УДК 62-50

МОДИФИКАЦИЯ АЛГОРИТМА ОБУЧЕНИЯ К2 ДЛЯ ЗАДАЧИ АДАПТАЦИИ БАЙЕСОВСКОЙ НЕЙРОННОЙ СЕТИ

С.А. КАТЕРИНИЧ

Сформулирована задача адаптации байесовской нейронной сети к новым данным. Произведен анализ задачи на соответствие с байесовским подходом с целью получения множества возможных стратегий решения. Для конкретной стратегии предложен алгоритм инкрементной адаптации сети к новым наблюдениям. В его основу положен пакетный алгоритм К2 обучения, на базе которого разработаны критерии, используемые при адаптации.

ВВЕДЕНИЕ

В современном мире информационных технологий для получения знаний широко используется задача обработки больших баз данных (БД). Вот уже более двух десятков лет байесовские нейронные сети (БНС) привлекают внимание исследователей как инструмент обработки БД в решении задач классификации и моделирования процессов, информация о которых собрана в этих базах.

Байесовская нейронная сеть — это вероятностная непараметрическая модель, которая состоит из двух элементов: графической структуры сети и ее вероятностной спецификации. Именно графическая структура, обеспечивающая преимущества при моделировании и визуализации его результата, обусловила большой интерес к данному инструменту. Этот вид моделей является непараметрическим в том смысле, что взаимосвязь между переменными определяется без ссылки на конкретный вид распределения, т.е. конфигурация модели однозначно определяет совместное распределение вероятностей всех переменных модели.

Графическая структура задается ориентированным ациклическим графом, в котором узлы представляют интересующие нас переменные, а дуги от родительских узлов к узлам-потомкам — вероятностные взаимосвязи между данными узлами [1, 2]. Узел X_j является родителем, или непосредственным предком, узла X_i, если существует дуга от X_j к X_i, что трактуется как «X_j непосредственно влечет появление X_i».

Вероятностная спецификация задается распределением условных вероятностей для значений каждого узла X_i сети относительно возможных инициализаций множества Π_i родительских узлов (непосредственных предшественников). Для корневых узлов, т.е. узлов без входных дуг, определяется априорное распределение вероятностей. Таким образом, неопределенность влияния родителей на потомка описывается вероятностями.

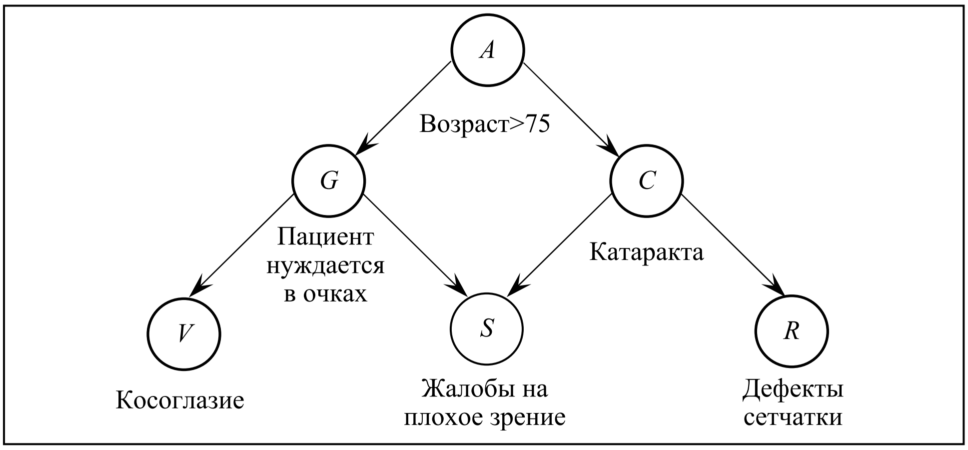
Одной из наиболее привлекательных возможностей такого подхода является графическая интерпретация результатов моделирования, интуитивно понятная лицу, принимающему решения (ЛПР). Графическая интерпретация — это визуализация причинно-следственных взаимосвязей между элементами процесса, моделируемого с помощью БНС (см. рисунок).

