TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN

1.1 Introduction and Motivation

Machine Learning 1: Foundations

Marius Kloft *(TUK)*

21–28 Apr 2020

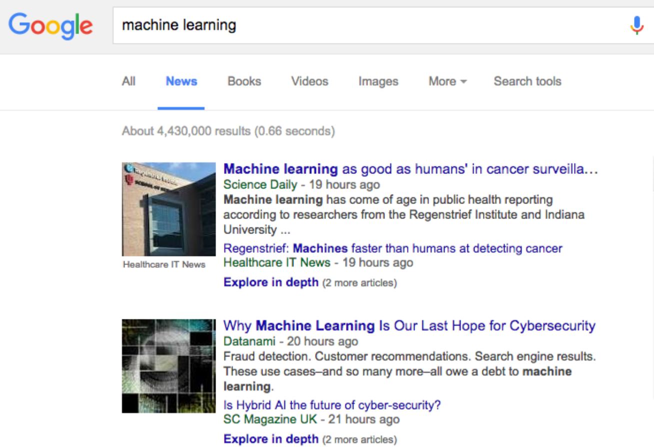
Contents

Machine Learning is Reshaping How We Live *

*You may not know it, but machine learning is all around you.**

* Domingos, Pedro (2015). The Master Algorithm: How Machine Learning is Reshaping How We Live. Pages xi-xiii.

Let's take a look on what Google News says about "Machine Learning"...


Google
machine learning
All
News
Books
Videos
Images
More
Search tools
About 4,430,000 results (0.66 seconds)
Machine learning as good as humans' in cancer surveilla...
Science Daily - 19 hours ago
Machine learning has come of age in public health reporting according to researchers from the Regenstrief Institute and Indiana University ...
Regenstrief: Machines faster than humans at detecting cancer
Healthcare IT News - 19 hours ago
Explore in depth (2 more articles)
Healthcare IT News
Why Machine Learning Is Our Last Hope for Cybersecurity
Datanami - 20 hours ago
Fraud detection. Customer recommendations. Search engine results. These use cases–and so many more–all owe a debt to machine learning.
Is Hybrid AI the future of cyber-security?
SC Magazine UK - 21 hours ago
Explore in depth (2 more articles)

ScienceDaily

Your source for the latest research news

Mobile | Follow | Subscribe

Breaking News: Zika Virus Structure: A Way to Target It

SD Health Tech Environment Society Quirky A- A A+ Search

Science News

from research organizations

Print | Email | Share

Machine learning as good as humans' in cancer surveillance, study shows

Date: April 21, 2016

Source: Indiana University

Summary: Machine learning has come of age in public health reporting. Researchers have found that existing algorithms and open source machine learning tools were as good as, or better than, human reviewers in detecting cancer cases using data from free-text pathology reports. The computerized approach was also faster and less resource intensive in comparison to human counterparts.

RELATED TOPICS

Health & Medicine

- Diseases and Conditions
- Health Policy
- Breast Cancer

Computers & Math

FULL STORY

Machine learning has come of age in public health reporting according to researchers from the Regenstrief Institute and Indiana University School of Informatics and Computing at Indiana University-Purdue University Indianapolis. They have found that existing algorithms and open source machine

Related Stories

Machine Learning Could Solve Riddles of Galaxy Formation

Nov. 11, 2015 — A new machine-learning simulation system promises cosmologists an expanded suite of galaxy models -- a necessary first step to developing more accurate and relevant insights into the formation of the ... **read more »**

'Machine Teaching' Holds the Power to Illuminate Human Learning

Aug. 11, 2015 — Human learning is a complex, sometimes mysterious process. Most of us have had experiences where we have struggled to learn something new, but also times when we've picked something up nearly ... **read more »**

Artificial Intelligence Improves Fine Wine Price Prediction

This app uses machine learning to guess who will die next in Game of Thrones

by ABHIMANYU GHOSHAL — 3 days ago in SHAREABLES

0 545 SHARES http://tnw.to/f507u

April 24 can't come soon enough for Game of Thrones fans eagerly awaiting the premiere of the hit show's sixth season. Naturally, most of us have been speculating wildly about the fate of our favorite characters for the past year, but now there's a clever app to help you withthat.

