

## PENGENALAN POLA

### PENDAHULUAN

Ali Akbar Septiandri

Universitas Al-Azhar Indonesia

*aliakbars@live.com*

February 21, 2020

### ① ADMINISTRASI

Tentang Perkuliahan  
Referensi

### ② PEMBELAJARAN MESIN & PENGENALAN POLA

Definisi

### ③ TUGAS DALAM MACHINE LEARNING

Klasifikasi  
Regresi  
Supervised Learning  
Clustering  
Representasi Data

## MATA KULIAH TERKAIT

### Prerequisites

- Statistika & Probabilitas, e.g. distribusi normal, ekspektasi
- Aljabar Linear, e.g. vektor dan nilai eigen, perkalian matriks
- Kalkulus, e.g. diferensial, integral
- Kecerdasan Buatan

### Paralel/Saran/Lanjutan

- Data/Text Mining
- Soft Computing

## ADMINISTRASI

## ATURAN PERKULIAHAN

- Materi bisa dilihat di <https://uai.aliakbars.id/mlpr/>
- Kuliah setiap hari Rabu, pukul 07.00-09.30
- Bahasa/teknologi pengantar: Python, PyLab, Jupyter Notebook, Google Colab

## ATURAN DALAM TUGAS

- Secara *default*, setiap tugas bersifat **individual**
- Silakan berdiskusi, tapi **jangan menyalin kode atau tulisan teman**
- **Keterlambatan pengumpulan** akan berakibat pada pengurangan nilai
- Pengumpulan tugas dilakukan melalui situs **e-learning**
- Kecurangan akan berakibat pada **nilai E** pada kuliah ini

## ATURAN PERKULIAHAN

- Terdapat 4 tugas
- Kuis yang tidak masuk komponen penilaian
- Ujian Tengah Semester dan Ujian Akhir Semester (tidak ada ujian perbaikan)
- Komponen nilai: 40% tugas, 30% UTS, 30% UAS

## REFERENSI

Buku dan materi daring yang bisa dijadikan referensi:

- ① VanderPlas, J. (2016). *Python Data Science Handbook*. O'Reilly Media. ([tersedia online](#))
- ② Deisenroth, M., Faisal, A.A., & Ong, C.S. (2019) *Mathematics for Machine Learning*. Cambridge University Press. ([tersedia online](#))
- ③ Bishop, C.M. (2007). *Pattern Recognition and Machine Learning*. Springer. ([tersedia online](#))
- ④ Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78-87. ([tersedia online](#))

Beberapa kuliah terkait:

- ① University of Edinburgh: Introductory Applied Machine Learning
- ② University of Edinburgh: Machine Learning and Pattern Recognition (graduate level)
- ③ Stanford CS229: Machine Learning
- ④ University of Oxford: Machine Learning (graduate level)
- ⑤ Max Planck Institute: Statistical Rethinking

### Sebelum UTS

- ① Intro to MLPR
- ② Probability
- ③ Discrete distributions
- ④ Gaussian distribution
- ⑤ Naïve Bayes
- ⑥ PCA
- ⑦ Linear models

### Setelah UTS

- ① Model evaluation
- ② SVM
- ③ Neural Networks
- ④ k-Nearest Neighbours
- ⑤ k-Means clustering
- ⑥ GMM
- ⑦ Kuliah tamu

## APA ITU MACHINE LEARNING?

## PEMBELAJARAN MESIN & PENGENALAN POLA

- Menemukan pola dalam data dan menggunakannya untuk melakukan prediksi
- Bagaimana cara menyelesaikan dengan menggunakan komputer?
  - Masalahnya, kita tidak tahu cara menulis programnya
  - ...tapi kita punya contoh data

## MENGAPA MACHINE LEARNING?

- Data ada di mana-mana
- ML mengombinasikan teori dan praktik
- Sudah berhasil menyelesaikan banyak kasus AI

Apa hubungannya dengan pengenalan pola?

## PENGENALAN POLA

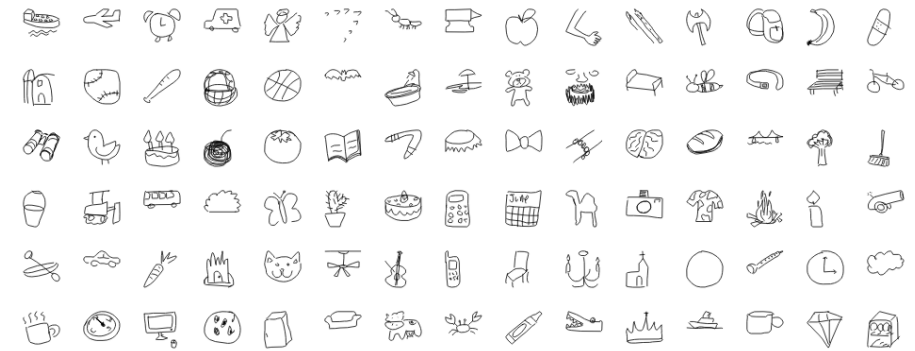
- Pengenalan pola, pembelajaran mesin, *data mining*, dan *knowledge discovery in databases* (KDD) terkadang sulit dibedakan
- Pengenalan pola awalnya dikenal dalam tugas-tugas *computer vision*
- *Pattern recognition*  $\neq$  *pattern matching*

