Advanced Machine Learning

Likhit Nayak

Naive Bayes Model

It is derived from two major ideas:

1. Independent Random Variables

Given a set of random variables ***{X₁,......Xₙ}***,

$$P(X_1, \ldots, X_n) = P(X_1)P(X_2) \cdots P(X_n)$$

Naive Bayes Model

It is derived from two major ideas:

2. Conditional Probability

Let's take a joint probability table as follows:

I	S	$P(I,S)$
i^0	s^0	0.665
i^0	s^1	0.035
i^1	s^0	0.06
i^1	s^1	0.24.

$$P(I,S) = P(I)P(S \mid I)$$

Naive Bayes Model

It is derived from two major ideas:

2. Conditional Probability

Let’s take a joint probability table as follows:

i^0	i^1
0.7	0.3

I	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

Naive Bayes Model

Conditional Independence:

Given three random variables *I*, *S*, and *G*:

$$\boxed{P(S, G \mid I) = P(S \mid I)P(G \mid I)}$$

$$P(I, S, G) = P(S, G \mid I)P(I)$$

$$P(I, S, G) = P(S \mid I)P(G \mid I)P(I)$$

Naive Bayes Model

Conditional Independence:

Given three random variables *I*, *S*, and *G*:

i^0	i^1
0.7	0.3

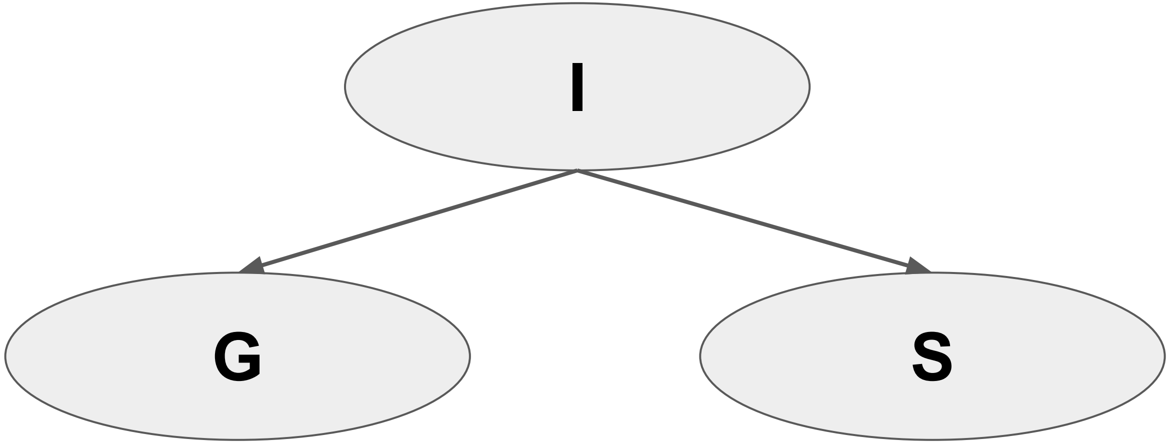
I	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

I	g^1	g^2	g^3
i^0	0.2	0.34	0.46
i^1	0.74	0.17	0.09

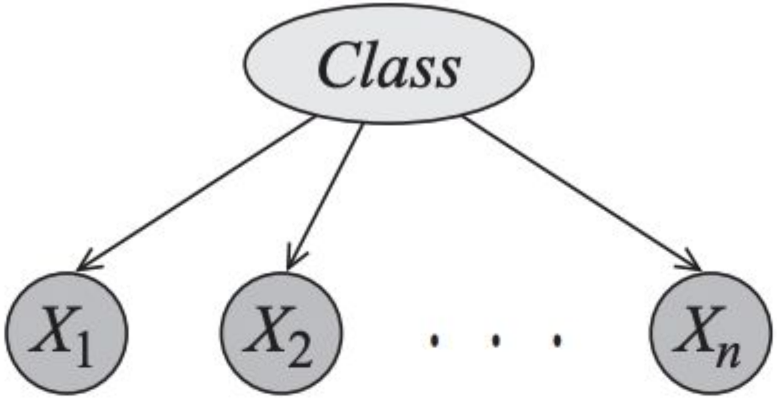
Naive Bayes Model

Conditional Independence:

Given three random variables *I*, *S*, and *G*:

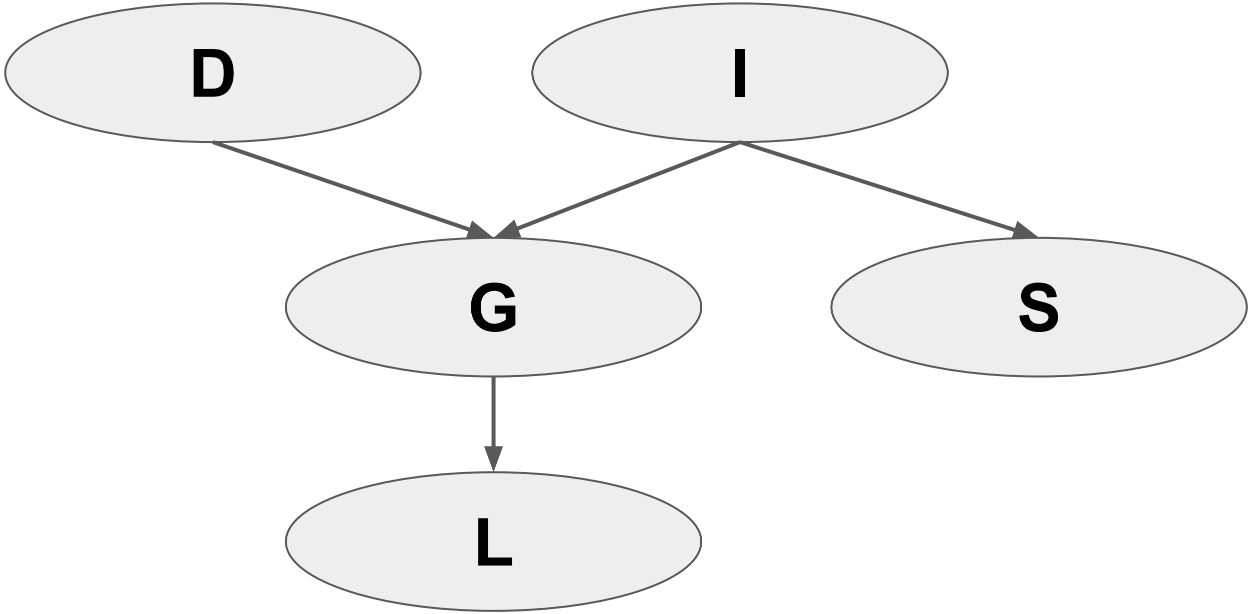


Naive Bayes Model



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

Directed Acyclic Graph (DAG)



$$P(I, D, G, S, L) = P(I)P(D)P(G \mid I, D)P(S \mid I)P(L \mid G).$$

Directed Acyclic Graph (DAG)