

ПРЕИМУЩЕСТВА ПРИМЕНЕНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ В АНАЛИЗЕ БИЗНЕС-ИНФОРМАЦИИ И ПРИЛОЖЕНИЯХ DATA-MINING

Анотація. Публікація присвячена доцільності застосування штучних нейронних мереж в аналізі, прогнозуванні та підтримці прийняття управлінських рішень. Розглянуто переваги штучних нейронних мереж для додатків добування даних (data-mining).

Ключові слова: штучні нейронні мережі, підтримка прийняття рішень, прогнозування, вилучення даних.

Summary. The publication is dedicated to the usefulness of artificial neural networks in the analysis, forecasting, and support decision making. Considered arguments about the benefits of artificial neural networks for data-mining applications.

Key words: artificial neural networks, decision support system, forecasting.

Постановка проблемы. Компьютерные информационные системы, используемые на предприятиях, прошли ряд этапов в своем развитии. Вначале были наиболее востребованы системы сбора данных, затем их формализации и хранения. В настоящее время перед любым предприятием возникает потребность анализа огромных объемов бизнес-информации разного рода. Наиболее актуальной становится проблема преобразования больших массивов табличной информации, обычно хранящейся в базе данных, в лаконичный набор правил бизнес-стратегии, четкий прогноз закупок товаров или поддержку оптимального управленческого решения. На данный момент большинство источников определяют этот процесс как data-mining (процесс аналитического исследования больших массивов экономической информации с целью выявления определенных закономерностей и систематических взаимосвязей между переменными, которые затем можно применить к новым совокупностям данных). Для краткости будем считать, что data-mining — это процесс добычи необходимых знаний из структурированных данных.

Основными методами добычи знаний являются анализ гипотез и разведочный анализ данных. Эти методы хорошо себя зарекомендовали, и их можно считать классическими, но существенным недостатком их использования является ориентированность на выяснение природы процесса распределения данных, а не на конечный результат прогноза. Также стоит обратить внимание на достаточно сложный математический аппарат этих методов и невысокий параллелизм полученных на их основании алгоритмов.

Квалифицированный пользователь компьютерной информационной системы при анализе данных ожидает увидеть достаточно простые и гибкие инструменты, позволяющие быстро добавить, удалить или модифицировать правила выборки данных. Одновременно с этим пользователи предъявляют достаточно жесткие требования к автоматизму и быстрдействию процесса работы с данны-

ми. Такие достаточно противоречивые требования к функционалу информационной системы заставляют искать новые методы добычи данных.

Как альтернативу описанным выше методам data-mining (добычи знаний) можно предложить использование искусственных нейронных сетей.

Анализ последних исследований и публикаций.

Тема различных практических применений искусственных нейронных сетей достаточно широко раскрыта в отечественной и зарубежной научной литературе. Наиболее известные авторы: Mohamad H. Hassoun, А. А. Ежов, С. А. Шумский, С. Хайкин, Л. Г. Комарцова, А. В. Максимов, А. И. Галушкин, Р. Дуда, П. Харт, В. А. Головкин, Г. Э. Яхьяева, Д. А. Тархов, В. В. Круглов, М. И. Дли, Р. Ю. Голунов.

В области нейронных сетей с общей регрессией и вероятностных нейронных сетей известны Burrascano, P., Schieler, H., Hartmann, U., Caudill, M. P. Tavan, H. Grubmuller, H. Kuhnel.

Публикации, посвященные практическому применению искусственных нейронных сетей, можно найти под авторством R. C. Frye, J. Vac. Sci. Technol, K. D. Cummings, E. A. Rietman, C. M. Bishop, B. Denby, С. В. Клименко и др.

20 марта 2012 года состоялась X Всероссийская научная конференция «Нейрокомпьютеры и их применение», посвященная актуальным проблемам применения нейронных сетей и других обучаемых структур для решения задач в области естественных, технических и гуманитарных наук. Ее цель — объединение специалистов, работающих в различных сферах применения нейросетевых алгоритмов, обсуждение и обобщение опыта их теоретических и практических разработок, определение перспектив развития обучаемых структур.

Цель данной публикации состоит в обосновании применения методологии искусственных нейронных сетей в компьютерных информационных системах добычи данных (data-mining) и поддержки принятия управленческих решений.

