ПРИМЕНЕНИЕ ТЕОРИИ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ В РЕШЕНИИ ЗАДАЧИ ПОИСКА ГЛОБАЛЬНОГО МИНИМУМА ФУНКЦИИ ОШИБКИ

Марьина О.А.
ГОУВПО «Мордовский государственный университет им. Н. П. Огарева», г. Саранск
Тел. +7 (927) 644-69-87, E-mail: fd_oxy@mail.ru

Аннотация. Сравниваются методы обучения многослойного персептрона, и обосновывается целесообразность применения вейвлет-преобразований в рамках решения поставленной задачи.
Ключевые слова: многослойный персептрон, обучение, вейвлет-преобразование, функция энергии.

Постановка задачи

Одной из актуальных проблем, которая встает при выборе метода обучения многослойного персептрона, является скорость и качество обучения.

Нейронные сети широко используются при решении самых разных задач, где обычные алгоритмические решения оказываются неэффективными или невозможными. Несмотря на упрощения, допускаемые при построении нейронных сетей, они демонстрируют такие свойства, как обучение на основе опыта, обобщение, извлечение существенных данных из избыточной информации. Обученная сеть может быть устойчивой к некоторым отклонениям входных данных.

При решении задачи ставится вопрос о выборе нейронной сети с подходящей конфигурацией и принципом функционирования. Широко используется для поиска закономерностей и классификации образов многослойная полносвязанная нейронную сеть прямого распространения. Наиболее успешным методом обучения сети является алгоритм обратного распространения ошибки, применяемый с целью минимизации ошибки работы многослойного персептрона и получения желаемого выхода.

Простейшее правило обучения соответствует методу наискорейшего спуска. Однако при подобном обучении нейронной сети всегда существует возможность попадания алгоритма в локальный минимум. Для этого используются специальные приёмы, позволяющие вывести найденное решение из локального экстремума. [1] К ним относятся градиентные методы первого и второго порядков, работающие существенно быстрее. Также одним из вариантов, помогающим в решении задачи определения глобального минимума на поверхности ошибок является сглаживание функции энергии. [2]

Было осуществлено сравнение модификаций методов градиентного спуска. Спроектированная нейронная сеть решала задачу сопоставления входной группы точек одному из трёх типов сигнала. Была выбрана архитектура многослойного персептрона с одним промежуточным слоем. Входной слой состоит из 100 нейронов, выходной – из 3, число нейронов в скрытом слое варьируется от 8 до 12. В качестве функции активации выбрана сигмоидальная функция, так как она легко дифференцируется. Обучение сети происходило разными методами: метод градиентного спуска, метод градиентного спуска с адаптивным обучением, метод градиентного спуска с учётом моментов и метод градиентного спуска с учётом моментов и с адаптивным обучением.

Исходя из полученные результатов можно сделать следующие выводы.

– Оптимальной для обучения нейронной сети представляется выборка из 100 групп точек на каждый тип сигнала. При увеличении количества обучающих групп точек наблюдается явление переобучения сети, что приводит к потери качества её работы.

– Наилучшим образом обучаются сети, методом тренировки которых является метод градиентного спуска с адаптивным обучением или метод градиентного спуска с учётом моментов и адаптивным обучением (последний алгоритм предпочтительнее).

– Наименьшая погрешность классификации признака (порядка 10^{-6}) наблюдается у нейронных сетей с 10 нейронами в промежуточном слое.

При обучении каждым из рассмотренных методов нейронная сеть считается наилучшим образом натренированной, если был найден глобальный минимум функции ошибок (или минимум функции энергии – выбор функции зависит от модели нейронной сети). Для определения его расположения потребовалось разное время. Представляется возможным ускорить процесс обучения нейронной сети и сделать результат задачи поиска глобального минимума более точным, если подвергнуть функцию ошибки быстрому вейвлет-преобразованию.

На рисунке 1 показан анализируемый сигнал. На рисунках 2 и 3 представлено разложение функции энергии под действием различных вейвлет-преобразований. Большие белые пропуски соответствуют области глобального минимума функции энергии.

