



О.П. Солдатова, А.М. Котова, И.В. Моисеева

ПРИМЕНЕНИЕ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ ПАЦИЕНТОВ ПРИ ЛЕЧЕНИИ МЕТОДАМИ ЭКО

(ФГБОУ ВПО «Самарский государственный аэрокосмический университет им. академика С.П. Королева (национальный исследовательский университет)»)

В современном мире оказание медицинских услуг не обходится без применения информационных технологий. Во многих отраслях медицины используется обработка большого количества различных данных о пациентах, которые часто невозможно систематизировать. Решением данной проблемы является либо использование различных технологий для извлечения закономерностей из большого набора данных, либо использование знаний квалифицированных специалистов в системах искусственного интеллекта, таких как экспертные или нейросетевые системы. Авторами исследуются возможности нейронной сети типа многослойный персептрон для решения задачи классификации пациентов по типам протоколов лечения методами ЭКО.

Актуальность поставленной задачи можно описать следующим образом:

- репродуктивные центры нуждаются в системах, реализующих поддержку принятия решений при назначении лечения;
- инструментальные средства, которыми пользуются врачи, не удовлетворяют всех потребностей;
- применение современных информационных технологий может обеспечить создание полноценной системы, которая будет способна удовлетворить данные потребности.

Методы ЭКО используют три типа протокола лечения: длинный, короткий и короткий с антагонистами. Они существенно отличаются длительностью лечения, днями начала стимуляции, пункций. Также при разных протоколах варьируются препараты, принимаемые пациентами. Конкретный вид протокола выбирается врачом в соответствии с некоторыми правилами, чаще всего интуитивно. Для того, чтобы помочь врачу принять правильное решение, авторы доклада предлагают использовать реальные данные о пациентах и протоколах лечения для обучения нейронной сети.

Многослойный персептрон – одна из наиболее распространённых нейросетевых моделей, способная эффективно решать задачи классификации, к типу которых и относится описанная выше проблема. Многослойный персептрон состоит как минимум из одного скрытого слоя нейронов и выходного слоя. Размер выходного слоя определяется числом классов для распознавания (в данной задаче число классов равно трём), число нейронов в скрытых слоях подбирается для каждой задачи индивидуально при проведении экспериментальных исследований. В многослойном персептроне, как правило, используются сигмоидальные нейроны. Структура многослойного персептрона



с двумя скрытыми слоями представлена на рисунке 1 [1].

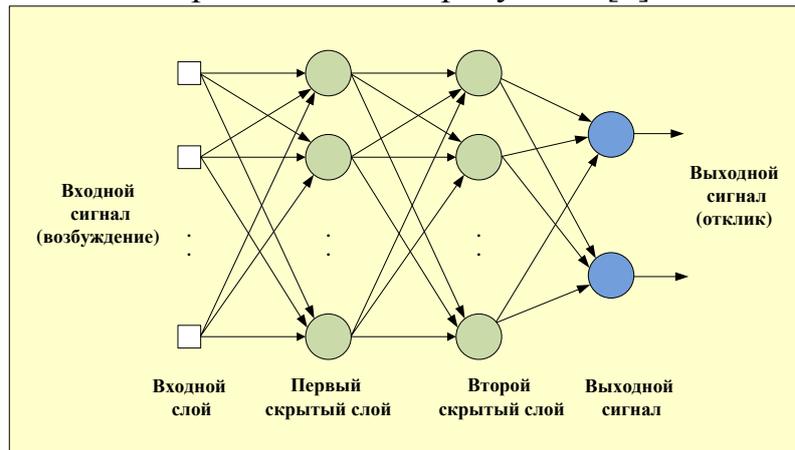


Рис. 1. Многослойный персептрон

Функционирование сигмоидального нейрона можно описать парой уравнений (1) и (2) [1].

$$u_k = \sum_{j=1}^N w_{kj} x_j, \quad (1)$$

$$y_k = \varphi(u_k + b_k), \quad (2)$$

где x_j – входные сигналы, w_{kj} – синаптические веса нейронов, u_k – линейная комбинация входных воздействий, φ – функция активации, y_k – выходной сигнал нейрона.

Нейронные сети не программируются в привычном смысле этого слова, они обучаются. Технически обучение заключается в нахождении коэффициентов синаптических связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными и выходными данными, а также обобщать полученные знания. Это значит, что в случае успешного обучения сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке, а также неполных и/или «зашумленных», частично искаженных данных.