Для краткости изложения некоторые условные вероятности не отражены на рисунке, однако их можно без труда получить из указанных, например:

$$P(G = F \mid A = T) = 1 - P(G = T \mid A = T) = 1 - 0{,}75 = 0{,}25\,.$$

Здесь вероятностная спецификация сети неявно задана с помощью таблиц условного распределения вероятностей. Явный вид таблицы условного распределения вероятностей для каждого узла может быть получен на основании выводов, аналогичных приведенному выше. Так, например, для узла S, значения которого обусловлены значениями узлов G и C, таблица будет иметь следующий вид:

G	C	$S=1$	$S=0$
1	1	0,95	0,05
1	0	0,75	0,25
0	1	0,40	0,60
0	0	0,05	0,95



Пример байесовской сети из области дефектов зрения

Условные вероятности событий, соответствующие рисунку:

$$P(A=T)=0{,}10\,;\ P(G=T\,|\,A=T)=0{,}75\,;\ P(G=T\,|\,A=F)=0{,}15\,;$$

$$P(C=T\,|\,A=T)=0{,}40\,;\ P(C=T\,|\,A=F)=0{,}01\,;$$

$$P(V=T\,|\,G=T)=0{,}80\,;\ P(V=T\,|\,G=F)=0{,}05\,;$$

$$P(S=T\,|\,G=T,\ C=T)=0{,}95\,;\ P(S=T\,|\,G=F,\ C=T)=0{,}40\,;$$

$$P(S=T\,|\,G=T,\ C=F)=0{,}75\,;\ P(S=T\,|\,G=F,\ C=F)=0{,}05\,;$$

$$P(R=T\,|\,C=F)=0{,}95\,;\ P(R=T\,|\,C=T)=0{,}25\,.$$

Теоретические аспекты обучения БНС изложены в работе [3].

Различными авторами было предложено несколько подходов к решению задачи построения БНС, каждый из которых воплотился в конкретный алгоритм обучения БНС. Среди них алгоритм на основе принципа минимальной длины описания сети (MDL) [4], алгоритмы обучения байесовских классификаторов [5], инкрементный алгоритм [6] и пр.

Одним из алгоритмов обучения стал К2, предложенный Edward Herskovits в работе [7]. Это пакетный (batch) алгоритм обработки БД наблюдений, т.е. алгоритм целостной обработки всей базы наблюдений в ходе построения структуры сети. Главной особенностью К2 является реализация байесовского подхода к задаче обучения БНС, а также предложенная Herskovits функция оценивания (критерий, функционал качества, scoring function) адекватности структуры БНС данным базы наблюдений. Наиболее интересным свойством этой функции является локальность вычислений, т.е. значение критерия адекватности состоит из значений, вычисляемых для каждого узла в отдельности.

Хотя поиск структуры БНС, используемый в алгоритме К2, реализуется в ускоренном варианте за счет дополнительного требования к подаче входной информации об узлах сети, такой вариант все-таки относится к типу greedy search, т.е. в конкретных случаях его вычислительная трудоемкость может достигать экспоненциальной сложности, так как задача обучения БНС изначально относится к NP-классу.

Таким образом, возникает объективная необходимость разработки модификации данного алгоритма, которая бы реализовывала инкрементный подход для задачи адаптации БНС, полученной в результате пакетной обработки БД алгоритмом К2. Инкрементный подход — это стратегия корректировки структуры построенной модели при получении очередного нового наблюдения с последующим добавлением его к исходной БД, используемой при построении текущей структуры модели.

Среди альтернативных подходов к задаче адаптации БНС наиболее фундаментальными являются направления, базирующиеся на принципе минимальной длины описания сети (MDL) [8] и на функционалах качества, позволяющих сравнить модели распределений из различных по объему БД [9], на концепции частичных и альтернативных теорий [10].

ПОСТАНОВКА ЗАДАЧИ

Необходимо разработать инкрементный алгоритм, позволяющий адаптировать байесовскую нейронную сеть, построенную по заданной базе наблюдений, к новым наблюдениям.

