

РОССИЙСКАЯ АКАДЕМИЯ НАУК
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ГОСУДАРСТВЕННАЯ КОРПОРАЦИЯ ПО АТОМНОЙ ЭНЕРГИИ «РОСАТОМ»
РОССИЙСКАЯ АССОЦИАЦИЯ НЕЙРОИНФОРМАТИКИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЯДЕРНЫЙ УНИВЕРСИТЕТ «МИФИ»
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ
СИСТЕМНЫХ ИССЛЕДОВАНИЙ РАН

–2010

НЕЙРОИНФОРМАТИКА – 2010

XII

-

По материалам Школы-семинара
«Современные проблемы нейроинформатики»

Москва 2010

УДК 001(06)+004.032.26 (06) Нейронные сети
ББК 72я5+32.818я5
М82

–2010. XII -
« -2010»: -
. – М.: НИЯУ МИФИ, 2010. – 328 с.

В книге публикуются тексты лекций, прочитанных на Школе-семинаре «Современные проблемы нейроинформатики», проходившей 26–29 января 2010 года в НИЯУ МИФИ в рамках XII Всероссийской конференции «Нейроинформатика–2010».

Материалы лекций связаны с рядом проблем, актуальных для современного этапа развития нейроинформатики, включая ее взаимодействие с другими научно-техническими областями.

Ответственный редактор
Ю. В. Тюменцев, кандидат технических наук

ISBN 978–5–7262–1225–8 © *Национальный исследовательский
ядерный университет «МИФИ», 2010*

И. Г. Персианцев.

	212
Введение	213
Основополагающие идеи метода	213
Алгоритм формирования групп классов	214
Алгоритм построения ИНК	215
Свойства алгоритма построения ИНК	216
Результаты численных экспериментов	217
Одномерное распределение Кантора	218
Распознавание текстур	221
Распознавание гласных	222
Дикторонезависимое распознавание изолированных слов	224
Нейросетевая сегментация временных рядов	227
Описание алгоритма анализа временных рядов с переключающейся динамикой	228
Описание алгоритма анализа временных рядов с дрейфовой динамикой	230
Эксперименты с псевдо-хаотическими временными рядами	232
Эксперименты с биомедицинскими данными	236
Работа с данными космофизического эксперимента	238
Заключение	239
Литература	240

НИИ Ядерной Физики МГУ, Москва,
E-mail: ipers@srd.sinp.msu.ru

Предложен алгоритм построения самоорганизующегося классификатора на основе многослойного персептрона. В результате работы алгоритма создается классификатор, имеющий древовидную структуру, каждый узел которой представляет собой персептрон. Структура дерева и параметры каждого персептрона формируются автоматически. При работе с данными, содержащими большое количество классов, предложенный алгоритм позволяет существенно снизить вычислительную стоимость решения задачи. Алгоритм был протестирован на ряде модельных, тестовых и практических задач. Предложена модификация алгоритма, допускающая его использование

Классификация данных и распознавание образов является одной из важнейших задач обработки результатов физических измерений в различных областях науки и техники. Нейронные сети (НС) являются одним из классов алгоритмов, часто применяемых для решения подобных задач. Отличительными чертами НС являются их способность обучаться на примерах и обобщать полученную информацию, невосприимчивость к шуму и присутствующая им параллельность архитектуры. При решении практических задач классификации и распознавания образов с помощью традиционных нейросетевых подходов существует ряд принципиально важных недостатков и ограничений, наиболее значимыми среди которых являются: сложность решения задач классификации в случае плохо разделимых классов, субъективность выбора алгоритма решения конкретной задачи и его параметров, возможность попадания в локальный минимум, заметное снижение качественных характеристик НС классификатора при большом количестве классов, неоправданно высокая вычислительная стоимость построения НС классификатора при большом количестве классов, проблема интерпретации признаков, по которым НС принимает решение.

Целью исследований являлось преодоление указанных выше недостатков путем разработки конструктивного нейросетевого алгоритма классификации, обладающего следующими свойствами: наращивание архитектуры НС по мере необходимости в процессе тренировки; сокращение вычислительной стоимости решения задачи классификации; обеспечение оптимального выбора признаков, используемых нейросетью для классификации данных, а также универсальность алгоритма, т. е. применимость его к возможно более широкому кругу задач.

Метод адаптивного построения

() основан на следующих соображениях [1–6].

Процесс тренировки многослойного персептрона (МСП) можно рассматривать как одновременное выделение признаков в скрытом слое и их декодирование в выходном слое. Для простейшей архитектуры МСП количество выделяемых признаков можно считать пропорциональным количеству нейронов в скрытом слое. При этом задание числа нейронов в скрытом

слое меньше некоторого порога приведет к ситуации, когда НС будет не способна выделить количество признаков, достаточное для классификации предложенных образов в требуемое количество классов. НС начнет путать образы из классов, похожих в терминах того небольшого количества признаков, которые выделяются в скрытом слое.

Можно предположить, что при некоторых условиях нейросеть будет в этой ситуации выделять наиболее существенные признаки во входных данных и, следовательно, эти признаки можно использовать для принятия решения о том, какие классы следует объединить в группы.

Основанный на вышеизложенной идее алгоритм формирования групп классов (АФГК) работает следующим образом.

Шаг 1. Вначале МСП обучается решать задачу классификации с использованием стандартного метода обратного распространения ошибки [7].

Шаг 2. По прошествии нескольких эпох тренировки, алгоритм производит анализ статистики ответов МСП на тренировочном наборе. В процессе анализа проводится «голосование» образов: для каждого образа (примера) определяется номер нейрона выходного слоя, имеющего максимальную амплитуду отклика и превышающего некий порог голосования. Этот номер трактуется как номер класса, за принадлежность к которому «проголосовал» данный образ. Если количество образов из класса k , проголосовавших за принадлежность к классу j , больше половины количества образов в классе k , то считается, что все представители класса k отныне принадлежат классу j . Другими словами, если ответом сети на образы из разных классов является стабильная активация одного и того же нейрона, то такая ситуация рассматривается как свидетельство в пользу того, что эти классы следует объединить.

Шаг 3. После окончания анализа всех классов, производится соответствующая модификация желаемых выходов. После этого тренировка продолжается с использованием измененной информации о классе, к которому принадлежит каждый образ.

Приведем формальное описание алгоритма АФГК.

Шаг 1. Обучение по методу обратного распространения ошибки

$$\forall p \in C_k \quad O_i(p) = \delta_{ik}$$

Шаг 2. Анализ статистики ответов

Возможный вариант — «голосование по порогу»:

$$\forall p \in C_k$$

IF $\operatorname{argmax}(Y(p)) = j$ AND $Y_j(p) > \Theta$ THEN

$$V_j = V_j + 1$$

Шаг 3. Модификация желаемого ответа

IF $V_j > N(C_k)/2$ THEN

$$\forall p \in C_k \quad O_i(p) = \delta_{ik}$$

Указанные выше стадии «тренировка–анализ–модификация желаемых ответов» повторяются, пока классы продолжают объединяться. В результате описанной процедуры исходный набор классов разбивается на некоторое количество групп и одновременно с этим создается классификатор, поддерживающий именно это разбиение.

Опишем как процесс построения дерева решений, узлы в котором реализованы в виде МСП с заданным размером скрытого слоя. Построение ИНК начинается с вызова АФГК, в результате работы которого происходит объединение исходных классов в небольшое количество групп. Таким образом, формируется базовый узел ИНК (первый уровень иерархии), в котором осуществляется наиболее «грубая» классификация. Очевидно, что полученные группы классов можно рассматривать как ветви дерева решений.

