

Выбор оптимальной модели классификации физической активности по измерениям акселерометра

Авторы: А. И. Задаянчук М. С. Попова, В. В. Стрижов

Аннотация: 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 прорезживания, генетический алгоритм.*

Selection of optimal physical activity classification model using measurements of accelerometer

Authors: Andrew Zadayanchuk, Maria Popova, Vadim Strijov

Annotation: 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.*

1 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 [4].

Оптимизировать размерность пространства можно на разных уровнях — на уровне нейронов (наборов параметров) [2] и отдельных параметров [10]. 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) [14], наращивание (network growing) [9] и пошаговое чередование наращивания и прореживания [15, 13, 10]. В настоящей работе рассматривается оптимизация размерности с помощью прореживания. Базовыми алгоритмами прореживания нейронных сетей являются оптимальное прореживание (англ. «optimal brain damage» [3] и «optimal brain surgery» [6]), основанные на вычислении вторых производных функции ошибки.

Устойчивая и оптимальная модель описывается с помощью генетического алгоритма на уровне нейронов [8, 11] путем оптимизации структурных параметров. Базовые алгоритмы [3, 6] 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 [8] набора нейронных сетей. Вероятность комбинирования и вариации структурных параметров нейронной сети тем меньше, чем больше показатель выпуклости, используемый в «optimal brain damage» [3].

В вычислительном эксперименте рассматривается задача классификации физической активности человека по измерениям акселерометра. Эта задача решалась в исследованиях [7, 12] 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 [7] и вводом метрики выравнивания временных рядов, с последующим выделением признаков, полученных как расстояние до центроидов классов [1].

2. Постановка задачи классификации

Дана выборка $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{t}_i)\}, i \in \mathcal{I} = \{1 \dots m\}$, 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 $\mathcal{I} = \mathcal{L} \sqcup \mathcal{C}$ 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:*

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

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

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:

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

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

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

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

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

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

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

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

Моделью \mathbf{f} является суперпозиция функций (3), (4), (5). В данной работе мы будем исследовать модели, отличающиеся друг от друга на структурном уровне. А именно содержащие разное количество 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 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):

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

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:

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

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} :

$$\mathbf{A}^{-1} = \text{cov}(\mathbf{W}) = \mathbb{E}(\mathbf{W}^T \mathbf{W}) - \mathbb{E}(\mathbf{W})\mathbb{E}(\mathbf{W}^T) = \mathbb{E}(\mathbf{W}^T \mathbf{W})$$

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:

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

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

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

Устойчивость модели будет дополнительным критерием качества.

3 Описание алгоритма

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

3.1 Алгоритм прореживания структуры нейронной сети NODE-OBD

Предлагаемый алгоритм определяет индекс нейрона, удаление которого приведет к минимизации приращения функции ошибки (7). Удаление нейрона эквивалентно занулению 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 соответствующая нейрону u_k , т.е. столбец матрицы \mathbf{W} , $\mathbf{W}_k = \mathbf{W}\mathbf{e}_k$. Удаление этого нейрона (присвоение всем его параметрам нулевого значения) эквивалентно выполнению условия

$$\Delta\mathbf{W}\mathbf{e}_k + \mathbf{W}_k = \mathbf{0},$$

Получаем задачу условной минимизации

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

Для решения этой задачи строим лагранжиан

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

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 -ый диагональный элемент этой матрицы. Критерию оптимального прореживания отвечает группа параметров $\mathbf{W}_{\hat{k}}$ с минимальным значением выпуклости:

$$\hat{k} = \underset{k \in \mathcal{A}}{\operatorname{argmin}} L_k.$$

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

3.2 Генетический алгоритм оптимизации структуры нейронной сети

Функция ошибки (7) является многоэкстремальной функцией вектора параметров \mathbf{w} . Поэтому при ее минимизации одним из детерминированных алгоритмом [3, 6] 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 [5]. 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

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

$$\text{где суммарная выпуклость } Q_l = \sum_{k=1}^K L_k^{\mathcal{A}_l}, \quad (10.1)$$

всех активных нейронов для нейронной сети со структурой \mathcal{A}_l , а

$$Q_{\max} = \max_{l \in \{1, \dots, M\}} Q_l.$$

В формуле (10.1) K – это число активных нейронов для нейронной сети со структурой \mathcal{A}_l .

