

Использование технологии глубокого машинного обучения для построения ИНС-модели профессиональной предрасположенности

А. А. Арзамасцев, e-mail: arz_sci@mail.ru¹

Н. А. Зенкова, e-mail: natulin@mail.ru²

О. В. Крючин, e-mail: kryuchov@gmail.com²

¹ Воронежский государственный университет

² Тамбовский государственный университет им. Г.Р. Державина

Аннотация. *Обсуждаются результаты глубокого машинного обучения искусственной нейронной сети на массиве данных, представляющих собой результаты тестирования более тысячи школьников города Тамбова в плане их профессиональной предрасположенности. В работе использована авторская программа, позволяющая автоматически определять архитектуру сети на основе конструктивного алгоритма. В докладе обсуждаются возможности использования указанного приема для извлечения новых знаний об объекте (Data Mining) и разработке компьютерных систем профессионального тестирования.*

Ключевые слова: *искусственные нейронные сети, машинное обучение, алгоритмы и методы глубокого обучения, ИНС-модель, анализ больших данных, извлечение знаний из данных, системы профессионального тестирования.*

Введение

В настоящее время наблюдается значительный рост числа объектов из области естественных наук, информационных технологий, а также социальной сферы, требующих изучения посредством исследования их математических моделей. В случае, если информация о подобных объектах представлена массивами эмпирических данных, общая методология таких исследований базируется на традиционных приемах построения математических моделей, их параметрической идентификации с последующим изучением модели аналитическими или компьютерными методами.

Значительные возможности и универсализм в плане построения моделей на основе эмпирических данных предоставляет аппарат искусственных нейронных сетей – ИНС-модели [1-4]. Развитие математического обеспечения для научных исследований на базе языка

Python в настоящее время позволяет исследователю сосредоточиться на понимании проблемы и использовать готовые, хорошо отлаженные и адаптированные алгоритмы и методы оптимизации [1], позволяющие в полной мере использовать потенциал нейронных сетей для создания генеративных ИНС-моделей [2]. Использование таких моделей уже приносит значительные результаты в науках о живых объектах - геномике, биологии, медицине [3], и вполне может быть использовано для объектов социальной сферы.

Ранее авторами данного доклада были построены ИНС-модели для различных систем психологического тестирования, позволившие вскрыть систему внутренних связей тестов, а также были разработаны методики психологического и профессионального тестирования, базирующиеся на обучении ИНС-моделей с использованием эмпирических данных для значительного числа респондентов и их независимых оценок [5].

В качестве инструмента получения знаний использовали аппарат искусственных нейронных сетей (ИНС-модели), включающий в себя конструктор структур, который позволяет оперативно изменять структуру ИНС-моделей и алгоритмы машинного обучения, что позволяет успешно решать задачи параметрической идентификации [5].

Выбор архитектуры сети является наиболее существенной проблемой при построении ИНС-модели какого-либо объекта [1]. На практике при конструировании архитектуры сети, предназначенной для описания некоторых эмпирических данных могут быть использованы следующие подходы: интуитивное представление о структуре связей объекта; теоремы Колмогорова о представимости функции многих переменных в виде суперпозиций и сумм функций одной переменной; возможность разложения функции многих переменных в ряд Тейлора; конструктивные (наращивание структуры сети от минимальной до оптимальной) или деструктивные (сокращение структуры сети от некоторой избыточной до оптимальной) алгоритмы.

В наших первых работах 2005-2010 года мы использовали архитектуры ИНС-моделей, вытекающие из классических теорем Вейерштрасса и Колмогорова, а также из разложения функции многих переменных в ряд Тейлора. Практика использования такого подхода показала, что полученные модели имеют следующие недостатки: аппроксимативный характер, не позволяющий увидеть "тонкие" стороны рассматриваемого объекта; незначительное число скрытых слоев, не позволяющее достичь глубокого обучения. В дальнейших работах мы использовали конструктивный алгоритм, обеспечивающий постепенное наращивание нейронов в скрытом слое и числа скрытых

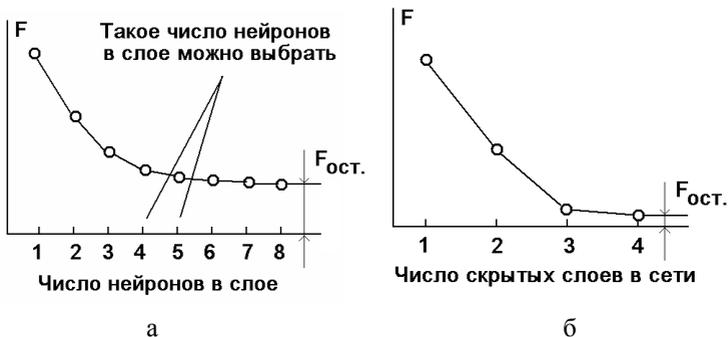
слоев, что позволяло получать многослойные полносвязные структуры и избегать переобучения ИНС-моделей ввиду поэтапного наращивания сети.

