

Экспериментальное сравнение рандомизированных и полносвязных нейросетей для выявления аномалий в промышленных системах

Е. О. Кузнецова

*Санкт-Петербургский государственный
электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)*

eokuznetsova@stud.etu.ru

Е. В. Федорченко

*Санкт-Петербургский государственный
электротехнический университет «ЛЭТИ»
им. В.И. Ульянова (Ленина);*

Китайский горно-технологический университет
evfedorchenko@etu.ru

Аннотация. В настоящее время актуальной является проблема выявления аномалий в промышленных системах в условиях ограниченных ресурсов. Утверждается, что рандомизированные нейросети позволяют сохранить качество выявления аномалий в условиях ограниченных ресурсов. В работе проводится экспериментальное сравнение рандомизированных нейросетей, на примере стохастически конфигурируемых нейросетей, и полносвязных нейросетей, на примере персептрона, с точки зрения качества и потребляемых ресурсов. Для экспериментов выбраны промышленный набор данных о передвижениях самосвалов и база образцов рукописного текста. По итогам проведенных экспериментов приводятся выводы и выбирается архитектура для дальнейших исследований.

Ключевые слова: *рандомизированные нейросети; стохастически конфигурируемые нейросети; персептрон; выявление аномалий; промышленные системы*

I. ВВЕДЕНИЕ

Машинное обучение дает много возможностей по применению уже накопленных данных для извлечения знаний и принятия решений в самых разных областях. Эффективность обученной модели во многом зависит от выбранной архитектуры нейросети, а именно, количества скрытых узлов и выходных весов, которая может быть построена с использованием детерминированных или рандомизированных алгоритмов. Построение оптимальной архитектуры является ресурсоемкой задачей [1]. В то же время, в промышленных системах часто используются небольшие устройства, ограниченные по ресурсам. Кроме того, в настоящее время активно исследуется концепция федеративного обучения, позволяющая сохранить конфиденциальность данных за счет локального обучения на таких устройствах. Для снижения ресурсозатрат были предложены рандомизированные алгоритмы [2, 3]. В [4] М. Ли и Д. Ванг показали, что построенные с применением

рандомизированных алгоритмов нейросети не обладают свойством универсальной аппроксимации. Поэтому в [5] М. Ли и Д. Ванг предложили управляющий механизм, накладывающий ограничения на случайные параметры и адаптирующий область определения параметров для обеспечения свойства универсальной аппроксимации. В данной статье приводятся результаты экспериментального сравнения полносвязных нейросетей, на примере персептрона, и рандомизированных нейросетей, на примере стохастически конфигурируемых нейросетей, по точности и временным затратам, на наборах данных двух типов – временные ряды и изображения. Целью исследования является экспериментальное подтверждение эффективности стохастически конфигурируемых нейросетей для решения задач анализа промышленных данных в условиях ограниченных ресурсов для последующего применения в федеративном режиме.

Для достижения поставленной цели были реализованы два алгоритма обучения стохастически конфигурируемых нейросетей и алгоритм обучения персептрона, и проведены эксперименты на двух наборах данных, содержащих временные ряды и изображения.

Вклад и новизна данного исследования состоит в полученных результатах экспериментального сравнения эффективности заявленных алгоритмов для решения задач анализа промышленных данных.

Работа организована следующим образом. В разделе II приведены основные определения и постановка задачи исследования. В разделе III описывается методология исследования. В разделе IV описаны проведенные эксперименты и их результаты.

II. ПОСТАНОВКА ЗАДАЧИ

A. Исходные данные

Набор данных Signal. Содержит данные о передвижении двух самосвалов АН и HL, собранные компанией Smartilizer Scandinavia AB с участка рекультивации грунта в окрестностях Гётеборга в

течение 6,5 часов: 5 часов данных от АН и 75 минут данных от HL¹. Содержимое набора данных приводится в табл. 1. Фрагмент набора данных приводится в табл. 2. Набор данных MNIST. Содержит изображения рукописных цифр². Включает обучающую выборку, содержащую 60000 экземпляров, и тестовую выборку, содержащую 10000 экземпляров. Пример изображения из набора данных MNIST представлен на рис. 1.

