

PELUANG

Ali Akbar Septiandri

Universitas Al-Azhar Indonesia

aliakbars@live.com

February 21, 2020

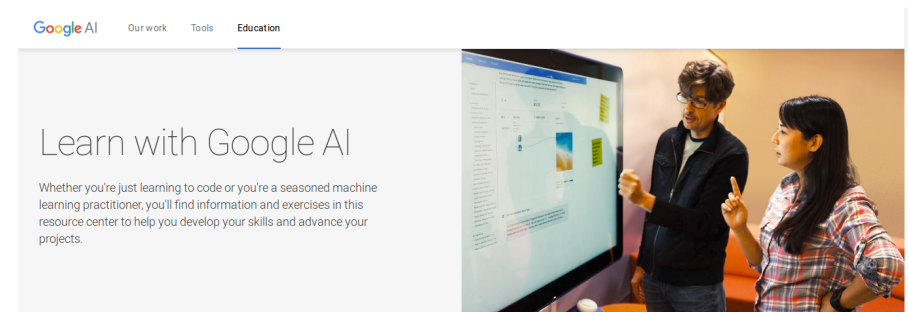
① PEUBAH ACAK

② PELUANG BERSYARAT

③ BAYES' RULE

COBA INI...

COBA INI...

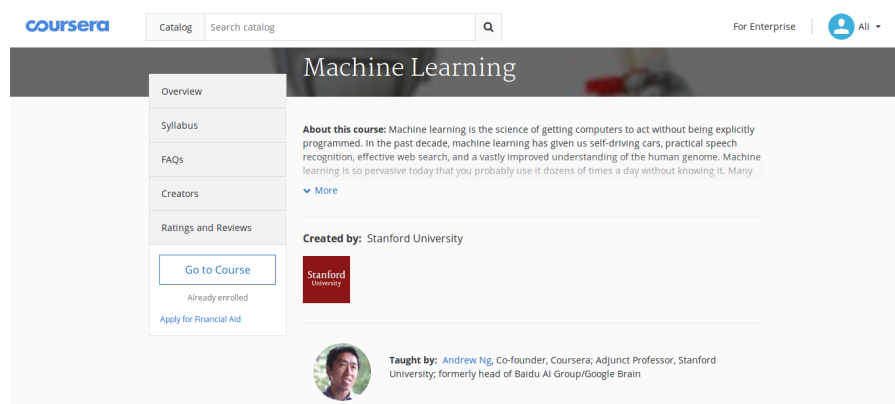


Educational resources from machine learning

GAMBAR: Learn with Google AI

GAMBAR: Machine Learning Flashcards

COBA INI...



GAMBAR: Machine Learning - Coursera

COBA INI...



GAMBAR: Statistics 110 - Harvard

PEUBAH ACAK

APA ITU?

S = himpunan dari semua keluaran yang mungkin terjadi

EXAMPLE

lemparan koin

$$S = \{Angka, Gambar\}$$

lemparan dua koin

$$S = \{(A, A), (A, G), (G, A), (G, G)\}$$

lemparan dadu

$$S = \{1, 2, 3, 4, 5, 6\}$$

jumlah email dalam satu hari

$$S = \mathbb{N}$$

jam bermain Mobile Legends

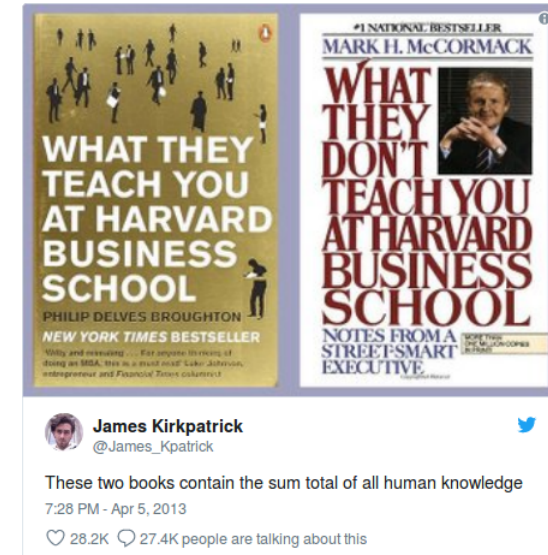
$$S = [0, 24]$$

APA ITU?

