

Тема 8

Обучение без учителя. Принцип «победитель забирает всё»

В этой лекции будут рассмотрены различные компоненты однородных (состоящих из нейронов одного типа) и неоднородных нейронных сетей. Большинство рассматриваемых далее архитектур НС предназначены, прежде всего, для использования их в задачах кластерного анализа. Общая идея кластерного анализа состоит в том, чтобы определять принадлежность вновь поступающих входных векторов в тот или иной класс, таким образом, чтобы сильно отличающиеся вектора принадлежали разным классам, а мало отличающиеся – одному классу. Как правило, нейронные сети, решающие такие задачи, обучаются «без учителя», то есть не требуют задания выходов в обучающей выборке.

Для начала рассмотрим некоторые простые нейроподобные элементы.

Звезды Гроссберга

Работы по исследованию возможностей биологических нейронных сетей являются неисчерпаемым источником новых знаний в нейроинформатике. Так, исследования Стефана Гроссберга в области биологической кибернетики оказались богаты новыми идеями и были положены в основу ряда новых моделей ИНС. Одной из ключевых идей этих исследований является наблюдение того, что в нервной ткани биологических нейронных сетей в большом количестве присутствуют нейроподобные элементы двух типов – «входные» и «выходные звёзды». Опишем математическую модель таких нейронов.

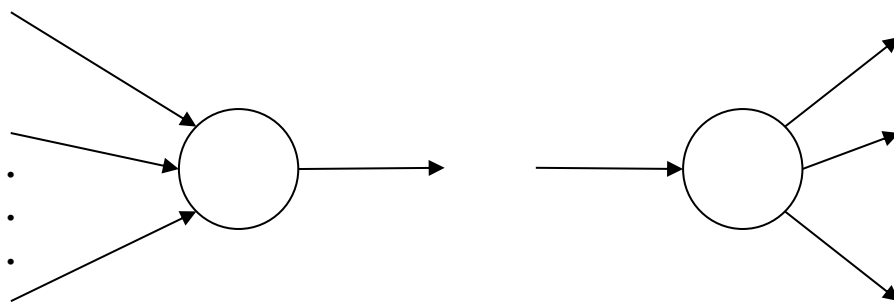


Рис 8.1. Входная и выходная звезда.

Нейрон в форме входной звезды имеет N входов x_1, \dots, x_N , которым соответствуют веса w_1, \dots, w_N , и один выход y , являющийся взвешенной суммой входов. Входная звезда обучается выдавать сигнал на выходе всякий раз, когда на входы поступает определенный

вектор. Таким образом, входная звезда является детектором совокупного состояния своих входов.

Формула функционирования входной звезды – формула функционирования стандартного формального нейрона. Однако есть следующие отличия:

1. функция активации, как правило, линейная;
2. входной вектор \bar{x} является нормированным (евклидовой нормой) вектором длины 1.

Процесс обучения представляется в следующей итерационной форме:

$$w_i(t+1) = w_i(t) + \alpha(x_i - w_i(t)).$$

Темп обучения $\alpha \in (0,1]$, часто имеет начальное значение масштаба 0.1 и постепенно уменьшается в процессе обучения. В процессе настройки нейрон учится усредненным обучающим векторам. Если, например, такой нейрон был обучен на одном векторе \bar{x}^1 , то при подаче на вход нейрона вектора \bar{x}^2 нейрон вычислит значение скалярного произведения этих векторов. А так как векторы единичные, то выходное значение будет равно косинусу угла между векторами: $y = |\bar{x}^1||\bar{x}^2| \cos(\alpha_{\bar{x}^1, \bar{x}^2}) = \cos(\alpha)$.

