НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ   
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»  
ФАКУЛЬТЕТ ІНФОРМАТИКИ ТА ОБЧИСЛЮВАЛЬНОЇ ТЕХНІКИ  
Кафедра автоматизованих систем обробки інформації та управління

**ЗВІТ**

до лабораторної роботи №4

з дисципліни «Інтелектуальний аналіз даних»

Виконав:

студент гр. ІС-32

Капорін Р. М.

Київ – 2016

**ЗАВДАННЯ**

В процессе предоставления кредита банки заинтересованы в изучении платежеспособности будущего потребителя кредита. Цель этого изучения - моделирование или предсказание вероятности, с которой претендент на кредит может быть отнесен к привлекательным или непривлекательным клиентам.

Методы, описанные в этом примере, демонстрируют построение кредитно-скоринговой модели с помощью *STATISTICA Data Miner* - системы для проведения интеллектуального анализа данных, позволяющей идентифицировать входы или предикторы, которые отделяют рискованных клиентов от всех остальных. Предиктивные методы, успешно применяемые на тестовых данных, и в дальнейшем могут быть использованы для предсказания новых рискованных клиентов.



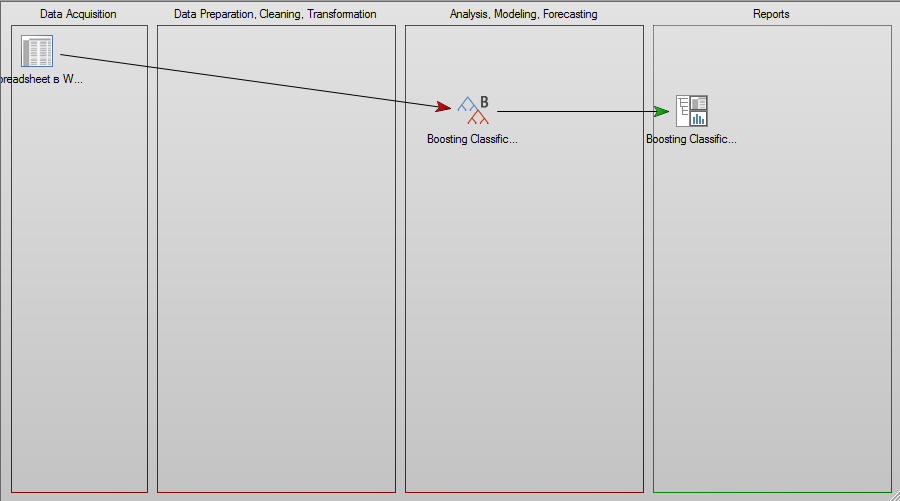


Рис. 1 - Графік значимості предикторів для залежної змінної «кредитоздатність»



Рис. 2 - Приклад структури робочої області DataMiner



Рис. 3 - Дерево рішень CHAID для кредитоздатності

****

Рис. 4 - Матриця класифікації



Рис. 4 - Візуалізація матриці класифікації



Рис. 5 - Карта виграшів для кредитоздатності = погано

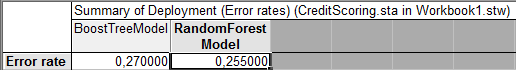
****

Рис. 6 - Загальний результат помилок

**СПОЖИВЧИЙ КРЕДИТНИЙ СКОРИНГ**

**1. Побудова моделі**

Відкриємо CreditScoring2.sta. Застосуємо аналіз нейронних мереж. Основною ціллю буде тип ризику, несуттєвими даними – вік клієнта, а важливими даними – Number of Loans (к-сть займів), How Paid (скільки заплатили), Monthly Income (Місячний прибуток), Income Range (діапазон прибутків), Mortgage (іпотека).

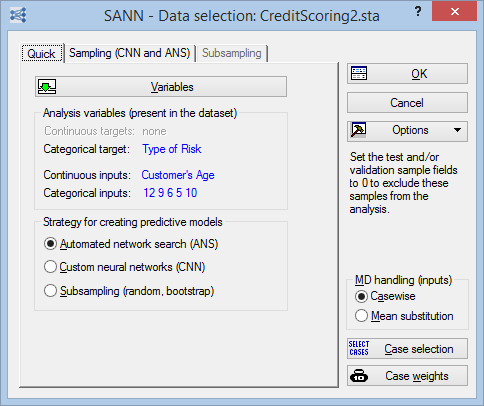


Рис. 1.1 – Налаштування інструменту «Побудова моделі»

**2. Налаштування параметрів**

Задамо значення мінімуму і максимуму нейронів: 2 і 13 відповідно. Вкажемо кількість мереж для навчання – 100. А для збереження – 5.

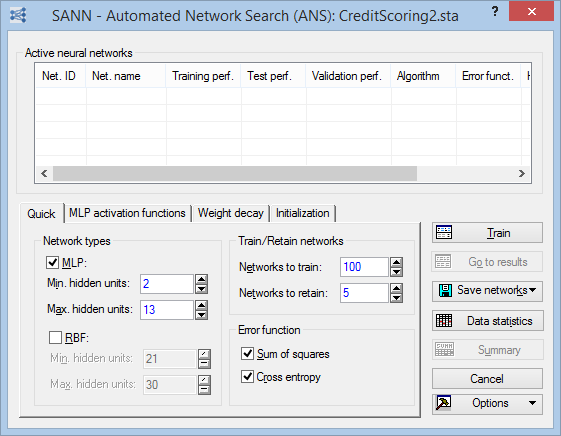


Рис. 2.1 – Завдання параметрів

Отримали результати: точність моделі – 65%.

