

Магістерська кваліфікаційна робота

Інформаційна технологія прогнозування завантаження банкоматів готівкою

Виконав

Кусяка Д. С.

Керівник: к.т.н., доц. Колесницький О.К.

МЕТА ТА ЗАВДАННЯ ДОСЛІДЖЕННЯ

Метою дослідження магістерської кваліфікаційної роботи є підвищення точності прогнозування завантаження банкоматів готівкою програмними засобами за рахунок застосування інтелектуальних технологій.

Для досягнення поставленої мети необхідно розв'язати такі завдання:

- провести аналіз проблеми розв'язання задачі прогнозування завантаження банкоматів готівкою;
- розглянути існуючі методи вирішення задачі прогнозування завантаження банкоматів готівкою та обрати й обґрунтувати вибір методу, який задовольняє мету даної магістерської кваліфікаційної роботи;
- розробити математичну модель прогнозування завантаження банкоматів готівкою;
- сформулювати стадії інформаційної технології, розробити структуру програмного засобу;
- виконати програмну реалізацію запропонованої інформаційної технології;
- провести тестування програмного продукту та виконати аналіз отриманих результатів..

ОБ'ЄКТ, ПРЕДМЕТ ТА МЕТОДИ ДОСЛІДЖЕННЯ

Об'єкт дослідження – процес прогнозування завантаження банкоматів готівкою з використанням інтелектуальних технологій.

Предмет дослідження – інформаційна технологія та програмні засоби прогнозування завантаження банкоматів готівкою з використанням інтелектуальних технологій та точність їх роботи.

Методи дослідження

У роботі використані наступні методи наукових досліджень:

- системного аналізу,
- інтелектуального аналізу даних,
- теорії баз даних та знань для реалізації інформаційної технології прогнозування завантаження банкоматів готівкою,
- методи математичної статистики для розробки процесу класифікації тональності речень та обрахунків результатів експериментів із програмним засобом,
- об'єктно-орієнтованого програмування.

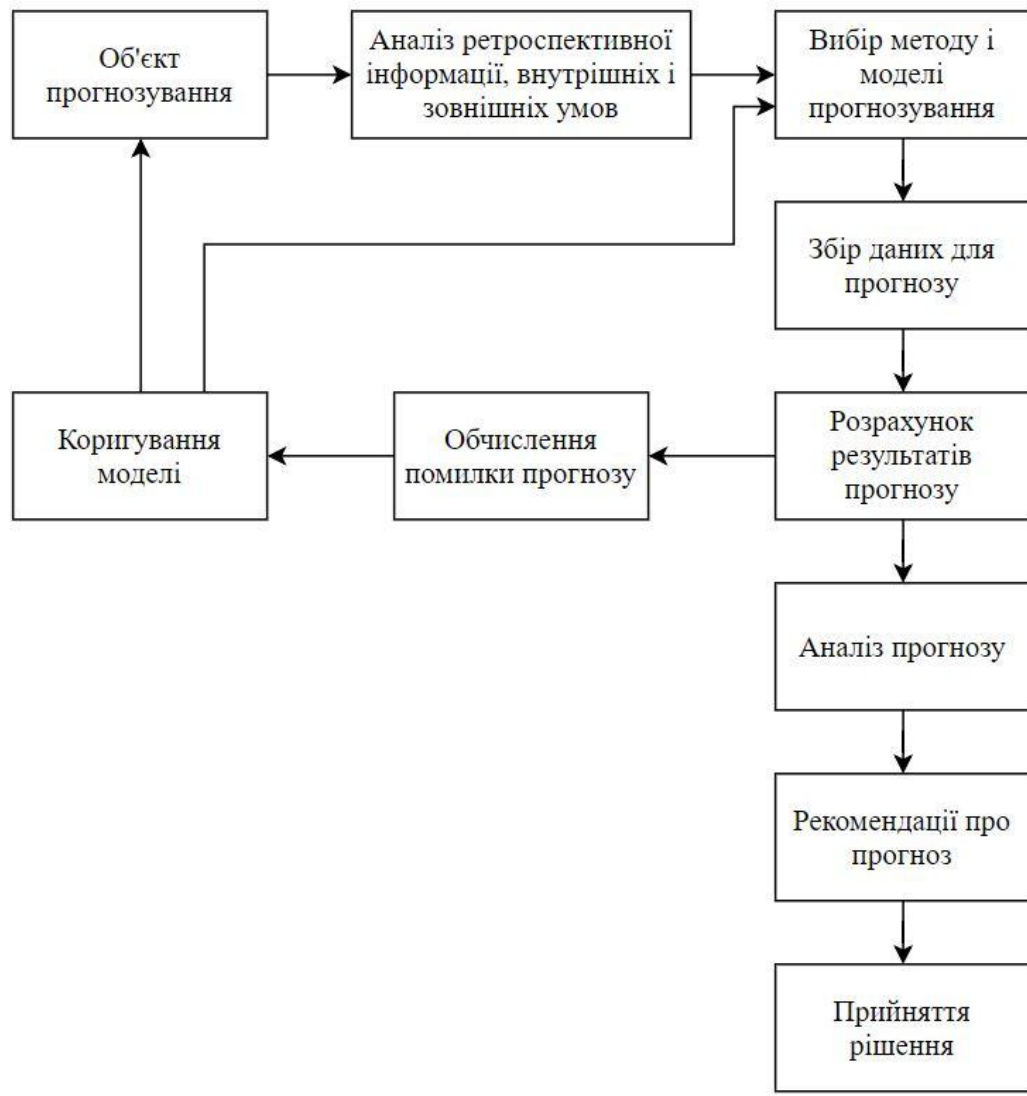
НАУКОВА НОВИЗНА ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

