О моделях искусственных нейронов агрегирующего типа

Чередников Д.Ю.[2] Шибзухов З.М.[1,2]

[1]Институт прикладной математики и автоматизации, Нальчик

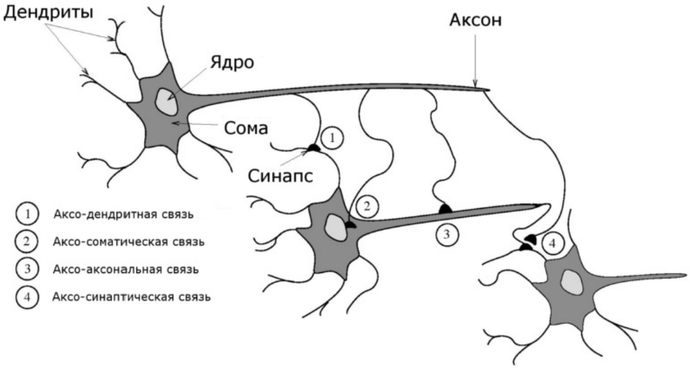
[2]Московский педагогический государственный унивеситет, Москва

ММРО - 17 2015

Содержание

- Биологическая нейронная сеть
- Модель классического нейрона
- Синаптические кластеры
- Выбор математической модели нейрона
- Модель$\sum\prod-$нейрона
- Модель нейрона с агрегирующими сложными синапсами
- Двуслойная нейронная сеть со сложными синапсами

Биологическая нейронная сеть



Природные нейроны имеют разветвленную дендритную систему, в которой сигналы поступают от других нейронов при помощи синапсов

Простейшая модель нейрона

Простейшая модель нейрона состоит из ряда допущений:

- синаптические связи - простые (с единственным входом)

- вклад синаптической связи пропорционален величине входа($w_i x_i$);
- величина суммарного входа нейрона линейно складывается из вкладов синаптических связей

Модель классического нейрона

Следующие утверждения приводят к модели с простыми синапсами:

$$y = sn(x) = out(\theta + \sum_{i=1}^{n} w_i x_i)$$

где y - выход, $x = (x_1 \ldots x_n)$ - входы, out(s) - функция выхода, θ- смещение, $w_1 \ldots w_n$- веса,
$sn(x)$ - функция преобразования нейрона

Суммарный потенциал нейрона есть арифметическая сумма вкладов простых синапсов и смещения θ

Выбор математической модели нейрона

Выбор математической модели нейрона влияет на характер процедуры настройки весов нейрона и на характеристики архитектуры нейронной сети

Слабая способность простых нейронов вида:

$$y = out(\theta + \sum w_i x_i)$$

приводит к необходимости использования многослойной архитектуры даже для простейших задач

Выбор математической модели нейрона

Наиболее оптимальны модели нейрона которые:

- лучше отражают процессы обработки информации в коре головного мозга
- обладают лучшими споособностями по представлению зависимостей

От линейной к полилинейной функции суммирования

Для того чтобы отразить особенности реальных нейронов в модель искусственного нейрона необходимо ослабить допущения и расширить её.

Простейший способ расширения модели состоит:

- уменьшение линейности вкладов;
- переход от простых синапсов к более сложным;
- независимость входов;
- переход к полилинейной функциям суммарного входного сигнала;

От линейной к полилинейной функции суммирования

Принятие полилинейной функции суммирования предполагает следующие допущения:

- синапсы могут быть как простые, так и сложные(синаптические кластеры);
- величина суммарного входа нейрона складывается из величин вкладов простых синапсов и синаптических кластеров;
- величина вклада синаптического кластера представляет полилинейную функцию от входов, учавствующих в образовании;

Синаптические кластеры

В природных нейронах встречаются синаптические кластеры, в образовании которых учавствуют несколько входов, так что вклад каждого входа зависит от величин других входов.

Модель $\sum\prod$-нейрона описывает преобразования в синаптическом кластере(сложном синапсе)

Акco-дендридная связь

Модель нейрона с полилинейной функцией суммарного входа

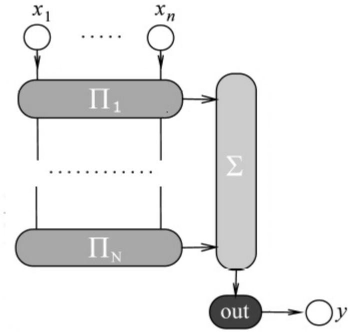
$$y = out(\theta + \sum_{k=1}^{m} w_i \prod_{i \in i_k} x_i)$$

Моделирует синаптические кластеры при помощи агрегирующей функции произведения.