The project, A Song of Ice and Data, was developed by a group of students of a JavaScript course at the Technical University of Munich. It scrapes information from

GoT Problem Setting

- We are given the storyline of GoT until now

GoT Problem Setting

- We are given the storyline of GoT until now
- We want to predict the likelihood of death of characters in Game of Thrones

GoT Problem Setting

- We are given the storyline of GoT until now
- We want to predict the likelihood of death of characters in Game of Thrones

How would you solve this problem?

GoT Problem Setting

- We are given the storyline of GoT until now
- We want to predict the likelihood of death of characters in Game of Thrones

How would you solve this problem?

Please **pause** your video here and think about this question for a few minutes...

Predicting the Death of GoT Characters—A Solution

1. From data to numerical attributes

Predicting the Death of GoT Characters—A Solution

1. From data to numerical attributes
2. Feed attributes into the computer

Predicting the Death of GoT Characters—A Solution

1. From data to numerical attributes
2. Feed attributes into the computer
3. Let the computer predict the future

Predicting the Death of GoT Characters—A Solution

1. From data to numerical attributes
2. Feed attributes into the computer
3. Let the computer predict the future

Step 1: From Data to Numerical Attributes

Step 1: From Data to Numerical Attributes

Idea—define character **attributes** that (potentially) might be relevant for prediction of a character's death:

- Age
- House Stark? House Targaryen? ...
- Married?
- How popular in GOT wiki?
- $\vdots$

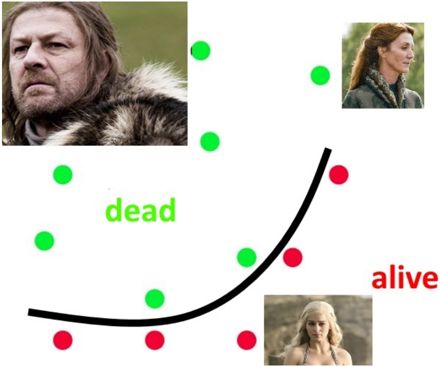
Each Character Is Now a Vector of Attributes:

$$\begin{array}{c} \textit{House Stark}? \\ \textit{Age} \\ \textit{Married} \\ \textit{Female} \\ \vdots \end{array} \qquad \begin{pmatrix} 1 \\ 36 \\ 1 \\ 0 \\ \vdots \end{pmatrix}$$

Color Encodes: Dead or Alive

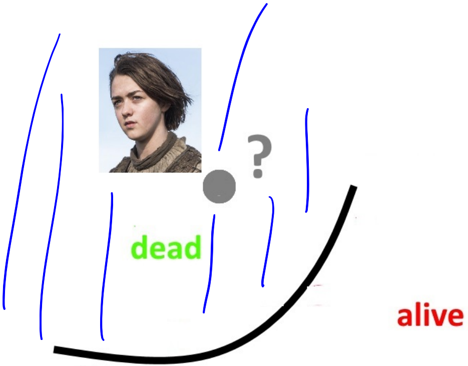
1. From data to numerical attributes
2. Feed attributes into the computer
3. Let the computer predict the future

Step 2: The Computer Now Computes a Smooth Separation of the Past Data



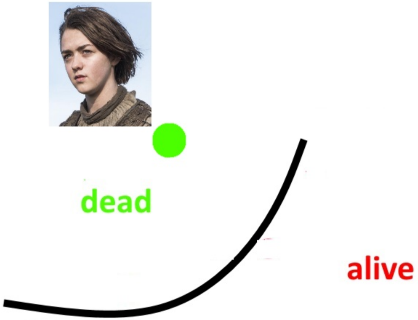
1. From data to numerical attributes
2. Feed attributes into the computer
3. Let the computer predict the future

Step 3: For a Character For Which We Want Predict the Likeliness of Death



Step 3: For a Character For Which We Want Predict the Likeliness of Death

...we use the separation computed by the computer



What is **learning**?

> "Learning is the act of acquiring new [...] knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information."

What is **learning**?

"Learning is the act of acquiring new [...] knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information."

