Probabilistic Machine Learning:
8. Belief Networks: Independence

Tomasz Kajdanowicz

Department of Computational Intelligence
Wroclaw University of Technology

HR EXCELLENCE IN RESEARCH

Wrocław University of Science and Technology

The presentation was inspired by Chapter 3 of D. Koller and N.Friedman book "Probabilistic Graphical Models - Principles and Techniques", 2009.

Conditional Independence

For sets of random variables X, Y, Z, $P \models (X \perp Y|Z)$ if:

- $P(X, Y|Z) = P(X|Z)P(Y|Z)$ or
- $P(X|Y, Z) = P(X|Z)$ or
- $P(Y|X, Z) = P(Y|Z)$

How to determine the Conditional Independece?

d-separation

Conditional Independence

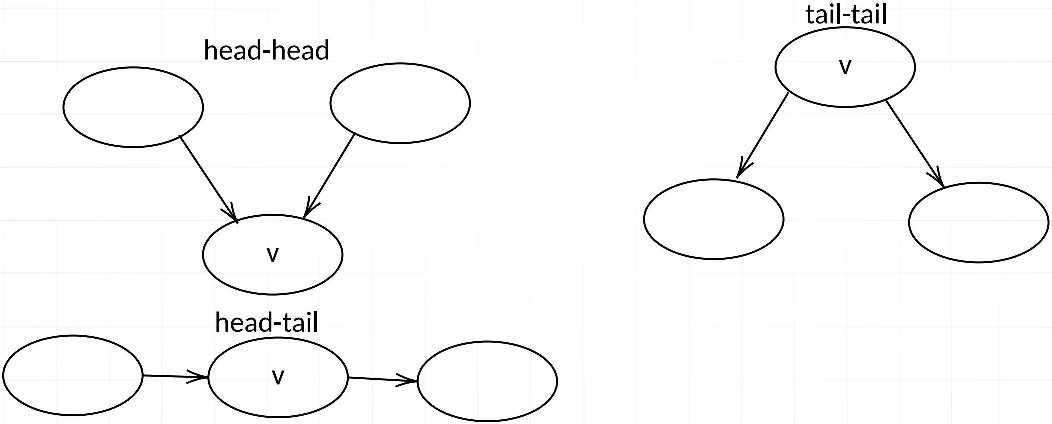
For random variables X, Y, Z if X and Y are ***d-separated*** by Z then $(X \perp Y|Z)$

d-separation

X and Y are d-separated by Z if all paths from a vertex of X to a vertex of Y are **blocked** with respect to v.

Types of connection on the path

with respect to Z



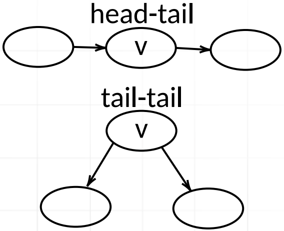
Blocked path

Blocked path

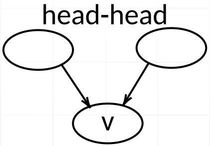
A path (not necessary directed) between two vertices is **blocked** with respect to C (even set of vertices) if it passes through a vertex v s.t. either:

- the connections are head-tail or tail-tail and $v \in C$
- the connections are head-head and $v \notin C$ and none of descendants of v are in C