Исходные данные задачи:

- БД наблюдений, соответствующая требованиям предположений алгоритма К2.
- БНС, построенная в результате пакетной обработки базы наблюдений алгоритмом К2.
- База новых наблюдений, не использованных в построении БНС.

ОБОЗНАЧЕНИЯ

Z — множество узлов БНС. Определяется набором переменных наблюдения из БД.

X_i — узел БНС, соответствующий одной переменной наблюдения из БД.

$n = |Z|$ — количество узлов БНС.

r_i — количество значений, принимаемых узлом X_i.

v_{ik} — k-е значение переменной X_i.

Π_i — множество узлов-предков узла X_i.

ϕ_i — множество возможных инициализаций Π_i.

$q_i = |\phi_i|$ — количество возможных инициализаций Π_i.

ϕ_{ij} — j-я инициализация множества узлов-предков Π_i узла X_i.

B_S — структура БНС.

B_P — вероятностная спецификация БНС, т.е. часть описания модели, представляющая вероятностные характеристики БНС.

$\theta_{ijk} = P(X_i = v_{ik} \mid \phi_{ij}, B_P)$, при этом $\sum_k \theta_{ijk} = 1$.

$f(\theta_{ij1}, ..., \theta_{ijr_i})$ — функция плотности распределения вероятностей для узла X_i и инициализации ϕ_{ij}.

D_0 — исходная база данных наблюдений.

S_0 — структура БНС, полученная в результате пакетной обработки базы D_0.

D_1 — БД новых наблюдений, не использованных при построении S_0.

S_1 — структура БНС, полученная после адаптации S_0 к новым данным D_1.

ОБЩИЕ ПОЛОЖЕНИЯ АДАПТАЦИОННОГО ВАРИАНТА АЛГОРИТМА К2

Поскольку в методе адаптации байесовской нейронной сети используется функционал качества, положенного в основу алгоритма К2 обучения байесовской нейронной сети, то необходимо рассмотреть исходные предположения алгоритма К2 [7].

1. Для построения модели используются только переменные из БД. Все переменные дискретные, поэтому понятия «узел» и «переменная» в отношении БНС употребляются как синонимы.

2. Наблюдения, накопленные в БД, появлялись независимо друг от друга. Структура вероятностной модели, по которой могли быть получены эти наблюдения, является неизменной на протяжении всего промежутка наблюдения. В случае, когда реальные наблюдения взаимосвязаны некоторыми характеристиками, можно увеличить количество характеристик наблю-

дения, добавив в БД дополнительные переменные, описывающие эти взаимосвязи.

3. Каждое наблюдение является полным, т.е. содержит информацию обо всех наблюдаемых переменных. Данное требование может быть смягчено [2].

4. Функции плотности распределения вероятностей $f(\theta_{ij1},...,\theta_{ijr_i})$ и $f(\theta_{i_0j_01},...,\theta_{i_0j_0r_i})$ являются независимыми для всех $1 \le i,\ i_0 \le n$, $1 \le j \le q$, $1 \le j_0 \le q_0$, $ij \ne i_0 j_0$. Это предположение представим в виде

$$f(B_P \mid B_S) = \prod_{i=1}^{n} \prod_{j=1}^{q_i} f(\theta_{ij1},...,\theta_{ijr_i}) .$$

5. $f(\theta_{ij1},...,\theta_{ijr_i})$ — равномерное распределение, т.е. $f(\theta_{ij1},...,\theta_{ijr_i}) = C_{ij}$, где $C_{ij} = \text{const}$. Это интерпретируется так: мы не отдаем предпочтения в выборе значений условных вероятностей $\theta_{ij1},...,\theta_{ijr_i}$ какому-то одному из наборов значений.

Далее определим дополнительные аспекты решения задачи адаптации, наследуемые от алгоритма К2, которые определяют как концептуальный подход к задаче адаптации, так и практические вопросы реализации алгоритма адаптации.

1. Задача адаптации байесовской нейронной сети рассматривается с использованием байесовского подхода по аналогии с задачей обучения БНС для алгоритма К2.