"The ability to learn is possessed by humans, animals and some machines. [...] learning may be viewed as a process, rather than a collection of factual and procedural knowledge."

What is **learning**?

"Learning is the act of acquiring new [...] knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information."

"The ability to learn is possessed by humans, animals and some machines. [...] learning may be viewed as a process, rather than a collection of factual and procedural knowledge."

"Learning produces changes in the organism and the changes produced are relatively permanent."

What is **learning**?

> “Learning is the act of acquiring new [...] knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information.”

> “The ability to learn is possessed by humans, animals and some machines. [...] learning may be viewed as a process, rather than a collection of factual and procedural knowledge.”

> “Learning produces changes in the organism and the changes produced are relatively permanent.”

— http://en.wikipedia.org/wiki/Learning, accessed July 25, 2014

What is **learning**?

> "Learning is the act of acquiring new [...] knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information."

> "The ability to learn is possessed by humans, animals and some **machines**. [...] learning may be viewed as a process, rather than a collection of factual and procedural knowledge."

> "Learning produces changes in the organism and the changes produced are relatively permanent."

— http://en.wikipedia.org/wiki/Learning, accessed July 25, 2014

Origins of **Machine** Learning

Origins of **Machine** Learning

Checkers

- strategy board game (8×8 or 10×10 board)

Origins of **Machine** Learning

Checkers

- strategy board game (8×8 or 10×10 board)

Computer Checkers (1952)

- historically one of the earliest computer games

Origins of **Machine** Learning

Checkers

- strategy board game (8×8 or 10×10 board)

Computer Checkers (1952)

- historically one of the earliest computer games
- unique feature: **intelligent** adaptation to the opponent's strategy

Origins of **Machine** Learning

Checkers

- strategy board game (8×8 or 10×10 board)

Computer Checkers (1952)

- historically one of the earliest computer games
- unique feature: **intelligent** adaptation to the opponent's strategy
- developed by **Arthur Samuel**

What is Machine Learning?

What is Machine Learning?

> "Field of study that gives computers the ability to learn without being explicitly programmed"

— Arthur Samuel (1959)

What is Machine Learning?

"Field of study that gives computers the ability to learn without being explicitly programmed"

— Arthur Samuel (1959)

"Computational methods using experience to make accurate predictions"

— Mehryar Mohri et al. (2012)

What is Machine Learning?

"Field of study that gives computers the ability to learn without being explicitly programmed"

— Arthur Samuel (1959)

"Computational methods using experience to make accurate predictions"

— Mehryar Mohri et al. (2012)

"Set of techniques that allow a computer to acquire or improve its ability to perform a task by automatically extracting knowledge from data."

— Yann LeCun (recently)

Examples of Machine Learning Problems

Robot learning

- Data: sensor data gained by robots roaming their environments
- Goal: robots learning to better navigate

Examples of Machine Learning Problems (2)

Data mining[1] of electronic health records

- Data: electronic health records of patients
- Goal: learn to predict which therapies work best for which diseases

[1] data mining = discovering patterns in large data sets

Examples of Machine Learning Problems (3)

Speech Recognition

- Data: annotated recordings of speech (smartphone, customer service hotlines, ...)
- Goal: better understand your speech based on experience listening to you
- State of the art: Dragon NaturallySpeaking

Examples of Machine Learning Problems (3)

Speech Recognition

- Data: annotated recordings of speech (smartphone, customer service hotlines, ...)
- Goal: better understand your speech based on experience listening to you
- State of the art: Dragon NaturallySpeaking

Optical Character Recognition

- Data: human-annotated handwritten digits or letters
- Goal: better recognize your handwriting

Examples of Machine Learning Problems (4)

