

Program Fresh Graduate Academy Digital Talent Scholarship 2019 | Machine Learning

Dasar Machine Learning : Intro

Nama pembicara dengan gelar





Bagian 1

Pengenalan Machine Learning



Apa itu Machine Learning

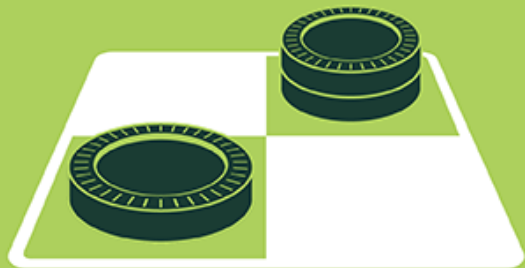
Machine Learning merupakan sub-area dari Ilmu Komputer yang mampu memberikan komputer **“kemampuan untuk belajar tanpa diprogram secara eksplisit”**

Arthur Samuel

Peneliti IBM di area Computer Gaming dan Artificial Intelligence, 1959

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.



AI vs. ML vs. DL

Artificial Intelligence

Setiap teknik yang membuat komputer dapat memiliki pengetahuan seperti manusia

Machine Learning

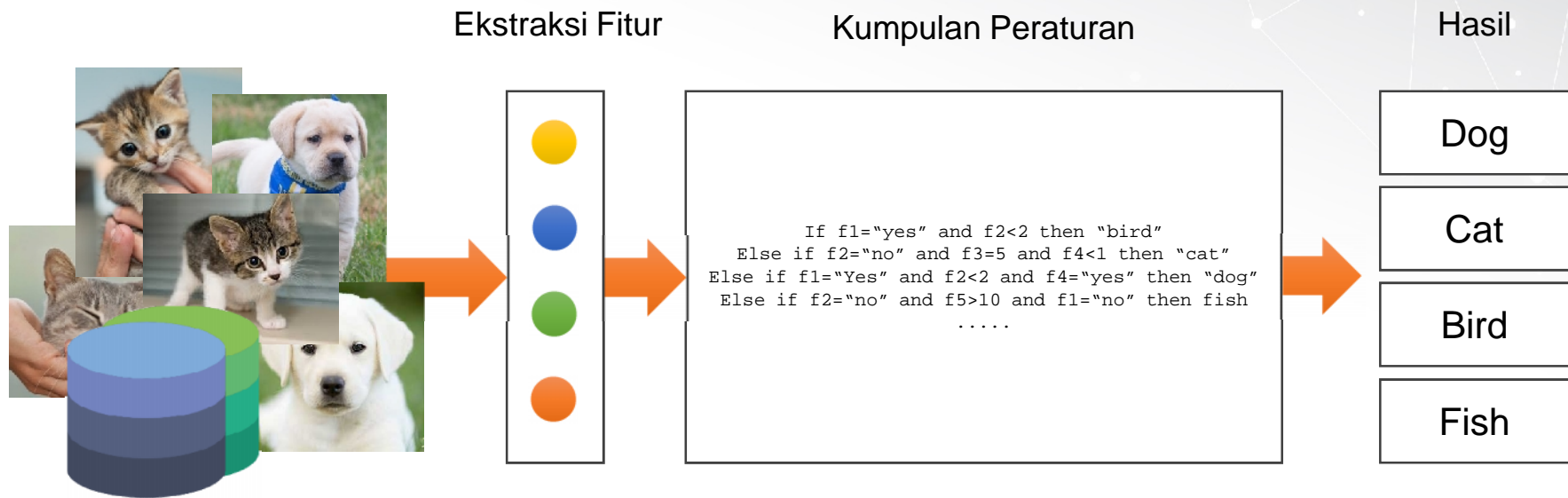
Teknik untuk mengajari komputer tanpa secara langsung memprogram

Deep Learning

Belajar untuk memahami fitur dari data-data dengan menggunakan neural network (Jaringan syaraf tiruan)

