



**PROGRAM STUDI  
TEKNIK INFORMATIKA – S1  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS DIAN NUSWANTORO**

MATA KULIAH  
**PEMBELAJARAN MESIN**



< a href='https://www.freepik.com/vectors/technology'>Technology vector created by sentavio - www.freepik.com</a>

# **PEMBELAJARAN MESIN**

## **“Pengantar Machine Learning”**

***TIM DOSEN PENGAMPU***

# KONTAK DOSEN

**ABU SALAM, M.KOM**

**EMAIL :**

[abu.salam@dsn.dinus.ac.id](mailto:abu.salam@dsn.dinus.ac.id)

**TELEGRAM:**

**0817244958**

**PROFIL :**

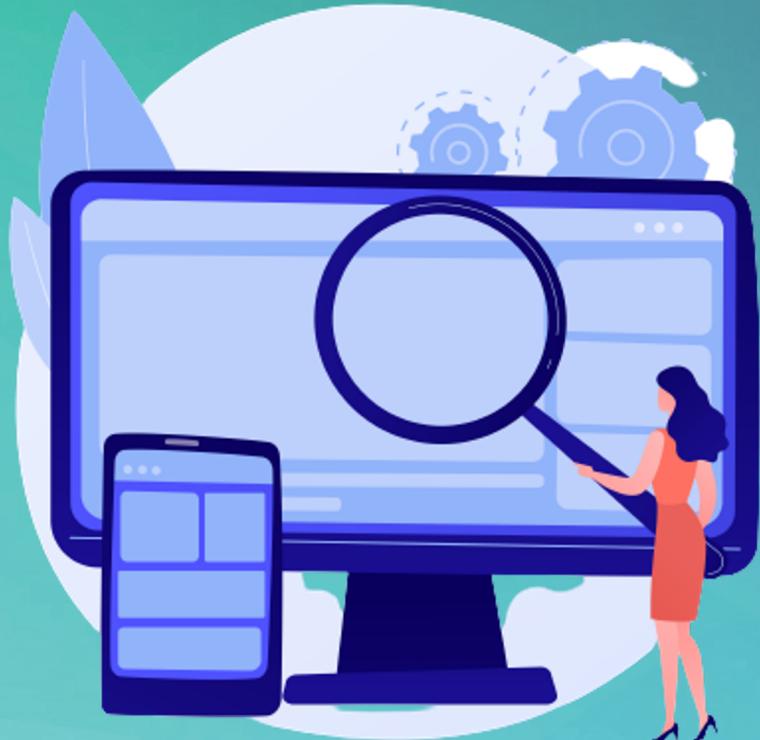
- Research Interest
  - Software Engineering, Machine Learning, Data Mining, Information Retrieval, Big Data, Data Science / Analysis.
- Activity
  - Dosen TI Fasilkom UDINUS (Bidang Kajian Sistem Cerdas)
  - Manager Departemen Software PT DINUSTEK
  - Direktur Teknis PT SSN (oncodoc.id)

## **MATERI PERKULIAHAN**

- Basic Concept (ML)
- Methode Learning (SL, USL, Semi SL, Reinforcement, Ensemble)
- SL (Class/Target/Output)
- Training – Validation – Testing
- Bias, Variance, Diversity, Outlier
- Underfitting, Overfitting
- Evaluasi Model (ConMat >> Accuracy, Recall, Precision, recall. >> rmse, dbi dll)
- SL (Klasifikasi – BackProp, Basic DL)
- USL (SOM)

# KONTRAK PERKULIAHAN

- **Kehadiran** : 75 % (Absensi melalui media Forum Diskusi KULINO)
- **Bobot Penilaian**
  - TUGAS : 40% (Akumulasi BOBOT Kehadiran, Keaktifan, Sikap dan Tugas)
  - UTS : 30%
  - UAS : 30%
- **Perkuliahan** : Synchronous (Kelas), Share Materi dan Penugasan (Kulino) dan Media Interaktif Grup Kelas (Telegram)
- **Tugas Akhir Makul**
  - Eksperimen Kasus menggunakan Python (Individu)



