

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ХАРЧОВИХ ТЕХНОЛОГІЙ

Кафедра інформаційних технологій, штучного інтелекту та кібербезпеки

ЛАБОРАТОРНА РОБОТА № 1

з дисципліни «Інтелектуальні системи підтримки прийняття рішень»

на тему: «Знайомство з програмним середовищем інтелектуальної обробки
даних Altair AI RapidMiner»

Виконав: Студент I курсу
групи КН-1-ЗМ

Кучерявий М. В.

Перевірив: _____

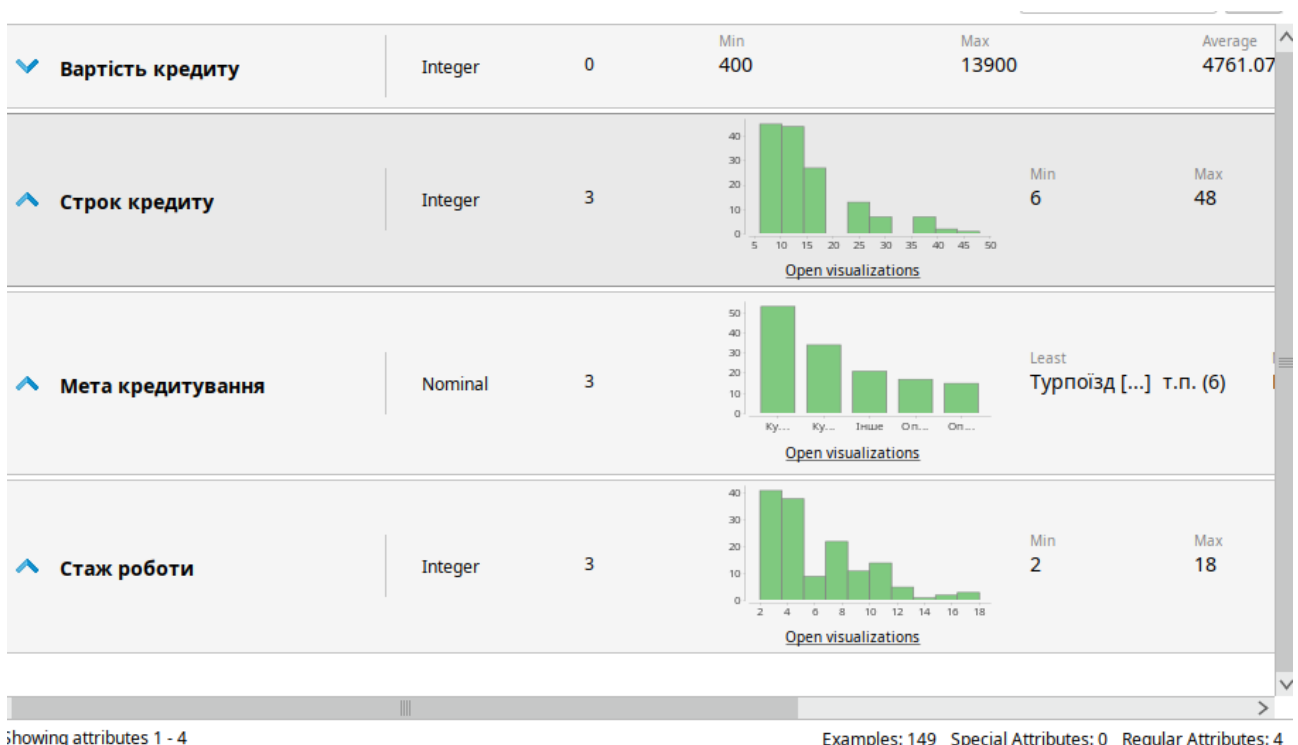
Мета роботи

Набуття навичок створення проєкту Altair AI RapidMiner та підготовки набору даних (очищення даних) для його інтелектуального аналізу.