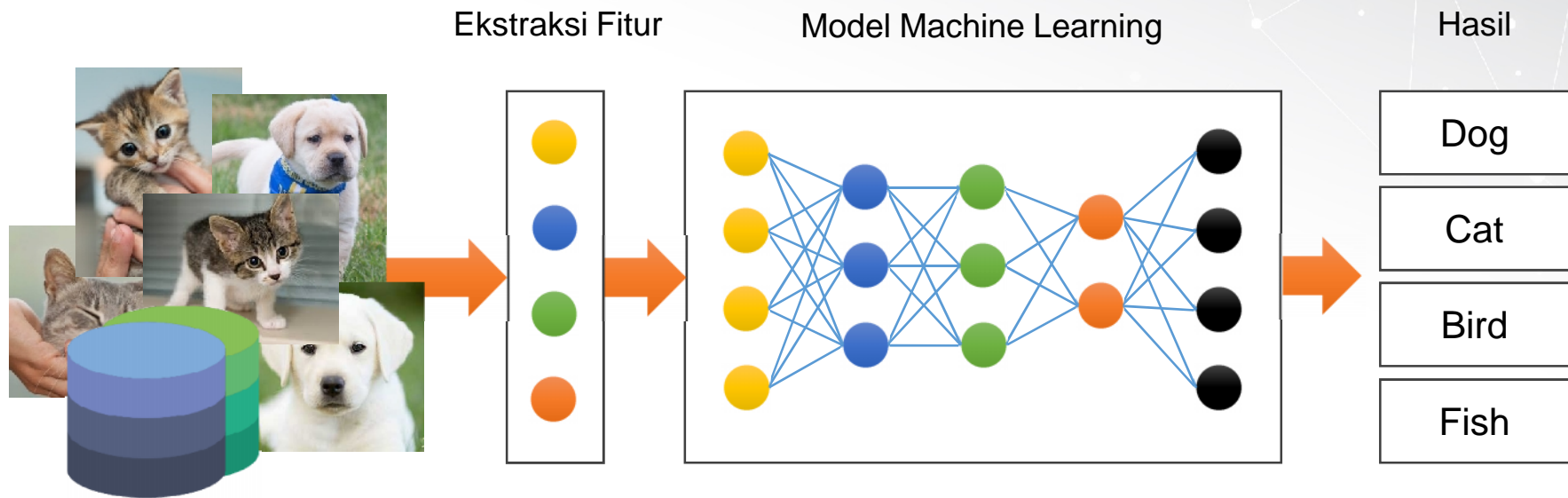


Pendekatan Tradisional





Pendekatan Machine Learning





Manfaat Machine Learning

- Menyederhanakan permasalahan
 - *Traditional approach*: menggunakan berbagai macam rule
 - *Machine learning approach*: menggunakan beberapa baris kode serta dapat diterapkan di problem lain
- Machine learning dapat beradaptasi dengan data baru sedangkan traditional approach mengharuskan merubah banyak rule
- Mendapatkan wawasan tentang masalah kompleks dan data dalam jumlah besar.



Bagian 2



Dataset Problem



Anjing Komondor atau Pel





Muffin atau Cihuahua





Kue atau Parrot





Trump atau Ayam





Bagian 3

Contoh Kasus Machine Learning



Contoh Kasus Bidang Kedokteran

Ini adalah sampel sel manusia yang diekstraksi dari seorang pasien

Is this a benign or malignant cell?

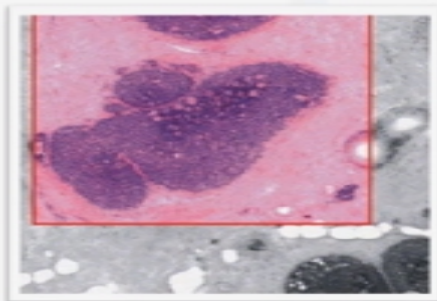
ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000015	6	1	1	1	7	1	3	1	1	



Contoh Kasus Bidang Kedokteran

Sel kanker memiliki beberapa karakteristik seperti besar, bentuk, adhesi marjinal, dll

Is this a benign or malignant cell?



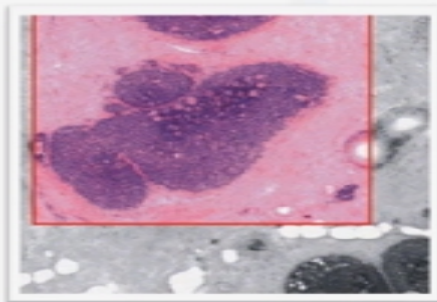
ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000015	6	1	1	1	7	1	3	1	1	



Contoh Kasus Bidang Kedokteran

Bagaimana cara kita mengenali sel ini jinak atau ganas?

