

Modele Bayesowskie

Opis i implementacja sieci Bayesowskiej

Wstęp

Modele Bayesowskie umożliwiają wnioskowanie probabilistyczne w warunkach niepewności za pomocą skierowanych grafów acyklicznych, odzwierciedlających zależności przyczynowo-skutkowe między zmiennymi. Znajdują zastosowanie m.in. w systemach wspomagania decyzji i sztucznej inteligencji, a ich kluczową zaletą jest efektywne wyznaczanie prawdopodobieństw na podstawie niepełnych lub niepewnych danych wejściowych.

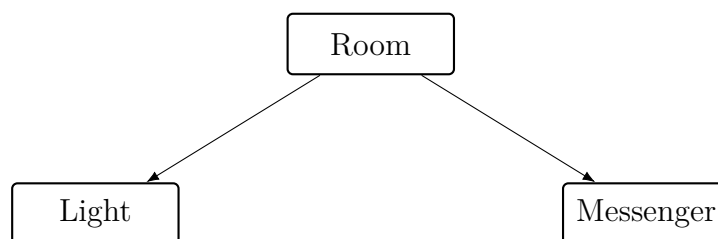
1. Architektura sieci Bayesowskiej

1.1 Rozpatrywany problem

Na potrzeby opisu sieci rozważmy następujący problem:

Pracownik spędza 40% czasu pracy w swoim pokoju na uczelni. Pozostałe 60% czasu pracuje zdalnie. Kiedy jest w swoim pokoju, połowę czasu ma wyłączone światło. Gdy nie ma go w pokoju, zostawia włączone światło tylko w 5% przypadków. 80% czasu, gdy jest w swoim biurze, jest zalogowany na komunikatorze. Ponieważ czasami loguje się z domu, w 5% przypadków, gdy nie ma go na uczelni, nadal jest zalogowany na komunikatorze.

Zależności występujące w zadaniu można przedstawić za pomocą sieci Bayesowskiej pokazanej poniżej.



W rozważanym modelu zdefiniowano trzy zmienne losowe:

- **Room** - informacja o miejscu pracy pracownika (0 - praca zdalna, 1 - biuro),
- **Light** - stan oświetlenia pokoju (0 - światło wyłączone, 1 - światło włączone),
- **Messenger** - status w komunikatorze (0 - offline, 1 - online).

Przyjęto założenie, że zmienna **Room** stanowi wspólną przyczynę dla pozostałych zmiennych. Obecność pracownika w biurze wpływa zarówno na prawdopodobieństwo włączenia światła, jak i na prawdopodobieństwo zalogowania do komunikatora. Zależności te zostały przedstawione za pomocą skierowanych krawędzi w sieci Bayesowskiej. Pozwala to na zwartą reprezentację rozkładu prawdopodobieństwa oraz efektywne wykonywanie wnioskowania probabilistycznego.

1.2 Tablice prawdopodobieństw warunkowych

Sama struktura sieci Bayesowskiej określa jedynie zależności pomiędzy zmiennymi losowymi. Aby możliwe było wykonywanie obliczeń probabilistycznych, konieczne jest przypisanie każdemu węzłowi odpowiedniego rozkładu prawdopodobieństwa.

W sieciach Bayesowskich informacje te przechowywane są w postaci tablic prawdopodobieństw warunkowych (*Conditional Probability Tables, CPT*). Każda tablica określa prawdopodobieństwo przyjęcia określonej wartości przez zmienną przy zadanych wartościach jej rodziców w grafie. Dla każdego możliwego zestawu wartości rodziców należy określić rozkład prawdopodobieństwa zmiennej potomnej. W praktyce oznacza to, że dla każdej konfiguracji rodziców definiowany jest osobny rozkład prawdopodobieństwa opisujący zachowanie rozważanej zmiennej.

Ponieważ zmienna **Room** nie posiada rodziców w sieci Bayesowskiej, jej zachowanie opisuje rozkład a priori. Rozkład ten określa nasze początkowe przekonanie o miejscu pracy pracownika przed uwzględnieniem jakichkolwiek dodatkowych informacji, takich jak stan światła czy status w komunikatorze:

<i>Room</i>	0	1
$P(Room)$	0.6	0.4

Dla zmiennej **Light** rozkład zależy od wartości zmiennej **Room**:

	<i>Room</i> = 0	<i>Room</i> = 1
$P(Light = 0)$	0.95	0.50
$P(Light = 1)$	0.05	0.50

Dla zmiennej **Messenger** rozkład zależy od wartości zmiennej **Room**:

	<i>Room</i> = 0	<i>Room</i> = 1
$P(Messenger = 0)$	0.95	0.20
$P(Messenger = 1)$	0.05	0.80

Można zauważyć, że zarówno zmienna **Light**, jak i **Messenger** posiadają jednego rodzica o dwóch możliwych wartościach. W rezultacie ich tablice CPT zawierają po dwa warunkowe rozkłady prawdopodobieństwa. W ogólnym przypadku rozmiar tablicy CPT rośnie wykładniczo wraz z liczbą rodziców węzła. Dla zmiennej binarnej posiadającej n rodziców konieczne jest określenie wartości dla 2^n różnych konfiguracji rodziców. Zjawisko to stanowi jedno z głównych ograniczeń praktycznego wykorzystania dużych sieci Bayesowskich.

