

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ АНАЛИЗА ДАННЫХ

А. В. ГАВРИЛОВ*, В. М. КАНГЛЕР♥

В данной работе рассматриваются основные задачи анализа реляционных баз данных и их решение с помощью нейронных сетей. Для их решения предлагается использовать нейронную сеть Хопфилда, и рассматриваются особенности ее обучения и функционирования. Описаны этапы анализа реляционных баз данных. Выделены основные достоинства предлагаемого подхода в сравнении с применением многослойных перцептронов и методов математической статистики.

1. ВВЕДЕНИЕ

Острая конкурентная борьба, стремление компаний получать отдачу от инвестиций в информационные технологии, рост числа сотрудников, принимающих решения, стимулировали активное развитие новой области информатики – обнаружение знаний в базах данных (KDD – Knowledge Discovery in Databases). Ее основное назначение - автоматизированный поиск ранее неизвестных закономерностей в базах данных, хранящих информацию (например, о деятельности компаний), и использование добытых знаний в процессе принятия решений. С помощью методов KDD можно выявить, например, профиль потребителей данного товара, предотвратить махинации с кредитными карточками или предсказать изменение ситуации на финансовом рынке.

Основной задачей систем KDD является проведение анализа данных, содержащихся в базах данных, с целью обнаружения скрытых или неочевидных закономерностей. Выделяют пять стандартных типов закономерностей, которые позволяют выявлять автоматические интеллектуальные методы анализа и представления данных: ассоциация, последовательность, классификация, кластеризация и прогнозирование [3].

* Доцент кафедры вычислительной техники, канд. техн. наук, докторант
♥ аспирант

Средства обнаружения знаний в базах данных в основном используют пять методов: индукция ассоциативных правил; деревья решений; К - ближайшие соседи; нейронные сети; генетические алгоритмы [3]. Иногда применяется их комбинация.

Доклад посвящен решению двух задач KDD - обнаружение ассоциаций и прогнозирование - методами искусственных нейронных сетей (ИНС).

2. АНАЛИЗ БАЗ ДАННЫХ С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ ХОПФИЛДА

Под анализом БД в данной работе понимается комплекс мероприятий, направленных на выявление скрытых закономерностей в базах данных. Рассматривается использование аппарата искусственных нейронных сетей для решения задачи, которая в рамках задач автоматизированного интеллектуального анализа данных может трактоваться следующим образом.

1. Качественное решение задачи определения значений полей БД, ассоциативно связанных с известными значениями других полей.
2. Решение задачи прогнозирования значений набора полей БД при заданных значениях другого набора полей.

Имея систему, схожую с ассоциативной памятью, а также формальные правила переноса информационного содержимого анализируемой базы данных в эту память, можно решить поставленную задачу. При этом используется ассоциативность такой системы и по значению некоторого фактора - "ключа" извлекаются ассоциативно близкие значения других факторов.

Один из возможных путей реализации ассоциативной памяти состоит в том, чтобы построить распределенную динамическую систему или сеть из дискретных элементов, аттракторами которой являются типичные картины-образы (в настоящее время принято сопоставлять различным образам, запомненным нейронной сетью, установившиеся режимы или аттракторы соответствующей динамической системы). Каждая такая картина будет обладать своей областью притяжения, и всякое начальное условие, представляющее собой какую-то допустимую картину, обязано попасть в одну из ее областей притяжения. С течением времени в ходе эволюции эта начальная структура трансформируется в наиболее близкую из хранящихся в памяти структур-аттракторов, а именно в ту, области притяжения которой она принадлежала.

Следовательно, подавая на вход в качестве начального условия для такой распределенной системы некоторую структуру, мы будем осуществлять ее автоматическое распознавание, которое будет параллельным. В роли такой распределенной динамической системы предлагается использовать искусственную нейронную сеть.

Предлагаемое нами решение в общем виде выглядит следующим образом. Искусственная нейронная сеть обучается на записях анализируемой реляционной базы данных. В процессе обучения нейронная сеть становится гносеологической моделью обучающей БД. Полученную таким образом модель пользователь может использовать для прогнозирования и исследования ассоциативных связей, скрытых в базе данных.

В качестве такой модели предлагается использовать сеть, представляющую собой конкретную реализацию динамической системы с ассоциативной памятью, предложенную Хопфилдом [2], [5], [6].

Поле факторов для данной задачи является содержание реляционной базы данных. В представлении совокупности факторов в виде содержания реляционной базы данных заключается первичная подготовка данных для решения поставленной задачи. Тем самым задаются структурные отношения и задача сводится к определению семантической близости содержимого полей заданной структуры. Так как бинарные искусственные нейронные сети оперируют двоичными векторами, то требуется дополнительная обработка исходных данных с целью приведения к виду, представляемому таким вектором. Необходимый для искусственной нейронной сети вектор получается конкатенацией двоичных векторов, представляющих поля реляционной базы данных (рис. 1).

