Министерство науки и высшего образования РФ

Федеральное государственное бюджетное

образовательное учреждение высшего образования

ВЯТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет автоматики и вычислительной техники

Отчет по лабораторной работе №3.1

**Исследование временных рядов с помощью среды R. Несбалансированные данные.**дисциплина «Теория и применение методов прогнозирования»

|  |  |
| --- | --- |
| Выполнил: студент группы ИВТм-1301 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_/ Савин Д.А. / |
|  |  |
| Проверил: к.т.н., доцент кафедры ЭВМ | \_\_\_\_\_\_\_\_\_\_\_\_\_\_/ Коржавина А.С. / |

Киров 2022

##### Практическое задание

1. Провести 10 опытов по определению «хороших» и «плохих» клиентов на различных случайных выборках. Использовать данные cs-test.csv из предыдущих работ. Результат представить в таблице:

|  |  |  |
| --- | --- | --- |
| № выборки (опыта) | Количество «плохих» клиентов | % «плохих» клиентов |
| 1 | 344 | 7,043 |
| 2 | 694 | 6,94 |
| 3 | 1285 | 6,425 |
| 4 | 1959 | 6,53 |
| 5 | 2755 | 6,8875 |
| 6 | 3377 | 6,754 |
| 7 | 4044 | 6,74 |
| 8 | 4622 | 6,6 |
| 9 | 5223 | 6,52 |
| 10 | 6024 | 6,69 |

1. Выполнить метод случайных сбалансированных подмножеств с использованием паспортов скользящего контроля.

Провести ряд опытов с различными значениями параметра alpha (0.75,0.8, 0.85,0.9, 0.95).

|  |  |  |  |
| --- | --- | --- | --- |
| alpha | размер подмножества | общее кол-во валидационных данных | AUC, соответствующий подмножеству |
| 0.75 | 9400 | 70600 | 0.8598 |
| 0.8 | 9469 | 70531 | 0.8599 |
| 0.85 | 9845 | 70155 | 0.8595 |
| 0.9 | 10295 | 69705 | 0.8599 |
| 0.95 | 10483 | 69517 | 0.8595 |

1. Построить модели нейронной сети с числом нейронов скрытого слоя 3, 10, 20. Вывести графики.

