

1. Моделирование фотонного пучка в сферическом микрорезонаторе / Ю.П. Ракович [и др.] // Вестник Брестского государственного технического университета – 2010. – № 5: Физика, математика, информатика. – С. 82-85.

2. Русаков, К.И. Фотонные пучки в сферических микрорезонаторах / К.И. Русаков [и др.] // Труды III Международной научно-практической конференции «Оптика неоднородных структур – 2011», Могилев, 2011. – С. 71-73.

3. Rusakov, K.I. Simulation of photonic jets in the microcylinders / K.I. Rusakov [et al.] // Актуальные научные проблемы теоретической и экспериментальной физики, астрономии и космонавтики: сборник материалов межвузовской научной конференции, посвященной 50-летию первого полета человека в космос. – Брест, 2011. – С. 56-60.

4. Русаков, К.И. Экспрессный метод определения функции передачи модуляции оптических систем / К.И. Русаков, М. Пфедфер // Научные проблемы современной физики: сборник материалов Республиканской научной конференции, посвященной 90-летию со дня рождения академика Н.А. Борисевича. – Брест, 2013. – С. 164-166.

5. Rusakov, K.I. Microresonators with whispering gallery modes / K.I. Rusakov [et al.] // Квантовая электроника: материалы IX Международной научно-технической конференции. – Минск, 2013. – С. 5.

6. Savateeva, D. Optical properties of Whispering Gallery Modes spherical microcavities with J-nanoaggregates and plasmonic hot spots / D. Savateeva [et al.] // In: Physics, Chemistry and Application of Nanostructures. (Nanomeeting 2013). – Eds. V.E. Borisenko, S.V. Gaponenko, V.S. Gurin, C.H. Kam (World Scientific, Singapore, 2013). – P. 70-73.

7. Melnikau, D. Optical resonators with whispering gallery modes and J-aggregates / D. Melnikau, D. Savateeva, K. I. Rusakov, Y. P. Rakovich // Proc. 15th International Conference on Transparent Optical Networks IEEE -2013. (in print). – P. 1-4.

8. Ferrand, P. Direct imaging of photonic nanojets / P. Ferrand [et al.] // Opt. Express. – 2008. – Vol.16, no.10. – P. 6930-6940.

УДК 004.89

Вышинская Н.В.

Научный руководитель: профессор Головки В.А.

ИССЛЕДОВАНИЕ АЛГОРИТМОВ ОБУЧЕНИЯ МНОГОСЛОЙНОГО АВТОЭНКODЕРА НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ ГЛУБОКОГО ДОВЕРИЯ

Введение

В настоящее время возникает необходимость анализировать огромное количество данных, например, получаемых со спутников изображения в численной модели производства, данных, получаемых от метеорологических и океанологических станций. Использование анализа главных компонент и связанных с ним методов подразумевает потенциальное упрощение анализа наборов данных [1].

Многие природные явления ведут себя нелинейно. Это означает, что полученные данные описывают кривую или изогнутое подпространство в исходном пространстве данных. Выявление такого нелинейного многообразия играет все более важную роль в различных областях науки, поэтому возникает необходимость анализировать огромное количество данных, описывающих нелинейные процессы. В настоящее время перспективным направлением в этой области являются нейронные сети глубокого доверия [2-9].

Цель настоящей работы – исследование нелинейного метода главных компонент, который базируется на нейронной сети глубокого доверия (deep belief autoencoder), а так же поиск наиболее оптимального алгоритма обучения многослойного автоэнкодера.

1. Архитектура многослойного автоэнкодера

Многослойный автоэнкодер – многослойный персептрон [10] с одинаковым количеством нейронов на входном и выходном слоях, где скрытый слой в каждой подсети служит видимым слоем для следующего слоя. Общая структура такой сети представлена на рис. 1.