# Pengantar Machine Learning

Pertemuan 1

## **Machine Learning (ML)?**

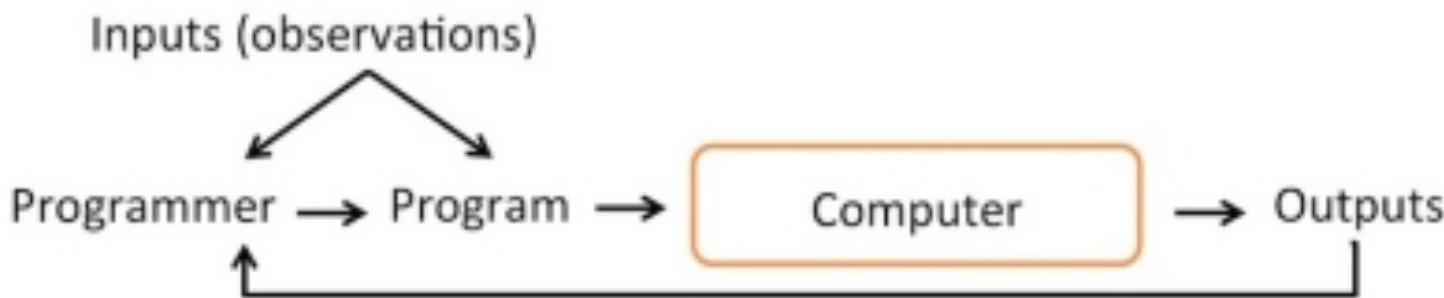
Machine = computer, computer program

Learning = improving performance on a given task, based on experience/examples

Instead of the programmer writing explicit rules for how to solve a given problem, the programmer instructs the computer how to learn from examples

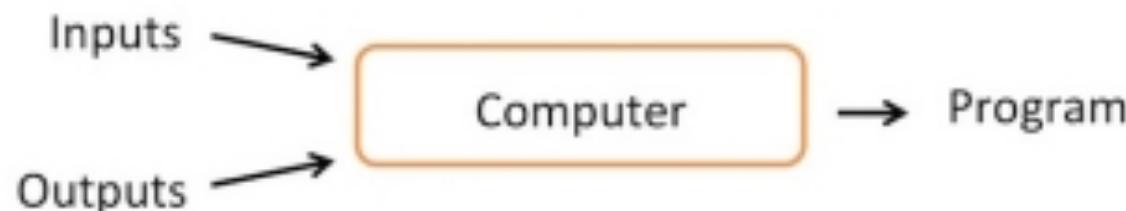
In many cases the computer program can even become better at the task than the programmer is!

## The Traditional Programming Paradigm



*Machine Learning is the field of study that gives computers  
the ability to learn without being explicitly programmed*  
– Arthur Samuel (1959)

## Machine Learning



## Machine Learning (ML)?

Suatu program komputer dikatakan belajar dari pengalaman  $E$  yang berhubungan dengan beberapa tugas  $T$  dan ukuran performansi  $P$  jika performansinya pada tugas-tugas  $T$ , sebagaimana diukur menggunakan  $P$ , meningkat dengan pengalaman  $E$ .

[Tom M. Mitchell, 1997]

# Era Machine Learning

- **Era Sebelum 1980:** hampir semua metode *learning* melakukan pembelajaran untuk menghasilkan *linear decision surfaces*. Metode-metode pembelajaran linier ini sudah memiliki pijakan teori yang kuat.
- **Era 1980-an:** *Decision trees* dan ANN menjadi pelopor dalam pembelajaran nonlinier. Namun, pijakan teorinya masih lemah. Kedua metode juga sering terjebak pada optimum lokal.
- **Era 1990 sampai sekarang:** telah dikembangkan metode-metode *learning* nonlinier yang efisien berbasis *computational learning theory*. Metode-metode pembelajaran nonlinier ini memiliki pijakan teori yang sudah mapan.

## Klasifikasi Metode Machine Learning

- Berdasarkan Dampak yang Diharapkan
- Diskrit atau Kontinu?
- *Offline* atau *Online*?
- Mudah atau Sulit Diinterpretasikan?
- Induktif atau Deduktif?

## Berdasarkan Dampak yang Diharapkan

- **Supervised Learning:** membangkitkan fungsi yang memetakan *input* ke *output*
- **Unsupervised Learning:** memodelkan *input* secara otomatis tanpa panduan
- **Semi-supervised Learning:** kombinasi Supervised dan Unsupervised
- **Reinforcement Learning:** mempelajari suatu kebijakan bagaimana melakukan aksi berdasarkan hasil pengamatan terhadap lingkungan
- **Transduction:** berlatih memprediksi *output* baru berdasarkan *training inputs*, *training outputs*, dan *test inputs* yang tersedia selama proses pelatihan (pembelajaran)
- **Learning to learn:** mempelajari bias induktifnya sendiri berdasarkan pengalaman

## Diskrit atau Kontinu?

- **Algoritma pembelajaran Diskrit:** menerima *input* diskrit dan menghasilkan model yang mengeluarkan *output* diskrit juga. Misal: *Decision Tree learning*.
- **Algoritma pembelajaran Kontinu:** menerima *input* yang kontinu dan menghasilkan model yang mengeluarkan *output* diskrit maupun kontinu. Misal: *Artificial Neural Networks (ANN)*, *support vector machine (SVM)*.

