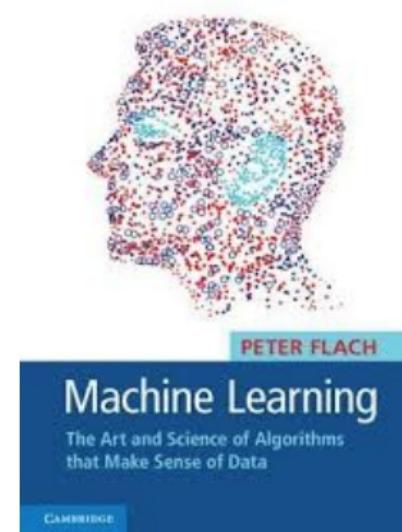
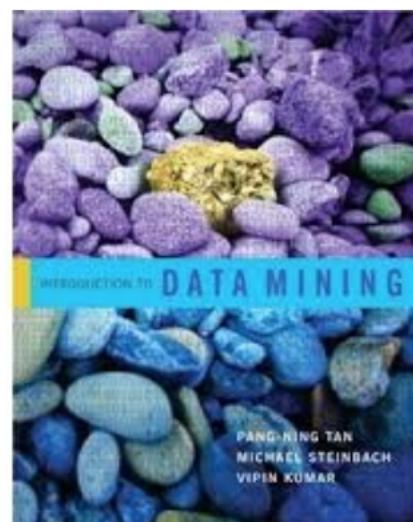
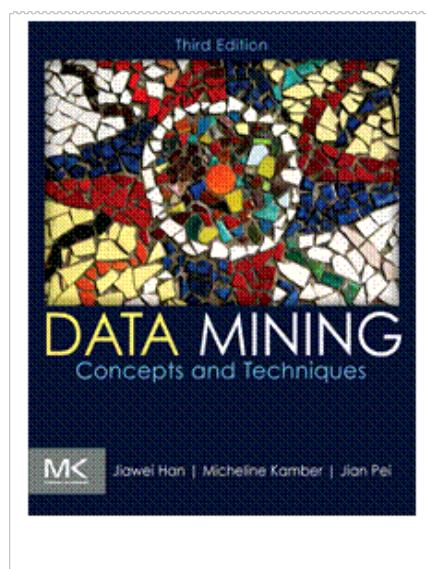


Aprendizado de Máquinas

Ronnie Alves (alvesrco@gmail.com)

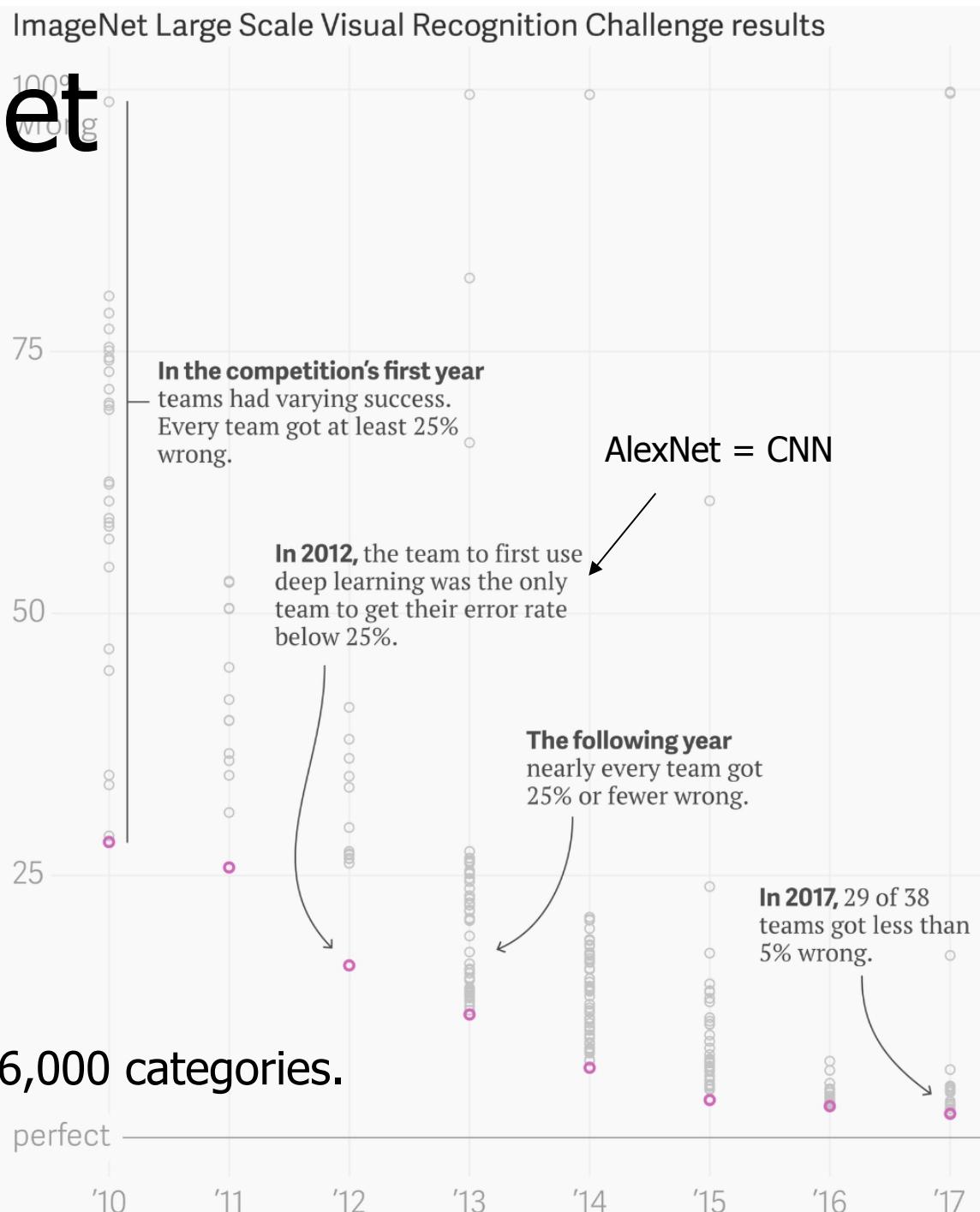
<https://sites.google.com/site/alvesrco/>