Выходная звезда Гроссберга выполняет противоположную функцию - функцию командного нейрона, выдавая на выходах определенный вектор при поступлении сигнала на вход. Нейрон этого типа имеет один вход x и M выходов y_1, \dots, y_M с весами w_1, \dots, w_M , которые обучаются по формуле:

$$w_i(t+1) = w_i(t) + \beta(y_i - w_i(t)).$$

Рекомендуется начать с β порядка единицы и постепенно уменьшать до нуля в процессе обучения. Итерационный процесс будет сходиться к собирательному образу, полученному из совокупности обучающих векторов.

Принцип «Победитель Забирает Все» в модели Липпмана-Хемминга.

Рассмотрим задачу о принадлежности образа \bar{x} некоторому классу X_K , определяемому заданными библиотечными образами $\bar{x}^j \in X_K$. Каждый из заданных образов обучающей выборки непосредственно определяет свой собственный класс, и таким образом, задача сводится к поиску "ближайшего" образа. В случае двух двоичных (0-1) образов расстояние между ними может быть определено по Хеммингу, как число несовпадающих компонент. Теперь после вычисления всех попарных расстояний $\rho_k = \rho(\bar{x}^j, \bar{x})$ искомым класс определяется по наименьшему из них.

Одно из нейросетевых решений этой задачи было получено на основе архитектуры Липпмана-Хемминга. Сеть имеет один слой из M одинаковых нейронов, число которых равно количеству классов. Таким образом, каждый нейрон "отвечает" за свой класс. Каждый нейрон связан с каждым из входов и со всеми остальными нейронами сети. Число входов каждого нейрона равно размерности рассматриваемых библиотечных образов. Веса связей полагаются равными нормированным библиотечным образам:

$$w_n^m = x_n^m / \left(\sum_j x_j^m \right).$$

Здесь w_n^m - значение веса связи от n -го входа к m -му нейрону. Процесс поступления информации о векторе \bar{x} в нейронную сеть является безитерационным. При этом входной вектор сначала нормируется:

$$s_n = \xi_n / \left(\sum_i \xi_i \right),$$

и нейроны принимают начальные уровни активности:

$$y^m(0) = f \left(\frac{1}{2} \sum_n s_n w_n^m - \theta \right).$$

Здесь $f(x)$ - функция активации нейрона:

$$f(x) = \begin{cases} 0, & \text{если } x < 0, \\ x, & \text{если } x > 0. \end{cases}$$

Пороги θ обычно полагаются равными нулю.