E = subhimpunan/subset dari S ($E \subseteq S$)

EXAMPLE

lemparan koin memunculkan angka	$E = \{Angka\}$
≥ 1 angka dari dua koin	$E = \{(A, A), (A, G), (G, A)\}$
lemparan dadu ≥ 3	$E = \{3, 4, 5, 6\}$
# email dalam sehari ≤ 5	$E = \{x \leq 5, x \in \mathbb{N}\}$
“hari-hari tidak produktif”	$E = [8, 24]$

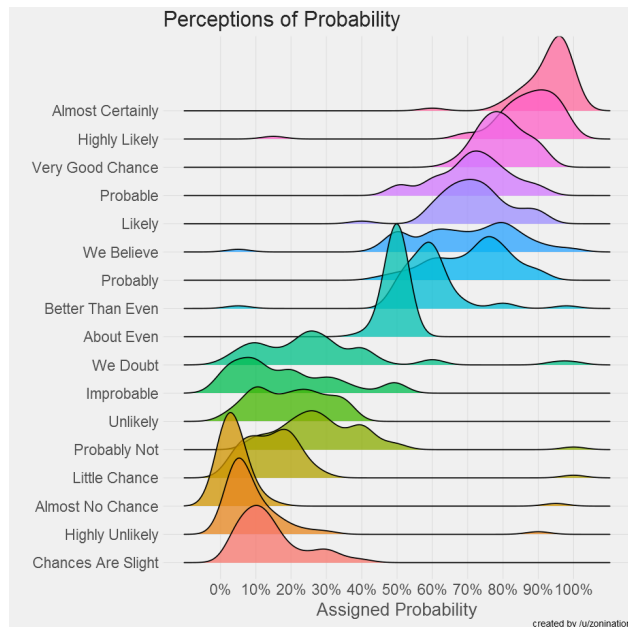


Mengapa?

$$E \cup E' = S$$

Jadi, apa itu peluang/probabilitas?

- Kuantifikasi dari ketidakpastian
- Nilai antara 0 dan 1 yang kita pautkan pada suatu kejadian
- Faktanya, persepsi kita terhadap ketidakpastian bisa berbeda-beda



GAMBAR: Persepsi akan probabilitas — Sumber: <https://github.com/zonination/perceptions>

FREQUENTIST VS BAYESIAN

INTERPRETASI FREQUENTIST

Frekuensi kemunculan kejadian dalam *jangka panjang*

EXAMPLE

Peluang kemunculan sisi angka dari suatu lemparan koin adalah 0.43

INTERPRETASI BAYESIAN

Kuantifikasi *derajat kepercayaan* terhadap sesuatu

EXAMPLE

Peluang besok¹ hujan adalah 0.3

¹Apakah mungkin mengulang “besok”?

$$P(E) = \lim_{n \rightarrow \infty} \frac{\#(E)}{n}$$

- ① $0 \leq P(E) \leq 1$
- ② $P(S) = 1$
- ③ Jika $E \cap F = \emptyset$, maka
 $P(E \cup F) = P(E) + P(F)$

AKIBATNYA...

PEUBAH ACAK

- ① $P(E') = 1 - P(E)$
- ② Jika $E \subseteq F$, maka $P(E) \leq P(F)$
- ③ $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

- Peubah acak atau *random variables* (RV) X menunjukkan sebuah nilai yang dapat berubah-ubah, tergantung kejadian
- Dapat berupa *hasil eksperimen* (e.g. lemparan koin) atau pengukuran kuantitas yang fluktuatif (e.g. temperatur)
- X menggambarkan RV, x menggambarkan nilai, e.g. $p(X = x)$
- Dapat disingkat menjadi $p(x)$
- Sebuah RV dapat bernilai *kontinu* maupun *diskrit*

CONTOH KASUS

EXAMPLE

Dua dadu dilempar bersamaan, berapa peluang munculnya sisi kedua dadu berjumlah 7?