The ImageNet Challenge

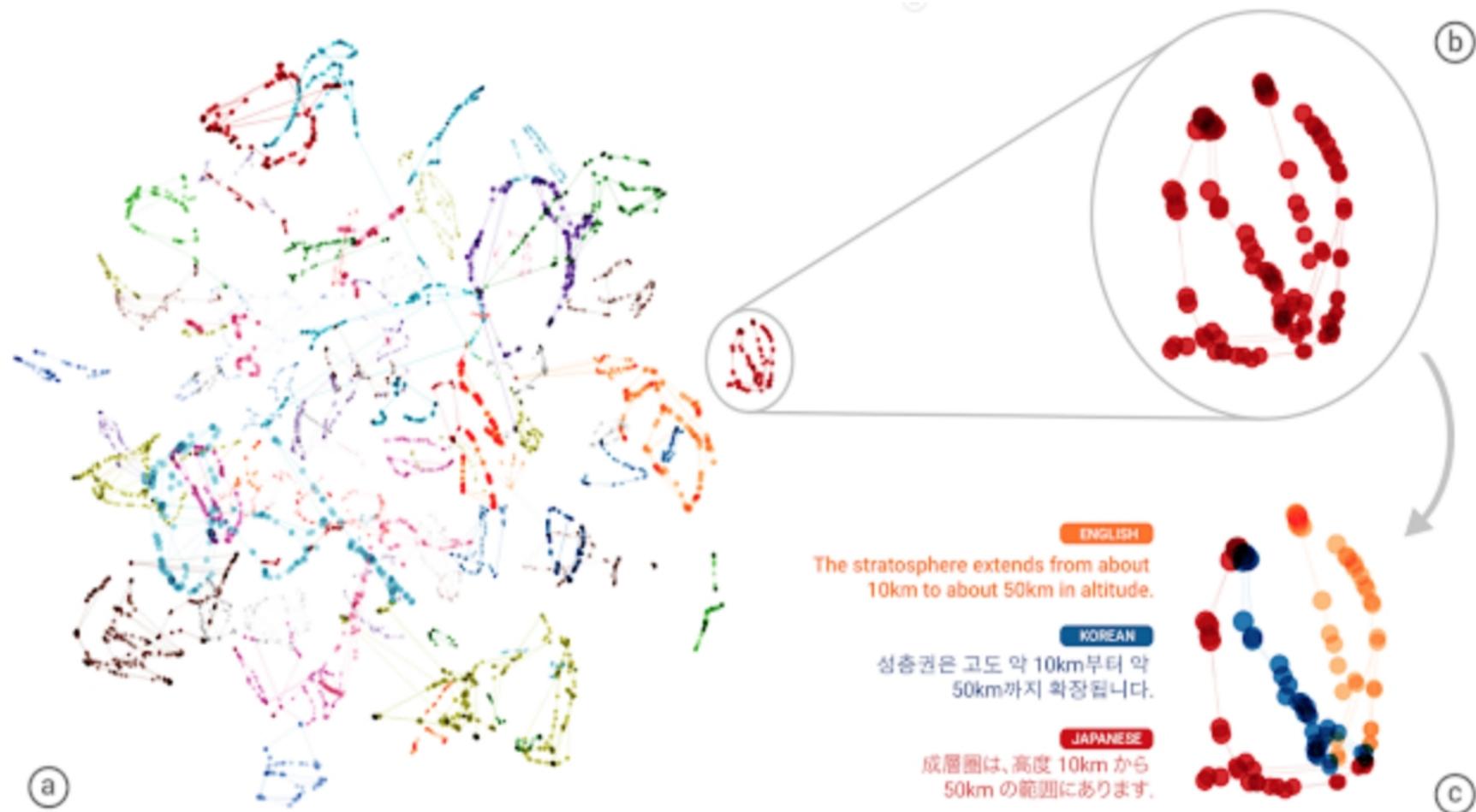
IT'S NOT ABOUT
THE ALGORITHM

The data that transformed
AI research—and possibly
the world



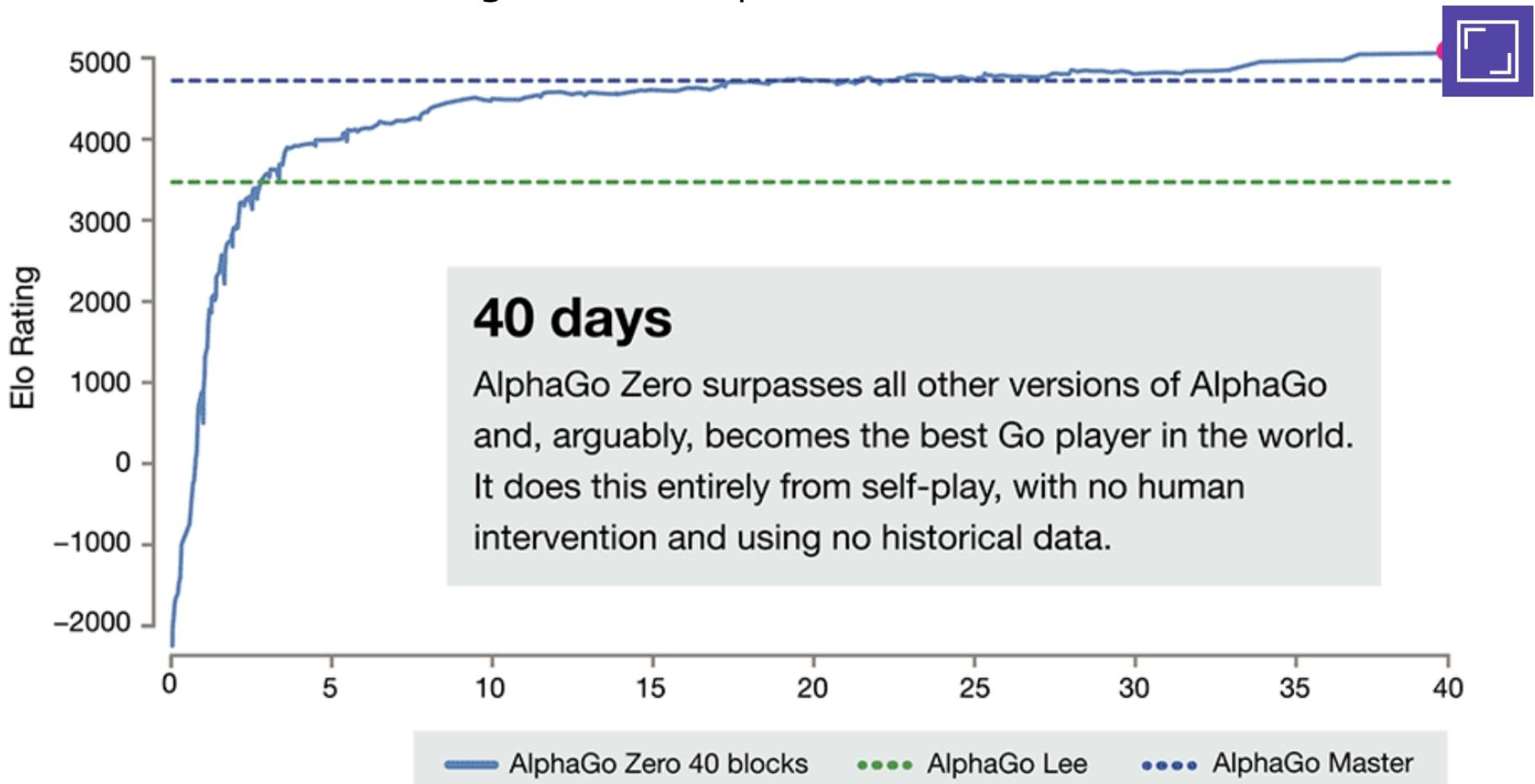
In 2016, Google released the
Open Images database,
containing 9 million images in 6,000 categories.

