

ЛЕКЦИЯ 11. НЕЙРОННЫЕ СЕТИ

Сергей Николенко

1. ОСНОВНЫЕ ПОНЯТИЯ

В этой лекции рассмотрим понятие нейронных сетей и способы их применения в задаче распознавания *речи*. Точнее говоря, мы изучим процесс обучения нейронной сети.

Итак, что же такое нейронная сеть?

Определение 1. *Нейронные сети — математические модели, а также их программные или аппаратные реализации, построенные по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма.*

Понятие нейронных сетей возникло при изучении процессов, протекающих в мозге при мышлении, и при попытке смоделировать эти процессы. Нейронная сеть представляет собой систему соединенных между собой простых процессоров. Они довольно просты, и каждый из них обрабатывает входящие сигналы и посылает их другим процессорам. Будучи соединенными в достаточно большую сеть с управляемым взаимодействием, такие локально простые процессоры вместе способны выполнять довольно сложные задачи.

Нейронные сети имеют возможность *обучаться*, в чем и заключается одно из главных их преимуществ перед традиционными алгоритмами. Технически обучение заключается в *нахождении коэффициентов* связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Это значит, что, в случае успешного обучения, сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке.

На рис. 1 представлена простая нейронная сеть. Нейроны в ней распределены по L уровням, на каждом из которых находятся I_k нейронов. Для наглядности каждый нейрон, кроме входных и выходных, представлен в виде *двух* узлов: один суммирует входящие сигналы, второй преобразует их.

Обозначение 1. *Нейроны и значение в каждом из них будем обозначать как y_{ij} , где i — номер уровня, а j — номер нейрона на уровне.*

Обозначение 2. *Запись w_{ij} будет обозначать связь между j -м нейроном уровня l с i -м нейроном уровня $l + 1$.*

Законспектировал Пенькин Даниил.

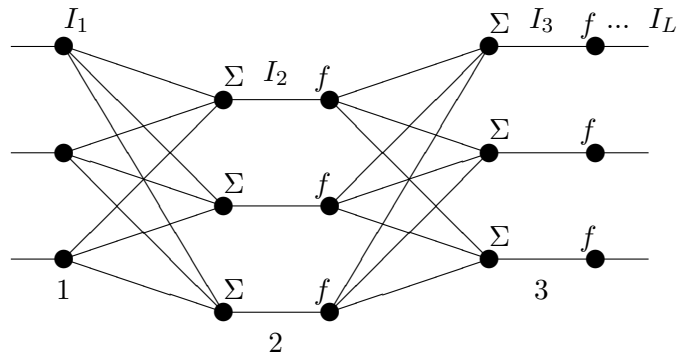


Рис. 1. Нейронная сеть.

ПРИМЕР 1

y_{5I_5} — значение в последнем нейроне пятого уровня; w_{111} — связь между y_{11} и y_{21} ; w_{132} — связь между y_{12} и y_{23} .

2. ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

Предположим, что наша нейронная сеть полностью связана между уровнями. В качестве функции преобразования f в нейроне возьмем $f = \frac{1}{1+e^{-x}}$, где x — входящий параметр, являющийся вещественным числом. В нашем случае x является суммой всех y . Итак, обобщая вышесказанное, запишем:

$$X_l = W_{l-1}^T \cdot y_{l-1}, \quad x_{li} = \sum_{k=1}^{I_{l-1}} w_{l-1,i,k} \cdot y_{i-1,k}$$

$$Y_l = f(X_l), \quad y_{li} = f(x_{li})$$

Теперь приступим к обучению: допустим, у нас есть набор входных данных, и мы хотим научить сеть описывать его. Будем действовать по схеме *градиентного спуска*. Для этого построим и будем минимизировать функцию ошибки, двигаясь в сторону ее обратного градиента.

Обозначение 3. Истинные выходы сети обозначим как $d_1 \dots d_{I_L}$

Имея на выходе сети значения $y_{L1} \dots y_{LI_L}$, можем записать *функцию ошибки*:

$$E = \sum_{k=1}^{I_L} (d_k - y_{Lk})^2$$

Эту функцию необходимо минимизировать от весов связей, то есть найти градиент, а именно, нужно посчитать производную:

$$\frac{\partial E}{\partial W_{lij}} \quad \forall lij$$

Введем правило обновления весов:

$$\delta w_{lij} = -\eta \cdot \frac{\partial E}{\partial w_{lij}}$$

Посчитаем для начала производную по связям с предпоследним уровнем. В ходе дальнейших рассуждений будем обозначать $e_i = d_i - y_{Li}$.

$$\begin{aligned} \frac{\partial E}{\partial w_{L-1,i,j}} &= \sum_{k=1}^{I_L} \underbrace{\frac{\partial E}{\partial y_{Lk}}}_{-2e_k} \cdot w_{L-1,i,j} = -2 \sum_{k=1}^{I_L} e_k \underbrace{\frac{\partial y_{Lk}}{\partial x_{L-1,k}}}_{f'(x_{L,k})} \cdot \frac{\partial x_{L-1,k}}{\partial w_{L-1,i,j}} = \\ &= -2 \sum_{k=1}^{I_L} e_k f'(x_{Lk}) \cdot y_{L-1,j} \cdot [i = k] = -2e_i f'(x_{Li}) \cdot y_{L-1,j} \end{aligned}$$

