Machine Learning
CSL7620

Dr. Harshal D. Akolekar

Dept. Of Mechanical Engineering

School of Artificial Intelligence & Data Science

IIT Jodhpur

Slides are prepared from numerous web sources and books

Selected Journal Publications

- Pacciani, R., Marconcini, M., Bertini, F., Taddei, S.R., Spano, E., Zhao, Y., **Akolekar, H.D.,** Sandberg, R.D., and Arnone, A., 2021, `Assessment of Machine-learnt Turbulence Models Trained for Improved Wake-mixing in Low Pressure Turbine Flows', Energies, 14(24), 8327
- **Akolekar, H.D.**, Zhao, Y., Sandberg, R.D., Pacciani, R., "Integration of Machine Learning and Computational Fluid Dynamics to Develop Turbulence Models for Improved Turbine Wake Mixing Prediction", ASME **Journal of Turbomachinery**).
- **Akolekar, H.D**., Waschkowski, F., Zhao, Y., Pacciani, R., and Sandberg, R.D., 2021, `Transition Modeling for Low Pressure Turbines Using Computational Fluid Dynamics Driven Machine Learning', Energies (MDPI), 14(15), 4680.
- Zhao, Y., **Akolekar, H. D.**, Weatheritt, J., Michelassi, V., and Sandberg, R. D., 2020. 'RANS Turbulence Model Development using CFD-Driven Machine Learning'. Elsevier **Journal of Computational Physics**, 411.
- **Akolekar, H. D.**, Sandberg, R. D., Hutchins, N., Michelassi, V., and Laskowski, G., 2019. 'Machine-Learnt Turbulence Closures for Low Pressure Turbines with Unsteady Inflow Conditions', ASME **Journal of Turbomachinery**, ISUAAAT15 Special Issue, 141 (10) p. 101009
- **Akolekar, H. D.**, Weatheritt, J., Hutchins, N., Sandberg, R. D., Laskowski, G., and Michelassi, V., 2019. 'Development and Use of Machine-Learnt Algebraic Reynolds Stress Models for Enhanced Prediction of Wake Mixing in Low Pressure Turbines', ASME **Journal of Turbomachinery**, 141 (4) p. 041010.

Selected International Conferences

- Fang Y, Zhao Y, **Akolekar HD**, Ooi, A, Sandberg Richard, Pacciani, R, Marconcini, M; A Data Driven Approach for Generalising the Laminar Kinetic Energy Model for Separation and Bypass Transition in Low- and High-Pressure Turbines, 67^{th} **ASME Turbo Expo**: Power for Land, Sea and Air, Boston, USA, June 2023 (Best Poster Presentation)
- **Akolekar HD**, Waschkowski F, Pacciani R, Zhao Y, Sandberg, RD; Multi-Objective Development of Machine Learnt Closures for Fully Integrated Transition and Wake Mixing Predictions in Low Pressure Turbines, 66^{th} **ASME Turbo Expo**: Power for Land, Sea and Air, Rotterdam, Netherlands, July 2022.
- **Akolekar, H.D.**, Zhao, Y., Sandberg, R.D., Pacciani, R., "Integration of Machine Learning and Computational Fluid Dynamics to Develop Turbulence Models for Improved Turbine Wake Mixing Prediction", **65^{th} ASME Turbo Expo Turbomach. Tech. Conf. Expo.**, June, 2020, **London, UK** (paper no. GT2020-14732) (Virtual Event: Sept. 2020).
- Zhao, Y., **Akolekar, H.D.**, Sandberg, R.D., "CFD-Ready Turbulence Models from Gene Expression Programming: Concepts", In Bulletin, **72nd DFD Meeting of American Physical Society**, November, 2019, **Seattle, USA**.
- **Akolekar, H. D.**, Zhao, Y., Sandberg, R. D., Hutchins, N., and Michelassi, V. 'Turbulence Model Development for Low & High Pressure Turbines Using a Machine-Learning Approach', **24th International Society for Air Breathing Engines (ISABE)**, September, 2019, **Canberra, Australia**,
- **Akolekar, H. D.**, Sandberg, R. D., Hutchins, N., Michelassi, V., and Laskowski, G. 'Machine-Learnt Turbulence Closures for LPTs with Unsteady Inflow Conditions'. **15th (ISUAAAT)**, September, 2018, **University of Oxford, UK** (paper no. ISUAAAT-019).
- **Akolekar, H.D.,** 'Machine Learning Based Turbulence Modeling for Low Pressure Turbines', Invited Seminar: Whittle Laboratory, **University of Cambridge, UK**, September 2018.

The Team

Instructor: Dr. Harshal Akolekar (www.harshalakolekar.com)
E: harshal.akolekar@iitj.ac.in

Teaching Assistants:

Bikash Dutta
PhD Student, Dept. of Computer Science and Engineering, IIT-J
E: d22cs051@iitj.ac.in

Talib Ansari
PhD Student, Dept. of Mechanical Engineering, IIT-J
E: p23me0011@iitj.ac.in

Course Objectives

- To understand various key paradigms for pattern classification and machine learning approaches
- To familiarize with the mathematical and statistical techniques used in pattern recognition and machine learning.
- To understand and differentiate among various pattern recognition and machine learning techniques.

Learning Outcomes

The students are expected to have the ability to:

- To formulate a machine learning problem
- Select an appropriate pattern analysis tool for analyzing data in a given feature space.
- Apply pattern recognition and machine learning techniques such as classification and feature selection to practical applications and detect patterns in the data.

Course Content

- Introduction: Definitions, Datasets for Machine Learning, Different Paradigms of Machine Learning, Data Normalization, Hypothesis Evaluation, VC-Dimensions and Distribution, Bias-Variance Tradeoff, Linear Regression, Classification (5-6 Lectures)
- Bayes Decision Theory: Bayes decision rule, Minimum error rate classification, Normal density and discriminant functions Parameter Estimation: Maximum Likelihood and Bayesian Parameter Estimation (3-4 Lectures)
- Discriminative Methods: SVM, Distance-based methods, Linear Discriminant Functions, Decision Tree, Random Decision Forest and Boosting (4 Lectures)
- Dimensionality Reduction: PCA, LDA, ICA, SFFS, SBFS (2-3 Lectures)
- Clustering: k-means clustering, Gaussian Mixture Modeling, EM-algorithm (3 Lectures)
- Kernels and Neural Networks, Kernel Tricks, SVMs (primal and dual forms), K-SVR, K-PCA (2 Lectures)
- Artificial Neural Networks: MLP, Backprop, and RBF-Net (3 Lectures)
- Foundations of Deep Learning: CNN, Autoencoders (2-3 lectures)
- Time series analysis

Evaluation

- Continuous Evaluation (50%)
- Assignments – 25% (2) – in groups
- Quizzes - 25 % (3) - individual

- End semester - 50%

Class conduct – and assignment / quiz protocol

- Please do not copy work of your peers or off the internet.

- If you use other material, please refer to it via citations.

- Collaboration policy: if you have discussed with anyone, you should acknowledge them.

Plagiarism Policy

- Cheating in assignments / quizzes/ projects /

– First offence: Zero in the evaluation component

– Second offense: Grade reduction/ F grade

- Cheating in exam: F grade

- Misbehavior: Institute guidelines

Some points to note …

- Mixture of theory and applied content
- ML does have a lot of theory and math – we will go through that systematically !
- Onus is on the students as well to explore the theory and write code
- Course is designed based on a median student level

What is Machine Learning

- Learning is any process by which a system improves performance from experience – Herbert Simon

- [Machine learning is the] field of study that gives computers the ability to learn without being explicitly programmed—Arthur Samuel, 1959

- A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E. —Tom Mitchell, 1997

E, T, P examples.

T: Playing checkers P: Percentage of games won against an arbitrary opponent E: Playing practice games against itself

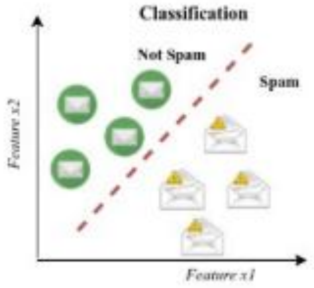
T: Recognizing hand-written words P: Percentage of words correctly classified E: Database of human-labeled images of handwritten words

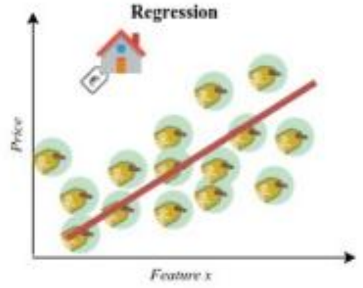
T: Driving on four-lane highways using vision sensors P: Average distance traveled before a human-judged error E: A sequence of images and steering commands recorded while observing a human driver.

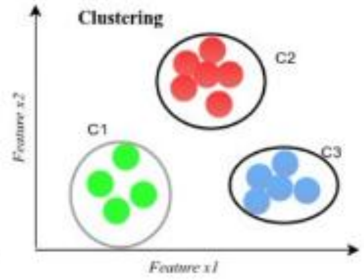
T: Categorize email messages as spam or legitimate. P: Percentage of email messages correctly classified. E: Database of emails, some with human-given labels

Task, T

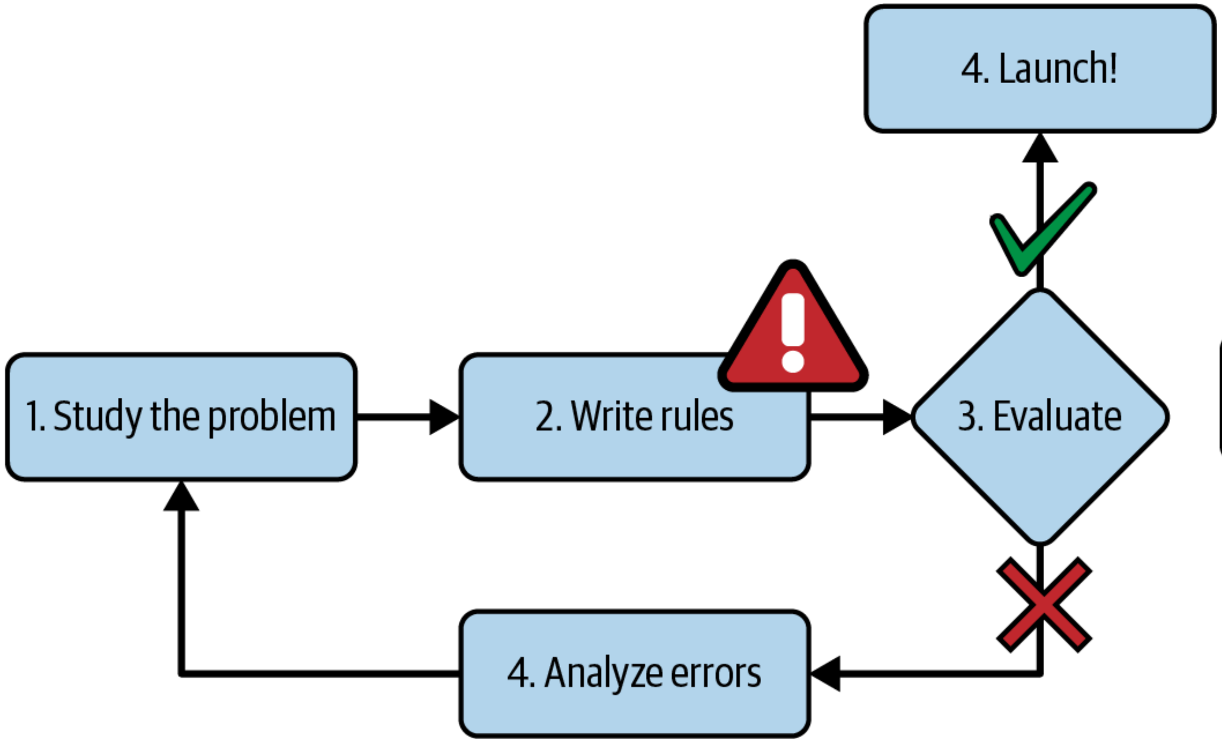
- Classification
- Regression
- Ranking
- Recommendation
- Clustering
- Density estimation
- ....