Хід виконання роботи

Імпортуємо данні з Anketa.csv. Створимо новий процес. Першою дією відфільтруємо атрибути оператором Select Attributes. Виділимо згідно завдання Вартість кредиту, Строк кредиту, Мета кредитування, Стаж роботи.

Згідно з аналізу даних визначено, що деякі поля містять відсутні значення:



Замінемо пропущені: в Стаж роботи замінемо на нульові, в Мета кредитування — на «Інше». В Строк кредиту відфільтруємо значення

Name	Type	Missing	Statistics	Filter (4 / 4 attributes):	Search for Attributes	
▼ Мета кредитування	Polynomial	0	Least Турпоїзд [...] т.п. (6)	Most Купівля товару (51)	Values Купівля	
▼ Стаж роботи	Integer	0	Min 0	Max 18	Average 6.027	
▼ Вартість кредиту	Integer	0	Min 400	Max 13900	Average 4800	
▼ Строк кредиту	Integer	0	Min 6	Max 48	Average 15	

Зменшимо кількість записів оператором Sample

Row No.	Мета кред...	Стаж роботи	Вартість к...	Строк кред...
1	Купівля това...	11	2900	12
2	Турпоїздки, ...	5	2300	6
3	Купівля това...	3	12300	30
4	Оплата посл...	9	2700	12
5	Купівля това...	0	1900	6
6	Купівля това...	8	5500	18
7	Купівля та р...	9	8100	24
8	Турпоїздки, ...	2	1700	6
9	Купівля това...	2	3300	12
10	Купівля та р...	4	11600	48
11	Купівля това...	3	4100	12
12	Інше	4	3700	6
13	Інше	6	5400	18
14	Купівля това...	17	3400	12
15	Купівля това...	3	2100	6

ExampleSet (73 examples,0 special attributes,4 regular attributes)

Відповіді на контрольні питання

1. Які особливості даних?

Особливості даних включають їхню структуру (структуровані, напівструктуровані, неструктуровані), якість (точність, повнота, актуальність), динамічність (статичні, змінні), масштабованість (малі, великі), взаємозв'язки (табличні, графові), безпеку (конфіденційність, шифрування) та формат зберігання (CSV, JSON, SQL, NoSQL). Вони визначають спосіб обробки, аналізу та використання даних у різних системах.

2. Що таке обробка даних?

Будь-які маніпуляції над набором даних, сортування фільтрація, перетворення, аналіз.

3. Які особливості реалізації сучасної обробки даних?

Сучасна обробка даних реалізується на основі використання сховищ даних відповідної інформації, її узагальнення та агрегації та багатовимірного аналізу тобто пошук функціональних і логічних закономірностей у накопиченій інформації, побудова моделей і правил, які пояснюють знайдені аномалії або прогнозують розвиток деяких процесів.

4. Які особливості Data Mining?

Основною ознакою Data Mining є те що отримані результати роботи над даними раніше були невідомі, є нетривіальними, є корисними в практичному застосуванні, можуть інтерпретуватися людиною та являються корисні для вибору рішень в різних галузях.

5. Що таке сценарій?

Це ієрархічна послідовність візуалізації та обробки даних.

6. Яка мета інтелектуальної обробки даних?

виявлення прихованих закономірностей, створення прогнозів, автоматизація прийняття рішень та отримання цінної інформації для оптимізації бізнес-процесів.

7. Що робить оператор ReplaceMissingValues?

Замінює відсутні значення на вказані

8. Що робить оператор Impute Missing Values?

Оператор оцінює значення для відсутніх значень, застосовуючи отриману модель для відсутніх значень.

9. Що робить оператор Declare Missing Value?

На відміну від операторів заміни відсутніх значень, цей оператор встановлює певні значення вибраних атрибутів на відсутні значення.

10. В чому особливість оператора Filter Examples?

Фільтрує значення на основі заданих фільтрів.