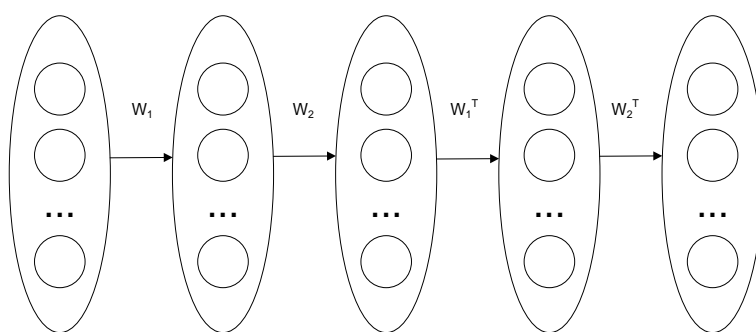


Рисунок 1 – Пример архитектуры автоэнкодера

Такие сети характеризуются как прямым, так и обратным преобразованием информации. Задачей такого преобразования является достижение наилучшего автопрогноза или самовоспроизводимости входного вектора при его отображении в пространство меньшей размерности. Таким образом, существует задача обучить сеть так, чтобы восстановление данных происходило с минимальными потерями, т.е. $|x - x'| \rightarrow 0$, где x и x' – входной и выходной вектор соответственно.

Одна из архитектур сети, используемая при исследовании алгоритмов обучения автоэнкодера, (рис. 2) представляет собой 5 слоев нейронных элементов: 1 входной слой, 3 скрытых слоя и 1 выходной слой.

- 1 слой – распределительный слой;
- 2 слой – слой кодирования;
- 3 слой – “узкий слой” (состоит из одного нейрона);
- 4 слой – слой декодирования;
- 5 слой – восстанавливающий слой.

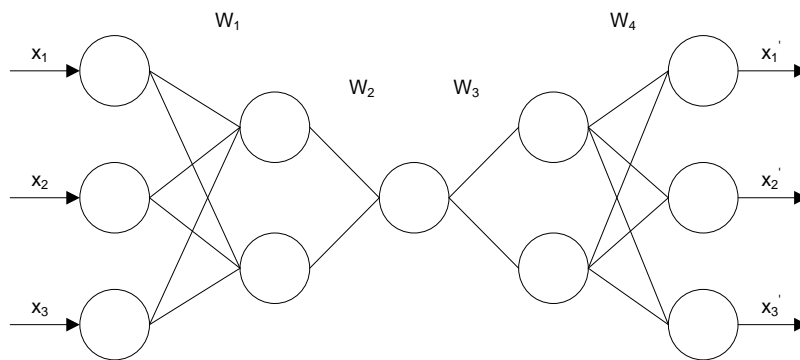


Рисунок 2 – Пример пятислойной сети

Нелинейная функция отображает из большей размерности входного пространства в пространство нижней размерности (узкое пространство), затем происходит обратное преобразование, т.е. узкое пространство возвращается к первоначальному пространству, представленному на выходах, которое должно быть как можно ближе к входному пространству, насколько это возможно путем минимизации ошибки.

Архитектура данной сети состоит из трех нейронов на входе и на выходе, первый и третий скрытые слои содержат по два нейрона, и, наконец, третий состоит из одного нейрона.

При обучении послойным алгоритмом будут использоваться архитектуры нейронных сетей, представленные на рисунках 3 и 4.