После создания базового узла, для каждой из полученных групп классов (ветвей) вновь вызывается АФГК. Важно, что при этом алгоритм анализирует лишь те исходные классы, которые попали в данную ветвь дерева. В результате формируется узел следующего уровня иерархии, а из анализируемых классов образуются новые группы (ветви). После формирования всех узлов данного уровня иерархии, АФГК вызывается для каждой из

вновь полученных ветвей. Ясно, что узлы более высоких уровней иерархии осуществляют все более и более детальную классификацию.

Алгоритм построения ИНК для данной ветви останавливается, если в ней содержится один исходный класс, или если дальнейшая детализация классов в этой ветви невозможна (например, если процент ошибочных классификаций в узле превысил некоторый порог).

Можно трактовать процесс построения ИНК как процесс подбора адекватной сложности задачи на каждом уровне иерархии. Пусть задана архитектура МСП, используемая при построении ИНК (для простоты единственным параметром архитектуры будем считать размер скрытого слоя). В процессе формирования очередного узла возможны три результата работы АФГК:

- 1) Заданная архитектура МСП способна разделить классы, входящие в данную ветвь, т. е. сложность поставленной в данном узле задачи классификации соответствует возможностям используемой НС.
- 2) Заданная архитектура МСП не может разделить классы, входящие в данную ветвь, но возможно образование групп классов, т. е. задача классификации, слишком сложная для решения в данном узле, разбивается на задачу разбиения классов по группам (причем сложность этой задачи адекватна возможностям используемой НС) и на набор задач для узлов следующего уровня иерархии. Заметим, что при использовании такой стратегии задача тренировки узла ИНК может быть решена достаточно быстро, поскольку предложенный метод модификации желаемых ответов сильно ускоряет сходимость обучения. С другой стороны, количество классов, подлежащих распознаванию, постепенно уменьшается с каждым следующим уровнем иерархии, что упрощает задачу классификации.
- 3) На некотором этапе построения ИНК может сформироваться ветвь, внутри которой границы между классами настолько сложны, что заданная архитектура МСП не в состоянии ни разделить классы, ни сформировать из них группы, что говорит о невозможности решить задачу классификации средствами заданной архитектуры.

Как видно из этого рассмотрения, первоначальная глобальная задача распознавания множества исходных классов (возможно, очень сложная)

разбивается на последовательность локальных задач со сложностью, адекватной возможностям нейросетевой архитектуры, используемой в узлах ИНК. Преимущества иерархического способа решения задачи перед использованием одиночного МСП с большим скрытым слоем достаточно очевидны. Хотя одиночный МСП и может решить задачу (по теореме об МСП как универсальном аппроксиматоре [8]), однако требуемый размер скрытого слоя заранее неизвестен, и его приходится находить методом подбора. В то же время, в предлагаемом иерархическом подходе архитектура МСП может быть практически произвольной, поскольку алгоритм наращивает структуру ИНК по мере необходимости, используя заданную архитектуру как конструкционные блоки.

Докажем теорему, позволяющую оценить количество скрытых нейронов МСП, достаточное для построения ИНК со 100% распознаванием.

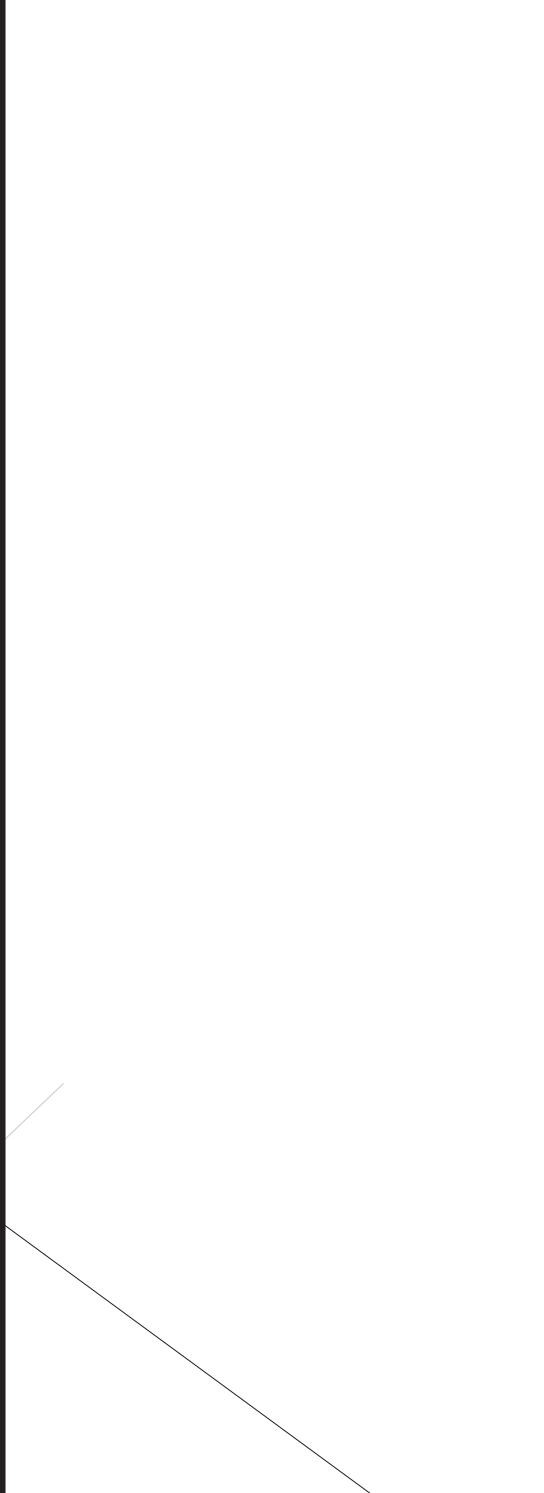
. Пусть требуется классифицировать P примеров $X \in R^N$ в M классов. Тогда для построения ИНК, дающего гарантированное 100% распознавание, достаточно использовать МСП с размером скрытого слоя, равным $\min(2 \cdot N, P - 1)$.

. Действительно, пусть в узле ИНК требуется классифицировать P примеров $X \in R^N$ в два класса \mathbf{K}_1 и \mathbf{K}_2 (если количество классов больше 2, то объединим эти классы на два новых класса произвольным образом). Рассмотрим пример $X \in \mathbf{K}_1$. В наихудшем случае, его требуется отделить от $(P - 1)$ примеров из \mathbf{K}_2 . Для этого достаточно:

- либо построить набор из $(P - 1)$ гиперплоскостей, отделяющих X от каждого из примеров из \mathbf{K}_2 , на что требуется $(P - 1)$ скрытых нейронов;
- либо для каждого из признаков X_j ($j = 1, N$) построить по 2 гиперплоскости, отделяющих данный пример от примеров с меньшими и с большими значениями этого признака, для чего требуется $(2 \cdot N)$ скрытых нейронов.

Таким образом, теорема доказана.

Для проверки алгоритма был выбран ряд модельных задач, а также известные задачи реального мира, чьи базы данных доступны по сети Интернет. Ниже приводятся основные результаты проверки алгоритма на некоторых из них.



пределения для одномерного распределения Кантора, пунктирной – пример классифицируемого образа.

Признаком, определяющим принадлежность образа к тому или иному классу, являлась координата его максимума. Каждый из 8 пиков на рис. 1 был разделен на два равных сегмента, что в сумме давало 16 исходных классов для базовой сети. Тренировочный и тестовый наборы содержали по 320 образов, по 20 образов для каждого из 16 классов. Количество нейронов в скрытом слое намеренно выбиралось минимальным (2–3) для получения возможно более сложной иерархической структуры с меньшим количеством выходных классов для каждого узла и с потенциально более высоким качеством распознавания.



Рис. 2. ИНК для одномерного распределения Кантора: $N = 2$, $\sigma=0.01$, $T = 5$ для базового узла, $T = 20$ для всех остальных узлов. Процент правильного распознавания на тренировочном наборе 98.75%, на тестовом наборе 100%

На нижеприведенных рисунках (рис. 2, рис. 3) представлены примеры ИНК для данной задачи, полученных для различных наборов параметров алгоритма (размер скрытого слоя N , скорость обучения σ и период анализа статистики ответов T).