ТАБЛИЦА I. СОДЕРЖИМОЕ НАБОРА ДАННЫХ SIGNAL

Параметр	Описание
time	Время в секундах
gFx	данные с гироскопа
wx	акселерометр
speed	скорость в км/ч
label	1 = холостой ход, 2 = движение, 3 = загрузка, 4 = сброс

ТАБЛИЦА II. ФРАГМЕНТ НАБОРА ДАННЫХ SIGNAL

time	0.0	0.0100000070759929	0.0200000141519858
gFx	-0.2727	-0.2727	-0.2895
gFy	0.0442	0.0442	0.0462
gFz	0.9382	0.9382	0.9934
wx	-0.0053	-0.0053	-0.001
wy	-0.0077	-0.0077	0.0069999999999999
wz	-0.0331	-0.0331	-0.0312
speed	0.0	0.0	0.0
label	1.0	1.0	1.0

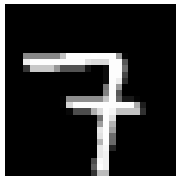


Рис. 1. Пример изображения из набора данных MNIST

В. Методы обучения нейронных сетей

Перцептрон. Данные нейросети были выбраны по причине относительно простой архитектуры с одним скрытым слоем, пороговой передаточной функцией и прямым распространением сигнала. Пороговая функция определяется на основе входных сигналов x как $f(x) = \text{sign}(\sum_{i=1}^d w_i x_i - b)$, где b – пороги, изначально выбираемые случайным образом; w_i – весовые коэффициенты, $w_i \in \{-1; 0; +1\}$, изначально выбираемые случайным образом, которые корректируются в процессе обучения в зависимости от того, относится ли входной сигнал к предъявленному образцу.

Стохастически конфигурируемая нейросеть. Данные нейросети относятся к классу рандомизированных нейросетей, то есть в отличие от предыдущей модели их архитектура формируется автоматически при обучении модели. Для обеспечения свойства универсальной аппроксимации рандомизированных нейросетей в [5] М. Ли и Д. Ванг предложили управляющий механизм, накладывающий ограничения на случайные параметры, то есть входные веса и смещения (w и b), и

адаптирующий область определения параметров. Ниже кратко описаны предлагаемые ими методы обучения стохастически конфигурируемой нейросети [5].

Предусловия: пусть $f: \square^d \rightarrow \square^m$ – целевая функция; построена стохастически конфигурируемая нейросеть SCN с $L-1$ скрытых узлов:

$$f_{L-1}(x) = \sum_{j=1}^{L-1} \beta_j g_j(w_j^T x + b_j) \quad (L = 1, 2, \dots, f_0 = 0),$$

где $\beta_j = [\beta_{j,1}, \dots, \beta_{j,m}]^T$ – выходные веса нейронов, $g_j(w_j^T x + b_j)$ – функция активации;

$e_{L-1} = f - f_{L-1} = [e_{L-1,1}, \dots, e_{L-1,m}]$ – текущая ошибка; и $\|e_{L-1}\|$ – норма ошибки, определенная в пространстве функций, измеримых по Лебегу, следующим образом

$$\|f\| := \left(\sum_{q=1}^m \int_D |f_q(x)|^2 dx \right)^{1/2} < \infty.$$

Тогда новая случайная функция активации (новые веса и смещения) формируется, если норма ошибки $\|e_{L-1}\|$ не попадает в заданный интервал. При этом используется управляющий механизм, который определяется следующим неравенством:

$$\langle e_{L-1,q}, g_L \rangle^2 \geq b_g^2 \delta_{L,q}, \quad q = 1, 2, \dots, m,$$

где $b_g \in \square^+$, $\forall g \in \Gamma$, Γ – Гильбертово пространство; $0 < \|g\| \leq b_g$, $\|g\|$ – мера Лебега для g ;

$\langle e_{L-1,q}, g_L \rangle$ – скалярное произведение $e_{L-1,q}$ и g_L , определяемое следующим образом

$$\langle f, \theta \rangle := \sum_{q=1}^m \langle f_q, \theta_q \rangle = \sum_{q=1}^m \int f_q(x) \theta_q(x) dx;$$

$\delta_{L,q} = (1 - r - \mu_L) \|e_{L-1,q}\|^2$, $0 < r < 1$, $\{\mu_L\}$ – последовательность неотрицательных вещественных чисел, $\lim_{L \rightarrow +\infty} \mu_L = 0$, $\mu_L \leq (1 - r)$, $L = 1, 2, \dots$, $\delta_L = \sum_{q=1}^m \delta_{L,q}$.