Каждое поле реляционной базы данных обобщенно может быть отнесено к одному из двух типов: числовое, символьное. Значение символьного поля заносится в соответствующий каждому полю словарь. С тем, чтобы повысить эффективность процедур поиска значений в словаре и их кодирование, содержимое словарей ранжируется, повторяющиеся значения исключаются. Значения числовых полей представляются иначе. Диапазон значений каждого из числовых полей (к числовым полям относятся и те, которые содержат даты) разбивается на интервалы, а каждое конкретное значение представляется интервалом, к которому оно принадлежит. Точку на числовой оси можно представить вырожденным интервалом. Замена числовых значений интервалами

обеспечивает большую гибкость представления. Аналогом словарей для числовых полей является перечень интервалов.

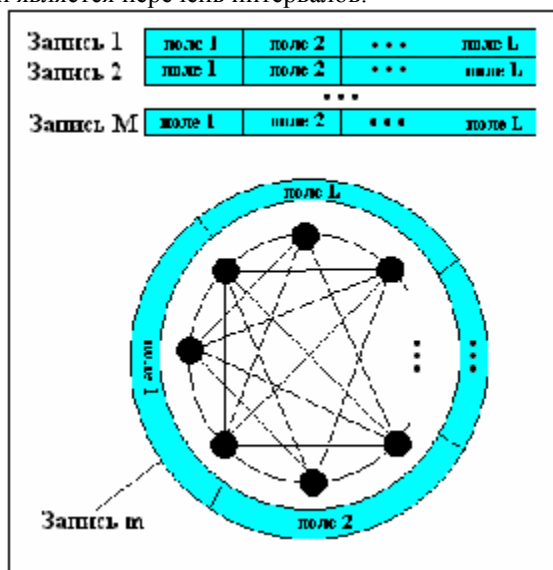


Рис. 1. Схематическое изображение представления структуры базы данных нейронной сетью.

После того, как были сформированы словари и наборы интервалов, они подвергаются кодированию. То есть каждому значению в словаре или интервалу в наборе ставится в соответствие двоичный код, представляющий впоследствии это значение или интервал в нейронной сети. Размерность кода (длина кода) для каждого словаря и набора (другими словами для каждого поля) определяется мощностью множества разнообразных значений в словаре или наборе, а также способом кодирования. Различные варианты кодирования могут существенно сказываться на качестве работы ассоциативной памяти, представленной нейронной сетью Хопфилда, что подтверждается результатами макетного моделирования.

Формирование словарей и их кодирование является подготовительным этапом обучения нейронной сети. Каждая запись базы данных представляется в виде двоичного вектора и участвует в формировании весовых коэффициентов связей между формальными

нейронами сети. Само обучение осуществляется в соответствии с правилом Хебба [2].

Из обучения желательно удалить повторяющиеся записи базы данных. Так как повторная запись вектора в ассоциативную память, представленную нейронной сетью неблагоприятно сказывается на энергетическом ландшафте этой сети. А именно, удваивается глубина энергетического минимума соответствующего такому вектору [5]. Более глубокий минимум обычно имеет более широкую область притяжения, а это уменьшает области притяжения соседних минимумов и искажает соответствующие притягивающие векторы. В предельном случае это может привести к тому, что в системе останется всего один очень глубокий минимум энергии, область притяжения которого захватывает все возможные спиновые конфигурации. У системы формируется как бы “навязчивая идея”, что, конечно же, крайне нежелательно.

Существуют различные версии сетей Хопфилда сходные в структуре, но несколько отличные в функционировании. В данной работе рассматриваются только бинарные (возможно два состояния нейрона: +1 и -1) нейросети Хопфилда с дискретным функционированием во времени. Спецификой объединяющей такие сети является их асинхронность функционирования. То есть в отличие от синхронного функционирования, при котором состояние всех нейронов сети определяется одновременно, асинхронное функционирование подразумевает в каждый конкретный момент времени возможность переключения только одного нейрона.

В модели Хопфилда могут быть реализованы детерминированный и/или стохастический (с имитацией “отжига”) алгоритмы. Процесс функционирования нейронной сети является многоитерационным. Каждая итерация включает в себя два шага: выбор нейрона-кандидата, формирование состояния выбранного нейрона-кандидата.

Отличие стохастического функционирования от детерминированного заключается в методе выбора нейрона-кандидата. При стохастическом кандидатом на переключение является нейрон, выбранный случайно с помощью датчика случайных чисел. Таким образом, возможны такие ситуации, что в процессе функционирования состояние некоторых нейронов не анализировалось. При детерминированном функционировании нейроны становятся кандидатами на изменение своего состояния в порядке следования своих номеров.

В алгоритмах с имитацией “отжига” изменение состояния нейрона-кандидата носит вероятностный характер. Если в результате изменения состояния на противоположное полная энергия сети понизится, то состояние нейрона меняется на противоположное. Иначе состояние нейрона меняется на противоположное с определенной вероятностью, зависящей от параметра, называемого “температурой” сети, имитирующей уровень теплового шума. На первой итерации функционирования нейросети устанавливается начальное значение температуры (максимальное значение), затем постепенно через некоторое количество итераций значение температуры уменьшается. Итерационный процесс прекращается, когда температура достигает конечного значения (минимальное значение).

