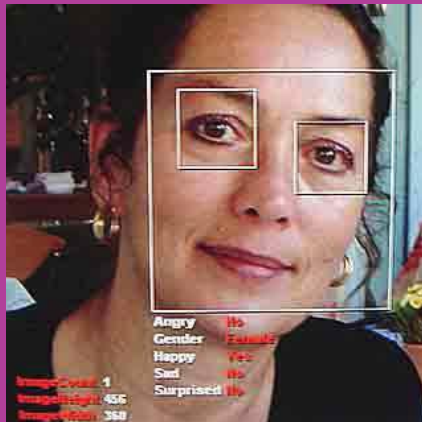


Gunadarma
UG University



INTRODUCTION TO MACHINE LEARNING

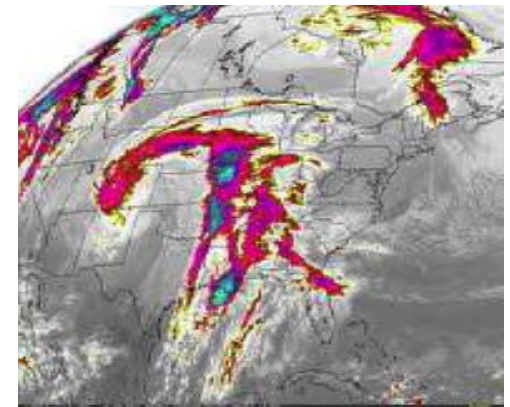
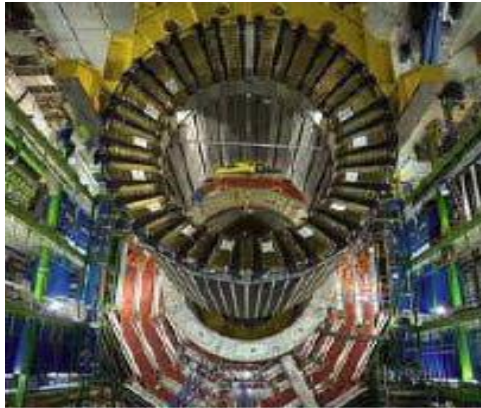
Pengantar Deep Learning

Belajar dari Data

Dunia didorong oleh data.

- Pusat penelitian iklim Jerman menghasilkan 10 petabyte per tahun
- Google memproses 24 petabyte per hari
- Large Hadron Collider menghasilkan 60 gigabyte per menit (~ 12 DVD)
- Ada lebih dari 50 juta transaksi kartu kredit sehari di USA saja.

Belajar dari Data



Belajar dari Data

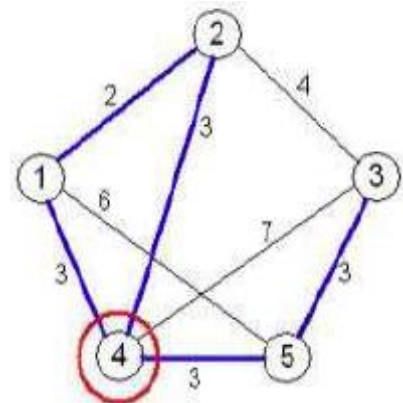
Data direkam dari beberapa fenomena dunia nyata.

Apa yang mungkin ingin kita lakukan dengan data itu?

- Ramalan
 - Apa yang bisa kita prediksi tentang fenomena ini?
- Deskripsi
 - Bagaimana kita bisa menggambarkan / memahami fenomena ini dengan cara baru?



AHRO Prevention Quality Indicators						
Dehydration Admission Rate (PQI 10)						
Counties highlighted in GREEN are significantly lower than the National Average.						
Counties highlighted in RED are significantly higher than the National Average.						
County Name	Cases	Population	Cure Rate	Risk Adj. Rate	Risk Adjusted Rate	Risk Adj. Rate
Adair	79	13,774	4.74	4.62	5.19	2.76
Allen	28	14,299	.96	1.41	2.89	2.59
Anderson	12	16,453	1.70	0.25	0.84	1.42
Ballard	8	6,538	.22	0.24	1.03	1.03
Barnes	102	31,112	3.28	2.56	2.93	3.31
Beth	15	8,943	.60	0.84	1.55	2.26
Bell	122	23,055	4.29	4.52	4.99	5.41
Boone	68	78,320	1.87	0.85	1.14	1.42
Bourbon	20	15,245	.31	0.73	1.28	1.81
Boyd	32	39,393	1.81	0.39	6.72	1.06
Boyle	32	22,387	.43	0.88	1.34	1.79
Bracken	18	6,760	2.69	1.79	2.63	3.47
Breathitt	40	12,381	3.23	2.84	3.59	4.15
Breckinridge	23	16,036	.53	0.94	1.60	2.07
Bullitt	23	52,112	4.44	0.23	6.55	0.93
Butler	9	10,366	1.87	0.18	0.86	1.54
Caldwell	13	10,281	.26	0.39	1.00	1.61
Callaway	28	29,186	1.95	0.53	0.90	1.38
Campbell	54	66,477	1.81	0.53	6.89	1.07
Carlisle	5	4,215	.19	0.00	0.93	1.89
Carroll	20	7,960	2.52	1.77	2.56	3.35
Carter	18	21,160	1.85	0.37	0.85	1.34
Cassidy	47	12,648	2.72	2.72	3.39	3.89



Belajar dari Data

Bagaimana cara mengekstrak pengetahuan dari data untuk membantu manusia mengambil keputusan?

Bagaimana kita bisa mengotomatisasi keputusan dari data?

Bagaimana kita dapat menyesuaikan sistem secara dinamis untuk memungkinkan pengalaman pengguna yang lebih baik?