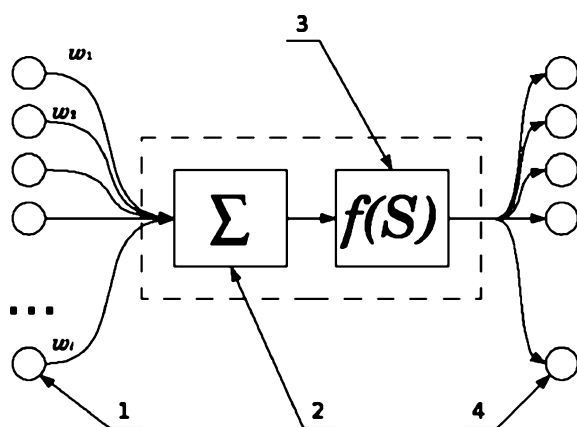
Изложение основного материала исследования.

Искусственная нейронная сеть представляет со-

© В. О. Лазарев, 2013

бой систему соединённых и взаимодействующих между собой искусственных нейронов [1; с. 9]. Математическая модель искусственного нейрона была предложена У. Маккалоком и У. Питтсом вместе с моделью сети, состоящей из этих нейронов. Авторы показали, что сеть на таких элементах может выполнять числовые и логические операции. Практически сеть была реализована Фрэнком Розенблаттом в 1958 году как компьютерная программа, а впоследствии — как электронное устройство — перцептрон.

Под искусственными нейронами подразумеваются простые процессоры, в обязанности которых входят прием сигнала, его обработка и отправка (рис. 1). Искусственный нейрон может иметь большое количество входных соединений, по которым он получает сигналы от других нейронов [2; с. 71]. Искусственный нейрон может быть представлен как некоторая функция (функция срабатывания) от одного аргумента, которым является комбинация входящих сигналов. Результат функции срабатывания передается на единственный выход искусственного нейрона, связанный со входами других нейронов сети или внешних устройств.



- 1 — нейроны, передающие сигнал;
- 2 — сумматор весов входящих сигналов;
- 3 — функция срабатывания;
- 4 — нейроны, принимающие сигнал;

Рис. 1. Схема работы искусственного нейрона

Диапазон значений функции срабатывания искусственного нейрона обычно определен как $[0,1]$ или $[-1,1]$. По виду функции срабатывания нейроны, в основном, можно разделить на 3 типа.

1. Пороговая передаточная функция (функция Хевисайда) представляет собой перепад. До тех пор, пока взвешенный сигнал на входе нейрона не достигает некоторого уровня, сигнал на выходе равен нулю. Как только сигнал на входе нейрона превышает указанный уровень, выходной сигнал скачкообразно изменяется на единицу.

$$f(x) = \begin{cases} 1 & \text{if } x \geq T \\ 0 & \text{else} \end{cases} \quad (1)$$

2. Линейная передаточная функция — сигнал на выходе нейрона линейно связан со взвешенной суммой сигналов на его входе.

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x \geq 1 \\ x & \text{else} \end{cases} \quad (2)$$

3. Сигмоидальная передаточная функция — один из самых часто используемых на данный момент типов передаточных функций. Введение функций сигмоидального типа было обусловлено ограниченностью нейронных сетей с пороговой функцией активации нейронов — при такой функции активации любой из выходов сети равен либо нулю, либо единице, что ограничивает использование сетей не в задачах классификации. Использование сигмоидальных функций позволило перейти от бинарных выходов нейрона к аналоговым.

$$\sigma(x) = \frac{1}{(1 + \exp(-tx))} \quad (3)$$

Существуют также другие виды передаточных функций, но они выходят за рамки данной публикации.

Таким образом, искусственный нейрон получает на входе сигналы одного или более нейронов, характеризующиеся меткой источника и весом, выполняет над полученным массивом сигналов функцию срабатывания и отправляет полученный результат одному или более нейронам, указывая источник и вес как результат функции срабатывания [3; с. 25]. Из вышесказанного видно, что искусственные нейронные сети имеют высочайшую степень масштабируемости и высокую гибкость конечной модели сети.

