



О.П. Солдатова, П.Д. Чайка

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ РАСПРЕДЕЛЁННОЙ ГИБРИДНОЙ СЕТЬЮ КОХОНЕНА

(Самарский государственный аэрокосмический университет)

В настоящее время использование нейронных сетей для решения задачи классификации становится обычной практикой. Различают множество видов архитектур нейронных сетей. В настоящее время наиболее часто используемой архитектурой нейросети является многослойный персептрон (MLP). Многослойные персептроны успешно применяются для решения разнообразных сложных задач. При этом обучение с учителем выполняется с помощью такого популярного алгоритма, как алгоритм обратного распространения ошибки.

В качестве отдельного большого класса нейронных сетей выделяют сети с самоорганизацией на основе конкуренции. Основу самоорганизации нейронных сетей составляет активация нейронов в ответ на поданный входной вектор, в результате чего активным оказывается один нейрон в сети (или в группе). Выходной нейрон, который выиграл соревнование, называется нейроном-победителем. Наиболее популярной моделью сети с самоорганизацией является карта Кохонена. В ней нейроны помещаются в узлах решетки, обычно одно- или двумерной.

Главная особенность сети с самоорганизацией – это очень высокая скорость обучения по сравнению с сетями, тренируемыми с учителем. Их недостатком считается сложность отображения пар обучающих данных (x, d) , поскольку сеть с самоорганизацией, выполняющая обработку только входного вектора x , не обладает свойствами хорошего аппроксиматора, присущими многослойному персептрону.

Очень хорошие результаты удалось получить при объединении самоорганизующегося слоя и персептронной сети, что позволяет совместить способности сети Кохонена к локализации и возможности аппроксимации, свойственные многослойному персептрону [1]. Подобная структура, называемая гибридной сетью Кохонена, изображена на рисунке 1. Она представляет собой каскадное подключение слоя Кохонена к персептронной сети. Самоорганизующийся слой улавливает значимые признаки процесса (локализует их на основе входных данных x), после чего им приписывается входной вектор в персептронном слое. Обучение гибридной сети состоит из двух отдельных этапов, следующих друг за другом. Сначала на множестве входных векторов x обучается слой Кохонена. В результате нейроны этого слоя организуются таким образом, что векторы их весов наилучшим образом отображают распределение данных обучающих векторов x . Обучение слоя Кохонена происходит по алгоритмам WTA или WTM [1].