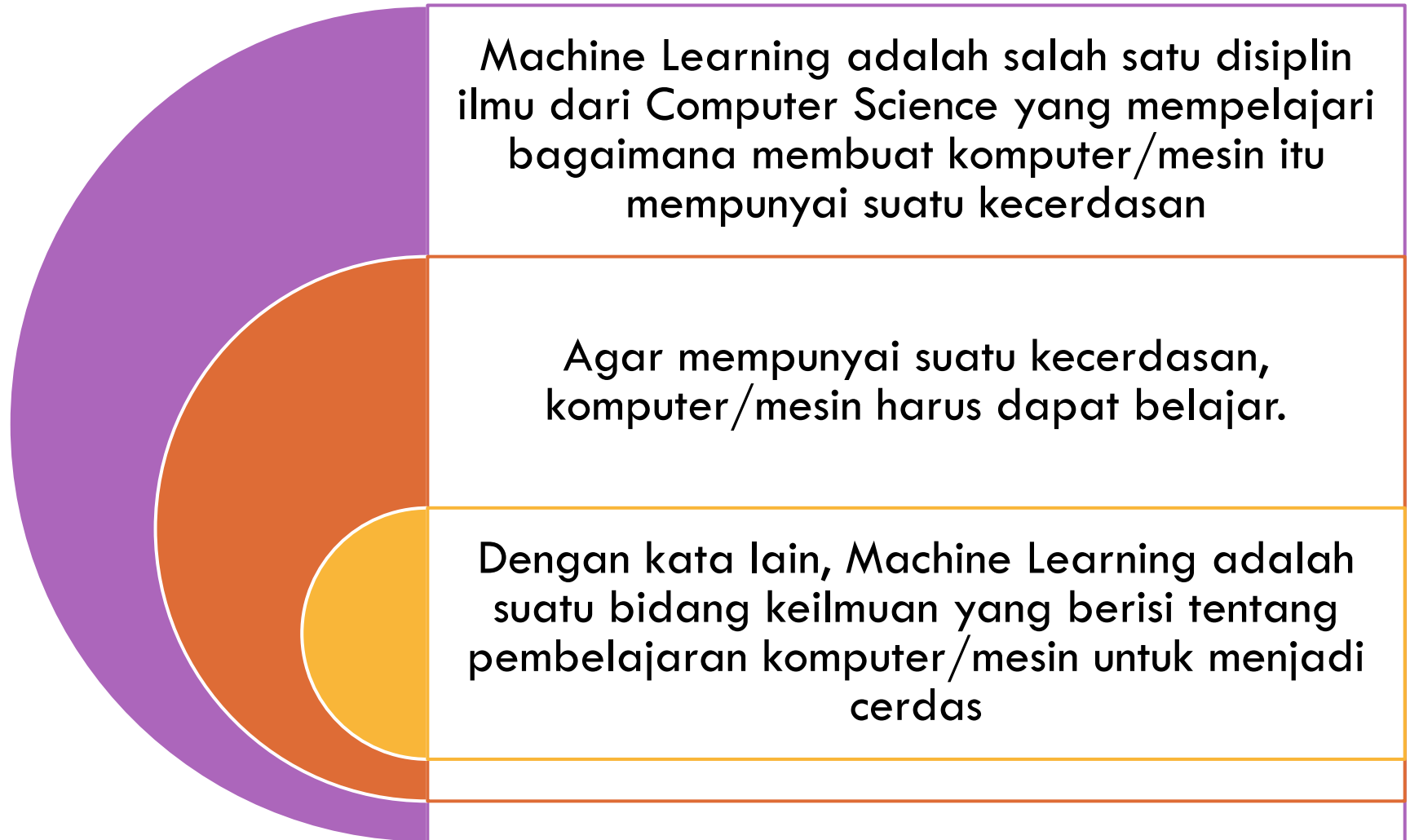
Write code to
explicitly do the
above tasks



Write code to make
the computer learn
how to do the tasks

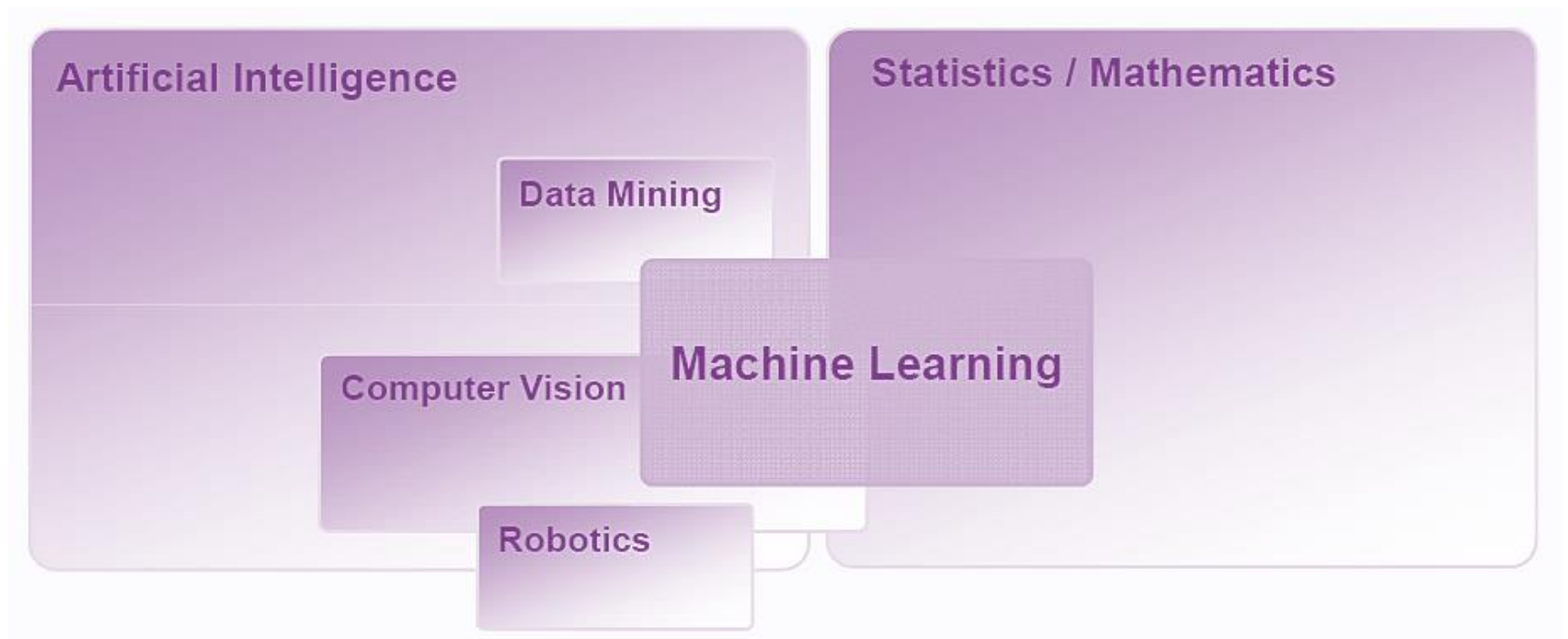


Apa itu Machine Learning?



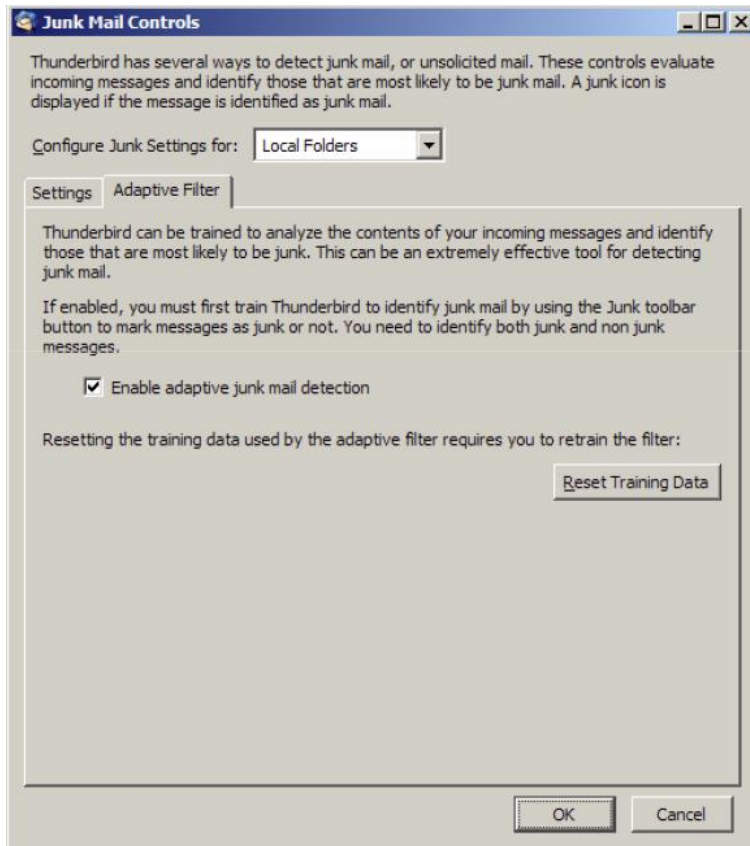
Machine Learning

Where does it fit? What is it not?