Конструктивный алгоритм модификации архитектуры ИНС-модели

Данный алгоритм использует полносвязные ИНС-модели с прямым распространением сигнала. Он базируется на двух эмпирических феноменах обучения ИНС-модели, которые обычно наблюдаются при изменении числа нейронов в скрытом слое и количества скрытых слоев (рис. 1 а, б).

На рис. 1 а показано как изменяется ошибка обучения сети при изменении числа нейронов в скрытом слое. Видно, что вначале, при добавлении новых нейронов ошибка убывает. Однако, начиная с некоторого числа нейронов, она стабилизируется так, что дальнейшее увеличение их количества не приводит к снижению ошибки обучения. По всей видимости, речь здесь может идти о непреодолимой погрешности, которую можно связать с несоответствием структуры ИНС-модели и моделируемого объекта. Можно сделать вывод, что увеличивать число нейронов в слое целесообразно лишь до тех пор, пока это приводит к уменьшению ошибки обучения (на рис. 1 а – это число 4 или 5).

На рис. 1 б показано, как изменяется ошибка обучения сети при увеличении числа скрытых слоев, при условии, что выбор количества нейронов в каждом слое производится так, как это показано на рис. 1 а. Из этого рисунка видно, что ошибка обучения сети быстро убывает до остаточного значения $F_{\text{ост}}$, которое не может быть уменьшено в дальнейшем. Такое значение $F_{\text{ост}}$ может быть связано с погрешностью эмпирических данных, используемых при обучении ИНС-модели. Из рис. 1 б также хорошо видно, что информацию об уменьшении $F_{\text{ост}}$ можно использовать для выбора числа скрытых слоев. Так, ошибка обучения практически не меняется при использовании 3-х или 4-х скрытых слоев, следовательно, такая величина и может быть выбрана для ИНС-модели (рис. 1 б).



а – выбор числа нейронов в скрытом слое, б – выбор количества скрытых слоев ИНС-модели

Рис. 1. К алгоритму выбора числа нейронов в скрытом слое и количества скрытых слоев ИНС-модели

Рассмотренная феноменология обучения ИНС использована нами для конструирования алгоритма самоорганизации структуры. Суть алгоритма заключается в постепенном наращивании числа нейронов в каждом слое и числа скрытых слоев до тех пор, пока ошибка обучения сети $F_{ост}$ не перестанет уменьшаться.

В блоке 1 задают число входов ИНС-модели – n (определяется постановкой задачи и зависит от числа факторов, оказывающих влияние на моделируемый объект), число выходов – k (определяется как число факторов, являющихся результатом моделирования и интересующих пользователя системы); начальное число слоев – S принимается равным единице.

В блоках 2 и 3 задаются начальные условия: начальная лучшая невязка для нейрона ($BestError$) – максимальное число используемого типа; начальная лучшая невязка для слоя ($LayerBestError$) – максимальное число используемого типа; начальная лучшая невязка для всей сети ($NetBestError$). В процессе конструирования сети данные невязки будут изменяться и достигать своего минимального значения.

Также пользователем задается величина $delta$ – незначительное изменение невязки. Если разница между двумя невязками (полученной в настоящий момент времени и предыдущей) становится меньше величины $delta$, то построение слоя (в случае незначительного изменения $BestError$) или построение всей сети (в случае незначительного изменения $LayerBestError$) будет прекращено.

В блоке 4 задается первоначальное число нейронов в каждом из предполагаемых слоев ИНС равным нулю.

В блоке 5 при переходе к построению нового слоя нейронов невязка сети получает то значение, которое получилось в конце построения предыдущего нейронного слоя.

В блоке 6 в начале построения нового нейронного слоя счетчик слоев в сети - S увеличивается на единицу.

В блоке 7 при подборе нового нейрона в текущем слое счетчик нейронов в этом слое - i увеличивается на единицу.

В блоке 8 при подборе нового нейрона в текущем слое невязка данного слоя получает то значение, которое получилось в результате подбора предыдущего нейрона в слое.

В блоке 9 при подборе активационной функции каждого нового нейрона в текущем слое осуществляется перебор всех уже имеющихся в наличии слоев ИНС.