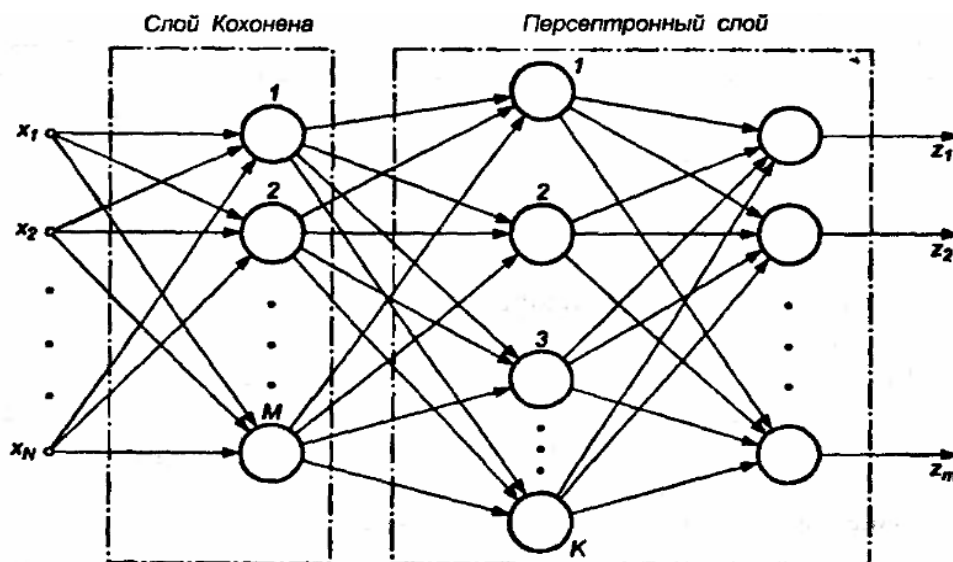


Рис. 1. Структура гибридной сети Кохонена

Обучающими сигналами для нее является множество пар (z_i, d_i) , где z_i —это вектор, составленный из выходных сигналов нейронов слоя Кохонена, а d_i —вектор, ожидаемых значений оригинального отображения (x_i, d_i) . Сеть обучается по методу обратного распространения ошибки. Если принять во внимание то, что каждому вектору x соответствует конкретный нейрон-победитель, то можно осуществить предварительную инициализацию весов персептронного слоя таким образом, что значения весов, соединяющих нейрон-победитель с выходными нейронами гибридной сети, будут усредняться по тем векторам d , на которых были одержаны победы. При такой предварительной инициализации весов, обучение персептронного слоя сводится к незначительной корректировке их значений, отражающей влияние проигравших нейронов слоя Кохонена на окончательный результат. Эта корректировка, как правило, требует небольшого количества итераций и ведет к достижению глобального минимума функции погрешности [2].

Однако острой остаётся проблема оптимизации архитектуры сети. Несмотря на хорошие результаты, гибридная сеть содержит большее число нейронов и весов, чем аналогичная сеть Кохонена и многослойный персептрон, что сказывается на времени обучения и объеме занимаемой памяти.

Поэтому авторами было выдвинуто предложение для уменьшения числа связей в гибридной сети Кохонена, путем объединения нейронов в слое Кохонена в группы. Такая сеть получила название распределенная гибридная сеть Кохонена. Ее структура представлена на рисунке 2.

В данной сети производится разделение обучающих данных на классы и в слое Кохонена создаётся своя группа нейронов для каждого отдельного класса. Затем каждая группа обучается только на своём классе при помощи алгоритмов WTA или WTM. После обучения слоя Кохонена обучается персептронная часть сети при помощи алгоритма обратного распространения ошибки. В качестве входного вектора для персептронного слоя используются не значения сигналов



от каждого нейрона слоя Кохонена, а усредненные значения выходных сигналов каждой группы нейронов.