Выходные веса определяются следующим образом:

$$\beta_{L,q} = \frac{\langle e_{L-1}, g_L \rangle}{\|g_L\|^2}, \quad q = 1, 2, \dots, m.$$

Недостатком данного алгоритма является низкая скорость сходимости модели. Для устранения данного недостатка в [5] предложена модифицированная версия алгоритма, предполагающая оценку части выходных весов путем минимизации ошибки методом наименьших квадратов, а не определение по формуле выше. Пусть L – номер скрытого узла, K – размер окна, тогда если $L \leq K$ выходные веса оцениваются путем минимизации ошибки методом наименьших квадратов:

$$[\beta_1^*, \beta_2^*, \dots, \beta_L^*] = \arg \min_{\beta} \left\| f - \sum_{j=1}^L \beta_j g_j \right\|. \quad \text{Если } L > K,$$

выходные веса оцениваются выборочно, т. е. $\beta_1^*, \beta_2^*, \dots, \beta_{L-K}^*$ остаются неизменными, а оставшиеся веса определяются следующим образом:

¹ <https://www.kaggle.com/datasets/smartilizer/commercial-vehicles-sensor-data-set>

² <https://www.kaggle.com/datasets/hojjatk/mnist-dataset>

$$[\beta_{L-K+1}^*, \beta_{L-K+2}^*, \dots, \beta_L^*] =$$

$$= \arg \min_{\beta_{L-K+1}, \dots, \beta_L} \left\| f - \sum_{j=1}^{L-K} \beta_j^* g_j - \sum_{j=L-K+1}^L \beta_j g_j \right\|$$

С. Задача

С использованием наборов данных, представленных в п.П.А и методов, описанных в п.П.В обучить перцептрон и стохастически конфигурируемую нейросеть, и получить значения следующих показателей: точность модели (Accurasy); среднеквадратическая ошибка (RMSE) и время обучения модели.

III. МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

В рамках данного исследования для проведения экспериментов были выбраны наборы данных разного типа: временные ряды (Signal) и изображения (MNIST). Далее на языке Python были реализованы алгоритмы обучения перцептрона и стохастически конфигурируемой нейросети. Для обучения перцептрона использовалась библиотека Scikit-learn. Для обучения стохастически конфигурируемой нейросети использовались алгоритмы, основанные на методах, описанных в разделе II. Использовалась открытая реализация первого алгоритма³ (фиксированная формула вычисления выходных весов) и дополнительно был реализован второй алгоритм (оценка части выходных весов путем минимизации ошибки методом наименьших квадратов). При обучении фиксировалось время обучения моделей. Последовательно проведены эксперименты на наборах данных Signal (временные ряды) и MNIST (изображения) с обученными моделями:

- Перцептрон
- Стохастически конфигурируемая нейросеть (фиксированная формула вычисления выходных весов)
- Стохастически конфигурируемая нейросеть (оценка части выходных весов путем минимизации ошибки методом наименьших квадратов), размер окна $K \in \{5; 10; 15; 25; 50; 100\}$

При этом фиксировалась точность каждой модели. По результатам экспериментов, на основе полученных показателей, были сделаны выводы о применимости выбранных моделей для выявления аномалий в промышленных системах и их обучения в федеративном режиме.

IV. ЭКСПЕРИМЕНТЫ И РЕЗУЛЬТАТЫ

Эксперименты и результаты представлены ниже в соответствии с методологией из раздела III.

На наборе данных Signal для перцептрона наилучшее полученное значение точности (Accurasy) на обучающей выборке составило 0.5599, на тестовой выборке – 0.5552. Время обучения составило 19 секунд.

Для стохастически конфигурируемой нейросети (фиксированная формула вычисления выходных весов) результат для всего набора данных не удалось получить за приемлемое время, поэтому последующие эксперименты были проведены на 10 % случайно выбранных записей набора данных, наилучшее

полученное значение точности (Accurasy) на обучающей выборке составило 0.7487, RMSE – 0.6211. Время обучения составило 3463.38 секунды.

Для стохастически конфигурируемой нейросети (оценка части выходных весов путем минимизации ошибки методом наименьших квадратов), наилучшие результаты по точности (RMSE и Accurasy) для размера окна $K = 100$ представлены на рис. 2.

