

**Кретов В.С., Аблов И.В., Котов Н.М.**

## **Подход к применению нейронных сетей к задачам классификации объектов**

**Аннотация:** Предлагается подход к решению задачи классификации объектов, описываемых по схеме «признак-значение» (международные конфликты, кризисные ситуации, террористические акты т.п.) с применением искусственных нейронных сетей.

**Ключевые слова:** искусственный интеллект, автоматизация классификации объектов, искусственные нейронные сети, препроцессинг, подбор весовых коэффициентов

### **Введение**

Решение задачи классификации является одним из важнейших применений нейронных сетей. Задача классификации представляет собой задачу отнесения классифицируемого объекта к одному из нескольких попарно не пересекающихся множеств. Примерами таких задач могут быть, например, определение кредитоспособности клиента банка, медицинская и техническая диагностика, управление портфелем ценных бумаг (продать, купить или «придержать» акции в зависимости от ситуации на рынке) и т.д. [1].

В докладе рассматривается такой объект как военно-политическая обстановка, уровень опасности которой может быть отнесен классификатором в зависимости от сочетания значений признаков объекта к одному из 4-х классов, условно названных «зеленый», «синий», «желтый», «красный».

Рассмотрим основные этапы создания нашего классификатора.

### **1. Подготовка исходных данных**

1.1. *Выбор множества признаков и их значений (не зависит от времени и места классифицируемой кризисной ситуации)*

Прежде всего, при построении классификатора необходимо определить, какие признаки влияют на принятие решения о том, к какому классу принадлежит объект. В нашей задаче политологом - экспертом было выбрано из [2] следующее множество признаков: 1) *Мирное сосуществование*; 2) *Диспут*; 3) *Конфликт*; 4) *Военные действия*; 5) *Урегулирование*. Из-за недостатка места в данной

статье приведем список возможных значений только одного признака «*Конфликт*»: Нарушение территориальной целостности государства; Угрозы применения силы; Военная мобилизация; Пограничные вопросы; Похищение военных; Эскалация взаимных обвинений государств [2]. Добавлены признаки: *Число пострадавших* и *Площадь территории конфликта*. Признаки и их значения записываются в виде двумерной таблицы.

1.2. *Формирование политологом-экспертом в данной предметной области обучающей выборки (зависит от времени и места классифицируемой кризисной ситуации*. Ввод обучающей выборки в терминах предметной области производится с клавиатуры. Политолог-эксперт для каждой строки обучающей выборки, представленной в виде таблицы, последовательно выбирает из таблицы «Признаки описания объектов и их значения» для каждого признака нужные значения, а после формирования значений всех признаков строки обучающей выборки вводит присвоенный этим экспертом класс военно-политической ситуации, описываемой в данной строке обучающей выборки.

## **2. Препроцессинг**

Основная задача препроцессинга – отображение данных в формат, пригодный для функционирования нейросети [3].

Сначала необходимо определить способ представления входных данных (исходная и обучающая выборки) нейронной сети, т.е. определить *способ нормирования*. Нормирование необходимо, поскольку нейронные сети работают с данными, представленными числами в диапазоне от 0,01 до 0,99, а исходные данные могут иметь произвольный диапазон (например, признак *Число пострадавших* может иметь значения *Малое* ( $a \leq 10$ ), или быть категориальными (например, признак *Объекты воздействия на население* может иметь значение: *Жизнь, здоровье, свобода и безопасность людей*).

При этом возможны различные способы нормирования: или путем простого линейного преобразования числа в требуемый диапазон (например, с помощью сигмоидальной функции), или путем замены булевых переменных на 1 или 0, или другим известным способом [1].

Строки исходной выборки должны иметь следующую классическую структуру: вектор значений признаков для каждой строки, содержащей информацию о классифицируемом объекте,  $X^{(i)} = \{x^{(i)}_1, x^{(i)}_2, \dots, x^{(i)}_n\}$ . В обучающей выборке кроме вектора значений признаков содержится также значение присвоенного политологом - экспертом значения признака класса.

