

КОРРЕКТНОСТЬ ПРИМЕНЕНИЯ НЕЙРОСЕТЕВОГО МОДЕЛИРОВАНИЯ В ЗАДАЧАХ КЛАССИФИКАЦИИ

Аннотация. В статье приводятся условия корректного применения нейросетевого моделирования в задачах классификации на примере разработки модуля классификации по принципу “свой/чужой” контрольно-пропускного пункта (КПП).

Ключевые слова: искусственные нейронные сети, нейросетевое моделирование, классификация, распознавание.

Введение

В настоящее время искусственные нейронные сети (ИНС) нашли широкое применение в различных сферах человеческой деятельности. Потенциальными областями применения ИНС являются те, где человеческий интеллект малоэффективен, а традиционные вычисления трудоемки или вовсе невозможны. Действительно, актуальность применения нейронных сетей многократно возрастает тогда, когда появляется необходимость решения плохо формализованных задач.

Одной из широко распространенных задач нейронной сети является распознавание образов. Она заключается в том, чтобы классифицировать входной образ, то есть отнести его к какому-либо известному сети классу. Классы объектов могут быть заранее выделены (задача классификации) или необходимо их выявить в процессе решения задачи (задача кластеризации). Под объектами в распознавании образов понимают: различные предметы и явления, процессы и ситуации, сигналы и т.п.

В настоящее время направление распознавание образов приобрело огромную популярность среди исследователей. Это связано с тем, что решение задач в рассматриваемой области, находит все больше прикладное применение в повседневной деятельности человека. В связи с усилением борьбы с

преступностью и терроризмом, подстегивающей рост индустрии безопасности, биометрическая идентификация личности – одно из самых перспективных и бурно развивающихся направлений. Это можно увидеть, наблюдая за новостями на различных IT порталах.

В Санкт-Петербурге стартовал пилотный проект по развертыванию в метро интеллектуальной системы видеонаблюдения, способной идентифицировать пассажиров. Система позволяет фиксировать фотоизображение каждого входящего в метро человека, сравнить его с базой данных интересных правоохранным органам персон и сформировать необходимый сигнал правохранителям, а также турникету, чтобы он не впускал в метро подозрительного пассажира. Также информация о его перемещениях становится доступна правоохранным органам. Тестовый продукт прошел успешно, и систему планировалось внедрить на 19 станциях петербургского метро. Однако, опробованная технология не получила дальнейшего распространения. Такая технология могла бы вовремя обнаружить, а возможно и просто не пустить в метро, самого террориста [1].

19 января 2017 г. Российский банковский сектор приступил к внедрению систем биометрической идентификации клиентов — по голосу, изображению лица, сетчатке глаза и отпечаткам пальцев. Также ВТБ24 готовится к запуску проекта голосовой идентификации клиентов, звонящих в call-центр. В ближайшем будущем биометрические технологии будут востребованы в кредитно-финансовой сфере (осуществление платежей, выдача кредитов, доступ к банковским ячейкам, удаленная идентификация и аутентификация клиентов), в сфере здравоохранения (идентификация пациентов и врачей в телемедицине, а также их удаленная идентификация), в сфере мобильных устройств (доступ к устройствам, голосовое управление устройствами). Также следует отметить активное применение в России биометрических технологий для обеспечения безопасности при проведении спортивно-массовых мероприятий, в транспортной безопасности, в криминалистике и в паспортно-визовой системе [2].

Определение ИНС

ИНС представляют собой системы искусственного интеллекта, позволяющие решать задачи прогнозирования, классификации, управления, выявлять зависимости и взаимосвязи, которые невозможно определить стандартными статистическими методами, например, корреляционным анализом или анализом таблиц сопряженности. Основой ИНС является искусственный нейрон.

