

А. И. Задаянчук¹, студент, e-mail: alex.goncharov@phystech.edu,
 М. С. Попова¹, студент, В. В. Стрижов², д-р физ.-мат. наук, вед. науч. сотр.

¹ Московский физико-технический институт

² Вычислительный центр им. А. А. Дородницына

Выбор оптимальной модели классификации временных рядов

Решается задача построения модели для точной и устойчивой классификации физической активности человека по временным рядам. Нейронные сети с недостаточным или избыточным числом нейронов обладают недостаточной обобщающей способностью и могут давать прогнозы, неустойчивые относительно изменения состава выборки. Исследуются модели из класса двухслойных нейронных сетей. Рассматриваются модели с избыточно сложной структурой. Структура модели оптимизируется путем удаления из нее наборов параметров — нейронов. Для оптимизации структуры нейронной сети и обеспечения устойчивости предлагается алгоритм генетического типа. Новизна работы заключается в том, что вероятность удаления наборов параметров определяется дисперсией параметров. В вычислительном эксперименте модели, порождаемые предложенной стратегией, сравниваются по двум критериям качества: точности и устойчивости. Модели оптимизируются на выборках, полученных путем выделения признаков из временных рядов.

Ключевые слова: классификация, нейронные сети, устойчивость, критерии прореживания, генетический алгоритм

Введение

Работа посвящается исследованию методов построения нейронной сети оптимальной структуры для решения задачи классификации временных рядов. В работе исследуются и сравниваются методы изменения размерности пространства параметров двухслойных нейронных сетей. При уменьшении размерности пространства параметров значительно уменьшается время оптимизации параметров, увеличивается обобщающая способность нейронной сети и, как следствие, уменьшается значение функции ошибки на контрольной выборке [1].

Оптимизировать размерность пространства можно на разных уровнях — на уровне нейронов (наборов параметров) [2] и отдельных параметров [3]. Структурные параметры модели — это параметры, управляющие включением нейрона в модель. В данной работе размерность пространства параметров оптимизируется на уровне нейронов путем изменения значений структурных параметров. Предложено несколько способов такой оптимизации: прореживание (*network pruning*) [4], наращивание (*network growing*) [5] и пошаговое чередование наращивания и прореживания [6, 7]. В настоящей работе рассматривается оптимизация размерности с помощью прореживания. Базовыми алгоритмами прореживания нейронных сетей являются оптимальные прореживания (англ. *optimal brain damage* [8] и *optimal brain surgery* [9]), основанные на вычислениях вторых производных функции ошибки.

Устойчивая и оптимальная модель описывается с помощью генетического алгоритма на уровне нейронов [10, 11] путем оптимизации структурных параметров. Базовые алгоритмы [8, 9] находят локальный минимум функции ошибки. В случае же, когда функция ошибки имеет значительное

число локальных минимумов, найденный минимум может не совпадать с глобальным. Для нахождения глобального минимума используется алгоритм отбора моделей путем случайного подбора, комбинирования и вариации структурных параметров [10] набора нейронных сетей. Вероятность комбинирования и вариации структурных параметров нейронной сети тем меньше, чем больше показатель выпуклости, используемый в *optimal brain damage* [8].

В вычислительном эксперименте рассматривается задача классификации физической активности человека по измерениям акселерометра. Эта задача решалась в исследованиях [12, 13] с помощью нейронных сетей. В вычислительном эксперименте оцениваются значения критериев качества для нейронных сетей, порождаемых предложенной стратегией. Временные ряды предварительно обрабатываются двумя способами — экспертным порождением признаков [12] и вводом метрики выравнивания временных рядов, с последующим выделением признаков, полученных как расстояние до центров классов [14].

1. Постановка задачи классификации

Дана выборка $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{t}_i)\}$, $i \in \mathcal{J} = \{1, \dots, m\}$, состоящая из m объектов, каждый из которых описывается n признаками $\mathbf{x}_i \in \mathbb{R}^n$ и принадлежит одному из z классов $\mathbf{t}_i \in \{0, 1\}^z$. Задано разбиение множества индексов выборки $\mathcal{J} = \mathcal{L} \sqcup \mathcal{C}$ на обучающую $(\mathbf{x}_l, \mathbf{t}_l)$, где $l \in \mathcal{L}$, и контрольную $(\mathbf{x}_c, \mathbf{t}_c)$, где $c \in \mathcal{C}$. Необходимо выбрать наиболее точную и при этом устойчивую модель классификации.