Visual Image Recognition

- Data: annotated images
- Goal: to annotate yet unannotated images
  - content-based image retrieval (Google image search)
  - sort your image collection according to topics (Beach, Uni, Skateboard, ...)
  - recognize faces in your images to link them with your friends' facebook pages

rw

SOCIAL

Facial Recognition Comes to Facebook

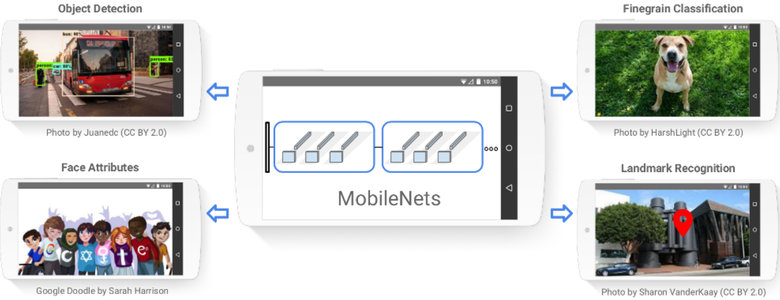
This morning, Face.com announced that they're bringing advanced facial recognition technology to Facebook by way of a new application called Photo Finder. Using proprietary facial scanning algorithms, this application scans through your photos and those public photos belonging to your friends in order to identify and suggest tags for the untagged...

Sarah Perez *on* March 24, 2009

Examples of Machine Learning Problems (5)

Low-Resource Object Detection

- Data: Images
- Goal: Identify objects within the images at low cost, e.g. on mobile devices
- State of the Art: MobileNets, Google, 2017



Examples of Machine Learning Problems (6)

Malware Detection

- Data: source code of computer programs
- Goal: detection of malicious code

Examples of Machine Learning Problems (6)

Malware Detection

- Data: source code of computer programs
- Goal: detection of malicious code

Network Security

- Data: annotated HTTP requests
- Goal: ~~image annotation~~ detect attacks

Examples of Machine Learning Problems (7)

Genome annotation

- Data: DNA sequences (strings)
- Goal: detection of interesting genetic markers (“genes”)

```
GGGTTTAGTT CTTTGAGAGT CACATCTCTT ATTTGGACCA GTATAGACAG
AAGTAAACCC ACCTGACTTG TTTCCTGGGA CAGTTGAGTT AAGGGATGGC
TTTCACAGAG CATTCACCGC TGACCCTTCA CCGTCGGGAC CTCTGTAGCC
GCTCTATCTG GCTAGCAAGG AAGATTCGTT CAGACCTGAC TGCTCTTACG
GAATCCTATG TAAGTTGCCT ATTTTGCTGT TATCTGAAAA CCCTTCATXX
XXXXXXXXXX XXCATGGGTA TGACAGAAGA TGTGGTGTTT TCCTGTATCC
TCGGCGAGGT GAAGCATCAG GGCCTGAACA AGAACATCAA CCTGGACTCT
GCGGATGGGA TGCCAGTGGC AAGCACTGAT CAGTGGAGTG AGCTGACCGA
GGCAGAGCCA CTCCAAGAGA ACCTTCAAGC TTATCGTACC TTCCATGTTT
TGTTGGCCAG GCTCTTAGAA GACCAGCAGG TGCATTTTAC CCCAACCGAA
GGTGACTTCC ATCAAGCTAT ACATACCCTT CTTCTCCAAG TCGCTGCCTT
TGCATACCAG ATAGAGGAGT TAATGATACT CCTGGAATAC AAGATCCCCG
CCAATGAGGC TGATGGGATG CCTATTAATG TTGGAGATGG TGGTCTCTTT
GAGAAGAAGC TGTGGGGCCT AAAGGTGCTG CAGGAGCTTT CACAGTGGAC
AGTAAGGTCC ATCCATGACC TTCGTTTCAT TTCTTCTCAT CAGACTGGGA
TCCCAGCACG TGGGAGCCAT TATATTGCTA ACAACAAGAA AATGTAGCAG
TTAGTCCCTT CTCTCTTCCT TGCTTTCTCT TCTAATGGAA TATGGGTAG
```

Examples of Machine Learning Problems (8)

Natural Language Translation

- Data: (Digital) Text
- Goal: Translate texts from one language to another
- State of the Art: Google Neural Machine Translation, Google, 2017

Input sentence:	***Translation (PBMT):***	***Translation (GNMT):***	***Translation (human):***
李克強此行將啟動中加總理年度對話機制，與加拿大總理杜魯多舉行兩國總理首次年度對話。	Li Keqiang premier added this line to start the annual dialogue mechanism with the Canadian Prime Minister Trudeau two prime ministers held its first annual session.	Li Keqiang will start the annual dialogue mechanism with Prime Minister Trudeau of Canada and hold the first annual dialogue between the two premiers.	Li Keqiang will initiate the annual dialogue mechanism between premiers of China and Canada during this visit, and hold the first annual dialogue with Premier Trudeau of Canada.