Zero-Shot Translation with Google's Multilingual Neural Machine Translation System / September 2016

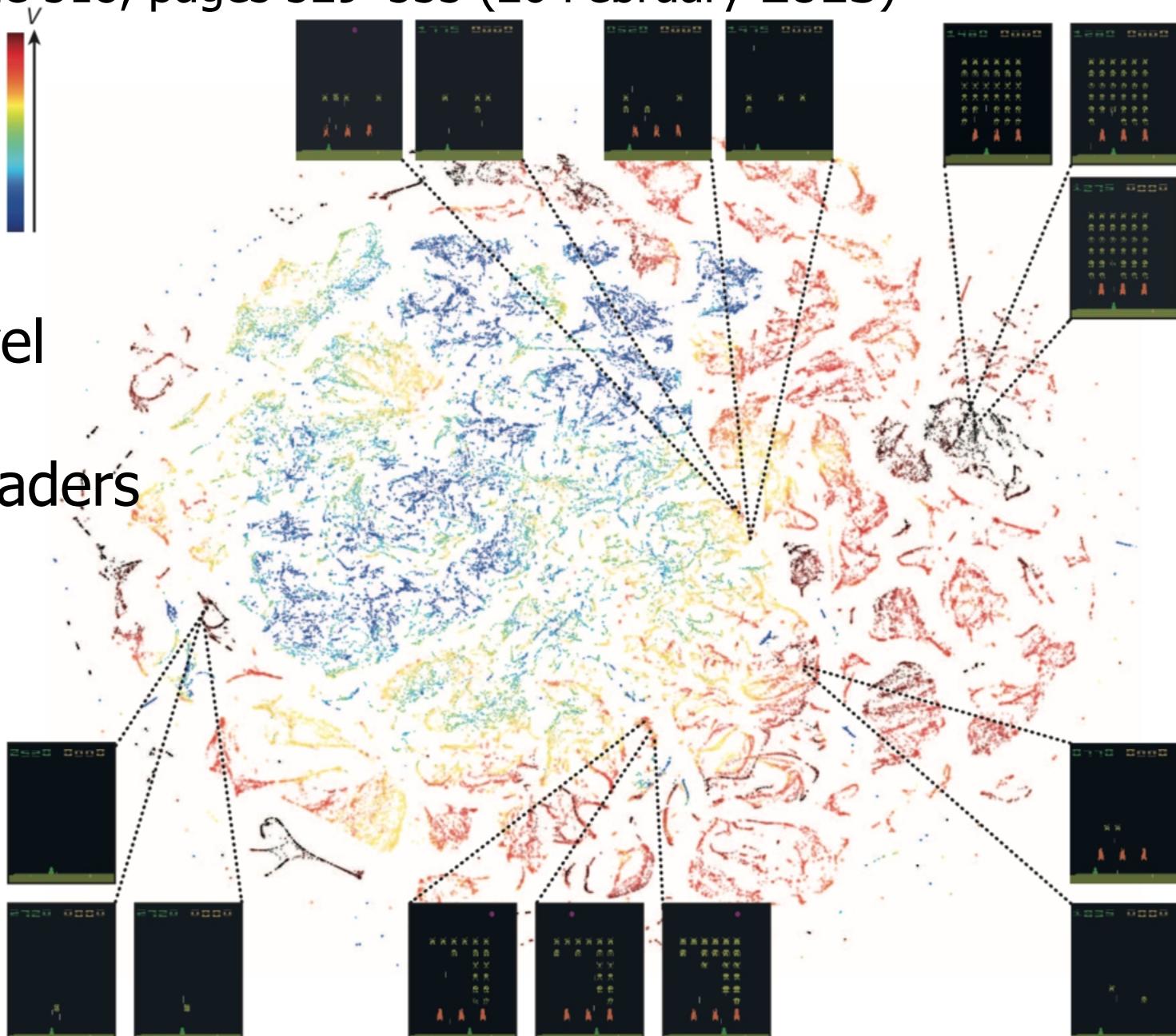


AlphaGo Zero Shows Machines Can Become Superhuman Without Any Help

Both AlphaGo and AlphaGo Zero use a machine-learning approach known as **reinforcement learning** as well as deep neural networks.



Human-level
Gaming
-Space Invaders



DeepMind's AI, AlphaStar Showcases Significant Progress Towards AGI

DeepMind's AI, AlphaStar sweeps StarCraft II players in head-to-head match displaying the heart of modern machine learning.



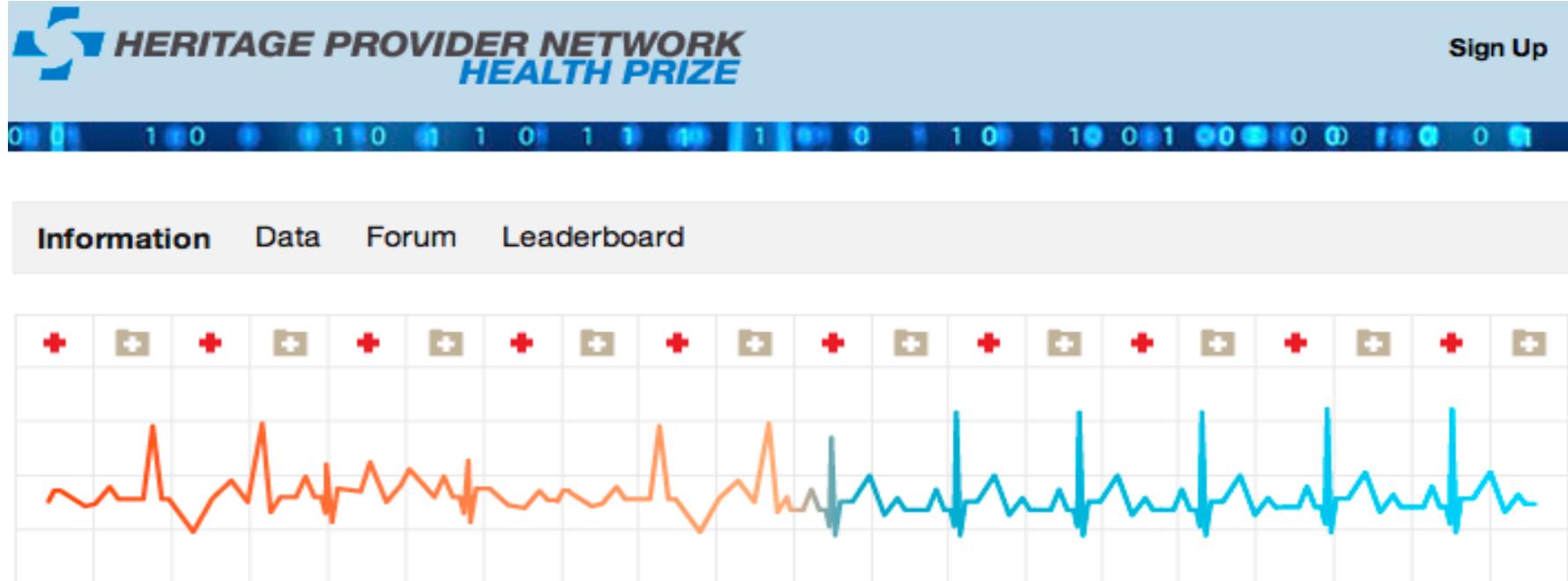
Stacy Stanford [Follow](#)

Jan 25 · 7 min read

January 25, 2019 by Stacy Stanford



Big Prize!