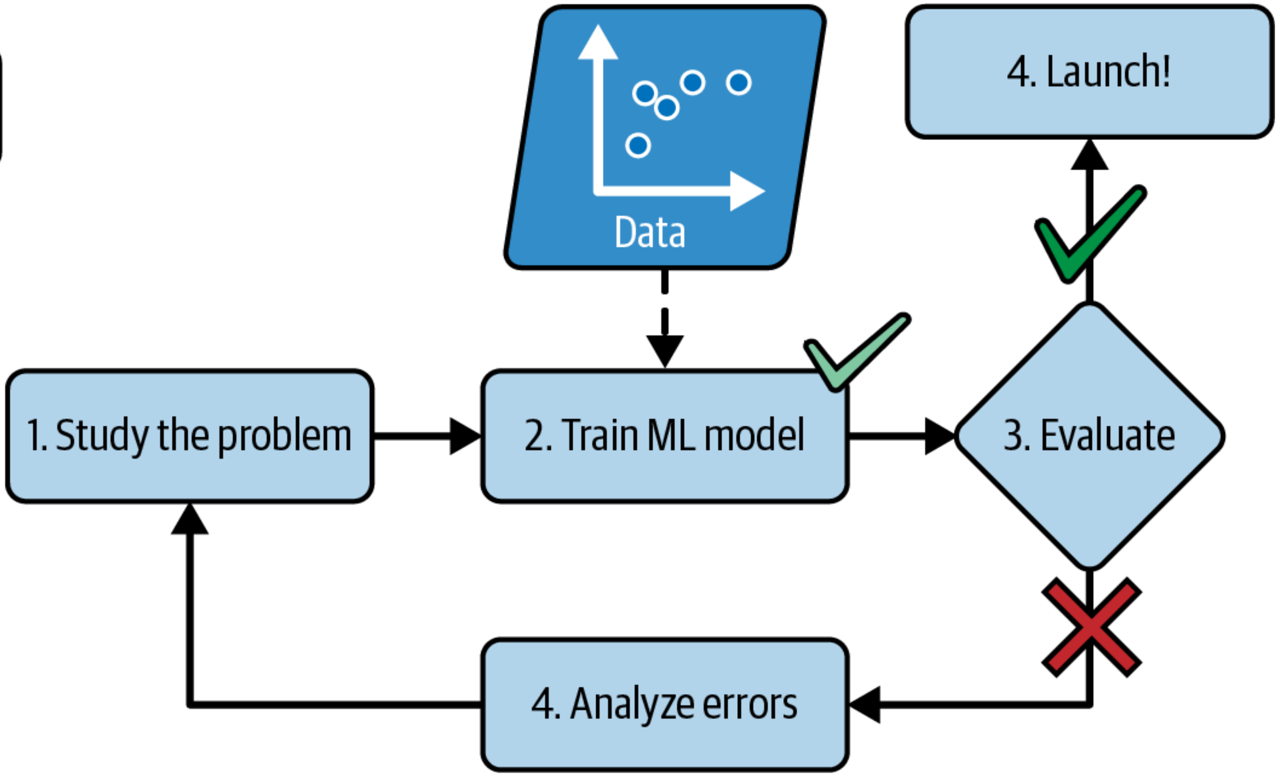




Traditional vs. ML approach

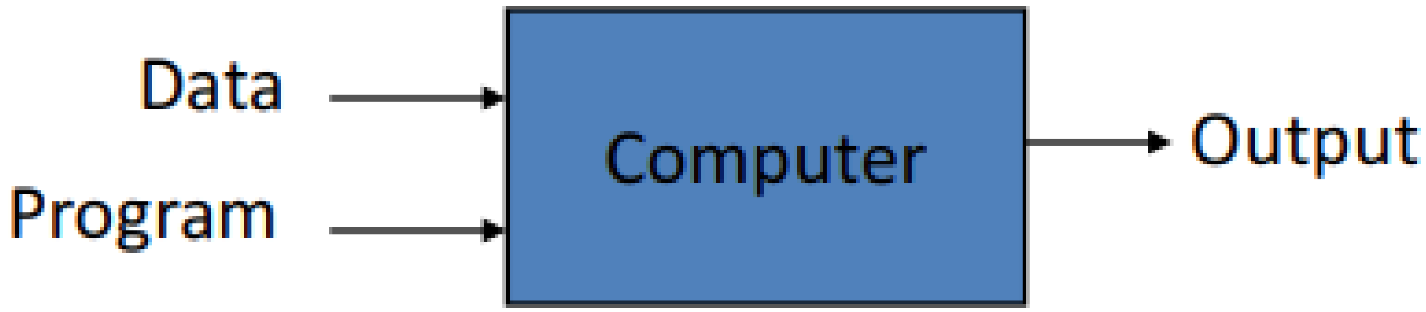


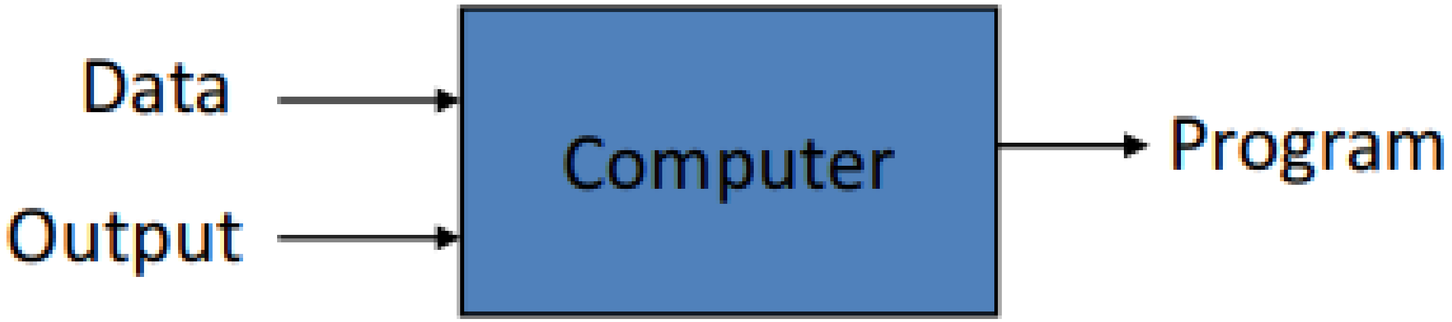
Traditional Approach



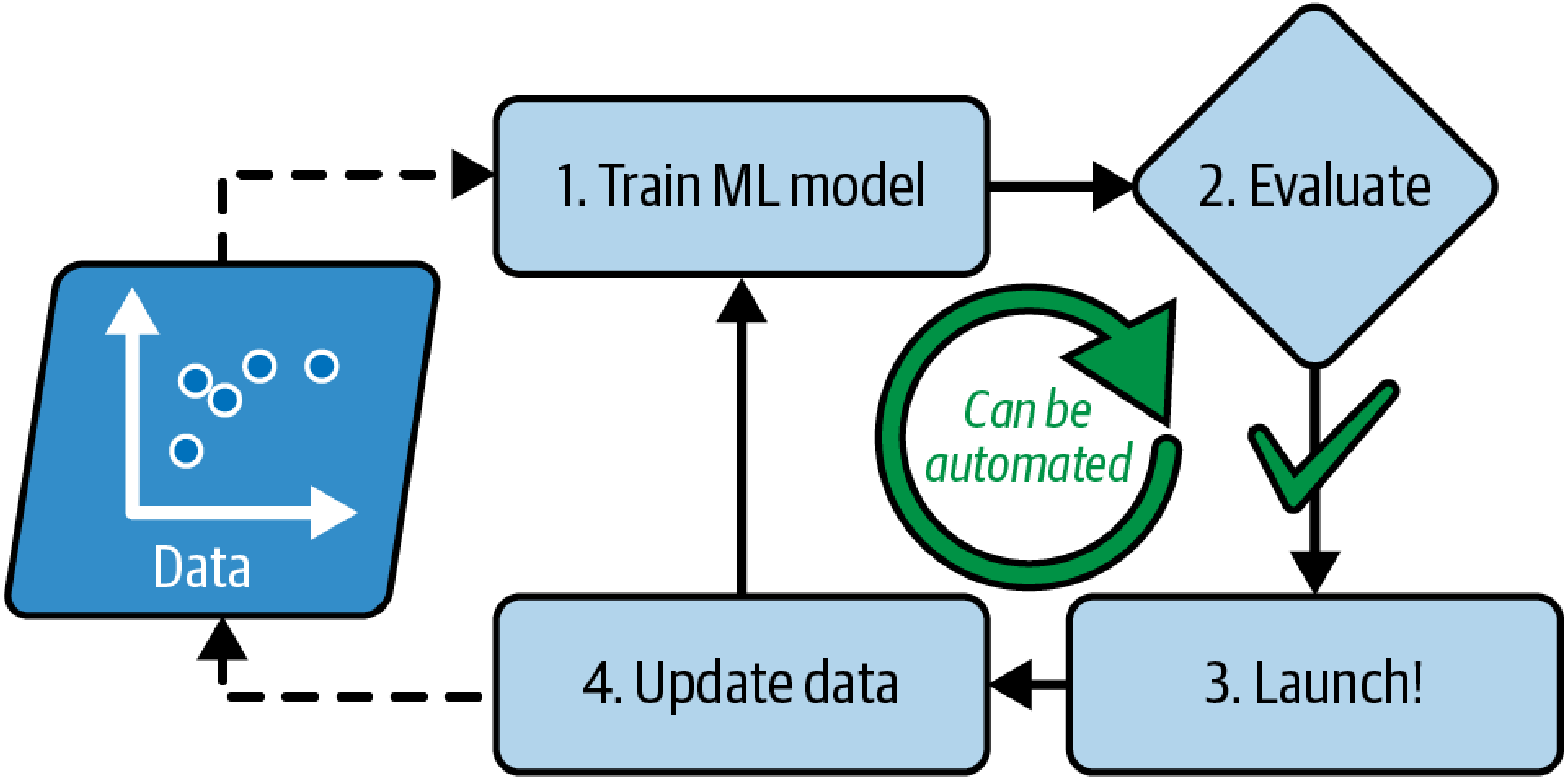
ML approach

Traditional vs. ML approach





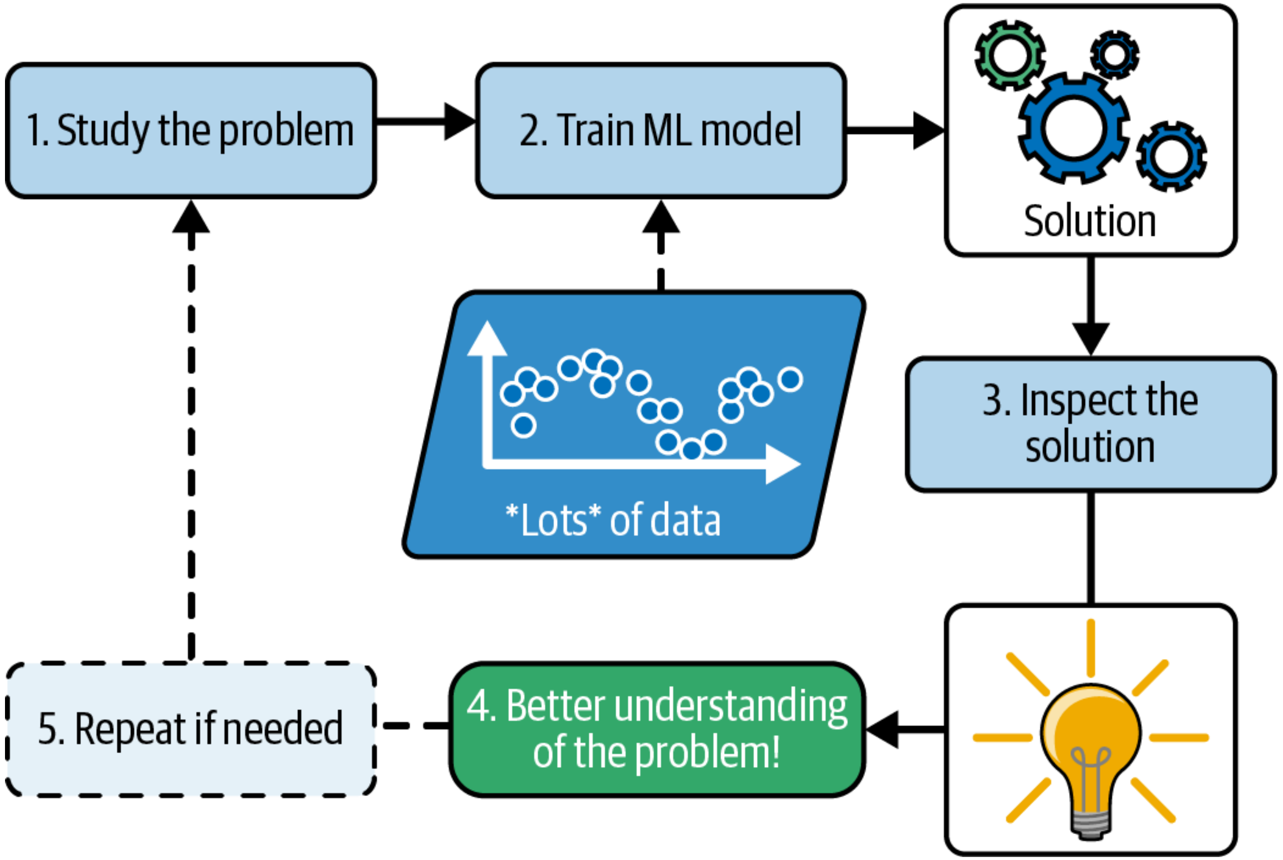
Data in the Loop



ML can help humans learn

ML is great for:

- Solutions which require a lot of fine tuning
- Complex problems for which the traditional approach is not good
- Fluctuating environments (new data being added to the mix)
- Getting insights from large amounts of data

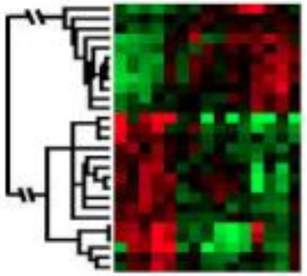


When do we use Machine Learning ?