2. В роли функционала качества используется функция, положенная в основу К2. Свойство локальности данной функции учитывается при построении алгоритма адаптации.

3. В качестве требования к входной информации сохраняется задание последовательности узлов БНС, определяющей причинно-следственные отношения между узлами.

БАЙЕСОВСКИЙ ПОДХОД К ЗАДАЧЕ АДАПТАЦИИ БНС К НОВЫМ ДАННЫМ

Пусть имеется некоторая база наблюдений. Необходимо предсказать следующие события, которые могли бы быть получены на основании того же распределения, которое описывается накопленной базой наблюдений. Пусть имеются априорные распределения по возможным структурам моделей и по значениям параметров данных моделей. Тогда байесовский подход требует оценки апостериорного распределения на пространстве моделей при условии наблюдения данных, накопленных в базе. В этом случае используются, в частности, преобразования на основе формулы условной вероятности.

Рассмотрим применение такого подхода к задаче адаптации БНС.

Пусть имеется исходная БД наблюдений D_0, по результатам обработки которой получена БНС со структурой S_0.

Пусть в следующий момент времени пришла информация D_1 о новом наблюдении, и необходимо адаптировать сеть S_0, т.е. получить S_1.

Исходная задача построения S_0 формулируется следующим образом [1]:

$$P(S \mid D_0) \to \max_S .$$

В свете байесовского подхода решение задачи обучения БНС имеет вид [1]

$$P(S_0 \mid D_0) = \arg\max_S \frac{P(S)\,P(D_0 \mid S)}{P(D_0)} .$$

При условии индифферентности с точки зрения предпочтений той или иной структуры БНС задача сводится к виду [1]

$$P(S_0 \mid D_0) = \arg\max_S P(D_0 \mid S) .$$

Аналогично задача адаптации БНС может быть представлена как оптимизационная задача

$$P(S \mid D_0, S_0, D_1) \to \max_S .$$

Используя правило цепи (chain rule) для представления совместной вероятности $P(S, D_0, S_0, D_1)$, можно получить в качестве одного из вариантов следующий результат:

$$P(S \mid D_1, D_0, S_0) = \frac{P(D_0)\,P(S \mid D_0)\,P(D_1 \mid S, D_0)\,P(S_0 \mid D_1, S, D_0)}{P(D_0)\,P(S_0 \mid D_0)\,P(D_1 \mid S_0, D_0)} =$$

$$= \frac{P(D_0)\,P(S \mid D_0)\,P(D_1 \mid S, D_0)\,P(S_0 \mid D_0)}{P(D_0)\,P(S_0 \mid D_0)\,P(D_1 \mid S_0, D_0)} = \frac{P(S \mid D_0)\,P(D_1 \mid S, D_0)}{P(D_1 \mid S_0, D_0)} .$$

Таким образом, решение оптимизационной задачи адаптации можно получить в виде

$$P(S_1 \mid D_1, D_0, S_0) = \arg\max_S \frac{P(S \mid D_0)\,P(D_1 \mid S, D_0)}{P(D_1 \mid S_0, D_0)} .$$

Такой вид оптимизируемого выражения может быть использован для определения стратегии адаптации БНС. Рассмотрим каждую составляющую выражения.

Знаменатель — это значение, не зависящее от аргумента оптимизации. Он выполняет некоторую нормирующую функцию, но как таковой для процесса оптимизации интереса не представляет. Однако учтем его логический смысл: знаменатель показывает, насколько вероятно появление наблюдений D_1 при условии истории процесса D_0 и наличия построенной модели этого процесса S_0. Т.е. он может быть интерпретирован как уровень адекватности новых данных уже имеющейся информации о процессе и как вклад новых данных в процесс обучения БНС.

Числитель оптимизируемой дроби — это произведение двух составляющих. Первая совпадает с исходной задачей построения БНС по исходной базе наблюдений. Решением этой задачи была сеть S_0, наиболее вероятная на момент отсутствия дополнительной информации D_1. Вторая составляющая отражает уровень адекватности новых наблюдений некоторой структуре сети. Если учесть ход алгоритма К2, то можно определить возможную тактику построения сети S_1. Для этого проанализируем функционал качества алгоритма К2.