Определение 1. Моделью назовем отображение:

$$\mathbf{f}(\mathbf{w}, \mathbf{X}) \mapsto \mathbf{y}, \quad (1)$$

$$\mathbf{w} = [w_1, \dots, w_j, \dots, w_J]^T, j \in \mathcal{J} = \{1, \dots, J\},$$

где \mathbf{w} — вектор параметров модели, $\mathbf{X} \in \mathbb{R}^{n \times m}$ — матрица объект-признак, $\mathbf{y} \in \{0, 1\}^z$ — зависимая переменная.

Предполагается, что переменная \mathbf{y} — мультиномиально распределенная случайная величина, а переменная \mathbf{w} имеет нормальное распределение:

$$\mathbf{w} \sim \mathcal{N}(0, \mathbf{A}^{-1}), \quad (2)$$

\mathbf{A}^{-1} — ковариационная матрица. В данной работе рассматриваются модели \mathbf{f} , принадлежащие классу двуслойных нейронных сетей с функциями активации **tanh** и **softmax**

$$\mathbf{a}(\mathbf{x}) = \mathbf{W}_2^T \tanh(\mathbf{W}_1^T \mathbf{x}), \quad (3)$$

$$\mathbf{p}(\mathbf{x}) = \frac{\exp(\mathbf{a}(\mathbf{x}))}{\sum_I \exp(a_I(\mathbf{x}))}. \quad (4)$$

Вектор \mathbf{p} интерпретируется как вектор вероятностей: p_ξ есть вероятность того, что вектор \mathbf{x} принадлежит классу с номером ξ :

$$\mathbf{p}(\mathbf{x}) = \{p_\xi\}, 0 \leq p_\xi \leq 1, \sum p_\xi = 1, \xi = 1, \dots, z.$$

Под вектором параметров двуслойной нейронной сети будем понимать $\mathbf{w} = \text{vec}(\mathbf{W}_1^T | \mathbf{W}_2^T)$, где $\mathbf{W}_1, \mathbf{W}_2$ — присоединенные матрицы весов первого и второго слоев нейронной сети. Вектор $\mathbf{y} = [y_1, \dots, y_\xi, \dots, y_z]^T$ определим следующим образом:

$$y_\xi = \begin{cases} 1, & \text{если } \xi = \arg \max_{\xi \in \{1, \dots, z\}} (p_\xi), \\ 0 & \text{иначе.} \end{cases} \quad (5)$$

Моделью \mathbf{f} является суперпозиция функций (3)–(5). В данной работе мы будем исследовать модели, отличающиеся друг от друга на структурном уровне, а именно — содержащие разное число групп связей — нейронов.

Определение 2. *Нейроном u_k назовем k -ю компоненту вектор-функции $\tanh(\mathbf{W}_1^T \mathbf{x})$ — сомножитель (3).*

Каждый нейрон задается весами в соответствующей строке матрицы \mathbf{W}_1^T .

Определение 3. *Нейрон назовем неактивным, если $u_k = 0$.*

Нейрон неактивен, если k -я строка матрицы \mathbf{W}_1^T нулевая.

Определение 4. *Нейронной структурой $\mathcal{A} = \{k, u_k \neq 0\}$ модели \mathbf{f} назовем множество активных нейронов.*

Каждая нейронная структура \mathcal{A} однозначно задает некоторую модель (1):

$$\mathbf{f}_{\mathcal{A}} \cdot \hat{\mathbf{w}}_{\mathcal{A}} \in \mathbb{R}^k,$$

где $\mathbf{f}_{\mathcal{A}}$ — модель со структурой \mathcal{A} , а $\hat{\mathbf{w}}_{\mathcal{A}} \in \mathbb{R}^k$ — оптимальный вектор параметров модели $\mathbf{f}_{\mathcal{A}}$, определение которому будет дано ниже. Объединение всех $\mathbf{f}_{\mathcal{A}}$ назовем множеством допустимых моделей:

$$\mathfrak{F} = \bigcup_{\mathcal{A} \subseteq J} \{\mathbf{f}_{\mathcal{A}}\}. \quad (6)$$

Оптимальную модель $\hat{\mathbf{f}}_{\mathcal{A}}$ будем выбирать из множества допустимых моделей $\mathbf{f}_{\mathcal{A}}$.