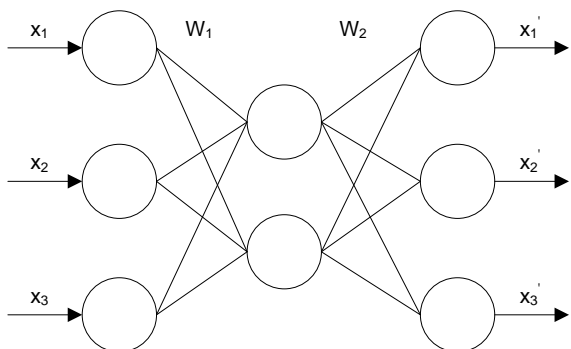


Рисунок 3 – Сеть с двумя нейронами в скрытом слое и тремя на входе

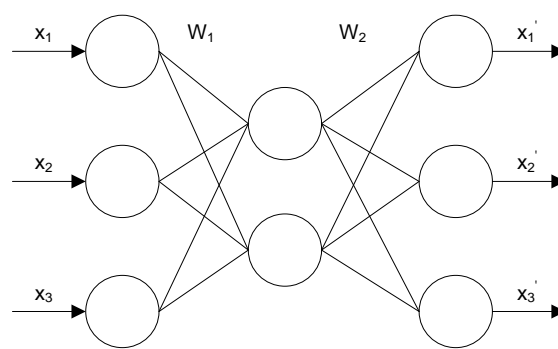


Рисунок 4 – Сеть с одним нейроном в скрытом слое и двумя на входе

2. Алгоритмы обучения

В данной работе произведено исследование различных алгоритмов обучения: алгоритм обратного распространения ошибки, послойное обучение сети, по методу обратного распространения ошибки и методу ограниченной машины Больцмана [11].

Алгоритм обратного распространения ошибки минимизирует среднеквадратичную ошибку нейронной сети. Для этого с целью настройки синоптических связей используется метод градиентного спуска в пространстве весовых коэффициентов и порогов нейронной сети. Согласно методу градиентного спуска, для рассматриваемой нами сети, изменение весовых коэффициентов нейронной сети происходит по следующему правилу:

$$w_{ij}(t + 1) = w_{ij}(t) - \alpha \frac{\partial E}{\partial w_{ij}}$$

где E – среднеквадратичная ошибка нейронной сети для одного образа.

Для выходного слоя модификация весовых коэффициентов производится согласно следующим выражениям:

$$\frac{\partial E}{\partial w'_{ij}} = \frac{\partial E}{\partial x'_i} \frac{\partial x'_i}{\partial S_i} \frac{\partial S_i}{\partial w'_{ij}} = (x'_i - x_i) F'(S_i) y_j$$

$$w'_{ij}(t+1) = w'_{ij}(t) - \alpha F'(S_i) y_j (x'_i - x_i) [10]$$

Алгоритм ограниченной машины Больцмана представляет собой незначительную модификацию стандартного алгоритма обратного распространения ошибки.

В процессе обучения сети для каждого входного образа производятся два цикла распространения информации. После этого производится настройка весовых коэффициентов сети. Пусть $x_i(0)$ – входной вектор, поступающий на вход сети в начальный момент времени. Тогда выходной вектор сети в момент $t = 1$ определяется в результате прямого преобразования информации:

$$y_j(1) = \sum_{i=1}^n w_{ij} x_i(0) + T_j, \text{ где } j = 1, p.$$

$$x_i(1) = \sum_{j=1}^p w_{ji} y_j(1) + T_i, \text{ где } i = 1, n.$$

На втором этапе распространения информации определяется вектор $y(2)$, при подаче на вход вектора $x(1)$:

$$y_j(2) = \sum_{i=1}^n w_{ij} x_i(1) + T_j, \text{ где } j = 1, p.$$

Тогда ошибка при подаче одного входного образа определяется по следующей формуле:

$$E = \frac{1}{2} \sum_{i=1}^n (x_i^k(1) - x_i^k(0))^2 + \sum_{j=1}^p (y_j^k(2) - y_j^k(1))^2$$

Дифференцируя по w_{ij}, T_i и T_j , можно получить следующее правило модификации весовых коэффициентов, для исследуемой сети.

$$w_{ij}(t+1) = w_{ij}(t) - \alpha \left((y_j(1) - y_j(2)) F'(S_j(1)) x_i(1) + (x_i(1) - x_i(0)) F'(S_i(1)) y_j(0) \right)$$

Данные правила, мы можем использовать при обучении исследуемой сети для любых данных: бинарных и вещественных.

