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:
студент гр. 820601
Шведов А.Р.

Проверила:
Герман Ю. О.

Минск 2022

1 ЦЕЛЬ РАБОТЫ

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

2 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, ..., H_n. Несовместность событий означает, что никакие из событий H_1, H_2, ..., H_n не могут произойти вместе (другими словами, вероятности их совместного наступления равны нулю). Известны вероятности этих событий: $P(H_1)$, $P(H_2)$, ..., $P(H_n)$, причём $P(H_1)+P(H_2)+...+P(H_n)=1$; это означает, что события H_1, H_2, ..., H_n образуют полную группу событий, то есть одно из них происходит обязательно. С событиями H_1, H_2, ..., H_n связано некоторое событие E. Известны вероятности события E при условиях того, что какое-либо из событий H_1, H_2, ..., H_n произошло: $P(E/H_1)$, $P(E/H_2)$, ..., $P(E/H_n)$. Пусть известно, что событие E произошло. Тогда вероятность того, что какое-либо из событий H_i (i=1, ..., n) произошло, можно найти по следующей формуле (формула Байеса):

$$P(H_i / E) = \frac{P(E / H_i)P(H_i)}{P(E / H_1)P(H_1) + P(E / H_2)P(H_2) + \ldots + P(E / H_n)P(H_n)} = \frac{P(EH_i)}{P(E)}.$$

События H_1, H_2, ..., H_n называются гипотезами, а событие E – свидетельством. Вероятности гипотез $P(H_i)$ без учёта свидетельства (то есть без учёта того, произошло событие E или нет) называются доопытными (априорными), а вероятности $P(H\text{i}/E)$ – послеопытными (апостериорными). Величина $P(EH_i)$ – совместная вероятность 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 $P(H_i)$ и $P(E/H_i)$ 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 А и В 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 все указанные события вместе):

$P(A_1A_2...A_n)=P(A_1)\ P(A_2/A_1)\ P(A_3/A_1,A_2)\ ...\ P(A_n/A_1,A_2,...,A_n\text{-}1)$.

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

Если события А и В 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:

$$P(A_1 + A_2 + ... + A_n) = 1 - P(\overline{A_1} + \overline{A_2} + ... + \overline{A_n}),$$

где $P(\overline{A_1} + \overline{A_2} + ... + \overline{A_n})$ – вероятность того, что не произойдёт ни одного из событий A_1, A_2, ..., A_n. Эту величину можно найти, например, по формуле умножения вероятностей.

3 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.

Вещество	Содержание в пробах	Количество случаев обнаружения полезного ископаемого	Количество случаев неудачного бурения
В1	Высокое Среднее Низкое	62 13 10	9 12 14
В2	Есть Нет	72 13	11 24
В3	Есть Нет	20 65	22 13

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):

$P(H_1) = 85/120=0,708;$

$P(H_2) = 35/120=0,292.$

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.

H1-Обнаружено полезное ископаемое	
H2-Не обнаружено полезное ископеамое	
количество бкрений когда обнаружено	85
кол-во бурений когда не обнаружено	35
P(H1)	0,708333
P(H2)	0,291667
ДЛЯ Ф-ЛЫ УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ	
P(E1/H1)	0,729412
P(E2/H1)	0,847059
P(E3/H1)	0,764706
P(E1/H2)	0,257143
P(E2/H2)	0,314286
P(E3/H2)	0,371429
Подставляем значения в ф-лу умножения вероятностей	
P(E/H1)	0,472477
P(E/H2)	0,030017
вероятность что при следующем бурении будет найдено ископаемое	
p(h1/e)	0,974507

Рисунок 2 – Результаты вычислений

ВЫВОД

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