Machine Learning

Penggunaan Machine Learning untuk mendeteksi SPAM



To: you@gmail.com

GET YOUR DIPLOMA TODAY!

If you are looking for a fast and cheap way to get a diploma, this is the best way out for you.

Choose the desired field and degree and call us right now: For US: 1.845.709.8044 Outside US:

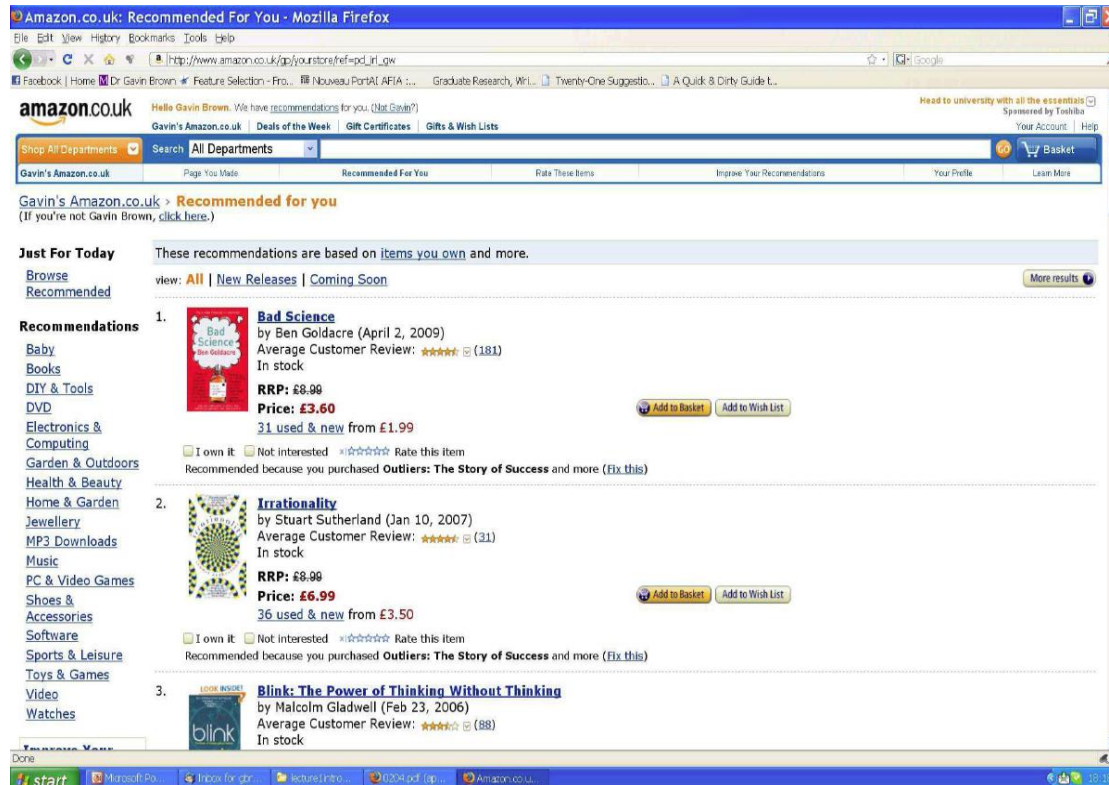
+1.845.709.8044 "Just leave your NAME & PHONE NO. (with CountryCode)" in the voicemail. Our staff will get back to you in next few days!

Algoritma yang dipakai:

- Naïve Bayes
- Rule Mining

Machine Learning

Penggunaan Machine Learning untuk rekomendasi buku

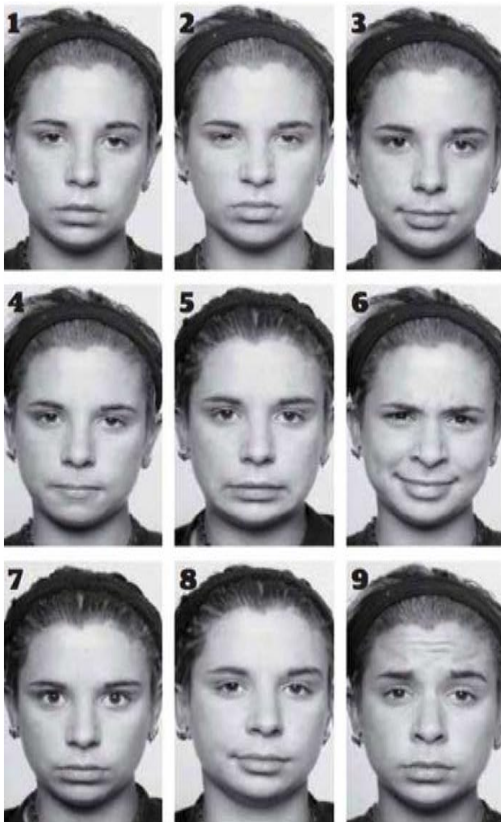


Algoritma yang digunakan

- Collaborative Filtering
- Nearest Neighbour
- Clustering

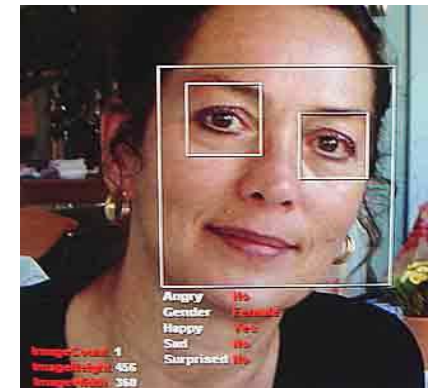
Machine Learning

Penggunaan Machine Learning untuk identifikasi wajah & ekspresi



Algoritma yang digunakan

- Decision Trees
- Adaboost

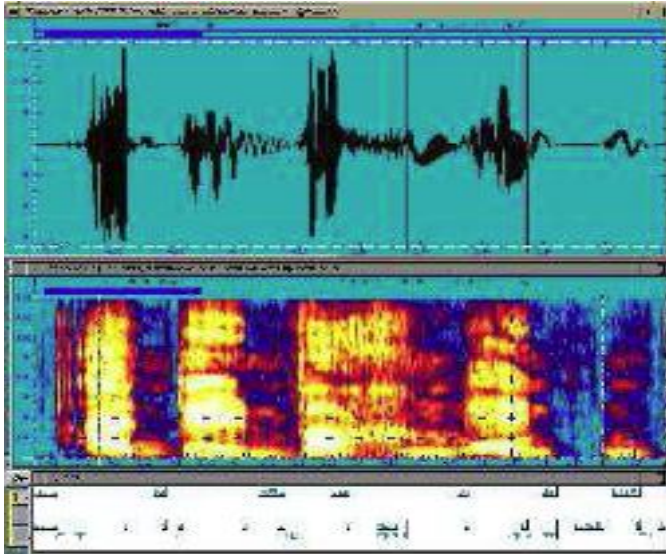


Angry	No
Gender	Female
Happy	Yes
Sad	No
Surprised	No

Image Count: 1
Image Height: 456
Image Width: 359

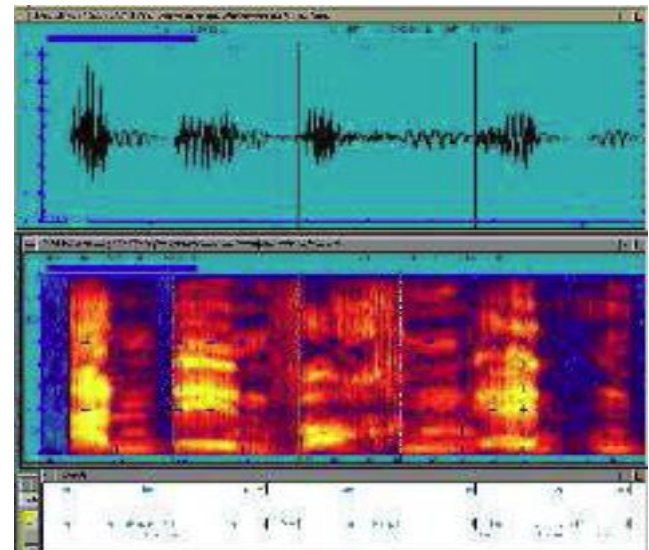
Machine Learning

Penggunaan Machine Learning untuk identifikasi pola suara



Algoritma yang digunakan

- Feature Extraction
- Probabilistic Classifiers
- Support Vector Machines



Machine Learning

ML untuk bekerja dengan data jaringan sosial:

Mendeteksi penipuan, memprediksi pola click-thru, iklan bertarget, dll.



Algoritma yang digunakan

- Support Vector Machines
- Collaborative filtering
- Rule mining algorithms
- dll

Machine Learning

- Mengendarai sebuah mobil
- Mengenali email spam
- Merekomendasikan buku
- Membaca tulisan tangan
- Mengenali ucapan, wajah, dll

Bagaimana akan menulis program-program ini?

Learning Process

Supervised
learning

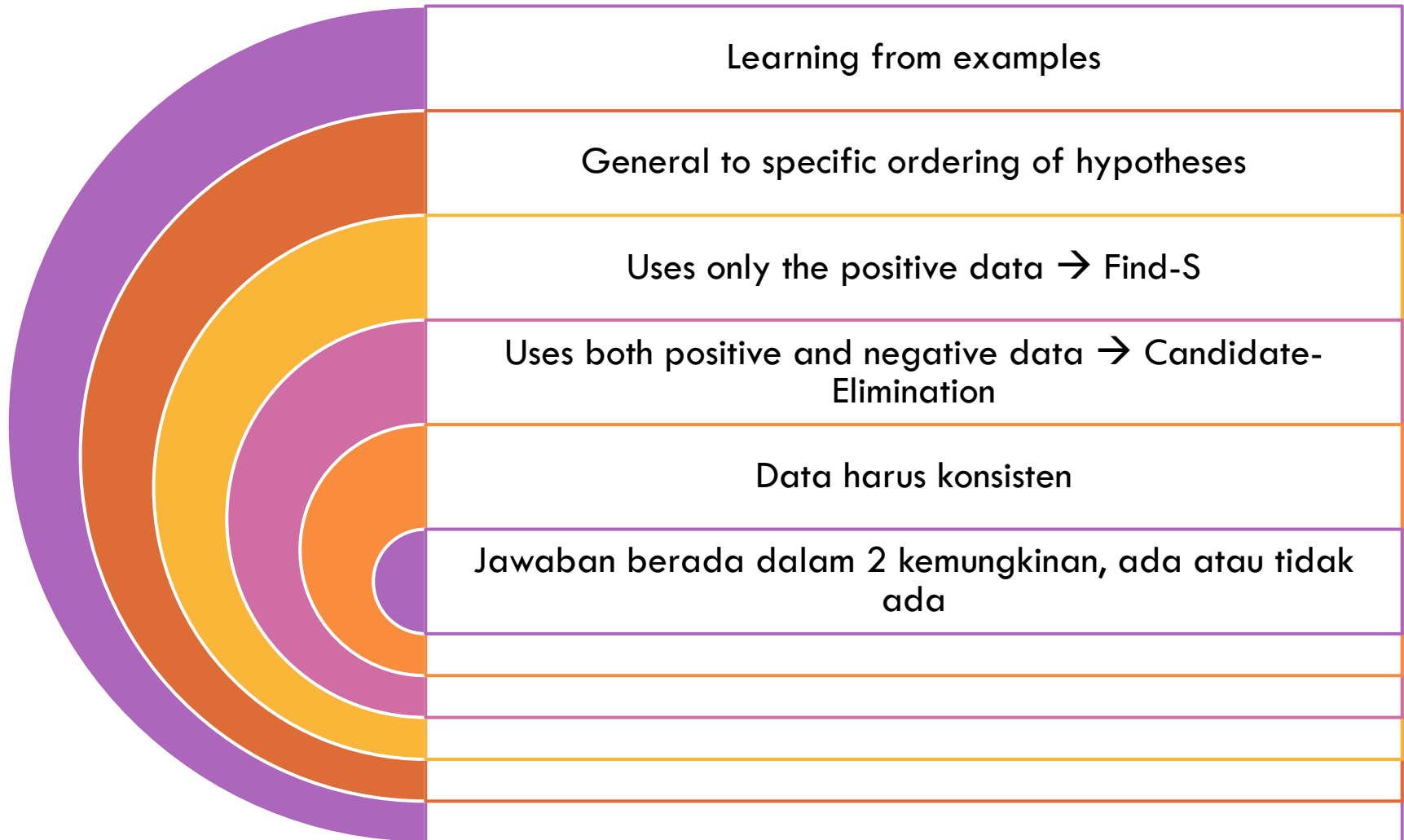
Unsupervised
learning

Reinforcement
learning

Pembahasan di dalam Machine Learning

- Concept learning
- Bayesian learning
- Instance based learning (clustering)
- Neural Networks
- Genetic Algorithm
- Reinforcement Learning
- Dan lain-lain

Concept Learning



Contoh Kasus

Data	Sky	Air Temp	Humidity	Wind	Water	Forecast	Enjoy Sport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

Bayesian Learning

- Learning from examples
- Termasuk ke dalam supervised learning
- Didasari pada Bayes Theorem
- Uses both positive and negative data
- Tidak mengharuskan data harus konsisten
- Jawaban ditunjukkan oleh nilai probabilitas
- Biasanya dipakai untuk fungsi-fungsi klasifikasi