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

2. Множество \mathcal{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 $(\mathbf{a}_s^T, \mathbf{a}_t^T)$ производится 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 \mathcal{F}' проводится операция модификации:
 - генерируется случайное число $\eta < K$;
 - инвертируется значение позиций η вектора \mathbf{a}_t , и получается вектор \mathbf{a}_t'' .

Полученное множество векторов обозначается как $\mathcal{F}_1 = \{\mathbf{a}_t''\}_{t=1}^P$ и является новой популяцией.

Таким образом, получен 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.

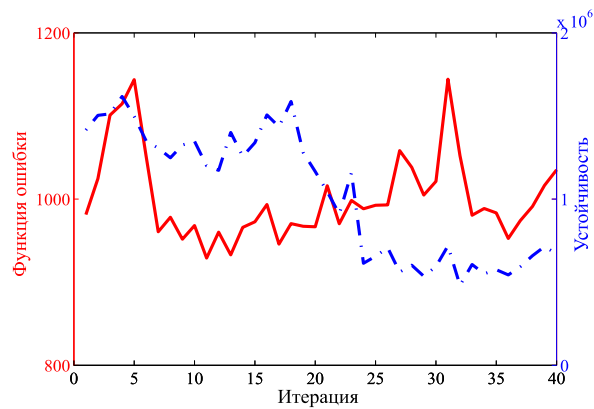
4 Вычислительный эксперимент

Целью вычислительного эксперимента ставилось сравнение функции ошибки (7) генетического алгоритма оптимизации структуры нейронной сети с алгоритмом NODE-OBD.

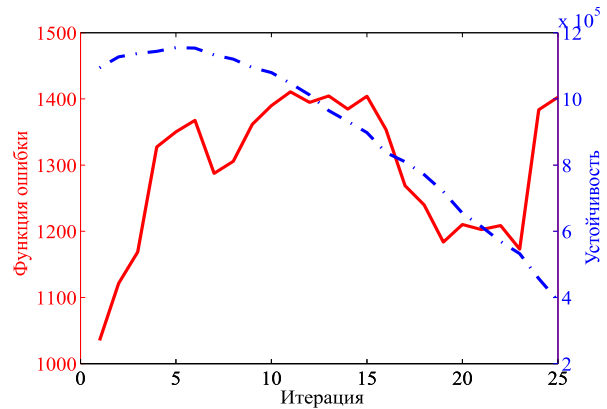
Использовались данные с акселерометра мобильного телефона. Показания акселерометра записывались при шести видах физической активности: ходьба, бег, сидение, стояние, подъем и спуск. Далее эти показания обрабатывались экспертным порождением признаков и метрическим алгоритмом. При использовании временных 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 [1].

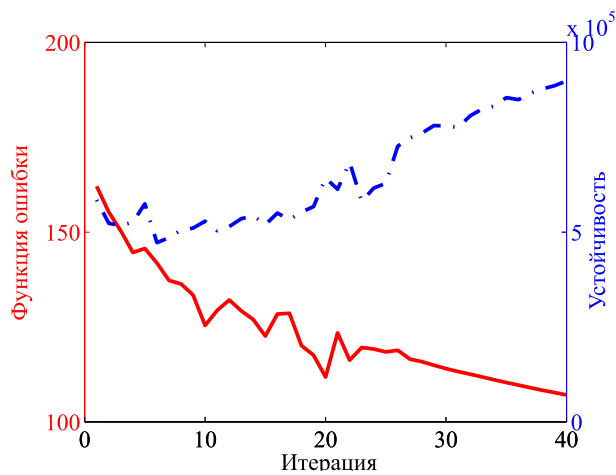
В вычислительном эксперименте оптимизировалась двуслойная нейронная сеть с 40 нейронами в скрытом слое. Оптимизация проводилась по модифицированному OBD из раздела 3.1, а также по генетическому алгоритму из раздела 3.2. Для полученных на каждой итерации моделей (для лучшей модели в поколении популяции) подсчитана функция (7) по значению которой и сравнивалось качество моделей. На Рис. 1 представлена зависимость