Znajomość struktury grafu oraz wszystkich tablic CPT pozwala wyznaczyć pełny rozkład łączny modelu. Dla rozważanej sieci ma on postać:

$$P(Room, Light, Messenger) = P(Room) \cdot P(Light|Room) \cdot P(Messenger|Room).$$

Oznacza to, że prawdopodobieństwo dowolnej konfiguracji zmiennych można obliczyć jako iloczyn odpowiednich wartości odczytanych z tablic CPT. Właśnie ta własność umożliwia efektywne wykonywanie wnioskowania probabilistycznego bez konieczności przechowywania pełnego rozkładu łącznego wszystkich zmiennych.

2. Wnioskowanie probabilistyczne

W celu zademonstrowania procesu wnioskowania probabilistycznego w sieci Bayesowskiej rozważmy sytuację, w której chcemy oszacować prawdopodobieństwo włączenia światła bez dodatkowych obserwacji, a następnie z uwzględnieniem informacji o aktywności pracownika w komunikatorze.

W pierwszej kolejności wyznaczamy prawdopodobieństwo a priori włączonego światła, czyli bez uwzględniania jakichkolwiek dodatkowych obserwacji. Korzystając z prawa prawdopodobieństwa całkowitego, otrzymujemy:

$$P(\text{Light} = 1) = P(\text{Light} = 1 \mid \text{Room} = 1)P(\text{Room} = 1) + P(\text{Light} = 1 \mid \text{Room} = 0)P(\text{Room} = 0)$$

$$P(\text{Light} = 1) = 0.5 \cdot 0.4 + 0.05 \cdot 0.6 = 0.23$$

Oznacza to, że bez dodatkowych informacji światło w pokoju pracownika jest włączone z prawdopodobieństwem równym 23%.

Następnie rozważmy sytuację, w której zaobserwowano aktywny status pracownika w komunikatorze, czyli zdarzenie $\text{Messenger} = 1$. W takim przypadku interesuje nas prawdopodobieństwo warunkowe, które zostało wyliczone algorytmem *Variable Elimination* przy użyciu biblioteki `pgmpy`:

$$P(\text{Light} = 1 \mid \text{Messenger} = 1) \approx 0.46$$

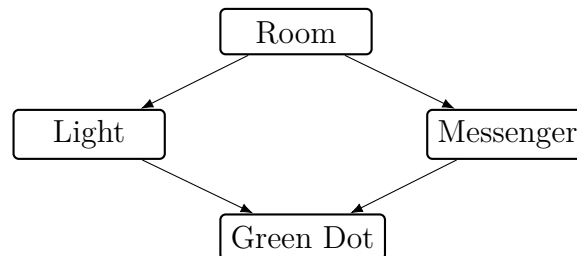
Interpretacja otrzymanego wyniku:

- Zmienne **Light** oraz **Messenger** nie są bezpośrednio połączone w sieci Bayesowskiej. Ich zależność wynika ze wspólnego rodzica – węzła **Room**, który reprezentuje obecność pracownika w biurze.
- Obserwacja zdarzenia $\text{Messenger} = 1$ zwiększa prawdopodobieństwo, że pracownik znajduje się w swoim biurze, ponieważ znacznie częściej korzysta on z komunikatora podczas pracy stacjonarnej niż zdalnej.
- Wzrost prawdopodobieństwa obecności pracownika w biurze prowadzi również do wzrostu prawdopodobieństwa włączonego światła, co skutkuje zmianą wartości $P(\text{Light} = 1)$ z 0.23 do około 0.46.
- Otrzymany wynik pokazuje podstawową zaletę sieci Bayesowskich – możliwość aktualizacji przekonań na temat nieobserwowanych zmiennych na podstawie nowych informacji (dowodów) oraz ukrytych zależności zapisanych w strukturze grafu.

Powyższy przykład doskonale uwydatnia zjawisko niebezpośredniego wpływu jednych zmiennych na drugie. Mimo całkowitego braku bezpośredniego połączenia między statusem komunikatora a oświetleniem biura, wprowadzony dowód (*evidence*) naturalnie propaguje się przez graf, modyfikując stan naszej wiedzy. To właśnie w tym celu buduje się sieci Bayesowskie – stanowią one dedykowane narzędzie matematyczne, które pozwala automatycznie i spójnie aktualizować rozkłady prawdopodobieństwa w całym modelu na podstawie wyrywkowych, pośrednich dowodów.

