

Лабораторна робота 2. Байєсовський аналіз та мережі Байєса

Згідно з байєсовським аналізом апіорні дані поєднуються з апостеріорними. Події, що відбивають дію «причин», називають гіпотезами, оскільки вони – гадані події, що спричинили дане. Безумовну імовірність справедливості гіпотези називають апіорною (наскільки ймовірна причина взагалі), а умовну - з урахуванням факту події, що відбулася, - апостеріорної (наскільки ймовірна причина виявилася з урахуванням даних про подію).

Формула Байєса виглядає наступним чином:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (2.1)$$

Де $P(A)$ - апіорна ймовірність гіпотези A ,

$P(A|B)$ - ймовірність гіпотези A при настанні події B (апостеріорна ймовірність),

$P(B|A)$ - ймовірність настання події B при істинності гіпотези A ,

$P(B)$ - повна ймовірність настання події B .

По суті формула Байєса витікає з визначення умовної ймовірності. Ймовірність сумісної події $P(AB)$ виражається через умовні ймовірності:

$$\begin{aligned} P(AB) &= P(A|B)P(B) = P(B|A)P(A) \\ \Rightarrow P(A|B) &= \frac{P(AB)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} \end{aligned} \quad (2.1)$$

Мережа Байєса представляє собою графічне відображення змінних та їх імовірнісних характеристик. Дана графічна модель складається з вузлів (змінних) та зв'язків між ними. Батьківський вузол представляє змінну, яка впливає на дочірню змінну.

Мережі Байєса використовують у діагностиці, аналізі зображень, генетиці, економічних дослідженнях, пошукових алгоритмах та інших галузях, які потребують дослідження невідомих змінних та зв'язків між даними.

Основними етапами дослідження є:

- Визначення змінних;
- Визначення причинних-наслідкових зв'язків;
- Визначення умовних та апіорних ймовірностей;
- Використання фактичних даних;
- Оновлення оцінок;
- Визначення апостеріорних довірчих оцінок.

Приклад 1. Класичним прикладом, що демонструє ефективність використання Байєсовського аналізу є перевірка достовірності тесту на вживання наркотичних речовин.

Наприклад, тест на вживання наркотичних речовин має чутливість 99% та специфічність 99% (тест даватиме 99% коректних позитивних результатів для тих, хто вживає наркотичні речовини, і 99% коректних негативних результатів для тих, хто не вживає. Тобто лише у 1% випадків тест дає неправильний результат і для тих хто вживає, і для тих хто не вживає наркотичні речовини).

Згідно з дослідженнями, приблизно 0.5% людей вживають наркотичні речовини. Завдання: Визначити, якою є ймовірність, що людина вживає наркотичні речовини, якщо при проходженні тесту він дав позитивний результат?

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0,99 \times 0,005}{0,99 \times 0,005 + 0,01 \times 0,995} \approx 33,2\%$$

Незважаючи на те, що на перший погляд тест виглядає досить точним, у результаті при отриманні позитивного тесту ймовірність того що людина не

вживає наркотичні речовини більша, ніж ймовірність того що вона їх вживає. Таким чином, при оцінці ризиків обов’язково враховувати базові рівні. У даному прикладі отримано несподіваний на перший погляд результат через те, що знехтували загальною ймовірністю отримати результати від людини, що вживає наркотичні речовини.

Тобто, якщо перевірити тисячу осіб, то з них тільки 5 будуть вживати наркотичні речовини, а 995 – ні. Із тих хто не вживає можна отримати приблизно 10 хибних результатів:

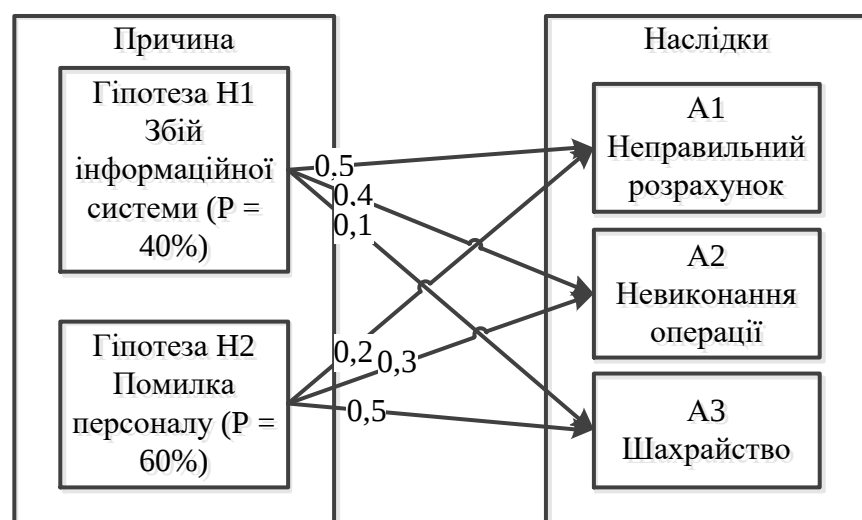
$$995 \times 0,01 = 9,95 \approx 10$$

Правильних результатів від тих хто вживає буде отримано приблизно 5:

$$5 \times 0,99 = 4,95 \approx 5$$

Тобто, із 15 отриманих позитивних результатів 5 будуть правильними, а 10 хибними.

Приклад 2. Системний адміністратор оцінює, що операційний ризик можливий або внаслідок збою інформаційних систем (з ймовірністю 40%), або внаслідок помилки персоналу (з ймовірністю 60%).



Наслідками операційного ризику може бути: неправильний розрахунок, шахрайство чи невиконання операції. Імовірність кожної з цих подій визначена для кожного з факторів операційного ризику на рисунку у вигляді мережі Байєса.

Наприклад, виникла помилка що призвела до неправильного розрахунку. Як зміниться ймовірність того, що це вийшло в результаті збою інформаційної системи або дій співробітників? Для вирішення цього завдання скористуємося формулою Байєса для знаходження умовної імовірності:

$$\begin{aligned} P(H1 | A1) &= \frac{P(H1)P_{H1}(A1)}{P(H1)P_{H1}(A1) + P(H2)P_{H2}(A1)} = \\ &= \frac{0,4 \times 0,5}{0,4 \times 0,5 + 0,6 \times 0,2} = 0,625 \end{aligned}$$

Таким чином, імовірність того, що помилка викликана збоєм інформаційної системи зростає до 62,5%.

У більш складних завданнях, коли створюються більш детальні, багаторівневі причинно-наслідкові мережі, такий аналіз може бути дуже корисним для аналізу причин збоїв, що відбулися. За допомогою подібних моделей можна з більшою ймовірністю визначити місце виникнення помилки, тому що в рамках такої задачі відтворюється вся послідовність бізнес-процесу.

Перевагами методу:

- Для використання методу достатньо знання апіорної інформації;
- логічно виведені твердження легкі розуміння;
- Застосування методу засноване на формулі Байєса;
- метод є способом використання суб'єктивних ймовірнісних оцінок.

Недоліки методу:

- Визначення всіх взаємодій в мережах Байєса для складних систем не завжди здійснено;

- підхід Байєса вимагає знання безлічі умовних ймовірностей, які зазвичай одержують експертними методами. Застосування програмного забезпечення засноване на експертних оцінках.

ЗАВДАННЯ.

Варіант 1.

Статистика запитів кредитів у банку така: 10% – державні органи, 85% – інші банки, 5% – фізичні особи. Імовірності неповернення взятого кредиту, відповідно, 0,01, 0,02, 0,2. Знайти можливість чергового запиту на кредит. Якою є ймовірність, що заданий кредит не повернула фізична особа?

Варіант 2.

З наявних на складі серверів 30% виготовлені фірмою 1, інші – фірмою 2. Ймовірність того, що сервер, виготовлений фірмою 1, не вийде з експлуатації протягом гарантійного терміну служби, дорівнює 0,9, для серверів, виготовлених фірмою 2, ця ймовірність дорівнює 0,8. Випадковим чином для перевірки зі складу вибрали сервер, який витримав гарантійний термін. Визначити ймовірність того, що його було вироблено першою фірмою.

