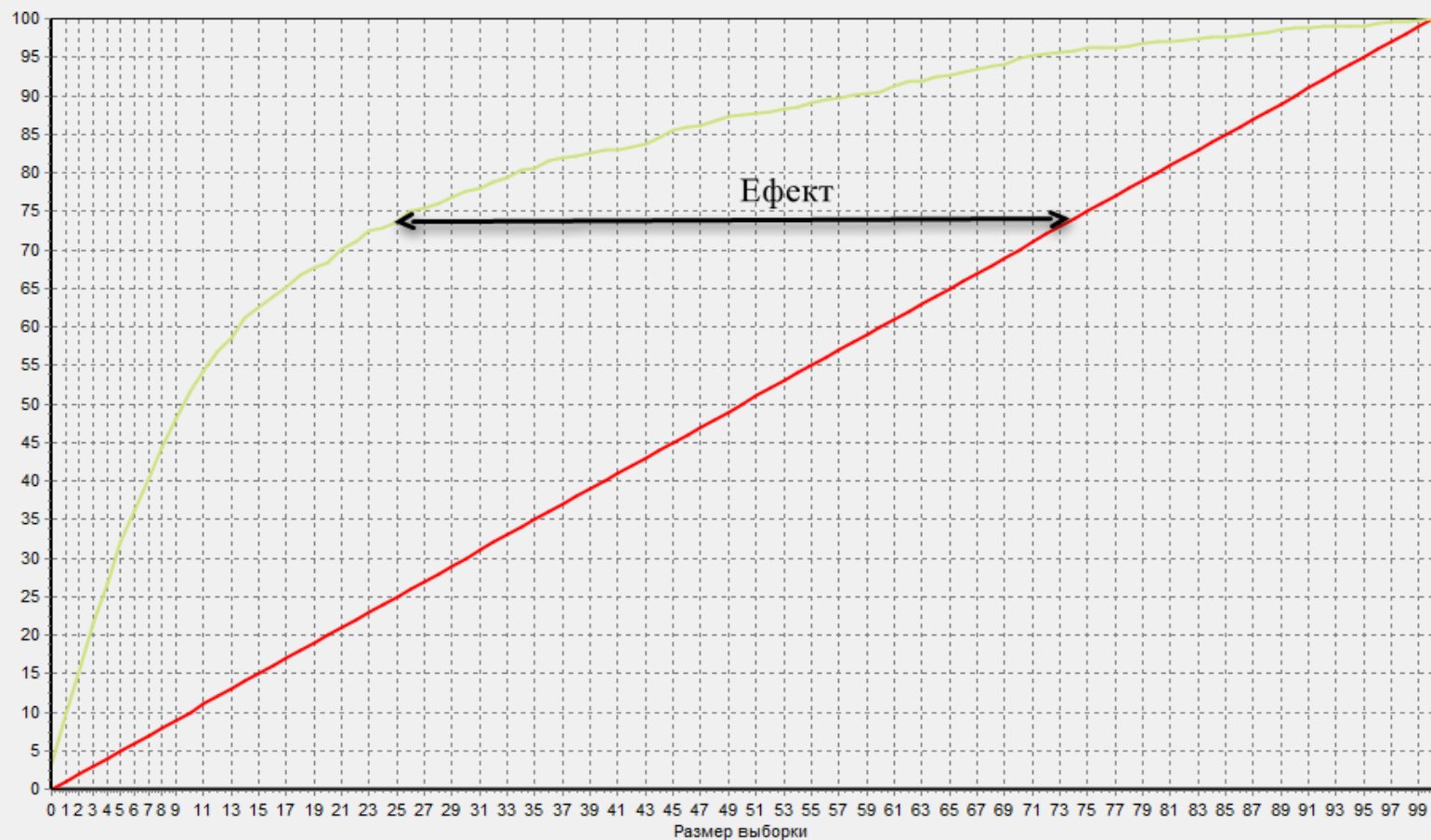
# Приклади застосування розробленого сценарію