Для нормирования обучающей выборки воспользуемся «Унитарным кодом» [4], что позволяет представить категориальные признаки в виде векторов в векторном пространстве размерностью, соответствующей количеству возможных значений признаков. При этом значение координаты этой категории берется за единицу, а все остальные координаты обнуляются. Поскольку нейронные сети работают с данными, представленными вещественными числами в диапазоне от 0,01 до 0,99, то нужно переформатировать представленные в булевой форме исходные данные (обучающая и исходная выборки), представив булевы переменные «1» как вещественную переменную «0,99» и «0» как вещественную переменную «0,01»).

При выборе объема сети будем применять конструктивный подход, предполагающий, что вначале берётся сеть минимального размера, который будем постепенно увеличивать до достижения требуемой точности, обучая сеть на каждом шаге.

При выборе архитектуры сети будем опробовать несколько конфигураций с различным количеством элементов. При этом основным показателем является объем обучающейся выборки и обобщающая способность сети.

### **3. Алгоритм построения нейросети**

Процесс построения нейросети включает в себя 2 главных этапа: инициализация сети, обучение сети [5].

#### **3.1. Инициализация сети**

Нам необходимо задать количество узлов *входного, скрытого и выходного узлов*. Эти данные определяют конфигурацию и размер нейронной сети. Вместо того, чтобы жёстко задавать их в коде, лучше предусмотреть установку соответствующих значений в виде параметров во время создания нейронной сети [5], что позволит без особого труда создавать новые нейронные сети различного размера.

Число слоев выбрано равным 3 (входной, скрытый и выходной слой).

Для начала зададим количество узлов входного слоя равным 16-ти (по числу различных значений признаков описания входного сигнала), количество узлов скрытого слоя, количество узлов выходного слоя равным 4-м (по числу классов объектов в нашем примере).

Коэффициент обучения зададим равным  $\eta=0,2$ .

Вид функции активации нейронов - «сигмоида».

Важной частью нейросети являются весовые коэффициенты связей (веса). Они используются для расчета распространения сигналов в прямом направлении, а также обратного распространения ошибок. Значения весовых коэффициентов первоначально назначаются случайными числами и уточняются при попытке улучшить характеристики нейросети.

### 3.2. Обучение сети

Процесс обучения сети можно разделить на 3 части:

1) расчет выходных сигналов сети для примеров обучающей выборки;

2) сравнение рассчитанных выходных сигналов с желаемым ответом;

3) обновление весовых коэффициентов связей между узлами на основе найденных при расчете различий.

#### 3.2.1. Расчет сигналов для всех примеров обучающей выборки

Нейронные сети работают лучше, если, например, входные данные конфигурируются таким образом, чтобы они оставались в диапазоне значений, оптимальном для функций активации узлов нейронной сети. Мы выбираем значение 0,01 в качестве нижней границы диапазона, чтобы избежать проблем с нулевыми входными значениями, поскольку они могут искусственно блокировать обновление весов. Выходные сигналы не могут превышать значение 0,99 [5].

При расчете сигналов нейросети будем помнить, что единственная задача первого слоя – просто представлять входной сигнал во входных узлах, функция активации к входным сигналам не применяется.

По формулам, приведенным в [5], проводится расчет сигналов в следующей последовательности: 1) *входящих сигналов для*

скрытого слоя, 2) выходных сигналов скрытого слоя, 3) входного сигнала третьего слоя, 4) выходного сигнала третьего слоя.

Чтобы отклик третьего слоя на входной сигнал как можно лучше имитировал аналогичный реальный процесс в мозгу человека мы, должны применить к узлам третьего слоя функцию активации.

Нам удалось успешно описать распространение сигналов по нейтронной сети, т.е. определить величину выходных сигналов при заданных величинах входных сигналов. Наш следующий шаг заключается в сравнении выходных сигналов нейронной сети с данными обучающей выборки для определения величины ошибки. Величину ошибки необходимо знать, чтобы улучшить выходные результаты путём изменения параметров сети.

### 3.2.2. Сравнение рассчитанных выходных сигналов с заданными в обучающей выборке выходными сигналами

Для определения подходящих значений весовых коэффициентов можно использовать метод градиентного спуска в том случае, если правильно выбрать функцию ошибки, которая представляет собой разность между заданными выходными значениями из обучающей выборки и фактическими рассчитанными выходными сигналами.

В качестве функции ошибки в [5] предложено использовать квадрат разности между заданными и рассчитанными выходными сигналами.