В общем случае искусственная нейронная сеть состоит из трех частей (рис. 2):

1. Слой входа — принимает данные от внешних источников и обеспечивает их фильтрацию.

2. Скрытый слой — в нем происходит основная обработка данных; скрытых слоев может быть достаточно много, это зависит от сложности решаемых сетью задач.

3. Слой выхода — содержит в себе результат обработки данных в искусственной нейронной сети, служит для конечного представления данных и передачи во внешние устройства.

Для выполнения своих функций искусственные нейронные сети обучаются [4; с. 64].

Обучение сети заключается в подстройке весов передаваемых искусственным нейронам сигналов. Процесс обучения сети может быть реализован двумя основными методами: обучение с учителем и обучение без учителя. В случае обучения искусственной нейронной сети с учителем

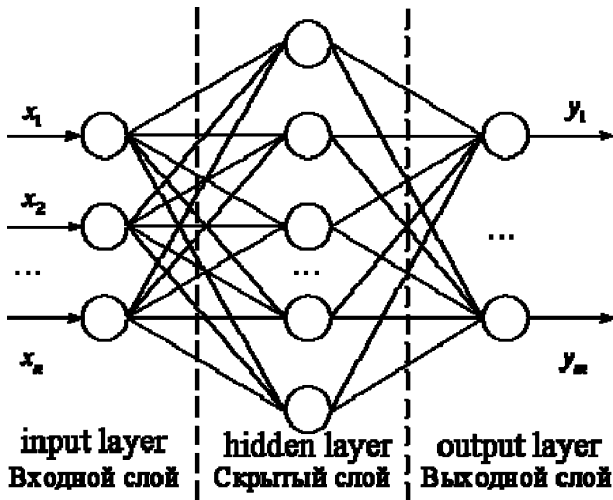


Рис. 2. Схема искусственной нейронной сети

создается специальная обучающая выборка, подающаяся на вход сети, после чего сеть настраивается таким образом, чтобы появились ожидаемые результаты на выходе.

Определив основные концепции искусственных нейронных сетей, перейдем к использованию их методологии в приложениях бизнес-аналитики, экономического прогнозирования, поддержки принятия управленческих решений и добычи данных (data-mining) [5; с. 89].

Одним из главных преимуществ использования искусственных нейронных сетей для приложений data-mining является принципиальное изменение поведения компьютерной информационной системы. В классических примерах систем на первом этапе производится ввод данных в хранилище (например базу данных), а затем производится преобразование, отсев, кластеризация и анализ данных. Эти операции обычно используют многочисленные операции считывания и записи данных с носителя, что часто снижает производительность и увеличивает износ оборудования. Применение искусственной нейронной сети на первых стадиях работы с данными позволяет решить вопросы фильтрации, кластеризации, классификации и ранжирования входящих данных в рамках единого процесса, который идеально использует параллелизм, что в наше время является одним из требований, налагаемым развитием аппаратных средств. В случае применения искусственных нейронных сетей информация, попадающая в базу данных, является более готовой к последующему использованию, снижается доля «информационного мусора» и бесполезных данных [6; с. 47].

Таким образом, проведение данных «сквозь строй» искусственной нейронной сети, при правильном выборе топологии и соответствующей настройке, позволяет уменьшить объем информации за счет более точного отсева, выполнить функции кластеризации буквально при вводе данных, уве-

личить производительность компьютерной информационной системы за счет использования параллельных вычислений и снизить износ аппаратного обеспечения устройств хранения данных.

Основными инструментами приложений data-mining являются кластеризация, классификация и ранжирование. При использовании этих приложений для решения бизнес-задач в современной модели конкурентной рыночной экономики неизбежны частые изменения величин критериев анализа данных, а нередко и появление новых критериев, с постепенным отмиранием старых. В классических компьютерных информационных системах зачастую достаточно узко заданы диапазоны изменения критериев, а возможностью создания новых могут похвастаться уже совсем не многие программы. Часто возникает ситуация, когда пользователи сталкиваются с необходимостью использовать внешние программы (например электронные таблицы) для более гибкого обращения с данными, но это влечет за собой неприятности экспорта и отсутствие единой логики.