Заметим, что все запуски дали сходные результаты: базовая сеть всегда группировала исходные классы по 4 (1–4, 5–8, 9–12 и 13–16); сети второго уровня иерархии имели по две или три выходных категории; сети третьего

уровня использовались для того, чтобы различить нераспознанные к этому моменту соседние классы.

Увеличение N (и в меньшей мере — увеличение T) приводит к увеличению количества классов, распознаваемых в каждом узле, и, как следствие, к уменьшению глубины дерева. Вся система оказывается достаточно чувствительной к начальной инициализации весов МСП, т. е. два последовательных запуска алгоритма с одними и теми же параметрами приводят, как правило, к построению различных по топологии деревьев.

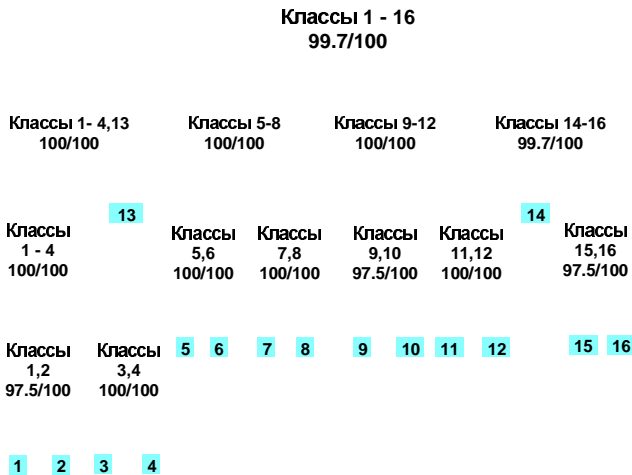


Рис. 3. Базовый и терминальный узлы $N = 2$, остальные $N = 3$, $\sigma=0.1$, $T = 5$. Процент распознавания на тренировочном наборе 98.75%, на тестовом наборе 100%

Увеличение σ дополнительно усиливает влияние начальной инициализации весов и приводит к построению менее сбалансированных деревьев. Впрочем, процент распознавания системы в целом весьма слабо зависит от конкретной топологии дерева (даже при его сильной асимметрии) и для данной задачи всегда оказывается весьма высоким (не менее 98.75%).

На рис. 4 показан пример эволюции принадлежности различных исходных классов с течением времени для базового узла ИНК.

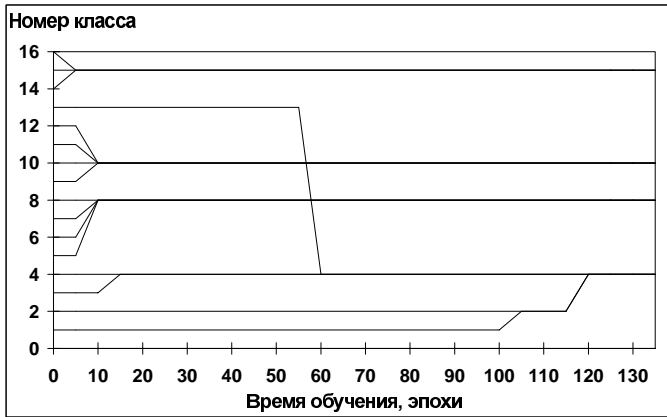


Рис. 4. Пример эволюции принадлежности различных исходных классов с течением времени для базового узла ИНК

Эта хорошо известная база данных, доступная по сети Интернет [9], позволила провести первое адекватное сравнение результатов работы алгоритма с результатами работы альтернативных алгоритмов. База данных `texture_CR.dat` содержит 11 классов, помеченных номерами 2, 3, 4, 6, 7, 8, 9, 10, 12, 13, 14. Тренировочный и тестовый наборы содержат по 250 образов в каждом классе. Данные нормировались к нулевому среднему и единичному стандартному отклонению. Каждый образ описывается 40 признаками.

Для этой задачи оказались справедливыми все наблюдения о влиянии параметров на получаемые деревья, сделанные выше для задачи с распределением Кантора. Отметим, что при любой разумной комбинации параметров процент распознавания получавшегося иерархического дерева оказывался не ниже 98.5%, что превышает наилучшие результаты, приводимые в [9] для того же способа предобработки данных (98%). Заметим, что для разных полученных деревьев существуют устойчивые группы классов, объединяемые алгоритмом на первой стадии. Такие группы, как правило, состоят из классов, имеющих, согласно [9], минимальную взаимную дисперсию.

Для дополнительной проверки возможностей алгоритма задача была модифицирована следующим образом. Каждый из 11 классов был случайным образом разбит на 4 подкласса равного размера. В результате работы алгоритма на втором уровне иерархического дерева 44 исходных класса были автоматически сгруппированы по 4 класса, образовав 11 реально различающихся групп. При этом уровень правильного распознавания этих 11 классов составил 99.4%. Стоит также заметить, что и в этом эксперименте на первом уровне дерева алгоритмом были объединены классы с минимальной взаимной дисперсией.

Задача дикторонезависимого распознавания 11 гласных звуков английского языка [10] отличается сложными для распознавания классами. Это хорошо видно из приведенной ниже таблицы, в которой показаны наилучшие результаты, достигнутые для этой задачи [10]. Образы описываются 10 признаками. Данные нормированы к нулевому среднему и единичному стандартному отклонению. Тренировочный набор содержит по 48 примеров для каждого класса, тестовый — по 42. На рис. 5 показан пример ИНК для данной задачи, полученный при $N = 1$, $\sigma=0.01$ и $T = 1$.

ТАБЛИЦА 1. Сравнительные результаты для задачи распознавания гласных

Алгоритм классификации	Нейронов в скрытом слое	% на тестовом наборе
Однослойный перцептрон	—	33
МСП	88	51
Модель Канервы	528	50
Радиальные базисные функции	528	53
Сеть с гауссовскими узлами	528	55
Сеть с квадратичными узлами	88	55
Метод ближайших соседей	—	56
ИНК	(1)	58

Классы 1-11
98.1/92.2

Классы 1-3
93.1/72.2

Классы 4-11
98.2/84.2

Классы 4-6,11
96.9/93.5

Классы 7-10
97.4/92.3

Классы 5,6,11
88.2/84.1

Классы 8-10
90.3/84.1

Классы 6,11
95.8/88.1

Классы 9,10
91.7/79.8

(сплошная линия) и для ИНК (отдельные точки) в зависимости от полного числа связей в полученной архитектуре. Анализ этой зависимости показывает, что ИНК не только существенно выигрывают по процентам распознавания, но также и весьма эффективны по вычислительной сложности. Наилучший ИНК, полученный для этой задачи, имеет такое же количество весов, как и МСП с 7 нейронами в скрытом слое.

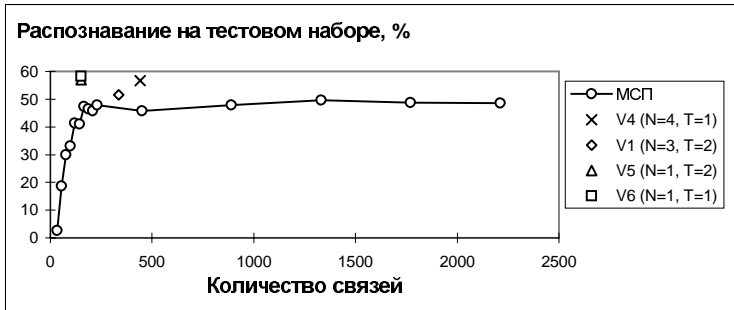


Рис. 6. Процент распознавания гласных на тестовом наборе для МСП (сплошная линия, MLP) и для некоторых ИНК (отдельные точки, V1, V4, V5, V6) в зависимости от полного числа связей в полученной архитектуре. Параметры алгоритма построения ИНК указаны в легенде. Скорость обучения во всех случаях $\sigma = 0.01$

Качество распознавания для тех же ИНК и МСП также сравнивалось на тестовом наборе, искаженном 10% и 20% белым шумом. На рис. 6 приведены результаты сравнения МСП и ИНК по устойчивости к шуму, накладываемому на тестовые данные. Даже в наихудшем случае 20% белого шума, ИНК имеют более высокий процент распознавания на тестовом наборе, что связано с меньшим количеством весов и большей устойчивостью решения.