В качестве функции ошибки выберем функцию:

$$S(\mathbf{w}|\mathcal{L}) = - \sum_{i \in \mathcal{L}} \sum_{\xi=1}^z t_{i\xi} \ln(p_\xi(\mathbf{x}_i, \mathbf{w})). \quad (7)$$

Определение 5. *Устойчивостью $\eta = \eta(\hat{\mathbf{w}})$ модели \mathbf{f} с вектором параметров \mathbf{w} назовем число η , равное числу*

обусловленности матрицы \mathbf{A} , т. е. $\eta(\hat{\mathbf{w}}) = \frac{\lambda_{\max}}{\lambda_{\min}}$, где

λ_{\max} — максимальное, а λ_{\min} — минимальное собственные числа матрицы \mathbf{A} .

Чем лучше обусловлена матрица \mathbf{A} , тем более устойчива модель.

Матрица ковариации вычисляется с учетом предположения (2) о нулевом математическом ожидании вектора параметров \mathbf{w} :

$$\begin{aligned} \mathbf{A}^{-1} &= \text{cov}(\mathbf{W}) = \\ &= \mathbb{E}(\mathbf{W}^T \mathbf{W}) - \mathbb{E}(\mathbf{W})\mathbb{E}(\mathbf{W}^T) = \mathbb{E}(\mathbf{W}^T \mathbf{W}), \end{aligned}$$

где \mathbf{W} — это матрица реализаций оптимального вектора параметров $\hat{\mathbf{w}}$.

Определение 6. *Под точностью S модели \mathbf{f} с вектором параметров $\hat{\mathbf{w}}$ будем понимать значение функции ошибки (7) на контрольной выборке.*

Чем больше значение функции ошибки, тем меньше точность модели.

Определение 7. *Оптимальным вектором параметров модели $\mathbf{f}_{\mathcal{A}}$ назовем такой вектор $\hat{\mathbf{w}}_{\mathcal{A}}$, который является решением следующей задачи оптимизации:*

$$\hat{\mathbf{w}}_{\mathcal{A}} = \arg \min_{\mathbf{w}_{\mathcal{A}} \in \mathbb{R}^k} S(\mathbf{w}_{\mathcal{A}}|\mathcal{L}, \mathbf{f}_{\mathcal{A}}, \hat{\mathbf{A}}). \quad (8)$$

Задача выбора оптимальной модели состоит в том, чтобы найти модель $\mathbf{f} \in \mathfrak{F}$, для которой функция ошибки будет минимальной.

$$\hat{\mathbf{f}}_{\mathcal{A}} = \arg \min_{\mathbf{f}_{\mathcal{A}} \in \mathfrak{F}} S(\mathbf{f}_{\mathcal{A}}|C). \quad (9)$$

Устойчивость модели будет дополнительным критерием качества.

2. Описание алгоритма

Для получения оптимальной структуры модели в работе предлагается генетический алгоритм оптимизации структуры нейронной сети. Для сравнения также реализован базовый алгоритм NODE-OBDD.

2.1. Алгоритм прореживания структуры нейронной сети NODE-OBDD

Предлагаемый алгоритм определяет индекс нейрона, удаление которого приведет к минимизации приращения функции ошибки (7). Удаление нейрона эквивалентно обнулению соответствующего столбца матрицы \mathbf{W}_2 , т. е. удалению сразу группы параметров вектора \mathbf{w} . В этом разделе, для краткости изложения, будем обозначать матрицу \mathbf{W}_2 как \mathbf{W} . Предполагаем, что удаляемый нейрон наименьшим образом влияет на функцию ошибки. Для нахождения таких нейронов аппроксимируем функцию ошибки вблизи локального минимума матрицы \mathbf{W}_0 :

$$S(\mathbf{W}_0 + \Delta\mathbf{W}) = S(\mathbf{W}_0) + \mathbf{g}^T(\mathbf{W}_0)\Delta\mathbf{W} + \frac{1}{2}\Delta\mathbf{W}^T\mathbf{H}\Delta\mathbf{W} + O(\|\Delta\mathbf{W}\|^3),$$