**Improve Healthcare,
Win \$3,000,000.**

Identify patients who will be admitted to a hospital within the next year, using historical claims data.

<http://www.kaggle.com/>

Big Prize!



[Home](#) [About](#) [Teams](#) [FAQs](#) [Contest Rules](#) [2017](#)

The Alexa Prize

Over \$3.5 Million to Advance Conversational Artificial Intelligence

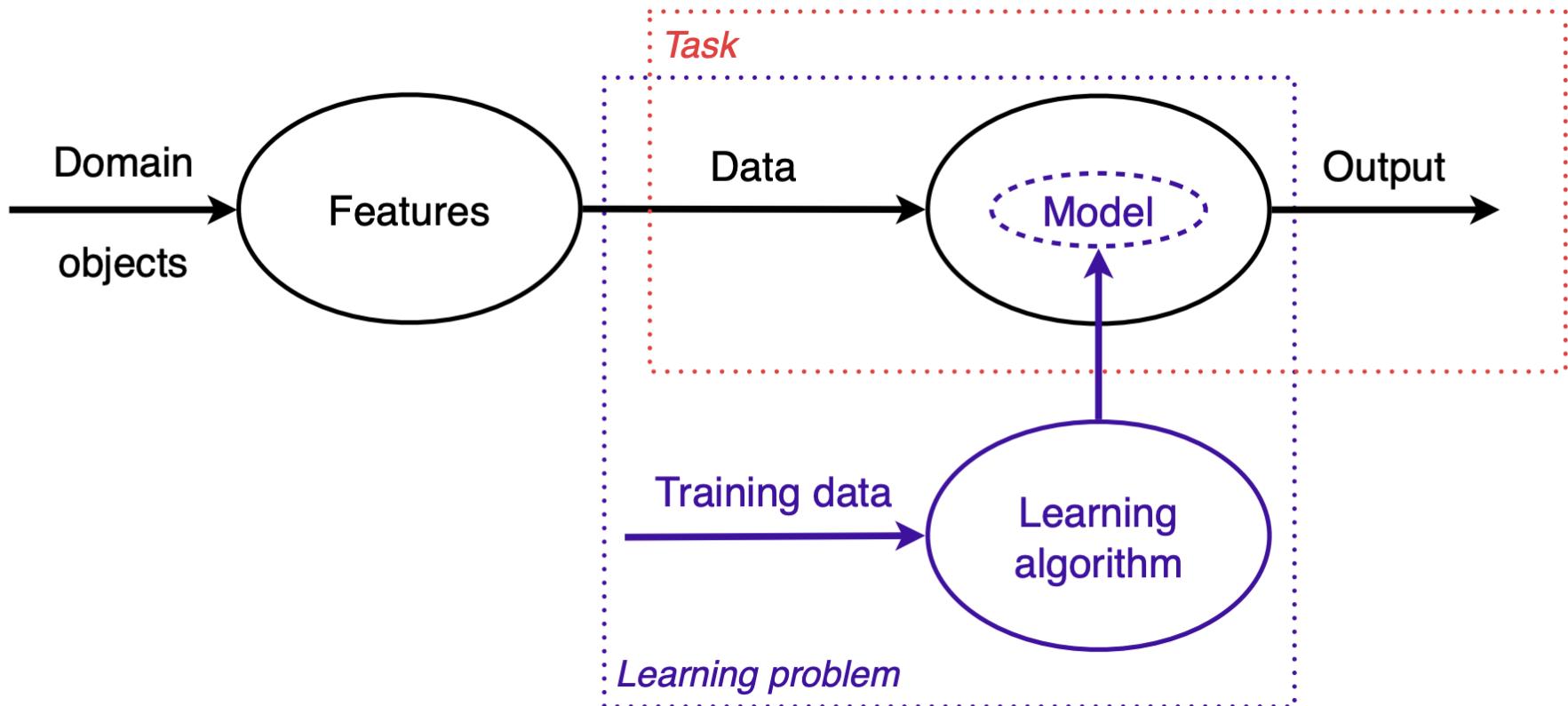
December 2017 - November 2018



It's Day One for 2018 Alexa Prize Participants

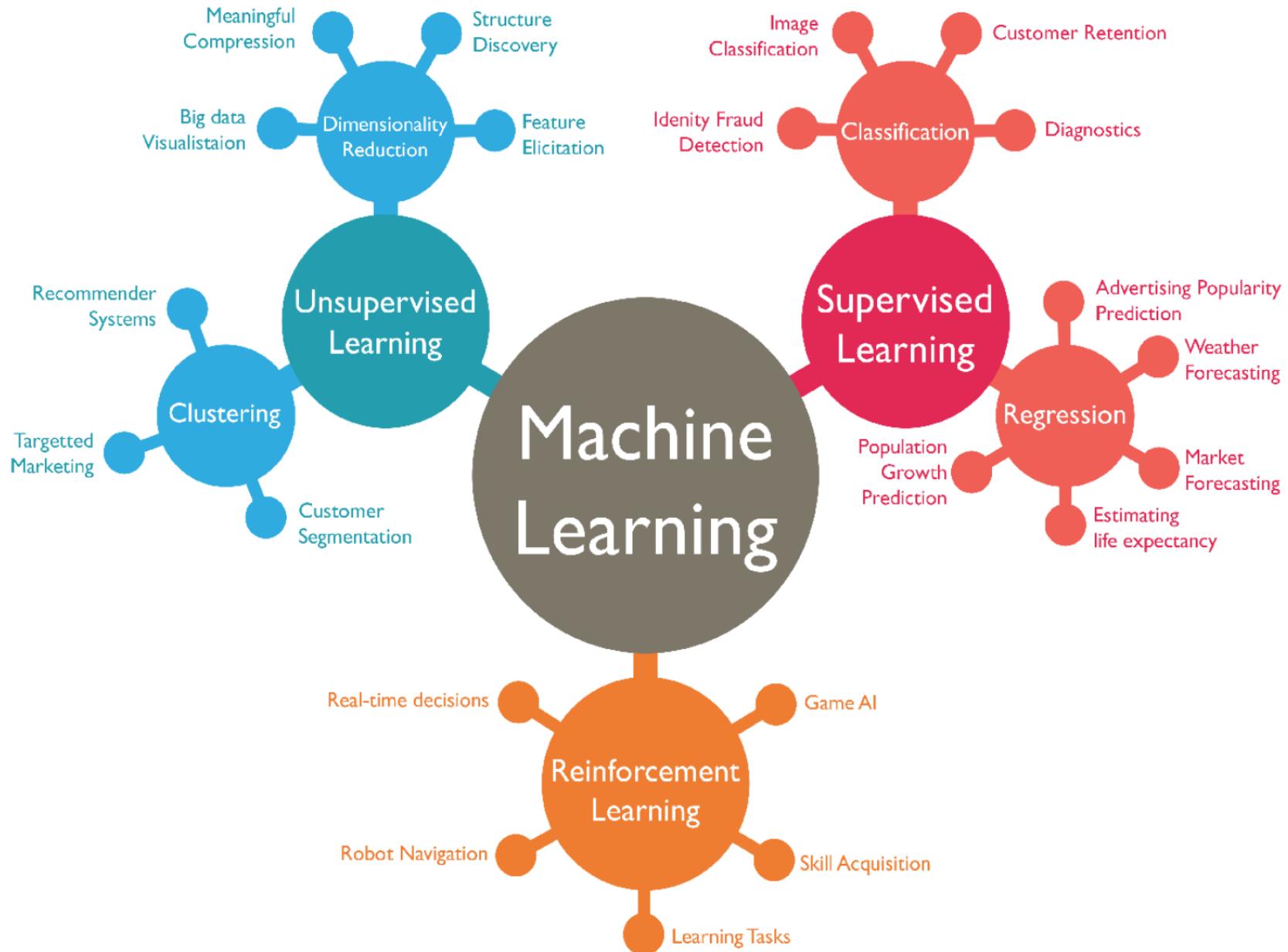
Amazon announces the 2018 teams selected to compete for the Alexa Prize, a \$3.5 million university challenge to advance human-computer interaction.