## Offline atau Online?

- *Offline learning* atau *batch learning* menggunakan sebagian atau semua data yang tersedia untuk proses pembelajaran atau sering disebut pembangunan model. Begitu model telah dibangun, Anda dapat melakukan prediksi pada data-data di masa depan. Model ini biasanya memberikan performansi yang bagus untuk data-data awal dan performansinya akan terus menurun seiring waktu untuk data-data yang akan datang.
- *Online learning* memperbarui model secara kontinu menggunakan sebuah data sampel. Anda tidak melakukan pembelajaran ulang (*retraining*) model seperti pada *offline learning*, tetapi Anda memperbarui model berdasarkan observasi-observasi baru secara terus menerus.

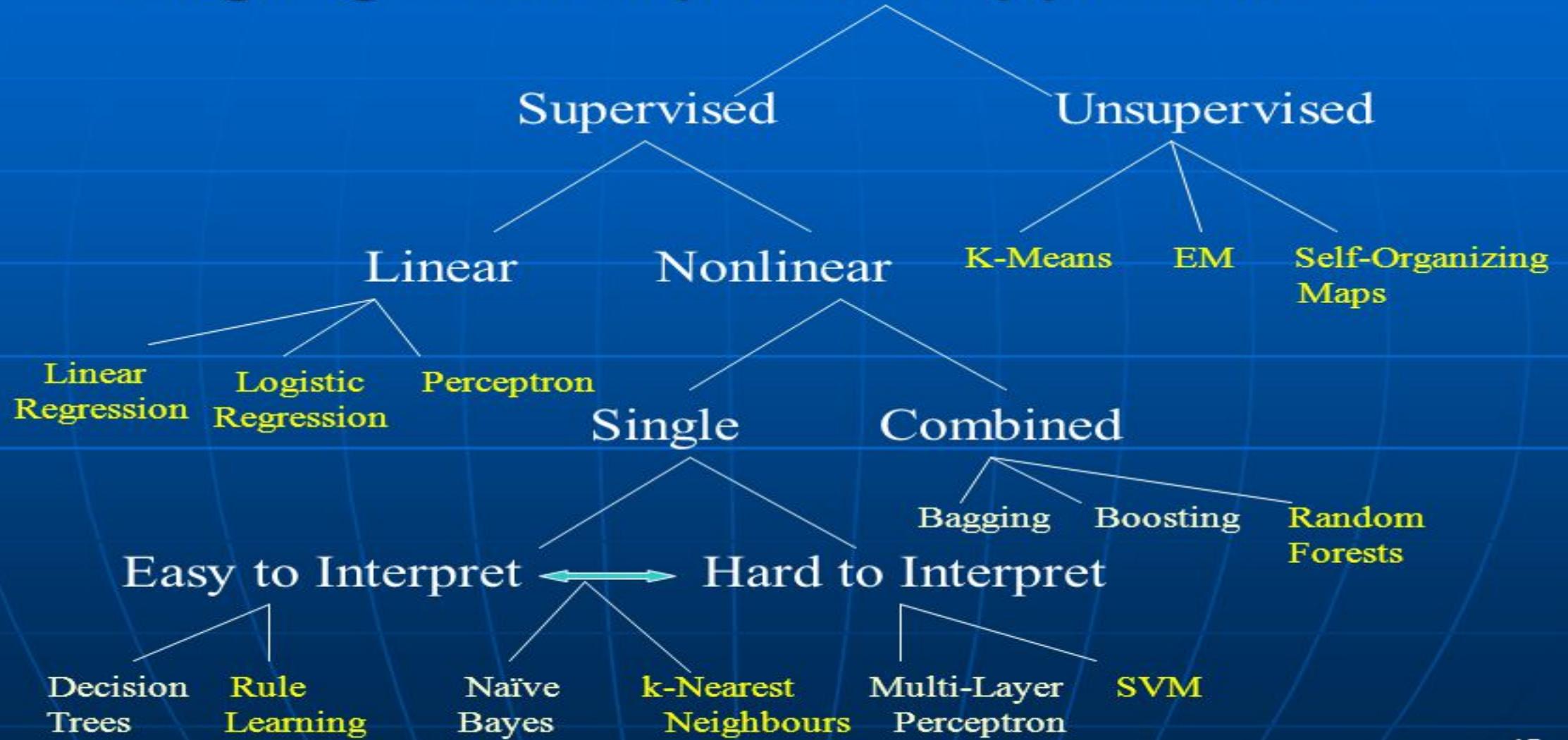
## Mudah atau Sulit Diinterpretasikan?

