Станислав Осовский

Нейронные сети для обработки информации

Оглавление

К читателю ... 3
Предисловие ... 4
Предисловие к российскому изданию ... 5

1. Введение ... 9
1.1. Биологические основы функционирования нейрона ... 9
1.2. Первые модели нейронной сети ... 13
1.3. Прикладные возможности нейронных сетей ... 15

2. Модели нейронов и методы их обучения ... 18
2.1. Персептрон ... 19
2.2. Сигмоидальный нейрон ... 21
2.3. Радиальный нейрон ... 26
2.4. Нейрон типа «адалайн» ... 27
2.5. Инстар и оутстар Гроссберга ... 29
2.6. Нейроны типа WTA ... 32
2.7. Модель нейрона Хебба ... 36
2.8. Стохастическая модель нейрона ... 40
2.9. Задания ... 41

3. Однонаправленные многослойные сети сигмоидального типа ... 42
3.1. Однослойная сеть ... 43
3.2. Многослойная персептронная сеть ... 46
3.2.1. Структура персептронной сети ... 46
3.2.2. Алгоритм обратного распространения ... 47
3.3. Применение потоковых графов для генерации градиента ... 51
3.4. Градиентные алгоритмы обучения сети ... 56
3.4.1. Алгоритм наискорейшего спуска ... 58
3.4.2. Алгоритм переменной метрики ... 60
3.4.3. Алгоритм Левенберга–Марквардта ... 61
3.4.4. Алгоритм сопряженных градиентов ... 63
3.5. Подбор коэффициента обучения ... 64
3.6. Эвристические методы обучения сети ... 68
3.6.1. Алгоритм Quickprop ... 69
3.6.2. Алгоритм RPROP ... 70
3.7. Компьютерная программа *MLP* для обучения персептронной сети ... 71

3.8. Сравнение эффективности алгоритмов обучения 72
3.9. Элементы глобальной оптимизации 78
3.9.1. Алгоритм имитации отжига 80
3.9.2. Элементы теории генетических алгоритмов 84
3.10. Методы инициализации весов 89
3.11. Задания 92
4. Проблемы практического использования искусственных нейронных сетей 94
4.1. Способность нейронной сети к обобщению 94
4.1.1. Основные зависимости 94
4.1.2. Мера VCdim 95
4.1.3. Зависимости между погрешностью обобщения и мерой VCdim 96
4.1.4. Обзор методов повышения способности нейронной сети к обобщению 99
4.2. Предварительный подбор архитектуры сети 104
4.3. Подбор архитектуры сети, оптимальной с позиций обобщения 107
4.3.1. Методы редуцирования сети с учетом чувствительности 108
4.3.2. Методы редуцирования сети с использованием штрафной функции 113
4.4. Добавление шума в обучающие выборки 116
4.5. Повышение способности к обобщению путем использования нескольких сетей 118
4.6. Примеры использования персептронной сети 121
4.6.1. Распознавание и классификация двоичных образов . 121
4.6.2. Распознавание образов на основе внешнего описания 130
4.6.3. Нейронная сеть для сжатия данных 136
4.6.4. Идентификация динамических объектов 140
4.6.5. Прогнозирование нагрузок энергетической системы . 144
4.7. Задания 148
5. Радиальные нейронные сети 150
5.1. Математические основы 151
5.2. Радиальная нейронная сеть 153
5.3. Методы обучения радиальных нейронных сетей 159
5.3.1. Применение процесса самоорганизации для адаптации параметров радиальных функций 161
5.3.2. Вероятностный алгоритм подбора параметров радиальных функций 164
5.3.3. Гибридный алгоритм обучения радиальных сетей ... 166
5.3.4. Алгоритмы обучения, основанные на обратном распространении 168

5.4. Методы подбора количества базовых функций 171
5.4.1. Эвристические методы 172
5.4.2. Метод ортогонализации Грэма–Шмидта 173
5.5. Компьютерная программа для обучения радиальных сетей 177
5.6. Пример использования радиальной сети для аппроксимации 179
5.7. Сравнение радиальных и сигмоидальных сетей 182
5.8. Задания 183
6. Сети SVM 185
6.1. Линейная сеть SVM для решения задачи классификации 186
6.1.1. Нелинейная сеть SVM для решения задачи классификации 192
6.2. Интерпретация множителей Лагранжа в структуре сети 201
6.3. Задача классификации при наличии нескольких классов 202
6.4. Сети SVM для решения задач регрессии 204
6.5. Обзор алгоритмов решения двойственной задачи 207
6.6. Компьютерная программа для обучения сети SVM .. 211
6.7. Примеры использования сетей SVM 214
6.7.1. Задача классификации двух спиралей 214
6.7.2. Распознавание текстур 216
6.7.3. Выявление неисправностей элементов в контуре электрического фильтра 219
6.8. Сравнение сети SVM с другими нейронными решениями .....
6.9. Задания 224
7. Специализированные структуры нейронных сетей . 225
7.1. Сеть каскадной корреляции Фальмана 225
7.2. Сеть Вольтерри 231
7.2.1. Структура и особенности обучения сети 232
7.2.2. Примеры использования сети Вольтерри 235
7.3. Задания 241
8. Рекуррентные сети как ассоциативные запоминающие устройства 243
8.1. Автоассоциативная сеть Хопфилда 245
8.1.1. Основные зависимости 245
8.1.2. Режим обучения сети Хопфилда 249
8.1.3. Режим распознавания сети Хопфилда 250