Предположим, что $F'(S_j(1)) = \frac{\partial y_j(1)}{\partial S_j(1)} = 1$ и $F'(S_i(1)) = \frac{\partial y_i(1)}{\partial S_i(1)} = 1$, то правила обучения можно привести к следующему виду

$$w_{ij}(t+1) = w_{ij}(t) - \alpha (y_j(1) x_i(0) - y_j(2) x_i(1)),$$

$$T_j(t+1) = T_j(t) - \alpha (y_j(1) - y_j(2)),$$

$$T_i(t+1) = T_i(t) - \alpha (x_i(0) - x_i(1)).$$

3. Эксперименты

В ходе исследования данной предметной области было разработано программное обеспечение и произведено исследование и сравнительный анализ описанных выше алгоритмов обучения многослойного автоэнкодера, критерием сравнения были количество эпох затраченных на обучение, а также достижение желаемой среднеквадратической ошибки. Также данные методы были испытаны на различных видах архитектур многослойного автоэнкодера.

Параметры конфигурации сети:

- ✓ Функция активации – гиперболической тангенс (диапазон значений лежит в промежутке [-1; 1]);
- ✓ Скорость обучения – 0.001;
- ✓ Максимальное количество эпох – 100 000;
- ✓ Минимальная ошибка – 0.01;
- ✓ Минимальное изменение ошибки – 0.000 000 001.

Для исследования была сгенерирована совокупность точек, описанная следующими уравнениями:

$$x_1 = t - 0,3t^2, \quad x_2 = t + 0,3t^3, \quad x_3 = t^2.,$$

которые затем претерпели правила нормализации по формуле:

$$\bar{x}_i = \frac{(x_i - x_{\min})(b - a)}{(x_{\max} - x_{\min})} + a,$$

где [a, b] – диапазон значений в соответствии с функцией активации.

Данный набор точек x_1 , x_2 и x_3 является входными значениями для исследуемых сетей.

На рисунках 5-7 представлено графическое представление точек в двумерном пространстве, т.е. зависимости функций x_2 от x_1 , x_3 от x_2 и x_3 от x_1 соответственно. На рисунке 8 представлена трехмерная зависимость данных точек.