где $\Delta\mathbf{W}$ — возмущение матрицы параметров в данной точке \mathbf{W}_0 ; $\mathbf{g}(\mathbf{W}_0)$ — вектор градиента, вычисленный в точке \mathbf{W}_0 ; $\mathbf{H} = \mathbf{H}(\mathbf{W}_0)$ — матрица вторых производных функции ошибки. Предполагается, что функция ошибки S находится в окрестности локального минимума. Тогда ее аппроксимация записывается в следующем виде:

$$\Delta S = \frac{1}{2}\Delta\mathbf{W}^T\mathbf{H}\Delta\mathbf{W}.$$

Пусть \mathbf{W}_k — набор параметров, соответствующий нейрону u_k , т. е. столбец матрицы \mathbf{W} , $\mathbf{W}_k = \mathbf{W}e_k$. Удаление этого нейрона (присвоение всем его параметрам нулевого значения) эквивалентно выполнению условия

$$\Delta\mathbf{W}e_k + \mathbf{W}_k = 0.$$

Получаем задачу условной минимизации

$$\Delta S = \frac{1}{2}\Delta\mathbf{W}^T\mathbf{H}\Delta\mathbf{W} \rightarrow \min, \Delta\mathbf{W}e_k + \mathbf{W}_k = 0.$$

Для решения этой задачи строим лагранжиан

$$L = \frac{1}{2}\Delta\mathbf{W}^T\mathbf{H}\Delta\mathbf{W} - \lambda(\Delta\mathbf{W}e_k + \mathbf{W}_k).$$

Продифференцировав L по $\Delta\mathbf{W}$, получаем значение выпуклости L_k для элемента \mathbf{W}_k :

$$L_k = \frac{\mathbf{w}_k^T\mathbf{w}_k}{2[\mathbf{H}^{-1}]_{k,k}},$$

где \mathbf{H}^{-1} — матрица, обратная гессиану \mathbf{H} ; $[\mathbf{H}^{-1}]_{k,k}$ — k -й диагональный элемент этой матрицы. Критерию оптимального прореживания отвечает группа параметров $\mathbf{W}_{\hat{k}}$ с минимальным значением выпуклости:

$$\hat{k} = \arg \min_{k \in \mathcal{A}} L_k.$$

Далее, используя функцию выпуклости L_k как величину, определяющую вероятность комбинирования и вариации структурных параметров, будет предложен недетерминированный вариант алгоритма *optimal brain damage*.

2.2. Генетический алгоритм оптимизации структуры нейронной сети

Функция ошибки (7) является многоэкстремальной функцией вектора параметров \mathbf{w} . Поэтому при ее минимизации одним из детерминированных алгоритмов [8, 9] определяется локальный минимум, который может не совпадать с глобальным минимумом. Для нахождения глобального минимума целесообразно воспользоваться недетерминированным генетическим алгоритмом.

Нейронная структура \mathcal{A} задается бинарным вектором $\mathbf{a} = [a_1, \dots, a_K]$:

$$\begin{cases} a_q = 1, & \text{если } k \in \mathcal{A}; \\ a_q = 0 & \text{иначе.} \end{cases}$$

Рассмотрим множество из M нейронных сетей с нейронными структурами \mathcal{A}_m , $m = [1, \dots, M]$, которому соответствует множество бинарных векторов $\mathcal{F}_0 = \{\mathbf{a}_m\}$, $m = [1, \dots, M]$. Назовем \mathcal{F}_0 популяцией. Для каждого вектора \mathbf{a}_m из множества \mathcal{F}_0 оценивается вектор параметров $\hat{\mathbf{w}}_{\mathcal{A}_m}$ соответствующей нейронной сети $\mathbf{f}_{\mathcal{A}_m}$ и вычисляется значение функции ошибки (7). Каждая из $\mathbf{f}_{\mathcal{A}_m}$ оптимизирована алгоритмом обратного распространения ошибки [14]. Опишем процедуру порождения новой популяции \mathcal{F}_1 из популяции \mathcal{F}_0 .

1. На множестве \mathcal{F}_0 задается случайная величина θ , которая принимает значение \mathbf{a}_m с вероятностью

$$p_m = \frac{\exp - \frac{Q_m}{Q_{\max}}}{\sum_{l=1}^K \exp - \frac{Q_l}{Q_{\max}}}, \quad (10)$$

где суммарная выпуклость $Q_l = \sum_{k=1}^K L_k^{\mathcal{A}_l}$ (здесь K — это число активных нейронов для нейронной сети

со структурой \mathcal{A}_l) всех активных нейронов для нейронной сети со структурой \mathcal{A}_p , а $Q_{\max} = \max_{l \in [1, \dots, M]} Q_l$.

Затем генерируется P реализаций случайной величины θ . Без ограничения общности будем считать, что P — четное число. Полученное множество векторов обозначим $F' = \{\mathbf{a}_1^T, \dots, \mathbf{a}_P^T\}$.