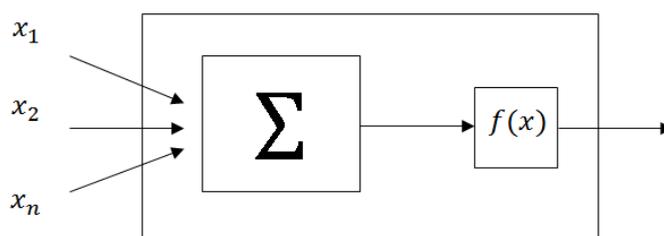


Рис. 1. Структура искусственного нейрона.

Таким образом, искусственный нейрон представляет собой упрощённую модель естественного нейрона [3]. На вход нейрона поступает множество сигналов, обозначаемое вектором. Эти сигналы соответствуют сигналам, приходящим на вход синапсов биологического нейрона. Каждый вход умножается на соответствующий "вес" – аналог силы синаптической связи.

Все полученные произведения суммируются в специальной структуре – сумматоре. Сумматор соответствует телу биологического элемента, а полученная сумма – уровню активации нейрона. Полученная сумма, как правило, преобразуется с помощью передаточной (активационной) функции. В качестве передаточной функции, например, может использоваться сигмоидальная, линейная, пороговая, радиально-базисная функции [4]:

Название функции	Передаточная функция	График передаточной функции
------------------	----------------------	-----------------------------

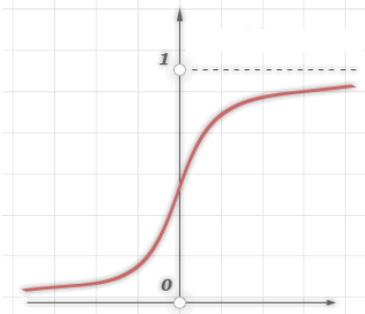
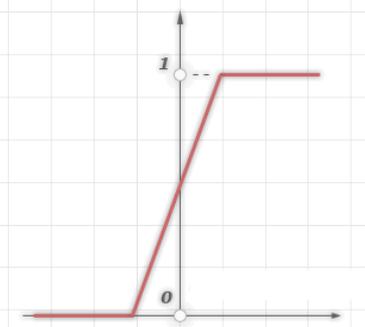
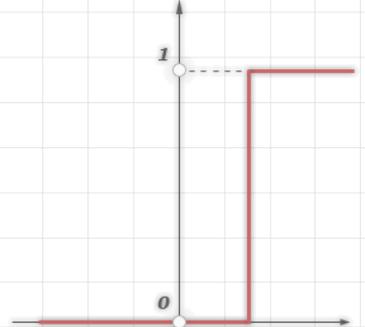
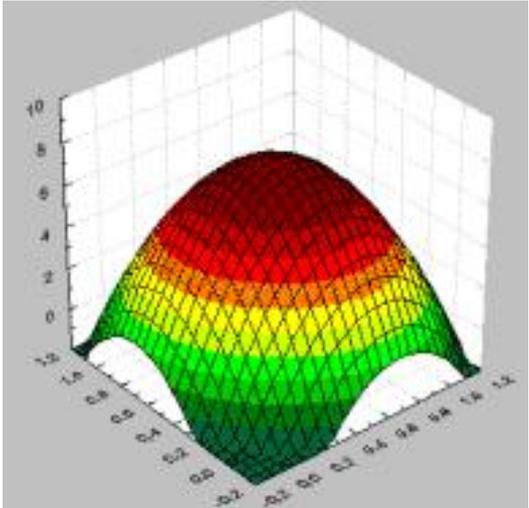
Сигмоидальная функция	$\sigma(x) = \frac{1}{1 + e^{-x}}$	
Линейная функция	$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$	
Пороговая функция	$f(x) = \begin{cases} 1 & \text{if } x \geq T \\ 0 & \text{else} \end{cases}$ <p>T – сдвиг функции активации относительно горизонтальной оси</p>	
Радиально-базисная функция	$y = \exp\left(-\frac{(S-R)^2}{2\sigma^2}\right)$, здесь $S = X - C $ – расстояние между центром C и вектором входных сигналов X . Скалярный параметр σ – ширина окна, R определяет сдвиг активационной функции по оси абсцисс.	

