

Нейросетевая технология оптимизационного синтеза и передачи баз знаний по распределённым в сети Internet базам данных

Тимофеев А.В., Азалецкий П.С.

Санкт-Петербургский институт информатики и автоматизации РАН

tav@iias.spb.su, eaglenk2@mail.ru

В настоящее время современные компьютерные сети позволяют накапливать большие массивы данных для решения задач, обработки информации и управления. Однако, компьютерная форма представления данных сама по себе содержит знания в скрытом виде. Для получения этих знаний в каком-либо формализованном виде используют различные методы поиска и анализа данных. Стоит отметить, что сейчас большинство тематически связанных данных хранятся не в одной базе данных и территориально распределены. Очевидно, необходимы методы, позволяющие интегрировать совокупность скрытых знаний отдельных БД в единую базу знаний определённой предметной области. Для оптимизационного синтеза и передачи баз знаний по распределённым базам данных можно использовать нейросетевые технологии, а в качестве транспорта для передачи этих знаний - глобальную сеть Internet.

Большой объем информации, с одной стороны, позволяет получить более точные расчеты и анализ, а с другой стороны – превращает синтез знаний и поиск решений в сложную задачу. В результате появился целый класс систем, выполняющих роль аналитиков или инженеров по знаниям. Они не просто анализируют хранящиеся данные, но и могут строить модели принятия решения на основе проделанного анализа данных. Основной задачей систем поиска знаний в базах данных, называемых также Knowledge Discovery Databases (KDD), является проведение анализа информации, содержащейся в базах данных (БД), с целью обнаружения скрытых, неочевидных или неизвестных закономерностей. Эти задачи усложняются тем, что современные БД гетерогенны и могут быть, как уже отмечалось выше, распределены по разным серверам. Кроме того, для каждого типа БД необходим свой способ доступа к данным. Вследствие этого, для каждой системы интеллектуального анализа данных необходимо разрабатывать свои алгоритмы применительно к реалиям разрабатываемой системы. Такие KDD-системы в процессе анализа должны учитывать еще одну особенность БД, которая заключается в человеческих ошибках при наборе данных оператором. Их наличие неизбежно с ростом объемов БД и система должна уметь их выявлять.

KDD-система позволяет организовать поиск полезных знаний в “сырых”

(например, экспериментальных) данных. Она реализует процессы подготовки данных, выбора информативных признаков, очистки данных, применения методов Data Mining (DM), постобработки данных и интерпретации полученных результатов. Ядром KDD-технологии являются методы DM или алгоритмы обучения нейронных сетей (НС), позволяющие обнаруживать знания и формировать нейрознания. Синтезируемые знания могут иметь различные формы: правила, описывающие связи между свойствами данных (деревья решений), часто встречающиеся шаблоны (ассоциативные правила), нейросетевые классификаторы и кластеризаторы данных (карты Кохонена) и т.д.

Описываемая технология предполагает определённое значение термина база знаний, а именно математическая модель принятия решения в какой-либо предметной области, то есть она представляет собой алгоритм вычисления результирующих данных по множествам значений входных признаков.

Нейросетевая KDD-технология определяет последовательность действий, которые необходимо сделать для того, чтобы из исходных данных получить нейрознания. Эта технология универсальна и не зависит от онтологии предметной области.

Описание основных понятий технологии:

В качестве математической модели знаний используется *полиномиальная нейронная сеть*. Она описывается арифметическими (диофантовыми) полиномами и предназначена для распознавания сложных (линейно неразделимых) классов образов, заданных в n -мерном пространстве двоичных или многозначных предикатов-признаков. Алгоритмы обучения таких НС впервые были предложены в [2]. Схематическое изображение данной сети представлено на рисунке 1. Данная модель нейронной сети обладает рядом преимуществ по сравнению с другими моделями нейронных сетей.

- Алгоритмы обучения позволяют за конечное число шагов, сравнимое с объемом обучающей выборки (специально созданной таблицы предметной области), синтезировать архитектуру НС.
- Рассчитать ее синоптические параметры.
- Процедуры обучения носят алгебраический характер и в отличие от градиентных процедур всегда сходятся к искомому решению [2-4]. Что позволяет гарантировать безошибочное распознавание объекта на обучающей выборке.

