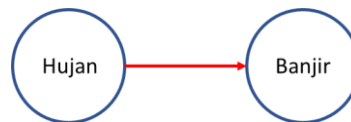


## Tutorial Bayesian Network

Bayesian Network itu digunakan untuk menggambarkan hubungan setiap variable dalam probabilitas. Sebagai contoh, misalkan ada dua variable yaitu hujan dan banjir. Di setiap variable ada nilai probabilitas. Nilai probabilitas tersebut dapat bergantung sama variable lain atau tidak bergantung sama sekali. Probabilitas banjir dapat bergantung terhadap pada cuaca (hujan atau tidak). Sedangkan probabilitas hujan tidak bergantung apapun.

Bayesian Network dari contoh di atas adalah sebagai berikut:



Contoh angka probabilitasnya adalah sebagai berikut:

- $P(Hujan) = 0.2$ 
  - Artinya kemungkinan hujan adalah 0.2 atau 20%
- $P(Banjir | Hujan) = 0.4$ 
  - Artinya kemungkinan terjadinya banjir apabila hujan turun adalah = 0.4 atau 40%
- $P(Banjir | \sim Hujan) = 0.05$ 
  - Artinya kemungkinan terjadinya banjir apabila hujan tidak turun adalah = 0.05 atau 5%

Dari Bayesian Network ini, kita dapat menghitung probabilitas keseluruhan berdasarkan semua variable. Sebagai contoh, Berapa probabilitas hujan tidak turun **dan** terjadi banjir? Cara menghitungnya adalah dengan mengalikan semua kemungkinan probabilitas yang ada (Cek Slide kelas).

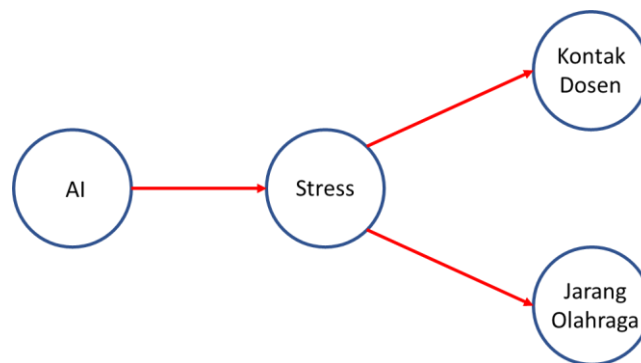
- $P(\sim Hujan) = 1 - 0.2 = 0.8$ 
  - Artinya kemungkinan tidak hujan adalah  $1 - 0.2 = 0.8$ .
- $P(Banjir | \sim Hujan) = 0.05$
- $P(Banjir, \sim Hujan) = P(Banjir | \sim Hujan)P(\sim Hujan) = 0.05 * 0.8 = 0.04$ 
  - Artinya kemungkinan hujan tidak turun dan terjadi banjir adalah 0.04 atau 4%

Sampai sini, cukup jelaskan? 😊

Mari kita tes di Bayesian Network yang lebih rumit.

Kemungkinan mahasiswa mengambil peminatan AI adalah 20%. Karena AI itu mudah, mahasiswa yang mengambil peminatan AI itu probabilitas mendapat stress adalah 10%, sedangkan mahasiswa yang bukan peminatan AI itu probabilitas mendapat stress adalah 30%. Kemungkinan mahasiswa yang stress untuk menghubungi dosen adalah 60%, sedangkan untuk yang tidak stress adalah 10%. Selain itu kemungkinan mahasiswa yang stress itu jarang olahraga adalah 70%, sedangkan untuk yang tidak stress adalah 50%. Gambarkan Bayesian Networknya! Selain itu tentukan probability dari PaoPao yang mengambil peminatan AI dan tidak stress, lalu suka menghubungi dosen dan sering olahraga?

### Bayesian Network



- $P(AI) = 0.2$
- $P(Stress|AI) = 0.1$
- $P(Stress|\sim AI) = 0.3$
- $P(KDosen|Stress) = 0.6$
- $P(KDosen|\sim Stress) = 0.1$
- $P(JOlahraga|Stress) = 0.7$
- $P(JOlahraga|\sim Stress) = 0.5$

### Probabilitas dari kasus PaoPao

- $P(KDosen, JOlahraga, \sim Stress, AI) = P(KDosen|\sim Stress)P(JOlahraga|\sim Stress)P(\sim Stress|AI)P(AI)$
- $P(KDosen, JOlahraga, \sim Stress, AI) = 0.1 * 0.5 * (1 - 0.1) * 0.2$
- $P(KDosen, JOlahraga, \sim Stress, AI) = 0.05 * 0.18 = \mathbf{0.009}$

Sebagai latihan, cobalah kalian mengganti variable untuk mengubah kasusnya!