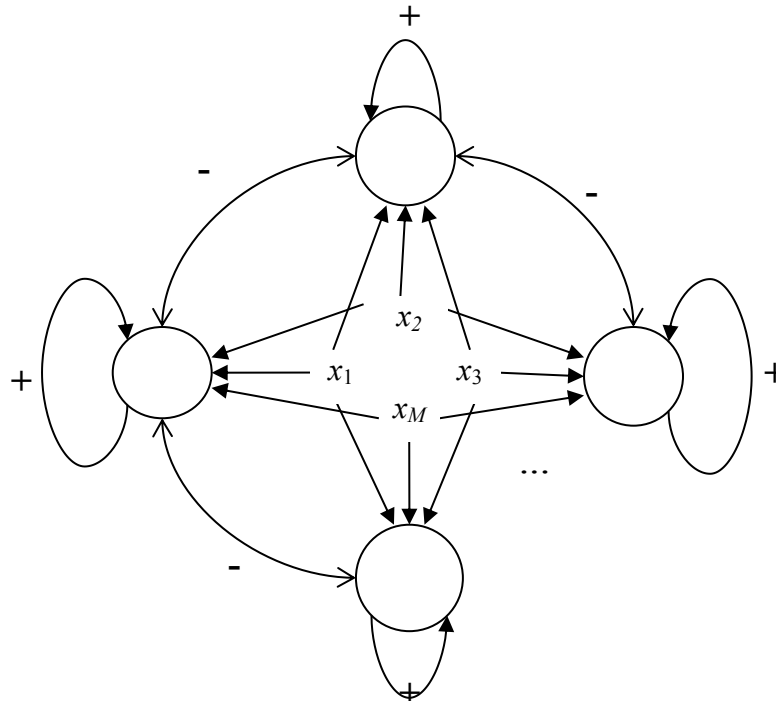


Рис. 8.2. ИНС Липпмана-Хеммнга

При поступлении входного вектора начальное возбуждение получают все нейроны, скалярное произведение векторов, памяти которых с входным вектором превышает порог. В дальнейшем среди них предстоит выбрать один, для которого оно максимально. Это достигается введением дополнительных обратных связей между нейронами, устроенных по принципу "латерального торможения". Каждый нейрон получает тормозящее (отрицательное) воздействие со стороны всех остальных нейронов, пропорционально степени их возбуждения, и испытывает возбуждающее (положительное) воздействие самого на себя. Веса латеральных связей в нейронном слое нормируются таким образом, что суммарный сигнал является возбуждающим только для нейрона с максимальной исходной активностью. Остальные нейроны испытывают торможение:

$$y^m(t+1) = f\left(y^m(t) - \left(\frac{1}{M+1}\right) \cdot \sum_{n \neq m} y^n(t)\right).$$

По выполнении некоторого числа итераций t для всех нейронов кроме одного значение аргумента функции $f(x)$ становится отрицательным, что обращает их активность y^m в нуль. Единственный, оставшийся активным, нейрон является победителем. Он и указывает на тот класс, к которому принадлежит введенный образ. Такой механизм получил название "Победитель-Забирает-Все" (Winner Take All - WTA). Механизм WTA используется и в других нейросетевых архитектурах. Заложенный в его

основе принцип латерального торможения имеет глубокие биологические основания и весьма широко распространен в нейронных сетях живых организмов.

Нейросетевая парадигма Липпмана-Хемминга является моделью с прямой структурой памяти. Информация, содержащаяся в библиотечных образах, никак не обобщается, а непосредственно запоминается в синаптических связях. Память здесь не является распределенной, так как при выходе из строя одного нейрона полностью теряется информация обо всем соответствующем ему образе памяти.

Карта самоорганизации Кохонена

Модель ИНС Кохонена [1] состоит из входных звезд Гроссберга и, в отличие от модели Липпмана-Хемминга, выполняет обобщение предъявляемой информации. В результате работы НС Кохонена получается образ, представляющий собой подобие карты распределения векторов из обучающей выборки. Таким образом, в модели Кохонена выполняется решение задачи нахождения кластеров в пространстве входных образов.

Данная сеть обучается без учителя на основе самоорганизации. По мере обучения, вектора весов нейронов стремятся к центрам кластеров - групп векторов обучающей выборки. На этапе решения информационных задач сеть относит новый предъявленный образ к одному из сформированных кластеров, указывая тем самым категорию, к которой он принадлежит.

Рассмотрим архитектуру НС Кохонена и правила обучения подробнее. Сеть Кохонена, также как и сеть Липпмана-Хемминга, состоит из одного слоя нейронов. Число входов каждого нейрона равно размерности входного образа. Количество нейронов определяется количеством кластеров, в исходной задаче классификации. При достаточном количестве нейронов и удачных параметрах обучения НС Кохонена может не только выделить основные группы образов, но и установить "тонкую структуру" полученных кластеров. При этом близким входным образам будет соответствовать близкие выходные вектора сети.

Функционирование НС Кохонена происходит по следующим правилам:

$$s_j = \sum_i w_{ij} x_i = \bar{w}_j \cdot \bar{x}.$$

Пусть $s_k = \max s_j$, тогда активность всех нейронов нейросети

$$y_j = F(s_j) = \begin{cases} 1, & \text{если } j = k; \\ 0, & \text{если } j \neq k, \end{cases}$$

где $j = \overline{1, m}$.