Алгоритм построения ИНК был также успешно опробован при разработке нейросетевой системы дикторонезависимого распознавания изолированных слов.

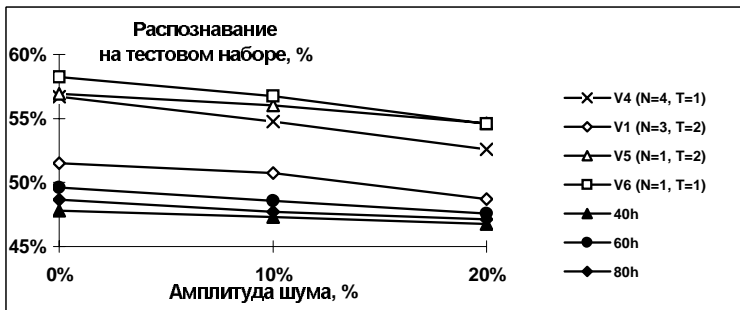


Рис. 7. Распознавание гласных. Сравнение МСП (40h, 60h, 80h обозначает число нейронов в скрытом слое) и ИНК (V1, V4, V5, V6) по устойчивости к шуму, накладываемому на тестовые данные. Зависимость процента распознавания на тестовом наборе от относительной амплитуды белого шума. Параметры алгоритма построения ИНК указаны в легенде. Скорость обучения во всех случаях $\sigma = 0.01$

С помощью этой системы была реализована возможность голосового управления работой стандартного калькулятора Windows, наглядно иллюстрируя эффективность ИНК.

Подготовка данных. Речевой сигнал оцифровывался с частотой 10 кГц 12-битовым аналого-цифровым преобразователем и разбивался на непрерывающиеся сегменты шириной 6.4 мс. После домножения сигнала в каждом сегменте на сглаживающее окно Хэмминга, производился расчет быстрого преобразования Фурье сигнальным процессором ADSP-1205 фирмы Analog Devices. Затем осуществлялся поиск начала/конца слова по алгоритму, основанному на алгоритме Рабинера-Шафера [11]. В результате определения начала и конца слова формировался массив векторов, образующий спектральный образ произнесенного слова. Частотный диапазон полученного массива преобразовывался в т. н. шкалу барков [12], для чего он разбивался на 20 интервалов и для каждого интервала вычислялась суммарная интенсивность попавших в него спектральных отсчетов. Границы интервалов вычислялись по формуле:

$$z = \frac{26.81f}{1960 + f} - 0.53, \quad (1)$$

где z — частота в барках, f — частота в герцах.

В результате для каждого слова получалась последовательность спектральных векторов (спектрограмма), длина которой зависела от длины произнесенного слова. Чтобы избавиться от временной зависимости, каждая такая последовательность приводилась к одинаковой длине путем линейной деформации по времени. Эта длина составляла 50 сегментов (640 мс) и соответствовала средней длине слова.

Амплитуды полученной спектрограммы переводились в шкалу децибел, и максимальное значение сигнала принималось за 0 дБ (при этом экспериментально определенный уровень шума соответствовал -17 дБ). После этого спектрограмма перенормировалась из диапазона $[-17$ дБ, 0 дБ] в диапазон $[0, 1]$. На вход системы распознавания подавались двумерные спектрограммы слов, состоящие из 1000 точек (50 точек по времени, 20 точек по частотной шкале), нормированные в диапазон $[0, 1]$.

База данных содержала 36 слов для управления калькулятором в русскоязычной версии оболочки Microsoft Windows, произнесенными 19 дикторами (14 мужских голосов, 5 женских голосов). Тренировочный набор включал 12 мужских и 3 женских голоса, тестовый набор — 2 мужских и 2 женских голоса.

Для искусственного расширения используемого при тренировке набора данных каждый предъявляемый образ мог искажаться одним из следующих способов. В одном случае к образцу добавлялся случайный белый шум, не превышающий заданной амплитуды A_{noise} , в другом случае образец подвергался нелинейным временным искажениям амплитуды $A_{distort}$ (чем имитировалось изменение темпа произнесения слова). Благодаря применению указанной методики эффективное количество образов, используемых при тренировке, было практически не ограничено и тем самым предотвращалась возможность эффекта «переучивания» нейросети.

Стратегия постепенного усложнения тренировки. Стратегия постепенного усложнения тренировки применительно к данной задаче заключалась в следующем. В начале обучения НС предъявлялись лишь неискаженные образы ($A_{noise} = 0$, $A_{distort} = 0$), и обучение продолжалось до достижения НС заранее заданного уровня правильного распознавания на тестовом наборе (90%). Затем амплитуда искажений A_{noise} и $A_{distort}$ устанавливалась равной некоторой малой величине, и, тем самым, тренировочный набор расширялся за счет предъявления слегка искаженных

образов, а задача классификации усложнялась. По достижении заданного уровня правильного распознавания амплитуды искажений еще увеличивались. Амплитуда искажений постепенно возрастала по мере обучения НС до тех пор, пока нейросеть была в состоянии поддерживать качество распознавания на заданном уровне.

Описание системы распознавания. Был построен ИНК, состоящий из двух уровней иерархии. Базовая НС имела 24 нейрона в скрытом слое. 36 исходных классов (36 слов) были объединены алгоритмом АФГК в 6 групп (1 группа из 4 классов, 3 группы по 6 классов и 2 группы по 7 классов). Для каждой из групп были построены вспомогательные НС, содержащие от 12 до 24 нейронов в скрытом слое. Для уменьшения вероятности ложного распознавания незнакомых слов в тренировочный набор каждой из НС второго уровня иерархии были включены не только представители классов, попавших в данную ветвь ИНК, но и все примеры из остальных классов. Для этих примеров из «посторонних» классов желаемым ответом НС была нулевая активность всех нейронов выходного слоя, что при распознавании трактовалось как ответ «не знаю». Общий процент правильного распознавания 36 слов для незнакомого диктора составил около 95%, что соответствовало типичным значениям, получаемым в нейросетевых системах распознавания [13].

Алгоритм построения ИНК может быть использован для анализа объектов, представимых в пространстве признаков в виде непрерывных или квазинепрерывных траекторий, в частности, для анализа временных рядов с переключающейся динамикой.

Пусть временной ряд y_t ($t = 1, 2, \dots$) генерируется источником $S(z_t)$, где z_t является переменным во времени параметром, принимающим значения из фиксированного множества параметров $\Theta = \{\theta_1, \theta_2, \dots, \theta_K\}$. В момент времени t значение y_t зависит от предыстории y_{t-1}, y_{t-2}, \dots и от текущего значения z_t . Значение z_t может скачком изменяться (переключаться) в произвольные моменты времени (т. е. значение изменяется за время, пренебрежимо малое по сравнению с интервалами между этими изменениями). Конкретные значения параметров из множества и их число K заранее неизвестны.

Задача анализа структуры временного ряда в рамках этой модели заключается в разбиении временного ряда на участки, внутри которых значение z_t не изменяется. На практике условие мгновенности изменения значения z_t (т. е. «переключение») выполняется далеко не всегда. Одним из наиболее очевидных предположений является наличие достаточно широкого участка «дрейфа». В этом случае анализ структуры временного ряда предполагает, помимо выделения участков с постоянным значением z_t , также определение участков дрейфа, на которых одновременно существует несколько значений z_t .