8.1.4. Программа Hop_win ... 251
8.2. Сеть Хемминга ... 254
8.2.1. Структура сети и алгоритм подбора весов ... 254
8.2.2. Функционирование сети Хемминга ... 256
8.2.3. Программа Shamming обучения сети ... 258
8.3. Сеть типа BAM ... 260
8.3.1. Описание функционирования сети ... 260
8.3.2. Модифицированный алгоритм обучения сети BAM ... 263
8.3.3. Модифицированная структура сети BAM ... 264
8.4. Задания ... 270
9. Рекуррентные сети на базе персептрона ... 272
9.1. Персептронная сеть с обратной связью ... 272
9.1.1. Структура сети RMLP ... 272
9.1.2. Алгоритм обучения сети RMLP ... 274
9.1.3. Подбор коэффициента обучения ... 276
9.1.4. Коэффициент усиления сигнала ... 277
9.1.5. Результаты компьютерного моделирования ... 277
9.2. Рекуррентная сеть Эльмана ... 282
9.2.1. Структура сети ... 282
9.2.2. Алгоритм обучения сети Эльмана ... 283
9.2.3. Обучение с учетом момента ... 286
9.2.4. Пример компьютерного моделирования сети Эльмана 287
9.3. Сеть RTRN ... 291
9.3.1. Структура сети и алгоритм обучения ... 291
9.3.2. Результаты численных экспериментов ... 293
9.4. Задания ... 297
10. Сети с самоорганизацией на основе конкуренции ... 298
10.1. Особенности сетей с самоорганизацией на основе конкуренции ... 299
10.1.1. Меры расстояния между векторами ... 300
10.1.2. Нормализация векторов ... 301
10.1.3. Проблема мертвых нейронов ... 302
10.2. Алгоритмы обучения сетей с самоорганизацией ... 304
10.2.1. Алгоритм Кохонена ... 305
10.2.2. Алгоритм нейронного газа ... 306
10.2.3. Программа Kohon ... 308
10.2.4. Сравнение алгоритмов самоорганизации ... 309
10.3. Сеть отображения одно- и двумерных данных ... 312
10.4. Отображение Сэммона ... 315
10.5. Применение сетей с самоорганизацией ... 316
10.5.1. Компрессия данных ... 316

10.5.2. Выявление неисправностей оборудования 319
10.5.3. Краткосрочное прогнозирование нагрузок энергетической системы 323
10.6. Гибридная сеть 327
10.7. Задания 331
11. Сети с самоорганизацией корреляционного типа 332
11.1. Энергетическая функция корреляционных сетей 332
11.2. Нейронные сети PCA 334
11.2.1. Математическое введение 334
11.2.2. Отношение между преобразованиями PCA и SVD . 338
11.2.3. Определение первого главного компонента 339
11.2.4. Алгоритмы определения нескольких главных компонентов 340
11.3. Нейронные сети для слепого разделения сигналов ... 343
11.3.1. Статистическая независимость сигналов 344
11.3.2. Рекуррентная структура разделяющей сети 344
11.3.3. Алгоритм Херольта–Джуттена для рекуррентной сети 346
11.3.4. Алгоритм Чихотского для обучения рекуррентной сети 347
11.3.5. Программа слепого разделения сигналов BS 348
11.3.6. Однонаправленная сеть для разделения сигналов .. 352
11.3.7. Toolbox ICALAB 357
11.4. Задания 359
12. Математические основы нечетких систем 360
12.1. Операции на нечетких множествах 362
12.2. Меры нечеткости нечетких множеств 364
12.3. Нечеткость и вероятность 365
12.4. Нечеткие правила вывода 367
12.5. Системы нечеткого вывода Мамдани–Заде 368
12.5.1. Фуззификатор 371
12.5.2. Дефуззификатор 374
12.5.3. Модель Мамдани–Заде как универсальный аппроксиматор 376
12.6. Модель вывода Такаги–Сугено–Канга 377
12.7. Задания 379
13. Нечеткие нейронные сети 381
13.1. Структура нечеткой сети TSK 381
13.2. Структура сети Ванга–Менделя 385
13.3. Алгоритмы самоорганизации для обучения нечеткой сети 387
13.3.1. Алгоритм пикового группирования 387

13.3.2. Алгоритм C-means 391
13.3.3. Алгоритм нечеткой самоорганизации Густафсона–Кесселя 393
13.4. Генерация правил вывода нечеткой сети 398
13.5. Гибридный алгоритм обучения нечеткой сети TSK .. 402
13.6. Модификации сети TSK 406
13.6.1. Алгоритм определения количества правил вывода . 407
13.6.2. Численный пример 409
13.6.3. Упрощенная сеть TSK 412
13.7. Гибридная нечеткая сеть 414
13.8. Примеры использования нечетких сетей 417
13.8.1. Оценивание концентрации компонентов газовой смеси 417
13.8.2. Распознавание компонентов газовой смеси 418
13.8.3. Распознавание сортов пива по запаху 421
13.9. Адаптивный алгоритм самоорганизации нечеткой сети 424
13.10. Задания 427
Литература 428
Предметный указатель 439