



И.А. Лёзин, С.П. Канабеев

ОПТИМИЗАЦИЯ ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ МЕТОДОМ КОМБИНАЦИИ АЛГОРИТМОВ ОБУЧЕНИЯ

(Самарский национальный исследовательский университет
имени академика С.П. Королёва)

Одно из главных преимуществ нейронных сетей перед традиционными алгоритмами является их способность к обучению. Обучение реализуется с помощью специально разработанных алгоритмов, и, технически, заключается в нахождении таких состояний весов нейронов, которые минимизируют выходную ошибку сети на обучающем и тестовом множествах. Задачи поиска оптимального состояния систем по минимуму или максимуму некоторой функции образуют отдельный класс и называются задачами оптимизации. В случае нейронной сети оптимизация ведется по функции ошибки, зависящей от состояния весов нейронов. Состояние системы, при котором обеспечивается минимум целевой функции (в данном случае — выходной ошибки сети), называется оптимальным. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. В случае успешного обучения сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке, а также неполных и/или «зашумленных», частично искаженных данных. Для обучения нейронной сети не требуется никакой априорной информации о структуре искомой функциональной зависимости, нужна лишь обучающая выборка в виде экспериментальных пар <входы-выходы> [1].

В качестве нейронной сети для обучения был выбран многослойный персептрон с одним скрытым слоем. Многослойный персептрон широко используется для поиска закономерностей и классификации образов. Нейронные сети, которые содержат более чем один нейрон в скрытом слое, имеют не одно, а несколько оптимальных состояний, обусловленных наличием у функции ошибки нескольких минимумов. Однако истинно оптимальным состоянием является только одно, где обеспечивается *глобальный минимум*. Остальные минимумы функции ошибки называются *локальными*.

Для преодоления локального, а также для достижения глобального минимума целевой функции наиболее быстрым способом многослойный персептрон обучается параллельно несколькими алгоритмами. Фактически, обучаются несколько нейронных сетей с одинаковой структурой. Если обучение еще не завершено, и глобальный минимум еще не был достигнут, на определенных стадиях (достаточное приближение к необходимой ошибке, фиксированное количество эпох обучения, время обучение), происходит подстройка коэффициентов связей нейронов обучаемых сетей, к коэффициентам той сети, у которой ошибка на данном этапе минимальна. Каждая из сетей продолжает обучение, но уже с новыми весовыми коэффициентами связи нейронов. За счёт принудительного



изменения таких коэффициентов предотвращаются возможные остановки в локальных минимумах и достигается более оптимальный процесс обучения нейронной сети.

В работе были использованы следующие алгоритмы: обратного распространения ошибки, имитации отжига, генетический алгоритм. Алгоритм обратного распространения ошибки – это итеративный градиентный алгоритм, который используется с целью минимизации ошибки работы многослойного персептрона и получения желаемого выхода. Основу алгоритма составляет целевая функция, формулируемая в виде квадратичной суммы разностей между фактическими y и ожидаемыми d значениями выходных сигналов [1]:

$$E(w) = \frac{1}{2} \sum_{j=1}^m (y_j - d_j)^2.$$

Алгоритм имитации отжига – общий алгоритмический метод решения задачи глобальной оптимизации, особенно дискретной и комбинаторной оптимизации.

Он основывается на имитации физического процесса, который происходит при кристаллизации вещества, в том числе при отжиге металлов. Предполагается, что атомы уже выстроились в кристаллическую решетку, но еще допустимы переходы отдельных атомов из одной ячейки в другую, и процесс протекает при постепенно понижающейся температуре. Переход атома из одной ячейки в другую происходит с некоторой вероятностью, причем вероятность уменьшается с понижением температуры. Устойчивая кристаллическая решетка соответствует минимуму энергии атомов, поэтому атом либо переходит в состояние с меньшим уровнем энергии, либо остается на месте [2].

$$p_a = \begin{cases} e^{-\frac{\Delta f}{T}}, & \Delta f > 0, \\ 1, & \Delta f < 0 \end{cases}$$

где $\Delta f = f(s_{new}) - f(s_{old})$, $T = f_T(T_0, n)$, а функция понижения температуры $f_T(T_0, n) = \frac{T_0}{1+n}$.

Генетический алгоритм – это стохастический алгоритм, используемый для оптимизации параметров сети путём случайного подбора, комбинирования и вариации искоемых параметров с использованием механизмов, аналогичных естественному отбору в природе [3].