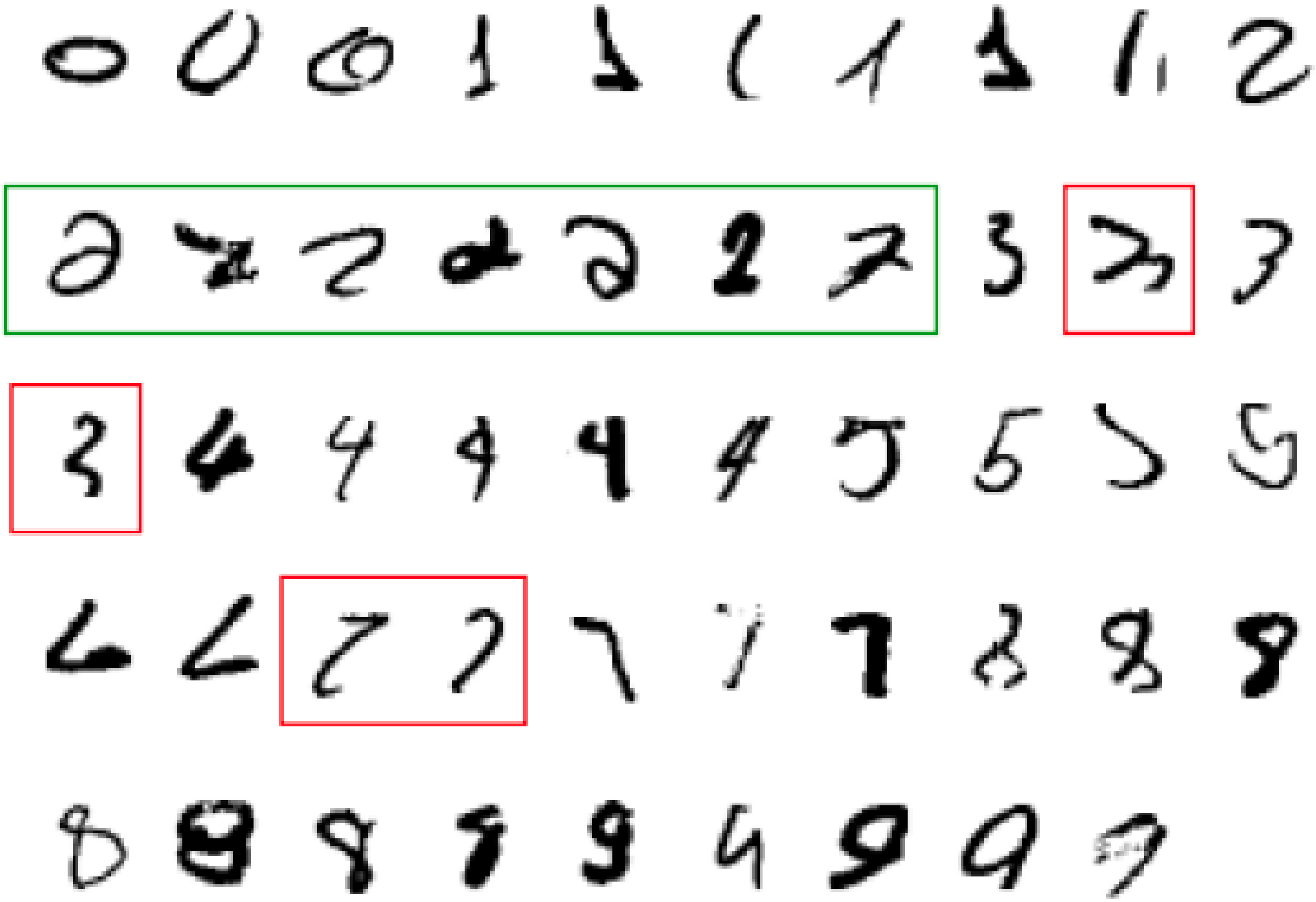
ML is used when:

- Human expertise does not exist (navigating on Mars)
- Humans can’t explain their expertise (speech recognition)
- Models must be customized (personalized medicine)
- Models are based on huge amounts of data (genomics)

- Learning isn't always useful



Classic ML task



What is 2 ?

Complex problems for which the traditional approach is not good

Learning Algorithm Tasks

- **Recognizing patterns:**

– Facial identities or facial expressions

– Handwritten or spoken words

– Medical images

- **Generating patterns:**

– Generating images or motion sequences

- **Recognizing anomalies:**

– Unusual credit card transactions

– Unusual patterns of sensor readings in a nuclear power plant

- **Prediction:**

– Future stock prices or currency exchange rates

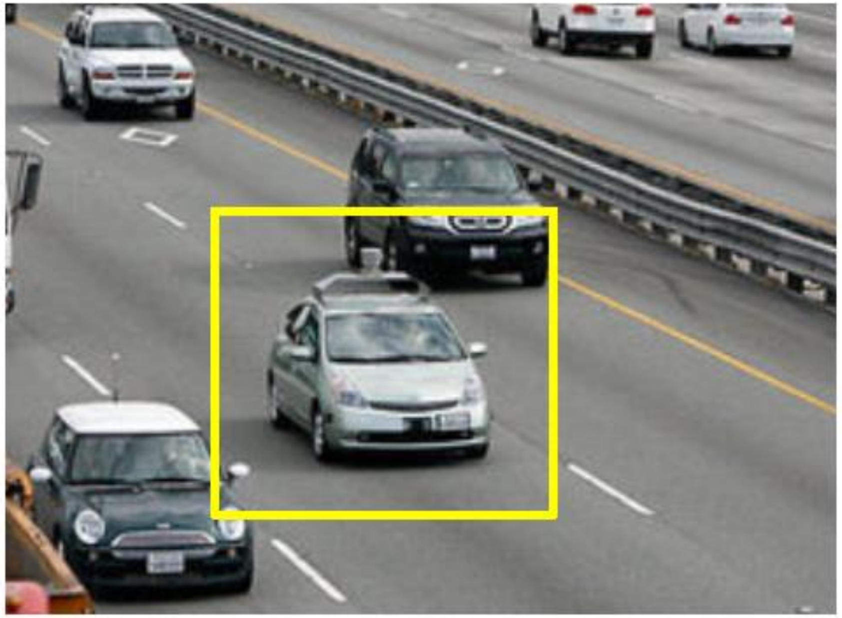
Sample Applications

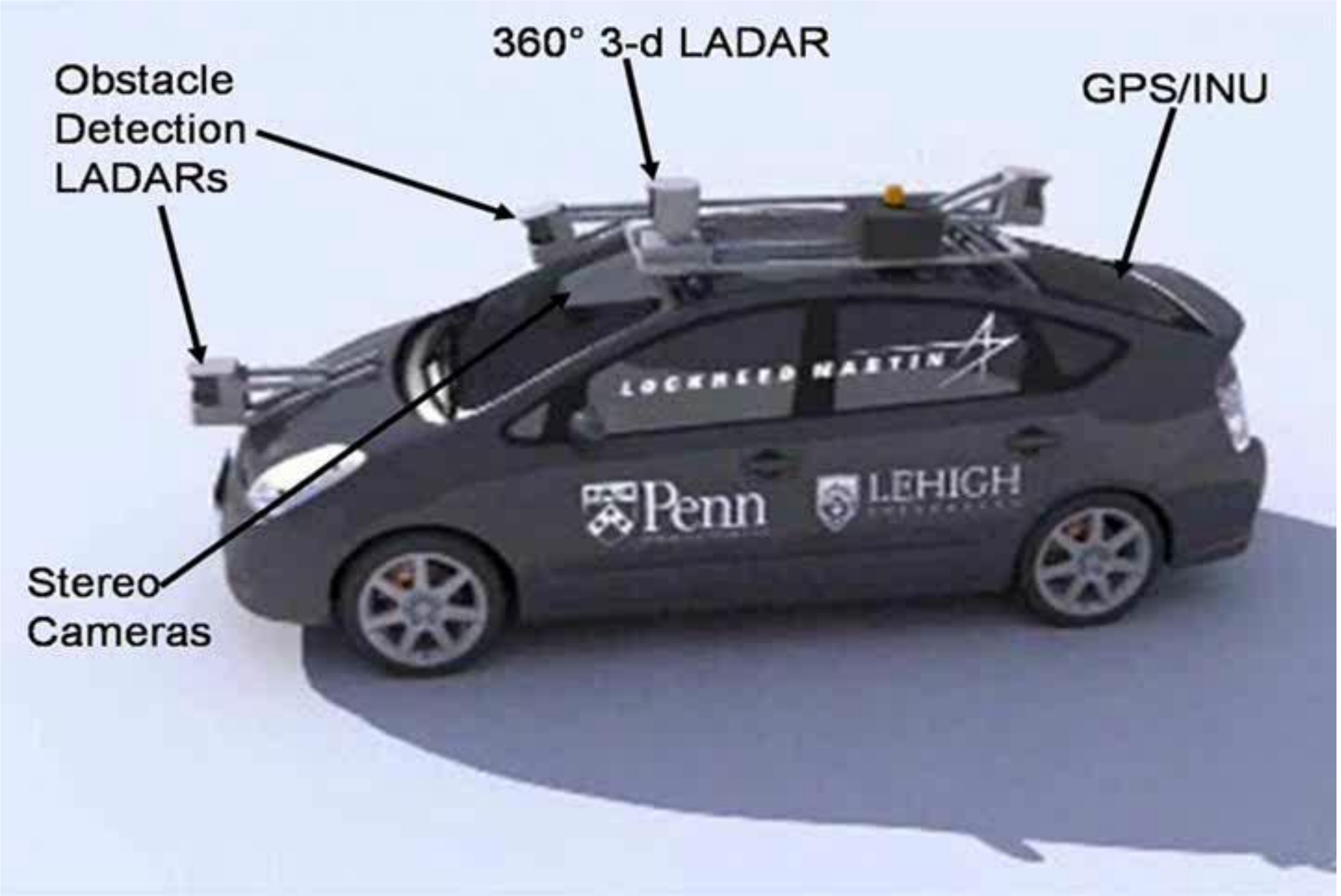
- Web search
- Computational biology
- Finance
- E-commerce
- Space exploration
- Robotics
- Information extraction
- Social networks
- Debugging software
- Medical imaging
- [Your favorite area]

Autonomous Cars

Nevada made it legal for autonomous cars to drive on roads in June 2011

As of 2013, four US states have legalized cars.

