INTRODUCTION TO Machine Learning

Why "Learn" ?

- Machine learning is programming computers to optimize a performance criterion using example data or past experience.
- There is no need to "learn" to calculate payroll
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

What We Talk About When We Talk About"Learning"

- Learning general models from a data of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:

  People who bought "Da Vinci Code" also bought "The Five People You Meet in Heaven" (www.amazon.com)

- Build a model that is *a good and useful approximation* to the data.

Data Mining

- Retail: Market basket analysis, Customer relationship management (CRM)
- Finance: Credit scoring, fraud detection
- Manufacturing: Optimization, troubleshooting
- Medicine: Medical diagnosis
- Telecommunications: Quality of service optimization
- Bioinformatics: Motifs, alignment
- Web mining: Search engines
- …

What is Machine Learning?

- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms to
  - Solve the optimization problem
  - Representing and evaluating the model for inference

Applications

- Association
- Supervised Learning
  - Classification
  - Regression
- Unsupervised Learning
- Reinforcement Learning

Learning Associations

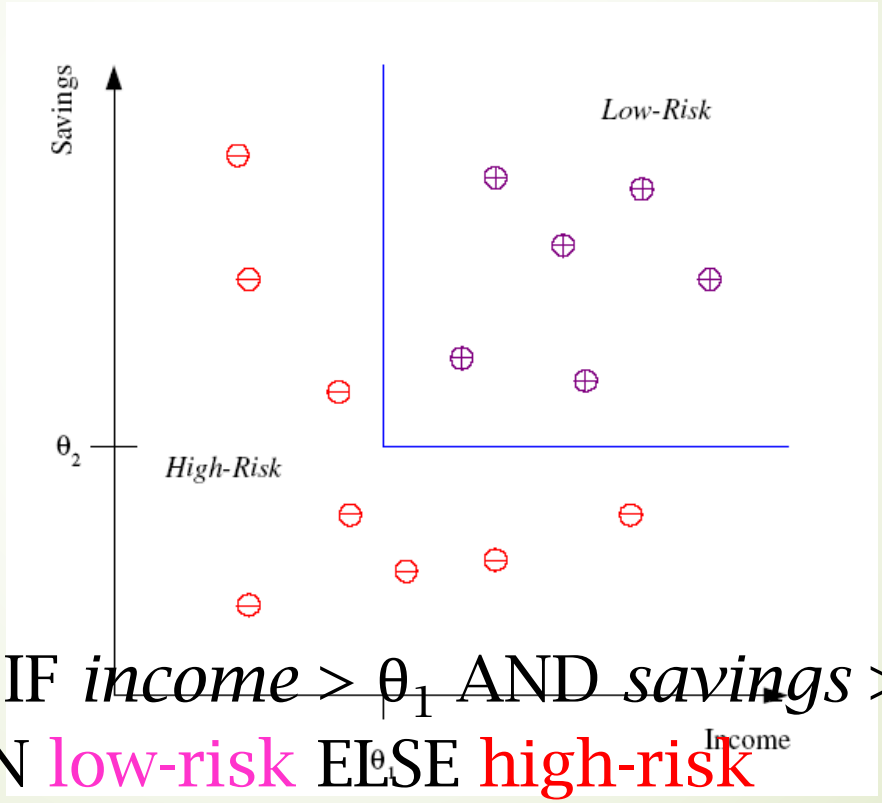
- Basket analysis:

  $P(Y \mid X)$ probability that somebody who buys X also buys Y where X and Y are products/services.

  Example: $P(\text{chips} \mid \text{beer}) = 0.7$

Classification

- Example: Credit scoring
- Differentiating between low-risk and high-risk customers from their *income* and *savings*



Discriminant: IF *income* $> \theta_1$ AND *savings* $> \theta_2$
THEN low-risk ELSE high-risk

Classification: Applications

- Pattern recognition
- Face recognition: Pose, lighting, occlusion (glasses, beard), make-up, hair style
- Character recognition: Different handwriting styles.
- Speech recognition: Temporal dependency.
  - Use of a dictionary or the syntax of the language.
  - Sensor fusion: Combine multiple modalities; eg, visual (lip image) and acoustic for speech
- Medical diagnosis: From symptoms to illnesses
- ...