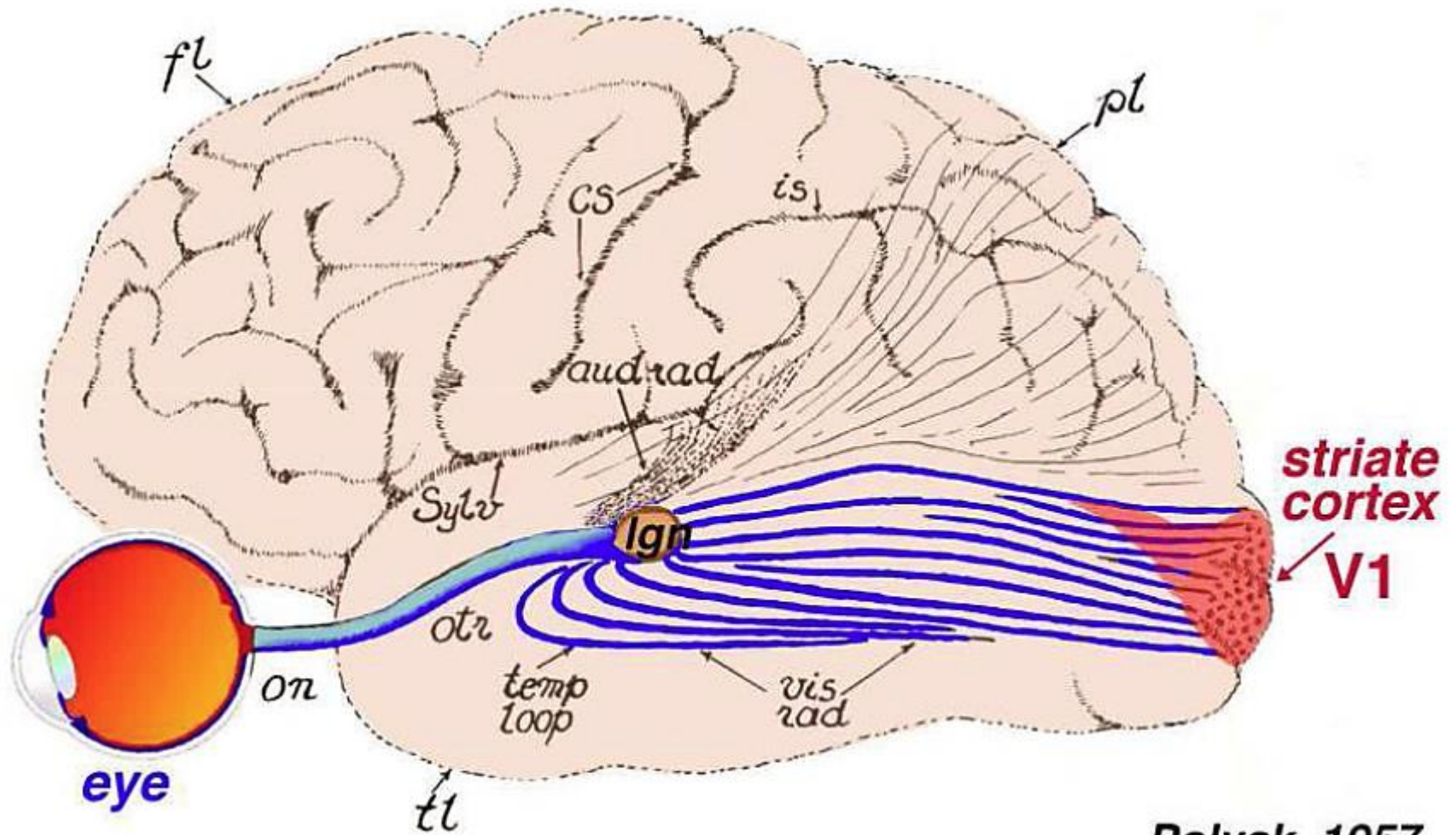
Instance Based Learning (Clustering)

- Tidak melibatkan jawaban dalam data
- Termasuk ke dalam unsupervised learning
- Hanya membuat suatu klasifikasi tanpa label/jawaban
- Labelisasi (pemberian jawaban) menjadi tanggung jawab user

Neural Networks

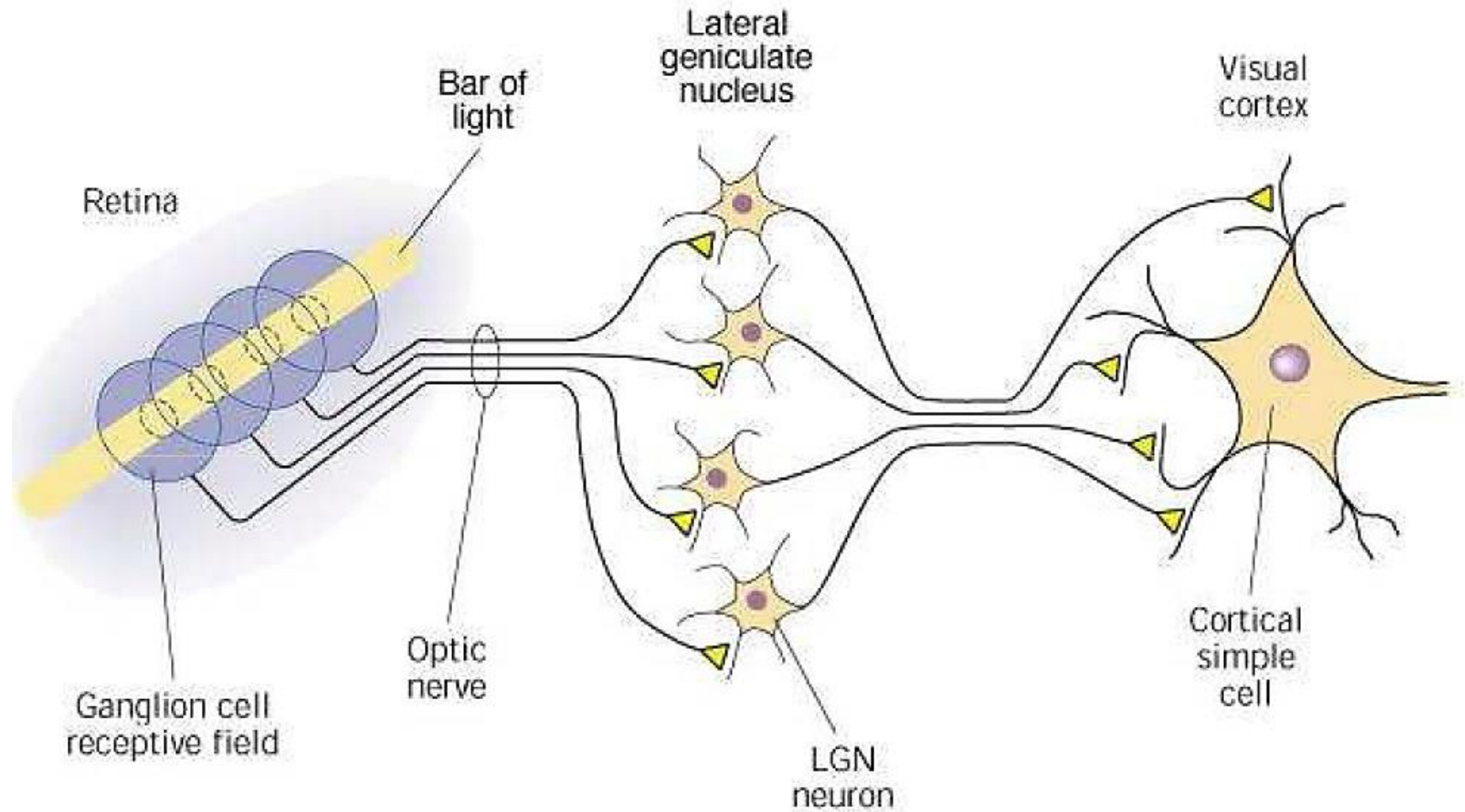
- Mensimulasikan kerja otak manusia
- Neuron adalah satuan unit pemroses terkecil pada otak
- Bentuk standard ini mungkin dikemudian hari akan berubah
- Jaringan otak manusia tersusun tidak kurang dari 10^{13} buah neuron yang masing-masing terhubung oleh sekitar 10^{15} buah dendrite
- Fungsi dendrite adalah sebagai penyampai sinyal dari neuron tersebut ke neuron yang terhubung dengannya
- Sebagai keluaran, setiap neuron memiliki axon, sedangkan bagian penerima sinyal disebut synapse
- Penjelasan lebih rinci tentang hal ini dapat diperoleh pada disiplin ilmu biology molecular
- Secara umum jaringan saraf terbentuk dari jutaan (bahkan lebih) struktur dasar neuron yang terinterkoneksi dan terintegrasi antara satu dengan yang lain sehingga dapat melaksanakan aktifitas secara teratur dan terus menerus sesuai dengan kebutuhan