Таблица 1. Передаточные (активационные) функции, x – результат работы сумматора.

Активационная функция является своеобразной нелинейной усилительной характеристикой. Результат работы передаточной функции выступает в качестве выходного значения ИНС.

Отличительной чертой ИНС является их обучаемость. Обучение заключается в корректировке весов в соответствии с множеством входов и ожидаемым выходным значением [3]. Необходимое число наблюдений для обучения нейронной сети в первую очередь зависит от числа связей и от заранее неизвестной сложности отображения. Однако известно, что с ростом количества входных переменных количество требуемых наблюдений растёт нелинейно. Эта зависимость носит название «проклятие размерности» [5].

Обучение может осуществляться как с учителем, так и без него. Обучение с учителем предполагает наличие целевого вектора, в соответствии с которым происходит корректировка весов. А обучение без учителя основывается на выявлении статистических свойств обучающего множества и группировке сходных векторов в классы [3], что позволяет каждый раз получать одинаковые выходные значения при предъявлении достаточно близких входных данных из определённого класса. Таким образом, процесс обучения позволяет определить, какой выходной вектор будет производиться некоторым классом входных данных.

Корректность применения ИНС в задачах классификации

ИНС позволяют решать задачи классификации и кластеризации, аппроксимации, управления, прогнозирования. Задача классификации подразумевает отнесение распознаваемого статического образа «к одному из нескольких попарно непересекающихся определённых множеств» [6]. Например, с классификацией связана задача распознавания диагноза, рукописного текста, стратегических объектов на фото, управления («пропустить/не пропустить», «да/нет», «активировать/не активировать») и пр.

На вход нейронной сети подаются параметры, характеризующие образец, а на выходе ИНС определены конкретные заранее известные классы, например, класс «активировать» и класс «не активировать».

Мода на использование ИНС в последнее время привела к тому, что применение нейросетевого моделирования не всегда бывает уместно и адекватно. В связи с этим перечислим условия, которыми должна обладать задача, чтобы было оправдано её решение с помощью ИНС:

- 1) задача должна быть плохо формализована или должна плохо поддаваться формализации;
- 2) задача предполагает предоставление на вход ИНС большого числа «неинформативных, избыточных, шумовых входных сигналов» [7];
- 3) задача связана с обработкой разнородных данных, измеренной в разных шкалах;
- 4) задача предполагает, что в её контексте найдены все линейные зависимости стандартными статистическими методами (регрессионным или линейным анализом, корреляционным анализом, анализом таблиц сопряжённости и пр.), но остаются подозрения, что существуют также и нелинейные зависимости, описываемые сложными поверхностями, которые сможет обнаружить ИНС;
- 5) требуется параллельное решение сразу нескольких подзадач. На основании результатов работы ИНС можно сформировать целостную картину, весь контекст, характеризующий задачу;
- 6) требуется выполнение функций эксперта при отсутствии данного эксперта. ИНС обучается, учится «мыслить» как эксперт и может служить основой для создания советующих или экспертных систем;
- 7) для решения задачи требуется использование относительно устойчивой структуры, способной предоставить хотя бы приближённое решение.

Пример применения ИНС в задачах классификации «свой/чужой» на примере КПП

Практически все современные системы контроля и управления доступом, используемые на российских предприятиях, регистрируют сам факт прохода человека, владеющего электронным пропуском, без идентификации личности

самого человека. Таким образом, по пропуску может пройти любой человек, что серьезно ставит под угрозу безопасность.

Проблема состоит в том, что на данный момент нет универсального набора алгоритмов качественно решающих задачу определения личности. Связано это с тем, что при съёмке камерой видеорегистрации, поступающий видеопоток подвергается различным искажениям, например, перспективным преобразованиям, связанным с перемещением камеры либо объектов в кадре, изменениям освещённости и зашумлением. Также меняется сам человек, его причёска, одежда, аксессуары.

