



ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И МАШИННОЕ ОБУЧЕНИЕ

Д.И. Алексеев, Д.С. Оплачко

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ОЦЕНКИ КРЕДИТОСПОСОБНОСТИ С ПРИМЕНЕНИЕМ НЕЙРОННОЙ СЕТИ

(Самарский университет)

В настоящее время кредиты – это привычное явление нашей жизни. По статистике, кредиты сегодня есть у каждого третьего гражданина России. Наиболее часто граждане берут кредиты для решения таких задач, как: покупка бытовой техники, транспортных средств; приобретение недвижимости; оплата дорогостоящего лечения; получение платного образования.

Для минимизации рисков банка или кредитной организации, связанных с невыплатой кредита в срок, производится анализ и оценка кредитоспособности клиента [1]. Кредитный скоринг – используемая банками система оценки заемщиков, в основе которой заложены статистические методы. Скоринговая система получила свое название от английского слова «score», что переводится как счёт или подсчёт очков. В работе таких систем может применяться анкетирование клиентов, подключение к базам данных бюро кредитных историй, государственных организаций, социальных сетей, а также отслеживание поведенческих факторов.

Кредитный скоринг широко используется как крупными банками, микрофинансовыми организациями, так и в потребительском (магазинном) экспресс-кредитовании на небольшие суммы.

Существует множество разных методов, однако применение нейронных сетей может сделать этот процесс более эффективным.

Структура нейросетевого классификатора кредитного скоринга

Система классификации кредитоспособности заемщиков состоит из частей, отвечающих за следующую функциональность: сбор данных, обработка данных и нормализация, нейросетевой классификатор, схема представлена на рисунке 1. На вход системы поступают данные – критерии оценки, полученные из их анкет: кредитный статус (наличие кредитного обязательства в настоящий момент); возраст; семейное положение; работа; стаж работы; недвижимость; кредитная история; доходы; расходы. Данные критерии выделены мировой и отечественной банковской практикой. На выход поступает результат классификации заемщика.

Для решения задачи приведения данных клиентов к основной форме используется метод нормализации минимакс, при котором линейное преобразо-



вание данных находится в диапазоне $[0...1]$, где минимальное и максимальное масштабируемые значения соответствуют 0 и 1 соответственно [2].