3. Paradoks Berksona

Rozpatrzmy sytuację, w której strona „Z kim można porozmawiać” wyświetla zieloną kropkę przy pracowniku, jeśli wykryje włączone światło w pokoju lub aktywność na komunikatorze. Zaktualizowanie sieci o dodatkowy węzeł **Green Dot**, który zależy od zmiennych **Light** oraz **Messenger**, wygląda następująco:



Zmienne **Light** oraz **Messenger** determinują zmienną **Green Dot**, zgodnie z zależnością:

$$Green = Light \vee Messenger.$$

Sprawdźmy prawdopodobieństwo, że światło jest włączone przy zaobserwowanej zielonej kropce oraz takie prawdopodobieństwo, gdy dodatkowo wiemy, że pracownik jest zalogowany do komunikatora:

$$P(Light = 1 \mid Green = 1) \approx 0.55,$$

$$P(Light = 1 \mid Green = 1, Messenger = 1) \approx 0.42.$$

Można zauważyć, że uwzględnienie dodatkowej informacji o aktywności w komunikatorze powoduje spadek prawdopodobieństwa świecenia światła. Zjawisko to wynika z faktu, że zarówno **Light**, jak i **Messenger** stanowią alternatywne przyczyny wystąpienia obserwacji **Green Dot**. W praktyce oznacza to, że:

- Gdy widzimy samą zieloną kropkę, wiemy, że *coś* musiało ją wywołać, dlatego prawdopodobieństwo obu potencjalnych przyczyn naturalnie rośnie.
- Gdy jednak otrzymujemy twardy dowód, że pracownik używa komunikatora, informacja ta staje się **wystarczającym uzasadnieniem** dla obecności zielonej kropki.
- W efekcie, hipoteza o włączonym świetle nie jest już potrzebna do wyjaśnienia stanu systemu. Jej prawdopodobieństwo maleje, ponieważ komunikator „przejmuje” odpowiedzialność za ten skutek.

Opisane zjawisko nosi nazwę **paradoksu Berksona** (ang. *Berkson's paradox*) lub efektu *explaining away*. Polega ono na tym, że warunkowanie na wspólnym skutku dwóch zmiennych wprowadza między nimi zależność, mimo że bez tej obserwacji mogą być one niezależne (po uwzględnieniu ich wspólnej przyczyny **Room**).

Jest to jedno z kluczowych, nieintuicyjnych zjawisk w sieciach Bayesowskich, pokazujące, że dodatkowa informacja może zmniejszać prawdopodobieństwo zdarzenia zamiast je zwiększać.

Podsumowanie

Przeprowadzona analiza oraz implementacja modelu probabilistycznego pozwoliły na praktyczne zbadanie własności sieci Bayesowskich jako narzędzia do wnioskowania w warunkach niepewności. Wykorzystanie prostego modelu środowiska umożliwiło przejrzyste pokazanie zarówno podstawowych mechanizmów propagacji informacji, jak i bardziej nieintuicyjnych zjawisk probabilistycznych.

- **Efektywna reprezentacja modelu:** Rozkład łączny został zdekomponowany do iloczynu lokalnych tablic CPT, co znacząco redukuje złożoność zapisu modelu i potwierdza przewagę podejścia grafowego nad pełną reprezentacją tabelaryczną.
- **Wnioskowanie na podstawie obserwacji:** Model pokazał zdolność do aktualizacji rozkładów zmiennych ukrytych na podstawie dowodów pośrednich. Obserwacja Messenger wpływa na ocenę Light poprzez wspólną przyczynę Room, co ilustruje propagację informacji w sieci.
- **Zależności nieintuicyjne:** Analiza węzła Green Dot pozwoliła zaobserwować efekt *explaining away* (paradoks Berksona). Wprowadzenie informacji o jednej z przyczyn zmniejsza potrzebę wyjaśniania skutku drugą, co prowadzi do zmiany rozkładów warunkowych.
- **Implementacja algorytmu:** Zastosowanie algorytmu eliminacji zmiennych (*Variable Elimination*) w bibliotece `pgmpy` umożliwiło automatyczne i dokładne wykonanie wnioskowania probabilistycznego bez enumeracji pełnej przestrzeni stanów.

Podsumowując, sieci Bayesowskie stanowią spójne połączenie graficznej reprezentacji zależności oraz formalizmu probabilistycznego. Dzięki temu pozwalają modelować niepewność w sposób strukturalny, a jednocześnie wykonywać efektywne wnioskowanie na podstawie częściowych obserwacji. W rezultacie są one praktycznym narzędziem w systemach decyzyjnych oraz modelach sztucznej inteligencji operujących w warunkach niepełnej informacji.