Для обучения сети использовались прямая (Incremental) и пакетная (BatchProp) версии градиентного алгоритма наискорейшего спуска с моментом. Для расчёта компонентов вектора градиента использовался алгоритм обратного распространения ошибки [1], в котором на вход сети циклически подаются примеры из множества $\{(x(n), d(n))\}_{n=1}^N$, где d_n – желаемый отклик. Входные данные были предварительно нормализованы таким образом, чтобы среднее значение по всему обучающему множеству было близко к нулю. Это позволило упростить сравнение со среднеквадратическим отклонением (СКО) [1]:

$$e_j(n) = d_j(n) - o_j(n), \quad (3)$$

$$E_{av}(n) = \frac{1}{2N} \sum_{n=1}^N e_j^2(n). \quad (4)$$

После начальной инициализации весов и подачи обучающего примера осуществляется прямой проход сигнала по сети для подсчета индуцированных



локальных полей нейронов (5), выходных сигналов (6) и ошибки (3) [1].

$$v_j^{(l)} = \sum_i w_{ji}^{(l)}(n) y_i^{(l-1)}(n), \quad (5)$$

$$y_j^{(l)}(n) = \varphi_j(v_j^{(l)}(n)), \quad (6)$$

$$y_j^{(0)}(n) = x_j(n).$$

Затем осуществляется обратный проход сигнала ошибки для коррекции весов. Локальные градиенты и изменение синаптических весов вычисляются с помощью формул (7) и (8) [1]:

$$\delta_j^{(l)}(n) = \begin{cases} e_j^{(L)}(n) \varphi_j'(v_j^{(L)}(n)) \\ \varphi_j'(v_j^{(l)}(n)) \sum_k \delta_k^{(l+1)}(n) \delta_{kj}^{(l+1)}(n) \end{cases}, \quad (7)$$

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha [w_{ji}^{(l)}(n-1)] + \eta \delta_j^{(l)}(n) y_i^{(l-1)}(n), \quad (8)$$

где α – коэффициент момента, η – коэффициент скорости обучения, φ_j' – обозначает дифференцирование по аргументу функции активации. Критерием останова обучения является достижение заданной погрешности.

Также для обучения нейронной сети были реализованы два эвристических алгоритма: RProp и QuickProp [2].

Авторами разработана автоматизированная система классификации пациентов по протоколам лечения. Для обучения использовались данные пациентов, уже прошедших лечение в ГБУЗ СО «Клинический центр клеточных технологий».

В процессе экспериментальных исследований были определены оптимальные параметры сети: число входных и выходных нейронов, число нейронов в скрытых слоях, число скрытых слоев, а так же параметры обучения – коэффициентов момента и обучения. Была исследована зависимость СКО результатов обучения сети от коэффициента обучения при следующих параметрах обучения: тестовая выборка - 80 индексов; число итераций обучения 500; коэффициент момента $\alpha = 0,2$; входной слой – 28 нейронов, скрытый слой - 10, выходной – 3.

Результаты исследований представлены в таблице 1.

Таблица 1. Зависимость ошибки от коэффициента обучения

Коэффициент обучения	СКО
0,8	0,000000441
0,6	0,000045
0,4	0,00000552
0,2	0,0000412
0,1	0,000038

Дальнейшие исследования проведены при следующих параметрах топологии сети и алгоритма обучения: тестовая выборка - 80 индексов; число итераций обучения 500; коэффициент обучения $\eta = 0,8$; входной слой – 28 нейронов, скрытый слой - 10, выходной – 3.

Результаты исследований зависимости СКО от вида алгоритма,



используемого при обучении, представлены в таблице 2.

Таблица 2. Зависимость СКО от выбранного алгоритма обучения нейронной сети

Алгоритм	СКО
Incremental	0.00007203
BatchProp	0.00001917
RProp	0.000337
QuickProp	0.000383

Результаты исследований зависимости СКО тестирования сети от числа скрытых нейронов представлены в таблице 3.

Результаты исследований зависимости СКО тестирования сети от вида функции активации нейронов скрытого слоя представлены в таблице 4.

В результате проведённых исследований оптимальная конфигурация сети для эффективного решения поставленной задачи классификации оказалась следующая: во входном слое – 5 нейронов, в скрытом – 11, в выходном – 3, коэффициент обучения – 0.8, алгоритм пакетного обучения BatchProp, функция активации скрытых нейронов – Эллиота.

Таблица 3. Зависимость СКО от числа нейронов в скрытом слое

Число нейронов	СКО
5	0.001508
7	0,000136
9	0.0000821
11	0.000034
12	0.0003406
13	0.0001003
15	0.0001025

Таблица 4. Зависимость ошибки от функции активации

Функция	СКО
Сигмоидальная	0.00001218
Линейная	0.07089
Гаусса	0.002918
Эллиота	0.00000032
sin	0.000649
cos	0.0012187

Литература

1. Хайкин С. Нейронные сети: полный курс.- М.: Вильямс, 2006. – 1104 с.
2. Осовский С. Нейронные сети для обработки информации / Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.