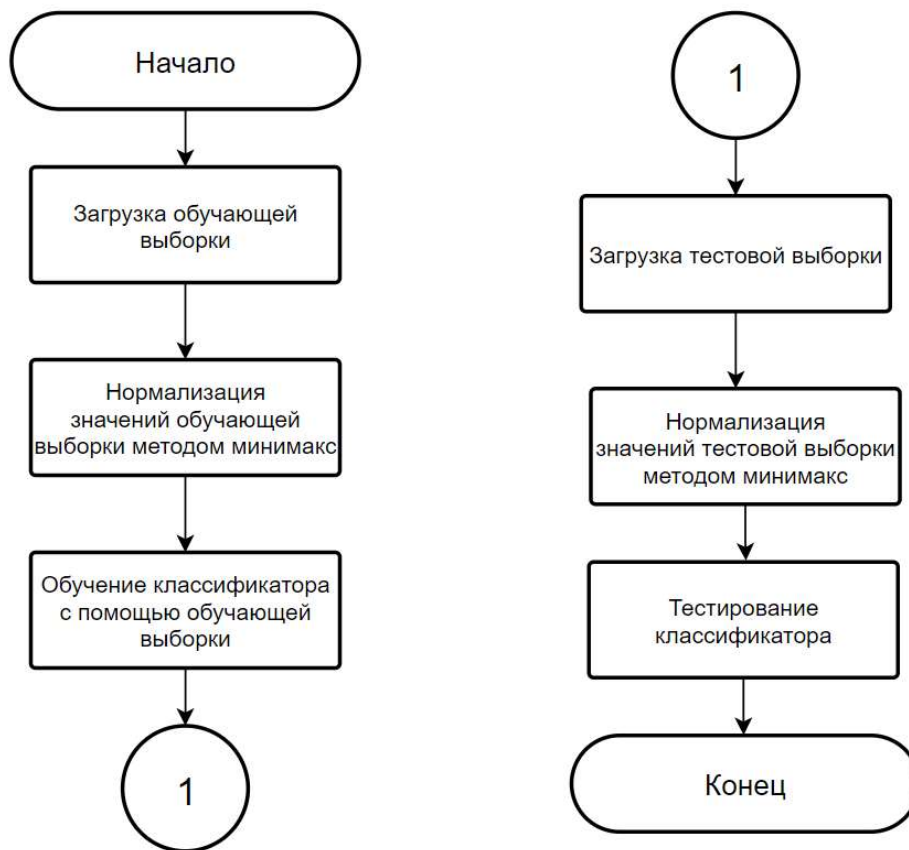


Рис. 1. Схема нейросетевого классификатора

Второй частью классификатора является нейронная сеть, которая классифицирует вектор характеристик критериев заёмщиков, т.е. разделяет все пространство признаков на определенное количество областей.

В качестве архитектуры нейронной сети была выбрана схема многослойного персептрона. Такая сеть состоит из множества наборов нейронов, называемых слоями. Множество входных узлов называют входным слоем сети, при этом сигнал движется от слоя к слою в прямом направлении.

На рисунке 2 представлен пример многослойного персептрона с двумя скрытыми слоями и с сигмоидами в качестве функций активации.

Число входных нейронов равно количеству критериев оценки, а число выходных – количеству классов, число скрытых слоёв и нейронов в них устанавливается при создании сети. Для обучения сети использовался метод обратного распространения ошибки [3].

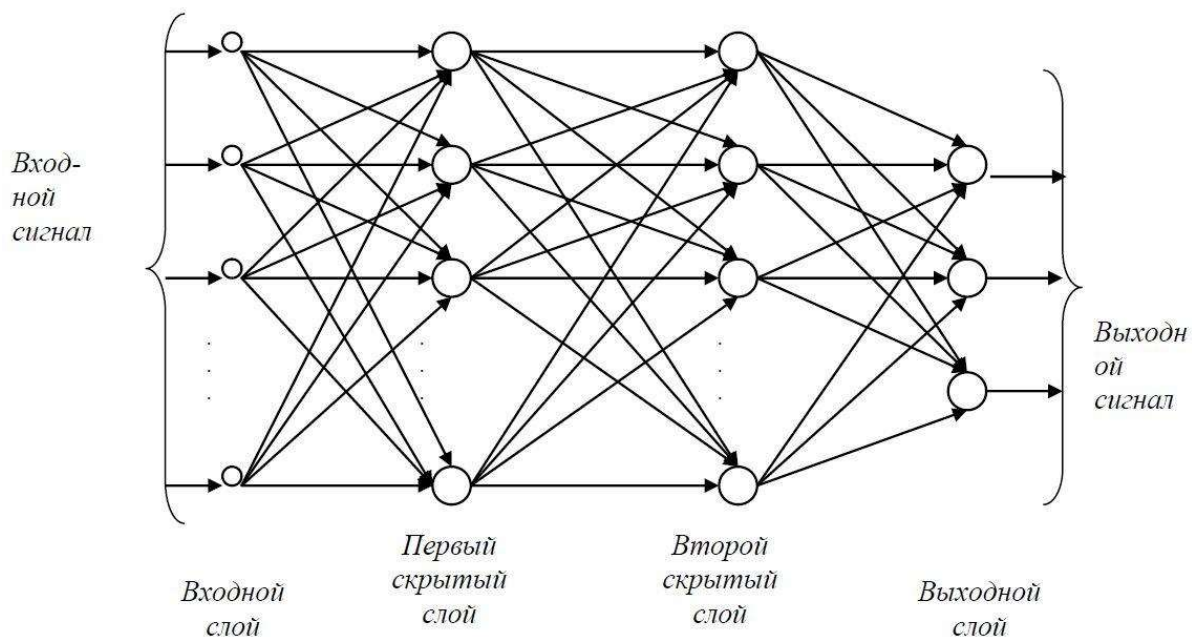


Рис. 2. Схема многослойного персептрона с двумя скрытыми слоями

Алгоритм обратного распространения ошибки основан на идее распространения сигналов ошибки от выходов нейросети к ее входам, в направлении, обратном прямому распространению сигналов в обычном режиме работы. Согласно методу наименьших квадратов, минимизируемой целевой функцией ошибки нейросети является величина:

$$E(w) = \frac{1}{2} \sum_{j,p} (y_{j,p}^{(N)} - d_{j,p})^2,$$

где $y_{j,p}^{(N)}$ – реальное выходное состояние нейрона j выходного слоя N нейронной сети при подаче на ее входы p -го образа, а $d_{j,p}$ – идеальное (желаемое) выходное состояние этого нейрона.

Суммирование ведется по всем нейронам выходного слоя и по всем обрабатываемым сетью образам. Минимизация ведется методом градиентного спуска [4].

Заключение

Разработанная система была апробирована на обучающей выборке из анкетных данных 2500 заёмщиков. На основании этих данных были определены оптимальные значения параметров нейронной сети, количество нейронов в скрытом слое, при которых максимальная точность классификации составляет 87%.

Литература

- 1 Оценка кредитоспособности заемщика [Электронный ресурс]: – URL: https://exbico.ru/metodika_otsenki_kreditosposobnosti_zaimshchika (дата обращения: 19.10.2020).
- 2 Нормализация на практике – методы и средства data preparation [Электронный ресурс]: <https://www.bigdataschool.ru/blog/нормализация-методы->



средства-data-preparation.html (дата обращения: 05.03.2021).

3 Осовский С. Нейронные сети для обработки информации [Текст]/ Осовский С. – М.: Научно-техническое издательство «Горячая линия–Телеком», 2017. 448 с.

4 Солдатова, О.П. Основы нейроинформатики. Самара, СГАУ, 2006. 131 с.

А.К. Алимуратов, П.П. Чураков

ПРИМЕНЕНИЕ МЕТОДА ДЕКОМПОЗИЦИИ НА ЭМПИРИЧЕСКИЕ МОДЫ ДЛЯ ИССЛЕДОВАНИЯ ВОКАЛИЗОВАННОЙ РЕЧИ В ЗАДАЧЕ ОБНАРУЖЕНИЯ СТРЕССОВЫХ ЭМОЦИЙ ЧЕЛОВЕКА

(Пензенский государственный университет)

Точность распознавания эмоций человека зависит от корректного выделения информативных параметров, отражающих эмоциональную составляющую речи. Анализ открытых источников отечественной и зарубежной литературы [1] выявил, что вопрос разработки высокоэффективного подходом выделения информативных параметров речи, релевантных нарушениям работы органов речевого аппарата остаются нерешенным и требует дальнейшей проработки. Данная статья посвящена исследованию вокализованной речи с применением уникальной адаптивной технологии анализа нестационарных сигналов - декомпозиции на эмпирические моды (ДЭМ) для обнаружения стрессовых эмоций человека [2].

Речь представляет собой нестационарный акустический сигнал сложной формы. Различные изменения в вегетативной нервной системе могут изменить речь человека. Речь человека состоит из вокализованных / невокализованных участков, участков пауз и дыхания. Информативные параметры вокализованной речи в полном объеме отражают нарушение работы органов речевого аппарата, вследствие эмоционального возбуждения человека. Вокализованная речь образуется в результате возбуждения голосового тракта, обусловленного вибрацией голосовых связок в области голосовой щели (глоттиса). Сила возбуждения во время глоттальной активности определяется в основном скоростью смыкания голосовых связок. Периодические колебания голосовых связок во время возбуждения голосового тракта называется основным тоном (ОТ). Величина обратная значению ОТ называется частотой основного тона (ЧОТ) и является важным информативным параметром речи.

ДЭМ представляет собой уникальную адаптивную технологию анализа нестационарных данных, не требующую никакой априорной информации для разложения сигнала на частотные составляющие [3]. Адаптивность ДЭМ позволяет эффективно применять ее для анализа естественных сигналов. Разложение с помощью ДЭМ обеспечивает извлечение из сигнала различных колебательных функций, называемых эмпирическими модами (ЭМ), каждая из кото-