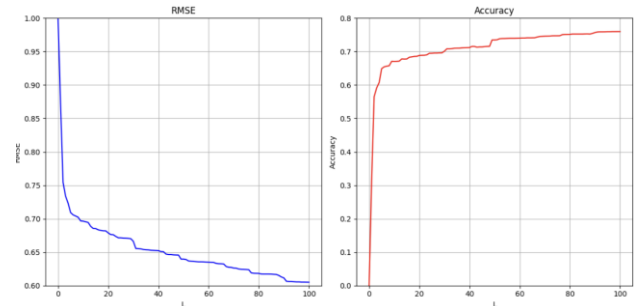


Рис. 2. RMSE и Accurasy стохастически конфигурируемой нейросети на наборе данных Signal, размер окна $K=100$

Для стохастически конфигурируемой нейросети (оценка части выходных весов путем минимизации ошибки методом наименьших квадратов), точность (Accurasy) в зависимости от размера окна $K \in \{5; 10; 15; 25; 50; 100\}$ представлена на рис. 3.

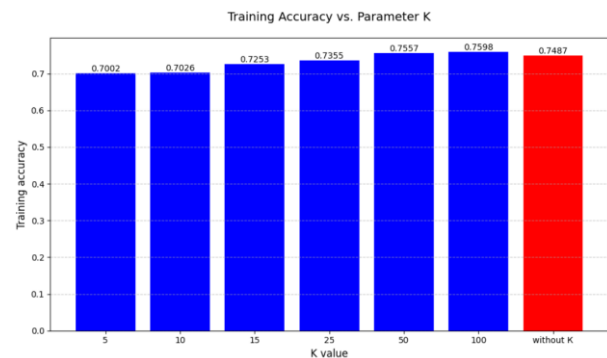


Рис. 3. Accurasy стохастически конфигурируемой нейросети на наборе данных Signal в зависимости от размера окна K

На рис. 4 представлена зависимость времени обучения от размера окна K .



Рис. 4. Зависимость времени обучения стохастически конфигурируемой нейросети на наборе данных Signal от размера окна K

На наборе данных MNIST для перцептрона наилучшее полученное значение точности (Accurasy) на

³ http://www.deepscn.com/SCN_release_v1_p.zip

обучающей выборке составило 0.8549, на тестовой выборке – 0.8630. Время обучения составило 15 секунд.

Для стохастически конфигурируемой нейросети (фиксированная формула вычисления выходных весов) наилучшее полученное значение точности (Accuracy) на обучающей выборке составило 0.8383, RMSE – 0.6492. Время обучения составило 342.18 секунд. Для стохастически конфигурируемой нейросети (оценка части выходных весов путем минимизации ошибки методом наименьших квадратов), наилучшие результаты по точности (Accuracy) и ошибке (RMSE) для размера окна $K=100$ представлены на рис. 5.

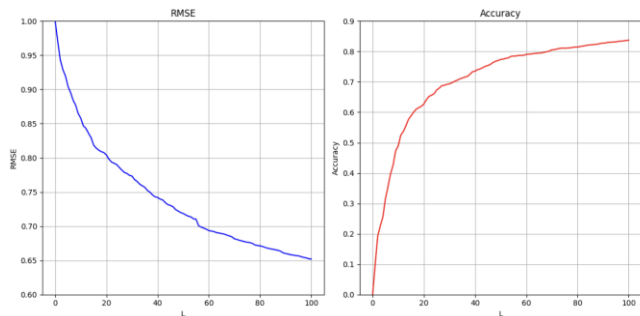


Рис. 5. RMSE и Accuracy стохастически конфигурируемой нейросети на наборе данных MNIST, размер окна $K=100$

Для стохастически конфигурируемой нейросети (оценка части выходных весов путем минимизации ошибки методом наименьших квадратов), точность (Accuracy) в зависимости от размера окна $K \in \{5; 10; 15; 25; 50; 100\}$ представлена на рис. 6.