АНАЛИЗ ФУНКЦИОНАЛА КАЧЕСТВА АЛГОРИТМА К2

Функционал качества алгоритма К2 имеет вид

$$P(D \mid B_P) = \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}!,$$

где $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$, N_{ijk} — количество наблюдений из базы данных D, в которых $\{X_i = v_{ik}; \pi_i = \phi_{ij}\}$.

Рассмотрим случай получения одного нового наблюдения, т.е. $|D_1| = 1$. Пусть данное наблюдение инициализирует узел X_i значением v_{ik_0}, а множество узлов-предков Π_i — набором ϕ_{ij_0}. Тогда

$$P(D_1, D_0 \mid S_0) = \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}! =$$

$$= \prod_{i=1}^{n} \frac{(r_i - 1)!}{(N_{ij_0} + r_i - 1)!} \prod_{\substack{j=1 \\ j \neq j_0}}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{\substack{k=1 \\ k \neq k_0}}^{r_i} N_{ijk}! N_{ij_0k_0}! =$$

$$= \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}! \, \frac{1}{N_{ij_0} + r_i} (N_{ij_0k_0} + 1) =$$

$$= P(D_0 \mid S_0) \prod_{i=1}^{n} \frac{N_{ij_0k_0} + 1}{N_{ij_0} + r_i}.$$

С другой стороны,

$$P(D_1, D_0 \mid S_0) = P(D_0 \mid S_0) P(D_1 \mid D_0, S_0).$$

Таким образом, получаем

$$P(D_1 \mid D_0, S_0) = \prod_{i=1}^{n} \frac{N_{ij_0k_0} + 1}{N_{ij_0} + r_i}.$$

Как указывалось ранее, данное значение может быть интерпретировано как вклад новых данных в процесс обучения БНС.

Теперь рассмотрим одну из итераций алгоритма К2 обучения БНС.

Пусть на момент начала текущей итерации построена структура сети S_0.

Пусть на момент окончания текущей итерации построена структура сети S_1 добавлением дуги от X_{i_1} к X_{i_2}. Тогда $P(D_0 \mid S_0) < P(D_0 \mid S_1)$.

Значения функционала качества для этих структур сетей будут отличаться значением множителя

$$\prod_{j=1}^{q_{i_2}} \frac{(r_{i_2}-1)!}{(N_{i_2 j}+r_{i_2}-1)!} \prod_{k=1}^{r_{i_2}} N_{i_2 jk}!$$

При этом $q_{i_2}^{S_1} = q_{i_2}^{S_0} r_{i_1}$.

Рассмотрим влияние дополнительного наблюдения D_1 на значение данного множителя. Пусть новое наблюдение инициирует узел X_{i_2} значением $v_{i_2 k_0}$, а множество узлов-предков Π_{i_2} — набором $\phi_{i_2 j_0}$. Тогда множитель принимает вид

$$\left[\prod_{j=1}^{q_{i_2}} \frac{(r_{i_2}-1)!}{(N_{i_2 j}+r_{i_2}-1)!} \prod_{k=1}^{r_{i_2}} N_{i_2 jk}!\right] \frac{N_{i_2 j_0 k_0}+1}{N_{i_2 j_0}+r_{i_2}}.$$

Обозначим появившийся коэффициент изменения

$$K_{\text{delete}} = \frac{N_{i_2 j_0 k_0}+1}{N_{i_2 j_0}+r_{i_2}}.$$

Данный коэффициент рассмотрим для ретроспективного анализа влияния новых наблюдений на ход обучения БНС.

Таким образом, используя локальность функционала качества алгоритма К2, получаем индикаторы необходимости адаптации той или иной дуги.

АДАПТАЦИОННЫЙ ВАРИАНТ АЛГОРИТМА К2

Адаптацию построенной сети проведем в следующем порядке.