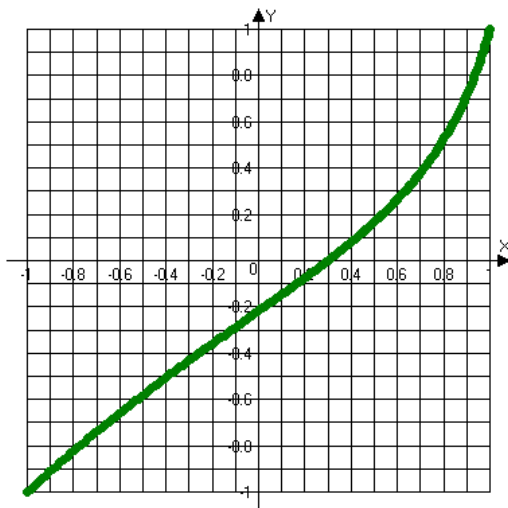


Рисунок 5 – Зависимость x_2 от x_1

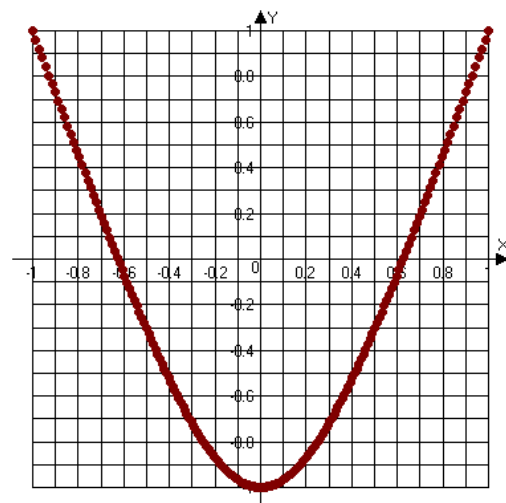


Рисунок 6 – Зависимость x_3 от x_2

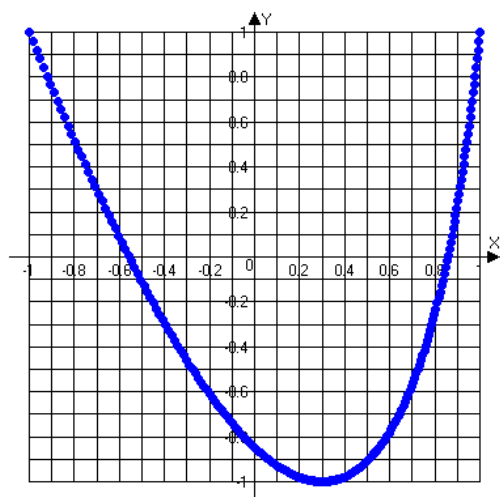


Рисунок 7 – Зависимость x_3 от x_1

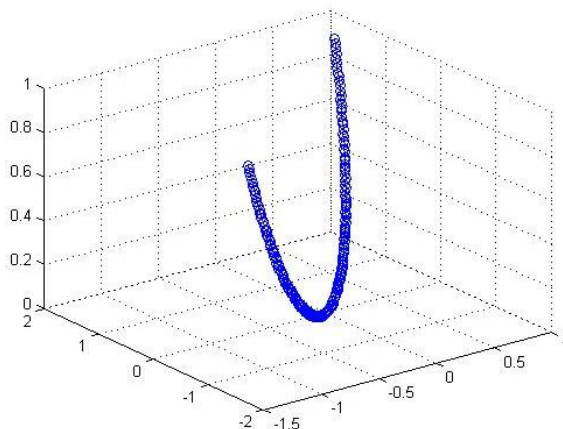


Рисунок 8 – Функция в 3D

Сравнительный анализ алгоритмов обучения на описанной выше архитектуре сети, представленной на рисунке 2, приведен в таблице 1.

Таблица 1 – Результаты исследования

№	Метод обучения	Количество итераций	Ошибка	Обучающая выборка
1	Back propagation	16 370	0.009968	100 точек
2	Послойный, back propagation	54 288	0.698235	100 точек
3	Послойный, restricted Boltzmann machine	20 003	0.050789	100 точек
4	Back propagation	600 000	4.567530	200 точек
5	Послойный, back propagation	115 058	17.44077	200 точек
6	Послойный, restricted Boltzmann machine	50 000	0.715596	200 точек

Самый лучший результат показал послойный алгоритм обучения – ограниченная машина Больцмана. Результаты, выдаваемые сетью, показаны ниже на рисунках 9-11, в качестве эксперимента исходные данные были предварительно зашумлены.

Все, исследованные в данной работе методы обучения показали, что вполне могут быть использованы при сжатии данных или визуализации информации, все зависит от размерности данных и архитектуры сети.