Face Recognition

Training examples of a person

Test images

AT&T Laboratories, Cambridge UK
http://www.uk.research.att.com/facedatabase.html

Regression

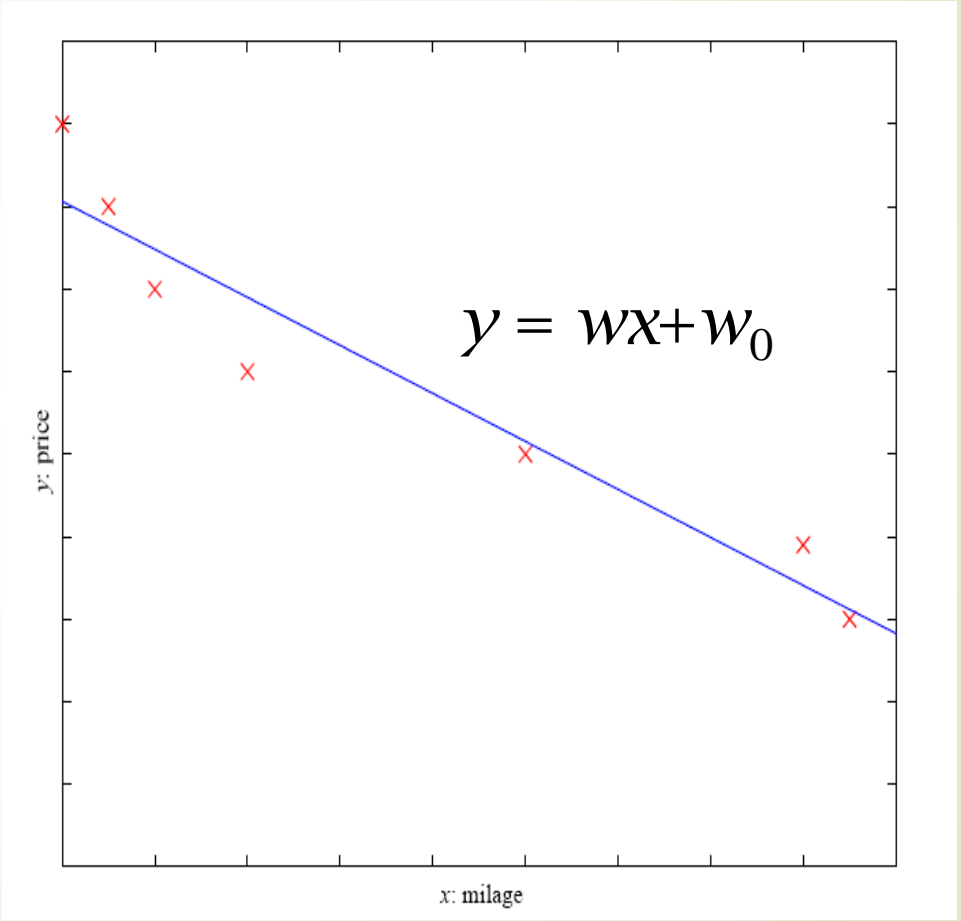
- Example: Price of a used car
- x : car attributes

  y : price

  $y = g(x \mid \theta)$

  $g(\,)$ model,

  θ parameters



Supervised Learning: Uses

- **Prediksi kasus di masa depan**: Gunakan aturan untuk memprediksi output dari input di masa depan.
- **Ekstraksi pengetahuan**: Aturan tersebut mudah dipahami.
- **Kompresi**: Aturan tersebut lebih sederhana dibandingkan data yang dijelaskannya.
- **Deteksi outlier**: Pengecualian yang tidak tercakup oleh aturan, misalnya, penipuan

Unsupervised Learning

- Learning "what normally happens"
- No output
- Clustering: Grouping similar instances
- Example applications
  - Customer segmentation in CRM
  - Image compression: Color quantization
  - Bioinformatics: Learning motifs

Reinforcement Learning

- Learning a policy: A sequence of outputs
- No supervised output but delayed reward
- Credit assignment problem
- Game playing
- Robot in a maze
- Multiple agents, partial observability, ...

What is Learning?

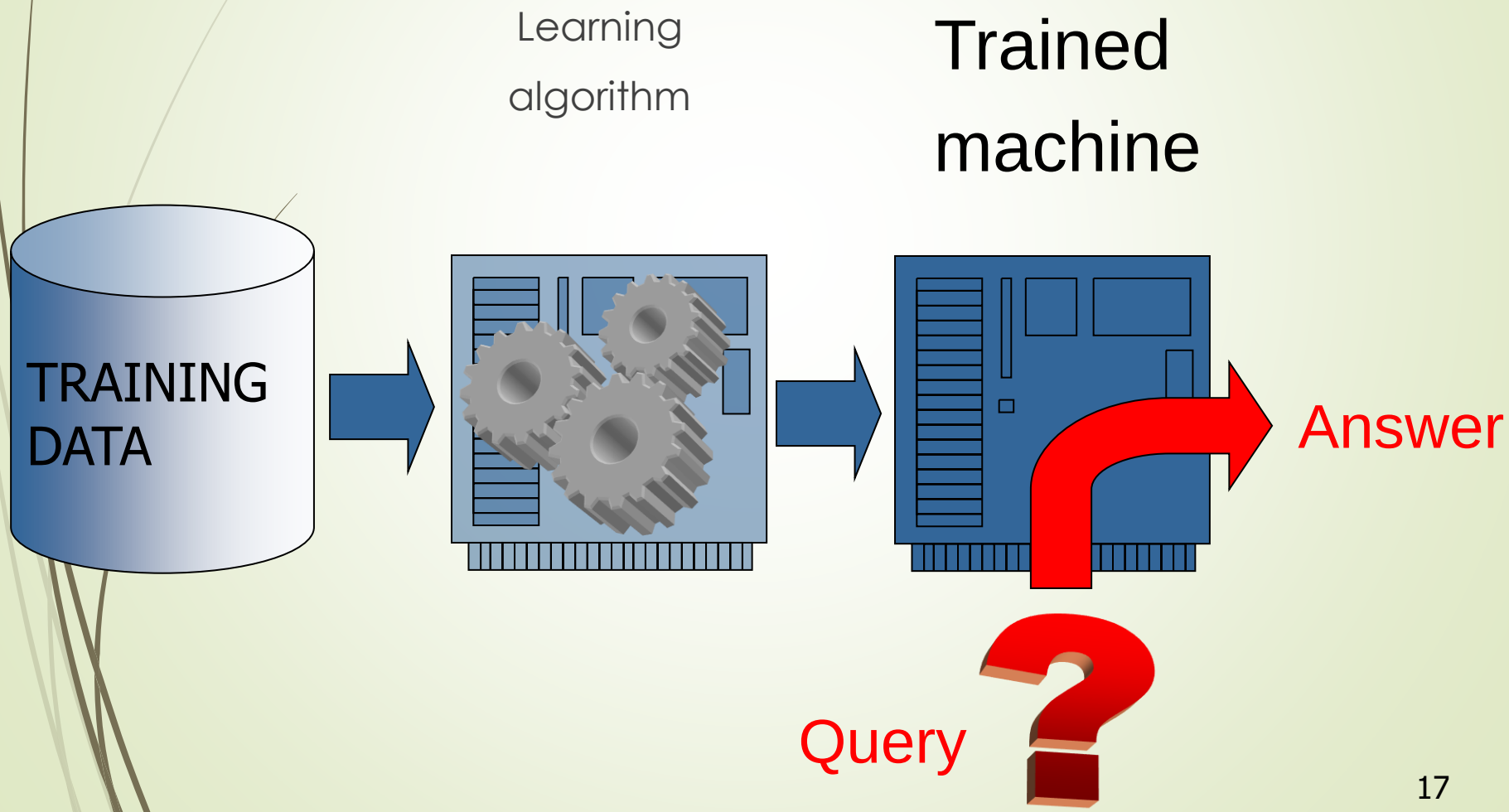
"To gain **knowledge** or **understanding** of, or **skill** in **by study**, **instruction** or **experience**"