Применение искусственных нейронных сетей в приложениях data-mining при правильной организации дает своеобразный конструктор новых критериев для инструментов извлечения данных [7; с. 15]. Практическая реализация данной функциональности сильно зависит от топологии сети. Например, используя многослойный персептрон получения новых критериев, можно добиться введения дополнительного скрытого слоя нейронов, в сети Word можно использовать дополнительные блоки нейронов. Другими словами, при разработке приложений data-mining, используя искусственные нейронные сети, целесообразно вводить инструменты, позволяющие экспертам и продвинутым пользователям в известных границах модернизировать архитектуру сети, реагируя на изменяющиеся бизнес-требования [8; с. 65]. Такой подход позволит продлить жизненный цикл разрабатываемых приложений. Другим положительным эффектом является создание вокруг подобного приложения сообщества конфигураторов и внедренцев, получивших в свое распоряжение эффективную платформу data-mining; действие этого эффекта можно проиллюстрировать, например, успешностью программного продукта 1С: Предприятие и других подобных.

Выводы. Технология достаточно часто использует решения, подсмотренные в природе. Искусственные нейронные сети не являются исключением, они копируют структуру нервных клеток человека. Несмотря на то, что идея искусственных нейронных сетей возникла достаточно давно, практические применения данного подхода раскрывают свои возможности только сейчас. Развитие процессоров в направлении увеличения количества ядер (исполнительных блоков) определяет выбор методов и алгоритмов с высоким параллеле-

лизмом вычислений. С этой точки зрения искусственные нейронные сети идеальны, так как они представляют собой массивы достаточно простых процессов (искусственных нейронов), выполняющихся одновременно (для каждого слоя сети).

Программы, использующие искусственные нейронные сети, обладают высокой степенью модульности, что является необходимым условием любых современных приложений. Широкая масштабируемость обуславливается тем, что для ввода новых критериев и условий обработки данных достаточно добавить новый скрытый слой искусственных нейронов, причем явных ограничений на количество этих слоев не существует. Также можно в широком диапазоне изменять функциональность искусственной нейронной сети, переучивая ее, то есть создавая новые обучающие выборки.

И наконец, использование искусственных нейронных сетей хорошо уживается с объектно-ориентированным программированием [9; с. 31], являющимся современным стандартом разработки. Каждый тип искусственного нейрона, использование которого планируется в данной сети, можно определить соответствующим классом, в таком случае конкретные нейроны станут экземплярами класса, а синапсы — интерфейсом класса.

Все сказанное выше позволяет сделать выводы о высокой целесообразности использования искусственных нейронных сетей в приложениях, связанных с обработкой больших массивов бизнес-информации.

Литература

1. Hassoun Mohamad H. Fundamentals of Artificial Neural Networks / Mohamad H. Hassoun. — Cambridge, Massachusetts : MIT Press, 1995. — 354 с.
2. Ежов А. А. Нейрокомпьютинг и его применения в экономике и бизнесе / А. А. Ежов, С. А. Шумский. — М. : МИФИ, 1998. — 460 с.
3. Хайкин С. Нейронные сети: полный курс / С. Хайкин. — [2-е изд.]. — М. : Вильямс, 2006. — 600 с.
4. Комарцова Л. Г. Нейрокомпьютеры / Л. Г. Комарцова, А. В. Максимов. — М. : Изд-во МГТУ им. Баумана, 2004. — 250 с.
5. Галушкин А. И. Нейронные сети. Основы теории / А. И. Галушкин. — М. : Горячая линия — Телеком, 2010. — 320 с.
6. Kosko B. Neural Networks and Fuzzy Systems. A Dynamical Systems Approach to Machine Intelligence / B. Kosko. — Prentice Hall : Englewood Cliffs, 1992. — 400 с.
7. Головкин В. А. Нейронные сети: обучение, организация и применение / В. А. Головкин. — М. : ИПРЖР, 2001. — 280 с.
8. Тархов Д. А. Нейронные сети. Модели и алгоритмы. (Справочник) / Д. А. Тархов. — М. : Радиотехника, 2005. — 370 с.
9. Круглов В. В. Нечёткая логика и искусственные нейронные сети / В. В. Круглов, М. И. Дли, Р. Ю. Голунов. — М. : Физматлит, 2001. — 415 с.