PERTANYAAN

Apa yang harus didefinisikan terlebih dahulu?

JAWAB

Apa yang menjadi ruang sampelnya? Apa pula kejadiannya?

CONTOH KASUS (LANJUTAN)

- $S = \{(1, 1), (1, 2), \dots, (6, 5), (6, 6)\}$
- $E = \{(1, 6), (2, 5), \dots, (5, 2), (6, 1)\}$
- $p(X_1 + X_2 = 7) = ?$
- $p((X_1 = 1 \cap X_2 = 6) \cup (X_1 = 2 \cap X_2 = 5) \cup \dots) = \frac{6}{36}$

CONTOH KASUS

EXAMPLE

Ada 3,200 mahasiswa UAI, Anda berteman dengan 40 orang di antaranya. Jika Anda pergi ke suatu acara yang didatangi 20 orang mahasiswa UAI, berapa peluang Anda menemukan *paling tidak* satu orang teman Anda?

DEFINISIKAN

$p(X \geq 1) = \dots$

Berapa banyak yang harus dihitung?

CONTOH KASUS (LANJUTAN)

- Hitung saja peluang tidak bertemu dengan teman sama sekali, i.e. $p(X = 0)$.
- Maka nilainya dapat dihitung dengan

$$\begin{aligned} p(X \geq 1) &= 1 - p(X = 0) \\ &= 1 - \frac{\binom{3200-40}{20}}{\binom{3200}{20}} = 0.2230 \end{aligned}$$

- Coba lihat: <http://web.stanford.edu/class/cs109/demos/serendipity.html>

- $\sum_x p(x) = 1$
- Dalam kasus RV kontinu, $\int p(x)dx = 1$
- $p(x)$ dalam kasus kontinu dikenal sebagai *probability density function (PDF)*
- Nilai $p(x)$ mungkin > 1 (mengapa?)

- Anggap kita punya fungsi $f(x)$ yang memetakan x ke nilai numerik

$$\begin{aligned}\mathbb{E}[f(x)] &= \sum_x f(x)p(x) \\ &= \int f(x)p(x)dx\end{aligned}$$

untuk variabel diskrit dan kontinu.

- Saat $f(x) = x$, kita akan mendapatkan *mean*, μ_x
- Saat $f(x) = (x - \mu_x)^2$, kita akan mendapatkan *variansi*

CONTOH KASUS

EXAMPLE

Saya akan melempar sebuah koin. Jika sisi yang keluar angka, maka saya akan memberikan Anda Rp 200,000. Jika keluarnya gambar, maka Anda harus memberikan saya Rp 100,000. Apakah Anda akan bermain?

SOLUSI

$$\begin{aligned}\mathbb{E}[f(x)] &= \sum_x f(x)p(x) \\ &= 200000 \cdot \frac{1}{2} + (-100000) \cdot \frac{1}{2} \\ &= 50000\end{aligned}$$

PELUANG BERSYARAT

JOINT DISTRIBUTIONS

- Kita akan lebih sering berurusan dengan banyak RV \rightarrow butuh **joint distributions**
- $p(X_1 = x_1, X_2 = x_2, \dots, X_D = x_D)$
- Saat ragu, selalu mulai dari sini
- Contoh (Koller & Friedman, 2009):

	Intelligence = low	Intelligence = high
Grade = A	0.07	0.18
Grade = B	0.28	0.09
Grade = C	0.35	0.03

CONDITIONAL PROBABILITY

- Peluang bersyarat:

$$p(X = x|Y = y) = p(x|y) = \frac{p(x, y)}{p(y)}$$

- **Product rule:**

$$p(x, y) = p(x)p(y|x) = p(y)p(x|y)$$

- Contoh: Tentukan nilai $p(\text{Intelligence} = \text{high}|\text{Grade} = A)$!