## TUGAS DALAM MACHINE LEARNING

## KLASIFIKASI

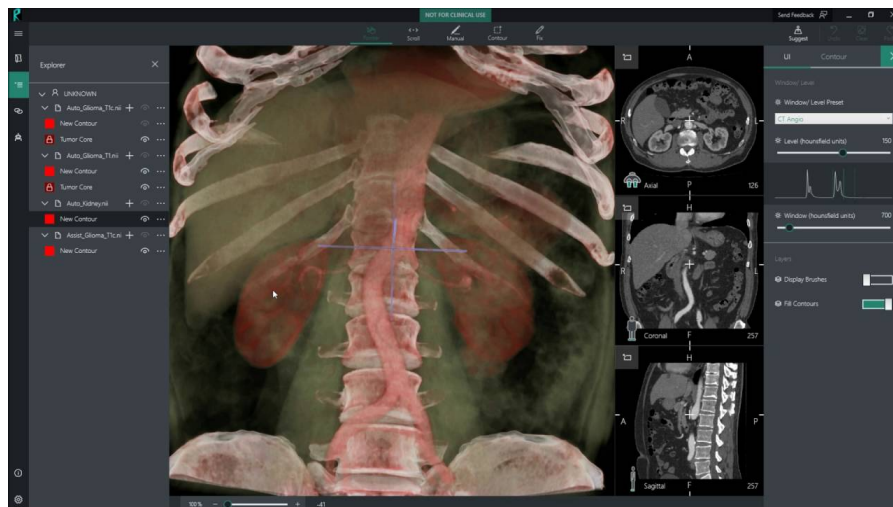
## QUICK, DRAW!

- ➊ Memprediksi nilai yang sudah pasti
- ➋ *Biasanya* direpresentasikan sebagai kelas biner  $\{0, 1\}$  atau  $\{-1, 1\}$
- ➌ Membutuhkan label
- ➍ Mempunyai *evaluation metrics* yang jelas, e.g. akurasi
- ➎ Contoh: identifikasi spam, MNIST digit recognition

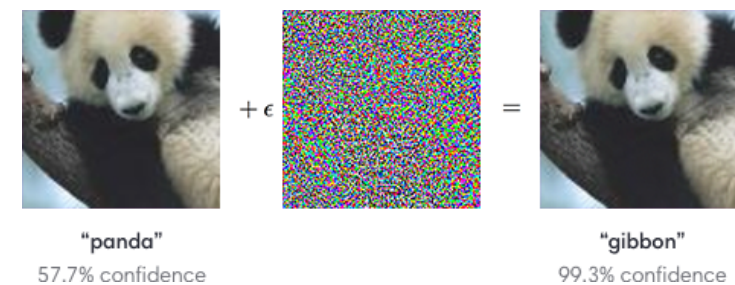


## GAMBAR: Quick, Draw! The Data dari Google

# INNEREYE RESEARCH PROJECT



# MISKLASIFIKASI DALAM PENGENALAN OBJEK



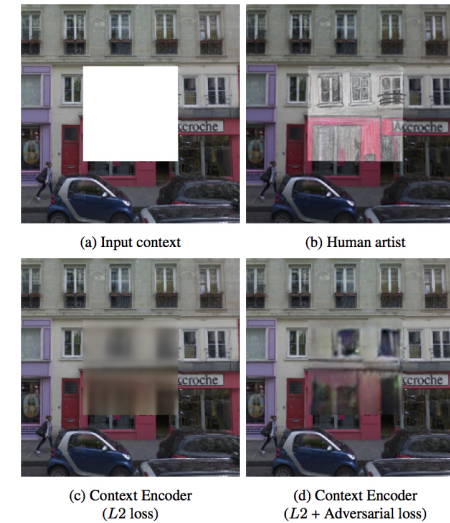
**GAMBAR:** Kesalahan model *machine learning* akibat *adversarial examples*