Варіант 3.

Розглядаються акції певної компанії. Якщо економіка країни буде на підйомі, то ймовірність зростання вартості акцій цієї компанії в наступному році дорівнюватиме 0,7; якщо ж розвиток економіки не буде успішним, то ця ймовірність дорівнює 0,2. На думку аналітиків, ймовірність зростання

економіки наступного року дорівнює 0,85. Оцінити ймовірність того, що наступного року акції компанії піднімуться в ціні.

Варіант 4.

Співвідношення кількості чоловіків, які переглядають рекламу на сайті, до жінок, які переглядають рекламу на цьому ж сайті, як 3:2. Імовірність того, що за рекламним посиланням перейде чоловік дорівнює 0,1; для жінки ця ймовірність дорівнює 0,2. За рекламним посиланням перейшов потенційний клієнт. Знайти ймовірність, що це чоловік.

Варіант 5.

Ви купили акцію підприємства, яке може виявитися «перспективним» або «неперспективним» з ймовірністю 0,5. Акція «перспективного» підприємства з ймовірністю 1 за рік зросте в ціні. Акція «неперспективного» підприємства через рік з ймовірністю $1/4$ виросте, а з ймовірністю $3/4$ впаде в ціні. За рік ціна купленої акції зросла. Якою є ймовірність того, що це акція «перспективного» підприємства?

Варіант 6.

Статистика запитів кредитів у банку така: 20% – державні органи, 75% – інші банки, 5% – фізичні особи. Імовірності неповернення взятого кредиту, відповідно, 0,01, 0,02, 0,2. Знайти можливість чергового запиту на кредит. Якою є ймовірність, що заданий кредит не повернула фізична особа?

Варіант 7.

З наявних на складі серверів 40% виготовлені фірмою 1, інші – фірмою 2. Ймовірність того, що сервер, виготовлений фірмою 1, не вийде з

експлуатації протягом гарантійного терміну служби, дорівнює 0,9, для серверів, виготовлених фірмою 2, ця ймовірність дорівнює 0,8. Випадковим чином для перевірки зі складу вибрали сервер, який витримав гарантійний термін. Визначити ймовірність того, що його було вироблено першою фірмою.

Варіант 8.

Розглядаються акції певної компанії. Якщо економіка країни буде на підйомі, то ймовірність зростання вартості акцій цієї компанії в наступному році дорівнюватиме 0,7; якщо ж розвиток економіки не буде успішним, то ця ймовірність дорівнює 0,2. На думку аналітиків, ймовірність зростання економіки наступного року дорівнює 0,75. Оцінити ймовірність того, що наступного року акції компанії піднімуться в ціні.

Варіант 9.

Співвідношення кількості чоловіків, які переглядають рекламу на сайті, до жінок, які переглядають рекламу на цьому ж сайті, як 5:3. Ймовірність того, що за рекламним посиланням перейде чоловік дорівнює 0,1; для жінки ця ймовірність дорівнює 0,2. За рекламним посиланням перейшов потенційний клієнт. Знайти ймовірність, що це чоловік.

Варіант 10.

Ви купили акцію підприємства, яке може виявитися «перспективним» або «неперспективним» з ймовірністю 0,5. Акція «перспективного» підприємства з ймовірністю 1 за рік зросте в ціні. Акція «неперспективного» підприємства через рік з ймовірністю $\frac{1}{5}$ виросте, а з ймовірністю $\frac{4}{5}$ впаде в ціні. За рік ціна купленої акції зросла. Якою є ймовірність того, що це акція «перспективного» підприємства?

Варіант 11.

Статистика запитів кредитів у банку така: 10% – державні органи, 85% – інші банки, 5% – фізичні особи. Імовірності неповернення взятого кредиту, відповідно, 0,005, 0,02, 0,2. Знайти можливість чергового запиту на кредит. Якою є ймовірність, що заданий кредит не повернула фізична особа?

Варіант 12.

