Лекция 6. **Машина логического вывода и база знаний**

При разработке экспертных систем делались усилия разделить задачу на две части - **машину вывода** и **базу знаний**.

Предполагая, что **машина логического вывода** является универсальной **думающей машиной**, а **база знаний** — это то, **над чем** ей **предстоит думать**.

Пример. Рассмотрим основные принципы построения блока логического вывода на основе байесовского подхода. Пример относится к области медицинской диагностики, однако байесовский подход может быть использован и в других областях.

В базе знаний используется **две разновидности формата для данных**.

Название болезни, p, **число** применимых симптомов **(j, p+, p⁻)**

1. В первом формате хранятся знания о конкретном заболевании:

Первый элемент — название болезни, **Второй** — **априорная вероятность** того, что это заболевание имеется у взятого на удачу члена популяции. При **байесовском подходе** — это **априорная вероятность** рассматриваемого заболевания. **Третий элемент** — число симптомов, которые могут быть использованы либо как признаки этой болезни, либо как противоречащие ей признаки.

Машина логического вывода и база знаний

2. **Второй тип данных** касается симптомов:
Номер симптома, **название** симптома, **вопрос**, который следует задать в отношении симптома

Здесь всего три поля. Первым идет номер симптома; это тот номер, который используется для ссылок на симптом в данных о заболевании. **Вторым** является имя симптома. **В третьем поле** содержится вопрос, который можно будет задать о пользователе системы при попытке определить, проявляется ли этот конкретный симптом у данного пользователя.

В результате ***приходим к некоторой структ*уре ввода данных**
ГРИПП, 0,001; 2; 1; 1; 0,01; 2; 0,9; 0,1
что достаточно хорошо суммирует информацию, которую можно получить о гриппе у любого врача.

В соответствии с этим **имеются следующие элементы базы знаний,** касающихся симптомов:
1. ТЕМПЕРАТУРА, ЕСТЬ ЛИ У ВАС ВЫСОКАЯ ТЕМПЕРАТУРА?
2. НАСМОРК, ЕСТЬ ЛИ У ВАС НАСМОРК?

Полученная **база знаний** представлена на **рис. 1.** На этом этапе **базу знаний** можно легко и быстро **модифицировать**, чтобы **уточнить** некоторые положения.

Машина логического вывода и база знаний

(1) Название болезни, p, число применимых симптомов (j, p^{+}, p^{-})
(2) Номер симптома, название симптома, вопрос, который следует задать

Пример: ГРИПП, 0,001, 2, 1, 1, 0,01, 2, 0,9, 0,1 и связанные с ним два симптома:

1. ТЕМПЕРАТУРА, ЕСТЬ ЛИ У ВАС ВЫСОКАЯ ТЕМПЕРАТУРА?
2. НАСМОРК, ЕСТЬ ЛИ У ВАС НАСМОРК?

Итак, априорная вероятность, что некто болен гриппом, равна 0,001. С гриппом связано два симптома.
Первый симптом – высокая температура. Вероятность высокой температуры при гриппе равна 1.
Вероятность высокой температуры при отсутствии гриппа равна 0,01.
Второй симптом – насморк. Вероятность насморка при гриппе равна 0,9.
Вероятность насморка при отсутствии гриппа равна 0,1.

Рис. 1. Формат данных в базе знаний

Машина логического вывода и база и знаний

Теорема байеса.

Рассматриваемый в примере подход к построению **логического вывода** является **байесовским**.

Однако, байесовский не является **единственно** возможным подходом к. **построению выводов, и**меются и другие, например **методы** классической статистики, **методы** распознавания образов.

В основе теории Байеса лежит предположение, что практически для **любого положения** имеется, какая бы **малая** она не была, **априорная вероятность** того, что данное положение **истинно**.

Наличие априорной вероятности некоторой гипотезы, дает возможность **привлечь** некоторые данные для **оправдания взглядов** по этому вопросу.

При наличии относящихся к делу сведений можно **модифицировать априорную вероятность** так, чтобы получить уже **апостериорную вероятность** той же самой **гипотезы** с учетом **поступивших новых данных**.

Машина логического вывода и база знаний

На рис.2 приведена теорема байеса.

$P(H)$ =априорная вероятность H при отсутствии каких-либо свидетельств
$P(H : E)$ = апостериорная вероятность H при наличии свидетельства E.
По определению:

$$P(H : E) = \frac{P(H \& E)}{P(E)} \text{ и } P(E : H) = \frac{P(E \& H)}{P(H)}$$

После некоторой перегруппировки получаем

$$P(H : E) = \frac{P(E : H) P(H)}{P(E)}$$

и

$$P(E) = P(E : H)\, P(H) + P(E : \text{не } H)\, P(\text{не } H)$$

Машина логического вывода и база знаний

Приведенные выражения дают возможность вычисления **вероятности гриппа** у данного, конкретного пациента.

