

רשתות בייסיאניות Bayesian Networks

רשת בייסיאנית היא מודל פורמלי שעוזר לחזק או להחליש השערות בהתאם לכל מיני נתונים בוליאניים.

התסברות מותנית

מתוך סט המצבים האפשריים שבהם B מקבל ערך true, בכמה גם A מקבל ערך true?

- נרשום: "ההסתברות של A המותנית ב" B "
- "ההסתברות של A בהינתן B "

למשל: H כאב ראש $P(H = \text{true}) = \frac{1}{10}$

F התחלה של שפעת $P(F = \text{true}) = \frac{1}{40}$

אם חלית בשפעת, הסיכוי שיהיה לך כאב ראש יותר גבוה - $P(H = \text{true}|F = \text{true}) = \frac{1}{2}$

Joint Probability Distribution

joint probability מוגדרת על סט של משתנים - למשל $P(A = \text{true}, B = \text{true}, C = \text{true})$. ניתן לקבוע הסתברות לכל קומבינציה של ערכים - והסכום של ההסתברויות של חייב להסתכם ל-1!
מתוך joint probability distribution ניתן לחשב את ההסתברויות המותנות.

בעיה

יש המון קומבינציות! אם יש k משתנים בוליאניים, יש 2^k אפשרויות! הפתרון - נשתמש באי תלות

אי תלות

המשתנים A, B בלתי תלויים אם מתקיים:

$$P(A, B) = P(A) \cdot P(B)$$

שקול: $P(A|B) = P(A)$

$P(B|A) = P(B)$

כלומר - ידיעת B לא מוסיפה לנו מידע על A

אם רוצים לחשב $P(C_1, \dots, C_n)$ צריך 2^n נתונים בטבלה - אבל אם ידוע ש C_1, \dots, C_n בלתי תלויים אזי

$$P(C_1, \dots, C_n) = \prod_{i=1}^n P(C_i)$$

ולכן אפשר להסתפק ב n ערכים!

אי תלות מותנית Conditional Independence

המשתנים A, B יהיו conditionally independent בהינתן C אם מתקיים

$$P(A, B|C) = P(A|C) \cdot P(B|C)$$

שקול: $P(A|B, C) = P(A|C)$ •
 $P(B|A, C) = P(B|C)$ •

כלומר אם כבר יודעים את C , ידיעת B לא מוסיפה לנו מידע על A

רשת בייסיאנית Bayesian Network

רשת בייסיאנית היא גרף כיווני לא מעגלי, שבו כל צומת היא משתנה אקראי(מאורע). לכל צומת יש טבלה שבה יש את ההסתברויות של כל משתנה בהינתן ערכי האבות $P(X_i | \text{Parents}(X_i))$. לכל ערך של אב, המצבים האפשריים של הבן צריכים להסתכם ל-1.

התפלגות Joint Distribution ברשת בייסיאנית

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i | \text{Parents}(X_i))$$

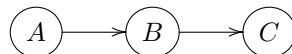
הסקה

השימוש ברשת בייסיאנית לצורך חישוב הסתברויות נקראית הסקה(inference), וכוללת שאילתות מסוג $P(X|E)$, כאשר X זה המאורע שרוצים למצוא ו- E הם העvidence, המאורעות הידועים לנו. לא חייבים להכניס ל- E את כל המאורעות ברשת הבייסיאנית.

סוגים של Evidence

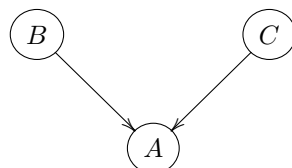
- Hard Evidence - ידוע לנו בוודאות שמאורע X מקבל ערך ספציפי כלשהו
- Soft Evidence - לא ידוע לנו הערך של X , אבל אנחנו יכולים לעדכן אותו בעקבות (Hard)evidence או (Soft) על מאורעות אחרים שמשפיעים עליו

Serial Connection



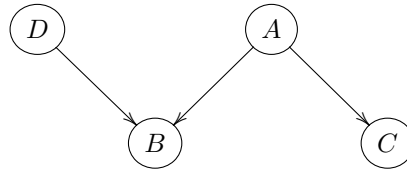
ידע על A משפיע על B ולכן גם על C . אבל אם יודעים גם את A וגם את B (עם Hard Evidence), אין ל- A השפעה על C , כי ההשפעה שלו קודם עברה דרך B

Converging Connection



מידע על B ישפיע על C ולהפך - אלא אם כן יש לנו Hard Evidence על A !

"Explaining Away" Evidence



אם יש לנו Hard Evidence על B , זה משפיע על ההסתברות של שתי הסיבות ל- A ו- D :

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

Matching Literature

יש (Multi Agent System) MAS. כל פעם זוג Agent ים מעריך אחד את השני, ואם שניהם אומרים כן - שניהם יוצאים מה-MAS. אם לפחות אחד מהם אומר לא - שניהם ממשיכים בחיפוש. יש עלות להמשך החיפוש, ואף אחד לא רוצה להמשיך לחפש עד אינסוף...

Benefits of Limited Information

יש initial cost לאינטרקציה עם סוכן אחר, שצריך לשלם אותה בין אם השידוך מצליח ובין אם לא, אבל אחרי שמשלמים אותה יודעים מה התועלת שהשידוך יתן לנו.

$$\text{utility} \in \{0, 1\} \text{ למשל}$$
$$\text{initial cost} = 0.2$$

אם הסיכוי הוא לכל utility הוא 50%, אז צריך במוצע 4 אינטרקציות כדי למצוא $\text{utility} = 1$ (כי צריך שגם הסוכן השני יאשר אותך), ואז:

$$\text{utility} = 1 - 4 \cdot 0.2 = 0.2$$

אבל מה אם לא יודעים מה utility לאחר שמשלמים את initial cost (למשל פוגשים את הסוכן האחר בבר עם מוזיקה רועשת ואורות עמומים)? במקרה כזה התוחלת לתועלת היא 0.5, ולא שווה לנו להגיד "לא" כי אין לנו אפשרות לדעת מתי לעצור, לכן:

$$\text{utility} = 0.5 - 0.2 = 0.3$$

כלומר הערך של המידע הוא -0.1 ! איך זה יכול להיות שהערך של המידע (VoI) שלילי? הערך של מידע הוא תמיד $VoI \geq 0$. מה שקרה פה זה שנתנו את המידע לכולם. מה היה קורא אם המידע היה מסופק רק לנו? במקרה הזה, הסוכן השני תמיד יקבל ואילו אנחנו נקבל רק כאשר $\text{utility} = 1$. במקרה הזה צריך רק 2 אינטרקציות, ולכן:

$$\text{utility} = 1 - 2 \cdot 0.2 = 0.6$$

וזה יותר טוב מאשר כאשר אין לאף אחד מידע. לעומת זאת, אם אנחנו היחידים שאין להם את המידע, מספר האינטרקציות הוא 2 ולכן:

$$\text{utility} = 0.5 - 2 \cdot 0.2 = 0.1$$

וזה פחות טוב מאשר כאשר לכולם יש מידע