Examples of Machine Learning Problems (9)

Image Captioning

- Data: Images
- Goal: Annotate photos with descriptive texts
- State of the Art: e.g. DenseCap, Johnson, Karpathy, Fei-Fei (Stanford University), 2016

... and Many More Examples!

- Protein function prediction (drug design, pharmaceutical industry)
- Learn to rank search queries (Google and friends)
- Learn to recommend you the books you love to read based on your shopping history (Amazon and friends)
- Text and document classification (advertising, wikipedia, internet, ...)
- Disease risk prediction
- Question answering
- Automatic code completion
- etc. (you name it!)

Machine Learning has a strong impact on all kinds of applications!

What can Machine Learning be for you?

What can Machine Learning be for you?

Core tasks in ML:

Theory $\rightarrow$ Algorithms $\rightarrow$ Applications

What can Machine Learning be for you?

Core tasks in ML:

Theory $\rightarrow$ Algorithms $\rightarrow$ Applications

- Theory: analyze learning algorithms using techniques from probability and statistical learning theory, in order to understand them better

What can Machine Learning be for you?

Core tasks in ML:

Theory $\rightarrow$ Algorithms $\rightarrow$ Applications

- Theory: analyze learning algorithms using techniques from probability and statistical learning theory, in order to understand them better
- New algorithms: design of algorithms that learn faster or more accurately

What can Machine Learning be for you?

Core tasks in ML:

Theory $\rightarrow$ Algorithms $\rightarrow$ Applications

- Theory: analyze learning algorithms using techniques from probability and statistical learning theory, in order to understand them better
- New algorithms: design of algorithms that learn faster or more accurately
- Applications: get learning algorithms working in applications

What can Machine Learning be for you?

Core tasks in ML:

Theory $\rightarrow$ Algorithms $\rightarrow$ Applications

- Theory: analyze learning algorithms using techniques from probability and statistical learning theory, in order to understand them better
- New algorithms: design of algorithms that learn faster or more accurately
- Applications: get learning algorithms working in applications

Machine learning is interdisciplinary (algorithms, AI, statistics, numerical mathematics, all kinds of applications, etc.

Basic Terminology

Data = Inputs + Labels

The data consists of **inputs** and **labels**:

- Inputs = the raw data instances $x_1, \ldots, x_n$ (e.g., source code of computer programs)
- Labels = their annotations $y_1, \ldots, y_n$ (malware: yes or no?!)

Basic Terminology

Data = Inputs + Labels

The data consists of **inputs** and **labels**:

- Inputs = the raw data instances $x_1, \ldots, x_n$ (e.g., source code of computer programs)
- Labels = their annotations $y_1, \ldots, y_n$ (malware: yes or no?!)

What are inputs and labels in the previous examples?

Basic Terminology

Data = Inputs + Labels

The data consists of **inputs** and **labels**:

- Inputs = the raw data instances $x_1, \ldots, x_n$ (e.g., source code of computer programs)
- Labels = their annotations $y_1, \ldots, y_n$ (malware: yes or no?!)

What are inputs and labels in the previous examples?

Please **pause** your video here and think about this question for a few minutes...

Training and prediction

Training

- use all data (inputs and labels) to train the computer

Prediction

- use trained computer to predict the (unknown) labels for new inputs

Formal Problem Setting

- Let $\mathcal{X}$ and $\mathcal{Y}$ be some sets (called *input space* and *label space*, respectively).
- Training data = $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\} \subset \mathcal{X} \times \mathcal{Y}$.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are called *inputs* and $y_1, \ldots, y_n$ are called *labels*.