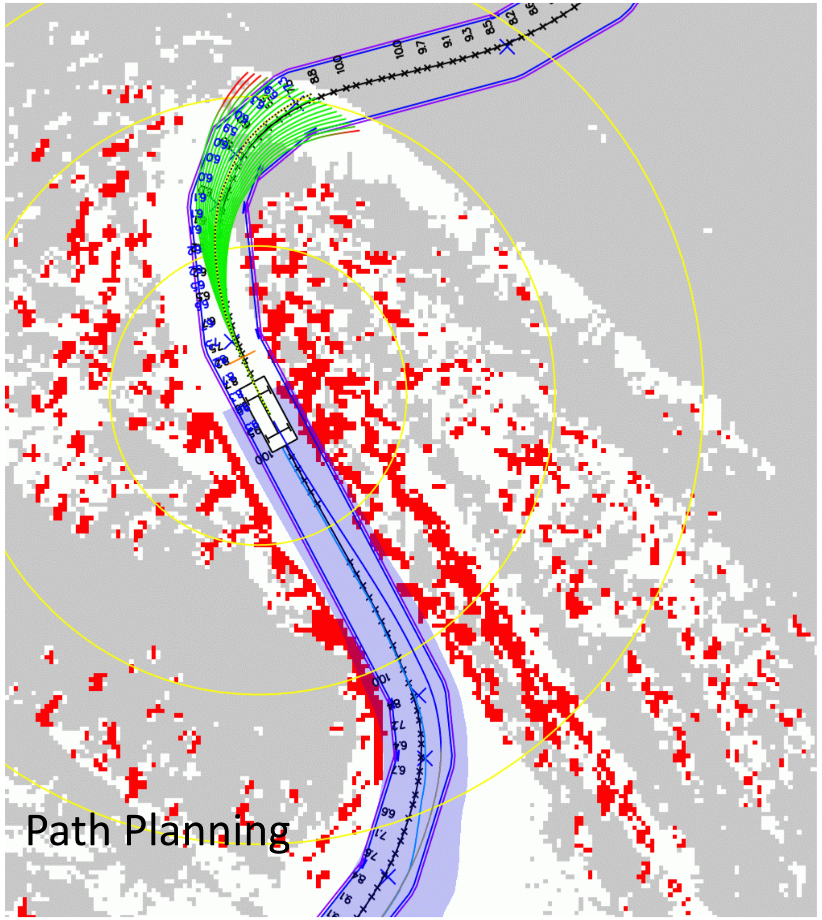




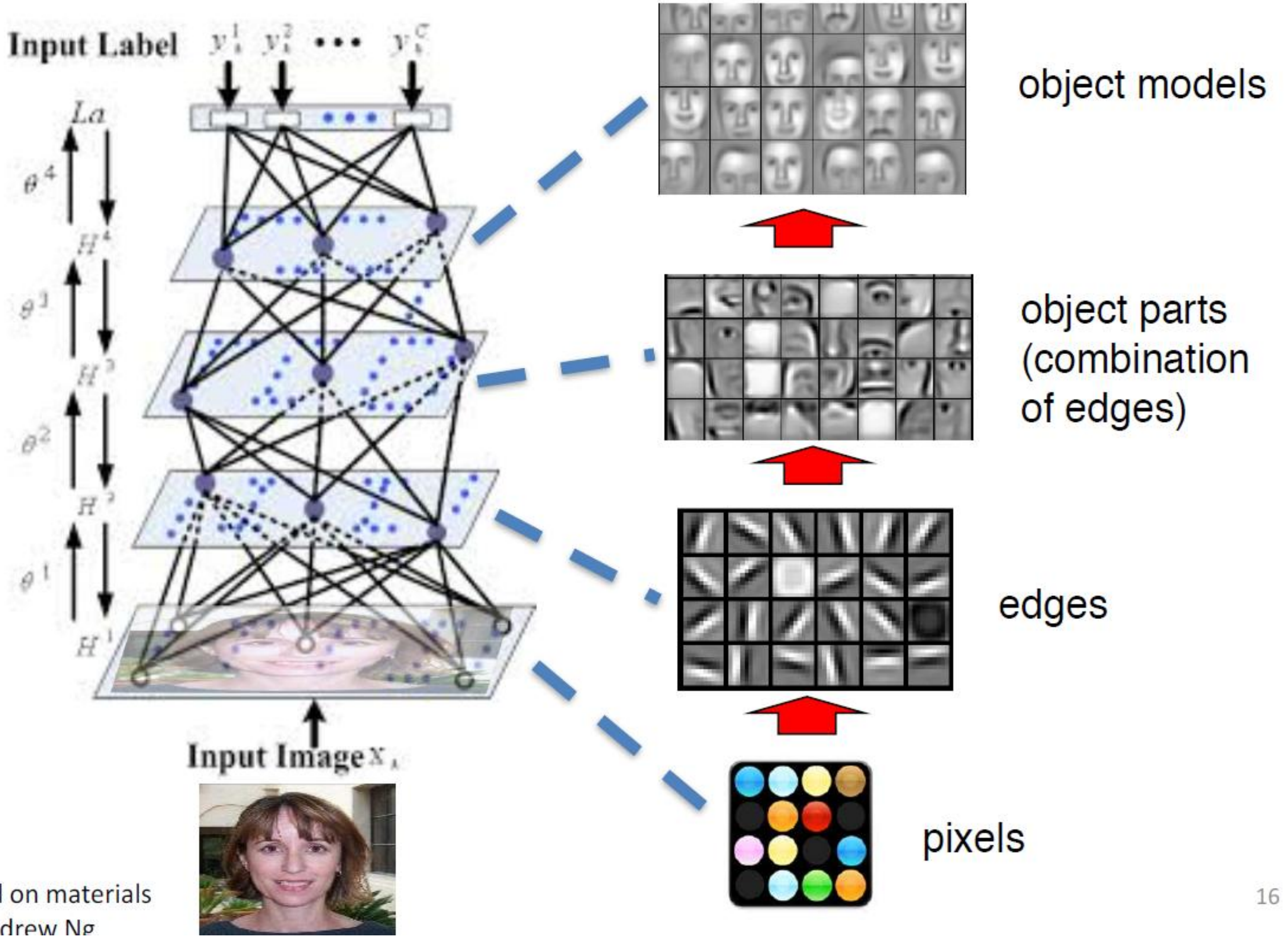
360° 3-d LADAR
Obstacle
Detection
LADARs
GPS/INU
LOCKHEED MARTIN
Penn
LEHIGH
Stereo
Cameras

Autonomous Cars Technology


Adaptive Vision

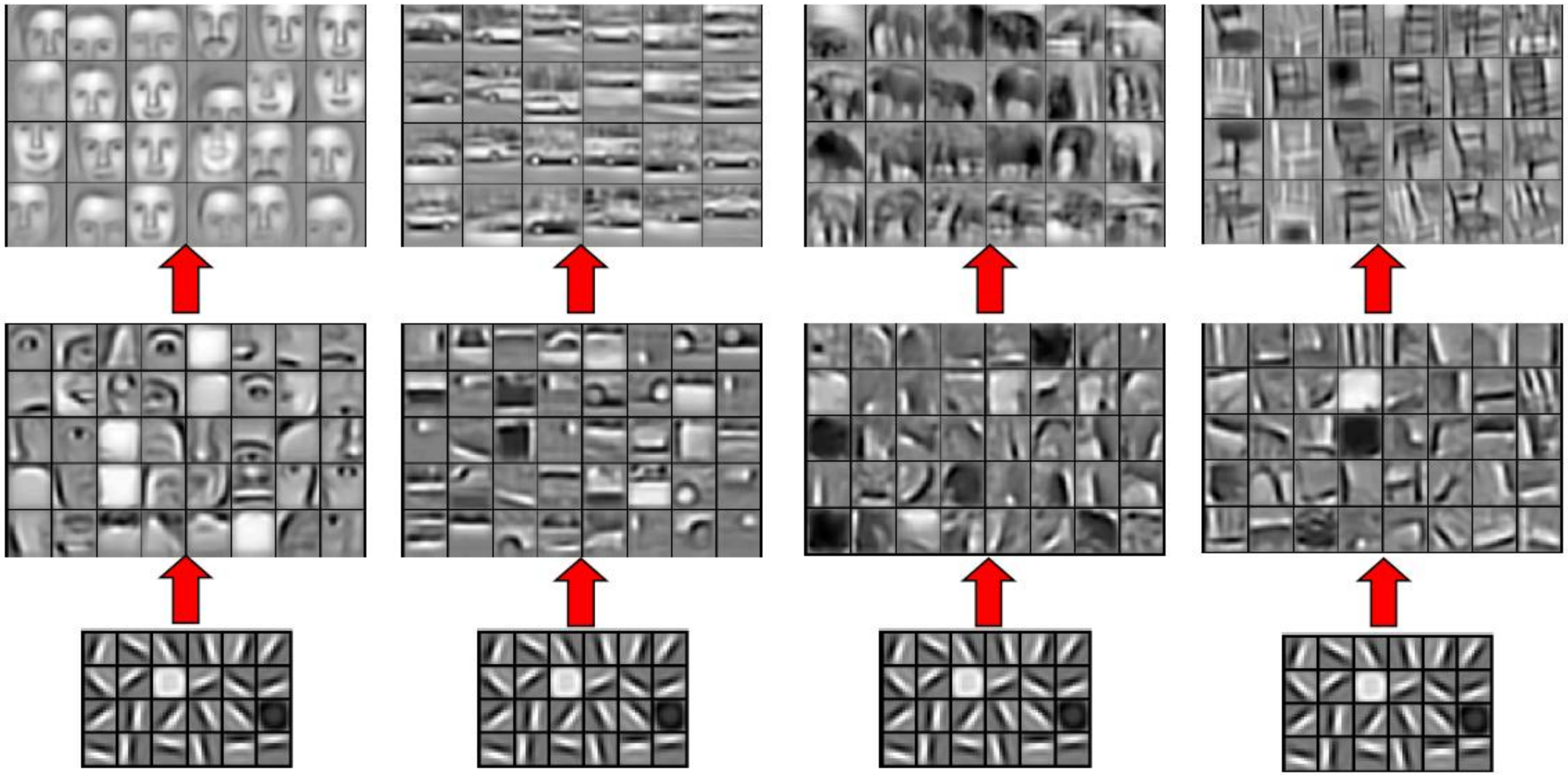


Facial Images



Based on materials
by Andrew Ng

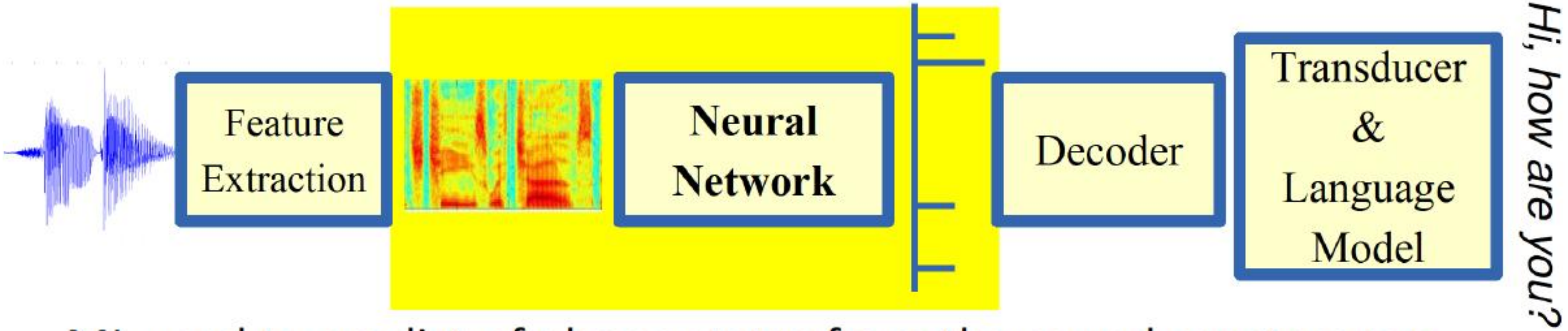
Learning of Object Parts



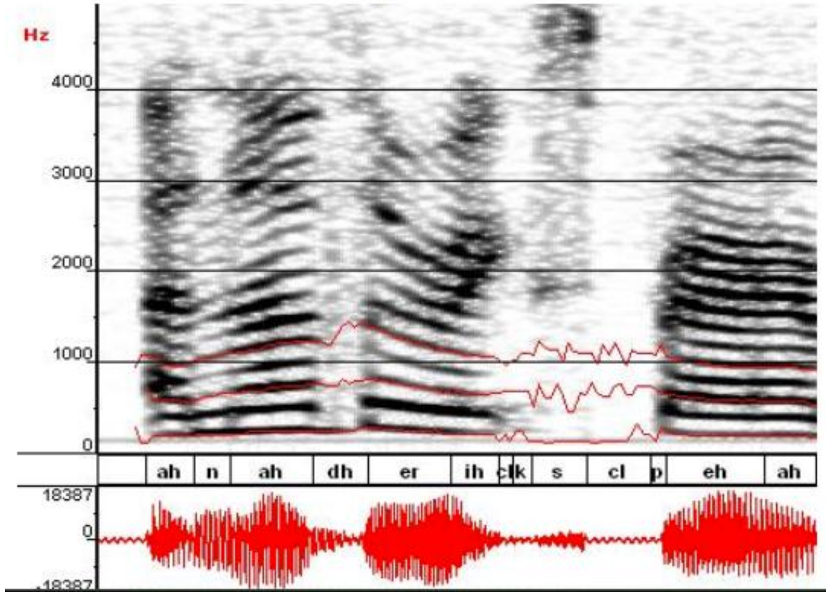
Scene Labelling via Deep Learning

ML in Automatic Speech Recognition

A Typical Speech Recognition System



ML used to predict of phone states from the sound spectrogram



Deep learning has state-of-the-art results

# Hidden Layers	1	2	4	8	10	12
Word Error Rate %	16.0	12.8	11.4	10.9	11.0	11.1

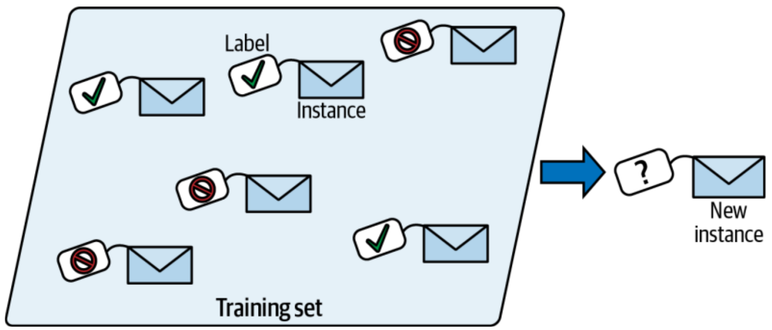
Baseline GMM performance = 15.4%

[Zeiler et al. “On rectified linear units for speech recognition” ICASSP 2013]