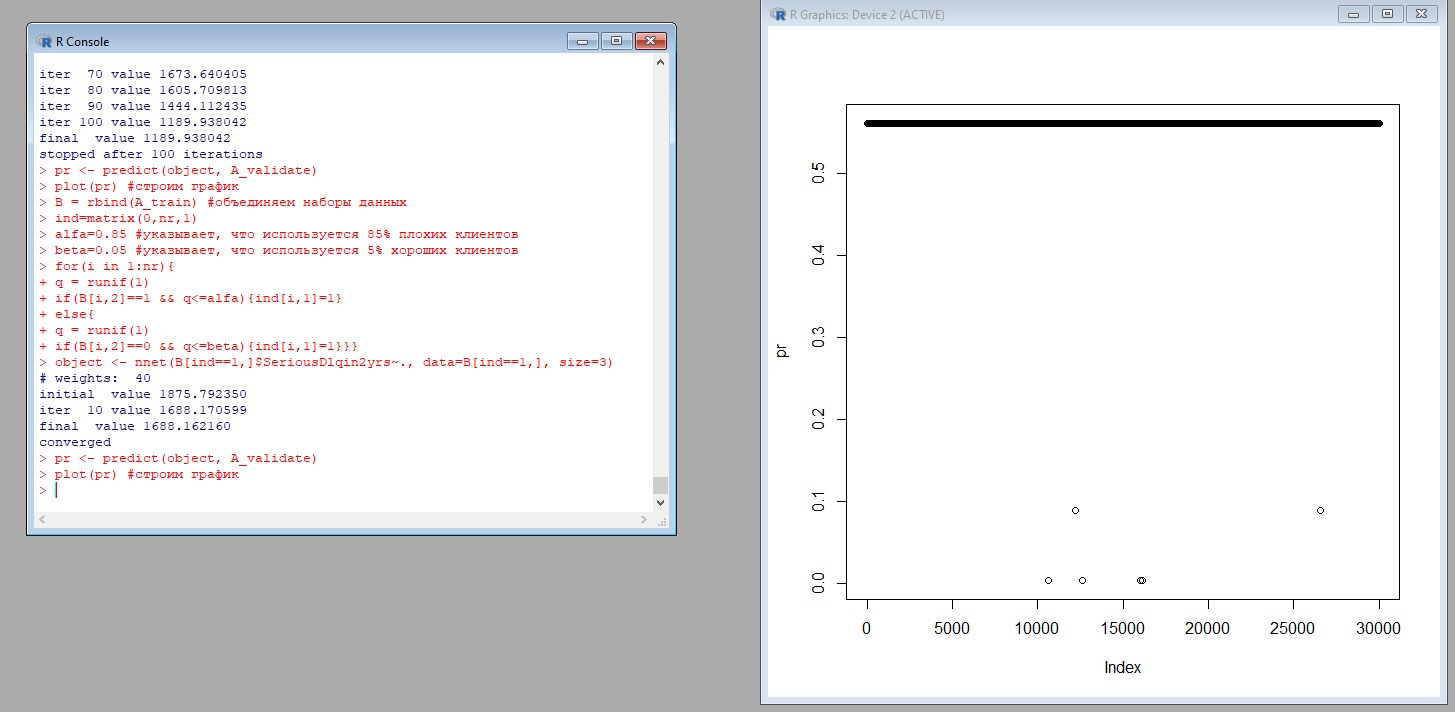


Рисунок 1 – Число нейронов 3.

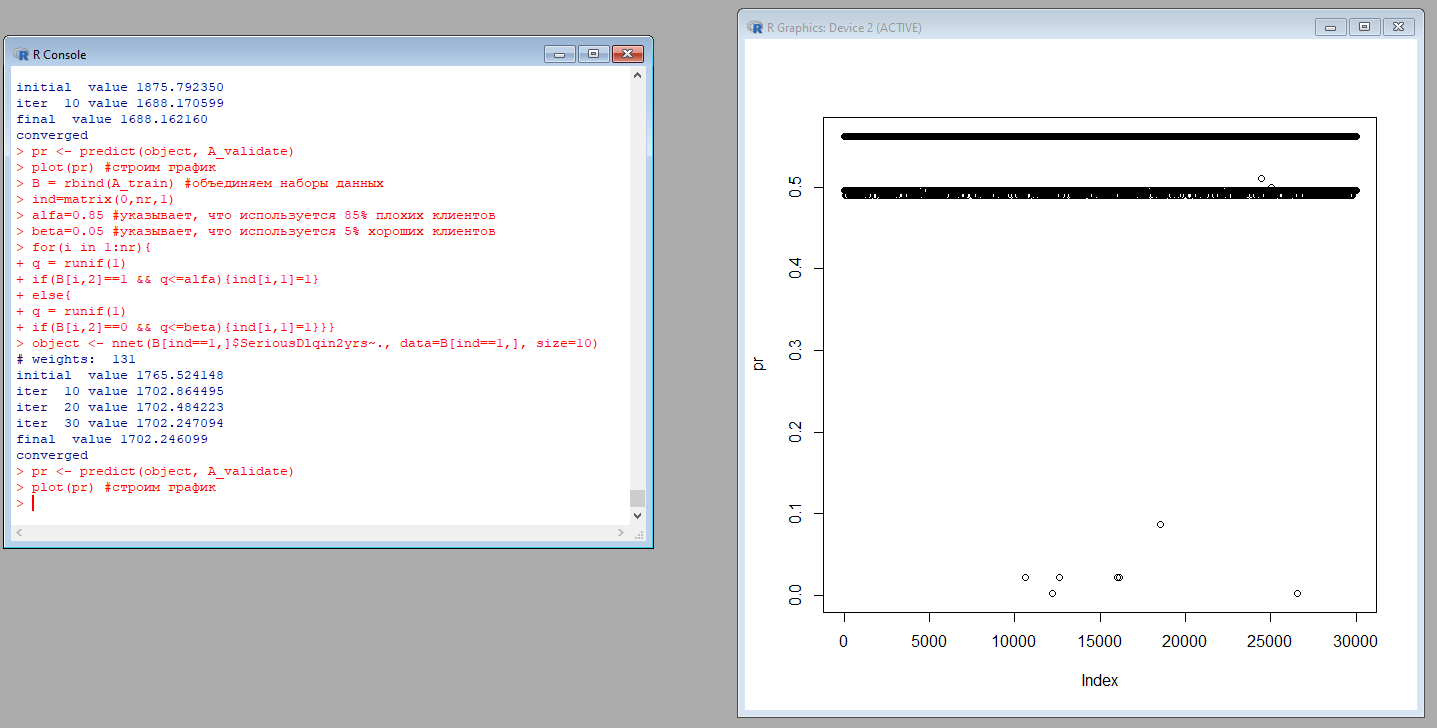


Рисунок 2 – Число нейронов 10.