Процедура обучения состоит из следующих шагов:

1. Задать начальные значения матрице весов связей w_n^m .
2. Подать на вход сети входной вектор $\bar{x}(0) = \bar{x}^\alpha$ из обучающей выборки.
3. Итерация самоорганизации. На этом этапе происходит модификация весов.
 - a. Для каждого нейрона определяется его расстояние до вектора входа:

$$d_m = \sum_{i=1}^N (x_i(t) - w_i^m(t))^2.$$

- b. Выбирается нейрон $m=m^*$, для которого это расстояние минимально. На текущем шаге обучения t будут модифицироваться только веса нейронов из окрестности нейрона m^* : $w_n^m(t+1) = w_n^m(t) + \eta(t)(x_n(t) - w_n^m)$.
 - c. Первоначально в окрестности любого из нейронов находятся все нейроны сети, в последствии эта окрестность сужается. В конце этапа обучения подстраиваются только веса самого ближайшего нейрона. Темп обучения $\eta(t) < 1$ с течением времени также уменьшается.
4. Образы обучающей выборки предъявляются последовательно, и каждый раз происходит подстройка весов.

Нейронная сеть Кохонена может обучаться и на искаженных версиях входных векторов, в процессе обучения искажения, если они не носят систематический характер, сглаживаются.

Для наглядности представления карты нейроны Кохонена могут быть упорядочены на плоскости, при этом нейронам с близкими векторами синаптических весов будут соответствовать близкие вектора на плоскости. Тогда в окрестности нейрона-победителя будут соседние (в геометрическом смысле) точки на плоскости. Результирующую карту весов сети можно представить в виде двумерного изображения, на котором различные степени возбуждения всех нейронов отображаются квадратами различной площади.

Каждый нейрон несет информацию о кластере - сгустке в пространстве входных образов, формируя для данной группы собирательный образ. Таким образом НС Кохонена способна к обобщению. Конкретному кластеру может соответствовать и несколько

нейронов с близкими значениями векторов весов, поэтому выход из строя одного нейрона не так критичен для функционирования НС Кохонена, как это имело место в случае хемминговой сети.

Главным свойством сети Кохонена считается компрессия данных, состоящая в том, что образующие кластер большие группы данных представляются единственным вектором весов нейрона-победителя. Это компрессия с потерями.

Нейронная сеть встречного распространения

Архитектура встречного распространения (counter propagation) удачно объединяет в себе преимущества возможности обобщения информации сети Кохонена и простоту обучения выходной звезды Гроссберга. Создатель сети встречного распространения Р. Хехт-Нильсен рекомендует использование этой архитектуры для быстрого моделирования систем на начальных этапах исследований с дальнейшим переходом, если это потребуется, на значительно более дорогой, но более точный метод обучения с обратным распространением ошибок [5,6].

НС встречного распространения (НСВР) обучается на выборке пар векторов (\bar{x}, \bar{y}) в задаче представления отображения $\varphi: X \rightarrow Y$. Замечательной особенностью этой сети является способность обучению также и отображению совокупности $X \times Y$ в себя. При этом благодаря обобщению, появляется возможность восстановления пары (\bar{x}, \bar{y}) по одной известной компоненте (\bar{x} или \bar{y}). При предъявлении на этапе распознавания только вектора \bar{x} (с нулевым начальным \bar{y}) производится прямое отображение - восстанавливается \bar{y} , и наоборот, при известном \bar{y} может быть восстановлен соответствующий ему \bar{x} . Возможность решения как прямой, так и обратной задачи, а также гибридной задачи по восстановлению отдельных недостающих компонент делает данную нейросетевую архитектуру уникальным инструментом.

НСВР состоит из двух слоев нейронов - слоя Кохонена и слоя Гроссберга. В режиме функционирования (распознавания) нейроны слоя Кохонена работают по принципу Победитель-Забирает-Все, определяя кластер, к которому принадлежит входной образ. Затем выходная звезда слоя Гроссберга по сигналу нейрона-победителя в слое Кохонена воспроизводит на выходах сети соответствующий образ.