- Learning a set of new facts.
- Learning HOW to do something .
- Improving ability of something already learned.

What is Machine Learning?

- Machine Learning is the study of methods for programming computers to learn.
- Building machines that automatically learn from experience.
- Machine learning usually refers to the changes in systems that perform tasks associated with artificial intelligence AI Such tasks involve recognition, diagnosis, planning, robot control, prediction, etc.

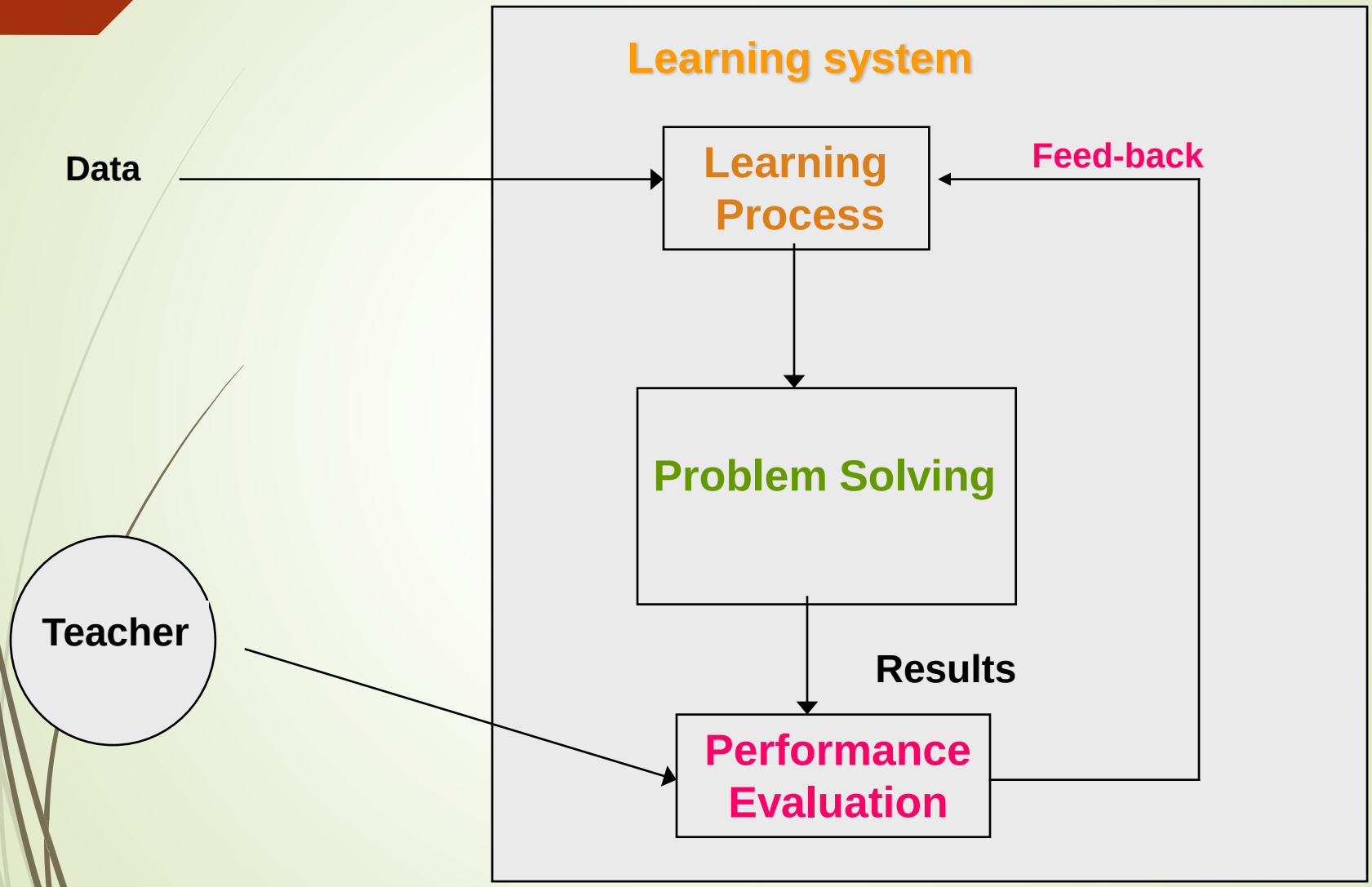
What is Machine Learning?



Steps in machine learning

1) Data collection.
2) Representation.
3) Modeling.
4) Estimation.
5) Validation.
6) Apply learned model to new “test” data

General structure of a learning system



Advantages of ML

1) Solving vision problems through statistical inference.