Можно предполагать, что в качестве объектов с описанными видами изменения типа динамики могут выступать временные ряды самой различной природы, от физических процессов (например, в задачах исследования динамики межпланетной плазмы, классификации типов электрических процессов в головном мозге, сегментации непрерывной речи) до финансовых зависимостей (например, при анализе состояния фондового рынка).

В работах [14–16] описывается нейросетевой подход к исследованию объектов с переключающейся динамикой, однако в этих работах производится подстройка и отбор нейросетей при заранее заданном их количестве, т. е. архитектура всего комплекса нейросетей в целом фиксирована и имеет параллельный, а не иерархический принцип организации. В нашем подходе используется принцип постепенного иерархического наращивания сложности структуры ИНК, что позволяет получать многоуровневые нейросетевые комплексы с очень сложными поверхностями разделения классов.

В работе [17] предлагается метод анализа временных рядов с дрейфовым характером изменения типа динамики, в котором для анализа участков дрейфа предлагается использование скрытых марковских моделей. В предложенном нами подходе при сегментации таких временных рядов используется чисто нейросетевой анализ участков дрейфа.

Описанные методики были применены для сегментации временных рядов с переключающейся и дрейфовой динамикой.

Алгоритм формирования групп классов (АФГК) для анализа и сегментации временных рядов с переключающейся динамикой работает следующим образом [18, 19]. Введем величину «минимального размера сегмента» (МРС), задавая ее значение, исходя из априорной оценки минимального времени

жизни одного типа динамики. Анализируемый временной ряд разбивается на одинаковые участки с размером, равным МРС, причем в пределах каждого из участков динамика объекта предполагается неизменной. В допущении, что переключение между различными типами динамики происходит достаточно редко, каждый из этих участков считается принадлежащим к отдельному классу со своей собственной динамикой. Таким образом, количество выходных нейронов в НС равно начальному количеству получившихся участков.

Методика тренировки МСП аналогична методике, использовавшейся для анализа статических объектов. Вначале обучение ведется стандартным методом обратного распространения ошибки. По прошествии нескольких эпох тренировки, алгоритм производит анализ статистики ответов МСП на тренировочном наборе. Так как данный алгоритм работает с временными рядами, то имеет смысл учитывать очевидное априорное соображение, что классы соседних образцов (в том числе и относящихся к разным участкам начального разбиения), скорее всего, совпадают. Таким образом, при рассмотрении голосования отдельного образца следует учитывать результаты голосования его соседей слева и справа, попадающих в т. н. «окно анализа». Размер окна анализа примем равным половине МРС. Будем считать, что рассматриваемый образец отнесен к классу j , если больше половины образцов в окне анализа проголосовало за принадлежность к классу j (в противном случае информация о принадлежности к классу не изменяется).

После окончания анализа голосования временной ряд разбивается на новые сегменты в соответствии с измененной принадлежностью к классам. Если количество образцов в каком-либо сегменте меньше, чем МРС, то такой сегмент присоединяется к сегменту, расположенному левее.

В результате, границы сегментов, изначально расположенные равномерно, могут быть передвинуты. После формирования нового разбиения на сегменты тренировка продолжается. Последовательность операций «тренировка–голосование–модификация сегментов» повторяется, пока разбиение временного ряда не перестанет изменяться. Итоговое количество классов, на которые был разбит временной ряд, определяется лишь по окончании тренировки МСП.

В результате работы АФГК участки временного ряда с одинаковой или близкой динамикой оказываются отнесенными к одному и тому же классу, а весь временной ряд разбивается на сегменты с указанием типа динамики.

Алгоритм построения ИНК при анализе временных рядов с переключающейся динамикой также аналогичен случаю анализа статических объ-

ектов. Сеть следующего уровня не строится, если после окончания тренировки все образцы объединены в один класс. В этом случае будем считать, что в рассматриваемом временном ряду выделение дополнительных типов динамики не представляется возможным. Сеть следующего уровня не строится и для классов, состоящих из сегментов, каждый из которых не превышает двойного размера MPC. Такие классы разделять на более мелкие сегменты нецелесообразно, так как их размер сравним со временем жизни динамики одного типа.

Для всех остальных классов строятся сети следующего уровня иерархии. Все образцы, принадлежащие одному такому классу, снова разбиваются на сегменты размера MPC. Количество выходных нейронов в сети следующего уровня равно количеству получившихся сегментов. Процедура обучения получившейся сети аналогична обучению базовой сети. Особенностью процедуры является то, что теперь мы рассматриваем временной ряд не как непрерывный, а как состоящий из выделенных на предыдущем уровне иерархии участков временного ряда, которые могут не являться смежными. Поэтому алгоритм сегментации применяется к каждому такому участку временного ряда по отдельности.

Одним из главных предположений, которое было положено в основу описанного выше алгоритма анализа временных рядов с переключающейся динамикой, являлось предположение о мгновенном переключении между различными типами динамики. Однако на практике данное условие выполняется далеко не всегда. Одним из наиболее очевидных предположений является наличие достаточно широкого участка «дрейфа». В рамках этой модели, на участке дрейфа предполагается одновременное существование нескольких типов динамики, причем типы динамики плавно трансформируются один в другой (дрейфуют) в течение некоторого достаточно значительного интервала времени. Например, в модельных задачах рассматривался дрейф вида:

$$f(t) = a(t) * f_L(t) + (1 - a(t)) * f_R(t), \quad (2)$$

где f_L и f_R — функции, описывающие динамику в сегментах, расположенных слева и справа от переходного участка, соответственно, а $a(t)$ — функция, линейно нарастающая от 0 до 1 на участке дрейфа.

Очевидно, что заложенные в существовавший алгоритм ограничения позволяли относить такие участки только к какому-либо одному классу, что приводило к тому, что при сегментации появлялись хаотические перескоки от класса к классу на коротком временном интервале

Обобщение алгоритма АФГК на случай дрейфового характера изменения типа динамики заключается в следующем.

Будем рассматривать ответ нейронной сети (множество активностей нейронов выходного слоя $O_j(S_i)$, $j = 1, \dots, K$) на предъявление примера S_i ($i = 1, \dots, N$) как оценку вероятности принадлежности данного примера к каждому из K классов. Учтем предположение о плавном характере изменения типа динамики во временном ряде с характерным масштабом, равным минимальному времени жизни одного типа динамики. С другой стороны, ответы НС на предъявление разных примеров являются статистически независимыми, поскольку используемая архитектура (МСП) не имеет внутренней памяти. Исходя из этого, можно оценить функцию распределения вероятностей принадлежности примера S к каждому из K классов путем усреднения ответов НС по окну анализа шириной W :

$$P_j = \sum O_j(S_i)/W, \quad i = 1, \dots, W, \quad j = 1, \dots, K,$$

где суммирование ведется по всем примерам S_i , попадающим в окно анализа.

Будем считать незначимыми классы

фрагментах будут доминировать различные типы динамики. Заметим, что анализ процесса изменения амплитуд ответов нейросети на участке дрейфа позволяет получать количественную информацию о характере дрейфа одного типа динамики в другой.