В блоке 10 осуществляется перебор всех нейронов в каждом из уже построенных слоев ИНС.

В блоке 11 осуществляется перебор всех имеющихся в распоряжении пользователя активационных функций нейронов. Таким образом, имеется возможность не просто подобрать активационную функцию только для одного (рассматриваемого в данный момент времени) нейрона при закрепленных активационных функциях остальных нейронов в сети, а произвести перебор всех имеющихся нейронов и, если это приведет к уменьшению значения невязки, поменять активационные функции некоторых из них.

В блоке 12 производится обучение сети по определенному алгоритму (один из методов определения минимума функции многих переменных) и находится значение лучшей текущей невязки - e_j .

В блоке 13 - если произошло уменьшение лучшей текущей невязки - e_j по сравнению с лучшей невязкой для подбираемого в данный момент времени нейрона - $BestError$, то управление передается к блоку 14, в противном случае осуществляется переход к новой активационной функции нейрона (блок 11).

В блоке 14 номер активационной функции, при которой произошло уменьшение e_j , записывается в массив $FBest$ в строку с номером S (этот номер соответствует номеру слоя) и в столбец с номером i (этот номер соответствует номеру нейрона, активационная функция которого подбирается в данный момент времени).

В блоке 15 значение невязки e_j записывается в переменную $BestError$ - теперь это значение будет лучшим для подбираемого нейрона.

В блоке 16 осуществляется запись текущих значений коэффициентов синаптических связей.

В блоке 17 - если произошло существенное уменьшение (на величину, большую, чем δ) величины лучшей невязки последнего нейрона ($BestError$) относительно лучшей невязки для последнего построенного слоя ($LayerBestError$), то можно добавить в текущий слой еще хотя бы один нейрон (см. блоки 18 и 19). В противном случае (если существенного уменьшения невязки не произошло), то добавление нового нейрона в текущий слой, вероятнее всего, не приведет к дальнейшему существенному уменьшению лучшей невязки, т.е. необходимо перейти к блоку 20.

В блоке 18 осуществляется запись количества нейронов на последнем построенном слое.

В блоке 19 производится восстановление начальных коэффициентов синаптических связей и осуществляется переход к блоку 7.

В блоке - если произошло существенное уменьшение (на величину, большую, чем δ) величины лучшей невязки последнего построенного слоя ($LayerBestError$) относительно лучшей невязки сети ($NetBestError$), то можно добавить в сеть еще хотя бы один нейронный слой: перейти к пункту 4. В противном случае (если существенного уменьшения невязки не произошло), то добавление нового слоя в сеть не приведет к дальнейшему уменьшению лучшей невязки сети, т.е. необходимо перейти к блоку 21.

В блоке 21 - сеть построена.

Данный алгоритм реализован в виде нескольких программ [7-8], которые позволяют осуществлять следующие функции:

- ввод исходных данных для обучения сети (векторы x и y) из отдельных файлов, в программе реализуется концепция «обучение с учителем»;
- обучение ИНС по методам сканирования, Монте-Карло и градиентному;
- выбор передаточной функции нейронов (линейная, параболическая второй степени, параболическая третьей степени, сигмоид);
- визуализацию ошибки обучения сети по мере увеличения числа нейронов в слое и увеличения числа слоев;
- визуализацию структуры сети в виде схемы;
- запись в соответствующие файлы результатов обучения сети - ее структуры, коэффициентов передачи (синаптических связей), функций нейронов и погрешностей обучения.

Глубокое обучение ИНС-модели на эмпирических данных по профессиональной предрасположенности школьников

Программа и алгоритм использованы для построения компьютерной модели профессиональной предрасположенности школьников старших классов города Тамбова на основе значительного по объему эмпирического материала [6]. Обучающая выборка, являющаяся результатом анкетирования более чем тысячи школьников, представляла собой прямоугольную матрицу размером 1052 строки и 253 столбца. Строки представляли собой записи, соответствующие респонденту, столбцы – варианты ответа респондентов на вопросы анкеты, касающейся предрасположенности к дальнейшему образованию в университете по различным профилям. В качестве выходной переменной использовали независимую оценку личностных качеств каждого респондента.

Обучение ИНС-модели проводили таким образом, что архитектура сети выбиралась автоматически по описанному ранее алгоритму. При этом, приведенная погрешность на контрольной выборке составила не более 10%, что позволило считать ИНС-модель адекватной реальному объекту (в данном случае объектом являлся социум, включающий учащихся старших классов школ г. Тамбова). Общая структура полученной сети показана на рис. 2, а ее коэффициенты синаптических связей приведены в презентации к докладу.