1. Набула подальшого розвитку інформаційна технологія прогнозування завантаження банкоматів готівкою, яка відрізняється спільним використанням технологій OLAP і Data Mining, що дозволило підвищити точність прогнозування завантаження банкоматів готівкою.

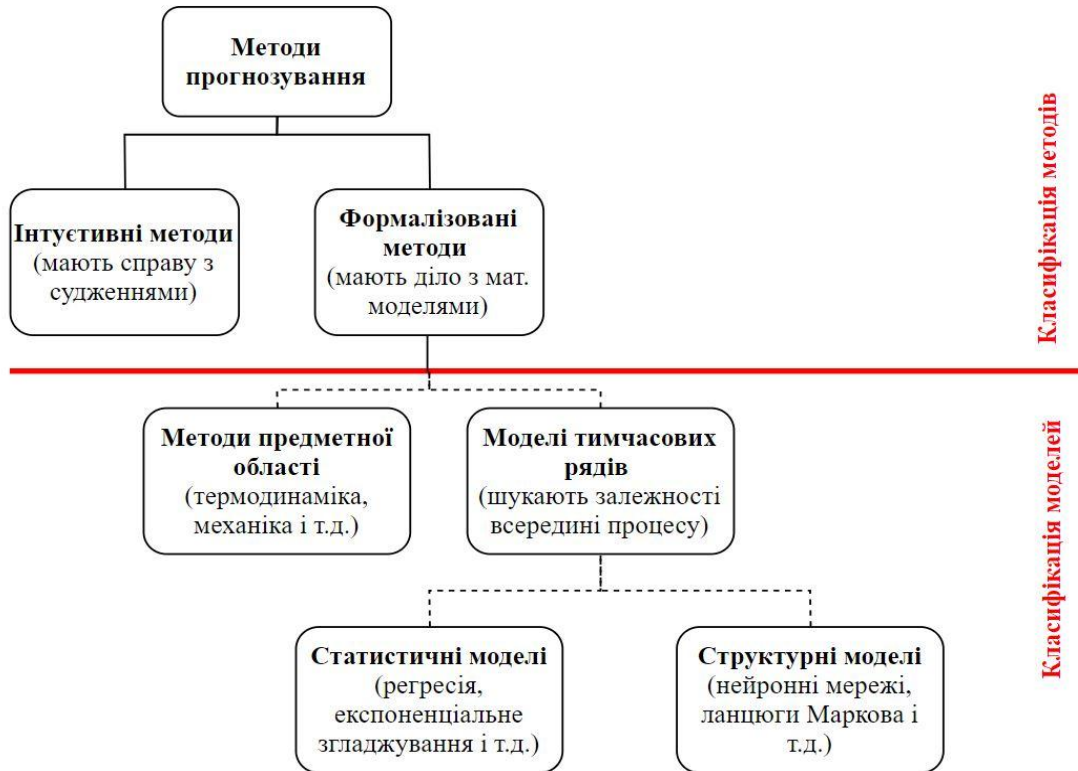
ПРАКТИЧНЕ ЗНАЧЕННЯ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

1. розроблено алгоритм роботи програмного забезпечення прогнозування завантаження банкоматів готівкою на основі технологій OLAP і Data Mining;
2. розроблено програмні засоби для прогнозування завантаження банкоматів готівкою на основі технологій OLAP і Data Mining

Загальна схема процесу прогнозування



Аналіз предметної області прогнозування завантаження банкоматів готівкою



Проаналізувавши всі ці методи, було обрано метод на основі Data Mining.