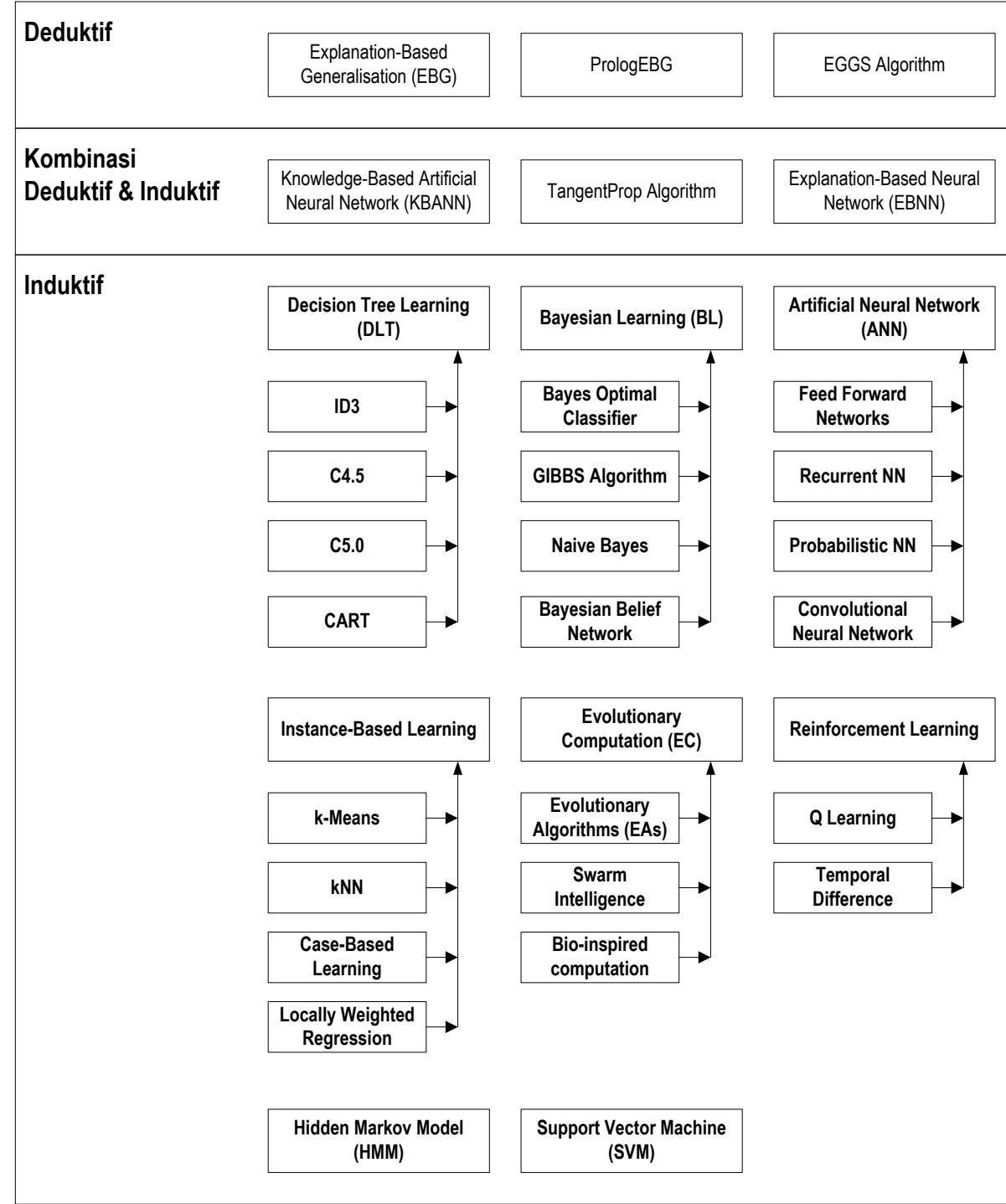
- **Mudah diinterpretasikan** artinya mudah dimengerti oleh pengguna. Misal, *Decision Tree Learning* menghasilkan aturan-aturan (*rules*) yang jelas dan sangat mudah dimengerti atau diinterpretasikan oleh pengguna
- **Sulit diinterpretasikan** artinya hasil pembelajaran sulit dimengerti pengguna Misal, *artificial neural networks* menghasilkan bobot-bobot sinaptik yang merepresentasikan kurva keputusan dan sulit dipahami pengguna. Demikian pula *support vector machine* (SVM) yang hanya menghasilkan *support vectors* yang merepresentasikan batas-batas keputusan dan sulit diinterpretasikan oleh pengguna.

## Induktif atau Deduktif?

- **Induktif (*inductive*)**: penalaran yang menghasilkan suatu kesimpulan berdasarkan observasi-observasi ganda (*multiple observations*)
- **Deduktif (*deductive*)**: penalaran yang menghasilkan suatu kesimpulan secara logis dari premis-premis tertentu

