МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение

высшего образования

«УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра «Информационные системы»

Дисциплина «Методы искусственного интеллекта для принятия кредитных решений»

**Лабораторная работа № 1**

Анализ источников данных

Выполнил:

студент гр. ИИмд-21

Шевалдов С.С

Проверил:

Доцент, к.т.н., доцент

Романов А.А.

Ульяновск,   
2023 г.

**Цель работы**

Цель работы: охарактеризовать датасет, содержащий данные о заёмщиках.

**Задачи работы**

Задачи работы:

1. Осуществить поиск наборов данных с информацией по заёмщикам в открытых источниках.
2. Провести анализ найденных данных.
3. Охарактеризовать полученные результаты.

**Порядок выполнения работы**

В рамках выполнения данной лабораторной работы был проанализирован датасета, содержащий в себе информацию о более чем 4,5 тысяч заемщиков. Данные датасет содержал в себе много характеристик заемщиков, такие как:

1. Возраст(age). Диапазон имеет нижнюю границу в 18 лет;
2. Сфера деятельности(job). Возможные значения – торговля, образование, управление и т.д.;
3. Информацию о семейном положении (marital). Варианты: холост, женат, разведен;
4. Уровень образования (education). Варианты: основное, среднее, высшее.
5. Информацию о наличии дефолта у заемщика (default). Варианты – да, нет;
6. Баланс счета(balance). Может быть отрицательным;
7. Информацию об имеющемся жилье (housing). Варианты – есть, нет;
8. Информацию о принятом решении(poutcome). Варианты – неизвестно, успешно, неуспешно;
9. Сведения о способе связи с заёмщиком(contact). Варианты – телефон, почта, нет;
10. Сведения о наличии кредита у заёмщика(loan). Варианты – да, нет.
11. Информация о времени с последнего кредита в месяцах(previous). Число больше нуля.

На рисунке 1 представлен набор столбцов в анализируемом датасете.



Рис.1 Набор столбцов в анализируемом датасете.

На основе этих данных представляется возможным сделать предположение о возможности выдачи кредита заёмщику, или же отказать в кредите, если велик риск просрочки. Основывая на простой логике, можно предположить, что заёмщик с возрастом 18 лет, с базовым образованием и маленьким личным счетом в банке имеет большую вероятность невыплаты кредита по сравнения с заёмщиком 33 лет, кандидатом наук, с большим личным счетом.

**Выводы по работе**

Основываясь на предоставленных в датасете критериях можно сделать предположения о вероятности невыплаты кредита заёмщиком и на основе этого принять решение о выдаче кредита или отказе в выдаче.

**Список использованных источников**

1. Bank credit scoring | Kaggel // Kaggle URL: https://www.kaggle.com/datasets/kapturovalexander/bank-credit-scoring (дата обращения: 06.09.2023).