Причинами выбора подобной функции ошибки являются следующие:

упрощаются вычисления, с помощью которых определяется величина наклона графика функции ошибки для метода градиентного спуска;

функция ошибки является непрерывно гладкой, что обеспечивает нормальную работу метода градиентного спуска ввиду отсутствия провалов и скачков значений функции ошибки;

по мере приближения к минимуму градиент уменьшается, что означает снижение риска перескока через минимум, если используется уменьшение величины шагов.

Подобным образом рассчитываются значения функции ошибок для всех узлов нейросети.

### 3.2.3. Обновление весовых коэффициентов связей между узлами на основе найденных различий [5]

Алгоритм обратного распространения ошибок в нейросети с помощью матричной алгебры.

Воспользуемся методом градиентного спуска, для чего нам нужно определить наклон функции ошибки по отношению к весовым коэффициентам, что требует применения дифференциального исчисления. Нас интересует зависимость функции ошибки от весовых коэффициентов связей внутри нейронной сети или, иначе говоря, то, насколько величина ошибки чувствительна к изменениям весовых коэффициентов.

Обновленный вес  $W_{jk}$  – это старый вес с учётом отрицательной поправки, величина которой пропорциональна производной функции ошибки. Данное выражение применяется к весовым коэффициентам связей не только между скрытым и выходным, но и между входным и скрытым слоями. Эти два случая различаются градиентами функции ошибки, выражения для которых приведены выше.

### **Расчет и отображение результатов классификации военно-политической обстановки**

Перейдем к опросу нейросети – получению результата классификации нового объекта исходной выборки, представленного в терминах предметной области пользователя в форме «признак-значение».

Выходные значения сигналов узлов нейросети представляются вектором. При интерпретации результата опроса будем считать, что номер класса определяется номером выхода сети, на котором появилось максимальное значение. При этом будем также определять *уверенность* сети в том, что тестовый пример относится к этому классу, как разность между максимальным значением выхода и ближайшим к максимальному значению другого выхода. Чем выше уверенность, тем больше вероятность того, что сеть дала правильный ответ [1]. В результате опроса нейросети на экран монитора выдается окно, где содержится номер и описание классифицируемого объекта, а также название класса

классифицируемого объекта: Класс 1 «ЗЕЛЕНЬ» – нет опасности; Класс 2 «СИНИЙ» – повышенный уровень опасности; Класс 3 «ЖЕЛТЫЙ» - высокий уровень опасности; Класс 4 «КРАСНЫЙ» – критический уровень опасности». При невозможности классифицировать объект на экран монитора выдается окно с номером и описанием классифицируемого объекта, а также предупреждением «КЛАССИФИКАЦИЯ ОБЪЕКТА НЕВОЗМОЖНА – ТРЕБУЕТСЯ ДООБУЧИТЬ НЕЙРОСЕТЬ».

### **Заключение**

Апробация нейросети на учебной задаче подтвердила её работоспособность. Однако серьезным препятствием для практического использования нейросети в реальных задачах классификации военно-политической обстановки является необходимость формирования вручную большого объема обучающей выборки, что определило для нас такое направление дальнейшего совершенствования данной нейросети, как автоматизация формирования обучающей выборки.

### **Литература:**

1. Применение нейронных сетей для задач классификации. Loginom DAY 2018/  
<https://basegroup.ru/community/articles/classification> (Дата обращения: 06.06.2019)
  2. *Lincoln P. Bloomfield and Allen Moulton* Managing International Conflict: From Theory to Policy. – New York Martin's Press, 1997.
  3. Краткий курс машинного обучения или как создать нейронную сеть для решения скоринг задачи <https://habr.com/post/340792/> (Дата обращения: 06.06.2019)
  4. Статья в Википедии «Gne-hot (рус.Унитарный Код)» [«https://en.wikipedia.org/wiki/One-hot](https://en.wikipedia.org/wiki/One-hot) (Дата обращения: 06.06.2019).
  5. *Рашид Тарик*. Создаем нейронную сеть. – ООО «Диалектика, 2017. – 274 с.
  6. *Роберт Каллан*. Нейронные сети. / Краткий справочник. Пер. с англ. – М.: ООО «И.Д. Вильямс», 2018. – 288 с.
-