Impact of Deep Learning in Speech Technology

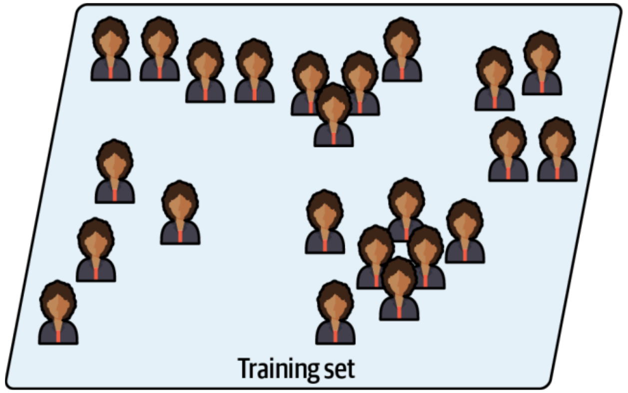
Slide credit: Li Deng, MS Research

Types of Learning



- **Supervised (inductive) learning**

– Given: training data + desired outputs (labels)

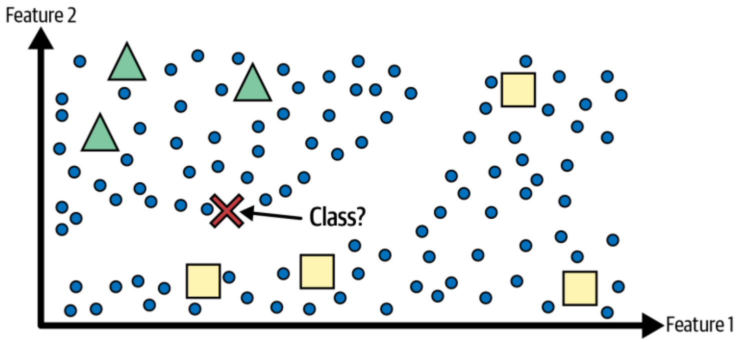


- **Unsupervised learning**

– Given: training data (without desired outputs)

- **Semi-supervised learning**

– Given: training data + a few desired outputs

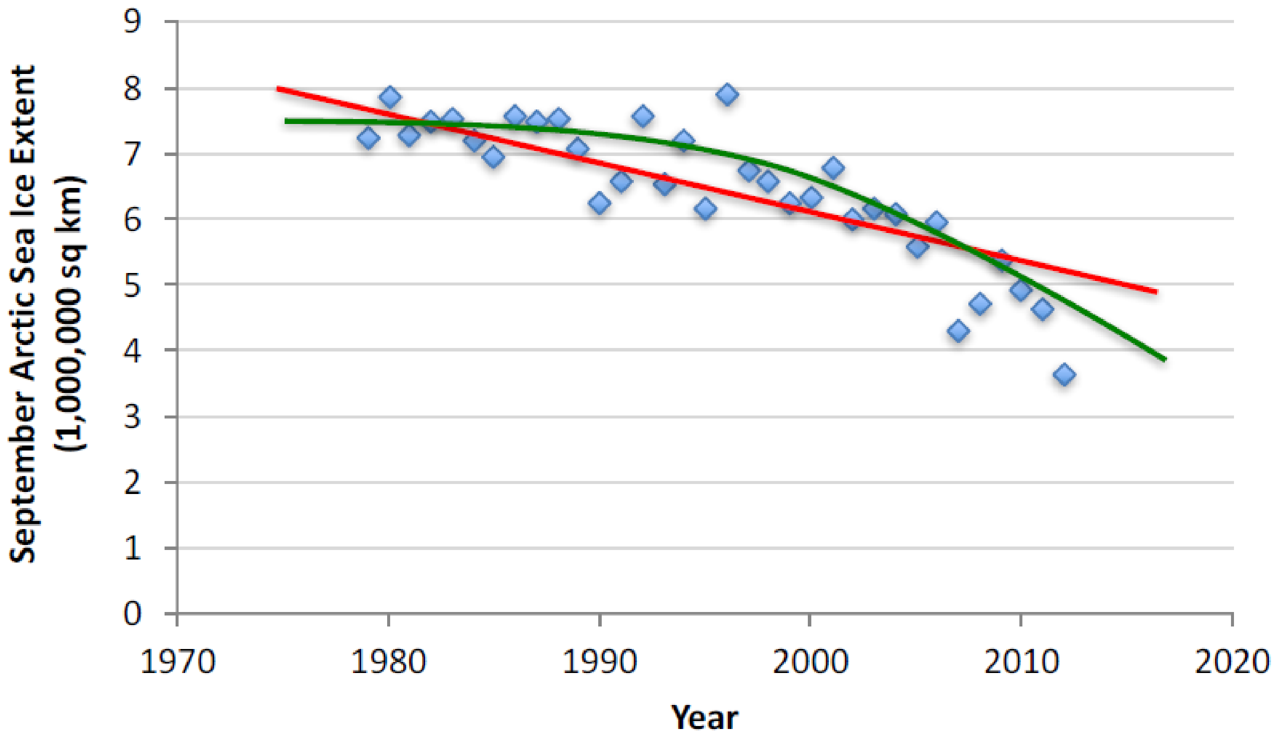


- **Reinforcement learning**

– Rewards from sequence of actions

Supervised Learning

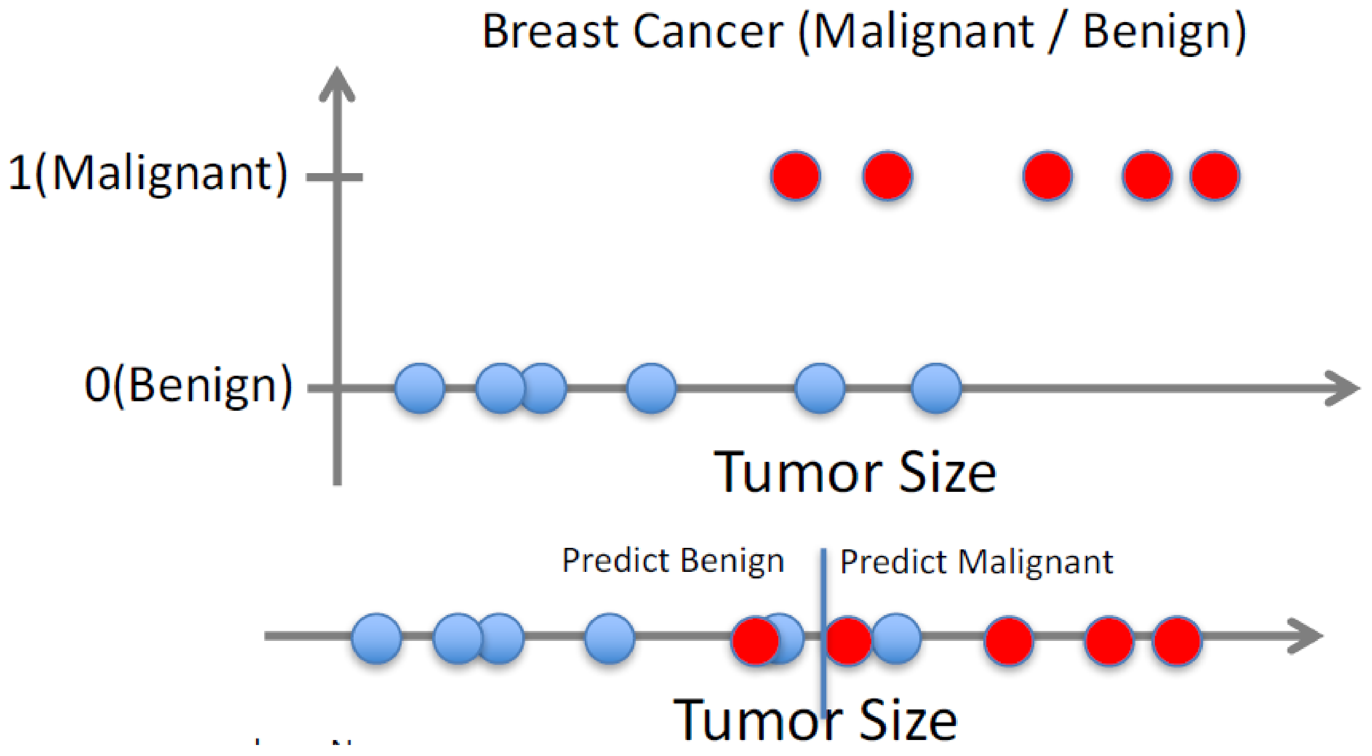
- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is real-valued == regression



Data from G. Witt. Journal of Statistics Education, Volume 21, Number 1 (2013)

Supervised Learning: Classification

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



Supervised Learning

- Given: training data + desired outputs (labels)
- (x1, y1), (x2, y2), ..., (xn, yn)
- Learn a function f(x) to predict y given x

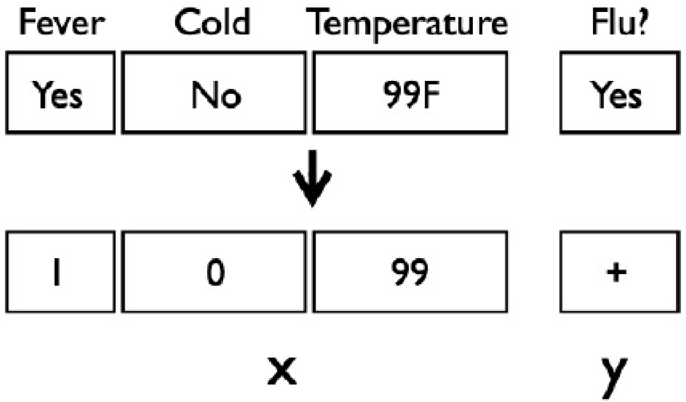
Cats

Dogs

Supervised Learning

Example 1: Predict if a **new** patient has flu or not, based on **existing** patient data

What is x and y?

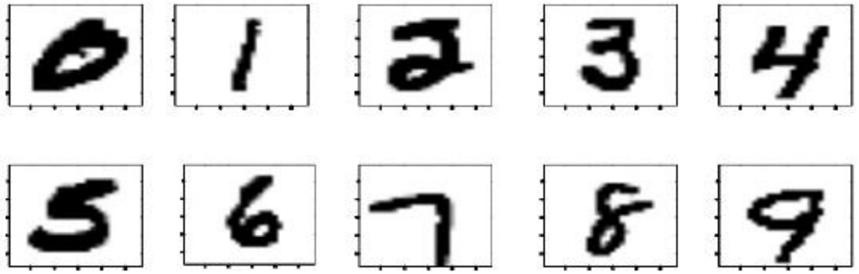


Features: Properties of patient

Label: Flu/No flu

A **binary** (two-label) classification problem

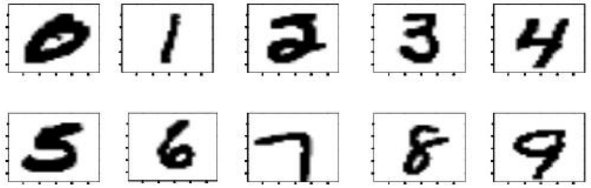
Example 2: Which digit in the image ?



Label: 0,1,..,9

What are the features?

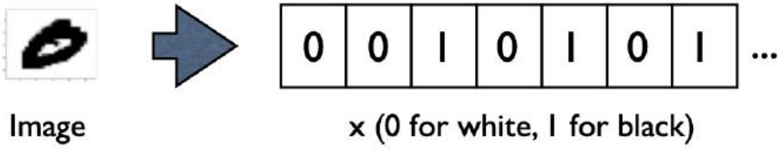
A **multiclass** classification problem



Label: 0,1,..,9

What are the features?

Option: vector of pixel colors



There are other options too

Lesson: Choosing features is non-trivial in real applications

Supervised Learning

Example 3: Spam or not?

Email 1
From: Canadian Pharmacy
Subject: Offer ends now!