З наявних на складі серверів 30% виготовлені фірмою 1, інші – фірмою 2. Ймовірність того, що сервер, виготовлений фірмою 1, не вийде з експлуатації протягом гарантійного терміну служби, дорівнює 0,8, для серверів, виготовлених фірмою 2, ця ймовірність дорівнює 0,7. Випадковим чином для перевірки зі складу вибрали сервер, який витримав гарантійний термін. Визначити ймовірність того, що його було вироблено першою фірмою.

Варіант 13.

Розглядаються акції певної компанії. Якщо економіка країни буде на підйомі, то ймовірність зростання вартості акцій цієї компанії в наступному році дорівнюватиме 0,7; якщо ж розвиток економіки не буде успішним, то ця ймовірність дорівнює 0,2. На думку аналітиків, ймовірність зростання економіки наступного року дорівнює 0,89. Оцінити ймовірність того, що наступного року акції компанії піднімуться в ціні.

Варіант 14.

Співвідношення кількості чоловіків, які переглядають рекламу на сайті, до жінок, які переглядають рекламу на цьому ж сайті, як 4:2. Імовірність того, що за рекламним посиланням перейде чоловік дорівнює 0,1; для жінки ця ймовірність дорівнює 0,2. За рекламним посиланням перейшов потенційний клієнт. Знайти ймовірність, що це чоловік.

Варіант 15.

Ви купили акцію підприємства, яке може виявитися «перспективним» або «неперспективним» з ймовірністю 0,5. Акція «перспективного» підприємства з ймовірністю 1 за рік зросте в ціні. Акція «неперспективного» підприємства через рік з ймовірністю $\frac{3}{4}$ виросте, а з ймовірністю $\frac{1}{4}$ впаде в ціні. За рік ціна купленої акції зросла. Якою є ймовірність того, що це акція «перспективного» підприємства?

Варіант 16.

Статистика запитів кредитів у банку така: 30% – державні органи, 65% – інші банки, 5% – фізичні особи. Імовірності неповернення взятого кредиту, відповідно, 0,01, 0,02, 0,2. Знайти можливість чергового запиту на кредит. Якою є ймовірність, що заданий кредит не повернула фізична особа?

Варіант 17.

З наявних на складі серверів 50% виготовлені фірмою 1, інші – фірмою 2. Ймовірність того, що сервер, виготовлений фірмою 1, не вийде з експлуатації протягом гарантійного терміну служби, дорівнює 0,9, для серверів, виготовлених фірмою 2, ця ймовірність дорівнює 0,8. Випадковим чином для перевірки зі складу вибрали сервер, який витримав гарантійний термін. Визначити ймовірність того, що його було вироблено першою фірмою.

Варіант 18.

Розглядаються акції певної компанії. Якщо економіка країни буде на підйомі, то ймовірність зростання вартості акцій цієї компанії в наступному році дорівнюватиме 0,7; якщо ж розвиток економіки не буде успішним, то ця ймовірність дорівнює 0,3. На думку аналітиків, ймовірність зростання економіки наступного року дорівнює 0,65. Оцінити ймовірність того, що наступного року акції компанії піднімуться в ціні.

Варіант 19.

Співвідношення кількості чоловіків, які переглядають рекламу на сайті, до жінок, які переглядають рекламу на цьому ж сайті, як 5:4. Імовірність того, що за рекламним посиланням перейде чоловік дорівнює 0,1; для жінки ця ймовірність дорівнює 0,2. За рекламним посиланням перейшов потенційний клієнт. Знайти ймовірність, що це чоловік.

Варіант 20.

Ви купили акцію підприємства, яке може виявитися «перспективним» або «неперспективним» з ймовірністю 0,4 та 0,6 відповідно. Акція «перспективного» підприємства з ймовірністю 1 за рік зросте в ціні. Акція «неперспективного» підприємства через рік з ймовірністю $\frac{1}{5}$ виросте, а з ймовірністю $\frac{4}{5}$ впаде в ціні. За рік ціна купленої акції зростає. Якою є ймовірність того, що це акція «перспективного» підприємства?