

Огромному количеству исследователей нейронных сетей за их творческий вклад,

всем критикам за их замечания,

моим многочисленным аспирантам за их живой интерес

и моей супруге Нэнси за ее терпение и понимание

Предисловие

Нейронные сети, или, точнее, искусственные нейронные сети, представляют собой технологию, уходящую корнями во множество дисциплин: нейрофизиологию, математику, статистику, физику, компьютерные науки и технику. Они находят свое применение в таких разнородных областях, как моделирование, анализ временных рядов, распознавание образов, обработка сигналов и управление благодаря одному важному свойству — способности *обучаться* на основе данных при участии учителя или без его вмешательства.

Эта книга предлагает исчерпывающее описание нейронных сетей, учитывая многодисциплинарную природу этого предмета. Представленный в книге материал насыщен примерами, описанием компьютерных экспериментов, содержит множество задач по каждому разделу, а также обширную библиографию.

Книга состоит из четырех частей, организованных следующим образом.

Вводная часть (главы 1 и 2). В главе 1 на качественном уровне рассматривается сущность нейронных сетей, их свойства, возможности и связь с другими дисциплинами искусственного интеллекта. Глава завершается некоторыми исторически заметками. В главе 2 представлен обзор основных подходов к обучению и их статистических свойств. В этой главе вводится важное понятие *измерения Вапника–Червоненкиса*, используемое в качестве меры возможности обучаемой машины реализовать семейство функций классификации.

Обучение с учителем — основному предмету обсуждения посвящены главы 3–7. В главе 3 рассматривается простейший класс нейронных сетей — сети с одним или несколькими выходными нейронами, не содержащие скрытого слоя. В этой главе рассматриваются алгоритм метода наименьших квадратов (очень популярный при создании линейных адаптивных фильтров) и теорема о сходимости персептрона. В главе 4 приводится исчерпывающее описание *многослойного персептрона*, обучаемого на основе *алгоритма обратного распространения*. Этот алгоритм (представляющий собой обобщение метода наименьших квадратов) стал основной рабочей лошадкой нейронных сетей. В главе 5 содержится подробное математическое представление другого класса многослойных нейронных сетей — *сетей на основе радиальных базисных функций* (RBF-сетей), включающих отдельный слой базисных функций. Основной акцент в этой главе делается на роли теории регуляризации при создании RBF-сетей. В гла-

ве 6 рассматривается сравнительно новый класс обучаемых систем, известных под названием *машины опорных векторов*. Теория таких систем базируется на материале главы 2, посвященном статистической теории обучения. Вторая часть книги завершается главой 7, в которой обсуждаются *ассоциативные машины*, включающие в качестве своих компонентов несколько отдельно обучаемых систем. В этой главе рассматриваются три различных подхода к построению ассоциативных машин — методы *усреднения по ансамблю, усиления и иерархического смешения мнений экспертов*.

Обучению без учителя посвящены главы 8–12. В главе 8 *Хеббовское правило обучения* применяется для решения задачи *анализа главных компонент*. В главе 9 описывается еще одна форма обучения на основе самоорганизации, — *конкурентное обучение*, — применяемое для построения вычислительных отображений, получивших название *карт самоорганизации*. Эти две парадигмы отличаются в основном правилами обучения, уходящими корнями в область нейробиологии. Глава 10 посвящена *теории информации*, применяемой для создания алгоритмов обучения без учителя, а также их применению к решению задач *моделирования, обработки изображений и анализа независимых компонент*. В главе 11 рассматриваются вопросы обучения машин на основе принципов *статистической механики*. Эта тематика близко при- мыкает к материалу предыдущей главы, посвященной вопросам теории информации. В главе 12, которая завершает третью часть книги, вводится понятие *динамического программирования* и анализируется его взаимосвязь с *обучением с подкреплением*.

Нелинейные динамические системы — предмет изучения глав 13–15. В главе 13 описывается класс динамических систем на основе краткосрочной памяти и многослойных нейросетевых структур прямого распространения. В главе 14 основное внимание уделяется вопросам устойчивости, возникающим в нелинейных динамических системах при наличии *обратной связи*. Приводятся примеры *ассоциативной памяти*. В главе 15 рассматривается еще один класс нелинейных динамических систем — *рекуррентные сети*, основанные на использовании обратной связи при построении отображения вход-выход.

В заключении кратко анализируется роль нейронных сетей при создании *интеллектуальных машин* распознавания образов, управления и обработки сигналов.