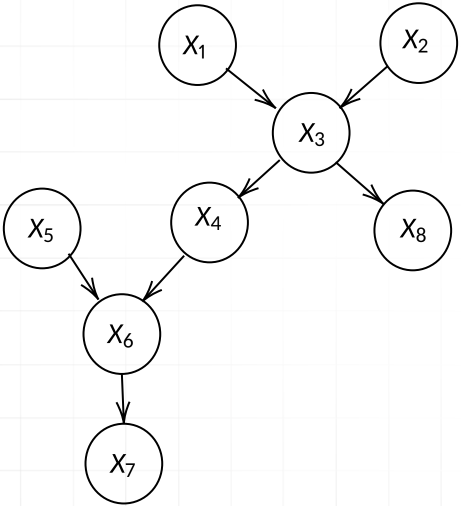
$v \in C$



$v \notin C$ and none of descendants of v are in C

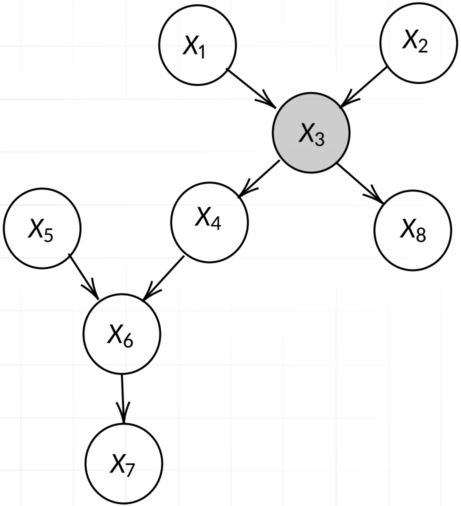


Example



- taking e.g.:
  - $C = \{X_3\}$
- $(X_i \perp X_j)|X_C$???

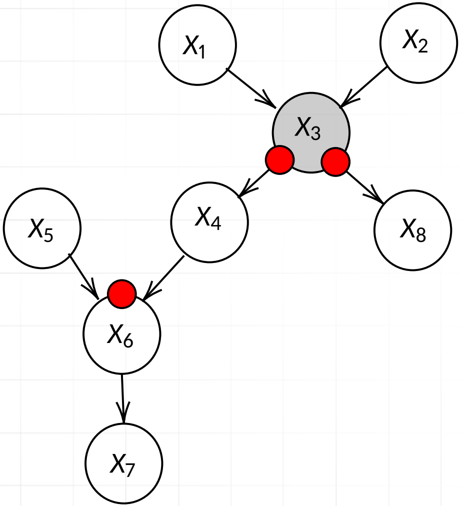
Example



- taking e.g.:
  - $C = \{X_3\}$
- $(X_i \perp X_j)|X_C$???

i	j	d-separated
1	4	yes
1	2	no
4	5	yes

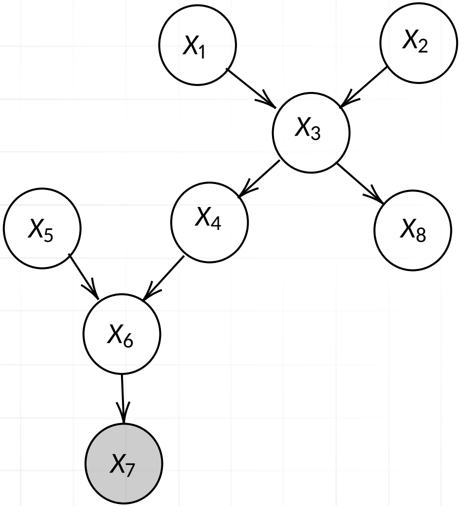
Example



- taking e.g.:
  - $C = \{X_3\}$
- $(X_i \perp X_j)|X_C$???

i	j	d-separated
1	4	yes
1	2	no
4	5	yes
4	7	no
4	8	yes
4	6	no
2	7	yes
2	5	yes
5	8	yes

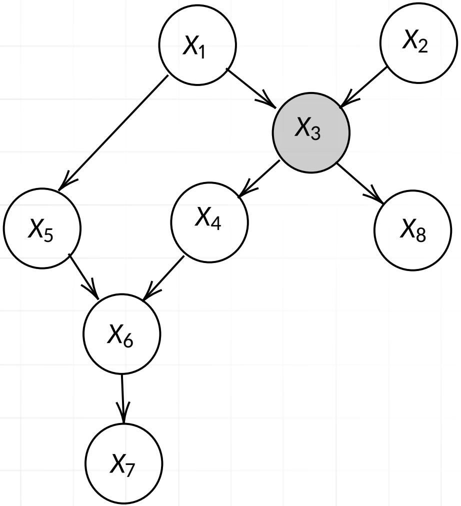
Example 2



- taking e.g.:
  - $C = \{X_7\}$
- $(X_i \perp X_j)|X_C$???