How machine learning helps to solve a task



Tasks are addressed by **models**, whereas learning problems are solved by **learning algorithms** that produce **models**.

So many tasks, so many learning algorithms



*Machine learning is concerned
with using the **right features**
to build the **right models** that
achieve the **right tasks.***

Peter Flach

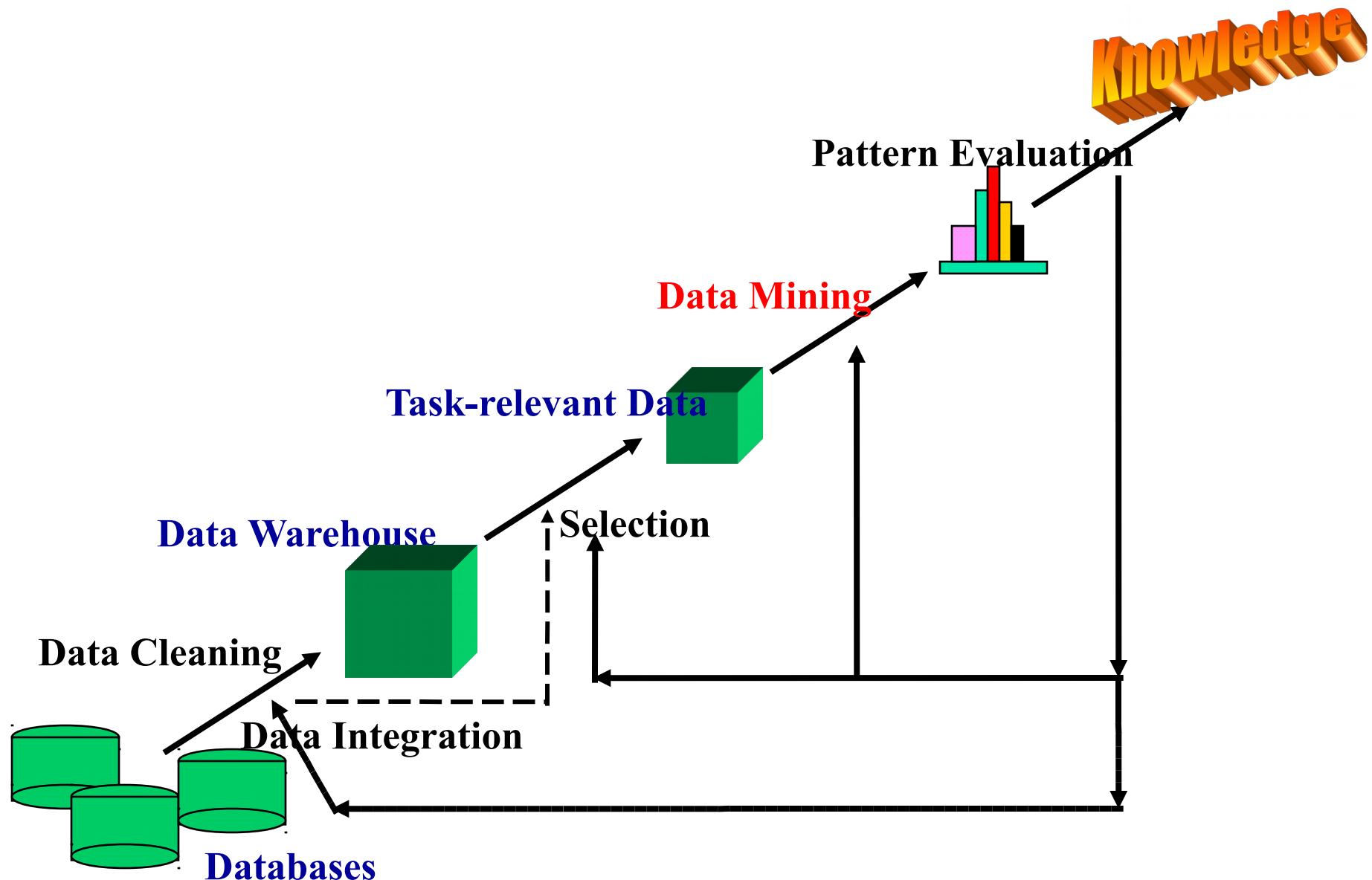
LEARNING = REPRESENTATION + EVALUATION + OPTIMIZATION

Table 1: The three components of learning algorithms.