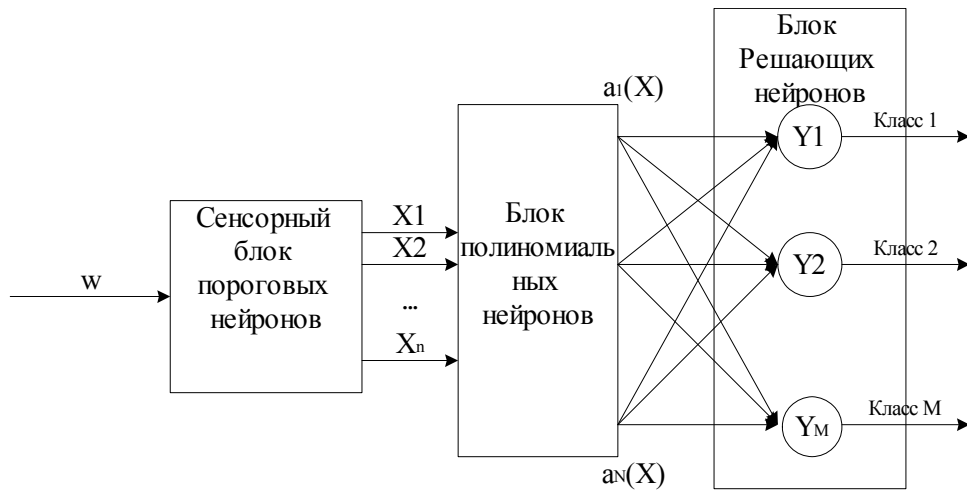


Рис. 1. Структурная схема порогово-полиномиальной НС.

В процессе обучения на входных данных НС корректирует свои синоптические веса и настраивает связи между нейронными элементами. Полученная НС выражает закономерности, присутствующие в исходных данных и позволяет принимать решения на новых данных. Синтезированная НС является функциональным эквивалентом некоторой модели знаний, подобной тем, которые строятся в традиционном моделировании или аппроксимации неизвестных закономерностей.

На входе НС имеется сенсорный блок пороговых нейроноподобных элементов (НПЭ), который кодирует объекты ω в виде вектора признаков, в виде двоичных кодов.

$$X(\omega) = |x_i(\omega)|_{i=1}^n$$

Полученный входной образ (двоичный вектор X) поступает на блок полиномиальных преобразователей (R-нейронов), формирующих N -мерный вектор вторичных (полиномиальных) признаков

$$A(X) = (a_1(X), a_2(X), \dots, a_m(X)).$$

Эти вторичные признаки определяют пространство полиномиальных $a_j(X), j = 1, 2, \dots, N$, признаков, называемое “спрямляющим пространством” [2]. Явный вид полиномиальных функции $a_j(X)$ выбирается адекватно решаемой задаче непосредственно по ОБ.

В выходном слое решающих нейронов расположены решающие пороговые нейроны вида:

$$Y(X) = \text{sign}(w_0 + \sum_{j=1}^N w_j a_j(X)),$$

где X – вектор входных сигналов, $W = (w_0, \dots, w_N)$ – вектор синоптических весовых коэффициентов

$$\text{sign}(u) = \begin{cases} 1, & \text{если } u \geq 0, \\ -1, & \text{если } u < 0. \end{cases}$$

Результат работы НС вычисляется по формулам

$$Y_i = \text{sign}(w_{i,0} + \sum_{j=1}^N w_{i,j} a_{i,j}(Y)), \quad i \in [1..M].$$

Технология получения базы знаний из базы данных предполагает определённую структуру таблиц, из которых возможно извлечение данных. То есть необходимо предварительное преобразование прикладных данных, определение признаков, по которым будет строиться база знаний. Это признаки-предикаты, определяющие значение целевого признака. Отсюда структура таблицы базы данных должна иметь следующий вид:

Результат	Признак 1	...	Признак N
-----------	-----------	-----	-----------

Рис. 2. Структура таблицы для построения базы знаний

Для формирования данной таблицы необходим эксперт в конкретной прикладной области, который определит, какую базу знаний он хочет создать, то есть определит целевой признак, значение которого будет вычислять нейронная сеть, и список возможных признаков, влияющих на формирование значения целевого. Данная таблица называется - обучающая выборка. Данное название вытекает из принципов работы нейронной сети, а именно, сначала обучается сеть, в последствие используется. В основе рассматриваемой технологии лежит сервис баз знаний, интернет ориентированная система, о которой будет сказано ниже. Рассматриваемая технология синтеза баз знаний и предполагает несколько этапов.