# A Taxonomy of Machine Learning Techniques: Highlight on Important Approaches





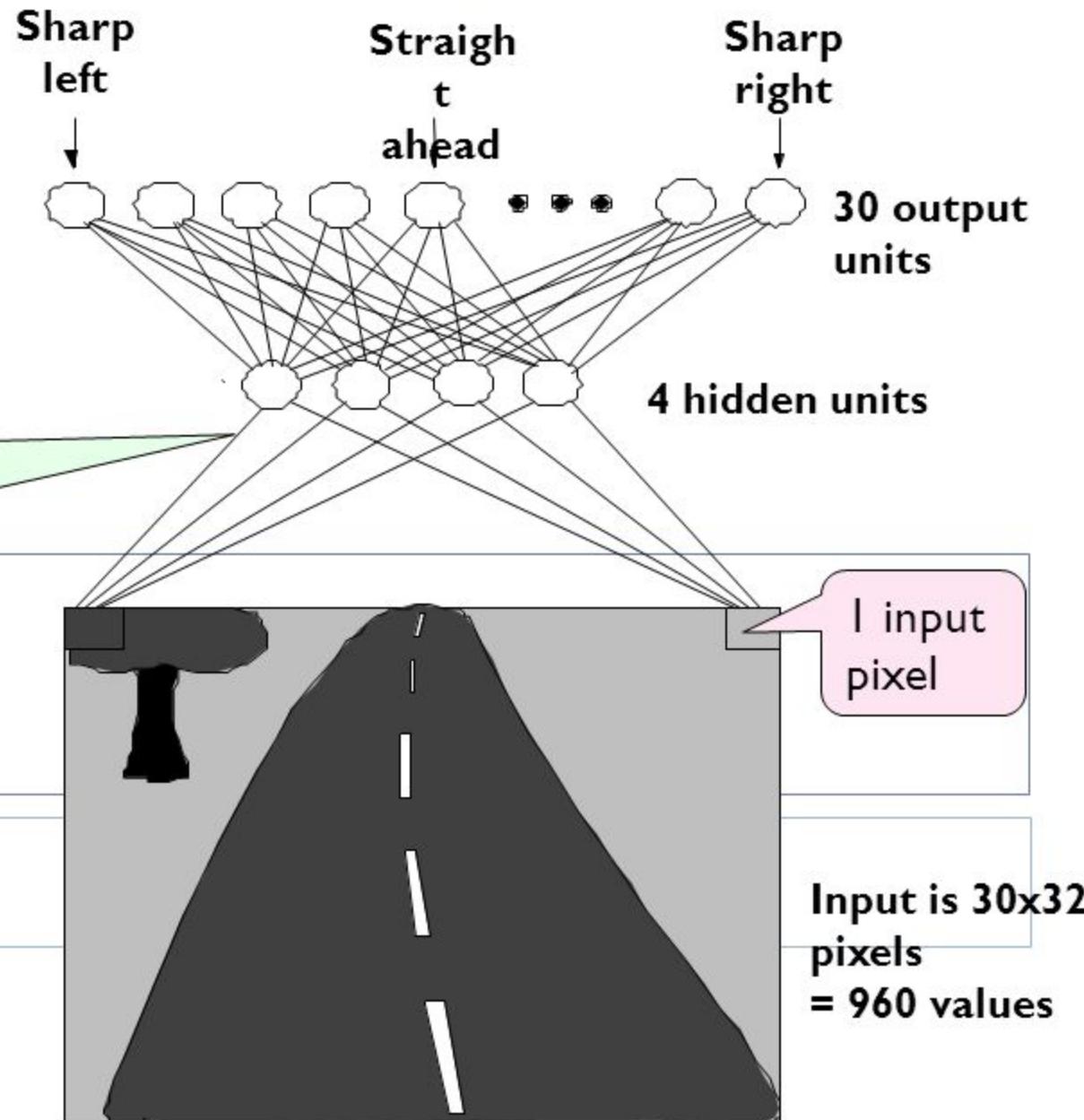
## Machine Learning untuk ICT

- **Prediksi *churn*:** prediksi berpindahnya pelanggan dari suatu operator telekomunikasi ke operator lain berdasarkan data histori;
- **Deteksi *hoax*:** memilah berita di internet sebagai *hoax* atau bukan;
- ***Spam filtering*:** email, pesan di media sosial;
- **Deteksi kecurangan (*fraud detection*):** internet, finansial, dsb;
- **Sistem perekомендasi (*recommender system*):** memberikan rekomendasi produk, buku, hotel, penerbangan, film, dsb.