Для решения данного вопроса предлагается создать программный модуль, идентифицирующий личность человека на КПП, используя видеопоток от камеры видеорегистрации.

Задача классификации пользователей по принципу «свой - чужой» в биометрических системах – это непрерывная аутентификация пользователей. Она может решаться как верификация или как идентификация. В первом случае система использует единственный шаблон «своего» пользователя, и сравнение параметров неизвестного пользователя осуществляется только с ним по принципу 1:1. Во втором случае решается общая задача сопоставления параметров неизвестного пользователя со всеми шаблонами зарегистрированных в системе пользователей по принципу 1:M.



Рис. 2. Непрерывная аутентификация пользователей.

Для решения задачи классификации часто используют классическую полносвязную нейронную сеть прямого распространения (многослойный персептрон), для обучения которого используется алгоритм обратного распространения ошибки. Этот алгоритм заключается в распространении сигналов ошибки от выходов сети к ее входам в направлении, обратном прямому распространению сигнала в обычном режиме работы. Таким образом, осуществляется корректировка весов сети. Однако эти сети обладают такими недостатками, как: длительность обучения; подбор правильной размерности сети. Длительность обучения персептрона прямо пропорциональна числу синапсов. Если сеть полносвязная, значит выход нейрона предыдущего слоя связан с входом каждого нейрона последующего слоя. Таким образом, если имеется входной образ размерностью 1000×1 , то первый слой должен содержать 1000 нейронов. Также в настоящее время не существует универсальных алгоритмов для того чтобы вычислить минимальное необходимое число нейронов в каждом слое. При достаточно большой размерности входного образа не следует использовать персептрон, так как в нашем случае потребуется использовать скрытый слой нейронов, размерность которого будет еще выше, что уменьшит скорость обучения сети. Кроме того, слишком простые сети могут не запомнить достаточное число образов, а погрешность сложных сетей будет велика за счет эффекта «насыщения», то есть при решении сеть не сможет достигнуть глобального минимума. В связи с этим логичнее будет использовать нейронную сеть Хопфилда, которая является разновидностью упрощенной модели человеческого мозга. Механизм ассоциативной памяти достигается за счет обратных связей.

Мы предлагаем использовать многослойную сеть Хопфилда с большим количеством нейронов. Чтобы исключить влияние случайно попавших в кадр артефактов, брать не всю фотографию, а только ключевые области – глаза, брови, рот и нос. Обучающей выборкой будет база лиц, составленная из фотографий сотрудников, причем для каждого человека приводится несколько фотографий, сделанных в разных позах.

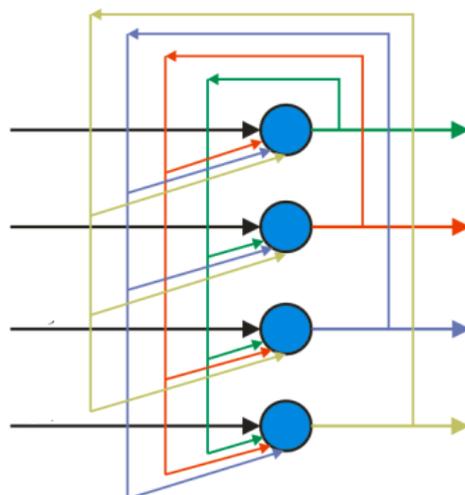


Рис. 3. Сеть Хопфилда.

Таким образом, модуль классификации по принципу «свой/чужой» будет выглядеть следующим образом.