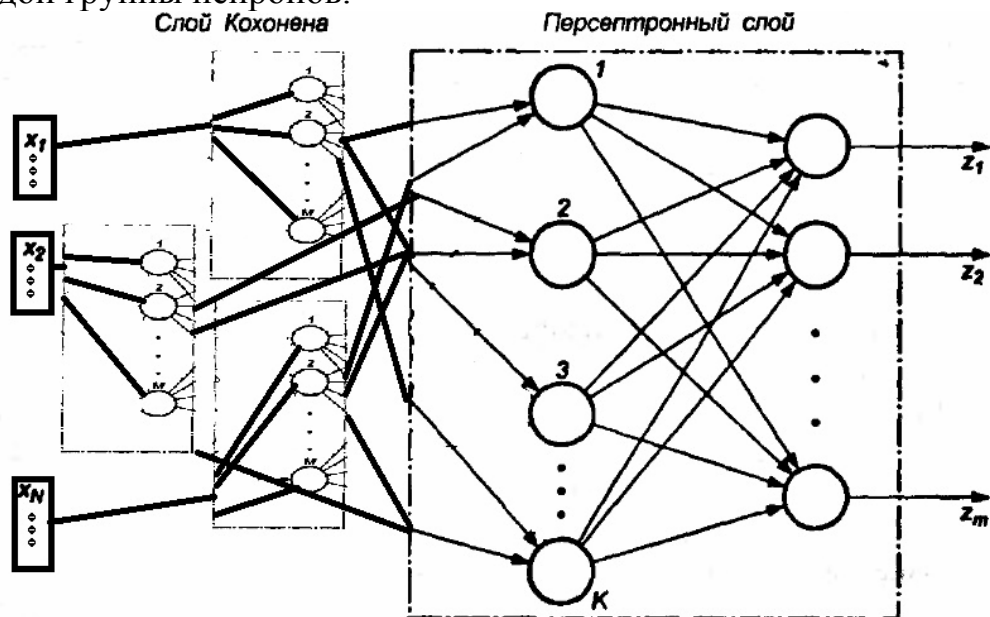


Рис. 2. Структура распределенной гибридной сети Кохонена

Так как каждая группа нейронов настроена только на свой класс, то в случае непересекающихся классов входных данных, в результате самоорганизации определяется не нейрон-победитель, а группа нейронов-победителей. Таким образом, количество групп нейронов равно числу классов, вследствие чего существенно уменьшается количество связей в сети, а значит, уменьшается количество операций на каждом шаге итерации. В данном случае вместо самоорганизации на уровне входных данных реализуется самоорганизация на уровне классов входных данных.

Данное действие не является эффективным в случае обычной карты Кохонена, так как неизвестно предварительное разбиение входных данных на классы. В случае гибридной карты это предварительное разбиение данных существует, поэтому его можно использовать для оптимизации структуры сети.

Однако пользователь может заранее не знать точного количества обучающих выборок для каждого класса. Поэтому для генерации структуры сети авторами вводится понятие так называемого коэффициента уменьшения. При использовании обычной карты Кохонена можно заметить, что минимально число нейронов, необходимое для эффективной работы сети равно количеству классов входных данных (по одному нейрону на класс). Меньшее количество нейронов приведет к отсутствию навыков распознавания некоторых классов входных данных. Максимальное количество нейронов будет равно количеству обучающих примеров (по нейрону на каждый пример). Больше количество нейронов приведет к неизбежному появлению мертвых нейронов. Поэтому оптимальное количество нейронов, необходимое для эффективного решения задачи классификации, лежит внутри этих границ. Для этих целей и вводится коэффициент уменьшения $a(0 < a \leq 1)$, при котором количество нейронов в каж-



дой группе $N_i = a * D_i$, где D_i количество обучающих примеров принадлежащих данному классу.

В качестве обучающих данных использовались четыре выборки с репозитория UCI– ирисы Фишера, классификация вин, заболевания печени и кредитный индекс Германии [3]. В таблице 1 показан процент неверно распознанных тестовых векторов в зависимости от структуры обычной гибридной сети Кохонена. В таблице 2 представлены аналогичные данные для распределенной гибридной сети Кохонена.

Таблица 1. Процент неверно распознанных векторов в зависимости от структуры обычной гибридной сети Кохонена

Количество нейронов в слое Кохонена	Количество весов	Название обучающей выборки	Процент ошибки классификации, %
100	615	Ирисы Фишера	0
144	879	Классификация вин	7.1
256	1546	Заболевания печени	10.2
400	2147	Кредитный индекс Германии	21.4

Таблица 2. Процент неверно распознанных векторов в зависимости от структуры распределенной гибридной сети Кохонена

Количество нейронов в слое Кохонена	Количество весов	Название обучающей выборки	Процент ошибки классификации, %
27	57	Ирисы Фишера	0.0
41	71	Классификация вин	5.4
181	251	Заболевания печени	8.3
231	384	Кредитный индекс Германии	15.7

Можно сделать вывод, что распределенная гибридная сеть Кохонена имеет более адекватную структуру для решения данных задач, чем обычная гибридная сеть Кохонена, при аналогичных погрешностях классификации. Это позволяет уменьшить количество операций, выполняемых за одну итерацию, время обучения и объем занимаемой памяти.

Литература

1. Осовский С., Нейронные сети для обработки информации [Текст] / С. Осовский, пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344с.
2. Хайкин С. Нейронные сети: полный курс. [Текст] / С. Хайкин, пер. с англ. – М.: Вильямс, 2006. – 1104 с.
3. Bache K., Lichman M. UCI Machine Learning Repository [available at: <http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science, 2013.