-

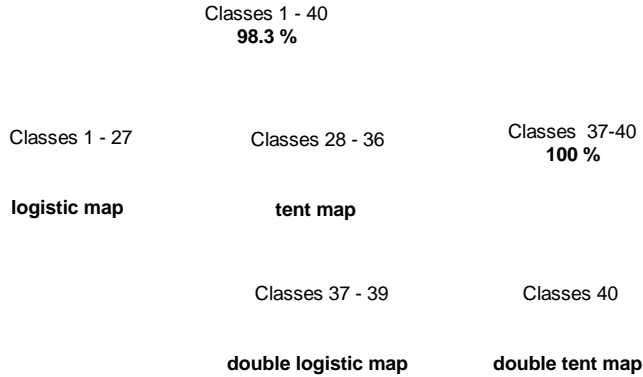
В модельных экспериментах использовалась задача определения границ псевдохаотических последовательностей $x(t+1) = f_i(x(t))$, $i = 1, 2, 3, 4$, где:

1. $f_1(x) = 4x(1-x)$, $x \in [0, 1]$; logistic map
2. $f_2(x) = 2x$, $x \in [0, 0.5]$, $2(1-x)$, $x \in [0.5, 1]$; tent map
3. $f_3(x) = f_1(f_1(x))$; double logistic map
4. $f_4(x) = f_2(f_2(x))$; double tent map

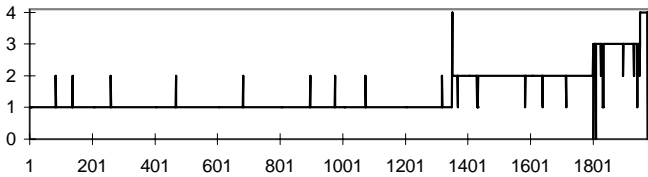
Эти последовательности чередовались, генерируя (в разных экспериментах) по 25–125 точек каждая, а общая длина тренировочного набора данных составляла 1000–2400 точек. Затем временной ряд линейно нормировался к нулевому среднему и единичному стандартному отклонению. Пример для тренировки нейросети формировался из пяти последовательных значений полученного временного ряда. Принадлежность примера к классу определялась по тому, к какому классу принадлежит последняя (самая правая) точка отрезка.

На начальной стадии исследовались временные ряды с переключающейся динамикой. Границы между сегментами в процессе работы алгоритма считались фиксированными.

В первом эксперименте каждая из псевдохаотических последовательностей поочередно генерировала по 50 точек, и этот процесс был повторен 5 раз. Таким образом, общая длина временного ряда составила 1000 точек. Этот ряд был поделен на 40 сегментов по 25 точек, поэтому первоначально в задаче было 40 классов. Общий процент распознавания этого ИНК составил 96.1%. Заметим, что классификация того же временного ряда в 4 класса, соответствующие 4 последовательностям, с помощью обыкновенного МСП дает лишь около 95%.



(a)



(b)

Рис. 8. Вверху – ИНК, построенный в эксперименте с сильным дисбалансом представительности классов (пропорция $2^7 : 9 : 3 : 1$). Общий процент распознавания ИНК 98.3%. Внизу — сегментация временного ряда, полученная с помощью построенного ИНК (номер псевдохаотической последовательности, к которой была отнесена каждая точка временного ряда, в зависимости от номера точки; нуль соответствует нераспознанным образам)

Был также проделан набор экспериментов с несбалансированными классами, когда некоторые псевдохаотические последовательности давали значительно больше точек, чем другие. Наиболее сильный дисбаланс использовался в эксперименте, результаты которого представлены на рис. 8.

Было взято 1350 точек logistic map, 450 точек tent map, 150 точек double logistic map и только 50 точек double tent map. Таким образом, пропорция

между последовательностями была $27 : 9 : 3 : 1$. Размер тренировочного набора составил 2000 примеров. Этот набор был разделен на 40 сегментов по 50 точек в каждом. Общий процент распознавания ИНК, представленного на рис. 8, составил 98.3%. Проценты распознавания для каждой из последовательностей были следующими: logistic — 99.33%, tent — 98.2%, double logistic — 91.4%, double tent — 92%.

В следующей серии экспериментов в процессе работы алгоритму разрешилось подстраивать положение границ сегментов. Опишем один из таких экспериментов. Все четыре последовательности поочередно генерировали по 100 точек. Процесс был повторен 5 раз, в результате общая длина временного ряда составила 2000 точек. Весь набор был разбит на 47 сегмента по 43 примера в каждом (кроме последнего сегмента, принадлежащего f_4 , в котором было 22 примера), причем 13 из 47 сегментов содержали примеры, принадлежащие различным последовательностям. Таким образом, во многих сегментах присутствовало изменение типа динамики, и ожидалось, что алгоритм передвинет границы сегментов соответственно. Значение MPC было выбрано равным 10, скорость обучения 0.01.

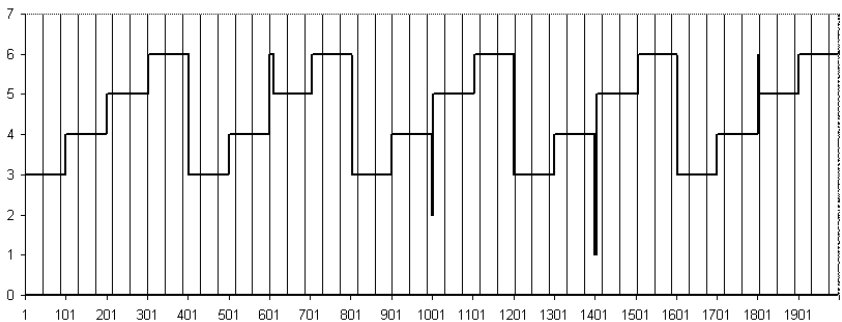


Рис. 9. Результат сегментации модельного временного ряда после подстройки положения границ сегментов. Вертикальные линии — границы сегментов при начальном разбиении

Общий процент распознавания ИНК составил 98% при точно определенных границах каждого типа динамики. Границы сегментов были передвинуты алгоритмом таким образом, что к каждому сегменту принадлежат примеры практически только одного типа динамики (рис. 9).

В последней серии экспериментов исследовались временные ряды с дрейфовым характером изменения типа динамики.

Последовательности f_1-f_4 поочередно генерировали по 100 точек, с переходными участками шириной 50 точек. Процесс был повторен 4 раза, и в результате общая длина временного ряда составила 2400 точек. Весь набор был разбит на 48 сегментов по 50 примеров в каждом.

Для анализа полученного временного ряда использовалась нейросеть с 5 входными нейронами, 20 нейронами в скрытом слое и 48 нейронами в выходном слое. Сеть тренировалась со скоростью 0.01. С шагом в 200 эпох для каждого примера из временного ряда вычислялась оценка вероятности его принадлежности к каждому из классов, путем усреднения ответов НС в окне анализа шириной 50 точек.

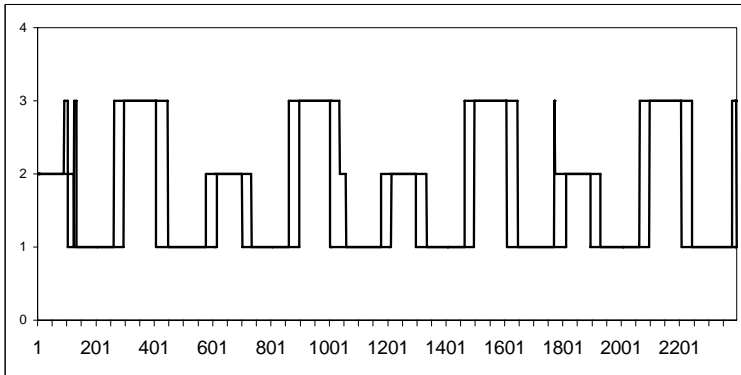


Рис. 10. Результат сегментации модельного временного ряда, в котором переключение типа динамики носит дрейфовый характер. Видно, что на участках дрейфа одновременно существуют два типа динамики

На рис. 10 приведен результат сегментации описанной псевдохаотической последовательности с помощью разработанного алгоритма. На графике показаны существующие в каждый момент времени типы динамики, определенные в результате работы алгоритма. Видно, что на участках дрейфа одновременно существуют два типа динамики.