2. Множество F' случайным образом разбивается на пары $(\mathbf{a}_s^T, \mathbf{a}_t^T)$, где $s, t = 1, \dots, P, s \neq t$.

3. С каждой парой $(\mathbf{a}_s^T, \mathbf{a}_t^T)$ проводится операция скрещивания:

- генерируется случайное число $\zeta \in \{1, \dots, K-1\}$;
- векторы $(\mathbf{a}_s^T, \mathbf{a}_t^T)$ разделяются на две части и смешиваются следующим образом:

$$[a_s^1, \dots, a_s^\zeta, a_t^{\zeta+1}, \dots, a_t^K] \rightarrow \mathbf{a}_s,$$

$$[a_t^1, \dots, a_t^\zeta, a_s^{\zeta+1}, \dots, a_s^K] \rightarrow \mathbf{a}_t.$$

4. С каждым вектором из F' проводится операция модификации:

- генерируется случайное число $\eta < K$;
- инвертируется значение позиций η вектора \mathbf{a}_p и определяется вектор $\mathbf{a}_{p'}$.

Полученное множество векторов обозначается как $F_1 = \{\mathbf{a}_{p'}\}_{l=1}^P$ и является новой популяцией.

Таким образом, найден алгоритм для нахождения глобального минимума функции ошибки (7). Этот алгоритм использует выпуклость для определения вероятности использования структуры сети для комбинирования и вариации. Тем самым модели с меньшей выпуклостью будут с большей вероятностью использоваться для дальнейшего поиска оптимальной структуры сети.

4. Вычислительный эксперимент

Целью вычислительного эксперимента ставилось сравнение функции ошибки (7) генетического

алгоритма оптимизации структуры нейронной сети с алгоритмом NODE-OBDD.

Использовались данные с акселерометра мобильного телефона. Показания акселерометра записывались при шести видах физической активности: ходьба, бег, сидение, стояние, подъем и спуск. Далее эти показания обрабатывались экспертным порождением признаков и метрическим алгоритмом. При использовании временных рядов порождались следующие признаки: проекции среднего ускорения на координатные оси, среднеквадратические отклонения от проекций среднего ускорения на каждую из трех координатных осей, время между пиками синусоидального сигнала в миллисекундах.

В метрической выборке признаки были получены как расстояния до центроидов классов. Расстояние между рядами задавалось с помощью метода выравнивания. С методом выделения центроидов и введения расстояния между временными рядами можно ознакомиться в работе [13].

В вычислительном эксперименте оптимизировалась двуслойная нейронная сеть с 40 нейронами в скрытом слое. Оптимизация проводилась по модифицированному OBD из раздела 2.1, а также по генетическому алгоритму из раздела 2.2. Для полученных на каждой итерации моделей (для лучшей модели в поколении популяции) подсчитана функция (7), по значению которой и сравнивалось качество моделей. На рис. 1 (см. третью сторону обложки) сплошной линией представлена зависимость функции ошибки (7) и штрих-пунктирной линией — зависимость устойчивости от числа итераций алгоритмов из разделов 2.1 и 2.2, рис. 1, а и б соответствуют функциям ошибки для генетического алгоритма и NODE-OBDD для выборки, полученной с помощью ручного выделения признаков, рис. 1, в и г соответствуют функциям ошибки для генетического алгоритма и NODE-OBDD для метрической выборки.

Как видно из рис. 1, генетический алгоритм позволяет получить большую точность за меньшее число итераций для обеих выборок. При этом в случае экспертного выделения признаков устойчивость уменьшается, тогда как в метрической выборке она увеличивается.

В случае базового алгоритма NODE-OBDD функция ошибки (7) уменьшается незначительно или даже возрастает в случае экспертного порождения признаков. Устойчивость при этом в обоих случаях уменьшается.