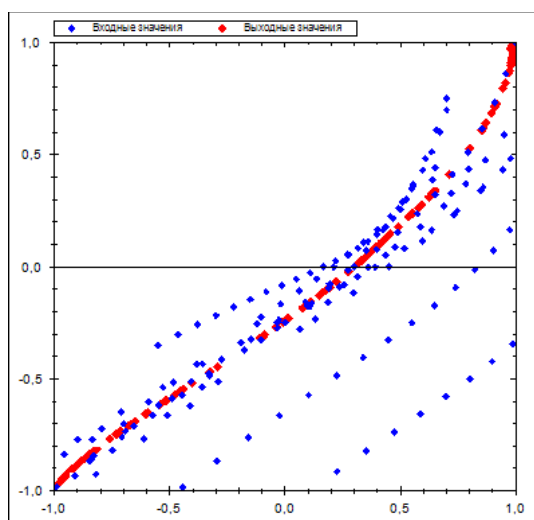


Рисунок 9 – Зависимость x_2 от x_1

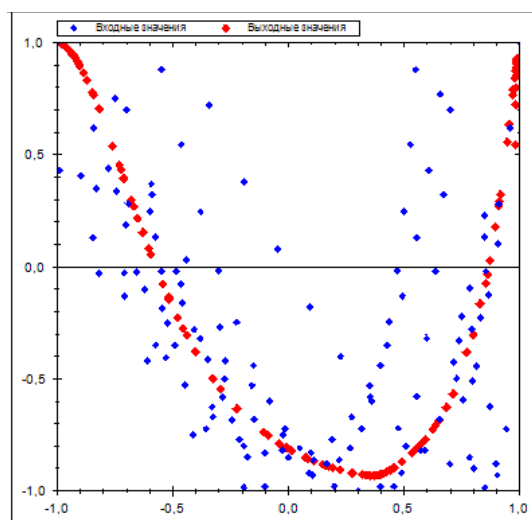


Рисунок 10 – Зависимость x_3 от x_1

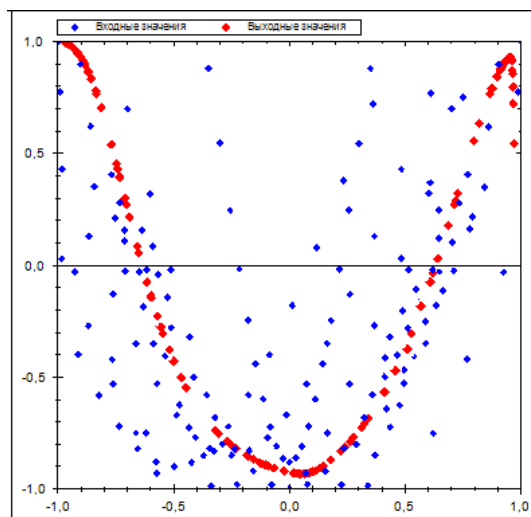


Рисунок 11 – Зависимость x_3 от x_1

Список цитированных источников

1. Scholz, M., Fraunholz, M., Selbig, J. Nonlinear principal component analysis: neural network models and applications. In book *Principal manifolds for data visualization and dimension reduction*. Springer Berlin Heidelberg, 44-67 (2008).
2. Hinton, G.E., Osindero, S., Teh, Y. A fast learning algorithm for deep belief nets. *Neural Computation*, 18, 1527-1554 (2006)
3. Hinton, G. Training products of experts by minimizing contrastive divergence. *Neural Computation*, 14, 1771-1800 (2002).
4. Hinton, G., Salakhutdinov, R. Reducing the dimensionality of data with neural networks. *Science*, 313 (5786), 504-507 (2006).
5. Hinton, G.E. A practical guide to training restricted Boltzmann machines. (Tech. Rep. 2010-000). Toronto: Machine Learning Group, University of Toronto (2010)
6. Bengio, Y. Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, 2(1), 1-127 (2009).
7. Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H. Greedy layer-wise training of deep networks. In B. Schölkopf, J. C. Platt, T. Hoffman (Eds.), *Advances in neural information processing systems*, 11, pp. 153-160. MA: MIT Press, Cambridge (2007).
8. Erhan, D., Bengio, Y., Courville, A., Manzagol, P.-A., Vincent, P., Bengio, S. Why does unsupervised pre-training help deep learning? *Journal of Machine Learning Research*, 11:625-660 (2010).
9. Golovko, V., Vaitsekhovich, H., Apanel, E., Mastykin, A. Neural network model for transient ischemic attacks diagnostics. *Optical Memory and Neural Networks (Information Optics)*, 21(3), 166-176 (2012).
10. Головкин, В.А. Нейронные сети: обучение, организация и применение: кн. 4: учеб. пособие для вузов / Общая ред. А.И. Галушкина. – М.: ИПРЖР, 2001. – 256 с.
11. Golovko V., Kroshchanka F., Rubanau U., Jankowski S. A Learning Technique for Deep Belief Neural Networks. *Lecture Notes in Computer Science*, Springer, 2014 (in press).

УДК 519.876:544.77.022:691.32

Дивинец А.А.

Научные руководители: к.т.н. Разумейчик В.С., к.т.н., доцент Дереченник С.С.

**ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ТРАНЗИТНОЙ ЗОНЫ
В ГЕТЕРОГЕННЫХ СИСТЕМАХ**

Гетерогенная система – это система, состоящая из двух или большего числа фаз с сильно развитой поверхностью раздела между ними. Частным случаем гетерогенного объекта является дисперсная система, в которой одна из фаз образует непрерывную дисперсионную среду, в объеме которой распределена дисперсная фаза (или несколько