MARGINAL PROBABILITY

- Berapa $p(\text{Grade} = A)$?
- Gunakan **sum rule**:

$$p(x) = \sum_y p(x, y)$$

- Ganti *sum* dengan integral untuk RV kontinu

CHAIN RULE

Aturan rantai (**chain rule**) didapatkan dengan mengaplikasikan *product rule* berulang kali.

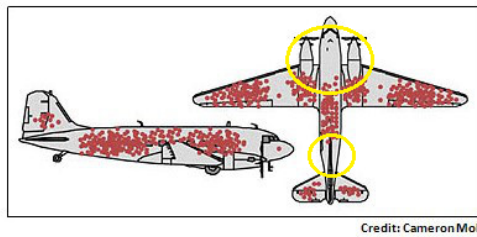
$$\begin{aligned} p(X_1, \dots, X_D) &= p(X_1, \dots, X_{D-1})p(X_D|X_1, \dots, X_{D-1}) \\ &= p(X_1, \dots, X_{D-2})p(X_{D-1}|X_1, \dots, X_{D-2})p(X_D|X_1, \dots, X_{D-1}) \\ &= \dots \\ &= p(X_1) \prod_{i=2}^D p(X_i|X_1, \dots, X_{i-1}) \end{aligned}$$

Break

BAYES' RULE

KISAH ABRAHAM WALD

BAYES' RULE



Gentlemen, you need to put more armour-plate where the holes aren't because that's where the holes were on the airplanes that didn't return - Abraham Wald 1942.

Berdasarkan *product rule*,

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

dengan bagian penyebut yang dapat dijabarkan dengan *sum rule* sebagai berikut

$$p(X) = \sum_Y p(X|Y)p(Y)$$

GAMBAR: Pendekatan matematisnya membuat sekutu berhasil menyelamatkan banyak pesawat pada Perang Dunia II

EXAMPLE

Terdapat 0.08% orang yang terkena virus Corona. Dari 1000 orang yang terkena virus Corona, 900 orang akan menunjukkan hasil tes positif. Di sisi lain, terdapat 7% orang tanpa virus Corona yang juga akan terdeteksi mengidap virus Corona berdasarkan tes yang sama. Jika seseorang menjalani tes tersebut dan dinyatakan positif, berapa peluangnya dia benar-benar mengidap virus Corona?

SOLUSI

- $p(C) = 8 \times 10^{-4}$
- $p(T|C) = 0.9$
- $p(T|C') = 0.07$
- $p(C|T) = \frac{p(T|C)p(C)}{p(T)} \approx 0.01$

TERMINOLOGI




$$\underbrace{P(C|T)}_{\text{posterior}} = \frac{\overbrace{P(C|T)}^{\text{likelihood}} \overbrace{P(C)}^{\text{prior}}}{\underbrace{P(C)}_{\text{normalizing constant}}}$$

Cari: Monty Hall problem

- Peubah acak, ruang sampel, dan kejadian
- Probabilitas dan kuantifikasi ketidakpastian
- Aksioma probabilitas, $0 \leq p(x) \leq 1$ dan $\sum_x p(x) = 1$
- Ekspektasi dan variansi
- Peluang bersyarat, aturan penjumlahan dan perkalian, dan aturan rantai
- Bayes' rule yang mengubah keyakinan berdasarkan observasi

- PMF
- Distribusi Bernoulli
- Distribusi Binomial
- Distribusi Poisson
- Maximum Likelihood Estimation

REFERENSI

-  Chris Piech (Sep. 2017)
Probability
<http://web.stanford.edu/class/archive/cs/cs109/cs109.1178/lectureHandouts/035-probability.pdf>
-  Chris Piech (Oct. 2017)
Conditional Probability
<http://web.stanford.edu/class/archive/cs/cs109/cs109.1178/lectureHandouts/040-cond-probability.pdf>
-  Chris Williams (Sep. 2015)
Probability - Machine Learning and Pattern Recognition
<https://www.inf.ed.ac.uk/teaching/courses/mlpr/2015/>

Terima kasih