Is this a benign or malignant cell?



Benign
or
Malignant?

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000015	6	1	1	1	7	1	3	1	1	



Contoh Kasus Bidang Kedokteran

Gunakan Dataset yang berisi karakteristik ribuan Sampel sel manusia diekstraksi dari pasien yang diyakini berisiko berkembang kanker.

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000025	5	1	1	1	2	1	3	1	1	benign
1002945	5	4	4	5	7	10	3	2	1	benign
1015425	3	1	1	1	2	2	3	1	1	malignant
1016277	6	8	8	1	3	4	3	7	1	benign
1017023	4	1	1	3	2	1	3	1	1	benign
1017122	8	10	10	8	7	10		7	1	malignant
1018099	1	1	1	1	2	10	3	1	1	benign
1018561	2	1	2	H	2	1	3	1	1	benign
1033078	2	1	1	1	2	1	1	1	5	benign
1033078	4	2	1	1	2	1	2	1	1	benign



Contoh Kasus Bidang Kedokteran

Analisis data asli yang menunjukkan bahwa terdapat karakteristik berbeda secara signifikan antara sampel jinak dan ganas.

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000025	5	1	1	1	2	1	3	1	1	benign
1002945	5	4	4	5	7	10	3	2	1	benign
1015425	3	1	1	1	2	2	3	1	1	malignant
1016277	6	8	8	1	3	4	3	7	1	benign
1017023	4	1	1	3	2	1	3	1	1	benign
1017122	8	10	10	8	7	10		7	1	malignant
1018099	1	1	1	1	2	10	3	1	1	benign
1018561	2	1	2	H	2	1	3	1	1	benign
1033078	2	1	1	1	2	1	1	1	5	benign
1033078	4	2	1	1	2	1	2	1	1	benien

Modeling

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000015	6	1	1	1	7	1	3	1	1	

Prediction 




Contoh Kasus Bidang Kedokteran

Setelah model telah dilatih secara berulang, dapat digunakan untuk memprediksi sel baru, dengan akurasi yang baik.

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000025	5	1	1	1	2	1	3	1	1	benign
1002945	5	4	4	5	7	10	3	2	1	benign
1015425	3	1	1	1	2	2	3	1	1	malignant
1016277	6	8	8	1	3	4	3	7	1	benign
1017023	4	1	1	3	2	1	3	1	1	benign
1017122	8	10	10	8	7	10		7	1	malignant
1018099	1	1	1	1	2	10	3	1	1	benign
1018561	2	1	2	H	2	1	3	1	1	benign
1033078	2	1	1	1	2	1	1	1	5	benign
1033078	4	2	1	1	2	1	2	1	1	benien

Modeling

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000015	6	1	1	1	7	1	3	1	1	Benign