Если сравнивать между собой алгоритмы без “отжига” и с имитацией “отжига”, то у первых множество преимуществ перед вторыми (проще, требуют меньшее количество итераций, обладают большей точностью - если энергетический ландшафт вблизи глобального минимума не изрезан локальными минимумами). Но, если энергетический ландшафт содержит множество локальных минимумов, то алгоритмы с имитацией “отжига” ведут себя лучше (с точки зрения критерия качества функционирования). Подтверждением последнего утверждения могут служить данные макетного моделирования.

3. ЭТАПЫ РАБОТЫ ПРИ АНАЛИЗЕ БАЗЫ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННОЙ СЕТИ

Использование нейронной сети Хопфилда для анализа базы данных включает в себя следующие этапы:

1. Формирование SQL-запроса для работы с нейронной сетью.
2. Формирование словаря содержимого полей SQL-запроса.
3. Обучение нейронной сети.
4. Поиск ассоциаций или прогноз значений полей БД.

Последний этап в свою очередь включает в себя следующие этапы:

1. Формирование запроса. На этом этапе пользователь выбирает часть полей БД и придает им некоторые значения (используя соответствующие этим полям словари значений и наборы интервалов). В процессе функционирования система будет прогнозировать значения других полей (отличных от задаваемых пользователем) подобно ассоциативной памяти.

2. Подготовка входного вектора сети. На этом этапе для нейронной сети формируется входной вектор. Он “собирается” из подвекторов, соответствующих полям базы данных. При формировании вектора нулевые разряды подвекторов заменяются на (-1). Те подвектора, чьи значения задает пользователь, называются “замороженными”. Значения других подвекторов формируются с использованием датчика случайных чисел.

3. Циклы функционирования нейронной сети. Этот этап включает в себя ряд циклов функционирования нейронной сети. Перед каждым циклом на нейронную сеть подается входной вектор, сформированный на предыдущем этапе функционирования. Нейронная сеть функционирует в соответствии с одним из четырех выше названных алгоритмов. Каждый цикл функционирования нейронной сети состоит из многократных итераций, при этом каждая итерация в свою очередь включает два вышеописанных этапа (выбор нейрона-кандидата, формирование состояния выбранного нейрона-кандидата). Следует заметить, что нейроны соответствующие “замороженным” подвекторам, не могут быть выбраны в качестве нейрона-кандидата, и, следовательно, они не меняют своего состояния в течении всего функционирования. По завершению первого цикла полученный нейронной сетью вектор и значение энергии сети запоминается в специальном буфере. В конце каждого последующего цикла значение энергии сети сравнивается с запомненным значением из буфера. Если текущее значение энергии меньше запомненного в буфере, то текущее состояние сети (состояние сети представляет собой вектор в качестве компонент которого выступают состояния нейронов) замещает в буфере ранее запомненный вектор.

4. Формирование выходного вектора. Это чисто формальный этап, когда вектор, сохраненный в буфере, выдается системой как вектор-ответ, восстановленный ассоциативной памятью по заданному пользователем фрагменту.

4. ЗАКЛЮЧЕНИЕ

К основным достоинствам предлагаемого нами подхода можно отнести следующие.

- Органичное объединение решений таких задач как прогнозирование и обнаружение ассоциаций в рамках единого подхода.
- Обучение нейронной сети с целью последующего моделирования требует всего двух сплошных проходов по моделируемой базе данных.

В то время как при традиционном подходе, а именно при накапливании статистик о комбинациях значений, встречающихся в БД, требуются многоитерационные сложные процедуры поиска.

- Решение задачи в тех ситуациях, когда традиционный подход оказывается статистически несостоятельным. Например, когда анализируемая БД содержит небольшое число похожих (нетождественных) записей.
- Универсальность нейронной сети Хопфилда, используемой в качестве основы построения ассоциативной памяти.
- Простота программной реализации модели Хопфилда.

Предложенный подход реализован в программе для анализа баз данных AnalDB. Исходными данными для работы с нейронной сетью в ней служит SQL-запрос. Программа реализована на Delphi 3.0 и CBuilder 3.0.

Данная работа поддержана программой "Университеты России - фундаментальные исследования" и грантом МОПО РФ по направлению "Информатика и вычислительная техника".

[1] ГАВРИЛОВ А. В. и др. *Обнаружение ассоциативных взаимосвязей между полями в базах данных с использованием нейронной сети* // Труды международной научно-технической конференции "Научные основы высоких технологий" 1997, том 2, стр. 210.

[2] ЛОСУТОВ А. Ю., Михайлов А. С. *Введение в синергетику*. - М.: Наука, 1990 год.

[3] *Программы обнаружения знаний - как выбрать?* // Банковские технологии июнь/97, стр. 91 - 93.

[4] *Продукты для интеллектуального анализа данных*. // Рынок программных средств №14,15/97, стр. 32 - 33.

[5] *Спиновые стекла и нейронные сети*. / под ред. Веденова Л. А. - М.: ВИНТИ - 1990, 384 с.

[6] УОССЕРМАН Ф. *Нейрокомпьютерная техника* - М.: Мир, 1992.