Formal Problem Setting

- Let $\mathcal{X}$ and $\mathcal{Y}$ be some sets (called *input space* and *label space*, respectively).
- Training data = $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\} \subset \mathcal{X} \times \mathcal{Y}$.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are called *inputs* and $y_1, \ldots, y_n$ are called *labels*.
- Goal of machine learning: write a computer program that learns from the training data a function $f : \mathcal{X} \to \mathcal{Y}$ that accurately predicts on future (yet unseen) data ("test data")

Formal Problem Setting

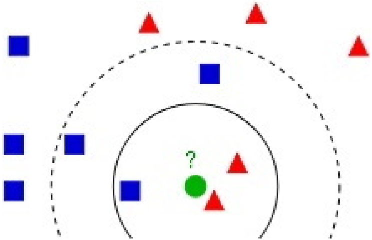
- Let $\mathcal{X}$ and $\mathcal{Y}$ be some sets (called *input space* and *label space*, respectively).
- Training data = $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\} \subset \mathcal{X} \times \mathcal{Y}$.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are called *inputs* and $y_1, \ldots, y_n$ are called *labels*.
- Goal of machine learning: write a computer program that learns from the training data a function $f : \mathcal{X} \to \mathcal{Y}$ that accurately predicts on future (yet unseen) data ("test data")
- f is called **prediction function**
  - or simply **classifier**, if $\mathcal{Y}$ is a finite set
  - the elements in $\mathcal{Y}$ are then called *classes*
- The computer program is synonymously also called **learning machine** or **learning algorithm**

Formal Problem Setting

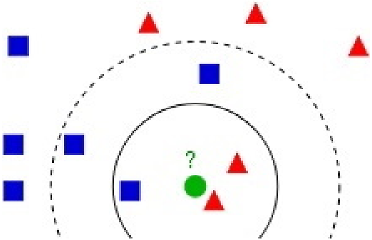
- Let $\mathcal{X}$ and $\mathcal{Y}$ be some sets (called *input space* and *label space*, respectively).
- Training data = $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\} \subset \mathcal{X} \times \mathcal{Y}$.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are called *inputs* and $y_1, \ldots, y_n$ are called *labels*.
- Goal of machine learning: write a computer program that learns from the training data a function $f : \mathcal{X} \rightarrow \mathcal{Y}$ that accurately predicts on future (yet unseen) data ("test data")
- f is called **prediction function**
  - or simply **classifier**, if $\mathcal{Y}$ is a finite set
  - the elements in $\mathcal{Y}$ are then called *classes*
- The computer program is synonymously also called **learning machine** or **learning algorithm**

Unless stated otherwise we will assume $\mathcal{X} = \mathbb{R}^d$ and $\mathcal{Y} = \{-1, 1\}$ (**binary classification**).

k-nearest neighbor learning algorithm



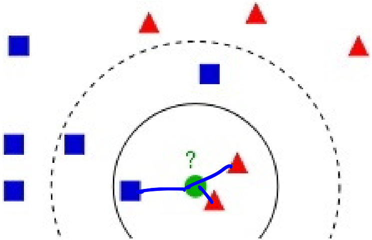
k-nearest neighbor learning algorithm



No training step

[Optional: store data in efficient data structure.]

k-nearest neighbor learning algorithm



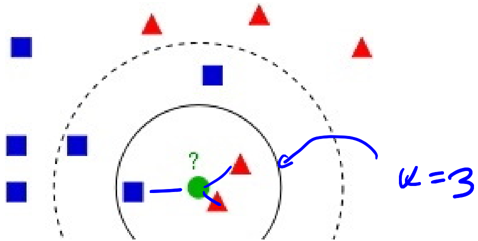
No training step

[Optional: store data in efficient data structure.]

Prediction

- Given a new input, compute the distances to the training inputs

k-nearest neighbor learning algorithm



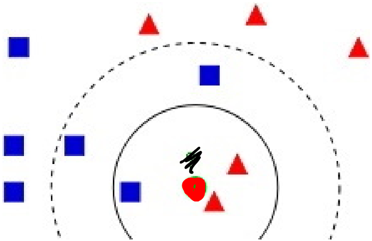
No training step

[Optional: store data in efficient data structure.]

Prediction

- Given a new input, compute the distances to the training inputs
- Find the k training inputs with the smallest distances

k-nearest neighbor learning algorithm



No training step

[Optional: store data in efficient data structure.]