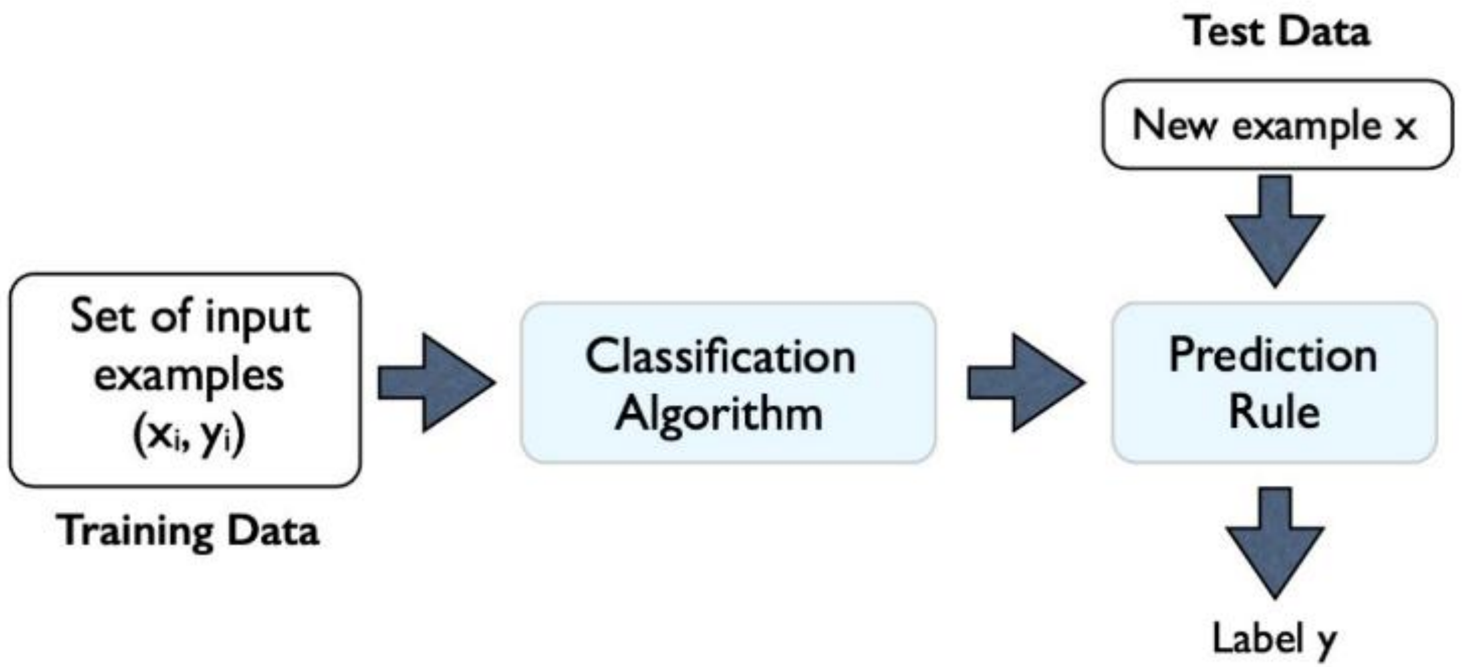
Email 2
From: Yuncong Chen
Subject: TA meeting

	Pharmacy	offer	meeting	TA	Spam?
Email 1	1	1	0	0	Yes
Email 2	0	0	1	1	No

Label: 0 (not spam), 1 (spam)

Features: Words in the email

Classification Algorithm

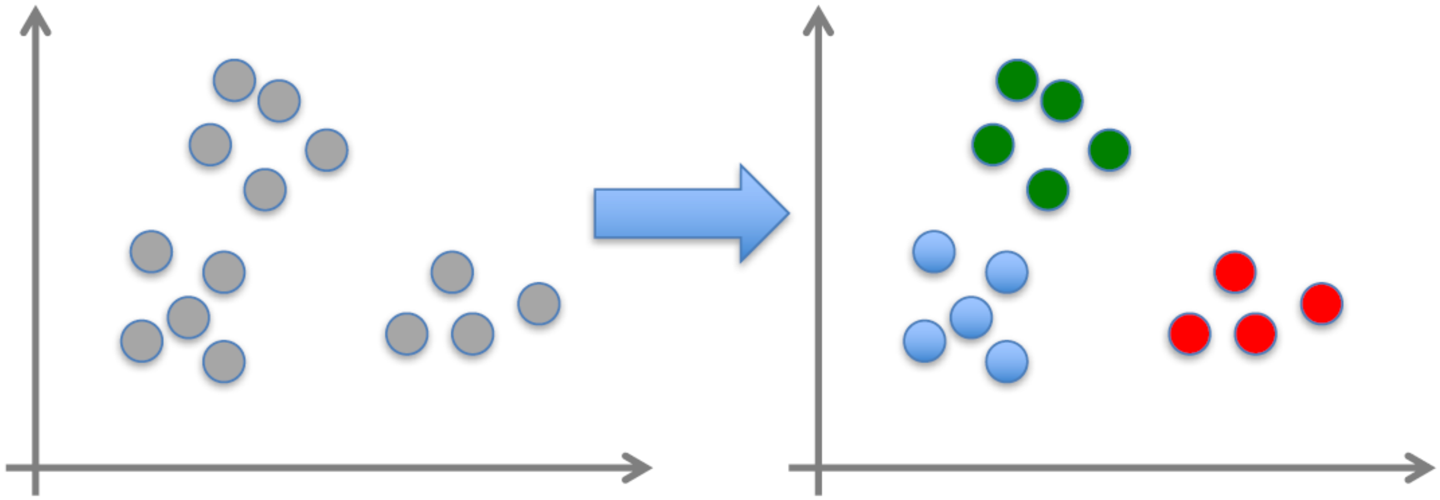


Training and test data must be **separate**!

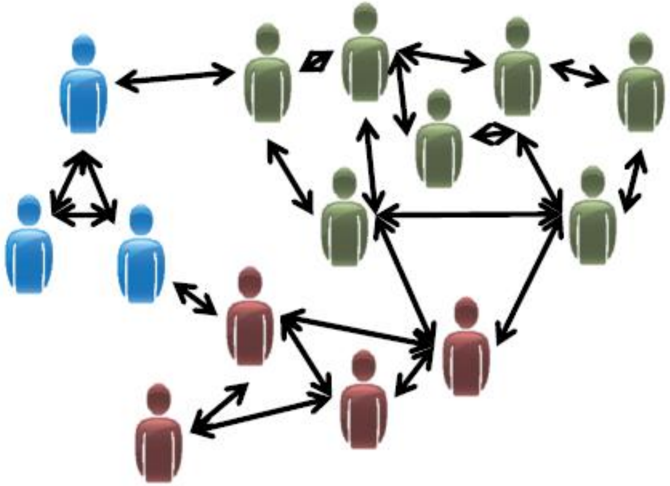
Performance Measure:
Accuracy (or fraction of correct answers) on **test data**

Unsupervised Learning

- Given x_1, x_2, …, x_n (without labels)
- Output hidden structure behind the x's
  - E.g., clustering

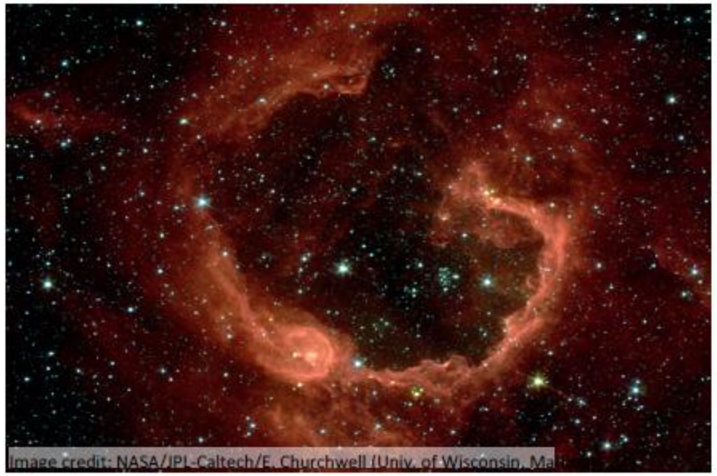


Organize computing clusters



Social network analysis

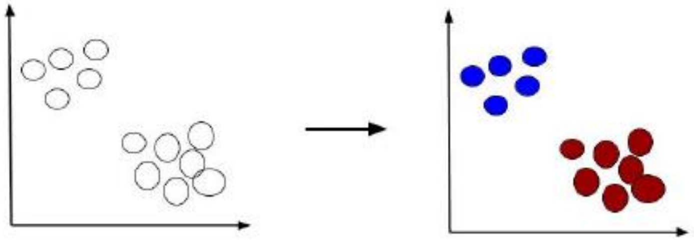
Market segmentation



Astronomical data analysis

Unsupervised Learning

- Given: training data (without desired outputs)
- x1, x2, ..., xn (without labels)
- Output hidden structure behind the x's – Group them by similarity



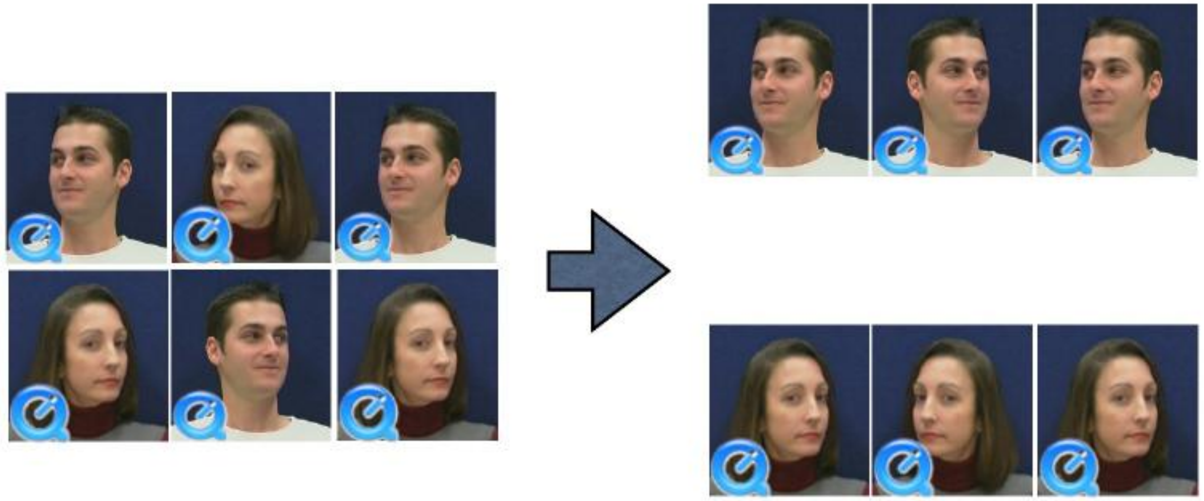
Clustering

Unsupervised Learning

Clustering

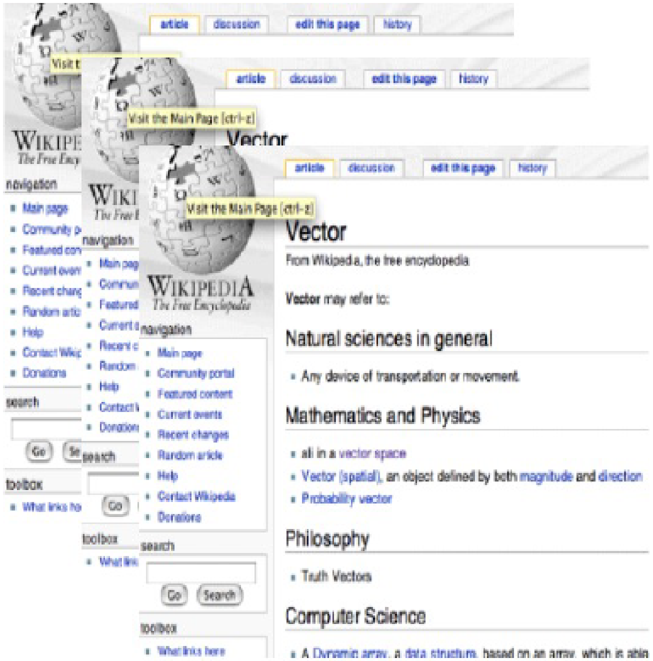
Given a set of input objects, group them to clusters by similarity