На рис. 2 отображено значение параметра выпуклости L_j (*Saliency*) для всех активных нейронов. На каждой итерации число активных нейронов уменьшается, а абсолютные значения выпуклости у оставшихся нейронов становятся больше по абсолютному значению, также увеличивается раз-

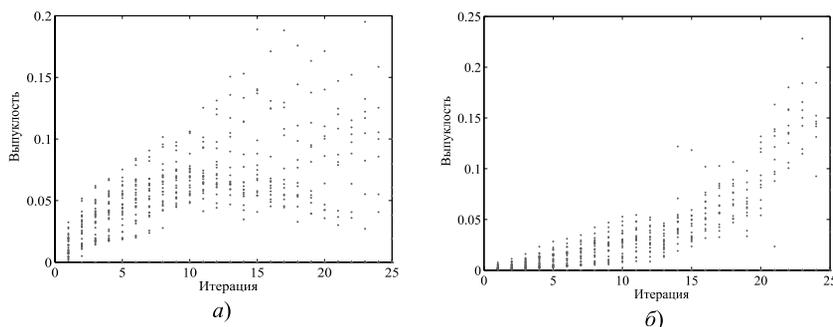


Рис. 2. Зависимость значений функции выпуклости нейронов сети от номера итерации:

а — экспертное порождение признаков; б — метрическая выборка

ность между значениями выпуклости активных нейронов.

На рис. 3 была визуализирована структура наиболее точной нейронной сети на каждой итерации. По горизонтали отложен номер итерации. Черная клетка означает, что нейрон активный, белая клетка — нейрон неактивный.

Представленные структуры получены при оптимизации сети, на вход которой подавались выборка с выделенными вручную 43 признаками [12] и выборка, полученная с помощью метода выравнивания расстояния до центроидов классов.

Для сравнения результатов, полученных с помощью исследуемых алгоритмов, с результатами других исследователей подсчитаны показатели Precision и Recall для наилучшей модели каждого из алгоритмов:

$$P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN},$$

где TP , FP , FN — это число истинно положительных, ложноположительных, ложноотрицательных объектов данного класса соответственно. Эти показатели для каждого из классов занесены в таблицу. В ней отображены показатели для двух выборок для каждого из экспериментов.

Заключение

В работе предложены два алгоритма оптимизации структуры нейронной сети — генетический алгоритм прореживания и алгоритм NODE-OBD. Эти алгоритмы сравнивались по значениям функции ошибки и устойчивости. Вычислительный эксперимент показал, что NODE-OBD позволяет значительно уменьшить число активных нейронов, не увеличивая функцию ошибки модели, а генетический алгоритм позволяет получить модель с таким же числом нейронов, как и NODE-OBD, при этом уменьшая значения функции ошибки модели. Проведено сравнение работы алгоритма для двух видов обработки временных рядов. Наиболее точные результаты, сравнимые с результатами [11], получаются при использовании метрической выборки [14] и генетического алгоритма.

Список литературы

1. Ghosh J., Turner K. Structural adaptation and generalization in supervised feed-forward networks // *Jl. of Artificial Neural Networks*, 1994. Vol. 1, N. 4. P. 431–458.
2. Chung F., Lee T. A node pruning algorithm for backpropagation networks // *Int. J. Neural Syst*, 1992. Vol. 3, N. 3. P. 301–314.

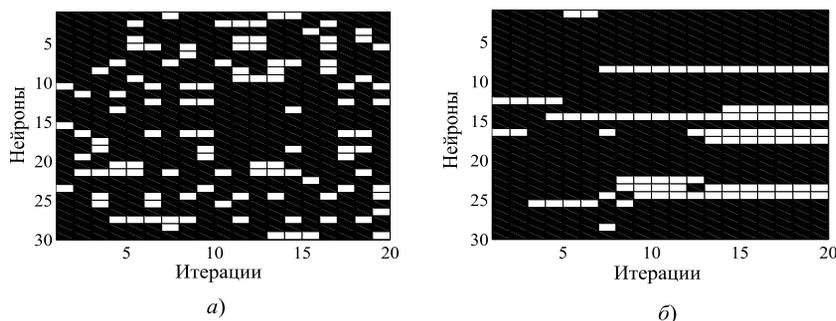


Рис. 3. Изменение структуры оптимальной сети из популяции: а — генетический алгоритм, экспертное порождение признаков; б — генетический алгоритм, метрическая выборка

Показатели самых точных моделей каждого из алгоритмов

Вид физической активности	OBD				Генетический алгоритм			
	Первая выборка		Метрическая выборка		Первая выборка		Метрическая выборка	
	P, %	R, %	P, %	R, %	P, %	R, %	P, %	R, %
Бег	83,2	84,9	100,0	98,0	88,3	86,9	100,0	98,0
Ходьба	95,1	96,3	90,9	89,2	98,0	96,7	89,1	90,7
Подъем	51,1	46,5	85,1	83,3	55,3	51,0	85,1	90,9
Спуск	46,7	46,5	82,6	90,4	40,2	50,3	93,4	89,6
Сидение	92,5	91,3	98,1	98,1	90,4	89,4	98,1	98,1
Стояние	93,1	92,1	100,0	98,0	93,1	92,1	100,0	98,0