Вибір і обґрунтування аналогу

ATM Cash Management Solution

CashLab

Головним недоліком обох аналогів є невисока точність прогнозування завантаження банкоматів готівкою

Аналіз процесу постачання банкоматної мережі готівкою

Дані, що використовуються:

дата;
місяць;
частина місяця;
день місяця;
день тижня;
історичні данні по сумах завантаження.

Прогнозоване значення:

загальна сума готівки для завантаження банкомата.

Подання інформації по банкоматах:

щоденне відображення даних;
загальна сума прогнозу готівки.

У даній роботі використовується метод обробки даних, заснований на спільному використанні OLAP і технології Data Mining

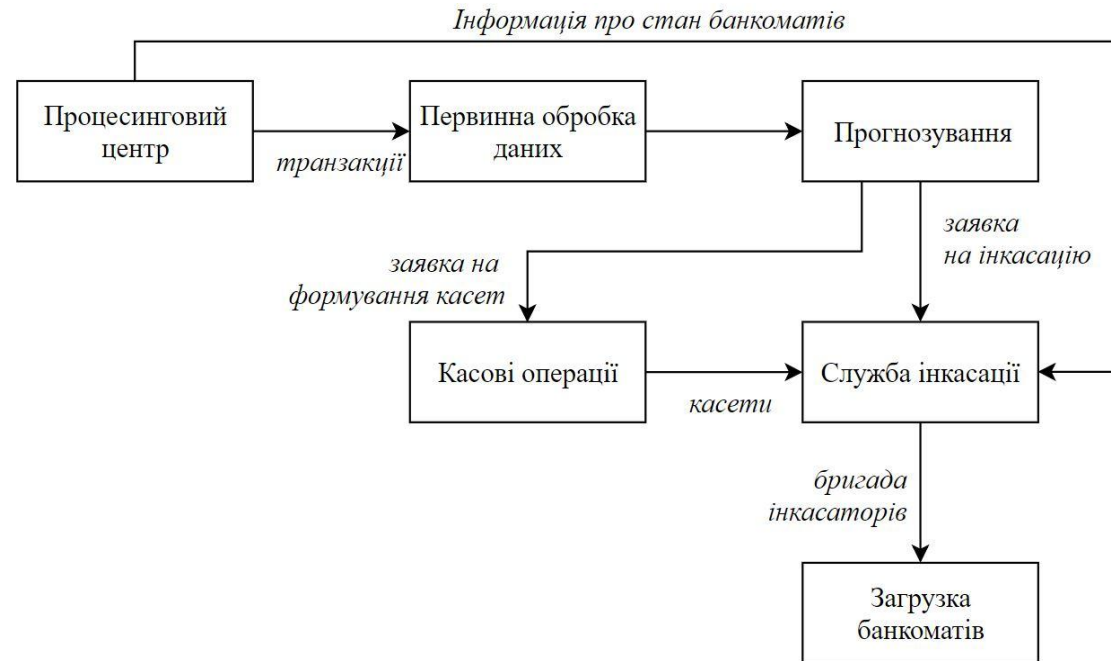


Схема проведення інкасації банкоматів

Розробка технології аналітичної обробки даних

Засобом представлення і багатовимірного аналізу даних є OLAP-куб.

Проектування OLAP-куба проводиться на основі атрибутів таблиць сховища даних в наступній послідовності [9,10]:

- Визначення та узгодження вимірювань. Вимірювання є критеріями для аналізу і визначають багатовимірний простір OLAP-куба.
- За рахунок фіксування значень одного або декількох вимірювань будуються зрізи. Кожен зріз являє собою запит до даних, що включає агрегації.
- Вибір набору заходів. Заходи - це атрибути вимірювання OLAP-куба, зіставляються з атрибутами таблиць сховища даних.
- Збереження результатів попередніх розрахунків в таблиці фактів.