Example 1: Cluster videos by people in them



Unsupervised Learning

Example 2: Cluster documents by topic



Physics

- Gravity
- Laws of Motion
- Electricity

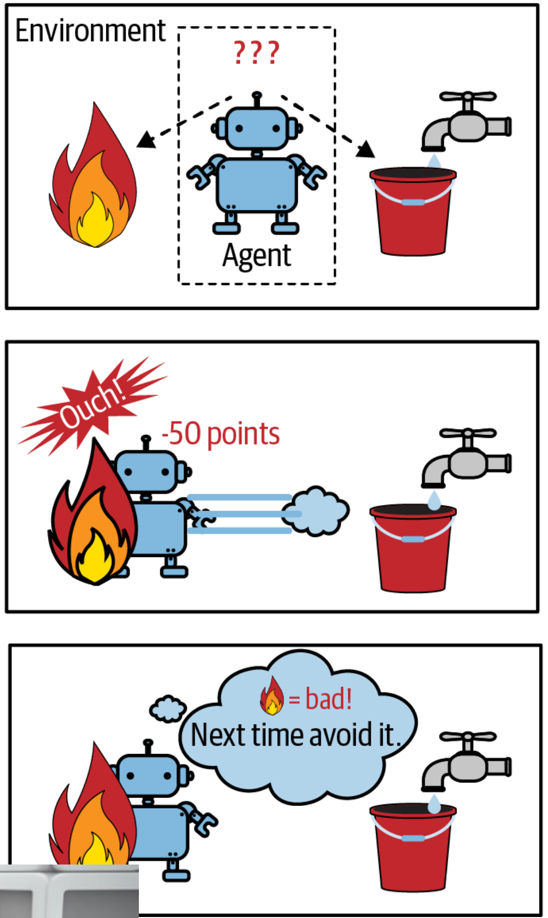
Math

- Geometry
- Algebra

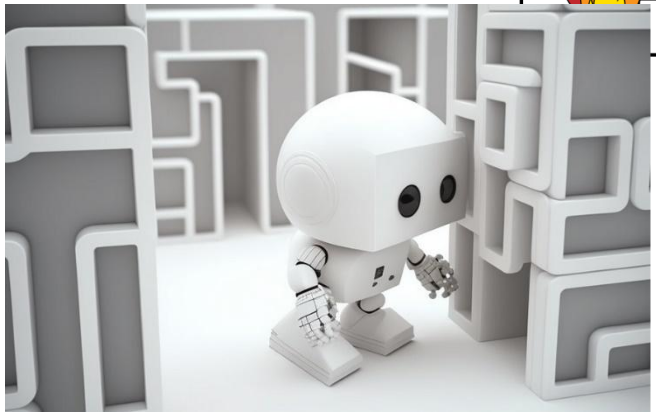
Features: Words in the document

Reinforcement Learning

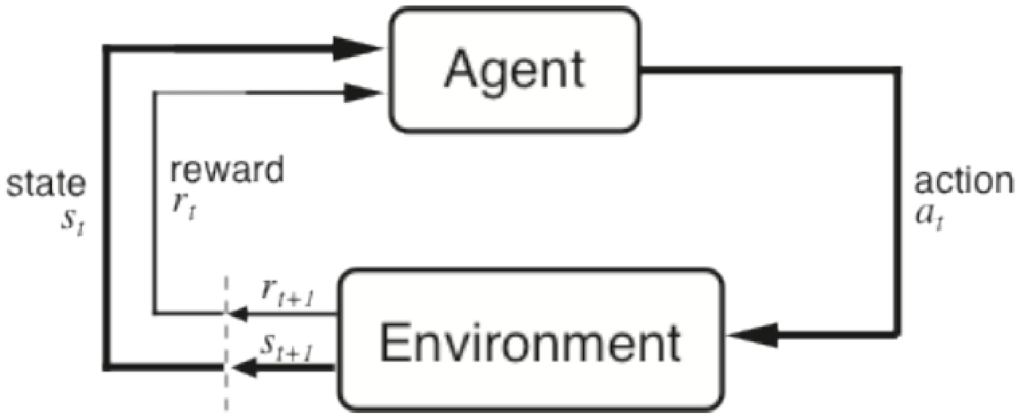
- Given a sequence of states and actions with (delayed) rewards, output a policy
  - Policy is a mapping from states → actions that tells you what to do in a given state
- Examples:
  - Credit assignment problem
  - Game playing
  - Robot in a maze
  - Balance a pole on your hand



1. Observe
2. Select action using policy
3. Action!
4. Get reward or penalty
5. Update policy (learning step)
6. Iterate until an optimal policy is found



Agent – Environment Interface



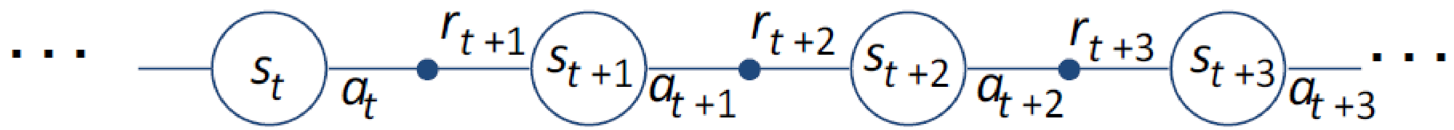
Agent and environment interact at discrete time steps : $t = 0, 1, 2, \ldots$

Agent observes state at step t: $s_t \in S$

produces action at step t: $a_t \in A(s_t)$

gets resulting reward : $r_{t+1} \in \Re$

and resulting next state : s_{t+1}



ChatGPT & Car Driving

- An example of Reinforcement Learning
- Learns from large human conversation datasets.
- Car driving example:
- https://www.youtube.com/watch?v=G-GpY7bevuw

- Any thoughts ?

History of ML

- 1950s
  - Samuel’s checker player
  - Selfridge’s Pandemonium
- 1960s:
  - Neural networks: Perceptron
  - Pattern recognition
  - Learning in the limit theory
  - Minsky and Papert prove limitations of Perceptron
- 1970s:
  - Symbolic concept induction
  - Winston’s arch learner
  - Expert systems and the knowledge acquisition bottleneck
  - Quinlan’s ID3
  - Michalski’s AQ and soybean diagnosis
  - Scientific discovery with BACON
  - Mathematical discovery with AM

History of ML

- 2000s
  - Support vector machines & kernel methods
  - Graphical models
  - Statistical relational learning
  - Transfer learning
  - Sequence labeling
  - Collective classification and structured outputs
  - Computer Systems Applications (Compilers, Debugging, Graphics, Security)
  - E-mail management
  - Personalized assistants that learn
  - Learning in robotics and vision
- 2010s
  - Deep learning systems
  - Learning for big data
  - Bayesian methods
  - Multi-task & lifelong learning
  - Applications to vision, speech, social networks, learning to read, etc.
- 2020s
  - Deep learning
  - ???

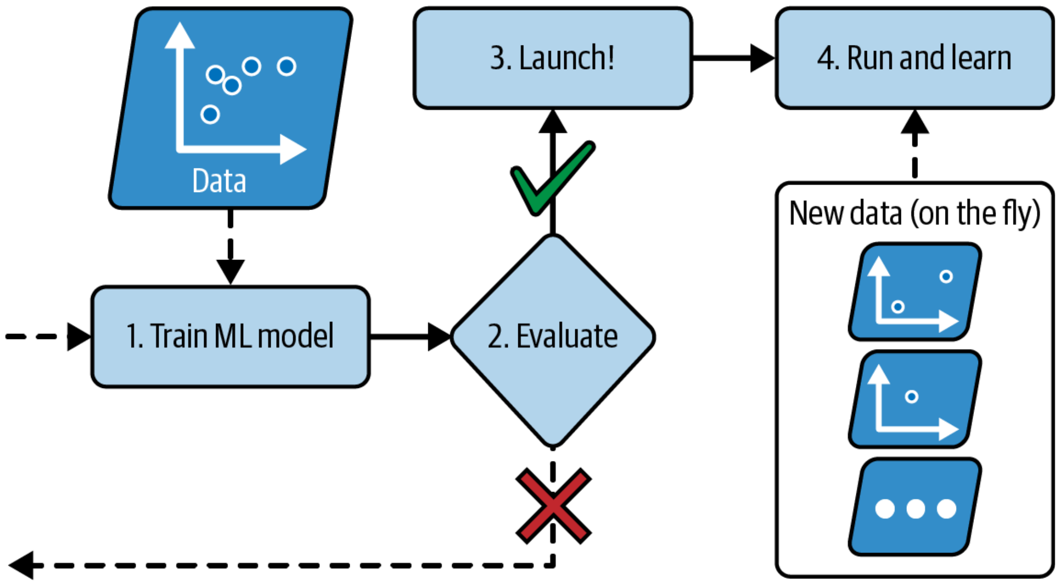
ML System Classification

- Batch vs. Online Learning

Batch:

- Incapable of learning incrementally
- Must be trained with all data.
- Trained and then launched into production.
- Aka offline learning.

Online learning exactly the opposite

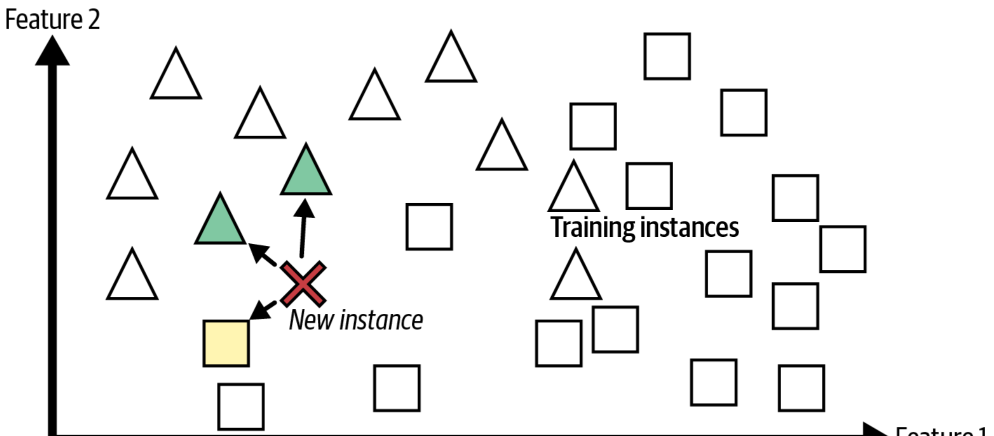


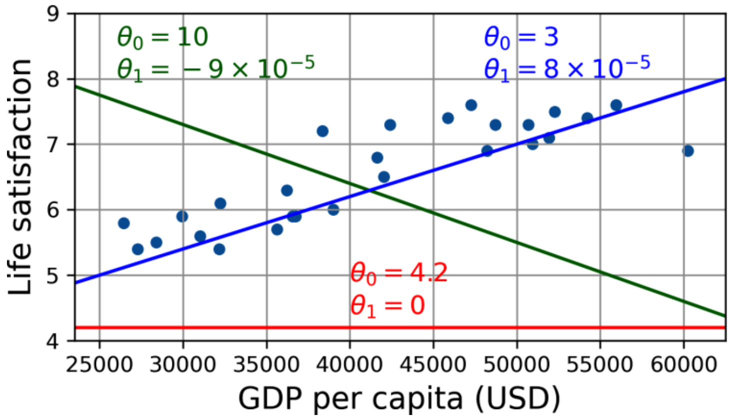
Online learning

ML System Classification

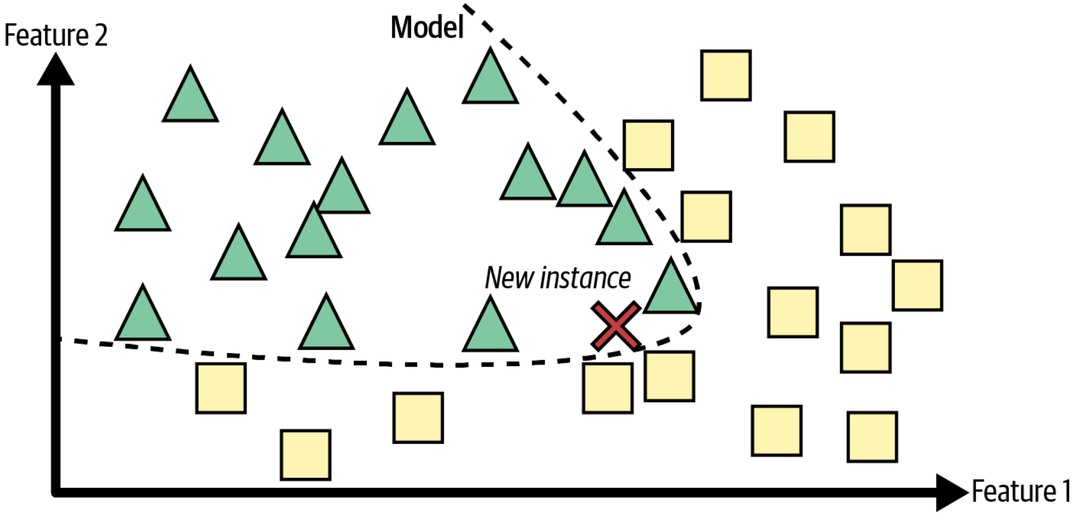
- Instance Based vs. Model-Based Learning

