Министерство образования Республики Беларусь

Учреждение образования

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Факультет информационных технологий и управления

Кафедра информационных технологий автоматизированных систем

Отчёт

по практической работе №7

«Байесовская стратегия оценки выводов»

по дисциплине «Экспертные Системы»

|  |  |
| --- | --- |
| Выполнил: | Проверила: |
| студент гр. 820601 Шведов А.Р. | Герман Ю. О. |
|  |  |

Минск 2022

# Цель работы

Изучение вероятностных методов оценки достоверности выводов в ЭС. По выданному заданию рассчитать вероятность указанной гипотезы на основе байесовской стратегии оценки.

# Теоретическая часть

Байесовская стратегия оценки выводов – одна из стратегий, применяемых для оценки достоверности выводов в ЭС, например, заключений продукционных правил. Основная идея байесовской стратегии заключается в оценке вероятности некоторого вывода с учётом фактов, подтверждающих или опровергающих этот вывод.

Формулировка теоремы Байеса, известная из теории вероятностей, следующая.

Пусть имеется *n* несовместных событий *H*1, *H*2, ..., *Hn*. Несовместность событий означает, что никакиеиз событий *H*1, *H*2, ..., *H*n не могут произойтивместе(другими словами, вероятности их совместного наступления равны нулю). Известны вероятности этих событий: *P*(*H*1), *P*(*H*2), ..., *P*(*Hn*), причём *P*(*H*1)+*P*(*H*2)+...+*P*(*Hn*)=1; это означает, что события *H*1, *H*2, ..., *Hn* образуют полную группусобытий, то есть одноиз них происходит обязательно. С событиями *H*1, *H*2, ..., *Hn* связано некоторое событие *E*. Известны вероятности события *E* при условияхтого, что какое-либо из событий *H*1, *H*2, ..., *Hn* произошло: *P*(*E*/*H*1), *P*(*E*/*H*2), ..., *P*(*E*/*Hn*). Пусть известно, что событие *E* произошло. Тогда вероятность того, что какое-либо из событий *Hi* (*i*=1, ..., *n*) произошло, можно найти по следующей формуле (формула Байеса):



События *H*1, *H*2, ..., *Hn* называются гипотезами, а событие *E* – свидетельством. Вероятности гипотез *P*(*Hi*) без учёта свидетельства (то есть без учёта того, произошло событие *E* или нет) называются доопытными (априорными), а вероятности *P*(*H*i/*E*) – послеопытными (апостериорными). Величина *P*(*EHi*) – совместнаявероятность событий *E* и *Hi*, то есть вероятность того, что произойдут оба события вместе. Величина *P*(*E*) – полная (безусловная) вероятность события *E*.

Формула Байеса позволяет уточнять вероятность гипотез с учётом новой информации, то есть данных о событиях (свидетельствах), подтверждающих или опровергающих гипотезу.

В ЭС формула Байеса может применяться для оценки вероятностей заключений продукционных правил на основе данных о достоверности их посылок. Заключения в этом случае соответствуют гипотезам в теореме Байеса, а посылки – свидетельствам. Обычно посылка правила в ЭС содержит несколько условий. Вероятности *P*(*Hi*) и *P*(*E*/*Hi*) определяются на основе статистических данных с использованием формул теории вероятностей. Основные из этих формул следующие.

Формула умножения вероятностей(произойдёт и событие *A*, и *B*):

*P*(*AB*)=*P*(*A*)*P*(*B*/*A*)=*P*(*B*)*P*(*A*/*B*), где *P*(*A*), *P*(*B*) – вероятности событий *A* и *B*; *P*(*B*/*A*) – условная вероятность события *B*, то есть вероятность события *B* при условии, что произошло событие *A*; *P*(*A*/*B*) – условная вероятность события *A*, то есть вероятность события *A* при условии, что произошло событие *B*.

Если события A и B независимы(то есть вероятность одного события не зависит от того, произошло ли другое событие), то формула умножения вероятностей записывается следующим образом: *P*(*AB*)=*P*(*A*)*P*(*B*).

Формула умножения вероятностей для нескольких событий(вероятность того, что произойдут всеуказанные события вместе):

*P*(*A*1*A*2...*An*)=*P*(*A*1) *P*(*A*2/*A*1) *P*(*A*3/*A*1,*A*2) ... *P*(*An*/*A*1,*A*2,...,*An*-1).