2) Intelligence from the common sense AI.

3) Reducing the constraints over time achieving complete autonomy.

Disadvantages of ML

1) Application specific algorithms.

2) Real world problems have too many variables and sensors might be too noisy.

3) Computational complexity.

Types of machine Learning

1) Unsupervised Learning .

2) Semi-Supervised (reinforcement).

3) Supervised Learning.

Unsupervised Learning

- Studies how input patterns can be represented to reflect the statistical structure of the overall collection of input patterns
- No outputs are used (unlike supervised learning and reinforcement learning)
- Learner is provided only unlabeled data.
- No feedback is provided from the environment

Unsupervised Learning

- **Advantage**
- Most of the laws of science were developed through unsupervised learning.

- **Disadvantage**
- The identification of the features itself is a complex problem in many situations.

24

Semi-Supervised (reinforcement)

- it is in between Supervised and Unsupervised learning techniques the amount of labeled and unlabelled data required for training.
- With the goal of reducing the amount of supervision required compared to supervised learning.
- At the same time improving the results of unsupervised clustering to the expectations of the user.

Semi-Supervised (reinforcement)

- Semi-supervised learning is an area of increasing importance in Machine Learning.
- Automatic methods of collecting data make it more important than ever to develop methods to make use of unlabeled data.

Supervised Learning

1) Analogical Learning.

2) Learning by Decision Tree.

27

Analogical Learning

instances of a problem and the learner has to form a concept that supports most of the positive and no negative instances. This demonstrates that a number of training instances are required to form a concept in inductive learning. Unlike this, analogical learning can be accomplished from a single example. For instance, given the following training instance, one has to determine the plural form of bacilus.

28

The main steps in analogical learning are now formalized below.

1. Identifying Analogy: Identify the similarity between an experienced problem instance and a new problem.

2. Determining the Mapping Function: Relevant parts of the experienced problem are selected and the mapping is determined.

3. Apply Mapping Function: Apply the mapping function to transform the new problem from the given domain to the target domain.

The main steps in analogical learning are now formalized below.

4. Validation: The newly constructed solution is validated for its applicability through its trial processes like theorem or

   simulation .

5. Learning: If the validation is found to work well, the new

   knowledge is encoded and saved for future usage.

Learning by Decision Tree

A decision tree receives a set of attributes (or properties) of the objects as

inputs and yields a binary decision of true or false values as output. Decision

trees, thus, generally represent Boolean functions. Besides a range of {0,1}

other non-binary ranges of outputs are also allowed. However, for the sake of

simplicity, we presume the restriction to Boolean outputs. Each node in a

decision tree represents 'a test of some attribute of the instance, and each

branch descending from that node corresponds to one of the possible values

for this attribute'

31

Learning by Decision Tree

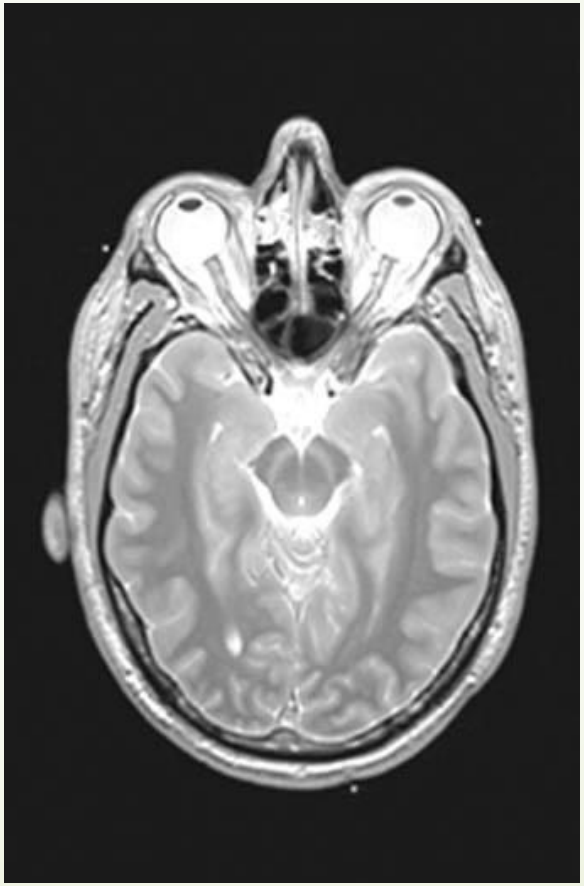
To illustrate the contribution of a decision tree, we consider a set of instances, some of which result in a true value for the decision. Those instances are called **positive instances. On the other hand, when the resulting decision** is false, we call the instance '**a negative instance'. We now consider** the learning problem of a bird's flying. Suppose a child sees different instances of birds as tabulated below.