Neural Networks

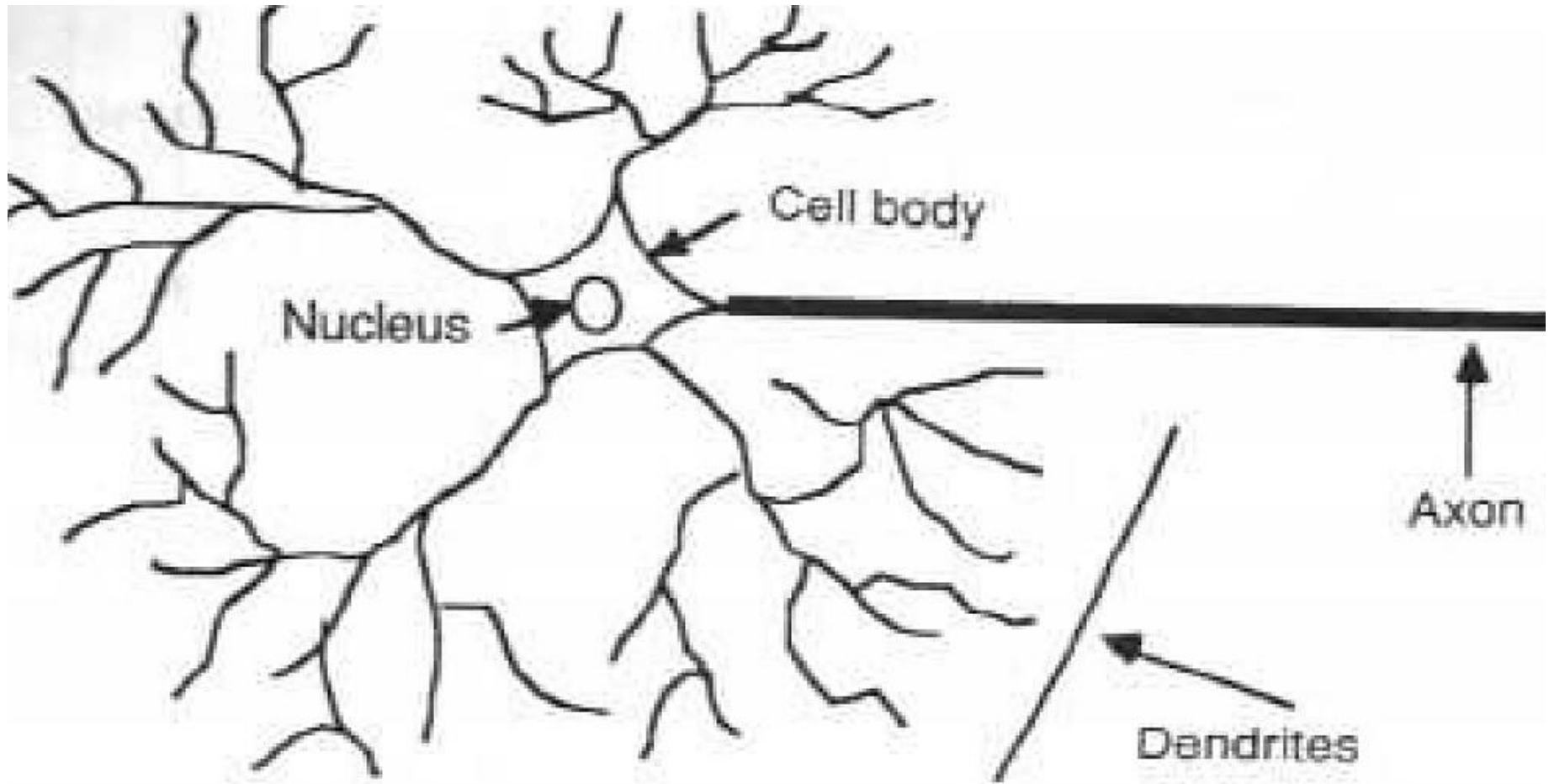


Polyak, 1957

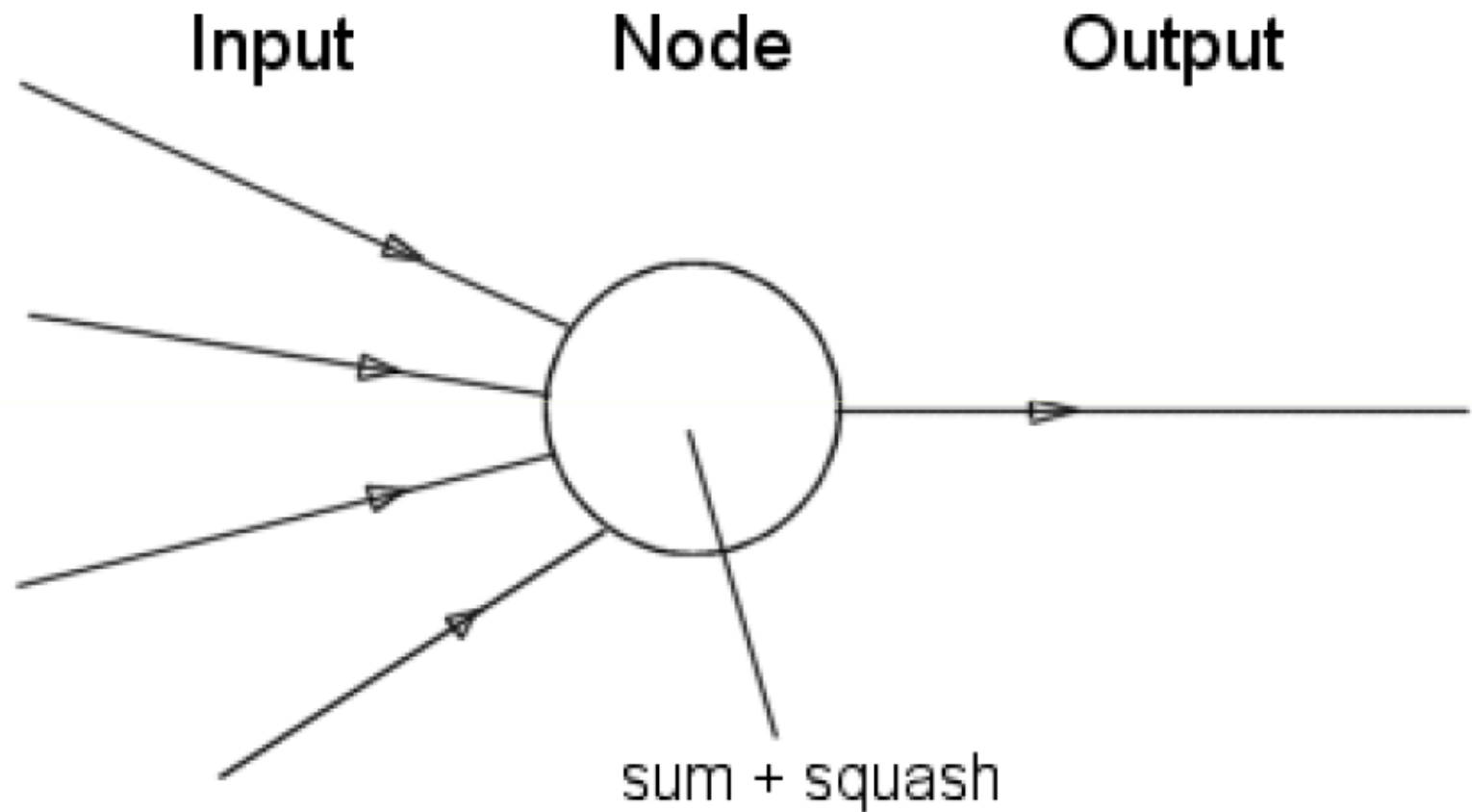
Neural Networks



Neural Networks



Neural Networks

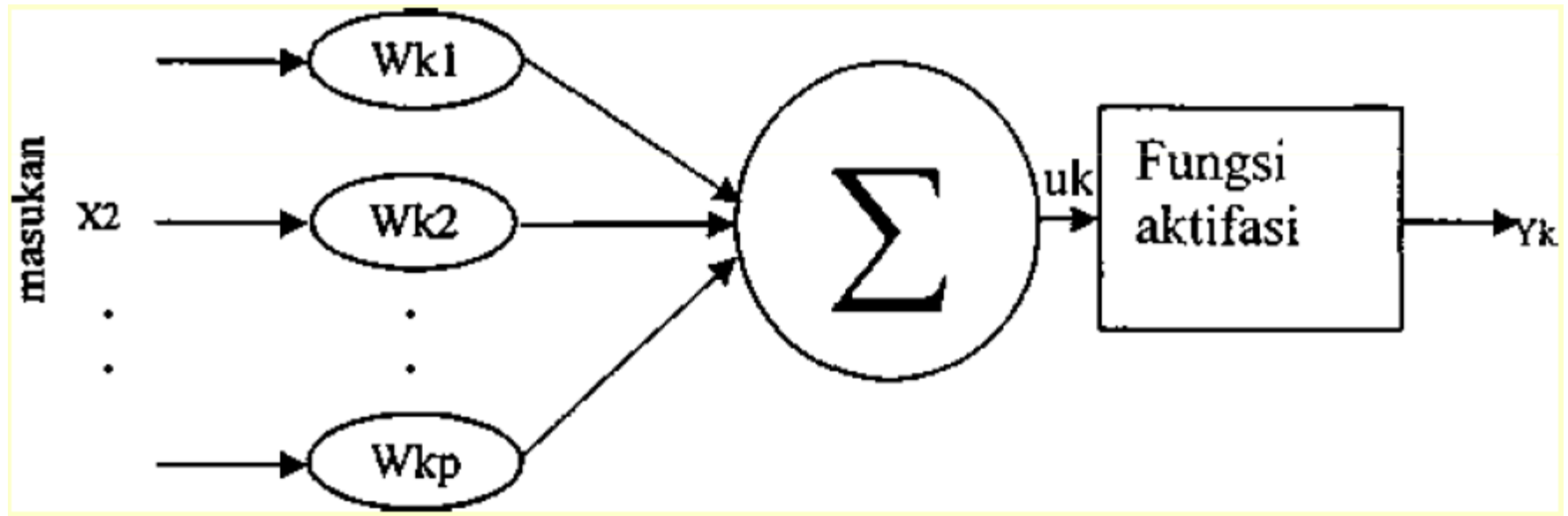


Neural Networks

PENIMBANG

PENJUMLAHAN

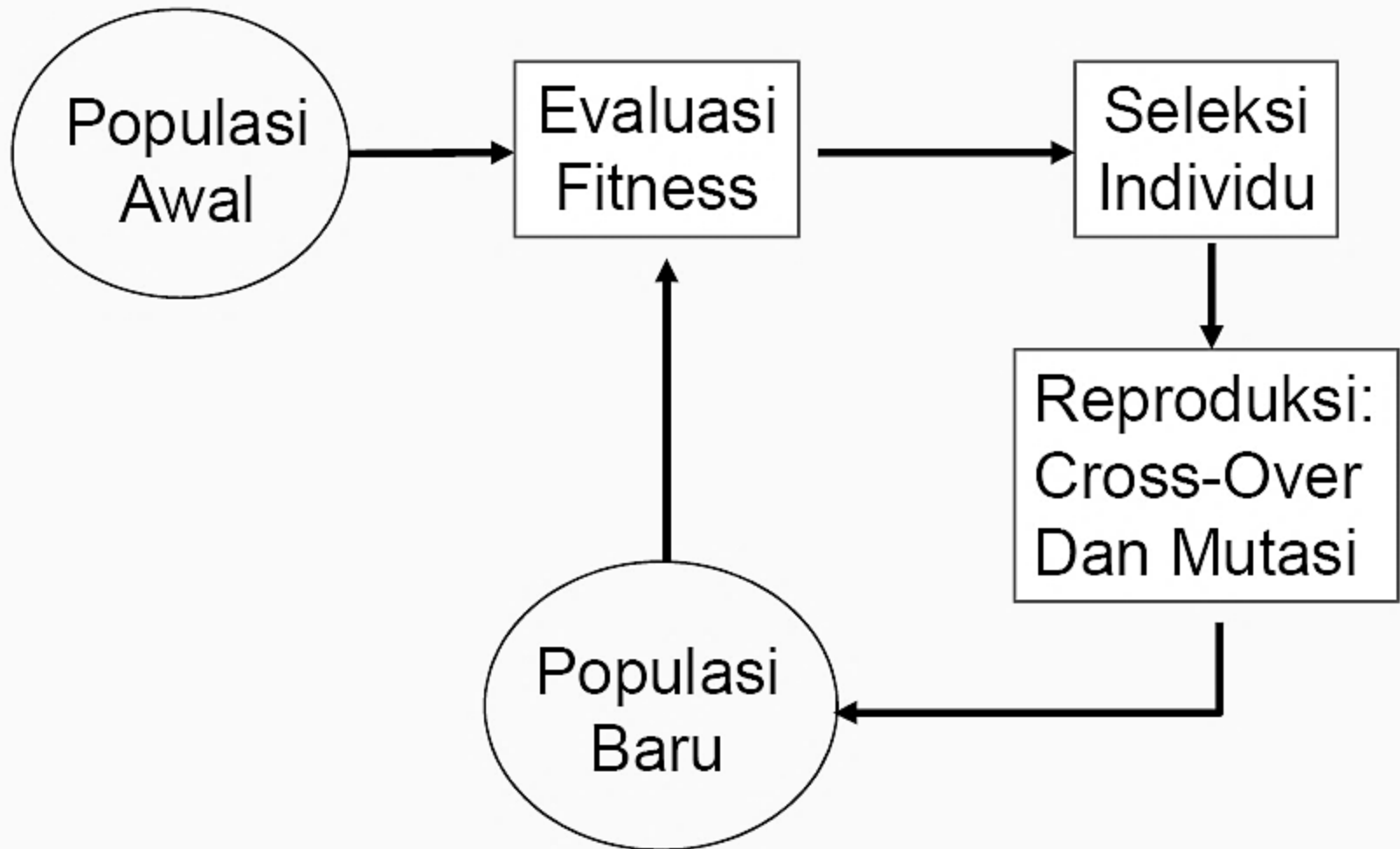
KELUARAN



Genetic Algorithm

- Algoritma Genetika adalah algoritma yang memanfaatkan proses seleksi alamiah yang dikenal dengan proses evolusi.
- Dalam proses evolusi, individu secara terus-menerus mengalami perubahan gen untuk menyesuaikan dengan lingkungan hidupnya. “Hanya individu-individu yang kuat yang mampu bertahan”.
- Proses seleksi alamiah ini melibatkan perubahan gen yang terjadi pada individu melalui proses perkembangbiakan. Dalam algoritma genetika ini, proses perkembang-biakan ini menjadi proses dasar yang menjadi perhatian utama, dengan dasar berpikir: “Bagaimana mendapatkan keturunan yang lebih baik”.