В такой модели произведение используется для отражения вклада группы зависимых входов с номерами $i \in i_k$

$\Sigma\Pi$- нейрон

$\sum\prod$ —нейрон можно представить в виде двуслойной сети со слоем из $\prod$-элементов и $\sum$- нейроном в выходном слое



$$y = out(\theta + \sum_{k=1}^{N} w_k s_k)$$

$$s_k = \prod_{i \in i_k} x_i,\ k = \overline{1N}$$

Свойства $\sum\prod$- нейрона

- $\sum\prod$- нейрон способен интерполировать произвольные дискретные функции, определенные на конечных подмножествах R^n
- Позволяет построить множество различных корректно обученных моделей
- В операциях типа бустинга или взвешенного голосования можно построить $\sum\prod$- нейрон который не ошибается более чем на половине обучающей выборки, используя при этом не более половины обучающей выборки

Конструктивное обучение

К моделям $\sum\prod$-нейрона можно применить конструктивную процедуру обучения.

Конструктивный подход к обучению ориенторован на:

- на построение потенциально оптимальной архитектуры сети
- осуществление адекватной настройки весовых параметров

Конструктивные методы обучения направлены на постепенное построение архитектур, способствующих:

- улучшению сходимости процесса обучения
- достижению корректного функционирования

Модель нейрона с агрегирующими сложными синапсами

Модель $\sum\prod$- нейрона является частным случаем следующей модели нейрона со сложными синапсами:

$$y = out(\theta + \sum_{k=1}^{m} u_k)$$

где u_k- вклад k-ого агрегирующего сложного синапса или синаптического кластера, состоящего из простых синапсов

$$u_k = Agg_{\text{п}}\{u_i : i \in i_k\}$$

где $Agg_{\text{п}}$- агрегирующая функция, которая удовлетворяет требованию

$$u_k = Agg_{\text{п}}\{u_i : i \in i_k\} = 0 \Leftrightarrow \exists i \in i : u_i = 0$$

Модель нейрона с агрегирующими сложными синапсами

Пусть out- корректная функция выхода, т.е для $\forall y \in Y$ найдется значение s такое что $y = out(s)$

Функция простого синапса, такая $syn(x - a, w)$, такая что:

- $syn(x - a, w) = 0$, только при $x \leqslant a$;
- $\forall w$ функция $syn_w(x) = syn(x, w)$ - монотонная;
- $\forall w$ функция $syn_x(w) = syn(x, w)$ - строго монотонная;
- $\forall x$ и $\forall u \, \exists w : \; u = syn_x(w)$, только при $x \geqslant a$;

Модель нейрона с агрегирующими сложными синапсами

Theorem

Для любого конечного непротиворечивого набора прецедентов $<\widetilde{X},\widetilde{Y}>$ можно построить агрегирующий нейрон

$$agn(x) = out(\theta + Agg_{\prod}\{u_i : i \in i_k\}$$

такой что $y = agn(x)$ для любой пары $<x,y>$ из $<\widetilde{X},\widetilde{Y}>$

Двуслойная модель нейрона со сложными синапсами

Модель агрегирующего нейрона можно обобщить, если заменить обычное суммирование на симметричную агрегирующую функцию:

$$y = out \circ Agg_{\sum}\{\theta, s_1, ..., s_m\},$$

$$s_k = Agg_{\prod}\{u_{ki} : i \in\}$$

$Agg_{\sum}$ —агрегирующая операция для вычисления суммарного потенциала нейрона,
$Agg_{\prod}$— агрегирующая операция для вычисления общего вклада простых синапсов, входящих в состав сложного синапса, в суммарный потенциал нейрона

Двуслойная модель нейрона со сложными синапсами

Theorem

Для любого конечного и непротиворечивого набора прецедентов $<\widetilde{X}, \widetilde{Y}>$ можно построить агрегирующий нейрон

$$agn(x) = out \circ Agg_{\sum}\{\theta, s_1, ..., s_m\},$$

$$s_k = Agg_{\prod}\{u_{ki} : i \in\}$$

такой что $y = agn(x)$ для любой пары $< x, y >$ из $<\widetilde{X}, \widetilde{Y}>$

Заключение

- В дальнейшем предстоит теоретическое и экспериментальное исследование возможностей моделей агрегирующих нейронов по апроксимации различных типов зависимостей.
- Также предстоит еще исследовать эффект от применения различных классов агрегирующих функций как для агрегирования вкладов в сложных синапсах и синаптических кластерах, так и для агрегирования всех вкладов в суммарный потенциал нейрона на качество и сложность аппроксимации зависимостей.

СПАСИБО ЗА ВНИМАНИЕ!