Prediction 

Accuracy = 89%

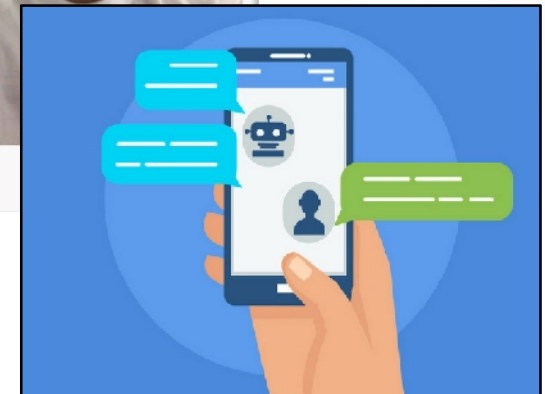
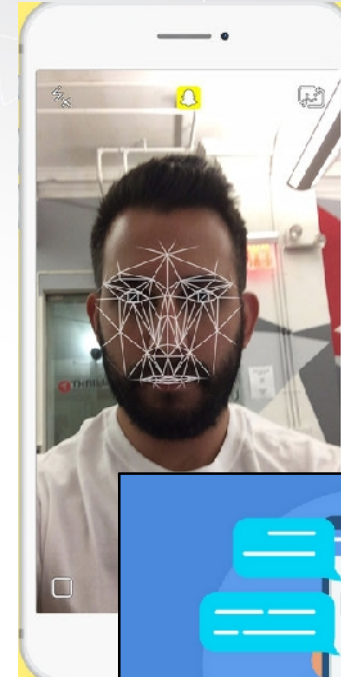
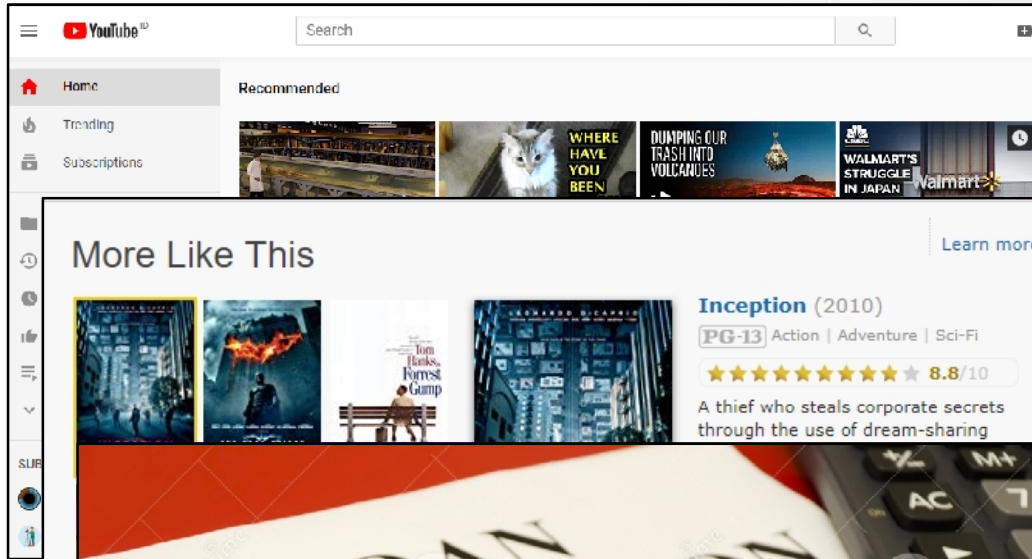


Bagian 4

Beberapa Aplikasi dan Teknik Machine Learning



Aplikasi Machine Learning





Beberapa Teknik Machine Learning

- **Regression**
 - Memprediksi nilai kontinyu.
 - Memprediksi harga rumah berdasarkan karakteristik.
- **Classification**
 - Mengelompokkan data ke dalam kategori tertentu berdasarkan label tertentu.
 - Mengelompokkan sel jinak dan ganas.
- **Clustering**
 - Menemukan struktur intrinsik dari data.
 - Pengelompokkan pasien rumah sakit yang serupa.
- **Associations**
 - Mengaitkan item yang sering terjadi bersama.
 - Barang grosir apa yang biasanya dibeli bersamaan.



Beberapa Teknik Machine Learning

- **Deteksi Anomali**
 - Menemukan kasus abnormal yang tidak biasanya terjadi dalam data.
 - Mendeteksi penipuan kartu kredit
- **Dimension Reduction**
 - Mengurangi besar dimensi dari data.
- **Sistem Rekomendasi**
 - Merekomendasikan item sesuai dengan perilaku pengguna.
 - Sistem rekomendasi video yang dilakukan youtube

Program Fresh Graduate Academy Digital Talent Scholarship 2019 | Machine Learning

Dasar Machine Learning : Python

Nama pembicara dengan gelar





Library Python untuk Machine Learning





Library Python untuk Machine Learning



- Numpy, merupakan library matematika untuk bekerja dengan array n-dimensional dalam Python.
- Numpy untuk melakukan perhitungan secara efisien dan efektif.
- Numpy lebih baik daripada python biasa karena kemampuannya yang luar biasa. Misalnya, untuk bekerja dengan array, kamus, fungsi, tipe data, dan bekerja dengan gambar.



Library Python untuk Machine Learning



- SciPy adalah kumpulan algoritma numerik dan toolbox domain spesifik, termasuk sinyal pemrosesan, optimisasi, statistik, dan banyak lagi.
- SciPy adalah library yang baik untuk perhitungan ilmiah dan berkinerja tinggi.