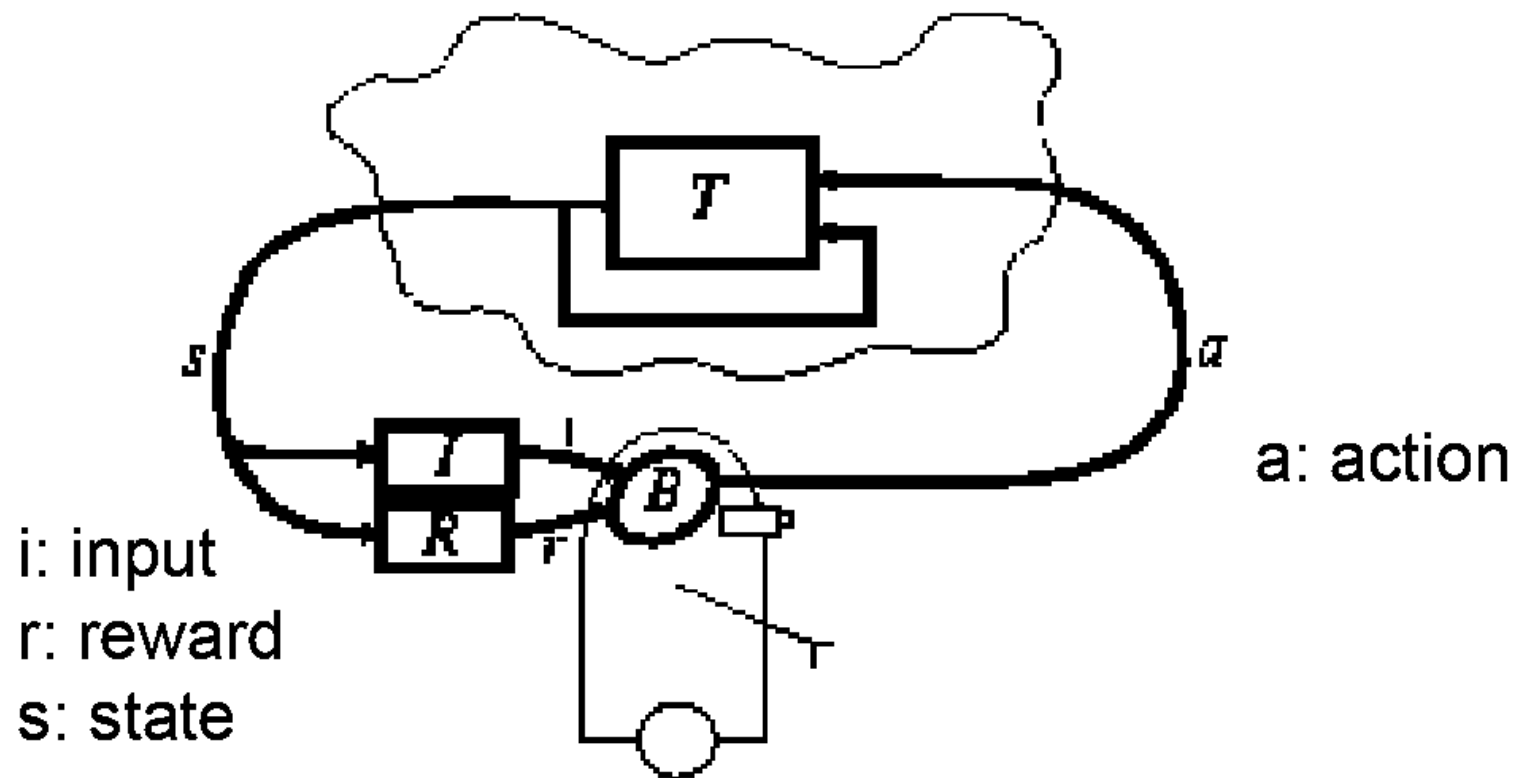
Siklus Genetic Algorithm



Reinforcement Learning

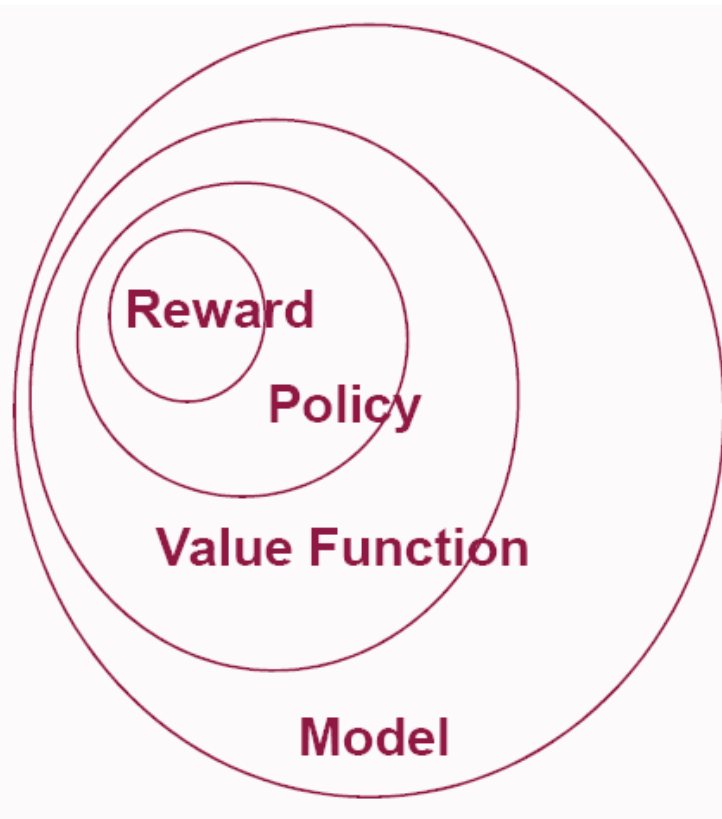
- Learning from experiences
- Memakai konsep reward dan punishment dalam proses learning

Reinforcement Learning



The standard reinforcement-learning model

Components of Reinforcement Learning



Reward

- Seberapa baik tindakan ini?

Policy

- Apa yang harus saya lakukan sekarang?

Value Function

- Seberapa baik keadaan ini?

Model

- Apa yang terjadi jika saya melakukan tindakan ini?

Referensi

- Introduction to Machine Learning
<http://www.cs.manchester.ac.uk/ugt/COMP24111>
- Machine Learning, Tom Mitchell, McGraw-Hill.
2008.
- Entin Martiana
<http://entin.lecturer.pens.ac.id/Machine%20Learning/Minggu%201%20Introduction%20to%20Machine%20Learning%202013.pdf>

TERIMA KASIH