Структура книги обеспечивает большую гибкость при разработке различных курсов обучения нейронным сетям. Выбор конкретных тем для изучения должен определяться только интересами аудитории, использующей эту книгу.

В книге описывается пятнадцать компьютерных экспериментов, тринадцать из которых рассчитаны на использование MATLAB. Файлы для выполнения этих экспериментов в среде MATLAB можно найти по адресу <ftp://ftp.mathworks.com/pub/books/haykin> или <http://www.mathworks.com/books/>

Во втором случае пользователь должен щелкнуть на ссылке Neural/Fuzzy, а затем — на названии книги. Такой подход обеспечивает более удобный интерфейс пользователя.

24 Предисловие

Каждая глава завершается несколькими упражнениями. Многие из этих задач являются творческими и призваны не только проверить качество усвоения материала, но и подтолкнуть к дальнейшему развитию соответствующей тематики. Решения всех задач содержатся в руководстве, копии которого могут получить преподаватели, использующие в своих курсах настоящую книгу. Для этого они должны обратиться в издательство Prentice Hall.

Книга рассчитана на инженеров, специалистов в области компьютерных наук и физиков. Автор надеется, что книга будет интересна также специалистам в других областях, таких как психология и нейробиология.

Саймон Хайкин,
Гамильтон, Онтарио,
февраль 1998 г.

Благодарности

Автор выражает глубокую благодарность многим рецензентам, которые потратили свое время на чтение этой книги или отдельной ее части. Хочу сказать большое спасибо доктору Кеннету Роузу (Kenneth Rose) из университета штата Калифорния в Санта-Барбара за существенный конструктивный вклад в написание книги и значительную помощь.

Я благодарен доктору С. Амари (S. Amari) из Японии; докторам Сью Беккер (Sue Becker) и Рону Рэйсину (Ron Racine) из Макмастерского университета; доктору Син Холден (Sean Holden) из университетского колледжа в Лондоне; доктору Майклу Турмону (Michael Turmon), Пасадена; доктору Бабаку Хассиби (Babak Hassibi) из Стэндфордского университета; доктору Раюлю Ёйи (Paul Yee), бывшему сотруднику Макмастерского университета, доктору Эдгару Осуна (Edgar Osuna) из Массачусетского технологического института; доктору Бернерду Шелкопфу (Bernard Schelkopf) из Института Макса Планка (Германия); доктору Майклу Джордану (Michael Jordan) из Массачусетского технологического института; докторам Редфорду Ниалу (Radford Neal) и Зубину Гархамани (Zoubin Gharhamani) из университета Торонто, доктору Джону Тситсиклису (John Tsitsiklis) из Массачусетского технологического института; доктору Джосу Принципу (Jose Principe) из Университета штата Флорида; Гинту Пушкориусу (Gint Puskorius) и доктору Ли Фелдкамп (Lee Feldkamp) из исследовательской лаборатории Форда; доктору Ли Гилесу (Lee Giles) из исследовательского института компании NEC в Принстоне, штат Нью-Джерси; доктору Микелю Форкада (Mikel Forcada) из университета Далаканта (Испания); доктору Эрику Вану (Eric Wan) из института науки и технологии штата Орегона; а также доктору Яну Лекуну (Yann LeCun) из компании AT&T Research, штат Нью-Джерси, за их неоценимую помощь в улучшении данной книги.

Мне также хочется поблагодарить доктора Ральфа Линскера (Ralph Linsker) из компании IBM; доктора Стюарта Гемана (Stuart Geman) из Браунского университета; доктора Алана Гелфорда (Alan Gelford) из университета штата Коннектикут; доктора Барта Коско (Bart Kosko) из университета Южной Калифорнии; доктора Нарыша Синха (Narish Sinha) из Макмастерского университета; доктора Костаса Диамантараса (Kostas Diamantaras) из университета г. Тессалоники имени Аристотеля (Греция); доктора Роберта Якобса (Robert Jacobs) из университета Рочестера; доктора Эндрю Барто (Andrew Barto) из Массачусетского университета; доктора Дона Хуша (Don Hush) из университета штата Нью-Мексико; доктора Йошуа Бенжио (Yoshua Bengio) из университета г. Монреаль; доктора Х. Янга (H. Yang) из института науки и технологий штата Орегона; доктора Скотта Дугласа (Scott Douglas) из университета штата Юта; доктора Натана Интратора (Nathan Intrator) из университета г. Тель-Авив (Израиль); доктора Владимира Вапника (Vladimir Vapnik) из компании AT&T Research, штат Нью-Джерси; доктора Тейво Кохонена (Teuvo Kohonen) из Хельсинского тех-

нологического университета (Финляндия); доктора Владимира Черкасского (Vladimir Cherkassky) из университета штата Миннесота; доктора Дэвида Лоува (David Lowe) из Астонского университета (Великобритания); доктора Джеймса Андерсона (James Anderson) из Браунского университета; доктора Андреаса Андрео (Andreas Andreou) из университета Джона Хопкина и доктора Томаса Анастасио (Thomas Anastasio) из университета штата Иллинойс.