Library Python untuk Machine Learning

- Matplotlib adalah library paket plot yang sangat populer yang menyediakan plot 2D dan juga plot 3D.
- Matplotlib dapat digunakan di dalam script Python, shell Python dan ipython (ala MATLAB or Mathematica), server aplikasi web, dan enam GUI toolkit.
- Dengan Matplotlib dapat menjadi lebih mudah membuat plot, histogram, power spectra, grafik batang, grafik error, scatterplot, dll, hanya dengan beberapa baris kode.

The image shows the 'matplotlib' logo in a blue, lowercase, sans-serif font. The letter 'o' is replaced by a circular icon containing a colorful, multi-segmented plot or chart.



Library Python untuk Machine Learning



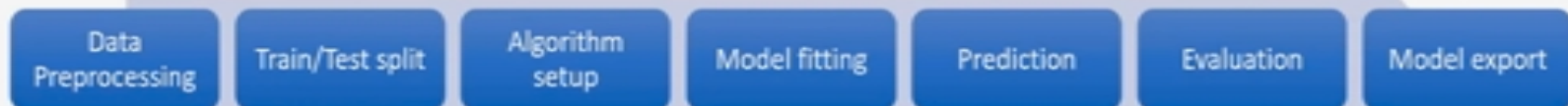
- Library Panda, adalah pustaka python tingkat tinggi yang memberikan kinerja tinggi, mudah untuk penggunaan data terstruktur.
- Library Panda memiliki banyak fungsi untuk mengimpor, memanipulasi, dan menganalisis data.
- Secara khusus, Panda menawarkan struktur data dan operasi untuk memanipulasi tabel numerik dan deret waktu.



Library Python untuk Machine Learning

Beberapa tentang Scikit Learn

1. Library Machine Learning free software
2. Terdiri dari algoritma: Klasifikasi, Regresi dan algoritma klustering
3. Bekerja dengan Numpy dan Scipy
4. Dokumentasi yang baik
5. Mudah diimplementasikan





Library Python untuk Machine Learning

Fungsi Scikit Learn

```
from sklearn import preprocessing  
X = preprocessing.StandardScaler().fit(X).transform(X)
```

Fungsi Scikit Learn untuk mengubah vektor fitur mentah menjadi vektor yang sesuai untuk pemodelan.



Library Python untuk Machine Learning

Fungsi Scikit Learn

```
from sklearn.model_selection import train_test_split  
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33)
```

Fungsi Scikit untuk membagi dataset menjadi data train dan data tes untuk melatih model, lalu uji akurasi model secara terpisah.



Library Python untuk Machine Learning

Fungsi Scikit Learn

```
from sklearn import preprocessing  
X = preprocessing.StandardScaler().fit(X).transform(X)
```

```
from sklearn.model_selection import train_test_split  
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33)
```

```
from sklearn import svm  
clf = svm.SVC(gamma=0.001, C=100.)
```

```
clf.fit(X_train, y_train)
```

Fungsi Classifier clf dapat melatih model dengan set data. Dengan melewati set pelatihan ke metode fit, model CLF belajar mengklasifikasikan kasus yang tidak diketahui.



Library Python untuk Machine Learning

Fungsi Scikit Learn

```
from sklearn import preprocessing  
X = preprocessing.StandardScaler().fit(X).transform(X)
```

```
from sklearn.model_selection import train_test_split  
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33)
```

```
from sklearn import svm  
clf = svm.SVC(gamma=0.001, C=100.)
```

```
clf.fit(X_train, y_train)
```

```
→ clf.predict(X_test)
```

Selanjutnya, kita bisa menggunakan data test untuk menjalankan prediksi. Dan, hasilnya diperoleh kelas apa dari masing-masing nilai.



Library Python untuk Machine Learning

Fungsi Scikit Learn

```
from sklearn import preprocessing  
X = preprocessing.StandardScaler().fit(X).transform(X)
```

```
from sklearn.model_selection import train_test_split  
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33)
```

```
from sklearn import svm  
clf = svm.SVC(gamma=0.001, C=100.)
```

```
clf.fit(X_train, y_train)
```

```
clf.predict(X_test)
```

```
from sklearn.metrics import confusion_matrix  
print(confusion_matrix(y_test, yhat, labels=[1,0]))
```

```
import pickle  
s = pickle.dumps(clf)
```