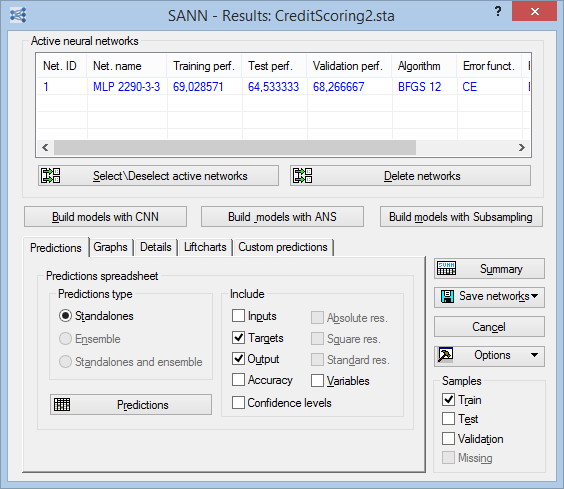


Рис. 2.2 – Отримання результатів з похибками

Отримали наступні прогнози.

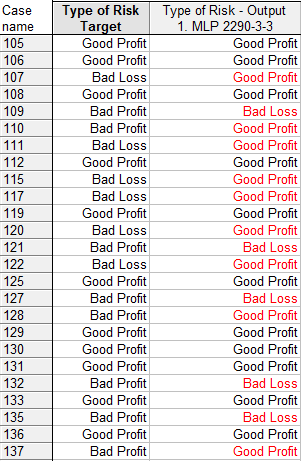


Рис. 2.3 – Прогноз

**3. Виконання аналогічних налаштувань моделі для CreditRisk.sta**

В даному випадку цільовим є Credit Standing. Важливими є статус активів, кредитна історія, кількість місяців вкладу активів, збережені активи.

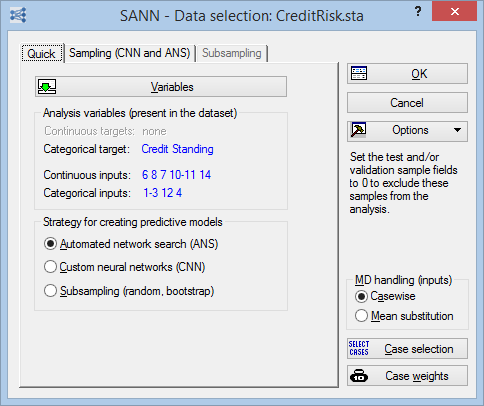


Рис. 3.1 – Завдання параметрів

Отже, точність моделі – 70%.

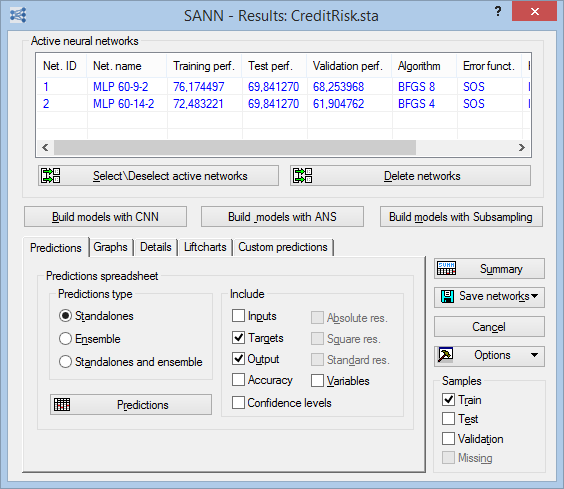


Рис. 3.2 – Отримання результатів з похибками

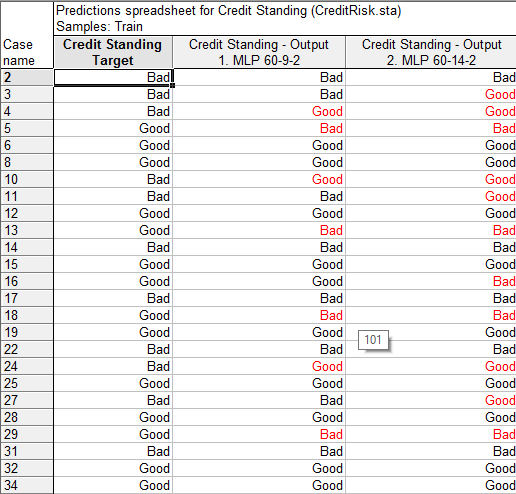


Рис. 3.3 – Прогноз

**ВИСНОВОК**

Під час виконання даної лабораторної роботи я навчилася використовувати інструменти Data Mining та застосувала набуті навички на побудові споживчої кредитної оцінки ризиків (скорингу).