# Machine Learning untuk Computer Vision

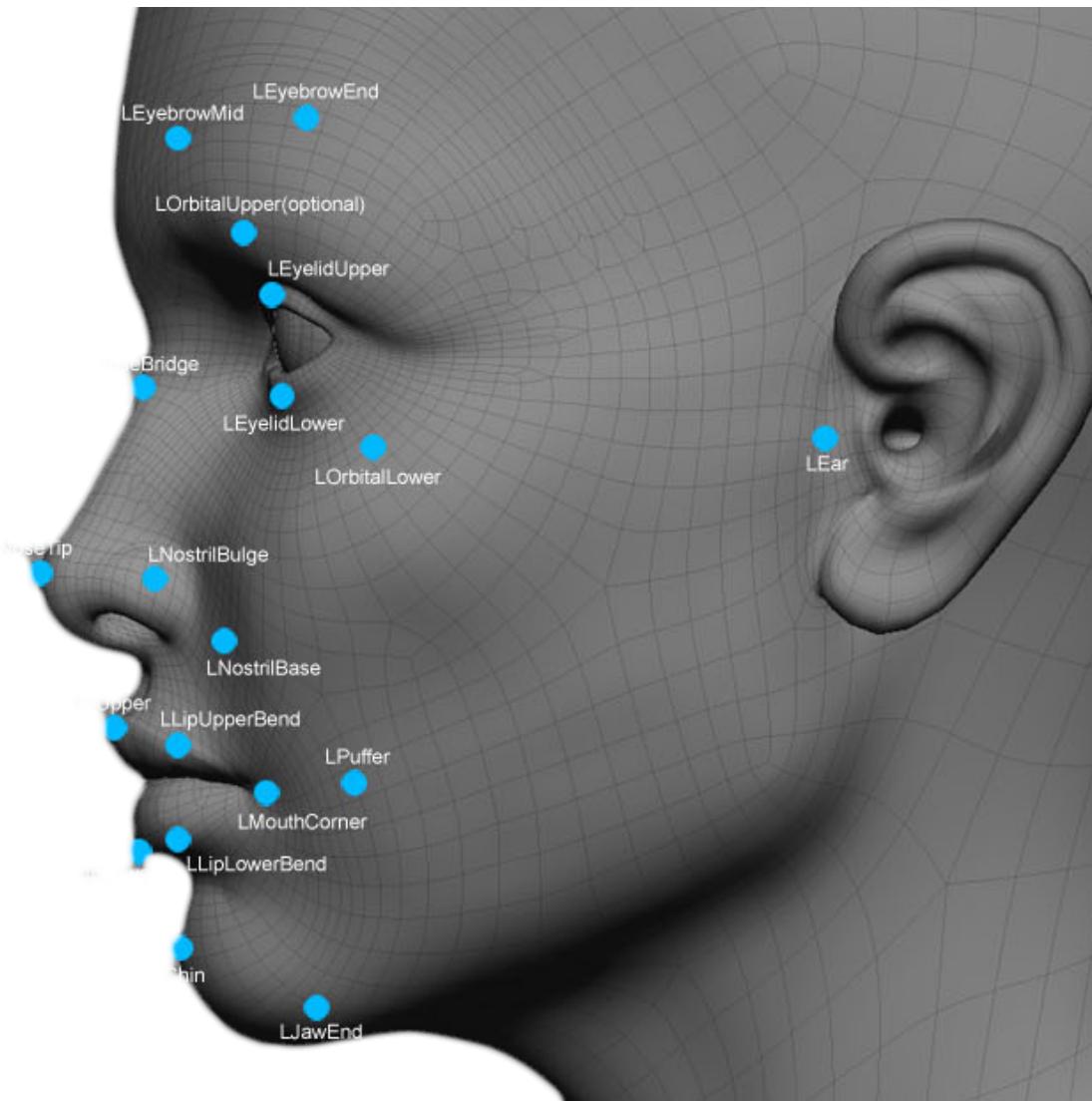
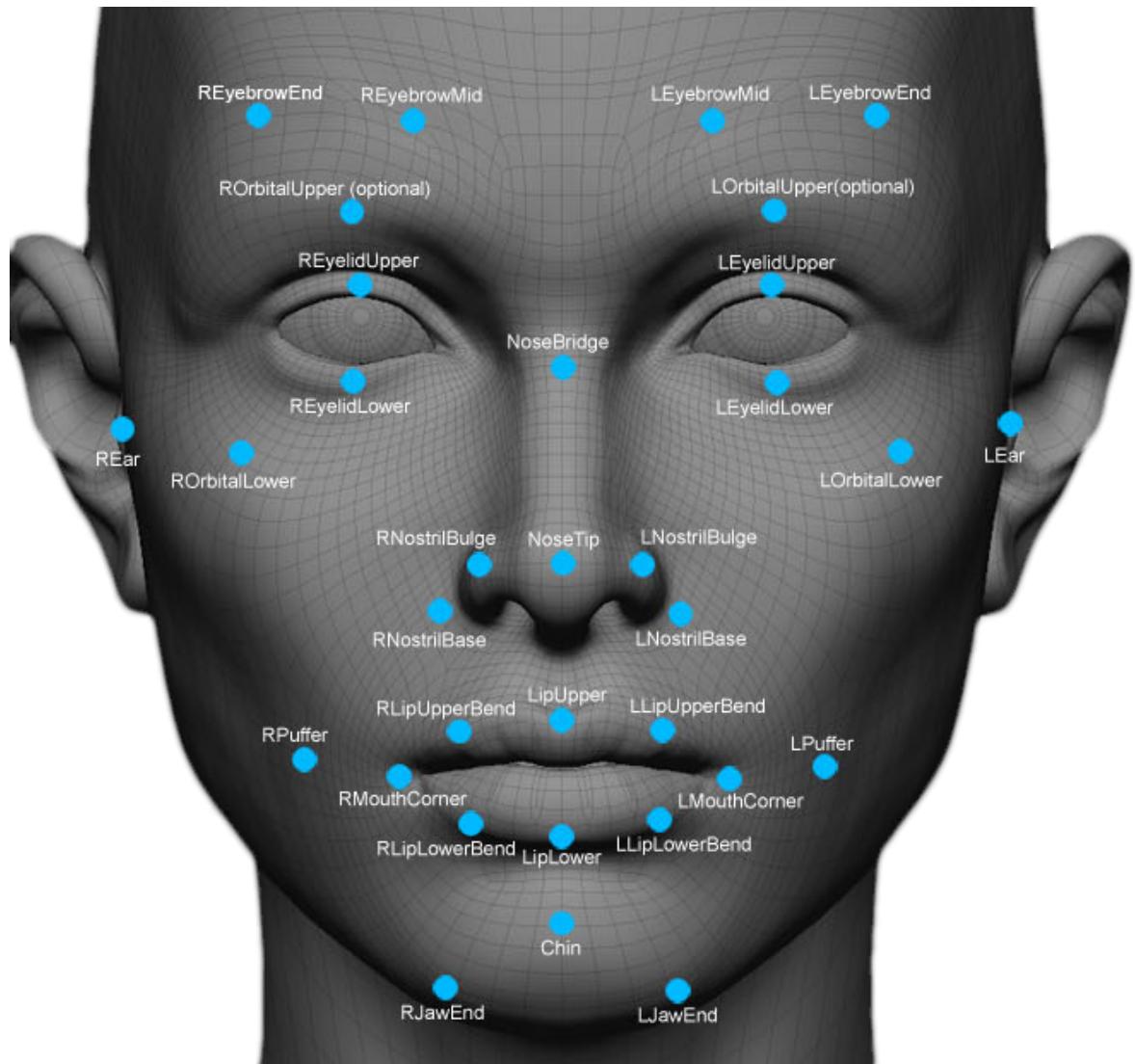
- **Autonomous car, driverless car, self-driving car, robotic car, atau unmanned ground vehicle:** sistem ini pertama kali dibuat pada tahun 1989 oleh Dean Pomerleau dari Carnegie Mellon University dengan produk bernama ALVINN (Autonomous Land Vehicle In a Neural Network). Saat ini banyak sekali autonomous car yang bermunculan di berbagai negara, seperti Jepang dan Finlandia, namun belum ada satupun yang diijinkan berjalan di jalanan umum yang dikendalikan penuh oleh mesin (bukan manusia). Siapa yang akan dihukum jika terjadi kecelakaan yang memakan korban? Mesin?