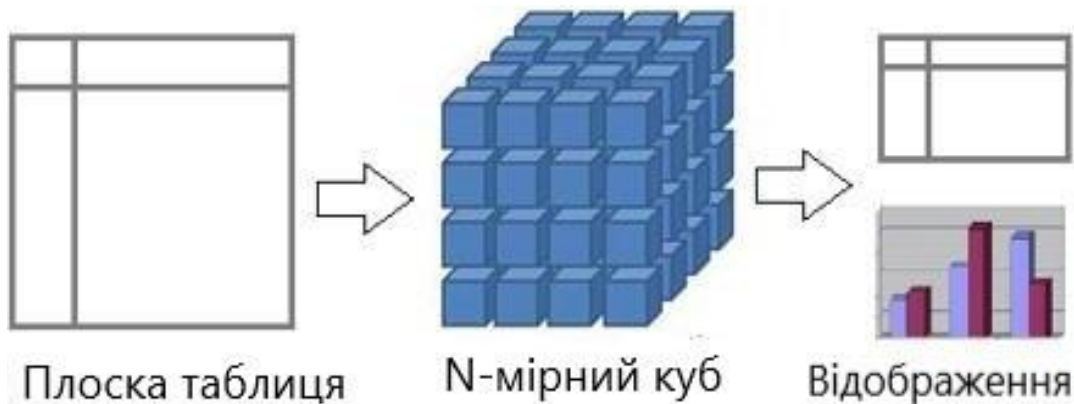


Схема багатовимірного аналізу даних

Прогнозування на основі методу дерев прийняття рішень

Вихідний набір даних має наступний вигляд: $(x, Y) = (x_1, x_2, \dots, x_k, Y)$,

Залежна змінна Y є цільовою неперервною змінною, значення якої необхідно прогнозувати. Вектор x складається з вхідних змінних x_1, x_2, x_3 і т. д., які використовуються для виконання цього завдання.

Неперервні змінні можуть бути спрогнозовані за допомогою формул лінійної регресії, заснованих на зазначених регресорах.

Класична формула лінійної регресії: $Y = a + b * X + e$

(де Y - неперервний атрибут, який прогнозується;

X - регресор; a, b - коефіцієнти регресії; e - випадкова помилка.

Розглянутий алгоритм в кожному вузлі дерева містить формулу лінійної регресії.

Прогнозування на основі методу часових рядів

1. Модель ARIMA(p, d, q):

$$\Delta^d Y_t = c + \sum_{i=1}^p a_i \Delta^d Y_{t-i} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t,$$

2. Модель авторегресії - це модель часових рядів, в якій значення часового ряду в даний момент лінійно залежать від попередніх значень цього ж ряду. Авторегресійний процес порядку p визначається наступним чином:

$$Y_t = \sum_{i=1}^p a_i Y_{t-i} + \varepsilon_t,$$

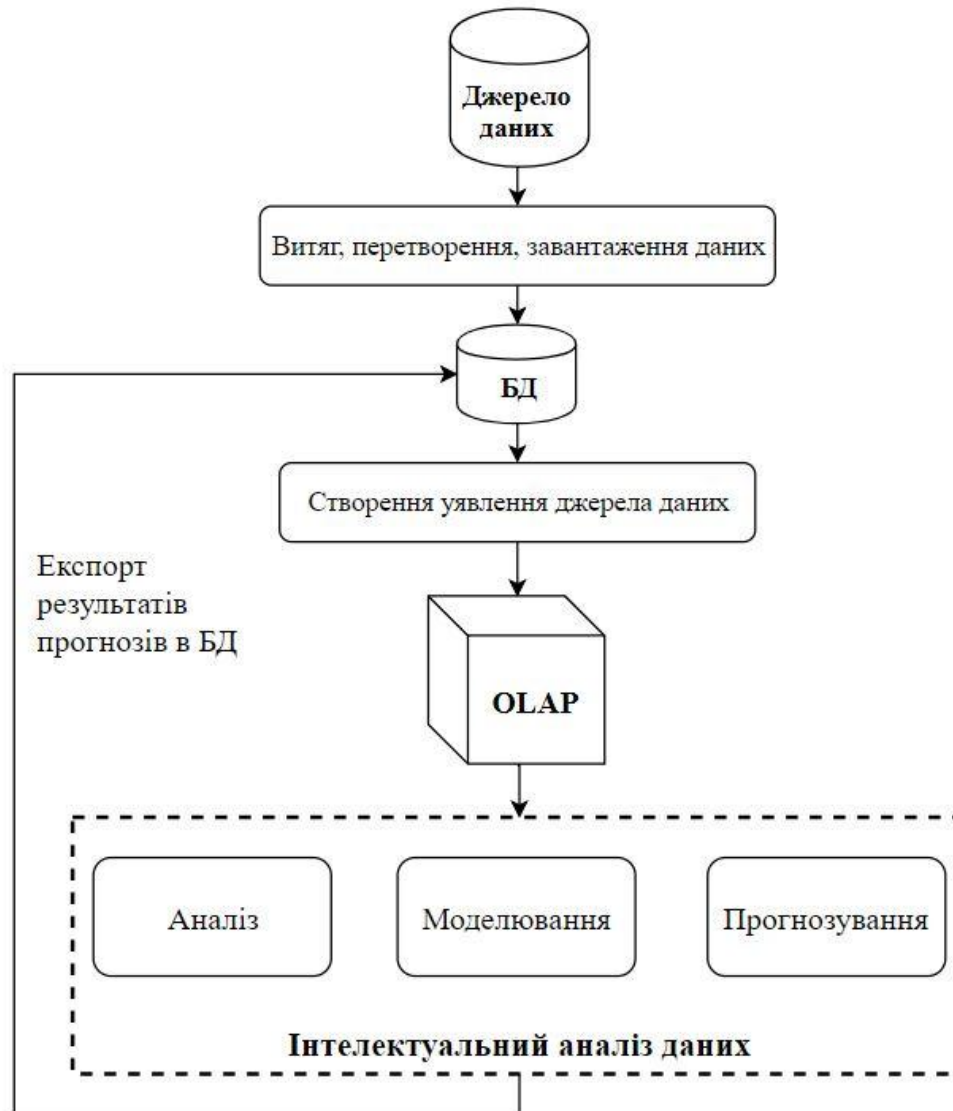
де a_1, a_2, \dots, a_p - коефіцієнти авторегресії; ε_t - випадкова помилка.

3. MA (параметр q) - ковзне середнє. На відміну від процесу авторегресії, в процесі ковзного середнього кожен елемент ряду схильний до сумарного впливу попередніх помилок. Модель ковзного середнього q -го порядку - модель часового ряду такого вигляду:

$$Y_t = \sum_{j=1}^q b_j \varepsilon_{t-j},$$

де b_1, b_2, \dots, b_q - параметри моделі; ε_t - випадкова помилка.

Розробка структури інформаційної технології прогнозування завантаження банкоматів готівкою



Обґрунтування вибору мови та середовища програмування

мови програмування C#
Microsoft Visual Studio

Сховище даних реалізоване за допомогою SQL Server, управляти БД і виконувати запити до неї - мовою T-SQL

Проектування бази даних прогнозування завантаження банкоматів

Проектування бази даних здійснено за допомогою методу «сутність - зв'язок». При проведенні аналізу предметної області виділено такі сутності:

- Банк
- Банкомат
- Інкасація
- Видача готівки

ER-діаграми предметної області

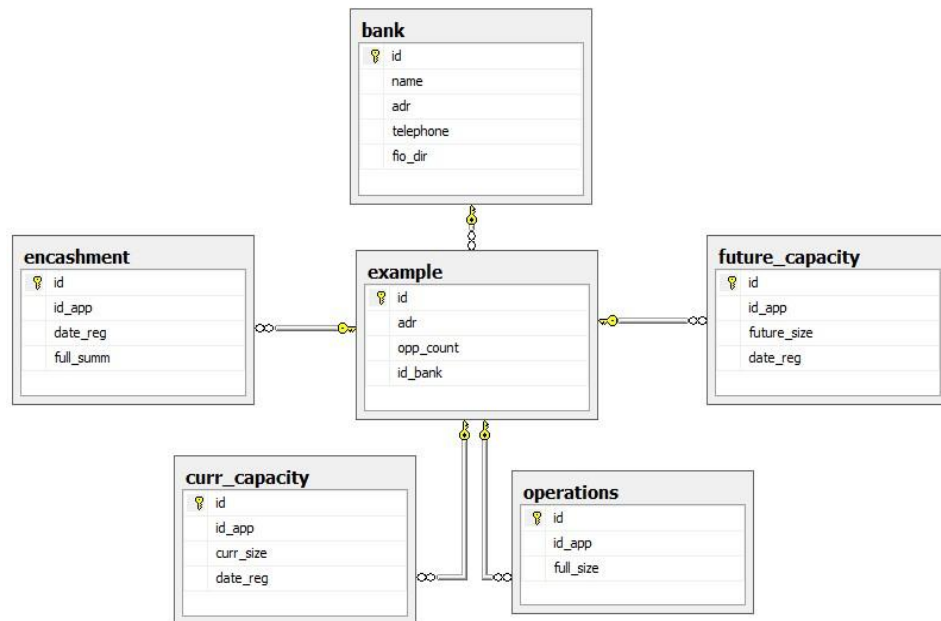
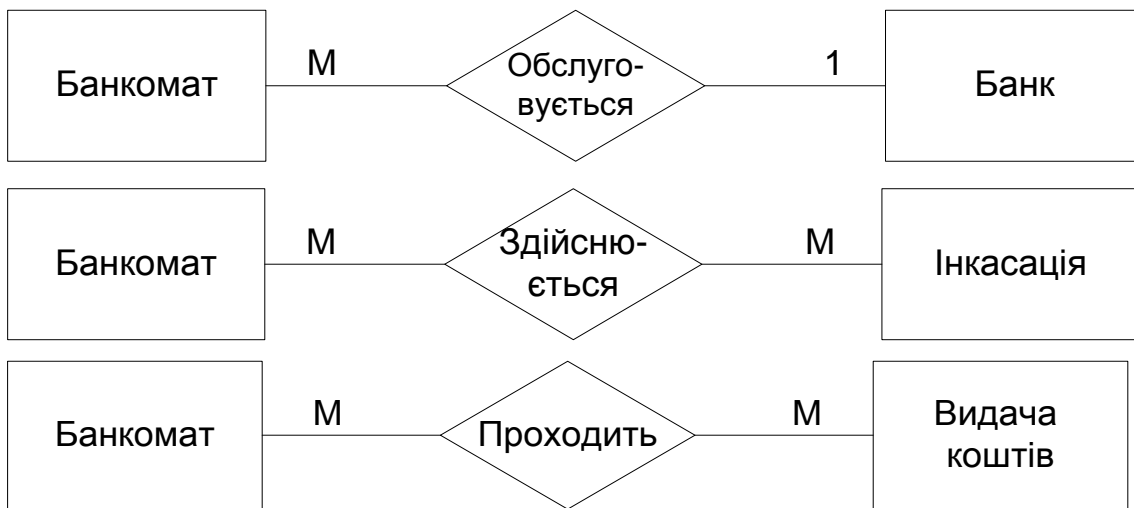
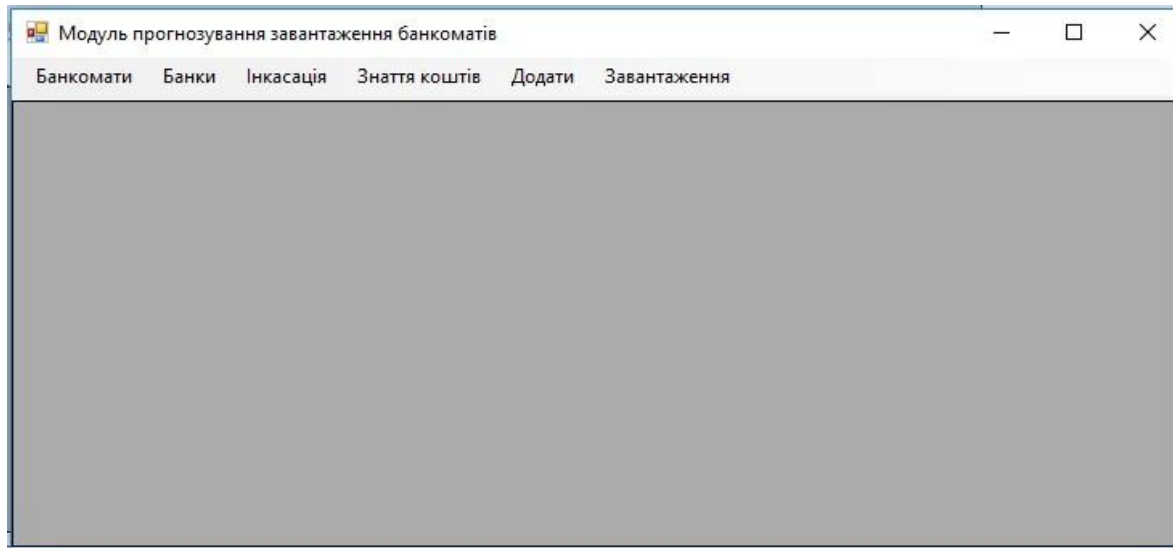