## Проведення кореляційного аналізу

Матриця кореляції		
Входные поля		Корреляция с выходными полями
№	Поле	Відгук
1	Вік	 0,000
2	Стать	 -0,007
3	Скільки років є клієнтом	 0,127
4	Кількість позицій товарів	 0,341
5	Дохід з клієнта, тис. од.	 0,363
6	Загальна кількість покупок	 0,391
7	Звернень в службу підтримки	 0,009
8	Затримки платежів	 0,011
9	Дисконтна карта	 -0,005

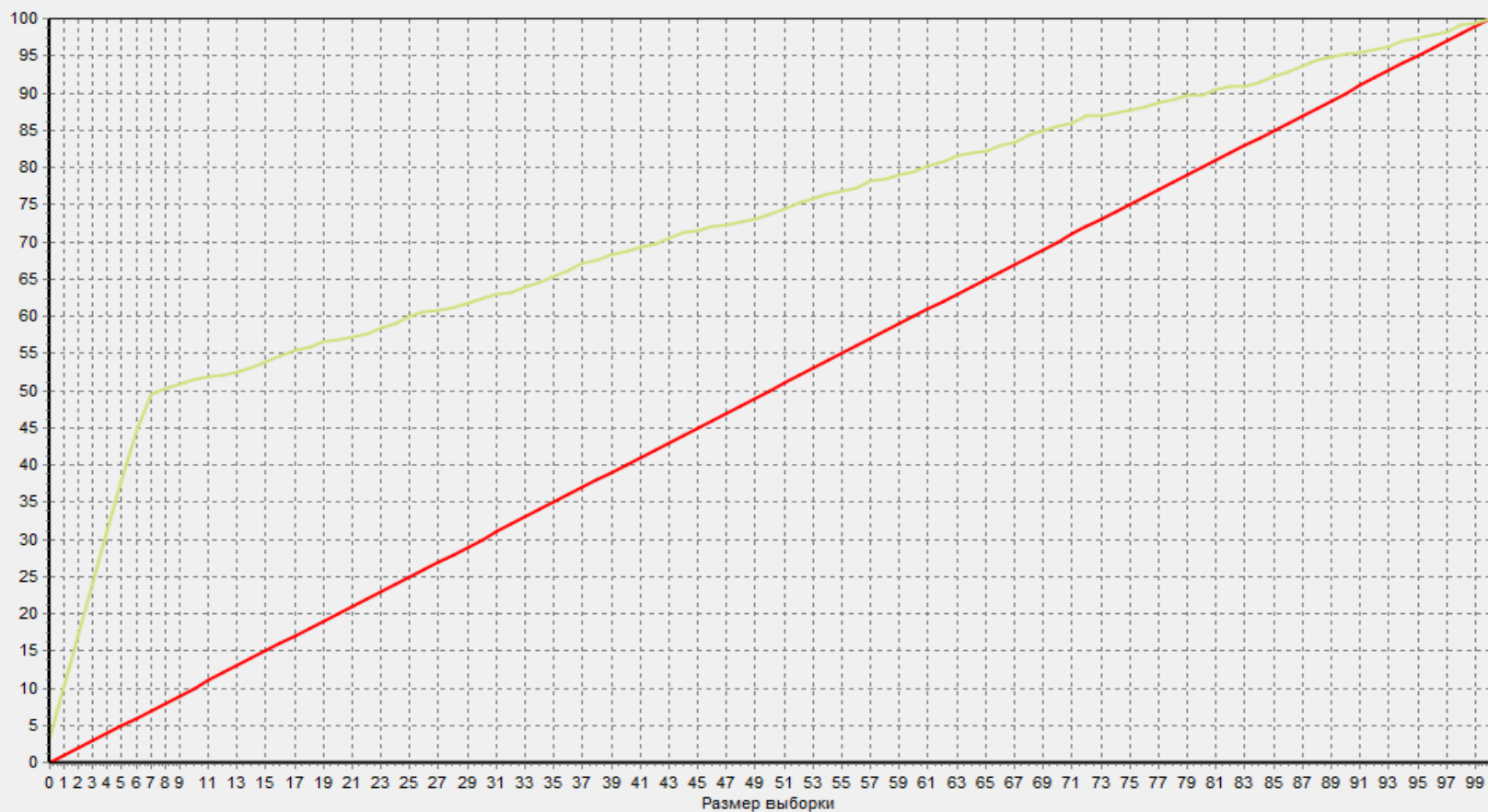
# Приклади застосування розробленого сценарію