Я глубоко благодарен моему студенту-дипломнику Хью Пасика (Hugh Pasika) за выполнение множества компьютерных экспериментов в среде MATLAB, а также за создание Web-узла этой книги. Большое спасибо также студенту-дипломнику Химешу Мадхуранату (Himesh Madhuranath); доктору Садасивану Путуссерипади (Sadasivan Puthusseripady); доктору Дж. Ни (J. Nie); доктору Паулю Йии (Paul Yee) и Гинту Пускориусу (Gint Puskorius) из лаборатории Ford Research за выполнение пяти экспериментов.

Большое спасибо также Хью Пасику (Hugh Pasika) за вычитку всей книги. В этой связи хочется также поблагодарить доктора Роберта Дони (Robert Dony), доктора Стефана Кремера (Stefan Kremer) и доктора Садасивана Путуссерипади (Sadasivan Puthusseripady) за вычитку отдельных глав книги.

Автор также выражает благодарность своему издателю Тому Роббинсу (Tom Robbins) и издателю Алис Дворкин (Alice Dworkin) за их полную поддержку и одобрение. Хочется поблагодарить Дженифер Маган (Jennifer Maughan) и сотрудников компании WestWorld, Inc. (штат Юта) за издание этой книги.

Автор выражает глубокую благодарность Бриджитт Майер (Brigitte Maier) из библиотеки Макмастерского университета за ее неоценимые усилия по поиску очень сложных ссылок и созданию столь полной библиографии. Следует отметить также помощь научного сотрудника библиотеки Пегги Финдлей (Peggy Findlay) и сотрудника библиотеки Регины Бендиг (Regina Bendig).

И наконец (но не в последнюю очередь), автор благодарен своему секретарю Лоле Брукс (Lola Brooks) за оформление различных версий рукописи. Без ее помощи написание этой книги и ее издание потребовали бы значительно больше времени.

Основные обозначения

Важные символы

a	Действие
$\mathbf{a}^T \mathbf{b}$	Скалярное произведение векторов \mathbf{a} и \mathbf{b}
$\mathbf{a} \mathbf{b}^T$	Внешнее произведение векторов \mathbf{a} и \mathbf{b}
$\binom{l}{m}$	Биномиальный коэффициент
$A \cup B$	Объединение A и B
B	Величина, обратная температуре
b_k	Внешнее смещение, применяемое к нейрону k
$\cos(\mathbf{a}, \mathbf{b})$	Косинус угла между векторами \mathbf{a} и \mathbf{b}
D	Глубина памяти
$D_{f g}$	Дивергенция Куллбека–Леблера между функциями плотности вероятности f и g
\mathbf{D}	Оператор, сопряженный \mathbf{D}
E	Функция энергии
E_i	Энергия состояния i в статистической механике
E	Статистический оператор ожидания
$\langle E \rangle$	Средняя энергия
erf	Функция ошибки
erfc	Дополнительная функция ошибки
exp	Экспонента
\mathbf{E}_{av}	Среднеквадратическая ошибка или сумма квадратических ошибок
$\mathbf{E}(n)$	Мгновенное значение суммы квадратических ошибок
F	Свободная энергия
$f_X(\mathbf{x})$	Функция плотности вероятности для случайного вектора \mathbf{X}
\mathbf{F}^*	Подмножество (сетей) с минимальным эмпирическим риском
\mathbf{H}	Матрица Гессиана
\mathbf{H}^{-1}	Матрица, обратная Гессиану
i	Квадратный корень из -1 , обозначаемый также как j
\mathbf{I}	Единичная матрица
\mathbf{I}	Информационная матрица Фишера
J	Среднеквадратическая ошибка
\mathbf{J}	Матрица Якобиана
$\mathbf{K}(n, n - 1)$	Матрица ковариации ошибки в теории фильтрации Калмана

