

После построения любой модели прогнозирования важно проверять, насколько адекватно она построена. Для этого можно, во-первых, провести визуальный анализ со сдвигом прогноза на несколько шагов назад. А во-вторых, воспользоваться анализом остатков – стандартным методом проверки адекватности любой построенной статистической модели.

В данной работе для решения различных задач классификации и прогнозирования использовались искусственные нейронные сети (Artificial Neural Networks, ANN) [114], а также машины опорных векторов (Support Vector Machines, SVM) [30]. Для их проектирования и настройки были применены разработанные и описанные ранее бионический метод оптимизации COBRA и две ее модификации COBRA-b и COBRA-c.

3.1. Нейросетевые классификаторы, генерируемые коллективными алгоритмами

Искусственная нейронная сеть (ИНС) – математическая модель, а также её программная реализация, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма [70]. Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы. Первой такой попыткой были нейронные сети У. Маккалока и У. Питтса [49]. Нейронные сети применяются для решения различных практических задач интеллектуального анализа данных: задач прогнозирования, задач распознавания образов, задач управления и др.

Искусственные нейронные сети представляют собой систему соединённых и взаимодействующих между собой простых процессоров – искусственных нейронов. Каждый нейрон подобной сети имеет дело только с сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим нейронам. И, тем не менее, будучи соединёнными в достаточно большую

сеть с управляемым взаимодействием, такие локально простые процессоры вместе способны выполнять довольно сложные задачи.

С точки зрения машинного обучения нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа, методов кластеризации и т.д. С математической точки зрения обучение нейронных сетей – это многопараметрическая задача нелинейной оптимизации. С точки зрения кибернетики нейронная сеть используется в задачах адаптивного управления и как алгоритмы для робототехники. С точки зрения развития вычислительной техники и программирования нейронная сеть – способ решения проблемы эффективного параллелизма [38].

Нейронные сети не программируются в привычном смысле этого слова, они обучаются. Возможность обучения – одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Это значит, что в случае успешного обучения сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке, а также неполных и/или «зашумленных», частично искаженных данных.

Чтобы отразить суть нейронных сетей, определение искусственного нейрона дается следующим образом:

- 1) Нейрон получает входные сигналы (исходные данные либо выходные сигналы других нейронов сети) через несколько входных каналов. Каждый входной сигнал проходит через соединение, имеющее определенный вес. Вычисляется взвешенная сумма входов и в результате получается величина активации нейрона.
- 2) Сигнал активации преобразуется с помощью функции активации и в результате получается выходной сигнал нейрона.

Нейроны сети соединены друг с другом. Если сеть предполагается для чего-то использовать, то у нее должны быть входы (нейроны, принимающие входные

данные, образуют, соответственно, входной слой) и выходы (прогнозы или управляющие сигналы). Кроме этого, в сети может быть еще много промежуточных (скрытых) слоев нейронов, выполняющих внутренние функции. Входные, скрытые и выходные нейроны должны быть связаны между собой.

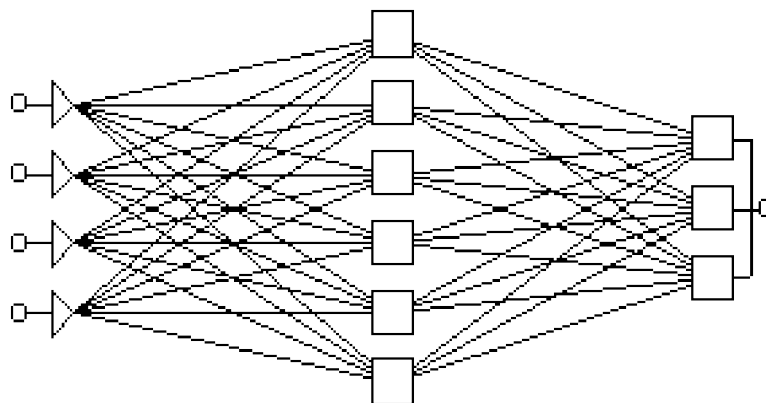


Рисунок 3.1. Пример сети с прямой передачей сигнала

Простейшая сеть имеет структуру прямой передачи сигнала, то есть сигналы проходят от входов через скрытые элементы и, в конце концов, приходят на выходные элементы. Такая структура имеет устойчивое поведение. Типичный пример сети с прямой передачей сигнала показан на рисунке 3.1.

Нейроны регулярным образом организованы в слои. Входной слой служит просто для ввода значений входных переменных. Каждый из скрытых и выходных нейронов соединен со всеми элементами предыдущего слоя. Сеть с подобной организацией может моделировать функцию практически любой степени сложности, причем число слоев и число элементов в каждом слое определяют сложность функции. Определение числа промежуточных слоев и числа элементов на них является важным вопросом при конструировании нейронных сетей.

Таким образом, для решения той или иной задачи при помощи нейронных сетей нужно выбрать ее структуру (количество скрытых слоев, количество нейронов на каждом слое, тип активационной функции на каждом нейроне), а затем для полученной сети настроить весовые коэффициенты связи нейронов.