# Neural Net example: ALVINN



# **Machine Learning untuk Computer Vision**

- **Pengenalan wajah manusia** telah lama digunakan dalam sistem kemanaan negara, seperti yang Anda lihat di film-film layar lebar;
- **Verifikasi sidik jari, iris mata, suara atau ciri-ciri biometrik lainnya** yang banyak digunakan dalam sistem keamanan;
- **Optical Character Recognition (OCR)**: mengenali tulisan tangan atau tulisan cetak hasil pemindaian (scanning).
- **Lip Reading**: membaca gerak bibir seseorang sehingga dapat diketahui apa yang sedang diucapkannya. Anda dapat membuat teks dari video seseorang yang sedang pidato walaupun di dalam video tersebut tidak ada suara sama sekali. Sistem lip reading dapat memetakan gerak bibir menjadi deretan fonem, kata, kalimat, paragraf, hingga pidato lengkap.



# What is visible?

- Stressed vowels are mostly identifiable
- Unstressed vowels are not identifiable.
- Consonants articulated at or near the back of the mouth are invisible (e.g., k, g, h).
- Other consonants are only identifiable within groups (visemes) that have different places of articulation.

## Phonemically

	oo, u		p, b, m
	aw, o		f, v
	ah, a		th, <u>th</u>
	eh		s, z, t, d n, l, r
	ee, i		sh, <u>zh</u> <u>ch, dge</u>

# Machine Learning untuk Natural Lang. Proc. (NLP)

- *Automatic summarization*: peringkasan teks secara otomatis;
- *Foreign language reading aid*: alat bantu membaca teks bahasa asing;
- *Foreign language writing aid*: alat bantu menulis teks bahasa asing;
- *Information extraction*: mengekstrak informasi dari suatu teks;
- *Information Retrieval (IR)*: bagaimana menyimpan, mencari dan menemukan kembali informasi.
- *Machine Translation (MT)*: sistem yang secara otomatis menerjemahkan suatu bahasa ke bahasa yang lain.