$\mathbf{K}^{1/2}$	Квадратный корень из матрицы \mathbf{K}
$\mathbf{K}^{T/2}$	Транспонированный квадратный корень из матрицы \mathbf{K}
k_B	Константа Больцмана
\log	Логарифм
$L(\mathbf{w})$	Функция логарифмического правдоподобия от вектора весов \mathbf{w}
$\mathbf{L}(\mathbf{w})$	Функция логарифмического правдоподобия от вектора весов \mathbf{w} , построенная на основе одного примера
\mathbf{M}_c	Матрица управляемости
\mathbf{M}_o	Матрица наблюдаемости
n	Дискретное время
p_i	Вероятность состояния i в статистической механике
p_{ij}	Вероятность перехода из состояния i в состояние j
\mathbf{P}	Стохастическая матрица
P_c	Вероятность корректной классификации
P_e	Вероятность ошибки
$P(e \mathbf{C})$	Условная вероятность ошибки e при условии выбора входного воздействия из класса \mathbf{C}
p_α^+	Вероятность нахождения видимого нейрона машины Больцмана в состоянии α при условии фиксированной сети (положительная фаза)
p_α^-	Вероятность нахождения видимого нейрона машины Больцмана в состоянии α при условии свободного функционирования сети (отрицательная фаза)
$\hat{r}_x(j, k; n)$	Оценка автокорреляционной функции $x_j(n)$ и $x_k(n)$
$\hat{r}_{dx}(k; n)$	Оценка функции взаимной корреляции $d(n)$ и $x_k(n)$
\mathbf{R}	Матрица корреляции входного вектора
t	Непрерывное время
T	Температура
\mathbf{T}	Обучающее множество (выборка)
tr	След матрицы
var	Оператор дисперсии
$V(\mathbf{x})$	Функция Ляпунова от вектора состояния \mathbf{x}
v_j	Индукционное локальное поле или активационный потенциал нейрона j
\mathbf{w}_o	Оптимальное значение вектора синаптических весов
w_{kj}	Синаптический вес синапса j , принадлежащего нейрону k
\mathbf{w}^*	Оптимальный вектор весов
$\bar{\mathbf{x}}$	Состояние равновесия вектора состояния \mathbf{x}
$\langle x_j \rangle$	Среднее значение состояния x_j в термальном смысле
\hat{x}	Оценка x , обозначенная символом “крышки”

$ x $	Абсолютное значение (амплитуда) x
x^*	Комплексно сопряженное значение x , обозначенное звездочкой
$\ \mathbf{x}\ $	Евклидова норма (длина) вектора \mathbf{x}
\mathbf{x}^T	Вектор, транспонированный к \mathbf{x}
z^{-1}	Оператор единичной задержки
Z	Функция разбиения
$\delta_j(n)$	Локальный градиент нейрона j в момент времени n
Δw	Малое изменение вектора весов w
∇	Оператор градиента
∇^2	Оператор Лапласа
$\nabla_w J$	Градиент J по w
$\nabla \cdot \mathbf{F}$	Дивергенция вектора \mathbf{F}
η	Коэффициент скорости обучения
κ	Семиинвариант
μ	Стратегия
θ_k	Пороговое значение нейрона k (т.е. величина смещения b_k с обратным знаком)
λ	Параметр регуляризации
λ_k	k -е собственное значение квадратной матрицы
$\Phi_k(\cdot)$	Нелинейная функция активации нейрона k
\in	Символ принадлежности множеству
\cup	Символ объединения множеств
\cap	Символ пересечения множеств
$*$	Символ конволюции
$+$	Символ псевдообращения матрицы

Открытые и закрытые интервалы

Открытый интервал (a, b) изменения переменной x означает, что $a < x < b$.

Закрытый интервал $[a, b]$ изменения переменной x означает, что $a \leq x \leq b$.

Полуоткрытый интервал $[a, b)$ изменения переменной x означает, что $a \leq x < b$.

Аналогично, интервал $(a, b]$ изменения переменной x означает, что $a < x \leq b$.

Минимумы и максимумы

Символ $\arg \min_{\mathbf{w}} f(\mathbf{w})$ обозначает значение вектора аргумента \mathbf{w} , доставляющее минимум функции $f(\mathbf{w})$.

Символ $\arg \max_{\mathbf{w}} f(\mathbf{w})$ обозначает значение вектора аргумента \mathbf{w} , доставляющее максимум функции $f(\mathbf{w})$.