3. Попова М. С., Стрижов В. В. Выбор оптимальной модели классификации физической активности по измерениям акселерометра // *Информатика и ее применения*, 2015. Т. 9. Вып. 1. С. 79–89.
4. Suisse M. V., Thimm G., Fiesler E., Thimm G., Fiesler E. Pruning of neural networks // *Technical report*, 2014.
5. MacLeod C., Maxwell G. M. Incremental evolution in ANNs: Neural nets which grow // *Artif. Intell. Rev*, 2001. Vol. 16, N. 3. P. 201–224.
6. Vukovic N., Miljkovic Z. A growing and pruning sequential learning algorithm of hyper basis function neural network for function approximation // *Neural Networks*, 2013. Vol. 46. P. 210–226.
7. Strijov V. V., Krymova E. V., Weber G. W. Evidence optimization for consequently generated models // *Mathematical and Computer Modelling*, 2013. Vol. 57, N. 1–2. P. 50–56.
8. Cun Y. L., Denker J. S., Solla S. A. Optimal brain damage // *Advances in neural information processing systems*, 1990. Vol. 2. P. 598–605.
9. Hassibi B., Stork D. G. Second order derivatives for network pruning: Optimal brain surgeon // *Advances in Neural Information Processing Systems*, 1993. Vol. 5, N. 1. P. 164–172.
10. Leung F. H., Lam H., Ling S., Tat P. K. Tuning of the structure and parameters of a neural network using an improved genetic algorithm // *IEEE Transactions on Neural Networks*, 2003. Vol. 14, N. 1. P. 79–88.
11. Oh I., Lee J., Moon B. R. Hybrid genetic algorithms for feature selection // *IEEE Trans. Pattern Anal. Mach. Intell*, 2004. Vol. 26, N. 11. P. 1424–1437.
12. Kwapisz J. R., Weiss G. M., Moore S. Activity recognition using cell phone accelerometers // *SIGKDD Explorations*, 2010. Vol. 12, N. 2. P. 74–82.
13. Rasekh A., Chen C., Lu Y. Human activity recognition using smartphone // *Technical report*, 2014.
14. Гончаров А. В., Стрижов В. В., Попова М. С. Метрическая классификация временных рядов с выравниванием относительно центроидов классов // *Системы и средства информатики*, 2015. Т. 1. № 4.

A. I. Zadayanchuk¹, Student, e-mail: alex.goncharov@phystech.edu,
M. S. Popova¹, Student, V. V. Strijov², Leading Scientist
¹ Moscow Institute of Physics and Technology
² Dorodnicyn Computing Center, Russian Academy of Sciences

Selection of Optimal Time Series Classification Model

This paper solves the problem of selecting optimal stable models for classification of physical activity. We select optimal models from the class of two-layer artificial neural networks. There are three different ways to change structure of neurons: network pruning, network growing, and their combination. We construct models by removing its neurons. Neural networks with insufficient or excess number of neurons have insufficient generalization ability and can make unstable predictions. Proposed genetic algorithm optimizes the neural network structure. The novelty of the work lies in the fact that the probability of removing neurons is determined by the variance of parameters. In the computing experiment, models are generated by optimization two quality criteria — accuracy and stability.

Keywords: classification, artificial neural networks, accuracy, genetic algorithm, pruning criteria