1. Предварительное преобразование экспертом прикладной базы данных до формата таблицы, пригодной для создания базы знаний. Этот этап называется экспертным, и данная таблица носит название обучающая выборка (ОВ).
2. Проверка правильности содержимого данной таблицы, на которую накладываются определённые ограничения:
 - а. Обоснованности – участие только высококвалифицированных специалистов (экспертов) на этапах построения ОВ.

- b. Полноты – все возможные прецеденты и результаты должны быть представлены в ОВ в достаточном объеме.
- с. Непротиворечивости – ОВ не должна содержать объектов с одинаковыми значениями входных признаков и разными значениями результата, т.е. объектов, принадлежащих к разным классам.
3. Построение по данной таблице нейросетевого представления знаний, заключенных в неё с помощью алгоритма синтеза структуры и настройки параметров нейронной сети.
 4. Проверка адекватности созданной базы знаний с помощью контрольной выборки (КВ). Данная выборка представляет собой таблицу той же структуры, но в данном случае, по ней проверяется правильность формируемых значений целевого признака. Данная выборка не должна содержать данные ОВ, чтобы можно было объективно оценить корректность её работы. Данную оценку должен проводить эксперт данной прикладной области, участвующий на первом этапе при формировании обучающей выборки
 5. Использование и передача базы знаний пользователям, тем кому необходимо получение значения данного целевого признака в данной предметной области по данным признакам.

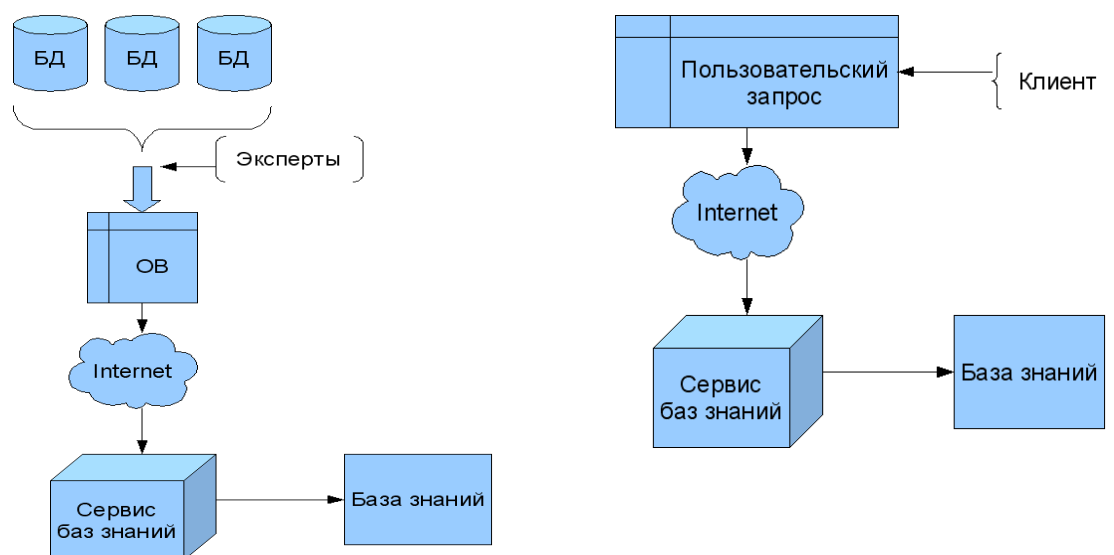


Рис. 3. Схемы экспертного и пользовательского запросов.

Сервис баз знаний

Сервис баз знаний, на использовании которого основана данная технология, представляет собой web-сервис, обращаясь к которому по Internet протоколам, HTTP

или e-mail возможно получение базы знаний, в виде программного продукта или дистанционно доступной программы (web-приложение). Сервис представляет собой клиент-серверную систему, и функционирует в двух режимах:

Первый режим называется Экспертным (ЭР), он включает в себя подготовку исходного набора данных и переработку, трансформацию, обучение и контроль качества полученной базы знаний.