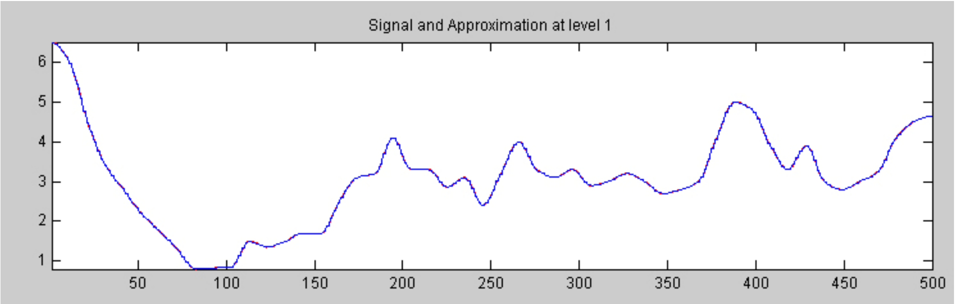

Рис. 1. Сигнал и его аппроксимация на уровне 1.

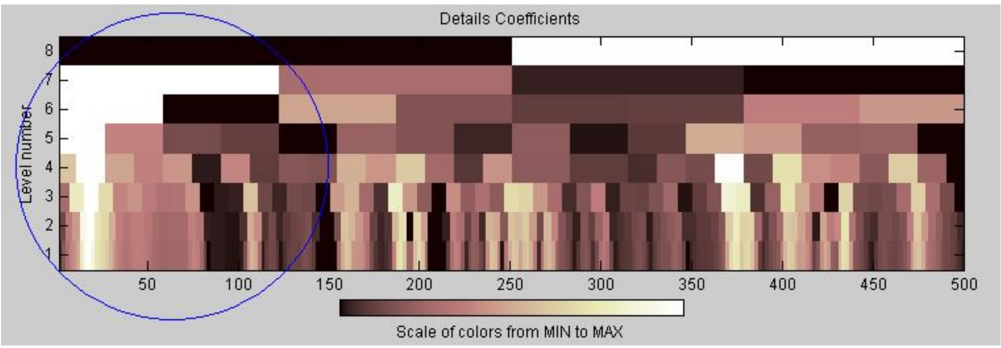

Рис. 2. Разложение функции энергии под действием вейвлет-функции Хаара (коэффициент равен 2, уровень 8).

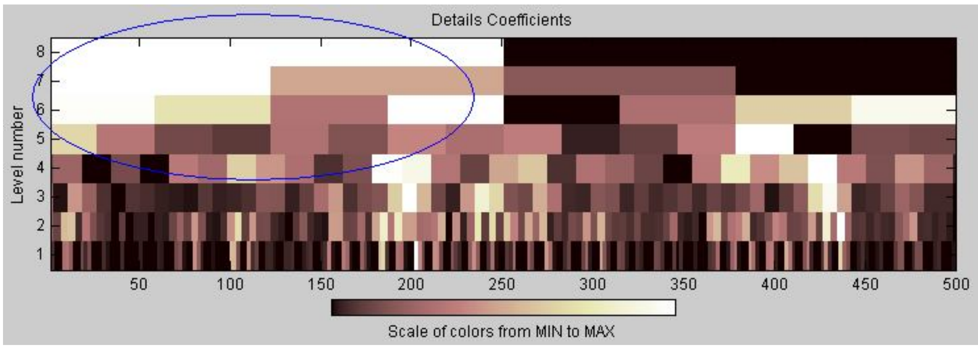

Рис. 3. Разложение функции энергии под действием вейвлет-функции Добеши (коэффициент равен 4, уровень 8).

Литература

1. Даниил Кальченко, «КомпьютерПресс» №1, январь 2005: http://www.neuroproject.ru/articles_dak_nn.php
2. http://www.conf.muh.ru/080130/thesis_Terehin.htm