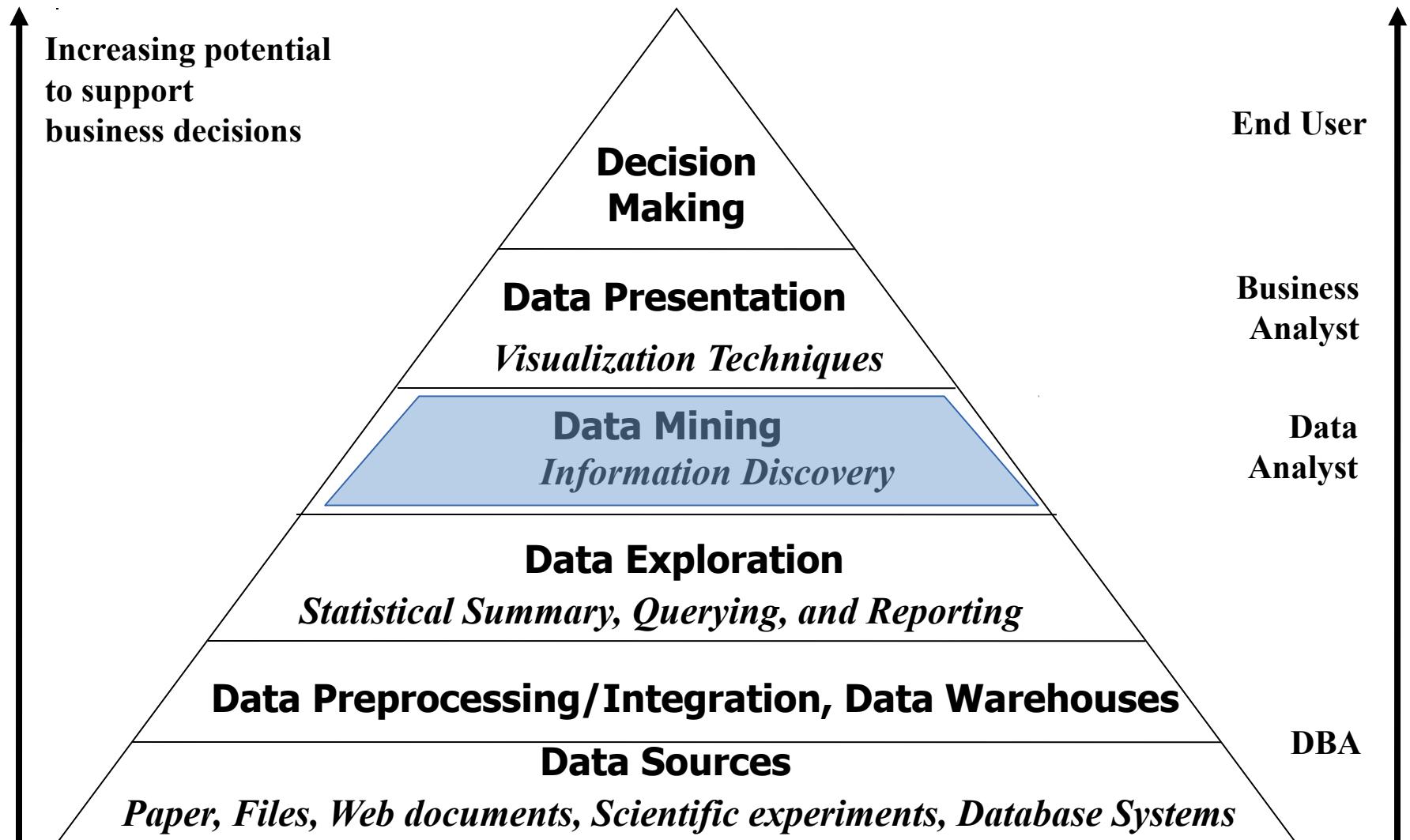
| Representation | Evaluation | Optimization |
|---|---|--|
| Instances <i>K</i> -nearest neighbor Support vector machines | Accuracy/Error rate Precision and recall Squared error Likelihood Posterior probability Information gain | Combinatorial optimization Greedy search Beam search Branch-and-bound Continuous optimization Unconstrained |
| Hyperplanes Naive Bayes Logistic regression | K-L divergence Cost/Utility Margin | Gradient descent Conjugate gradient Quasi-Newton methods |
| Decision trees Sets of rules Propositional rules Logic programs | | Constrained Linear programming Quadratic programming |
| Neural networks Graphical models Bayesian networks Conditional random fields | | |

A Few Useful Things to Know about Machine Learning
by Prof. Pedro Domingos

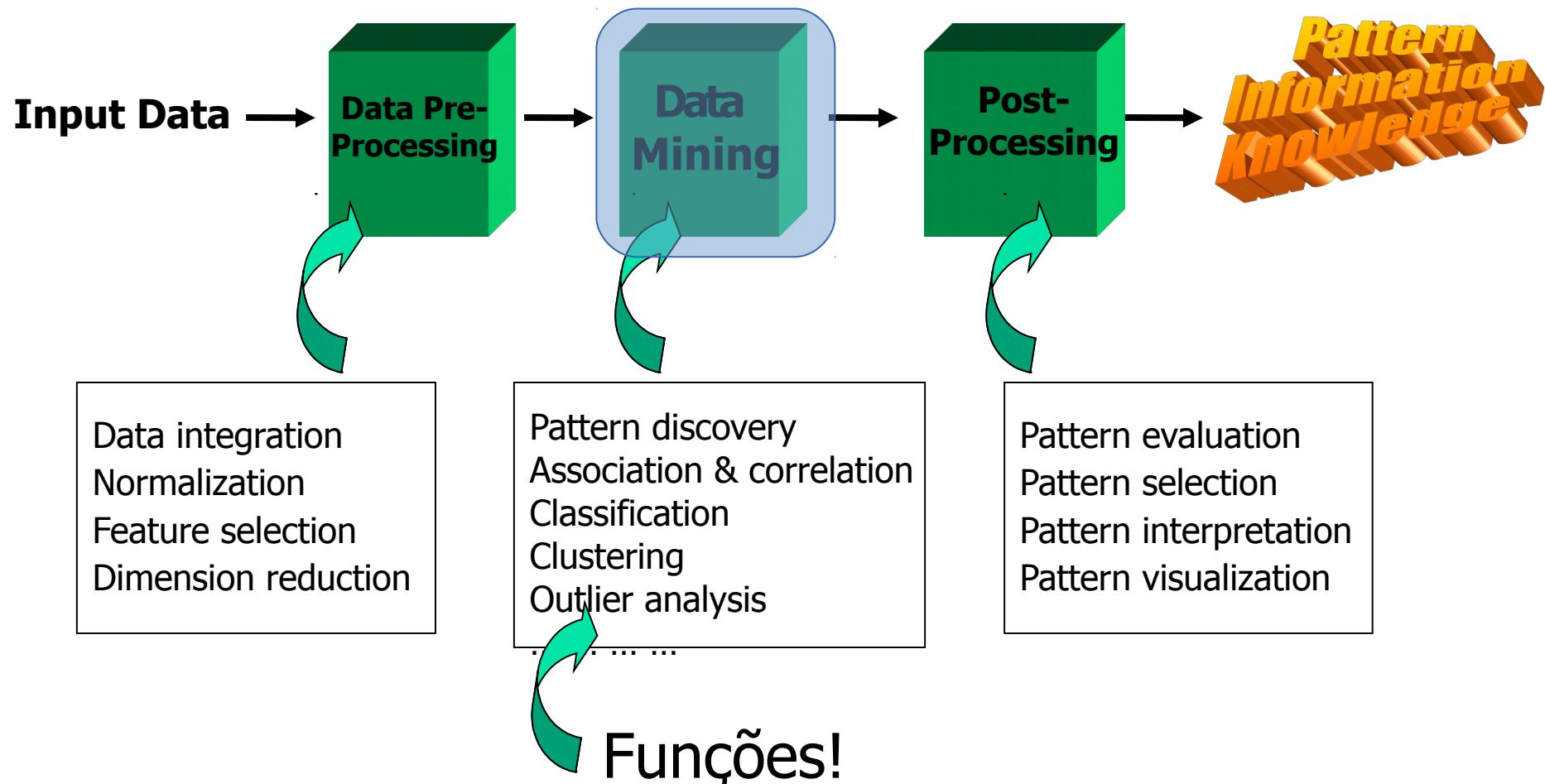
Processo KDD: Visão Banco de dados



Processo KDD: Visão Business Intel.