Побудова Lift-кривої для моделі відгуку на основі логістичної регресії. Поріг відсікання після врівноваження – 0,15.



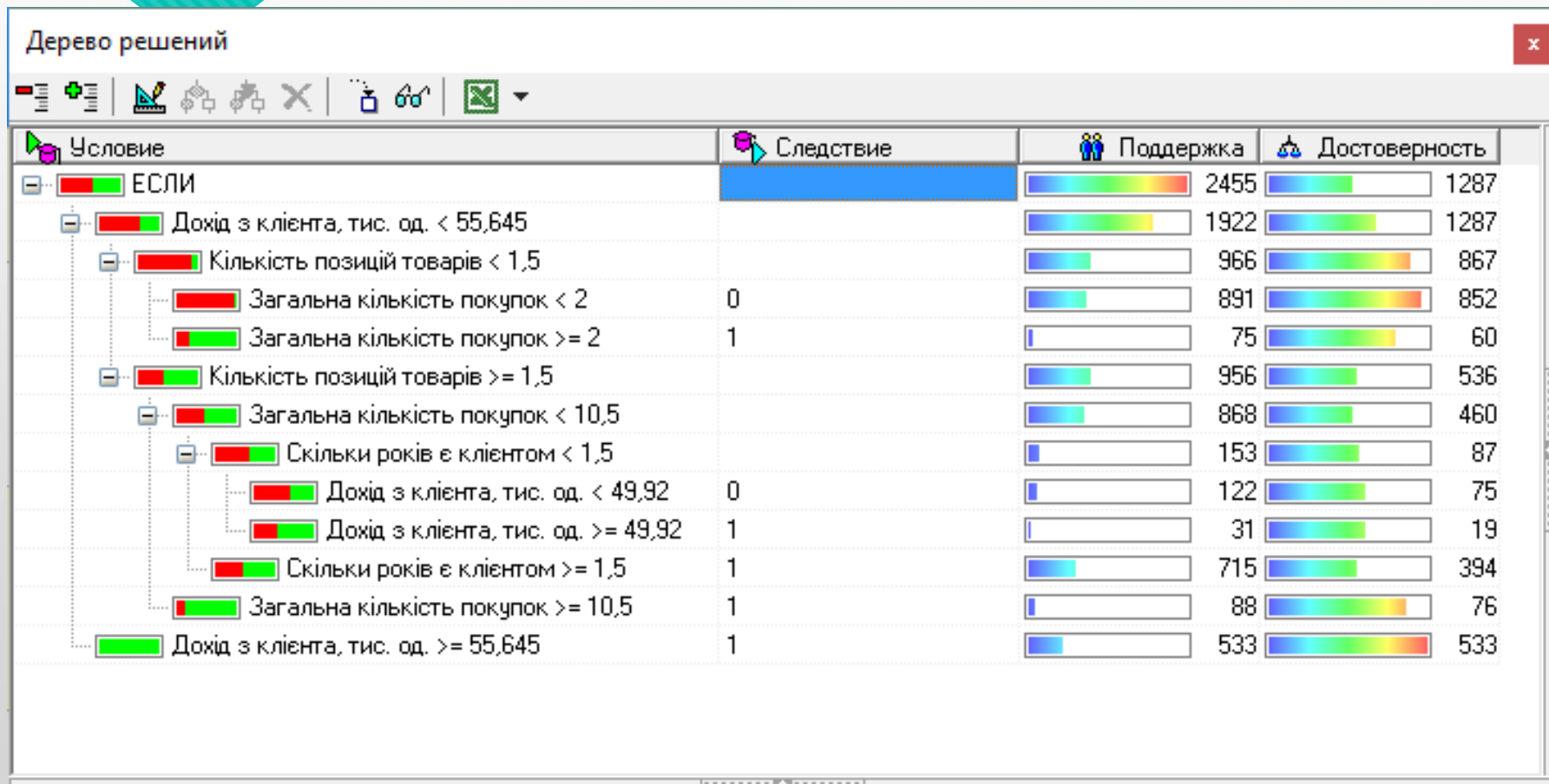
# Приклади застосування розробленого сценарію

