

Сеть встречного распространения

Введение

Сеть встречного распространения (Contrastive Divergence, CD) - это метод, используемый в обучении ограниченных машин Больцмана и глубоких нейронных сетей. Этот метод предназначен для ускорения процесса обучения, позволяя модели быстро сходиться к локальному оптимуму.

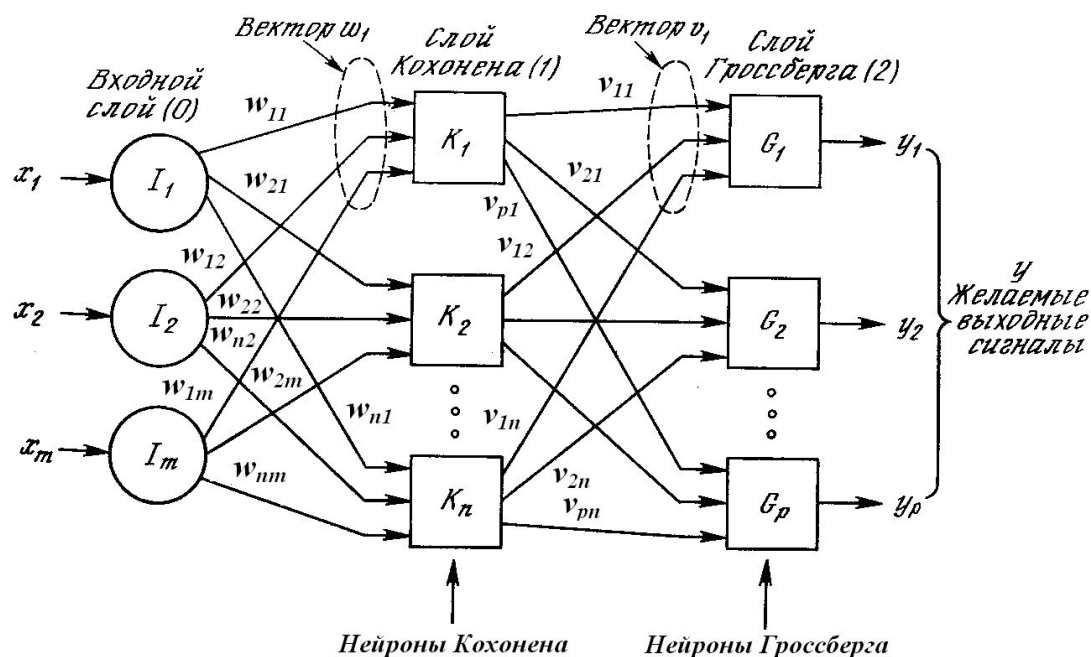
Основные особенности

1. Обучаются без учителя, т.к. нейроны сети не требуют маркированных входных данных для обучения, а вместо этого учатся распознавать и воспроизводить паттерны в данных.
2. Сеть состоит из входного слоя, выходного слоя и одного или нескольких слоев ассоциативных нейронов между ними, то есть сеть обладает двухслойной архитектурой.
3. Веса нейронов корректируются для минимизации ошибки на выходе, при этом корректировка происходит одновременно для всех нейронов.
4. Могут быть использованы в решении различных задач, связанных с распознаванием образов, сжатием данных и ассоциативной памятью.

Структура сети

Структура сети

- Нейроны слоя 0 (показанные кружками) служат лишь точками разветвления и не выполняют вычислений
- Каждый нейрон слоя 0 соединен с каждым нейроном слоя 1 (называемого слоем Кохонена) отдельным весом w_{mn} (веса в целом рассматриваются как матрица весов W)
- Аналогично, каждый нейрон в слое Кохонена (слое 1) соединен с каждым нейроном в слое Гроссберга (слое 2) весом v_{np} (веса образуют матрицу весов V)



Особенность состоит в операциях, выполняемых нейронами Кохонена и Гроссберга. Как и многие другие сети, встречное распространение функционирует в двух режимах: в нормальном режиме, при котором принимается входной вектор X и выдается выходной вектор Y , и в режиме обучения, при котором подается входной вектор и веса корректируются, чтобы дать требуемый выходной вектор.