Исследования проводились на данных, взятых из репозитория UCI Калифорнийского университета, для задач классификации винных напитков [4].

По результатам проведенных исследований видно, что при подстройке весовых коэффициентов связей нейронов сетей к коэффициентам сети, у которой ошибка на выходе на данной итерации обучения наименьшая, ошибка на последующих итерациях обучения этих сетей уменьшается. Можно сделать вывод о пригодности обучения методом комбинации алгоритмов. Хотя такой способ требует больших вычислительных затрат, он позволяет избежать попадания в локальные минимумы целевой функции и в общем случае достичь глобального минимума за меньшее время.

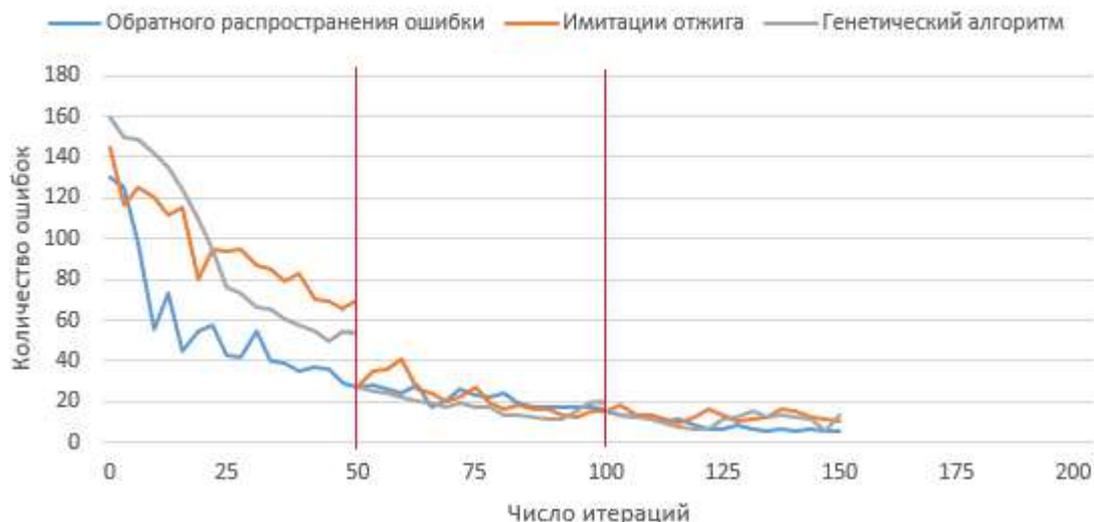


Рис. 1. Обучение нейронной сети методом комбинации алгоритмов обучения

Литература

- 1 Осовский С. Нейронные сети для обработки информации [пер. с польского И.Д.Рудинского] [Текст] –М.: издательский дом «Финансы и статистика», 2002. -344 с.
- 2 Алгоритм имитации отжига [Электронный ресурс]. – URL: <http://habrahabr.ru/post/112189> (Дата обращения: 11.10.2016)
- 3 Ротштейн, А.П. Интеллектуальные технологии идентификации: нечеткая логика, генетические алгоритмы, нейронные сети [Текст] — Винница: УНИВЕРСУМ-Винница, 1999. -320 с.
- 4 .UCI Machine Learning Repository [Электронный ресурс]. – URL: <http://archive.ics.uci.edu/ml/> (Дата обращения: 14.11.2016)

Н.И. Лиманова, М.Н. Седов

АЛГОРИТМ АВТОМАТИЗИРОВАННОГО ПОИСКА ПЕРСОНАЛЬНЫХ ДАННЫХ НА ОСНОВЕ МЕТРИКИ ЛЕВЕНШТЕЙНА

(Поволжский государственный университет
телекоммуникаций и информатики)

Введение

В связи с отсутствием единых реестров и баз данных о гражданах РФ на региональных и муниципальных уровнях различным уполномоченным учреждениям при обмене персональными данными приходится затрачивать дополнительные ресурсы на сопоставление одного набора данных другому. Сложность данной задачи прямо пропорционально зависит от количества записей в каждом из наборов, что в случае прямого и/или ручного сравнения приводит к невозможности решения данной проблемы с приемлемым уровнем качества. Для однозначного результата необходимо выполнять автоматизированный поиск рек-