В данном случае, если **обнаруживается**, что у пациента **жар** необходимо узнать **Р(Н:Е)** — **вероятность** того, что у пациента **грипп**, **при условии**, что у него есть **температура**. Для выполнения необходимых вычислений **предполагаем**, что если у больного **грипп**, то **вероятность у него высокой** температуры **равна единице**.

Для **вычисления Р(Е) – вероятности свидетельства**, в нашем случае **вероятности**, что у человека **имеется высокая температура** используем выражение:

Р(Е) = Р(Е : Н) Р(Н) + Р(Е : не Н) Р(не Н) .

Приведенную формулу можно **трактовать** как **вероятность** того, что у произвольного человека **имеется высокая температура**, которая **равна вероятности температуры** у больного **гриппом**, **умноженной** на **вероятность** гриппа у произвольно взятого человека, **плюс вероятность высокой температуры** у пациента **не больного гриппом, умноженная на** вероятность того, что данное лицо не болеет гриппом

$$P(H:E) = \frac{p^{+}p}{p^{+}p + p^{-}(1-p)}.$$

Е... ым определением базы знаний, в котором величина **Р** **(Е : Н)** — это **р**$^{+}$, связанная с гриппом, а **Р(Е : не Н)** — это **р**$^{-}$ получаем

Машина логического вывода и база знаний

Преимущество байесовского метода проявляется в том, что воспользовавшись **первоначальным значением P(H),** с некоторой **априорной** вероятностью **p,** имеем возможность **вычислить новую** величину **P(H : E)** (т.е. задав, скажем, вопрос о наличии у больного высокой температуры) и воспользоваться значением **P(H : E)** как **обновленным** значением **P(H**).

При таком подходе **процесс можно повторять** несколько раз **увеличивая** или **уменьшая вероятность** заболевания, но каждый раз обращаясь **к одной и той же формуле Байеса**, **подставляя** в нее каждый раз **новую** априорную вероятность, получаемую **из апостериорной вероятности**, имевшейся на **предыдущем** шаге.

Собрав все сведения, касающиеся всех гипотез, можем прийти к окончательному заключению, выведя, что верная гипотеза является истинной.

Машина логического вывода и база знаний

Однако относительно данного метода высказываются **критические высказывания**, связанные с тем, что рассмотренное **вычисление** является **точным только** в том случае, если **известно P(E : не H),** а это **значение не всегда** может быть точно **известно**.

Одна из возможностей обойти это затруднение заключается в использовании общей формулы для вероятности некоторого свидетельства

$$P(E) = \sum_{i=1}^{n} P(E : H_i)\, P(H_i)$$

с учетом того, что формула, использованная ранее, является частным случаем этого выражения.

Если **P(H)** постоянно уточняется в ходе работы, то **P(E)** теоретически также может уточняться с течением времени.

Машина логического вывода и база знаний

Второе критическое замечание, которое может быть выдвинуто, касается того, что в этом методе используется предположение о независимости переменных. С точки зрения теории это замечание очень серьезно.

Использование предположения о независимости переменных вызывает определенные трудности потому, что большинство симптомов (или "показаний", если речь идет о других областях) всегда в какой-то степени коррелируют друг с другом. И из этого положения нет никакого "теоретически приемлемого" выхода.

Однако, если для обоснования истинности решения необходимо знать **порядок отношения вероятностей**, то вся проблема, по-видимому, **исчезнет**.

До тех пор пока информация в базе знаний одинаково ошибочна в отношении каждого ее элемента и числа свидетельств, имеющихся для каждой гипотезы, считается, что и **относительный порядок ошибок** в целом примерно **одинаков**.

Машина логического вывода и база знаний

В статье Крис Нейлор [Экспертные системы. Принципы работы и ПРИМЕРЫ. М.: “Радио и связь”,1987] отмечается, что высказывание можно сделать более сильным, если оценивать высказывание с позиции использования шансов.

Шансы и вероятности связаны следующей формулой:

$$O(H) = \frac{P(H)}{1 - P(H)},$$

в этом случае гипотезы с вероятностью 0,5 шансы равны.

Некоторые отдают предпочтение вычислению байесовской суммы с использованием шансов, а не вероятностей только потому, что это оказывается проще для вычислений. Переходя к шансам, получаем из предыдущих формул

$$O(H : E) = \frac{P(E : H)}{P(E : \text{не } H)} O(H) = \frac{p^{+}}{p^{-}} O(H).$$

В этом случае, шансы являются простыми функциями P(E : H) и P(E : не H).