Нормальное функционирование

Слой Кохонена

В своей простейшей форме слой Кохонена функционирует так, что для данного входного вектора один и только один нейрон Кохонена выдает на выходе логическую единицу, все остальные выдают ноль.

Нейроны Кохонена можно воспринимать как набор электрических лампочек, так что для любого входного вектора загорается одна из них. Ассоциированное с каждым нейроном Кохонена множество весов соединяет его с каждым входом.

Выход каждого нейрона Кохонена высчитывается как:

$$NET_i = w_{1j}x_1 + w_{2j}x_2 + \dots + w_{mj}x_m$$

где NET_j - это выход NET нейрона Кохонена

$$N = XW$$

где N - вектор выходов NET слоя Кохонена

Нейрон Кохонена с максимальным значением NET является "победителем". Его выход равен единице, у остальных он равен нулю.

Нормальное функционирование

Слой Гроссберга

Слой Гроссберга функционирует в сходной манере. Его выход NET является взвешенной суммой выходов k_1, k_2, \dots, k_n слоя Кохонена, образующих вектор K.

Вектор соединяющих весов, обозначенный через V, состоит из весов $v_{11}, v_{21}, \dots, v_{np}$. Тогда выход NET каждого нейрона Гроссберга есть:

$$NET_j = \sum_{i=1}^n k_i w_{ij}$$

где NET_j - выход j-го нейрона Гроссберга, или в векторной форме

$$Y = KV$$

где Y - выходной вектор слоя Гроссберга, K - выходной вектор слоя Кохонена, V - матрица весов слоя Гроссберга. Если слой Кохонена функционирует таким образом, что лишь у одного нейрона величина NET равна единице, а у остальных равна нулю, то лишь один элемент вектора K отличен от нуля, и вычисления очень просты. Фактически каждый нейрон слоя Гроссберга лишь выдает величину веса, который связывает этот нейрон с единственным ненулевым нейроном Кохонена.

Обучение слоя Кохонена

1. Входной сигнал подается на входной слой нейронов, которые передают сигнал на слой Кохонена
2. На слое Кохонена каждый нейрон имеет свой вес (координату в пространстве признаков), который определяет его положение на карте Кохонена. Эти веса представляют собой матрицу, где каждая строка соответствует одному нейрону, а каждый столбец соответствует одной координате в пространстве признаков.
3. Каждый нейрон на слое Кохонена вычисляет свое расстояние до входного сигнала, используя функцию расстояния, такую как евклидово расстояние. Нейрон с наименьшим расстоянием называется победителем или наиболее активным нейроном
4. Веса нейронов на слое Кохонена обновляются в направлении к входному сигналу, используя алгоритм обучения Кохонена, такой как обучение с конкуренцией. В результате карта Кохонена самоорганизуется и разбивает пространство признаков на кластеры
5. Выходной слой нейронов получает информацию о наиболее активном нейроне на слое Кохонена и выводит соответствующий класс или метку

Обучение слоя Гроссберга

1. Входной сигнал распространяется по слою Гроссберга, и каждый нейрон вычисляет свой выходной сигнал, используя свою функцию активации

2. Ошибка вычисляется на выходном слое, и она распространяется обратно через сеть в направлении входного слоя
3. Веса связей между нейронами корректируются таким образом, чтобы уменьшить ошибку на выходном слое. Это достигается путем изменения весов в соответствии с градиентом ошибки, который вычисляется с использованием алгоритма обратного распространения ошибки
4. Этот процесс повторяется для каждого входного сигнала и соответствующей целевой величины, пока не будет достигнута требуемая точность или количество эпох обучения не будет выполнено