Описанный алгоритм построения ИНК был применен для анализа записей физиологического изменения параметров дыхания (респираторные данные), снятых у здоровых людей во время дневного сна. Задача заключалась в разбиении временного ряда (10–15 тысяч точек, рис. 11, верхний график) на сегменты с различными типами динамики, соответствующими определенным стадиям сна. Сегментация, полученная с помощью ИНК (рис. 11, нижний график), сравнивалась с ручной разметкой временного ряда медицинским экспертом (рис. 11, средний график). Экспертная разметка включала 5 классов (W1/W2 — две фазы бодрствования, S1/S2 — две фазы сна, и артефакт) [20]. Данные для этого эксперимента доступны в Интернете [21].

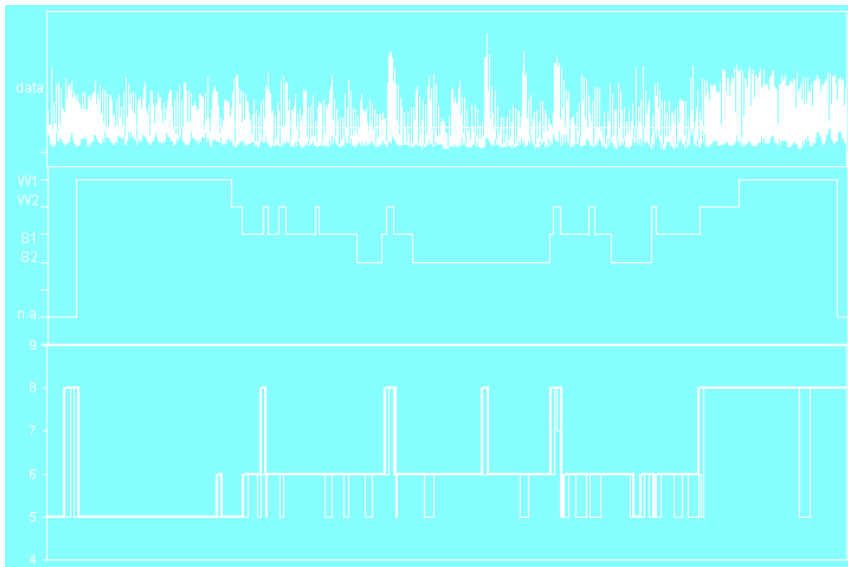


Рис. 11. Анализ биомедицинских данных. Сверху вниз: анализируемый временной ряд (респираторные данные); ручное разбиение экспертом-медиком; результат сегментации данных разработанным алгоритмом

Данные нормировались к нулевому среднему и единичному стандартному отклонению. На вход нейросети подавались 16 точек временного

ряда, взятых с шагом в 6 точек. Временной ряд разбивался на сегменты по 360 точек в каждом. Значение MPC было выбрано равным 180, что соответствовало характерному времени существования одного типа динамики (фазы сна) по заключениям экспертов.

В табл. 2 приведены средние проценты соответствия разбиения, полученного с помощью ИНК, с экспертным разбиением. Соответствие ручной разметки с машинной сегментацией измеряется на различных уровнях точности. Грубое разбиение включает в себя два класса: $W = W1 \cup W2$ и $S = S1 \cup S2$. В более тонком разбиении используются три класса — $W1$, $W2$ и $S = S1 \cup S2$. В качестве «тестового набора» использовались две другие записи респираторных данных, полученные для того же пациента. Для сравнения приведены также результаты из статьи.

ТАБЛИЦА 2. Средние проценты соответствия разбиения, полученного с помощью ИНК, с экспертным разбиением

	Результат ИНК	Результат ИНК	Результаты [6]	Результаты [6]
Набор данных	W1-W2-S	W-S	W1-W2-S	W-S
Тренировочный	81.30%	86.93%	79.08%	86.33%
Тестовый	59.10%	71.35%	63.92 %	75.24%

Сильное различие между процентами, полученными для тренировочного и тестового наборов при сегментации W1-W2-S, могут свидетельствовать о некоторой «переученности» ИНК. Этот эффект связан, по-видимому, как с недостаточной представительностью данных, так и с тем, что одним из основных положений алгоритма является отсутствие ограничений сверху на выявляемое количество различных типов динамик.

Полученные проценты, тем не менее, демонстрируют перспективность алгоритма. Важно заметить, что эксперт давал оценку не только по респираторным данным, но и на основе других физиологических сигналов, регистрировавшихся одновременно с респираторными данными (ЭЭГ, ЭКГ, электроокулограмма, артериальное давление, частота сердечных сокращений), в то время как ИНК строился лишь на основе респираторных данных.

В качестве объекта анализа использовался временной ряд, состоящий из среднечасовых значений скорости солнечного ветра за 1974 г. (8760 точек). Поскольку в данных имелись значительные (около 20%) пропуски, для их заполнения использовалась следующая процедура: пропуски сначала заполнялись линейно с добавлением небольшого шума и вычислялся десятичный логарифм полученной зависимости. Затем рассчитывались коэффициенты вейвлет-преобразования Добеши 4-го порядка по 512 точкам (т.е. 512 часов, что составляет около 3 недель), и проводилось обратное преобразование с отбрасыванием коэффициентов с амплитудой ниже 30% от максимальной. Отфильтрованная кривая сглаживалась путем усреднения по 6 часам. Анализируемый временной ряд представлен на верхнем графике рис. 12. Расстояние между вертикальными линиями соответствует одному обороту Солнца.

Для компактного представления динамики полученного временного ряда в качестве входов для нейронной сети использовались 4 первые коэффициента вейвлет-преобразования Добеши 4-го порядка по 128 точкам (128 часов, что составляет около 5 суток) и 13 задержанных значений каждого из этих коэффициентов (с шагом в 48 часов). Максимальная задержка составляла 624 часа, а размер временного окна — около 27 суток.

Для определения участков различных типов динамики временной ряд разбивался на 27 сегментов по 324 точки в каждом (рис. 12, верхний график). Окно анализа составляло 162 точки. Нейронная сеть имела 56 входных, 27 выходных нейронов и 5 нейронов в скрытом слое. Скорость обучения сети составляла 0.5.

В результате работы алгоритма была получена сегментация анализируемого временного ряда, представленная на рис. 12.

На среднем графике показаны амплитуды ответов наиболее активных нейронов выходного слоя нейросети в зависимости от номера предъявленного примера. На нижнем графике показано разбиение временного ряда, полученное в результате работы алгоритма (нумерация классов произвольная). Из рисунка видно, что был выделен один основной тип динамики (класс № 3), один тип динамики с периодом проявления 4-5 оборотов Солнца (класс № 2), и два короткоживущих типа динамики с временем существования около половины оборота Солнца (классы № 1 и № 5). Длительности переходных периодов между различными типами динамики составили от 1/5 до половины оборота Солнца.

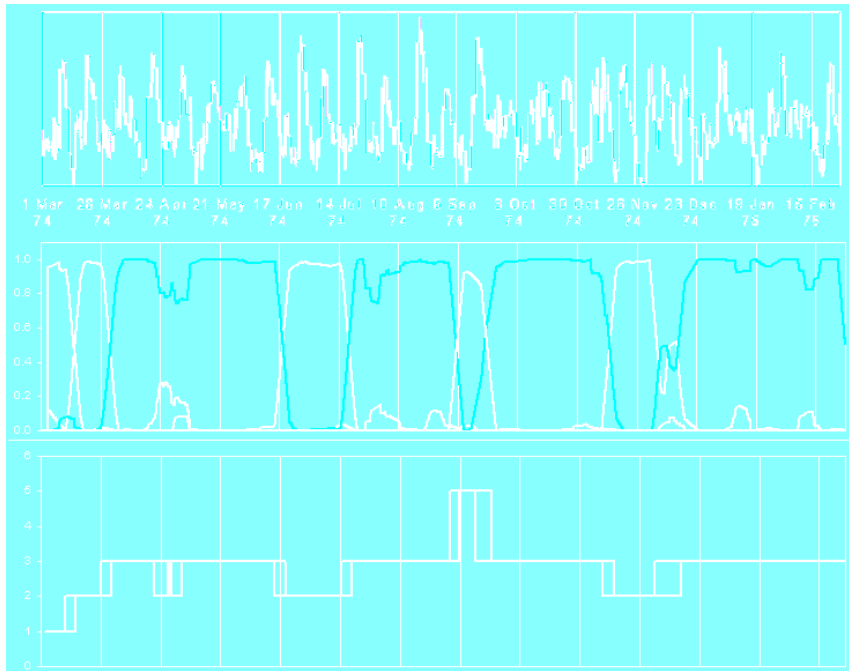


Рис. 12. Анализ космофизических данных. Сверху вниз: анализируемый временной ряд (скорость солнечного ветра в логарифмическом масштабе); усредненные амплитуды ответов нейронной сети; полученная сегментация временного ряда. Вертикальные линии — начальные границы нечетных сегментов