**GAMBAR:** Project InnerEye untuk membantu dokter mendeteksi penyakit pada organ dalam dari Microsoft

## REGRESI

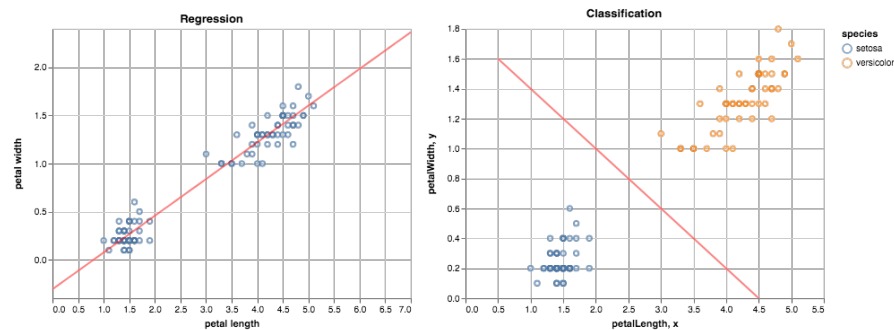
- 1 Membutuhkan label
- 2 Memprediksi nilai kontinu
- 3 *Evaluation metrics* berupa *error*, e.g. Mean Squared Error (MSE), Mean Absolute Error (MAE)
- 4 Contoh: prediksi nilai saham, jumlah RT dari suatu *tweet*

## INPAINTING



GAMBAR: Mengisi potongan gambar yang hilang [Pathak, 2016]

## KLASIFIKASI VS REGRESI



GAMBAR: Perbedaan klasifikasi dan regresi

## KLASIFIKASI DAN REGRESI

### FUNGSI

Kedua tugas ini dapat dilihat sebagai fungsi  $f$  yang memetakan atribut  $x$  ke label  $y$ .

## PROBABILITAS DALAM MACHINE LEARNING

- Asumsikan Anda diberi kasus berupa klasifikasi artikel berita. Anggap  $\mathbf{x}$  adalah dokumen, dan  $y$  adalah label.  $y \in \{\text{"Olahraga"}, \text{"Politik"}\}$
- Anda diminta membuat fungsi  $f$  dalam Java yang menerima masukan berupa  $\mathbf{x}$  dan mengeluarkan  $y$
- Jika saya akan membayar Anda Rp 1000 setiap artikel politik yang benar, dan Rp 1 juta untuk setiap artikel olahraga yang Anda dapat temukan dengan benar, bagaimana Anda akan membuat  $f$ ?

## PROBABILITAS DALAM MACHINE LEARNING

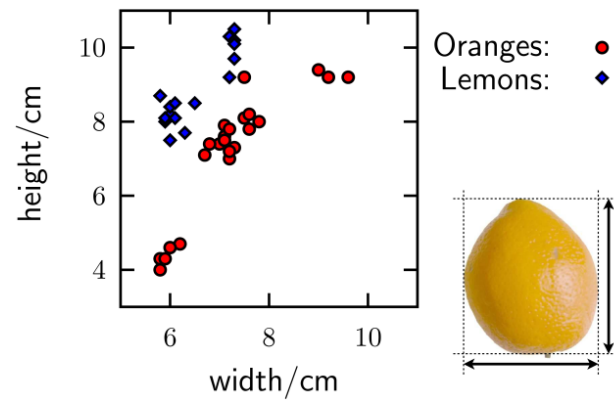
- Agar lebih sulit, bagaimana kalau Anda akan didenda Rp 10.000 setiap ada dokumen yang salah diklasifikasi? Apa yang Anda akan lakukan?
- Jawabannya: **Jangan buat fungsi**. Definisikan probabilitas  $p(y|\mathbf{x})$ . Lalu, ambil keputusan yang dapat memaksimalkan keuntungan Anda.

## CLUSTERING

“No silver bullet”

- 1 Mencoba memberikan deskripsi terhadap data
- 2 Tidak berhubungan dengan label
- 3 Menemukan pola yang “menarik” dalam data
- 4 Tidak mempunyai *evaluation metrics* yang pasti

## CONTOH CLUSTERING



GAMBAR: *Clustering* buah lemon dan jeruk [Murray, 2011]

Bagaimana merepresentasikan datanya?

## REPRESENTASI DATA

- dokumen?
- gambar?
- video?
- suara?

## REPRESENTASI DATA




- dokumen? *bag of words, word embedding*
- gambar? *pixels, regions*
- video? *pixels in frames*
- suara? *MFCC*



- ML & PR ada di mana-mana dan berguna untuk prediksi dalam skala besar
- Butuh pemahaman teoretis untuk dapat memahami algoritma dan membuat yang baru
- “No silver bullet”
- Buat model, bukan algoritma
- Mungkin butuh pandangan probabilistik

- Probabilitas
- Peubah acak
- Ekspektasi
- Peluang bersyarat
- Bayes’ rule

## REFERENSI

-  Ian J. Goodfellow et al. (2015)  
Explaining and Harnessing Adversarial Examples  
ICLR 2015
-  Deepak Pathak et al. (2016)  
Context Encoders: Feature Learning by Inpainting  
CVPR 2016
-  Iain Murray (2011)  
Oranges, Lemons and Apples dataset  
[http://homepages.inf.ed.ac.uk/imurray2/teaching/oranges\\_and\\_lemons/](http://homepages.inf.ed.ac.uk/imurray2/teaching/oranges_and_lemons/)

Terima kasih