# Machine Learning untuk Natural Lang. Proc. (NLP)

- ***Named Entity Recognition (NER)***: jika diberikan suatu teks, bagaimana menentukan kata-kata mana saja yang merupakan nama tempat atau nama orang sehingga penulisan kata-kata tersebut dapat dikoreksi (dimana huruf pertamanya seharusnya menggunakan huruf besar).
- ***Text Proofing***: pengecekan suatu teks sehingga diketahui kesalahan ketik, kesalahan tata bahasa, dan sebagainya. Misalnya: <http://www.grammarly.com>
- ***Text-to-Speech***: mengubah teks menjadi suara, sehingga komputer bisa membacakan pesan WA, pemanggilan nama antrian, dan banyak lagi aplikasi yang bisa dibangun.

# Machine Learning untuk Natural Lang. Proc. (NLP)

- ***Automatic Speech Recognition (ASR)***: mengenali ucapan manusia, sehingga komputer bisa menuliskan teks yang didiktekan manusia kepadanya. ASR bisa menangani *call center*, reservasi tiket melalui telepon, dan sebagainya. Saat ini ASR mulai dikembangkan ke *Audiovisual ASR* yang menerima masukan berupa video. Jika melihat Youtube, Anda bisa mendapatkan teks *subtitle* dalam sebuah video yang dibangkitkan secara otomatis melalui sistem *Audiovisual ASR*.
- ***Natural Language Understanding***: memahami teks bahasa alami, sehingga komputer bisa mengerti isi teks tersebut.

# Machine Learning untuk Bioinformatics

- **Diagnosis berbagai macam penyakit**, seperti tumor, kanker, diabetes, menggunakan teknik klasifikasi berdasarkan masukan yang berupa hasil pemeriksaan darah, rontgen, USG, dan tes laboratorium lainnya.
- **Verifikasi *deoxyribonucleic acid (DNA)***, menguji apakah benar Si A adalah putra kandung Si B berdasarkan struktur DNA kedua orang tersebut.

## Machine Learning untuk Robotic

- *Robot learning* merupakan gabungan dari *machine learning* dan *robotics*. Istilah *robot learning* mengacu pada pembelajaran untuk menjalankan tugas-tugas seperti menghindari halangan (*obstacle avoidance*), kontrol dan tugas-tugas lain yang berhubungan dengan gerakan. *Robot learning* sangat erat kaitannya dengan kontrol adaptif dan *reinforcement learning*.

## Machine Learning untuk Robotic

- *Humanoid* adalah robot yang strukturnya mirip manusia, berjalan menggunakan dua kaki, bisa melihat, berbicara, meraih, dan sebagainya. Satu contoh proyek *humanoid* adalah yang dikerjakan oleh para peneliti di Chalmers University of Technology, Swedia pada tahun 1999. Mereka membangun tiga *humanoid* yang diberi nama Elvis, Priscilla, dan Elvina (Wolff and Nordin, 2001). Contoh *humanoid* robot lainnya adalah Asimo (singkatan dari *Advanced Step in Innovative Mobility*) yang dibangun oleh sebuah perusahaan besar Jepang, Honda. Asimo adalah robot cerdas dan responsif yang mampu mengenali suara, wajah, dan gestur manusia serta mampu mendeteksi jarak dan arah objek bergerak (Honda, 2017).

# Machine Learning untuk Games

- **Neurogammon 1.0** adalah program komputer permainan *backgammon* berbasis *multilayer neural networks*. Neurogammon dilatih menggunakan algoritma pembelajaran *Back Propagation* dengan sekumpulan data dari ahli *backgammon*. Pada olimpiade komputer pertama di London, Neurogammon memenangkan kompetisi permainan *backgammon* dengan catatan sempurna: lima kali menang, tanpa kalah. Neurogammon diklaim sebagai program komputer pertama yang sanggup belajar dan memenangkan kompetisi (Tesauro, 1990).

## Machine Learning untuk Games

- **NeuroGo (Enzenberger, 2004)** adalah contoh program komputer untuk permainan Go yang menggunakan *neural network*. Go adalah permainan strategi, mirip catur, untuk mendapatkan wilayah teritorial. Pada permainan profesional, ukuran papannya sebesar 19 x 19 untuk dua pemain: Hitam dan Putih. Permainan ini sudah dimainkan sejak 3000-an tahun yang lalu di Cina Kuno (Jeffrey, 2007). NeuroGo dan WinHonte dilatih menggunakan basisdata sangat besar dari permainan-permainan profesional atau memainkan banyak permainan melawan dirinya sendiri atau melawan manusia atau program lain untuk meningkatkan performansinya.

## **Referensi**

1. Ethem Alpaydin : Machine Learning, Second Edition, The MIT Press, 2010
2. Jiawei Han, Micheline Kamber, Jian Pei, Data mining : concepts and techniques – 3<sup>rd</sup> ed, Elsevier, 2012.
3. Ian H. Witten, Frank Eibe, Mark A. Hall, Data mining: Practical Machine Learning Tools and Techniques 4th Edition, *Elsevier*, 2017.



# THANKS

ANY QUESTIONS?