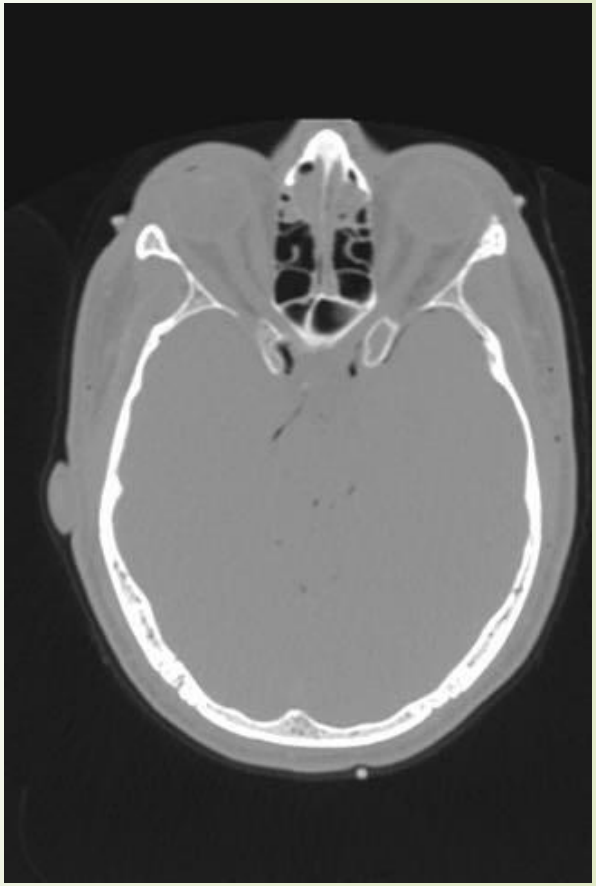
32

Applications of Machine Learning

Drug discovery

Medical diagnosis





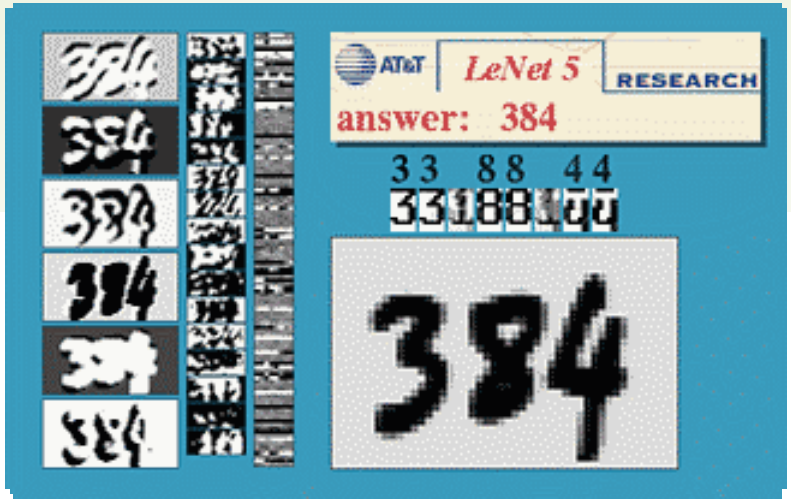
Photo

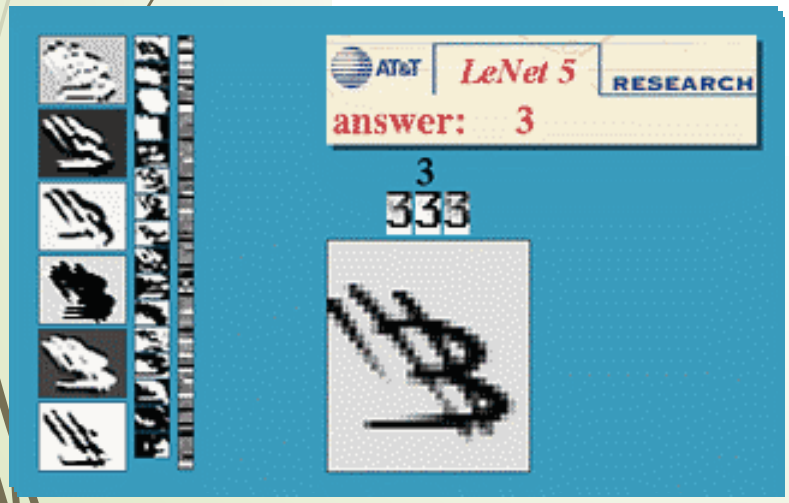
MRI

CT

Iris verification

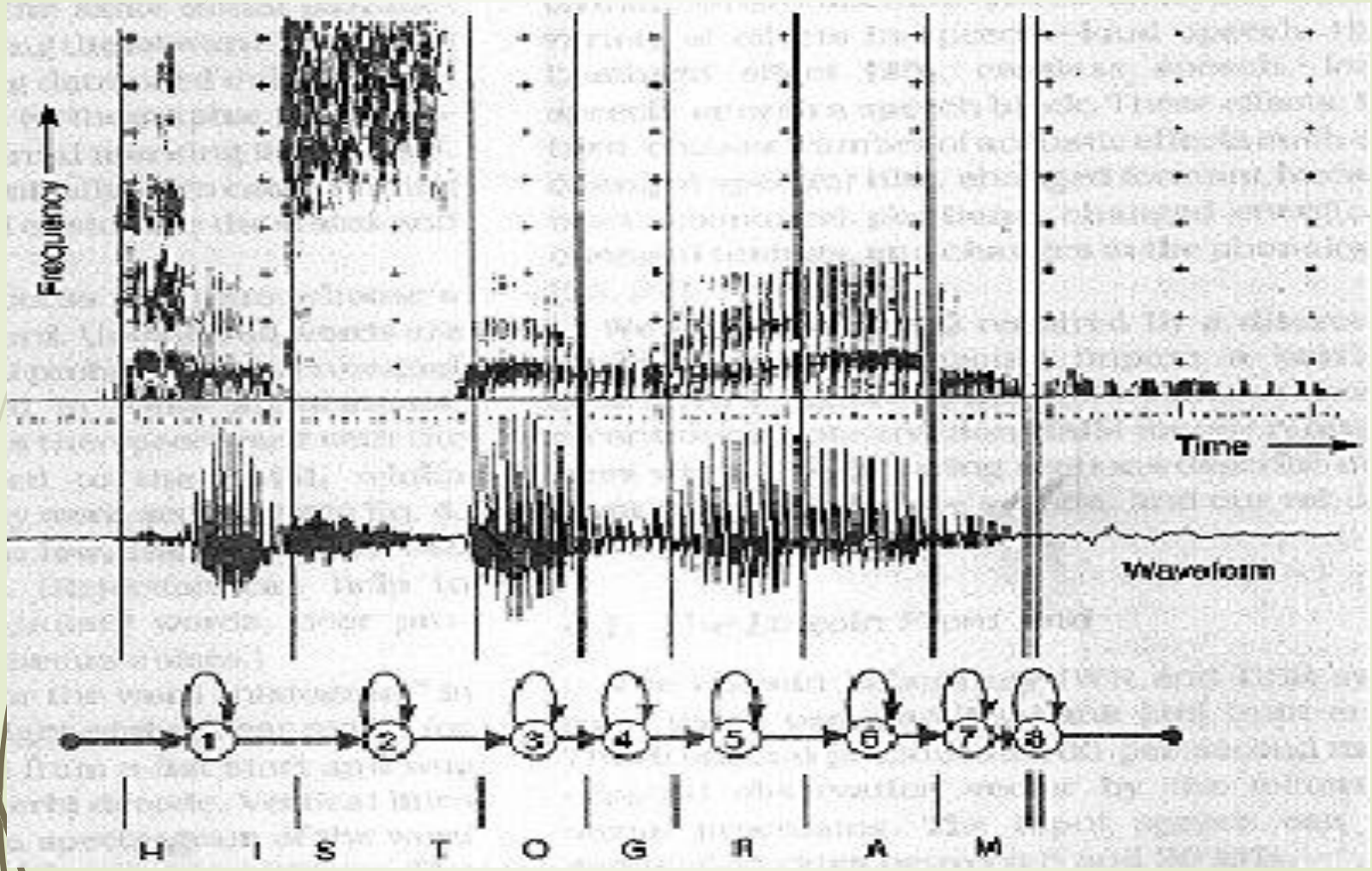
Hand-written digits



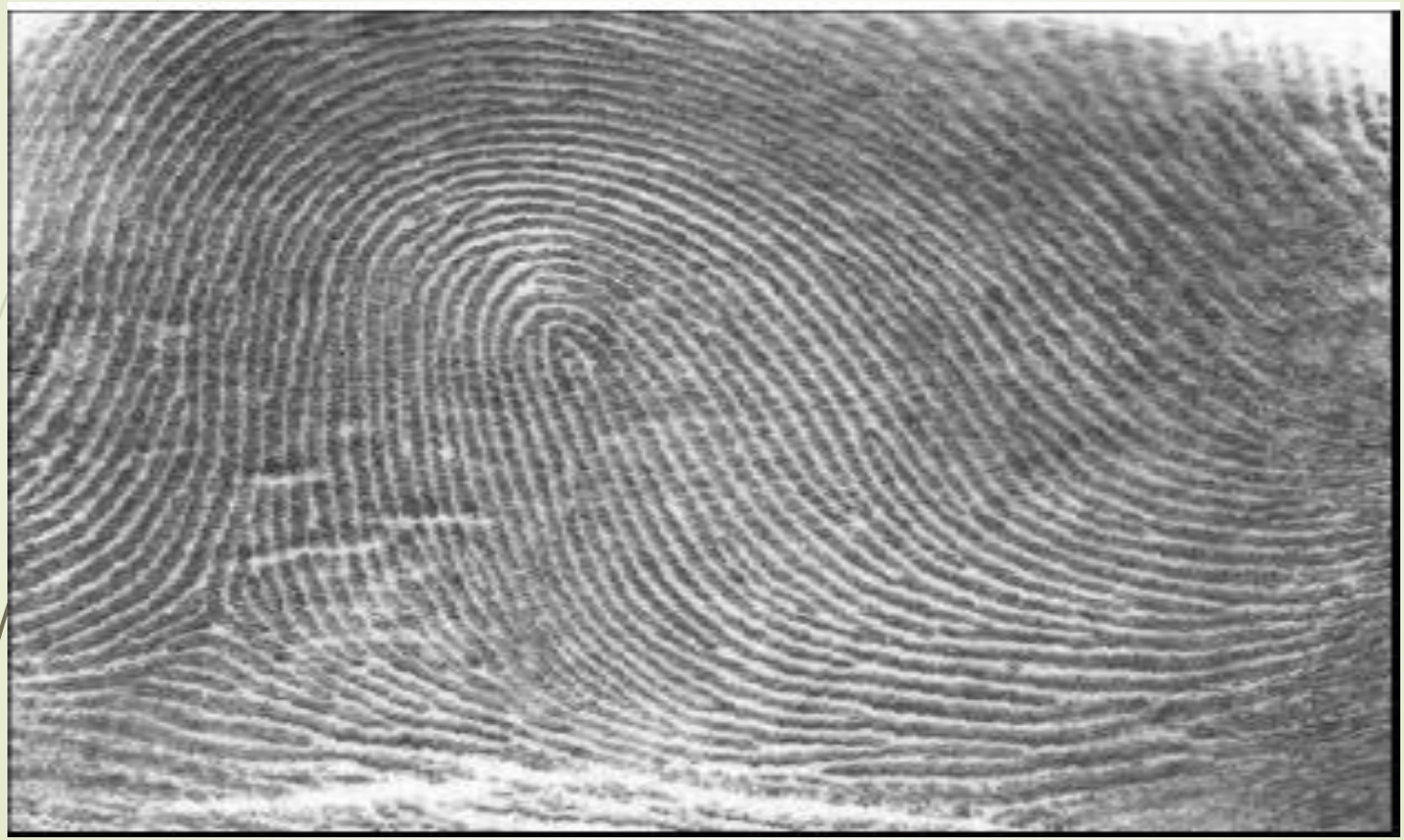


Radar Imaging

Speech Recognition

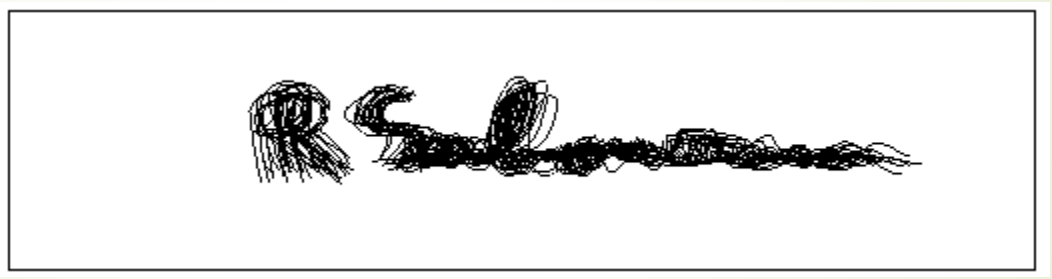


Finger print