Instance-based learning





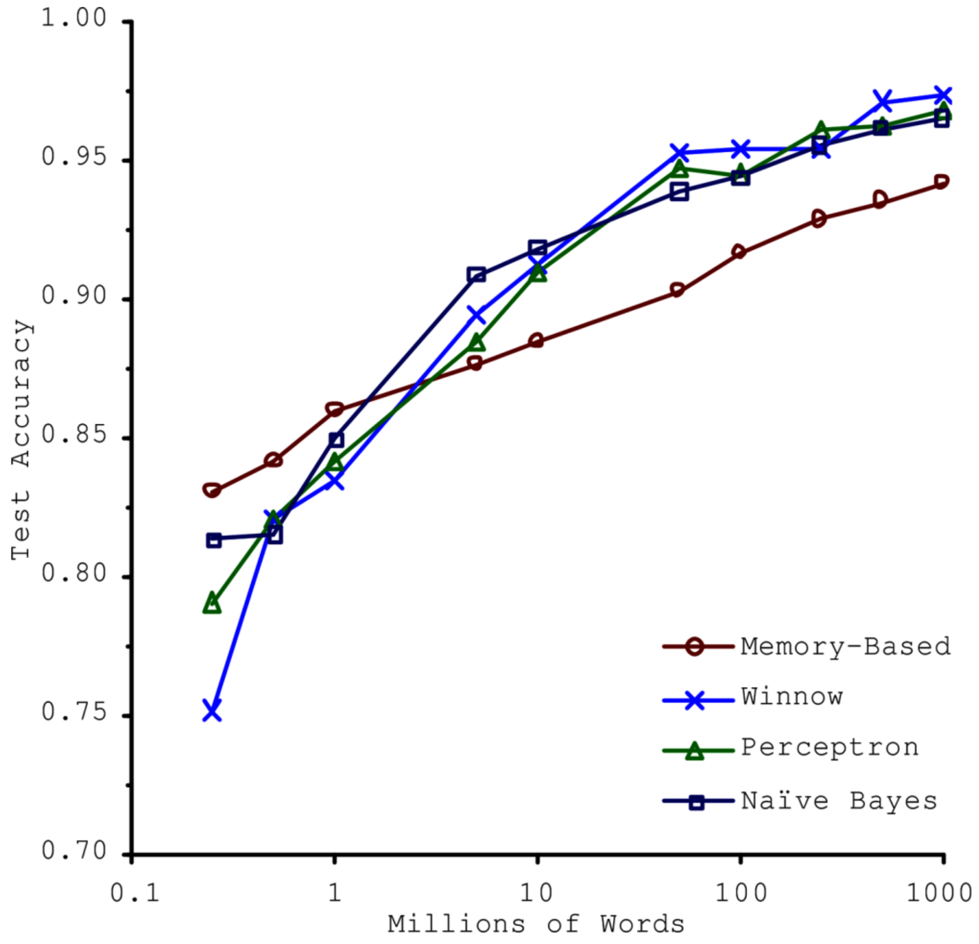
A simple regression model



Model-based learning

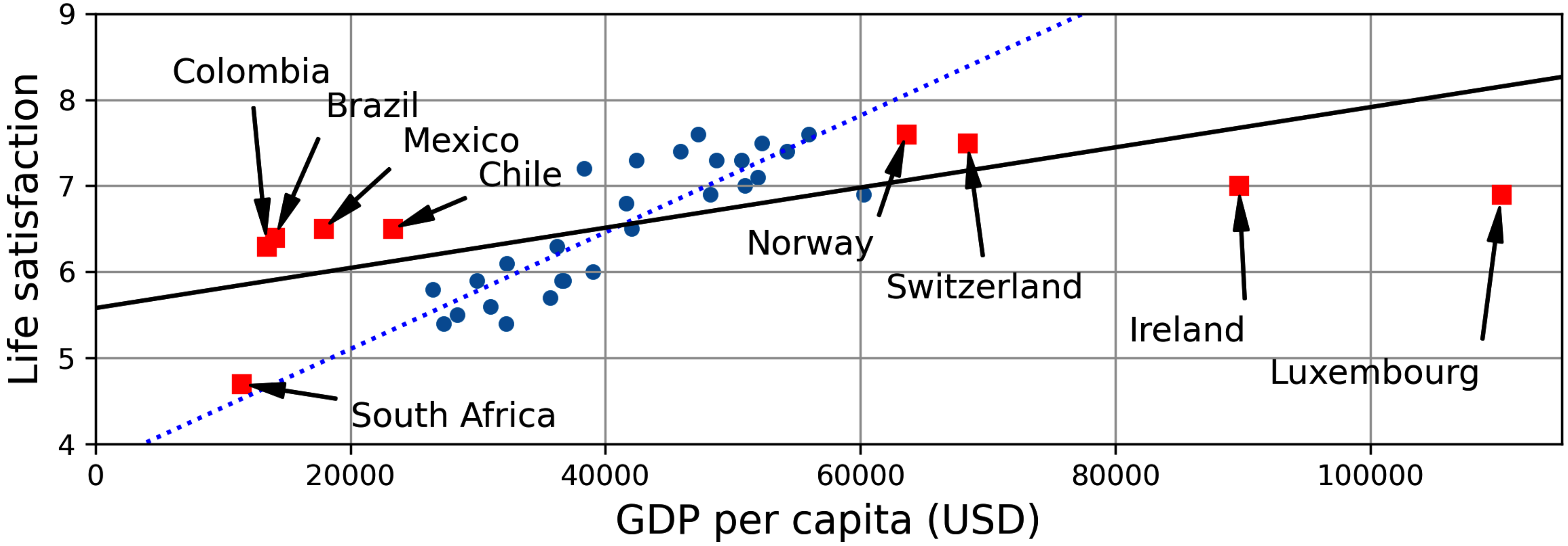
Challenges of Machine Learning

- Insufficient Quantity of Data.



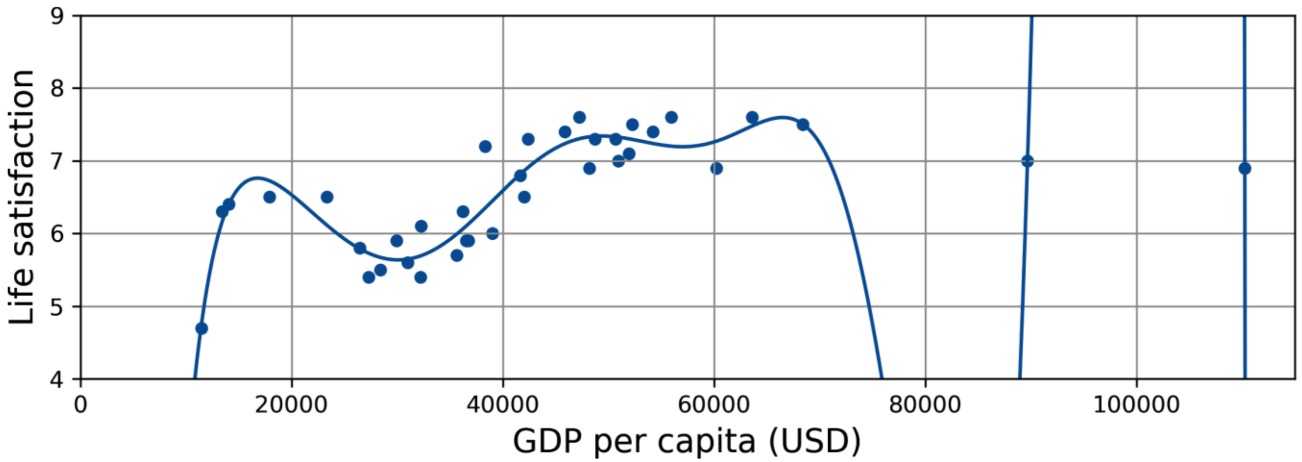
Challenges of Machine Learning

- Nonrepresentative Training Data
- Sampling Bias

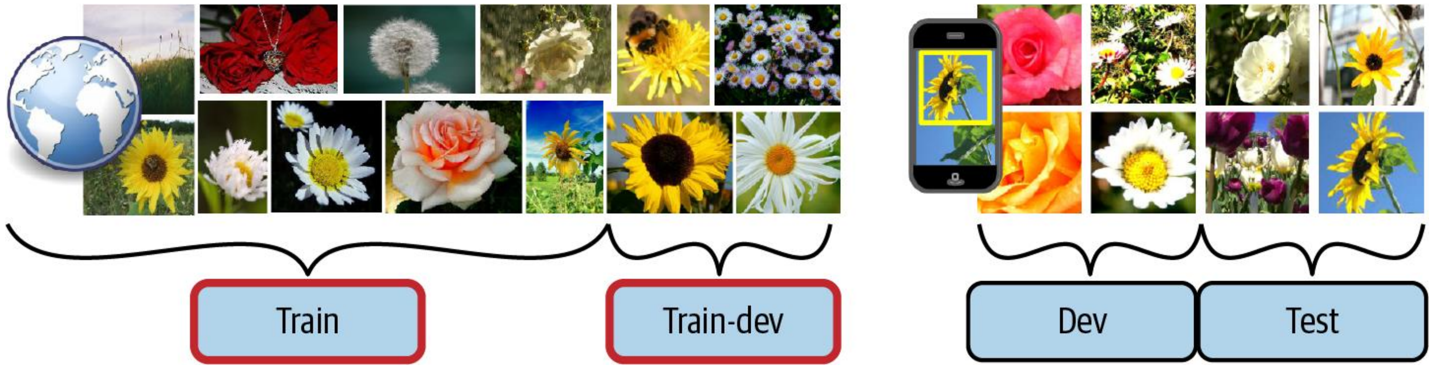


Challenges of Machine Learning

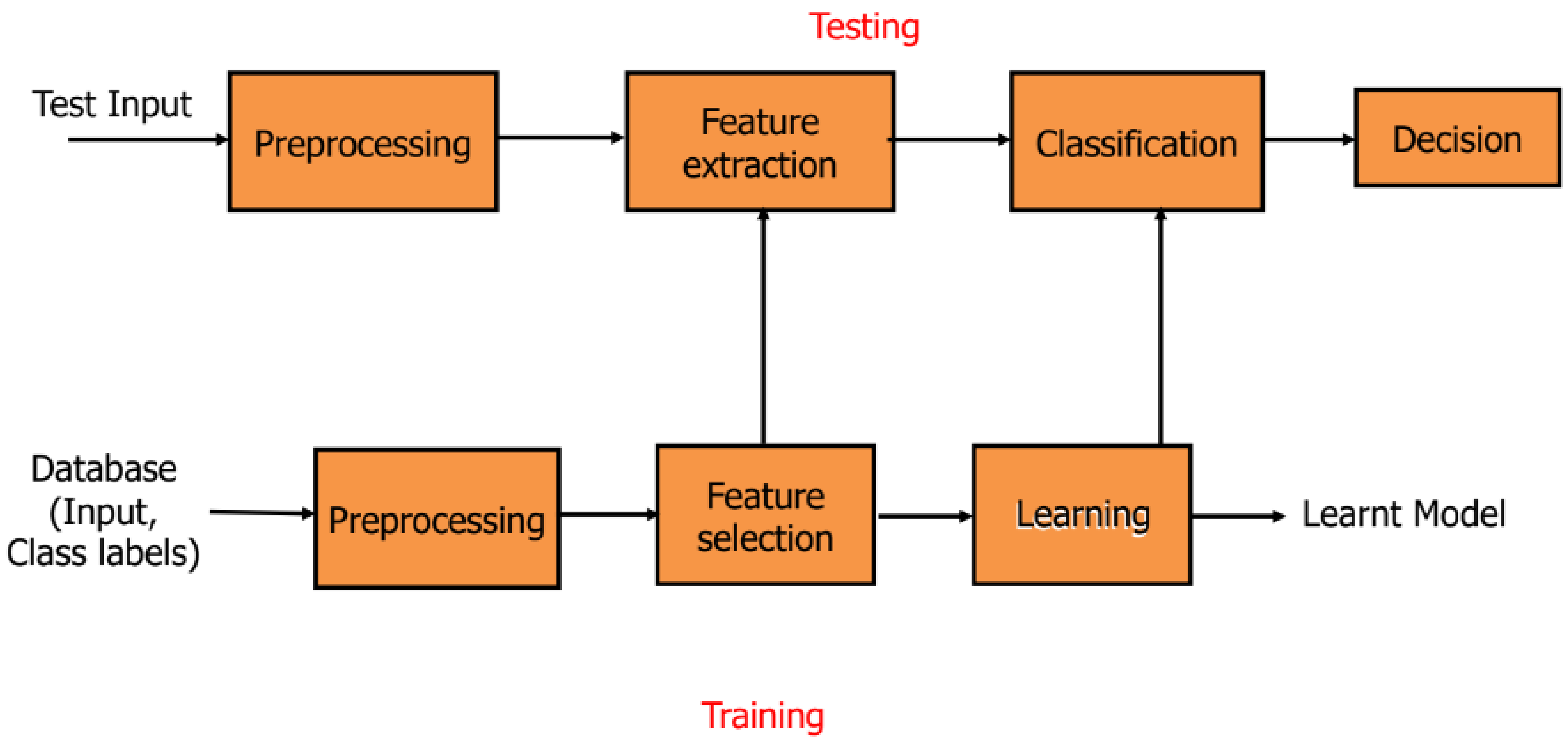
- Poor Data Quality
- Irrelevant Features
- Overfitting Training Data
- Underfitting the data
- Testing and Validating
- Hyperparameter Tuning and Model Selection
- Data Mismatch



Overfitting training



ML Pipeline


Testing
Test Input
Preprocessing
Feature extraction
Classification
Decision
Database (Input, Class labels)
Preprocessing
Feature selection
Learning
Learnt Model
Training