i	j	d-separated

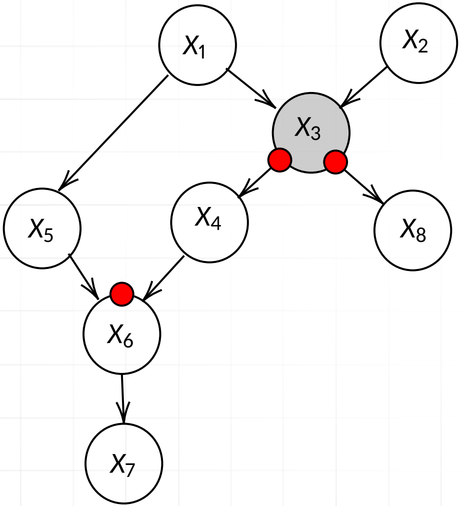
Example 3



- taking e.g.:
  - $C = \{X_3\}$
- $(X_i \perp X_j)|X_C$???

i	j	d-separated

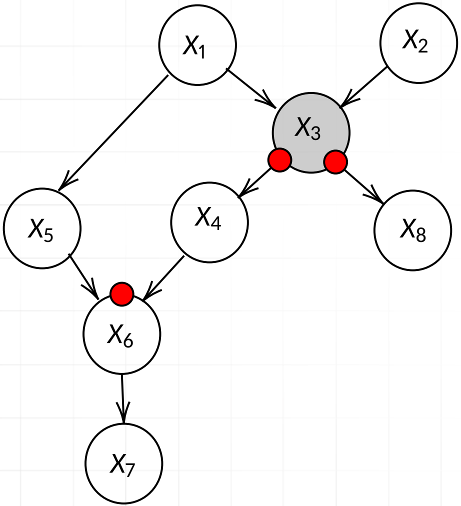
Example 3



- taking e.g.:
  - $C = \{X_3\}$
- $(X_i \perp X_j)|X_C$???

i	j	d-separated

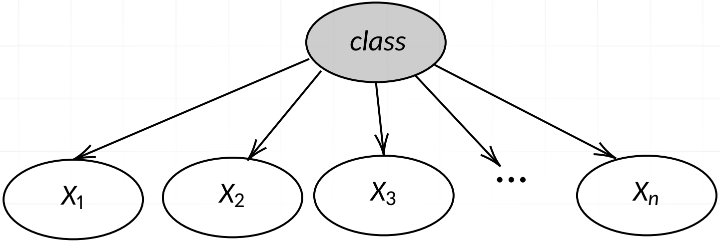
Example 3



- taking e.g.:
  - $C = \{X_3\}$
- $(X_i \perp X_j)|X_C$???

i	j	d-separated
2	5	no
2	7	no
2	4	yes
7	8	yes

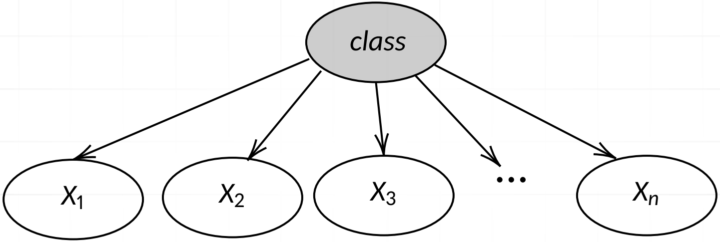
Naive Bayes



- we infer *class*
- based on features $X_1, X_2, X_3, ..., X_n$
- model acomplishes: $(X_i \perp X_j)|X_C$ for all X_i,X_j

$$P(C, X_1, \ldots, X_n) = P(C) \prod_{i=1}^{n} P(X_i|C)$$

Naive Bayes



$$\frac{P(C=c^1|x_1,\ldots,x_n)}{P(C=c^2|x_1,\ldots,x_n)} = \frac{P(C=c^1)}{P(C=c^2)}\prod_{i=1}^{n}\frac{P(x_i|C=c^1)}{P(x_i|C=c^2)}$$