Стоит отметить, что количество весовых коэффициентов зависит от числа входных параметров заданной задачи и числа нейронов сети.

Итак, для настройки нейронной сети необходимо решить две задачи безусловной оптимизации: одна задача связана с выбором структуры сети, а другая – подбором весовых коэффициентов [61].

В данной работе для нейронной сети генерировалось множество структур, причем каждая структура была представлена в виде бинарной строки. Длина бинарной строки определялась следующим образом:

1. Устанавливалось максимальное число скрытых слоев для сети (в данной работе – 5 слоев).
2. Устанавливалось максимальное число нейронов на каждом скрытом слое нейронной сети (в данной работе 5 нейронов на слое).
3. Каждый нейрон кодировался бинарной строкой заданной размерности. Длина бинарной строки для нейрона определялась количеством активационных функций (в данной работе рассматривались 15 функций, поэтому каждый нейрон кодировался бинарной строкой, длина которой была равна 4).

Как уже было сказано, в данной работе использовались 15 активационных функций, среди которых были тангенс гиперболический, сигмоидальная, линейная функция, пороговая функция, и т.д. Ниже представлен полный список использованных активационных функций:

1. $f(x) = \frac{1}{1 + \exp(-x)}$;
2. $f(x) = 1$;
3. $f(x) = th(x)$;
4. $f(x) = \exp\left(\frac{-x^2}{2}\right)$;
5. $f(x) = 1 - \exp\left(\frac{-x^2}{2}\right)$;
6. $f(x) = x^2$;

$$7. f(x) = x^2;$$

$$8. f(x) = \sin(x);$$

$$9. f(x) = \exp(x);$$

$$10. f(x) = |x|;$$

$$11. f(x) = \begin{cases} -1, & x < -1 \\ x, & -1 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

$$12. f(x) = \begin{cases} 0, & x < -0.5 \\ x + 0.5, & -0.5 \leq x \leq 0.5 \\ 1, & x > 0.5 \end{cases}$$

$$13. f(x) = \frac{2}{1 + \exp(x)} - 1;$$

$$14. f(x) = \frac{1}{x};$$

$$15. f(x) = \text{sign}(x).$$

Если нейрон закодирован бинарной строкой вида «0000», значит, на заданном скрытом слое его не было. Для того чтобы определить, какая именно активационная функция закреплена за заданным нейроном, необходимо двоичное число, которым он кодировался, перевести в десятичное число. Последнее определяло номер активационной функции нейрона согласно приведенному выше списку.

Таким образом, в данной работе каждая структура нейронной сети представлялась бинарной строкой длиной $5 \times 5 \times 4 = 100$. То есть проблему выбора структуры нейронной сети можно рассматривать как задачу безусловной оптимизации с бинарными переменными, и размерность задачи в данной работе была равна 100. Для решения данной задачи (выбора структуры нейронной сети) был использован один из разработанных оптимизационных методов, а именно алгоритм COBRA-b [44].

В итоге при генерировании нейронной сети алгоритмом COBRA-b сначала инициализировалась популяция структур в виде заданного множества бинарных

строк. Для каждой структуры подбирались весовые коэффициенты, затем выполнялась «основная» работа эвристики, и в конечном итоге выбиралась та структура, которая решала бы задачу интеллектуального анализа данных, допустим задачу классификации, лучше всего.

Для каждой нейронной сети подбирались весовые коэффициенты в зависимости от полученной для нее структуры. Именно на этом этапе и происходит обучение сети. По своей сути настройка весовых коэффициентов нейронной сети – это задача безусловной оптимизации с вещественными переменными. Одним из наиболее известных и часто используемых методов «обучения» сети является алгоритм обратного распространения [36]. Однако в настоящее время разрабатывается все больше новых методов настройки весов ИНС, например, в статье [123] описан один из них. В данной работе для подбора весовых коэффициентов нейронной сети был применен один из разработанных методов, а именно алгоритм COBRA [23, 22].

Таким образом, после того как алгоритмом COBRA-b будет сгенерирована популяция структур (множество бинарных строк), для каждой структуры алгоритмом COBRA подбираются весовые коэффициенты. Затем выполняются остальные этапы метода COBRA-b. В итоге определяется наилучшая структура нейронной сети и для этой структуры вновь выполняется настройка весовых коэффициентов разработанной эвристикой COBRA.

3.2. Генерирование машин опорных векторов бионическими алгоритмами

Метод опорных векторов (Support Vector Machine, SVM) – набор схожих алгоритмов обучения с учителем, использующихся для задач классификации и регрессионного анализа [30]. Принадлежит к семейству линейных классификаторов. Особым свойством метода опорных векторов является непрерывное уменьшение эмпирической ошибки.

Основная идея метода опорных векторов заключается в переводе исходных векторов в пространство более высокой размерности и поиске разделяющей