Terkadang, kita tidak mempunyai kesempatan untuk mengetahui semua informasi mengenai variable. Oleh karena itu, kita dapat menghitung probabilitas dari sebuah kondisi berdasarkan evidence/observation yang kita lakukan. Sebagai contoh di kasus PaoPao, PaoPao suka menghubungi dosen, jadi berapakah probabilitas bila PaoPao peminatan AI? Kita hanya mengetahui kondisi dimana PaoPao suka menghubungi dosen atau  $KDosen = True$ . Lalu kita ditanya kondisi dimana PaoPao merupakan peminatan AI atau  $AI = True$ . Jadi pertanyaan dalam bentuk equation adalah  $P(AI|KDosen) = ??$

Untuk menjawab pertanyaan di atas, kita dapat menghitung dengan **menjumlahkan** segala kemungkinan yang mana kondisi  $KDosen = True$  dan  $AI = True$  terpenuhi.

$$\begin{aligned} P(AI|KDosen) &= P(AI|Stress, Jolahraga, KDosen) + \\ &P(AI|Stress, \sim Jolahraga, KDosen) + P(AI|\sim Stress, Jolahraga, KDosen) + \\ &P(AI|\sim Stress, \sim Jolahraga, KDosen) \end{aligned}$$

Bila diperhatikan, kita mempunyai informasi  $P(Stress|AI)$  tapi kita tidak mempunyai informasi  $P(AI|Stress)$ . Untuk menghitungnya, buka kembali materi pertemuan sebelum UTS yang berupa naïve Bayes.

$$\begin{aligned} P(AI|Stress, Jolahraga, KDosen) &= P(AI) * P(Stress, Jolahraga, KDosen|AI) \\ &= P(AI) * P(Stress|AI) * P(Jolahraga|Stress) * \\ &\quad P(KDosen|Stress) \\ &= 0.2 * 0.1 * 0.7 * 0.6 \\ &= 0.0084 \\ P(AI|Stress, \sim Jolahraga, KDosen) &= P(AI) * P(Stress, \sim Jolahraga, KDosen|AI) \\ &= P(AI) * P(Stress|AI) * P(\sim Jolahraga|Stress) * \\ &\quad P(KDosen|Stress) \\ &= 0.2 * 0.1 * (1 - 0.7) * 0.6 \\ &= 0.0036 \\ P(AI|\sim Stress, Jolahraga, KDosen) &= P(AI) * P(\sim Stress, Jolahraga, KDosen|AI) \\ &= P(AI) * P(\sim Stress|AI) * P(Jolahraga|\sim Stress) * \\ &\quad P(KDosen|\sim Stress) \\ &= 0.2 * (1 - 0.1) * 0.5 * 0.1 \\ &= 0.009 \\ P(AI|\sim Stress, \sim Jolahraga, KDosen) &= P(AI) * P(\sim Stress, \sim Jolahraga, KDosen|AI) \\ &= P(AI) * P(\sim Stress|AI) * P(\sim Jolahraga|\sim Stress) * \\ &\quad P(KDosen|\sim Stress) \\ &= 0.2 * (1 - 0.1) * (1 - 0.5) * 0.1 \\ &= 0.009 \end{aligned}$$

Jadi, berapa probabilitas dari PaoPao peminatan AI adalah sebagai berikut:

$$P(AI|KDosen) = 0.0084 + 0.0036 + 0.009 + 0.009 = 0.03$$

**NAMUN**, nilai probabilitas ini BELUM dinormalisasi. Untuk normalisasi probabilitasnya, kita harus **menghitung nilai PaoPao bukan peminatan AI dan PaoPao suka mengontak dosen** ( $KDosen = True$  dan  $AI = False$ )

$$\begin{aligned} P(\sim AI|KDosen) &= P(\sim AI|Stress, Jolahraga, KDosen) + \\ &P(\sim AI|Stress, \sim Jolahraga, KDosen) + P(\sim AI|\sim Stress, Jolahraga, KDosen) + \\ &P(\sim AI|\sim Stress, \sim Jolahraga, KDosen) \end{aligned}$$