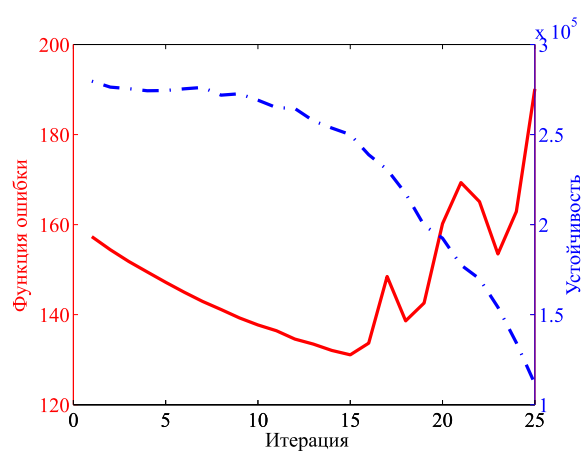
(а) Генетический алгоритм.
Экспетрное порождение признаков



(б) NODE-OBD. Экспетрное порождение признаков



(а) Генетический алгоритм.
Метрическая выборка



(б) NODE-OBD.
Метрическая выборка

функции ошибки (7) и устойчивости от числа итераций алгоритмов из разделов 3.1 и 3.2.

Рисунок 1. Зависимость значений функции ошибки от номера итерации

Как видно из Рис. 1, генетический алгоритм позволяет получить большую точность за меньшее количество итераций для обоих выборок. При этом в случае экспертного выделения признаков устойчивость уменьшается, тогда как в метрической выборке она увеличивается. В случае базового алгоритма NODE-OBД 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 количество активных нейронов уменьшается, а абсолютные значения выпуклости у оставшихся нейронов становятся больше по абсолютному значению, также увеличивается разность между значениями выпуклости активных нейронов

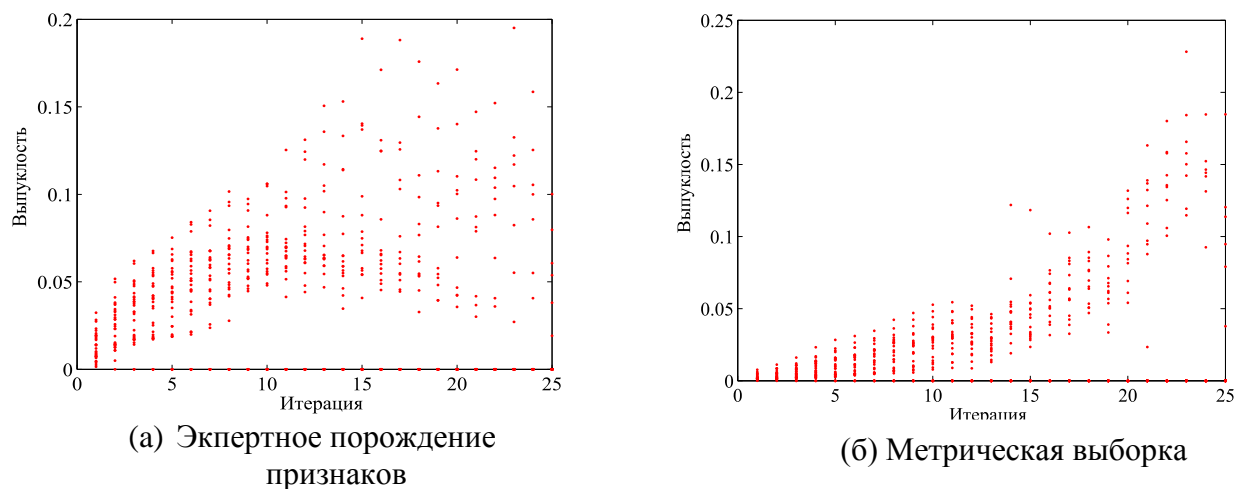


Рисунок 2. Зависимость значений функции выпуклости нейронов сети от номера итерации

Также 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.

На Рис. 3 представлены структуры, полученные при оптимизации сети, на вход которой подавались выборка с выделенными вручную 43 признаками [7] и выборка полученная с помощью метода выравнивания расстояния до центроидов классов.