Проведенные исследования показали, что предложенный метод решения задач классификации путем построения ИНК обеспечивает высокую эффективность как при решении задачи с хорошо разделяемыми классами, так и при решении сложной задачи с сильно перекрывающимися классами. Перечислим разработанный подход:

- Каждый узел ИНК представляет собой обычный МСП, традиционно обучаемый с учителем. Предложенная модификация алгоритма тре-

нировки фактически переводит обучение МСП в режим обучения без учителя.

- Образование групп классов в узлах ИНК происходит одновременно с тренировкой соответствующего классификатора, обеспечивая оптимальный выбор признаков.
- Алгоритм формирования групп классов наследует все хорошо известные преимущества МСП, в частности, возможность тренировки на примерах, что позволяет учесть естественным образом априорную информацию о сходстве группируемых образов.
- Структура ИНК не задается заранее. Рекурсивное применение алгоритма обеспечивает адаптивное наращивание иерархических структур.
- Параметры алгоритма позволяют управлять топологией получаемых ИНК.
- Применение этого подхода позволяет получить структуры с высоким процентом распознавания, достаточно устойчивые к наличию шума в тестовых данных.
- Модификация алгоритма позволяет его использовать для сегментации временных рядов как с переключающейся динамикой, так и с дрейфовым характером изменения динамики.

1. *Dolenko S. A., Orlov Yu. V., Persiantsev I. G., Shugai J. S., Eremin E. K.* Adaptive method of construction of a perceptron-based hierarchical structure of classifiers // The 4th Open Russian-German Workshop on Pattern Recognition and Image Analysis, Valdai, Russia, March 4-8, 1996, pp. 45-49.
2. *Dolenko S. A., Orlov Yu. V., Persiantsev I. G., Shugai J. S., Eremin E. K.* The perceptron-based hierarchical structure of classifiers constructed by the adaptive method // *Pattern Recognition and Image Analysis*, 1997, Vol.7, No.1, pp.24-28.
3. *Dolenko S. A., Orlov Yu. V., Persiantsev I. G., Shugai J. S., Eremin E. K.* Self-organising construction of hierarchical structure of multi-layer perceptrons // Proc. 5th Int. Conf. on Artificial Neural Networks, Churchill College, University of Cambridge, UK, July 7-9, 1997, p. 285-290.
4. *Доленко С. А., Орлов Ю. В., Персианцев И. Г., Шугай Ю. С.* Использование принципов самоорганизации для построения иерархических нейросетевых классификаторов // 3-я конференция «Распознавание образов и анализ изображений:

- Новые информационные технологии», 1–7 декабря 1997 г., Н. Новгород. Тезисы докладов в 2 ч., ч. 1. с. 31–35.
5. *Dolenko S. A., Orlov Yu. V., Persiantsev I. G., Shugai J. S.* Use of self-organization principles in construction of hierarchical neural network classifiers // *Pattern Recognition and Image Analysis*, 1998, Vol. 8, No. 2, pp. 122–124.
 6. *Dolenko S. A., Orlov Yu. V., Persiantsev I. G., Shugai J. S.* Construction of hierarchical neural classifiers by self-organizing error back-propagation // Proc. Int.ICSC/IFAC Symposium on Neural Computation, Sept 23–25 1998, Vienna, Austria. Ed. *M.Heiss*, ICSC Academic Press International Computer Science Conventions Canada/Switzerland, 1998, pp. 113–116.
 7. *Rumelhart D. E., Hinton G. E., Williams R. J.* Learning internal representations by error propagation // In: *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, Vol. 1: Foundations, MIT Press, 1986, pp. 318–362.
 8. *Hornik K., Stinchcombe M., White H.* Multilayer feedforward networks are universal approximators // *Neural Networks*, 1989, Vol. 2, No. 5, pp. 359–366.
 9. ELENA – Enhanced Learning for Evolutive Neural Architecture, ESPRIT Basic Research Project Number 6891, June 1995,
URL: <ftp://ftp.dice.ucl.ac.be/pub/neural-nets/ELENA/databases>
 10. Carnegie-Mellon University Repository of Neural Network Benchmarks.
URL: <http://www.boltz.cs.cmu.edu>
 11. *Рабинер Л., Шафнер Р.* Цифровая обработка речевых сигналов. – М: Радио и связь, 1981.
 12. *Traunmuller H.* Analytical expressions for the tonotopic sensory scale // *J. Acoust. Soc. Am.* 1990, Vol. 88, pp. 97–100.
 13. *Lee G. E., Tattersall G. D., Smyth S. G.* Isolated word speech recognition using a neural network based source model // *BT Technol. J.*, 1992, Vol. 10, No. 3, pp. 38–47.
 14. *Pawelzik K., Kohlmorgen J., Muller K.-R.* Annealed competition of experts for a segmentation and classification of switching dynamics // *Neural Computation*, 1996, Vol. 8, No. 2, pp. 340–356.
 15. *Kehagias A., Petridis V.* Time-series segmentation using predictive modular neural networks // *Neural Computation*, 1997, Vol. 9, No. 8, pp. 1691–1709.
 16. *Watanabe E., Nakasako N., Mitani Y.* A prediction method of non-stationary time series data by using a modular structured neural network // *IEICE Trans. Fundamentals*, 1997, Vol. E80-A, No. 6, pp. 971–976.
 17. *Kohlmorgen J., Muller K.-R., Rittweger J., Pawelzik K.* Identification of nonstationary dynamics in physiological recordings // *Biological Cybernetics*, 2000, Vol. 83, pp. 73–84.

18. *Dolenko S. A., Orlov Yu. V., Persiantsev I. G., Shugai Yu. S.* Time series analysis using hierarchical neural network classifiers // Proc. 5th International Conference on Computational Intelligence And Neuroscience, Atlantic City, NJ USA, Ed. *P. P. Wang*, Association for Intelligent Machinery, Inc., 2000, Vol. 1, pp. 908–911.
19. *Dolenko S. A., Orlov Yu. V., Persiantsev I. G., Shugai Yu. S.* Solar wind data analysis using self-organizing hierarchical neural network classifiers // *Lecture Notes in Computer Science*, Eds. *J. Kittler* and *F. Roli*. Springer-Verlag Berlin Heidelberg, 2001, Vol. 2096, pp. 289–298.
20. *Hornik K., Stinchcombe M., White H.* Multilayer feedforward networks are universal approximation // *Neural Networks*, 1990, Vol. 2, No. 5, pp. 359–366.
21. URL: <http://www.first.gmd.de/persons/kohlmorgen.jens/physiodata.tar.gz>

, ведущий научный сотрудник Научно-исследовательского института ядерной физики имени Д. В. Скобельцына МГУ имени М. В. Ломоносова, доктор физико-математических наук, профессор. Область научных интересов: физика плазмы, спектроскопия, газовые лазеры, адаптивные методы обработки данных физического эксперимента, искусственные нейронные сети, генетические алгоритмы. Имеет более 120 научных публикаций.