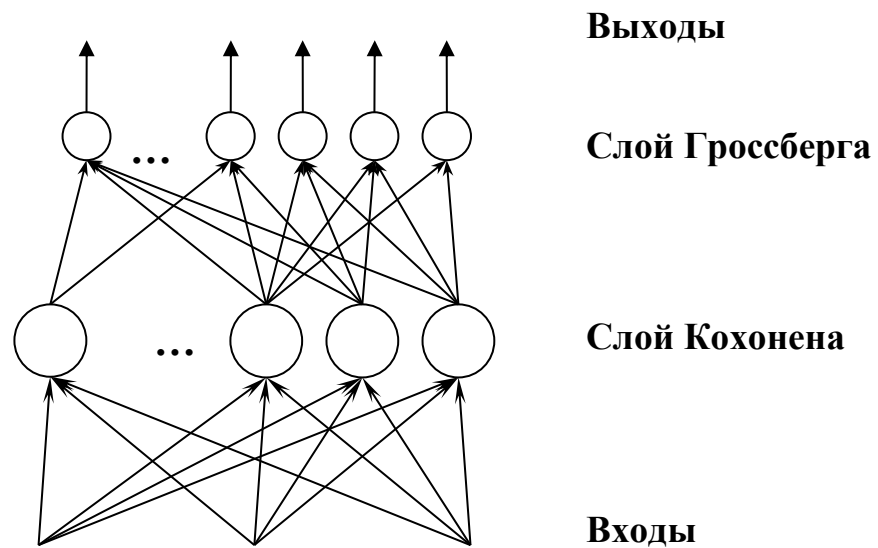


Рис. 8.3. Архитектура сети встречного распространения.

Обучение весов слоя Кохонена выполняется без учителя на основе самоорганизации по алгоритму описанному ранее. Входной вектор вначале нормируется, сохраняя направление. После выполнения одной итерации обучения определяется нейрон победитель, состояние его возбуждения устанавливается равным единице, и теперь могут быть модифицированы веса соответствующей ему звезды Гроссберга. Темпы обучения нейронов Кохонена и Гроссберга должны быть согласованы. В слое Кохонена обучаются веса всех нейронов в окрестности победителя, которая постепенно сужается до одного нейрона.

Обученная нейронная сеть ВР может функционировать и в режиме интерполяции, когда в слое Кохонена оставляется не один, а несколько победителей. Тогда уровни их активности пропорционально нормируются, чтобы в сумме составлять единицу, а выходной вектор определяется по сумме выходных векторов каждой из активных звезд Гроссберга. Таким образом НС производит линейную интерполяцию между значениями выходных векторов, отвечающих нескольким кластерам. Однако режим интерполяции в сети встречного распространения изучен не столь достаточно, чтобы можно было рекомендовать его широкое использование.

Гибридная сеть

Главная особенность сети с самоорганизацией на основе конкуренции – это очень высокая скорость обучения, многократно большая, чем у сетей, тренируемых 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 или радиальной сети. Слой Гроссберга в сети ВР не может считаться хорошим аппроксиматором и естественным шагом для улучшения свойств НС является замена слоя Гроссберга на многослойный персептрон. Причем для получения хороших результатов, часто достаточно одного слоя. Нейронная сеть с такой архитектурой называется гибридной. В гибридной сети слой Кохонена имеет роль локализации главных признаков входного вектора, после чего осуществляется преобразование выделенных признаков в другое значение персептронной частью сети.