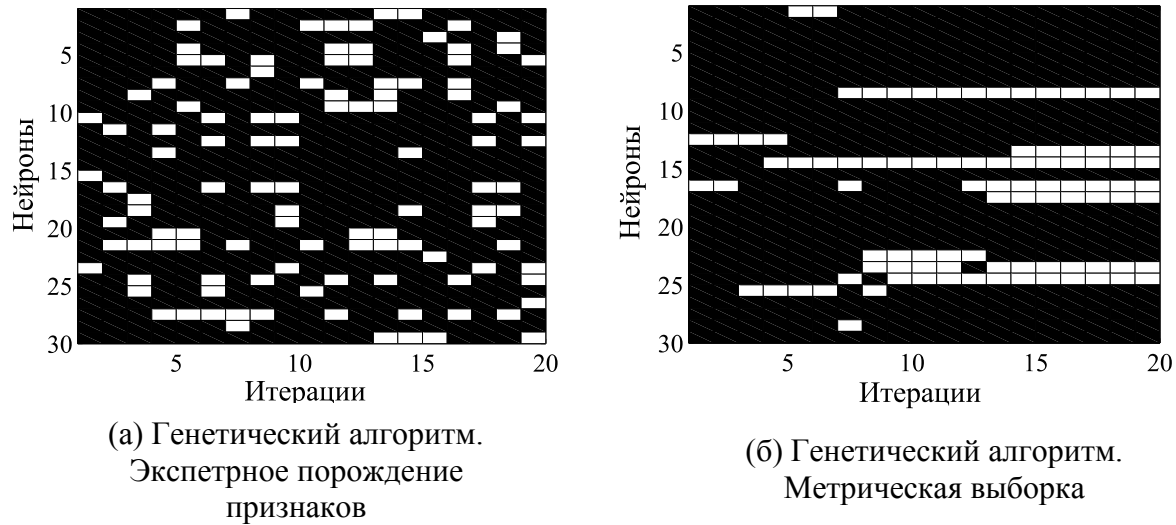


Рисунок 3. Изменение структуры оптимальной сети из популяции.

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

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

где TP, FP, FN — это количество истинно-положительных, ложно-положительных, ложно-отрицательных объектов данного класса, соответственно. Эти показатели для каждого из классов занесены в таблицу 1. В ней отображены показатели для двух выборок для каждого из экспериментов.

Таблица 1. Показатели самых точных моделей каждого из алгоритмов

Алгоритм	OBD				Генетический			
	Первая		Метрическая		Первая		Метрическая	
Выборка	P,%	R,%	P,%	R,%	P,%	R,%	P,%	R,%
Бег	83,2	84,9	100	98,0	88,3	86,9	100	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	98,0	93,1	92,1	100	98,0

5 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 количество активных нейронов, не увеличивая функцию ошибки модели, а генетический алгоритм позволяет получить модель с таким же количеством 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 [1] и генетического алгоритма.

Литература

1. Гончаров А. В., Стрижов В.В., Попова М.В. Метрическая классификация временных рядов с выравниванием относительно центроидов классов // Системы и средства информатики, 2015. Vol.1. №4.
2. Chung F., Lee T. A node pruning algorithm for backpropagation networks // Int. J. Neural Syst , 1992. Vol.3, No.3.P. 301–314.
3. Cun Y.L., Denker J.S., Solla S.A. Optimal brain damage // Advances in neural information processing systems, 1990. Vol.2. P. 598–605.
4. Ghosh J, Tumer K. Structural adaptation and generalization in supervised feed-forward networks // JI. of Artificial Neural Networks, 1994. Vol.1. No.4. P. 431–458.
5. Goh A. Back-propagation neural networks for modeling complex systems // Artificial Intelligence in Engineering, 1995.Vol.9. No.3. P. 143–151.
6. 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. P. 164–172.
7. Kwapisz J. R., Weiss G. M., Moore S. Activity recognition using cell phone accelerometers // SIGKDD Explorations, 2010. Vol.12. No.2. P. 74–82.
8. 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. P. 79–88.
9. MacLeod C., Maxwell G.M. Incremental evolution in ANNs: Neural nets which grow // Artif. Intell. Rev, 2001. Vol.16. No.3. P. 201–224.
10. Strijov V.V., Popova M.V. Selection of optimal physical activity classification model using measurements of accelerometer // Informatics and applications, 2015. Vol.9. No.1. P. 79–89.
11. Ohl, Lee J., Moon B. R., Hybrid genetic algorithms for feature selection // IEEE Trans. Pattern Anal. Mach. Intell, 2004.Vol.26. No.11. P.1424–1437.

12. *Rasekh A., Chen C., Lu Y.* Human activity recognition using smartphone // Technical report, 2014.

13. *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. P.50–56.

14. *Suisse M.V., Thimm G., Fiesler E., Thimm G., Fiesler E.* Pruning of neural networks // Technical report, 2014.

15. *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.