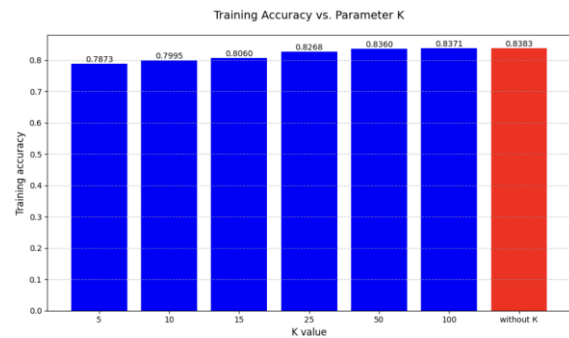


Рис. 6. Accuracy стохастически конфигурируемой нейросети на наборе данных MNIST в зависимости от размера окна K

На рис. 7 представлена зависимость времени обучения от размера окна K . Результаты для персептрона и стохастически конфигурируемой нейросети (СКН) сведены в табл. 3. Из табл. 3 видно, что для временных рядов наилучшие результаты по точности (Accuracy) показала СКН с K , однако для этого потребовалось больше всего времени. Наилучшие результаты по ошибке (RMSE) показала СКН без K , за меньшее время.

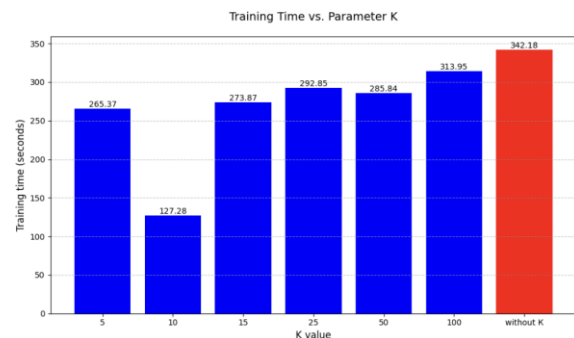


Рис. 7. Зависимость времени обучения стохастически конфигурируемой нейросети на наборе данных MNIST от размера окна

Для изображений наилучшие результаты по точности (Accuracy) показал персептрон со значительным выигрышем по времени.

ТАБЛИЦА III. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Показатели Модель	Наборы данных	Accuracy (обучающая выборка)	Accuracy (тестовая)	RMSE (обучающая)	Время обучения, секунды
Персептрон	Signal	0.5599	0.5552		19
	MNIST	0.8549	0.8630		15
СКН (без K)	Signal	0.7487		0.6211	3463.38
	MNIST	0.8383		0.6492	342.18
СКН ($K=100$)	Signal	0.7557		0.6185	4527.82
	MNIST	0.8371		0.6520	313.95

V. ЗАКЛЮЧЕНИЕ

В рамках исследования было проведено экспериментальное сравнение персептрона и стохастически конфигурируемых нейросетей по точности и временным затратам на наборах данных двух типов – временные ряды и изображения. Эксперименты показали, что стохастически конфигурируемые нейросети выигрывают по точности на временных рядах, однако для этого требуется значительно больше времени, кроме того, стандартные реализации алгоритмов не позволили обучить модель за приемлемое

время. Для изображений наилучшие результаты по точности показал персептрон со значительным выигрышем по времени. Таким образом, на текущий момент можно сделать вывод, что в задачах выявления аномалий в промышленных системах стохастически конфигурируемые нейросети дают выигрыш по точности, однако требуют значительных временных затрат на обучение. Для выбора модели для обучения в федеративном режиме требуется проведение экспериментов с другими архитектурами.

СПИСОК ЛИТЕРАТУРЫ

- [1] T.-Y. Kwok and D.-Y. Yeung, "Objective functions for training new hidden units in constructive neural networks," *IEEE Trans. Neural Netw.*, vol. 8, no. 5, pp. 1131–1148, Sep. 1997
- [2] D. Wang, "Editorial: Randomized algorithms for training neural networks," *Inf. Sci.*, vols. 364–365, pp. 126–128, Oct. 2016
- [3] S. Scardapane and D. Wang, "Randomness in neural networks: An overview," *WIREs Data Min. Knowl. Disc.*, vol. 7, no. 2, 2017, Art. no. e1200, doi: 10.1002/widm.1200
- [4] M. Li and D. Wang, "Insights into randomized algorithms for neural networks: Practical issues and common pitfalls," *Inf. Sci.*, vols. 382–383, pp. 170–178, Mar. 2017
- [5] D. Wang and M. Li, "Stochastic Configuration Networks: Fundamentals and Algorithms," in *IEEE Transactions on Cybernetics*, vol. 47, no. 10, pp. 3466–3479, Oct. 2017, doi: 10.1109/TCYB.2017.27340