Saya berikan kepada kalian untuk menghitung nilai probabilitas yang belum dinormalisasi menggunakan cara di atas. Hasilnya adalah sebagai berikut

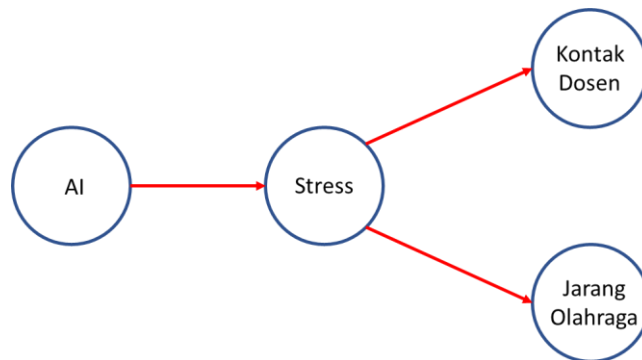
- $P(\sim AI|KDosen) = 0.1008 + 0.0432 + 0.028 + 0.028 = 0.2$

Setelah mendapat dua kondisi tersebut, kita dapat mernormalisasi nya dengan cara sebagai berikut

- $P(AI|KDosen) = \frac{0.03}{0.2+0.03} = 0.13$
- $P(\sim AI|KDosen) = \frac{0.2}{0.2+0.03} = 0.87$

Cara 2:

Untuk menyelesaikan soal ini, terdapat juga cara lain yaitu menggunakan table. Table digunakan untuk menampilkan semua jenis kemungkinan joint probability distribution. Yang harus dilakukan pertama kali adalah menentukan variable apa saja yang dimasukkan ke dalam table. Karena variable KDosen = True (sudah pasti true), maka tersisa 3 variable lain yang dapat dimasukkan ke table yaitu (AI, Stress, JOlahraga). Kombinasi dari 3 variable adalah  $2^3 = 8$ .



- $P(AI) = 0.2$
- $P(Stress|AI) = 0.1$
- $P(Stress|\sim AI) = 0.3$
- $P(KDosen|Stress) = 0.6$
- $P(KDosen|\sim Stress) = 0.1$
- $P(JOlahraga|Stress) = 0.7$
- $P(JOlahraga|\sim Stress) = 0.5$

AI	Stress	JOlahraga	$P(AI) * P(Stress AI) * P(JOlahraga Stress) * P(KDosen = T Stress)$
T	T	T	$0.2 * 0.1 * 0.7 * 0.6 = 0.0084$
T	T	F	$0.2 * 0.1 * (1-0.7) * 0.6 = 0.0036$
T	F	T	$0.2 * (1-0.1) * 0.5 * 0.1 = 0.009$
T	F	F	$0.2 * (1-0.1) * (1-0.5) * 0.1 = 0.009$
F	T	T	$(1-0.2) * 0.3 * 0.7 * 0.6 = 0.1008$
F	T	F	$(1-0.2) * 0.3 * (1-0.7) * 0.6 = 0.0432$
F	F	T	$(1-0.2) * (1-0.3) * 0.5 * 0.1 = 0.028$
F	F	F	$(1-0.2) * (1-0.3) * (1-0.5) * 0.1 = 0.028$
Total			= 0.23

Setelah kita membuat table joint probability distributionnya, kita jumlahkan semuanya: = **0.23**

Lalu kita jawab pertanyaan dari soalnya yaitu  $P(AI = T | KDosen = T)$  atau berapa probabilitas PaoPao merupakan mahasiswa peminatan AI apabila dia suka menghubungi dosennya. Kita dapat menghitungnya dengan cara menjumlah angka-angka dari table yang mempunyai kondisi dimana  $AI = True$ . \*Dapat dilihat pada warna merah di table. Setelah dijumlahkan, kita dapat membaginya dengan total nilai yang sudah dihitung yaitu (0.23).

$$P(AI = T | KDosen = T) = \frac{0.0084 + 0.0036 + 0.009 + 0.009}{0.23} = \frac{0.03}{0.23} = 0.13$$

Selain itu berapa probabilitas PaoPao **bukan** merupakan mahasiswa peminatan AI apabila dia suka menghubungi dosennya adalah sebagai berikut:

$$P(AI = F | KDosen = T) = 1 - 0.13 = 0.87$$

Coba lakukan kombinasi observasi yang dapat kalian lakukan sebagai latihan!

Semoga latihannya berguna!