Побудова Lift-кривої для випадку першочергової розсилки клієнтам, від яких був найбільший дохід.



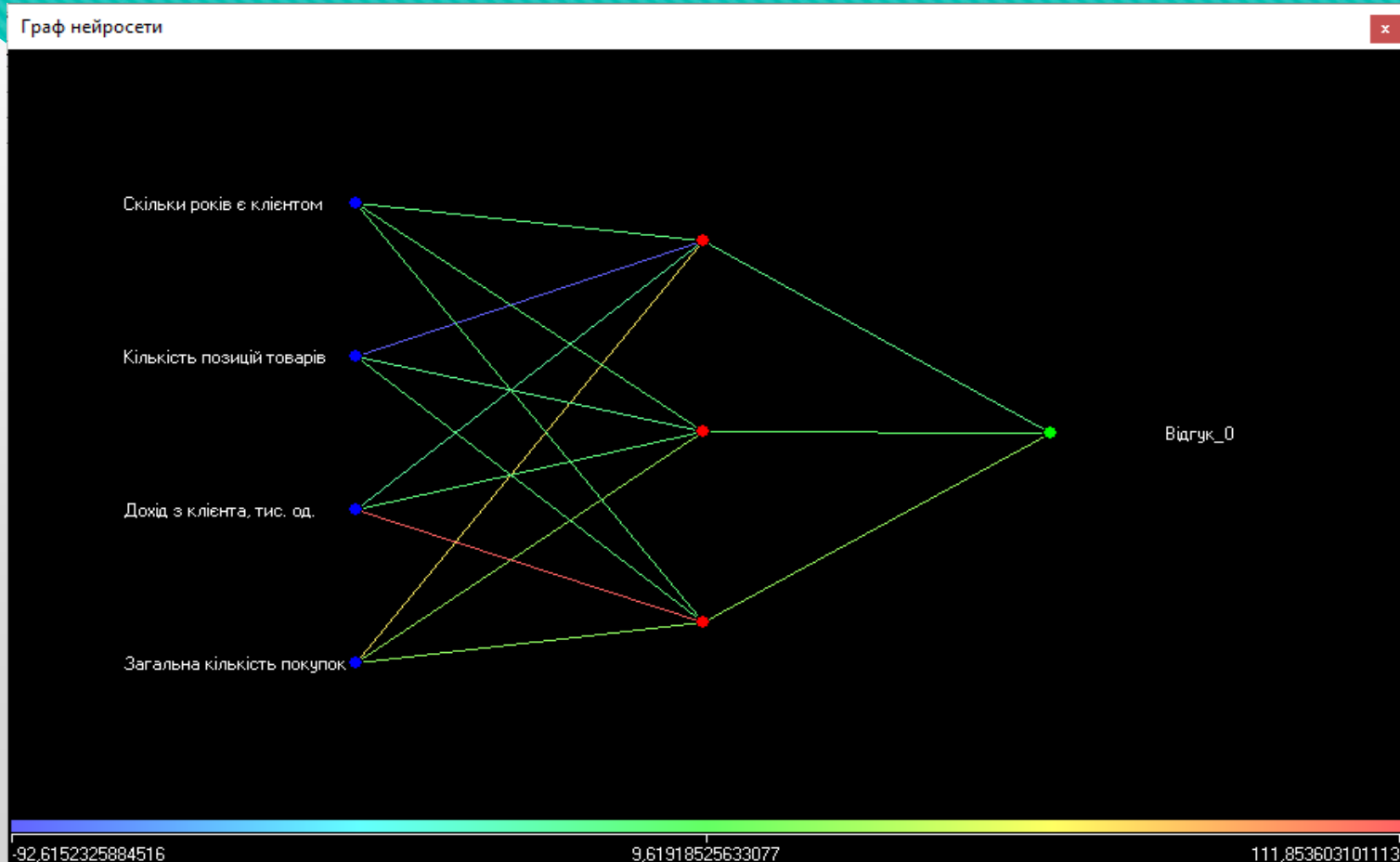
# Приклади застосування розробленого сценарію