Схема бази даних

Розробка інтерфейсу програми



Вигляд головного вікна

Інкасація

ID банкомата:

Адреса:

Дата: 20 апреля 2018 г.

Сума: 10000

OK

Cash

ID банкомата:

Адреса:

Дата: 20 апреля 2018 г.

Сума: 100

OK

Завантаження

ID банкомата:

Адреса банкомата:

Дата: 20 апреля 2018 г.

Прогнозоване завантаження:

OK

Інтерфейси вікон для введення даних

Проектування OLAP-куба

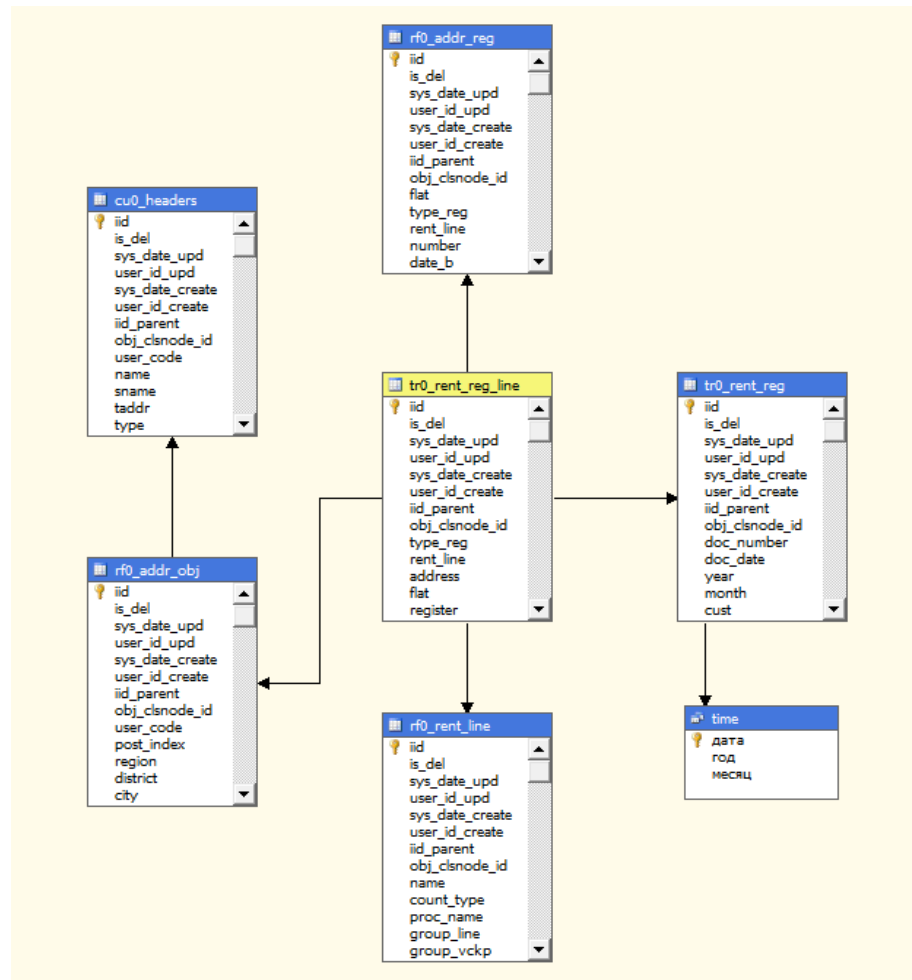


Схема даних OLAP-куба

Тестування роботи програми прогнозування завантаження банкоматів готівкою

Модуль прогнозування завантаження банкоматів

Банкомати Банки Інкасація Знання коштів Додати Завантаження

	id	adr	name	telephone	fio_dir
▶	1	Хмельницьке шоссе, 45	Приват	0672372717	Іванов І.В.
	2	Набережна 34	ПУМБ	0992382761	Петров В.В.
	3	Шевченка 67	Аваль	0675278276	Ткаченко С.О.
	4	Сумська 117	Укрсоцбанк	0633434632	Алісов В.К.
	5	Франка 89	Альфа-банк	0986544324	Калінін С.Д.
*					