Формула сложения вероятностей(вероятность того, что произойдет хотя бы одноиз событий): *P*(*A*+*B*)=*P*(*A*)+*P*(*B*)-*P*(*AB*).

Если события A и B несовместны(то есть не могут произойти вместе), то *P*(*AB*)=0, и формула сложения вероятностей принимает следующий вид: *P*(*A*+*B*)=*P*(*A*)+*P*(*B*).

Формула сложения вероятностей для нескольких событий обычно записывается следующим образом:



где  – вероятность того, что не произойдёт ни одного из событий *A*1, *A*2, ..., *An*. Эту величину можно найти, например, по формуле умножения вероятностей.

# Ход работы

Рассматривается ЭС, используемая для анализа данных геологической разведки и принятия решения о бурении скважин. Решение принимается на основе информации о содержании в пробах трех веществ (В1, В2, В3). Имеются статистические данные о результатах 120 бурений (из них в 85 случаях было обнаружено полезное ископаемое). Результаты бурений представлены на рисунке 1.

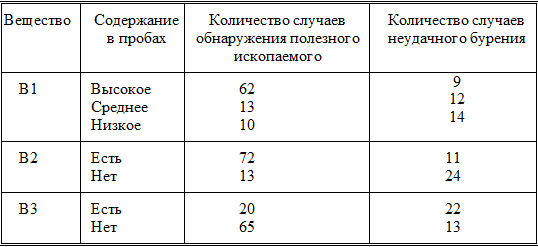


Рисунок 1 – Результаты бурений

В пробе обнаружено высокое содержание вещества В1; вещество В2 обнаружено, В3 - не обнаружено. Оценить вероятность того, что при бурении будет обнаружено полезное ископаемое.

В данном случае, гипотезы следующие: *H*1 – обнаружено полезное ископаемое, *H*2 – полезное ископаемое не обнаружено. Свидетельствами здесь являются вещества В1, В2, В3. Обозначим их как *E*1, *E*2, *E*3.

Определим вероятности, необходимые для расчётов по формуле Байеса. Априорные вероятности гипотез (вероятности удачного или неудачного бурения):

*P(H1) = 85/120=0,708;*

*P(H2) = 35/120=0,292.*

Далее находим величины, необходимые для формулы умножения вероятностей:

*P(E1/H1) = 62/85 = 0,729;*

*P(E2/H1) = 72/85 = 0,847;*

*P(E3/H1) = 65/85 = 0,765;*

*P(E1/H2) = 9/35 = 0,257;*

*P(E2/H2) = 11/35 = 0,314;*

*P(E3/H2) = 13/35 = 0,371;*

При подставке данных значений в формулу умножения вероятностей получаем:

*P(E/H1) = P(E1/H1)\*P(E2/H1)\*P(E3/H1) = 0,472;*

*P(E/H2) = P(E1/H2)\*P(E2/H2)\*P(E3/H2) = 0,030;*

Вычислим апостериорную вероятность:

*P(H1/E) = (P(E/H1)\*P(H1))/(( P(E/H1)\*P(H1))+(P(E/H2)\*P(H2)) = 0,975;*

Полученная апостериорная вероятность (0,975) больше, чем априорная (0,708). Это означает, что при попытке бурения полезное ископаемое будет найдено. Результаты вычислений представлены на рисунке 2.

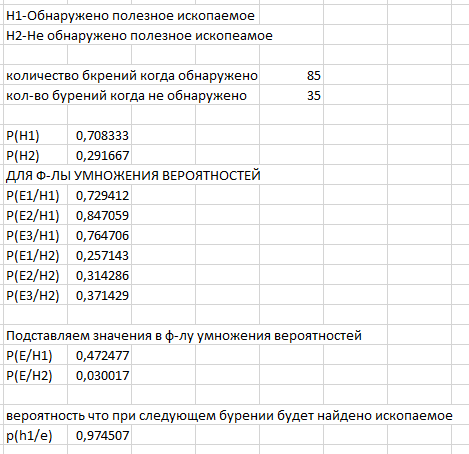


Рисунок 2 – Результаты вычислений

# вывод

В ходе успешно проделанной работы были получены навыки применения байесовской стратегии оценки выводов.