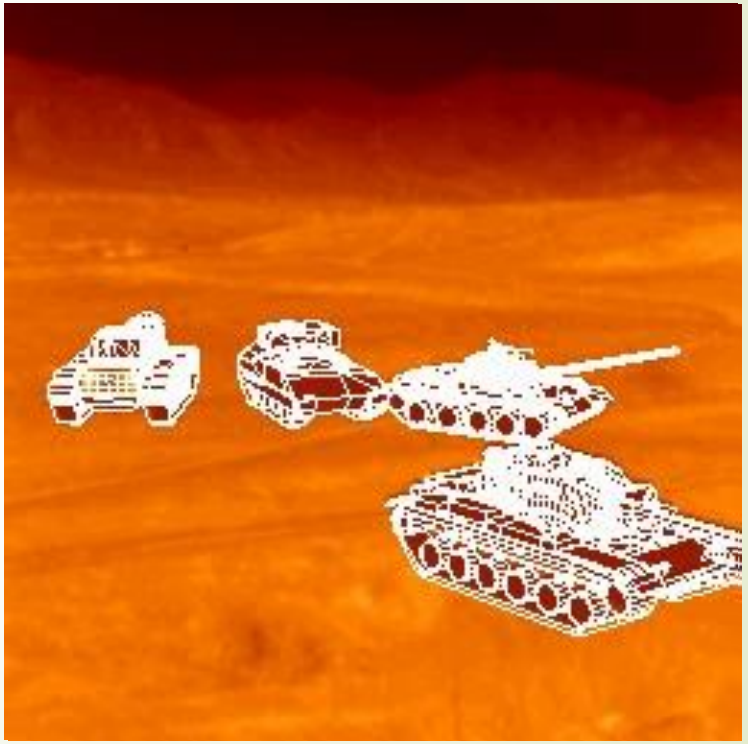
fingerprint image

Signature Verification

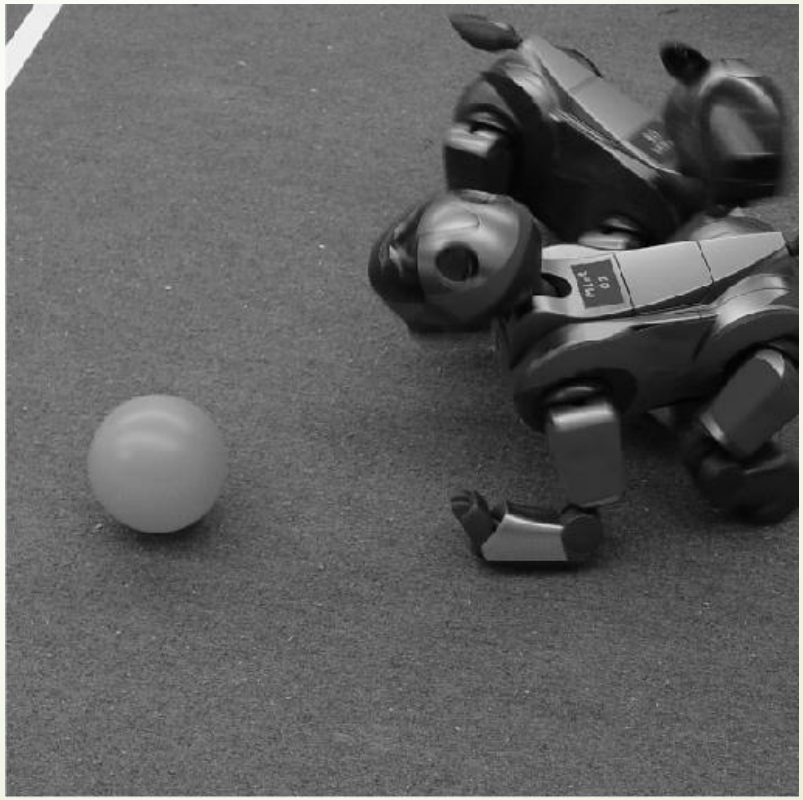


Face Recognition

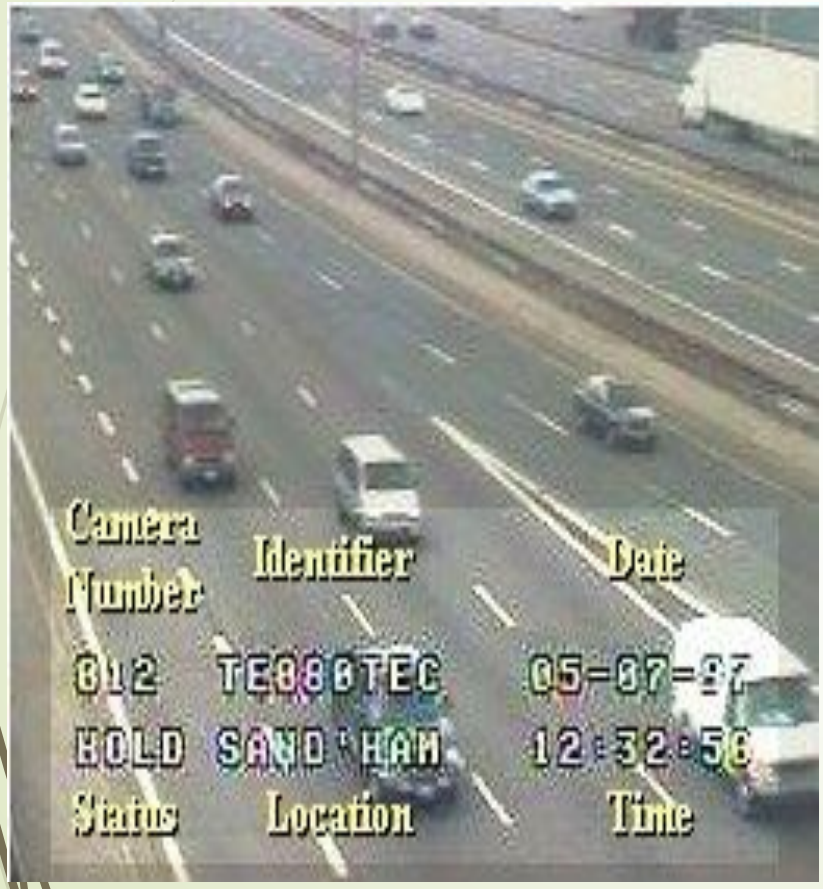
Target Recognition

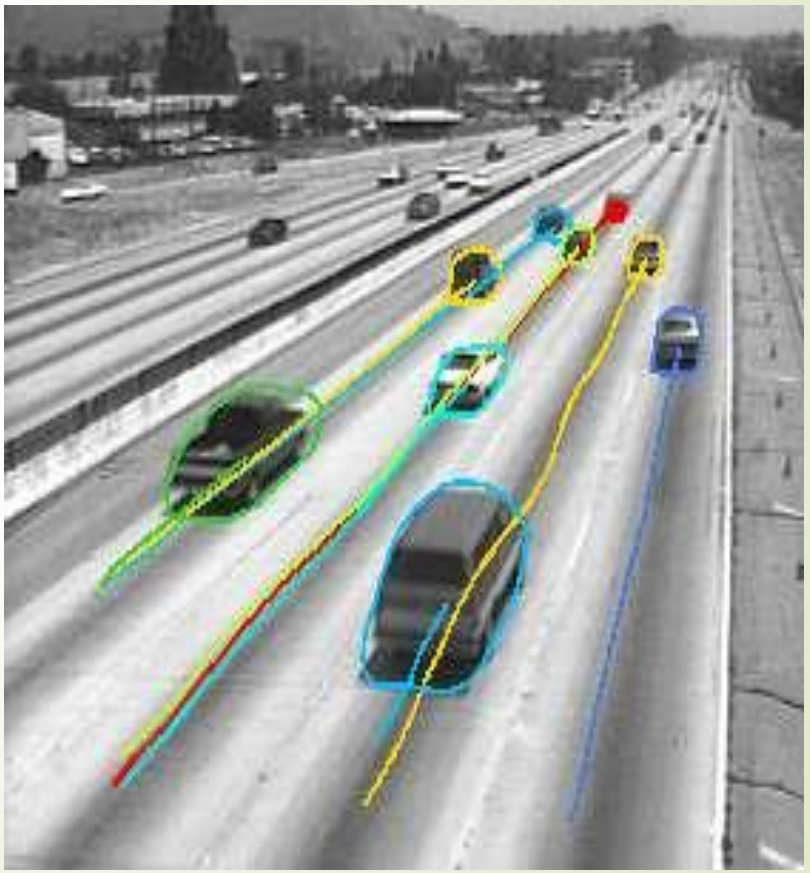


Robotics vision



Traffic Monitoring





Resources: Datasets

- UCI Repository: http://www.ics.uci.edu/~mlearn/MLRepository.html
- UCI KDD Archive: http://kdd.ics.uci.edu/summary.data.application.html
- Statlib: http://lib.stat.cmu.edu/
- Delve: http://www.cs.utoronto.ca/~delve/

Resources: Journals

- Journal of Machine Learning Research www.jmlr.org
- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association
- ...

Resources: Conferences

- International Conference on Machine Learning (ICML)
  - ICML05: http://icml.ais.fraunhofer.de/
- European Conference on Machine Learning (ECML)
  - ECML05: http://ecmlpkdd05.liacc.up.pt/
- Neural Information Processing Systems (NIPS)
  - NIPS05: http://nips.cc/
- Uncertainty in Artificial Intelligence (UAI)
  - UAI05: http://www.cs.toronto.edu/uai2005/
- Computational Learning Theory (COLT)
  - COLT05: http://learningtheory.org/colt2005/
- International Joint Conference on Artificial Intelligence (IJCAI)
  - IJCAI05: http://ijcai05.csd.abdn.ac.uk/
- International Conference on Neural Networks (Europe)
  - ICANN05: http://www.ibspan.waw.pl/ICANN-2005/
- ...