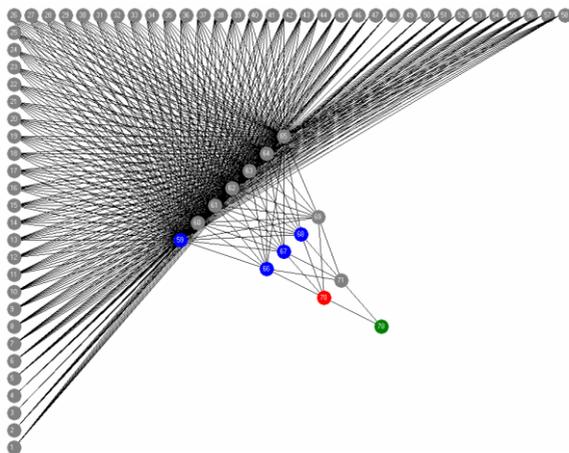


Рис. 2. Структура ИНС-модели: 1-58, 60-65, 69, 71 – входные нейроны и нейроны с линейной активационной функцией; 59, 66, 67, 68 – нейроны с активационной функцией типа «сигмоид» - $(f(s)=1/(1+s^2))$; 70 – нейрон с параболической активационной функцией $(f(s)=s^2)$; 78 – выходной нейрон с параболической активационной функцией $(f(s)=s^3)$. Общее число связей (весовых коэффициентов сети) – 444

Заключение

Таким образом, на основе значительного массива эмпирических данных нами получена многослойная ИНС-модель профессиональной предрасположенности. Из рис. 2 видно, что данная модель имеет четыре скрытых внутренних слоя, содержащих как линейные, так и нелинейные элементы, которые были найдены автоматически в соответствии с приведенным алгоритмом. Данные обстоятельства, а также значения весовых коэффициентов, приведенные в докладе, позволяют классифицировать модель как нелинейную со значительным количеством связей между входными и внутренними параметрами.

Полученная ИНС-модель может использоваться и в качестве генеративной, т.к. обобщая эмпирические данные с погрешностью, не превышающей 10%, она может предсказывать выходные значения для респондентов, результаты тестирования которых не содержатся в обучающей выборке.

Эти и другие свойства полученной ИНС-модели позволили говорить о ее глубоком обучении в данной предметной области.

Указанные обстоятельства позволяют использовать ее как для извлечения новых знаний об объекте (Data Mining), так и в качестве ядра при разработке компьютерных систем профессионального тестирования.

Список литературы

1. Будума Н. Основы глубокого обучения. Создание алгоритмов для искусственного интеллекта следующего поколения / Нихиль Будума, Николас Локашо; пер. с англ. А. Коробейникова; [науч.ред. А. Созыкин]. – М. : Манн, Иванов и Фербер, 2020. — 304 с.

2. Фостер Д. Генеративное глубокое обучение. Творческий потенциал нейронных сетей. — СПб. : Питер, 2020. — 336 с.

3. Глубокое обучение в биологии и медицине / Рамсундар Б. [и др.]; пер. с англ. В. С. Яценкова. – М. : ДМК Пресс, 2020. – 200 с.

4. Харрисон М. Машинное обучение: карманный справочник. Краткое руководство по методам структурированного машинного обучения на Python. : Пер. с англ. — СПб. : ООО «Диалектика», 2020 – 320 с.

5. Арзамасцев А.А., Зенкова Н.А. Моделирование в психологии на основе искусственных нейронных сетей. – Тамбов: ИМФИ ТГУ им. Г.Р. Державина, 2003. – 106 с.

6. Арзамасцев А. А., Безрученко И. Е., Зенкова Н. А. Личностные качества и профессиональная предрасположенность школьников старших классов города Тамбова. – Тамбов : ИМФИ ТГУ имени Г. Р. Державина, 2004. – 103 с.

7. Свидетельство о государственной регистрации программы для ЭВМ № 2007610622. Многофункциональный программный комплекс для компьютерного моделирования на основе искусственной нейронной сети с самоорганизацией структуры / А. А. Арзамасцев, О. В. Крючин, А. Н. Королев, Н. А. Зенкова. – 08.02.2007.

8. Свидетельство о государственной регистрации программы для ЭВМ № 2017619671. Программный комплекс для моделирования нейронных сетей с параметрически изменяющимися активационными функциями нейронов / А. А. Арзамасцев, М. А. Кисляков, Н. А. Зенкова. - 01.09.2017.