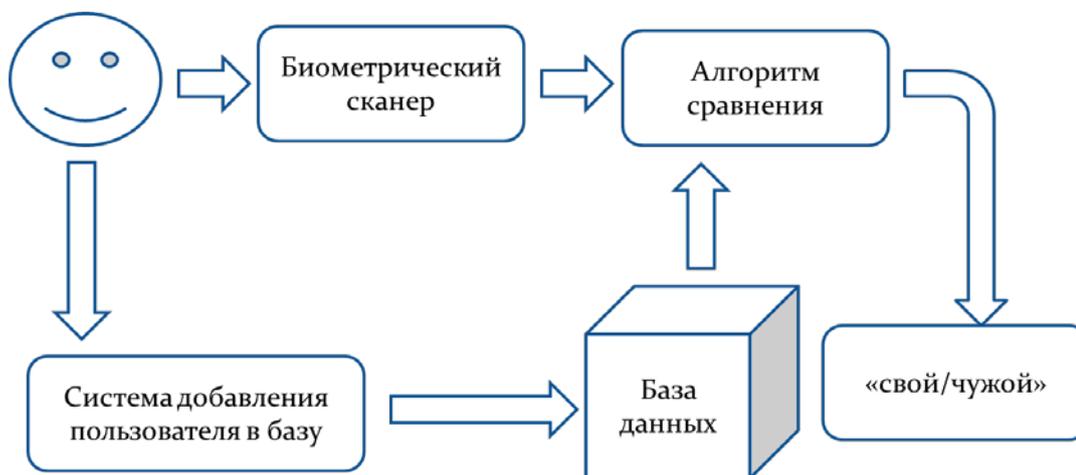


Рис. 4. Модель классификации по принципу «свой/чужой».

Планируется апробация приложения на базе ООО «ТННЦ» и ОАО МРСК Урала.

Заключение

За последние годы произошел значительный технологический рост в области проектирования нейронных сетей и нейрокомпьютеров. Открыто немало новых возможностей нейронных сетей, а работы в данной области становятся важным вкладом в науку и технологии.

Грядущий взрывной рост в области нейрокомпьютерных технологий будет связан с новыми открытиями в области нейронного моделирования – как только будет разгадана тайна функционирования хотя бы одной области мозга,

так сразу будет получено представление о работе многих других его областей. Открытие биологических основ обработки информации вызовет существенную активизацию работ в построении искусственного мозга. Появится возможность синтезировать «умные» машины и системы, способные вместо людей выполнять монотонные, скучные и опасные задания.

СПИСОК ЛИТЕРАТУРЫ

1. Грязневич, В. Современные методы борьбы с терроризмом Петербургу оказались не нужны: РБК [Электронный ресурс]. – Режим доступа: http://www.rbc.ru/spb_sz/04/04/2017/58e3409b9a79470e152260ae (дата обращения 14.04.2017);

2. Маркелов, Р. Неевклидова биометрия: как российские банки теперь распознают клиентов: РБК [Электронный ресурс]. – Режим доступа: <http://www.rbc.ru/finances/19/01/2017/587e2c689a79470b502835ba> (дата обращения 14.04.2017);

3. Яхьяева, Г.Э. Основы искусственных нейронных сетей. Материал из курса: Основы теории нейронных сетей [Электронный ресурс]. – Режим доступа: <http://cogsys.ru/sites/default/files/library/NeuroNets.pdf> (дата обращения 09.04.2017);

4. Горбачевская, Е. Н. Классификация нейронных сетей // Вестник ВУиТ. 2012. №2 [19]. – Режим доступа: <http://cyberleninka.ru/article/n/klassifikatsiya-neuronnyh-setey> (дата обращения: 09.04.2017);

5. StatSoft. Нейронные сети. – Режим доступа: <http://www.statsoft.ru/home/textbook/modules/stneunet.html> (дата обращения: 09.04.2017);

6. Применение нейронных сетей для задач классификации [Электронный ресурс]. – Режим доступа: <https://basegroup.ru/community/articles/classification> (дата обращения 09.04.2017);

7. NeuroPro, нейронные сети, методы обработки и анализа данных: от исследований до разработок и внедрений [Электронный ресурс]. – Режим доступа: <http://neuropro.ru/neu2.shtml> (дата обращения 09.04.2017).