References

1. Ghosh J., Turner K. Structural adaptation and generalization in supervised feed-forward Networks, *Jl. of Artificial Neural Networks*, 1994, vol. 1, no. 4, pp. 431–458.
2. Chung F., Lee T. A node pruning algorithm for backpropagation networks, *Int. J. Neural Syst.*, 1992, vol. 3, no. 3, pp. 301–314.
3. Strijov V. V., Popova M. S. Vybor optimal'noj modeli klasifikacii fizicheskoy aktivnosti po izmerenijam akselerometra, *Informatika i ee primenenija*, 2015, vol. 9, no. 1, pp. 79–89.
4. Suisse M. V., Thimm G., Fiesler E., Thimm G., Fiesler E. Pruning of neural networks, *Technical report*, 2014.
5. MacLeod C., Maxwell G. M. Incremental evolution in ANNs: Neural nets which grow, *Artif. Intell. Rev.*, 2001, vol. 16, no. 3, pp. 201–224.
6. Vukovic N., Miljkovic Z. A growing and pruning sequential learning algorithm of hyper basis function neural network for function approximation, *Neural Networks*, 2013, vol. 46, pp. 210–226.
7. Strijov V. V., Krymova E. V., Weber G. W. Evidence optimization for consequently generated models, *Mathematical and Computer Modelling*, 2013, vol. 57, no. 1–2, pp. 50–56.
8. Cun Y. L., Denker J. S., Solla S. A. Optimal brain damage, *Advances in neural information processing systems*, 1990, vol. 2, pp. 598–605.
9. Hassibi B., Stork D. G. Second order derivatives for network pruning: Optimal brain surgeon, *Advances in Neural Information Processing Systems*, 1993, vol. 5, no. 1, pp. 164–172.
10. Leung F. H., Lam H., Ling S., Tam P. K. Tuning of the structure and parameters of a neural network using an improved genetic algorithm, *IEEE Transactions on Neural Networks*, 2003, vol. 14, no. 1, pp. 79–88.
11. Oh I., Lee J., Moon B. R. Hybrid genetic algorithms for feature selection, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2004, vol. 26, no. 11, pp. 1424–1437.
12. Kwapisz J. R., Weiss G. M., Moore S. Activity recognition using cell phone accelerometers, *SIGKDD Explorations*, 2010, vol. 12, no. 2, pp. 74–82.
13. Rasekh A., Chen C., Lu Y. Human activity recognition using smartphone, *Technical report*, 2014.
14. Goncharov A. V., Strijov V. V., Popova M. S. Metricheskaja klassifikacija vremennyh rjadov s vyvavnivaniem odnositel'no centroidov klassov, *Sistemy i sredstva informatiki*, 2015, vol. 1, no. 4.

ИНФОРМАЦИЯ

В Волгограде с 1 по 3 июня 2016 года состоится выставка-форум "Мир ТелеКом" —

ХIII выставка информационных технологий.

Выставка зарекомендовала себя высокоэффективной площадкой для проведения деловых встреч и переговоров и демонстрации технических достижений.

Разделы выставки:

IT городской среды: Облачные решения. Телефония. Системы доставки корреспонденции. Электронное правительство. Системы электронных торгов. Геоинформационные системы. Управление ИТ. ИТ системы для библиотек. ИТ технологии коммунального хозяйства, обращения с отходами, охраны окружающей среды. Городское видеонаблюдение. "Зеленые" ИТ технологии.

Телекоммуникационные технологии: Проводные, беспроводные, мобильные, спутниковые технологии. Проектирование, строительство, эксплуатация сетей. Интернет технологии. Веб-дизайн. Телевидение. IP-технологии / VoIP. Data Center Технологии и решения. Телематика. М2М — межкомпьютерное взаимодействие. РС, ноутбуки, нетбуки, планшетные ПК. Hardware & Software, приложения. Электронная коммерция. Цифровой маркетинг. Биометрия.

Развитие ИТ: Start-up/Spin-off. Образование, здравоохранение, культура. Электронное обучение. Венчурный капитал. Инвестиции и финансирование. Web-Медиа. Интерактивное телевидение. Умный дом. Автоматизация. Автомобильные ИТ решения. Робототехника (игровые и сервис роботы). Прикладные и фундаментальные исследования.

ИТ для бизнеса: ИТ-услуги / Аутсорсинг. Программные продукты для предприятия. Системы планирования на предприятиях (ERP). Управление взаимоотношениями с клиентами (CRM). Управление бизнес-процессами (BPM, ESM). Управление человеческими ресурсами (HR). Системы автоматической идентификации пользователей. Логистика ИТ. Open Source системы. Виртуализация, облачные вычисления. Интеграция и корпоративные решения.

ИТ-безопасность: ИТ технологии общественной безопасности и внутренней безопасности. Mobile Security. Коммуникации, защита. Anti-Spam & Anti-Virus — решения. Безопасность облачных решений. Web-безопасность. Сетевая безопасность.

Подробная информация по выставке: <http://www.regionex.ru/exhibits/2016/telekom/>

По вопросу участия и посещения обращаться в Волгоградский Выставочный Центр "Регион".

С уважением,

Волгоградский Выставочный Центр "Регион"

Тел./факс: (8442) 26-51-86, 24-26-02, 23-28-99

E-mail: bez2@regionex.ru