Побудова дерева рішень. Для врівноваження відкидаємо 14/17 прикладів з клієнтами, від яких не було відгуку.



# Приклади застосування розробленого сценарію

Побудова нейронної мережі. Для врівноваження відкидаємо 14/17 прикладів з клієнтами, від яких не було відгуку.



# Точність класифікації побудованих моделей

Модель	Точність класифікації за моделлю	
	Начальна вибірка	Тестова вибірка
Логістична регресія	81,03 %	82,01 %
Дерево рішень	82,66 %	76,47 %
Нейронна мережа	80,08 %	87,86 %

# Вибір кращої моделі

Модель	TN	TP	FN	FP	Загальна помилка	Прибуток, од.
Логістична регресія (поріг 0,15)	3859	571	213	759	17,99 %	7430
Дерево рішень (undersampling, 17/3)	3430	701	83	1188	23,53 %	8353
Нейронна мережа (undersampling, 17/3)	4206	540	244	412	12,14 %	7944
«Розіслати всім»	0	784	0	4618	—	-526

Із наведеної вище таблиці видно, що найкращою моделлю для передбачення відгуку клієнта є дерево рішень, що дає прибуток 8353 од.

При використанні цієї моделі ефективність розсилки підвищується на  $8353 - (-526) = 8879$  од.



# Наукова новизна



Побудовано нові моделі для поставленої задачі класифікації відгуку клієнтів на основі обраного набору статистичних даних:

- Логістична регресія;
- Дерево рішень;
- Нейронна мережа.



При проведенні рекламної розсилки пропонується використовувати не стратегію «Розіслати всім» чи першочергової розсилки клієнтам, від яких був найбільший дохід, а стратегію, що заснована на моделі «Дерево рішень».

# ВИСНОВКИ



Розроблено сценарій розв'язку поставленої задачі підвищення ефективності рекламної розсилки у середовищі Deductor.



Розроблений сценарій застосовано до набору статистичних даних щодо клієнтів деякої торгової компанії, та проведений порівняльний аналіз побудованих 4 моделей:

- Логістична регресія;
- Дерево рішень;
- Нейронна мережа;
- Модель «Розіслати всім».



Рекомендовано використовувати для подальших розсилок не стратегію «Розіслати всім», а стратегію, що заснована на моделі «Дерево рішень».

# Перспективи для подальших досліджень

Побудова нових класифікаційних моделей, наприклад, алгоритмів «k-means», SVM та ін.

Подальше вдосконалення побудованих моделей з метою покращення якості класифікації.

Виконання перевірки побудованих моделей на реальних даних, щоб підтвердити їхню якість та адекватність.

# Наукові публікації

На основі дипломної роботи була написана наукова стаття «Підвищення ефективності надання рекламної інформації шляхом масової розсилки», що буде опублікована в збірнику «Системні науки та кібернетика».



A teal banner with a fine, repeating geometric pattern, featuring a small notch at the bottom center.

**Дякую за увагу !**