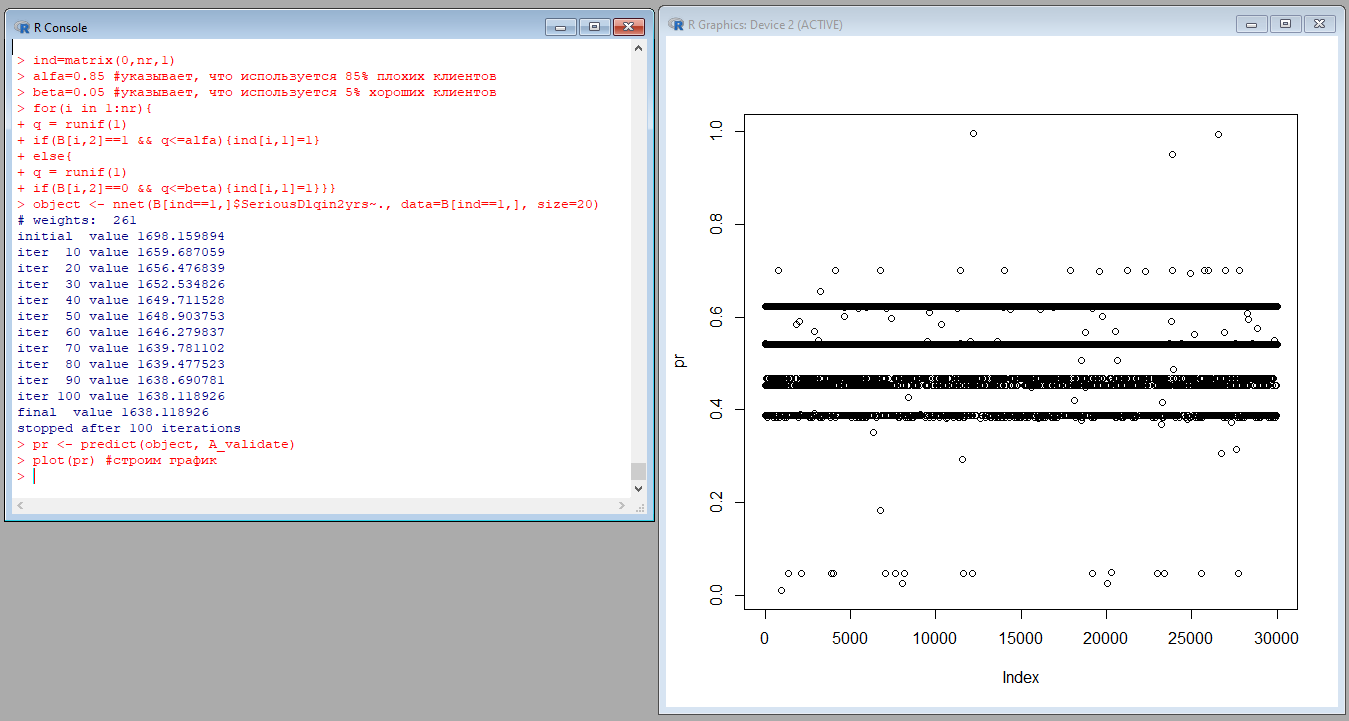


Рисунок 3 – Число нейронов 20.

4) Определить незначимые переменные для данных Credits. Сравнить AUC построенной модели и полной модели.

AUC полной модели = 0.5422

AUC построенной модели = 0.6832

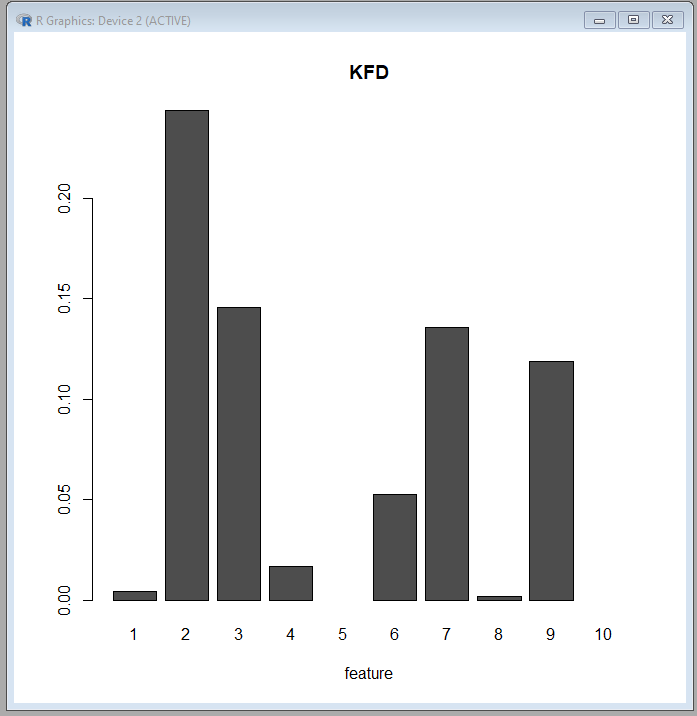


Рисунок 4 – Значимые и незначимые переменные.

**Вывод**

Были изучены временные ряды, несбалансированные данные, а также использование нейронных сетей. Были построены нейронные сети с разным количеством скрытых нейронов, а также с разным значением alpha.