Processo KDD: Visão Aprendizado de Máquinas

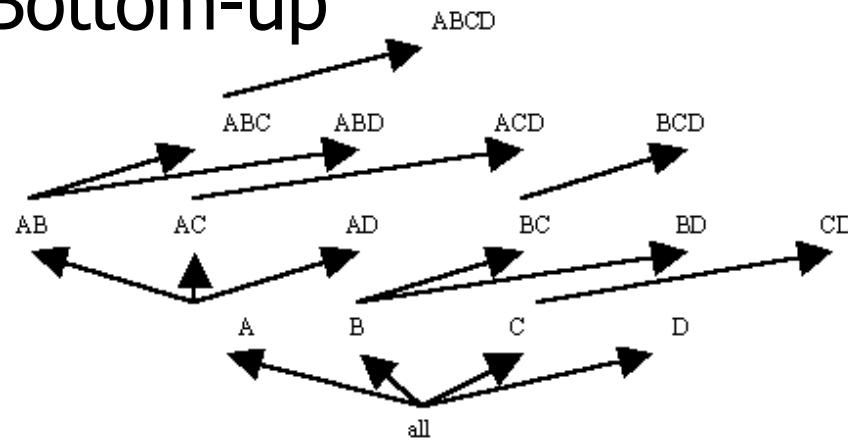


Função: (1) Generalização

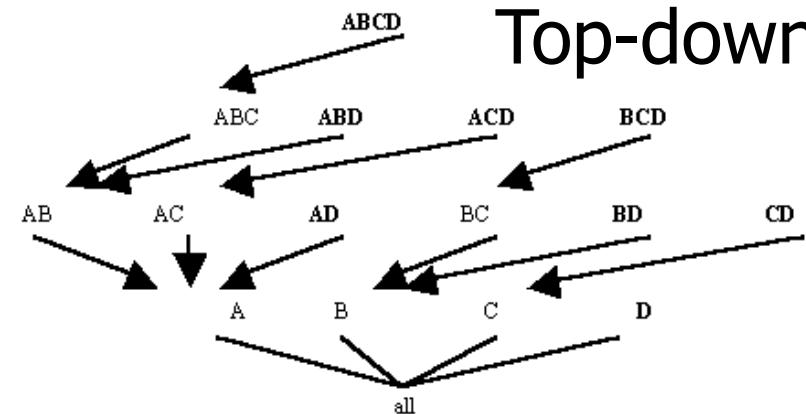
- Integração e construção de DWs
 - Data cleaning, transformação, integração, e modelagem multidimensional
- Data cubing
 - Métodos escaláveis (i.e., materialização) agregados multidimensionais
 - OLAP (online analytical processing)
- *Multidimensional concept description*: Caracterização e discriminação
 - Generalizar, sumariar, e contrastar características dos dados, ex., seco vs. molhado (região)

Função: (1) Generalização

Bottom-up



Top-down



cube

dimensions

| | Part | | | |
|----|-----------|---------|---------|----------|
| P1 | 2 | 5 | 7 | 6 |
| P2 | 4 | | 4 | 13 |
| P3 | 1 | 8 | 7 | 9 |
| P4 | | 3 | 4 | 2 |
| P5 | 9 | 2 | 5 | 7 |
| | 7 | 15 | 5 | 27 |
| | Vancouver | Calgary | Toronto | Montreal |
| | | | | Vance |

| | | | |
|----|---|----|----|
| 3 | 2 | 2 | 2 |
| 11 | 1 | 3 | 5 |
| 9 | 4 | 12 | 13 |
| 2 | | 2 | 7 |
| 9 | | 9 | 2 |
| 31 | 7 | 38 | 16 |

Bob

| | | | | |
|----|----|----|----|-----|
| 3 | 4 | 7 | 3 | 17 |
| 11 | 7 | 3 | 21 | |
| 14 | 6 | 14 | 8 | 42 |
| 2 | | 2 | 7 | 11 |
| 10 | | 2 | 8 | 20 |
| 40 | 17 | 28 | 26 | 111 |

Allison

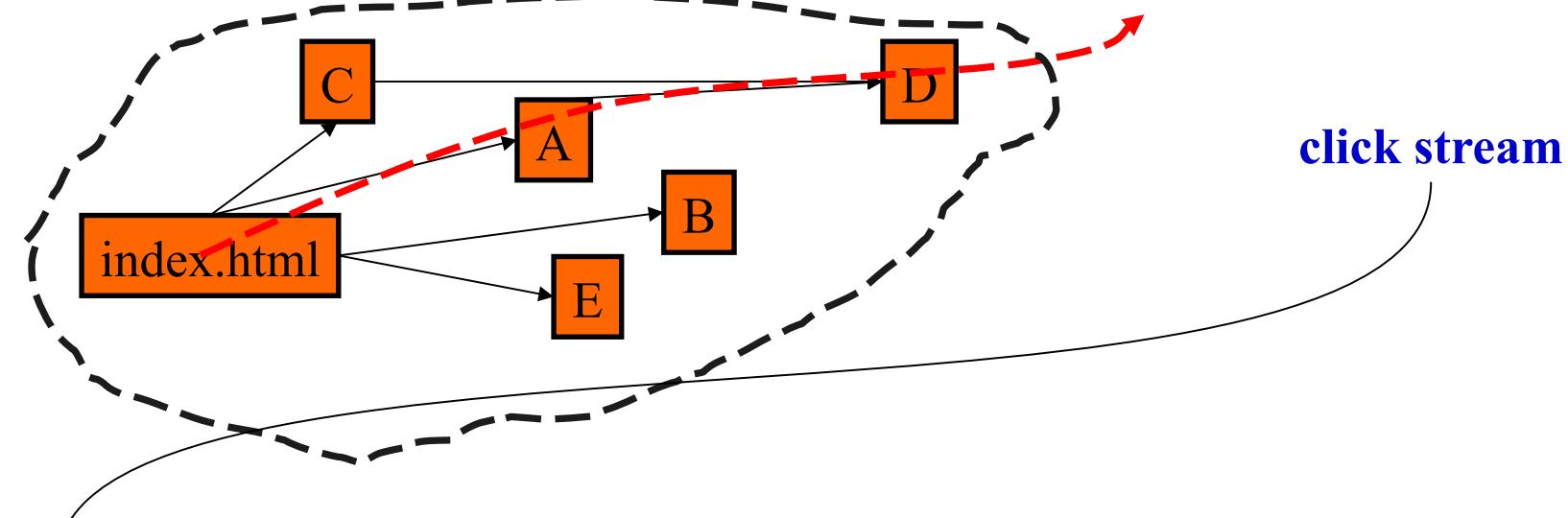
Richard

cells

Função: (2) Associação e Correlação

- Padrões frequentes (or itemsets frequentes)
 - Quais items são comprados com que frequência num supermercado?
- Associação, correlação vs. causalidade
 - Uma regra clássica de associação
 - *Diaper → Beer [0.5%, 75%] (support, confidence)*
 - **Forte associação significa forte correlação?**
- Como minerar estes tipos de padrões de forma eficiente?
- Como utilizar estes padrões para outros tipos de análise, ex., classificação, agrupamentos...outras aplicações?