1. Корректировка вероятностной части модели.
2. Корректировка структурной части модели:
   2.1. Удаление несоответствующих дуг.
   2.2. Добавление дуг.

С целью облегчения проведения корректировки вероятностной части модели полезно хранить не таблицы распределения условных вероятностей, а значения N_{ijk}. Это позволит быстрее обновлять данные о распределении условных вероятностей, а их значения вычислять, пользуясь формулой Дирихле

$$P(X_i = v_{ik} \mid \Pi_i = \phi_{ij}) = \frac{N_{ijk} + 1}{N_{ij} + r_i}.$$

При корректировке структуры БНС порядок обхода узлов определим вкладом каждого узла в значение $P(D_1 \mid D_0, S_0) = \prod_{i=1}^{n} \frac{N_{ij_0k_0} + 1}{N_{ij_0} + r_i}$.

Для каждого узла на этапе проверки дуг на необходимость удаления вычисляется значение $K_{\text{delete}}(S_0)$ для текущей конфигурации множества узлов-предков и $K_{\text{delete}}(S_{-1}^{m})$ для конфигураций, получаемых при удалении одной из M входящих в текущий узел дуг, $1 \le m \le M$. Если выполняется условие $K_{\text{delete}}(S_{-1}^{m}) \le K_{\text{delete}}(S_0)$, то m-я дуга остается в структуре сети, так как удаление данной дуги приводит к уменьшению локального значения функционала качества (для текущего узла). Иначе дуга заносится в список дуг, подлежащих проверке на удаление. Список может быть отсортирован по возрастанию значения $K_{\text{delete}}(S_{-1}^{m})$. Дуги из списка рассматриваются последовательно. Дальнейшая проверка заключается в вычислении локального значения функционала качества при исходной конфигурации и конфигурациях, получаемых при удалении одной из дуг, оставшихся в списке.

Если учесть вид решения оптимизационной задачи адаптации БНС, то тактика удаления дуг должна приводить к тому, что первая составляющая числителя $P(S \mid D_0)$ будет уменьшаться, поскольку она достигает максимума при $S = S_0$ как результат начального построения БНС. Таким образом, для получения эффекта от адаптации необходимо потери от удаления дуги компенсировать эффектом от добавления новой дуги.

Имея упорядоченную последовательность узлов (входное требование алгоритма К2), поиск дуги претендента на добавление следует осуществлять в следующем порядке. Оценка дуги производится вычислением локального значения функционала качества. Соответственно, претендент на добавление должен определять конфигурацию входных дуг, имеющую наибольшее локальное значение функционала качества.

После каждого удаления следует попытка добавить одну из возможных дуг набора предков-претендентов, в который также добавляется предок удаленной дуги. Если в результате такой попытки было определено, что удаленная дуга максимально увеличивает локальное значение функционала качества по сравнению с прочими претендентами, то эта дуга возвращается в структуру, и переходим к оценке удаления следующей дуги.

Если после удаления дуги, первой дугой, выбранной для добавления, является не удаленная дуга, то итерация добавления проводится до тех пор, пока не наступит один из исходов:

1) все претенденты добавлены, включая удаленную дугу, добавленную последней;

2) на очередной итерации выбрана ранее удаленная дуга, которая возвращается в структуру;

3) на очередной итерации нельзя повысить локальное значение функционала качества добавлением новой дуги из набора предков-претендентов. Если этот исход наступает при анализе последней дуги из претендентов на удаление, то для данного узла нет надобности проводить процедуру 2.2.

Для узлов, не имеющих дуг претендентов на удаление, производится только проверка необходимости добавления новой дуги. В соответствии с алгоритмом К2 процедура добавления дуг проводится по каждому узлу в порядке его местонахождения во входящей последовательности узлов до тех пор, пока ни одна из оставшихся дуг не будет способна увеличить локальное значение функционала качества.

АНАЛИЗ РЕЗУЛЬТАТОВ АДАПТАЦИИ БНС

Критерием эффективности адаптации может служить соотношение величин $P(D_1, D_0 \mid S_0)$ и $P(D_1, D_0 \mid S_1)$.