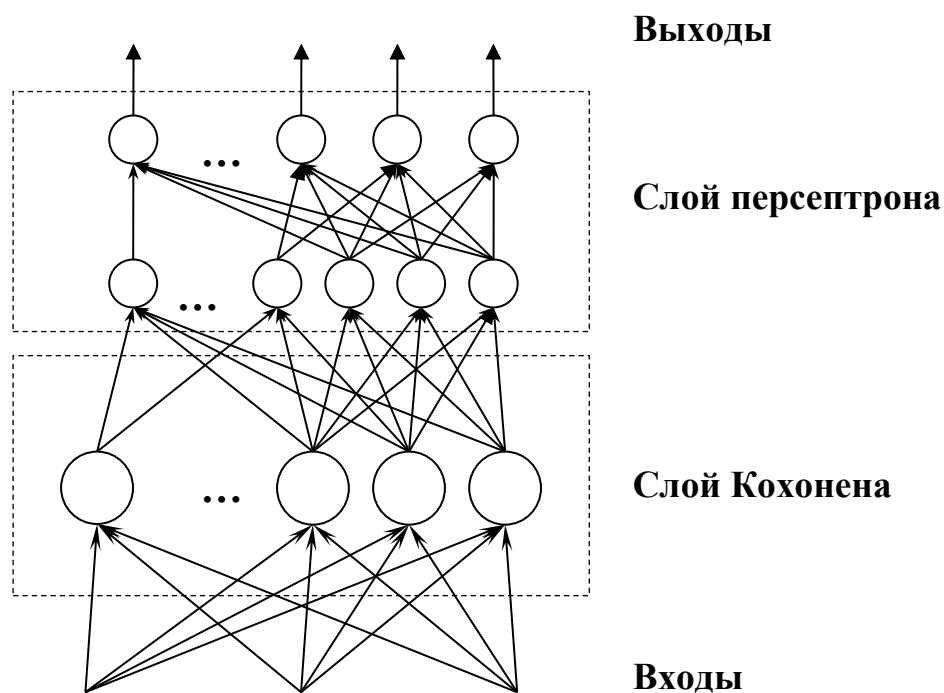
Обучение гибридной сети состоит из двух этапов, выполняющихся последовательно.

1. Сначала на всём множестве входных векторов обучается слой Кохонена, в результате чего нейроны этого слоя организуются оптимальным образом. По завершении обучения слоя Кохонена, веса его нейронов замораживаются и проводится анализ их выходных сигналов при подаче на вход сети последовательности сигналов x из обучающего множества. Победитель переводится в единичное состояние, а остальным нейронам приписываются состояния из интервала $(0, 1)$. Другими словами, выходные сигналы слоя Кохонена нормализуются. Это можно сделать, например, с помощью функции

$$y_j = \exp\left(-\frac{|u_{\max} - u_i|^2}{\sigma^2}\right).$$

2. Персептронная часть сети обучается с учителем по завершении тренинга слоя Кохонена. Обучающей выборкой для неё будет множество пар (y_j, \bar{d}_j) , где \bar{y}_j - это

вектор

**Рис. 7.4. Архитектура гибридной сети**

составленный из выходных сигналов нейронов слоя Кохонена, а \bar{d}_j - вектор ожидаемых значений оригинального отображения (\bar{x}_i, \bar{d}_i) , которому соответствует вектор \bar{y}_i . Обучение можно проводить как алгоритмом обратного распространения, так и любым другим градиентным методом.

Следует отметить, что градиентное обучение персептронной части сети проходит в несколько раз быстрее, чем в случае обучения многослойного персептрона решению задачи аппроксимации функции заданной обучающей выборкой (\bar{x}_i, \bar{d}_i) .

Литература

1. T. Kohonen, "Self-organized formation of topologically correct feature maps", Biological Cybernetics, Vol. 43, pp.59-69, 1982.
2. T. Kohonen, Self-Organizing Maps, Springer, 1995.
3. Ф.Уоссермэн. Нейрокомпьютерная техника. М. Мир, 1992.
4. Grossberg S. 1982. Studies of mind and brain. Boston: Reidel.
5. Hecht-Nielsen R. 1987a. Counterpropagation networks. In Proceedings of the IEEE First International Conference on Neural Networks, eds. M. Caudill and C. Butler, vol. 2, pp. 19-32. San Diego, CA: SOS Printing.
6. Hecht-Nielsen R. 1987b. Counterpropagation networks. Applied Optics 26(23): 4979-84.