Второй - пользовательский (ПР), он включает этап использования базы знаний. И как следствие, два запроса: экспертный (ЭЗ) и пользовательский. (ПЗ).

Клиентская сторона – сторона, формирующая запрос. Представляет собой персональный компьютер со специализированным программным обеспечением (ПО). Такое ПО может и отсутствовать – в этом случае требуется наличие только почтового клиента или Web-браузера. Клиентское ПО необходимо для удобства создания запроса. В случае эксперта – это заполнение ОВ и КВ, в случае пользователя – это автоматическое формирование полей для выбранной БЗ и последующий ввод значений признаков.

Серверная сторона – сторона со специализированным ПО формирующим ответы на запросы клиента. Она располагается удалённо от пользователя системы.

По схеме прохождения запросов видно, что клиент может получить доступ по:

- HTTP-протоколу
- SMTP-протоколу
- зайти на Web-сервер и с помощью предоставленных серверных приложений сформировать ОВ и КВ непосредственно на сервере;

Вне зависимости от того каким способом пользователь отправляет запрос системе, он проходит четыре стадии:

1. Авторизация (проверка имени пользователя).
2. Ассоциация пользователя с переданными им на сервер файлами.
3. Проверка корректности формата присланного файла и выполнения требований к ОВ, а также формирование отчета в случае возникновения ошибок.
4. Передача файлов на вход блока формирования ответа.

Система позволяет предоставить свою БЗ в общее пользование.

Ответы пользователь может получить по тем же каналам, по которым он формировал запрос: по электронной почте, по HTTP-протоколу или Web интерфейса.

Формирование ответа на ЭЗ происходит за несколько этапов:

1. По ОВ синтезируется архитектура НС и осуществляется ее обучение. Формируется имитационная модель НС.

2. Проверяется качество работы НС на КВ.
3. На основе модели НС строится файл в формате XML, описывающий структуру НС.
4. При необходимости строится представление БЗ.

Формирование ответа системы на ПЗ происходит также за несколько этапов:

1. Читается сформированный экспертом файл описания НС.
2. Формируется имитационная модель НС.
3. Читаются новые значения признаков НС.
4. Моделируется работа НС на заданном кортеже признаков.
5. Формируется ответ на ПЗ в необходимой форме.

Описанная технология служит средством удалённого синтеза нейронного представления знаний и обобщения опыта территориально разделённых специалистов-экспертов (в своей предметной области).

Технология позволяет строить модели принятия интеллектуальных решений независимо от предметной области. При этом она настраивается на любой класс интеллектуальных задач. Это делает технологию универсальным средством автоматического формирования и передачи нейронных БЗ по удалённым БД пользователей.

В задачах робототехники данная технология может быть применима, в качестве удалённого программатора логической зависимости какого-либо целевого параметра от параметров признаков. К примеру, робот накопил определённый объем данных с сенсоров и ему необходимо быстро получить алгоритм вычисления значения какого-либо целевого параметра, он формирует запрос сервису баз знаний и получает в ответ программный модуль готовый к использованию.

Список литературы:

1. Люггер Д.Ф., Искусственный интеллект: стратегии и методы решения сложных проблем, 4-е издание. Пер с англ.: Издательский дом «Вильямс», 2003.-864 с.
2. Каляев А.В., Тимофеев А.В. Методы обучения и минимизации сложности когнитивных нейромодулей супер-макро-нейрокомпьютера с программируемой архитектурой. – Докл. РАН. – 1994, т. 237, с. 180-183.
3. Тимофеев А.В. Методы синтеза диофантовых нейросетей минимальной сложности. – Докл. РАН. – 1995, т. 345, №1, с. 32-35.

4. Timofeev A.V. Parallelism and Self-Organization in Polynomial Neural Networks for Image Recognition.- Proceedings of the 7th International Conference on Pattern Recognition and Image Analysis: New Information Technologies (18–23 October, 2004, St.Petersburg)., p. 97-100, 2004.
5. <http://www.basegroup.ru>
6. <http://www.getinfo.ru>