Наиболее интересен такой результат адаптации, когда $P(D_1, D_0 \mid S_0) >$ $> P(D_1, D_0 \mid S_1)$, что возможно в случае нарушения исходных предположений, на которые опирается алгоритм К2. Изменение характера моделируемого процесса, т.е. нестационарность потенциальной структуры вероятностной модели, является наиболее значимым фактором, способным вызывать подобный результат. В данном случае наилучшим решением будет построение модели с самого начала на основании дополненной БД.

В принципе, в качестве альтернативы адаптации БНС можно видеть периодическое переобучение БНС по обновляемой БД. Критерием необходимости переобучения может служить $P(D_1 \mid D_0, S_0)$, значение которого интерпретируется как вклад новых данных в процесс обучения БНС. На практике этот вариант может применяться в случаях работы с процессами, характеризующимися высокой интенсивностью поступления новых наблюдений, когда временные затраты на адаптацию несравнимо велики по отношению к усредненному периоду обновления данных.

ВЫВОДЫ

Рассмотрен вариант построения алгоритма инкрементной адаптации ранее построенных БНС к новым данным на основе анализа базовых положений алгоритма К2 обучения БНС.

В результате применения байесовского подхода к задаче адаптации получены возможные стратегии адаптации, одна из которых применена для разработки конкретной процедуры инкрементной адаптации.

Использование декомпозиции функционала качества с последующим ретроспективным анализом позволяет получить критерии, необходимые для процедуры инкрементной адаптации модели к новым данным.

На основании теоретических положений разработанного алгоритма адаптации проведен анализ возможных результатов его работы.

Дальнейшие разработки, на наш взгляд, следует проводить в таких направлениях:

- Смягчение исходных требований алгоритма.
- Разработка алгоритма пакетной адаптации.

Пакетный подход подразумевает такой алгоритм адаптации, при котором принятие решения о необходимой корректировке какой-либо составляющей модели происходит на основе целостного анализа всего пакета новых наблюдений, тогда как при инкрементной адаптации каждое наблюдение из нового пакета анализируется последовательно для определения необходимости изменения структуры модели. В общем случае инкрементный подход — частный случай пакетного, однако при разработке конкретных алгоритмов для конкретных предметных областей инкрементный подход может оказаться предпочтительней.

ЛИТЕРАТУРА

1. *Cooper G.F., Herskovits E.* A Bayesian method for the introduction of probabilistic networks from data // Technical Report KSL-91-02. Knowledge Systems Laboratory. — Stanford: Stanford University, 1991. — 39 p.
2. *Murphy K.* A Brief Introduction to Graphical Models and Bayesian Networks. — http://www.cs.berkeley.edu/~murphyk/Papers/intel.ps.gz.
3. *Heckerman D.* A tutorial on learning Bayesian networks / Technical report MSN-TR-95-06, Microsoft Research, Advanced Technology Division, 1995. — 52 p.
4. *Lam W., Bacchus F.* Learning Bayesian belief networks. An approach based on the MDL principle // Computational intelligence. — 1994. — № 4 — P. 104–127.
5. *Sacha J.P.* New synthesis of Bayesian network classifiers and cardiac SPECT image interpretation. — http://jbnc.sourceforge.net/.
6. *Roure J., Sanguesa R.* Incremental methods for Bayesian network learning. — http://citeseer.ist.psu.edu/.
7. *Herskovits E.* Computer-Based Probabilistic-Network Construction. – http://search.microsoft.com/.
8. *Lam W., Bacchus F.* Using new data to refine Bayesian network // Proc. of the Tenth Conference on Uncertainty in Artificial Intelligence. — 1994. — P. 383–390.
9. *Friedman N., Goldszmidt M.* Sequential update of Bayesian network structure // Proc. of the Thirteenth Conference on Uncertainty in Artificial Intelligence. — 1997. — P. 273–280.
10. *Buntine W.* Theory refinement on Bayesian networks // Proc. of the Seventh Conference on Uncertainty in Artificial Intelligence. — 1991. — P. 52–60.

Поступила 07.04.2006