Prediction

- Given a new input, compute the distances to the training inputs
- Find the k training inputs with the smallest distances
- The label of the new input is obtained by majority vote over the labels of the k training inputs determined in the previous step

How can we formally capture whether or not the computer predicts well?

- Let f be the classifier output after training the computer using the training data
- For a new input x with label y the classifier f errs when $f(x) \neq y$

How can we formally capture whether or not the computer predicts well?

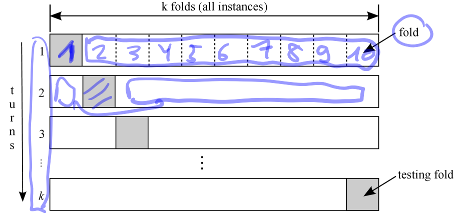
- Let f be the classifier output after training the computer using the training data
- For a new input x with label y the classifier f errs when $f(x) \neq y$
- The error probability $P[f(x) \neq y]$ measures the quality of the classifier f

How can we formally capture whether or not the computer predicts well?

- Let f be the classifier output after training the computer using the training data
- For a new input x with label y the classifier f errs when $f(x) \neq y$
- The error probability $P[f(x) \neq y]$ measures the quality of the classifier f

We would like to design computer programs that compute f with error probability $P[f(x) \neq y]$ as small as possible!

Evaluation: estimator for error probability



t-times k-fold cross validation (CV)

function CV(t, k, training_data)
    for $i = 1 : t$ **do**
        Randomly split the samples into k sets of the same size ("folds")
        for $j = 1 : k$ **do**
            Use jth fold as **test set** and union of all remaining folds as **training set**
            Train classifier on training set and predict on test set
        end for
    end for
    return average classification accuracy and standard deviation (over the $k \cdot t$ many runs)
end function

Conclusion

- Learning from Experience
- Machine Learning = making computers learn from data (usually to make accurate predictions of the future)
- Example: k-nearest neighbor algorithm
- Measure quality of algorithms using cross validation
- ML: Strong impact on all kinds of applications
- Many opportunities in this interdisciplinary field (algorithms, theory, applications, ...)

Machine learning plays a key role in technology

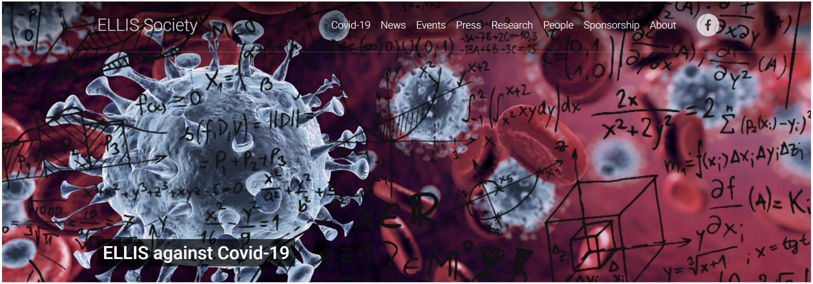
Conclusion

- Learning from Experience
- Machine Learning = making computers learn from data (usually to make accurate predictions of the future)
- Example: k-nearest neighbor algorithm
- Measure quality of algorithms using cross validation
- ML: Strong impact on all kinds of applications
- Many opportunities in this interdisciplinary field (algorithms, theory, applications, ...)

Machine learning plays a key role in technology

Suggested reading:

Duda, Hart, and Stork: Pattern Classification, Chapter 1.
`https://kplus.ub.uni-kl.de/Record/KLU01-000924363`.



Overview

The world is facing a public health emergency caused by the Covid-19 pandemic. We all need to take this on, and science can make a major contribution. This online workshop will present projects on how to tackle Covid-19 using methods of machine learning and AI, carried out by leading international researchers.

Research topics include outbreak prediction, epidemiological modelling, drug development, viral and host genome sequencing, and health care management.

`http://ellis.eu/covid-19`

Questions?

Refs I

M. Mohri, A. Rostamizadeh, and A. Talwalkar, Foundations of machine learning. MIT press, 2012.