Очевидно, что переходы здесь законны, поскольку:

$$y_{Lk} = f(x_{Lk}); \quad x_{Lk} = \sum_{s=1}^{I_{L-1}} w_{L-1,k,s} \cdot y_{L-1,s}$$

Займемся теперь нахождением требуемой производной. Для этого сначала введем новые обозначения:

$$\delta_{li} = \frac{\partial E}{\partial x_{li}}; \quad \delta_{Li} = \frac{\partial E}{\partial x_{Li}} = -2y_{Li} \cdot f'(x_{Li})$$

Найти требуемую производную не составляет труда:

$$\frac{\partial E}{\partial w_{L,i,j}} = \frac{\partial E}{\partial x_{l+1,i}} \cdot \underbrace{\frac{\partial x_{l+1,i}}{\partial w_{lij}}}_{y_j} = y_j \cdot \delta_{l+1,i}$$

Для полноты картины осталось лишь найти выражение для δ_{li} . Оно выглядит следующим образом:

$$\delta_{li} = \frac{\partial E}{\partial x_{li}} = \sum_{s=1}^{I_L} \underbrace{\frac{\partial E}{\partial x_{l+1,s}}}_{\delta_{l+1,s}} \cdot \underbrace{\frac{\partial x_{l+1,s}}{\partial x_{li}}}_{\frac{\partial x_{l+1,s}}{\partial y_{li}} \cdot \frac{\partial y_{li}}{\partial x_{li}} = f'(x_{li}) w_{lsi}}$$

Итак, запишем отдельно полученные результаты:

$$\begin{cases} \delta_{Li} &= -2 \cdot y_{Li} \cdot f'(x_{Li}) \\ \delta_{li} &= \sum_{s=1}^{I_l} f'(x_{li}) \cdot w_{lsi} \cdot \delta_{l+1,s} \end{cases}$$

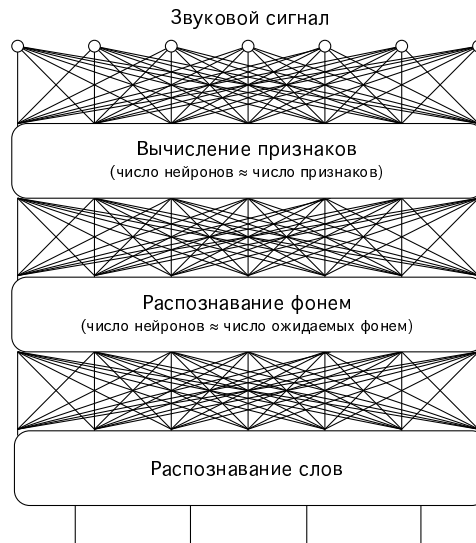


Рис. 2. Статическая нейронная сеть с логически разделенными уровнями

Используя полученные выражения, мы можем записать градиент, который искали, следующим образом:

$$\frac{\partial E}{\partial w_{lj}} = \delta_{l+1,i} \cdot y_{lj}$$

3. ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ

Не будем забывать, что основной темой курса являются вопросы распознавания речи. В этом разделе мы бегло рассмотрим применение нейронных сетей в данной области.

- (1) Можно обучать *статические* нейронные сети. Лучше, если уровни в ней имеют смысл. Пример такого смыслового разделения представлен на рис. 2.

Однако при своей относительной простоте, применение статических нейронных сетей во многих случаях бессмысленно по ряду причин. Главная из них — ожидаемость количества выходов. То есть, если мы ждем всего одно слово, работать такая схема будет неплохо, но в этом случае неизбежно придется фрагментировать входной сигнал. Вторая причина заключается в том, что огромное количество выходов, обусловленное временем на произношение слова, влечет за собой необходимость в большом числе сэмплов входного сигнала.

- (2) Попробуем построить сеть, которой на вход подается изменяющийся звуковой сигнал, то есть на вход сигналы подаются один за другим, таким же образом они появляются на выходах. Тогда обработка последующих сигналов будет зависеть от предыдущих. Отсюда возникает идея *TDNN — Time Delay Neural Network*.

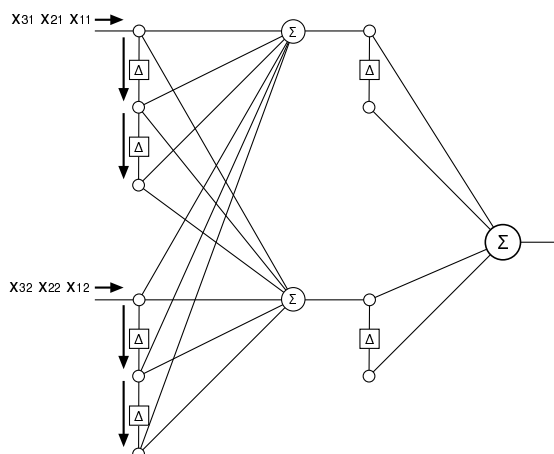


Рис. 3. Нейронная сеть с временной задержкой

Идея TDNN заключается во введении задержки Δ при обработке сигнала в нейроне. Схема такой сети представлена на рис. 3. Каждый сигнал не просто замещается следующим, а “смещается” на соседний узел, оставаясь в том же уровне и том же нейроне. При этом он продолжает участвовать в получении сигнала на следующем уровне наравне со вновь поступившими сигналами. Таким образом, обработка последующих сигналов зависит, в частности, от предыдущих сигналов.

Общее временное влияние такой сети равно сумме всех задержек в ней.