Перегляд таблиці «Банк»

Перегляд таблиці «Банкомат»

Модуль прогнозування завантаження банкоматів

Банкомати Банки Інкасація Знання коштів Додати Завантаження

	id	adr	opp_count	id_bank
▶	2	Вишнева 112	500000	1
	3	Межова 87	500000	3
	4	Хмельницьке шоссе 95	700000	2
	5	Набережна 56	750000	2
	6	Коцюбинського 89	500000	4
	8	Хмельницьке шоссе	800000	5
*				

Аналіз результатів роботи програми прогнозування завантаження банкоматів готівкою

Для аналізу результатів було обрано такі показники:

• середня абсолютна помилка (MAE):

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - y_t^*|,$$

де y_t - реальне значення потрібної банкомату готівки в момент часу t ; y_t^* - прогнозоване значення потрібної банкомату готівки в момент часу t , отримане в результаті роботи програми; n - кількість ретроспективних спостережень.
 - коефіцієнт невідповідності КТ, рівний нулю в разі досконалого прогнозу

$$K_T = \sqrt{\frac{\sum_{t=1}^n (y_t - y_t^*)^2}{\sum_{t=1}^n y_t^2 + \sum_{t=1}^n y_t^{*2}}}$$

• час роботи програми.

Із табл. видно, що розроблена програма має меншу помилку MAE ($11,27 < 15,14$; відношення $15,14/11,27=1,34$, тобто на 34% краще) та менше значення коефіцієнта КТ ($0,033 < 0,041$; відношення $0,041/0,033=1,24$, тобто на 24% краще). Це дозволяє стверджувати, що розроблена програма щонайменше на 24% точніша за аналог, а значить мета роботи досягнута – точність прогнозування завантаження банкоматів готівкою підвищена

Програма \ Показник якості	MAE	КТ	T (с)
Розроблена програма	11,27	0,033	13
Аналог (ATM Cash Management Solution)	15,14	0,041	24

ЕКОНОМІЧНА ЧАСТИНА

Було проведено економічне обґрунтування доцільності розробки програми для прогнозування завантаження банкоматів готівкою. Було визначено, що нова розробка має високий рівень комерційного потенціалу, оскільки середньоарифметична сума балів становить 43. Загальна сума витрат на виконання означених робіт склала 27682,61 грн. Прогнозоване значення загальних витрат на виконання та впровадження результатів виконаної наукової роботи – 30698,45 грн. Визначено, що абсолютна ефективність вкладених інвестицій становить 317293,89 грн, і це свідчить про те, що вкладання коштів на виконання та впровадження результатів НДДКР є доцільним. Відносна (щорічна) ефективність вкладених в наукову розробку інвестицій – 124 %, терміну окупності складає 0.8 року (менше 10 місяців). Тобто фінансування розробки програми для прогнозування завантаження банкоматів готівкою є економічно доцільним проектом

АПРОБАЦІЯ РЕЗУЛЬТАТІВ РОБОТИ ТА ПУБЛІКАЦІЇ

Апробація результатів роботи.

Результати досліджень апробовані на конференції «Молодь в науці: дослідження, проблеми, перспективи-2019», Вінниця, 2019, м. Вінниця та «Молодь в науці: дослідження, проблеми, перспективи-2020», Вінниця, 2019, м. Вінниця

Публікації.

За результатами досліджень опубліковано 2 тез доповідей на науково-технічній конференції

ВИСНОВОК

В результаті виконання МКР розроблено інформаційну технологію та програмне забезпечення для прогнозування завантаження банкоматів готівкою на основі технологій Data Mining. Програмне забезпечення створено мовою програмування C# в середовищі розробки MS Visual Studio Community 2017 з використанням для бази даних технології OLEDB . Програма має вищу точність прогнозування, ніж аналогічна програма на 24%. Таким чином, мета роботи досягнута – точність прогнозування завантаження банкоматів готівкою підвищена.