Função: (2) Associação e Correlação



Obs.: A D

→ **Rules:**



→ **D** (conf: 0,8)



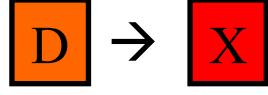
→ **D** (conf: 0,7)



→ **F** (conf: 0,6)



→ **D** (conf: 0,5)



→ **X** (conf: 0,4)

Recommendations (top 2):



(0,6)



(0,4)

Função: (3) Classificação

- Classificação, predição de classes
 - Construir modelos (funções) como base em casos de exemplo
 - Descrever e diferenciar classes ou conceitos para predição futura
 - Ex., classificar pacientes com base no (perfil clínico), or classificar modelos de carros com base nas características
 - **Predizer uma classe desconhecida**
- Técnicas clássicas
 - Árvores de decisão, modelos bayesianos, support vector machines (SVM), redes neurais, classificação com base em regras, classificação com base em padrões, modelos de regressão, ...
- Aplicações
 - Detecção de fraudes, marketing direto, classificar estrelas, estágios doenças, páginas webs, ...

IRIS dataset

| | Sepal.Length | Sepal.Width | Petal.Length | Petal.Width |
|--|--------------|-------------|--------------|-------------|
|--|--------------|-------------|--------------|-------------|

| | | | | |
|---|-----|-----|-----|-----|
| 1 | 5.1 | 3.5 | 1.4 | 0.2 |
| 2 | 4.9 | 3.0 | 1.4 | 0.2 |
| 3 | 4.7 | 3.2 | 1.3 | 0.2 |
| 4 | 4.6 | 3.1 | 1.5 | 0.2 |
| 5 | 5.0 | 3.6 | 1.4 | 0.2 |
| 6 | 5.4 | 3.9 | 1.7 | 0.4 |

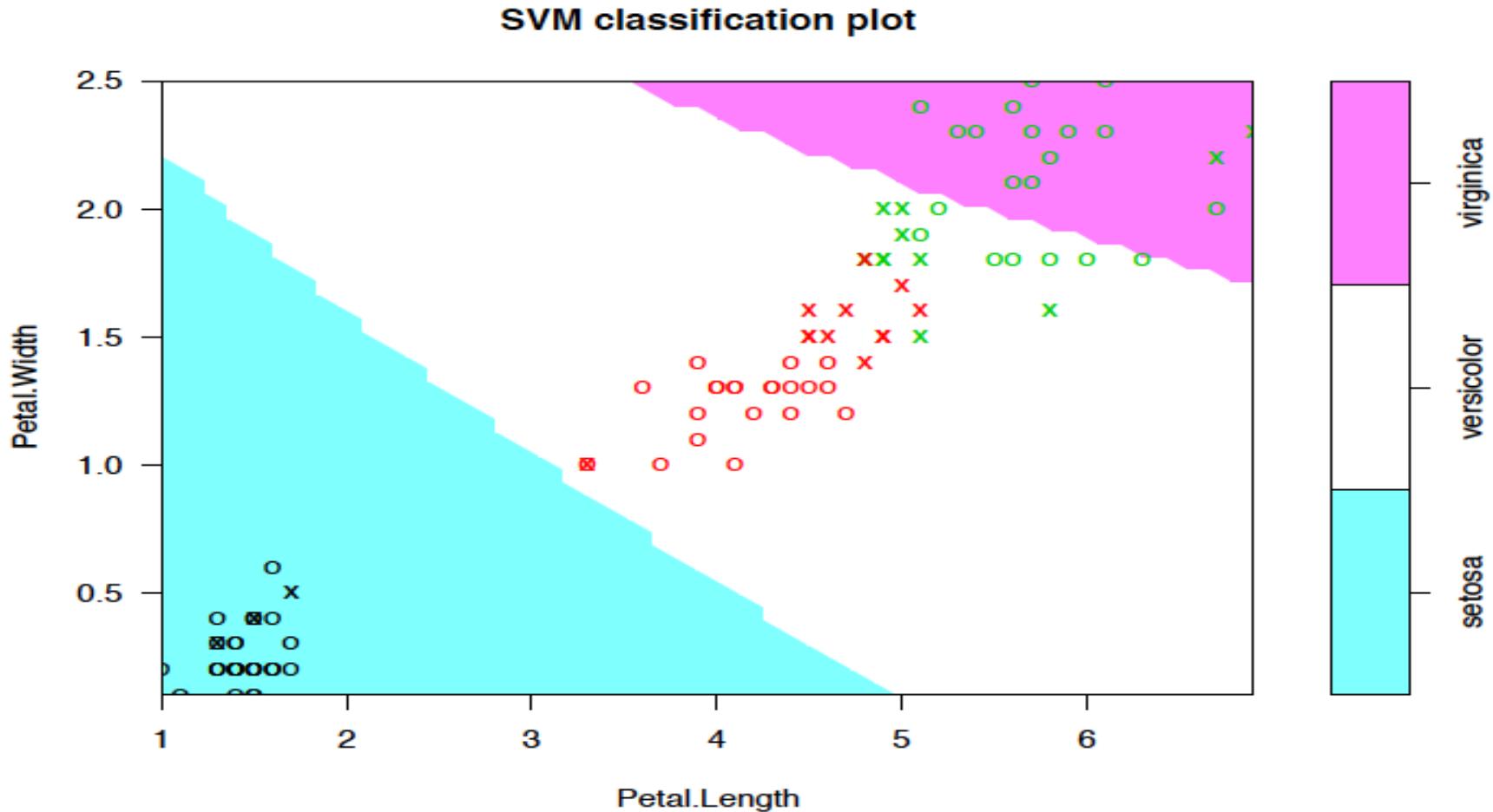
Species

1 setosa
2 setosa
3 setosa
4 setosa
5 setosa
6 setosa

A diagram illustrating the mapping from index to species name. On the left, indices 1 through 6 are listed next to their corresponding species names: 'setosa'. An arrow points from the 'setosa' label to the word 'Classe' below it. Another arrow points from the 'setosa' label to the first three species names ('setosa', 'versicolor', and 'virginica') listed on the right.

| | | |
|--------|------------|-----------|
| setosa | versicolor | virginica |
| 50 | 50 | 50 |

SVM ideal – IRIS dataset



Exemplo SVM: Estágios Câncer

■ Matrix confusão

[1] "Confusion matrix"

Classes

rawPred3 Blue Green Pink

Blue 7 0 2

Green 0 4 0

Pink 7 1 19

Performance

[1] "Accuracy"

\$diag

[1] 0.75

\$kappa

[1] 0.5490417

\$rand

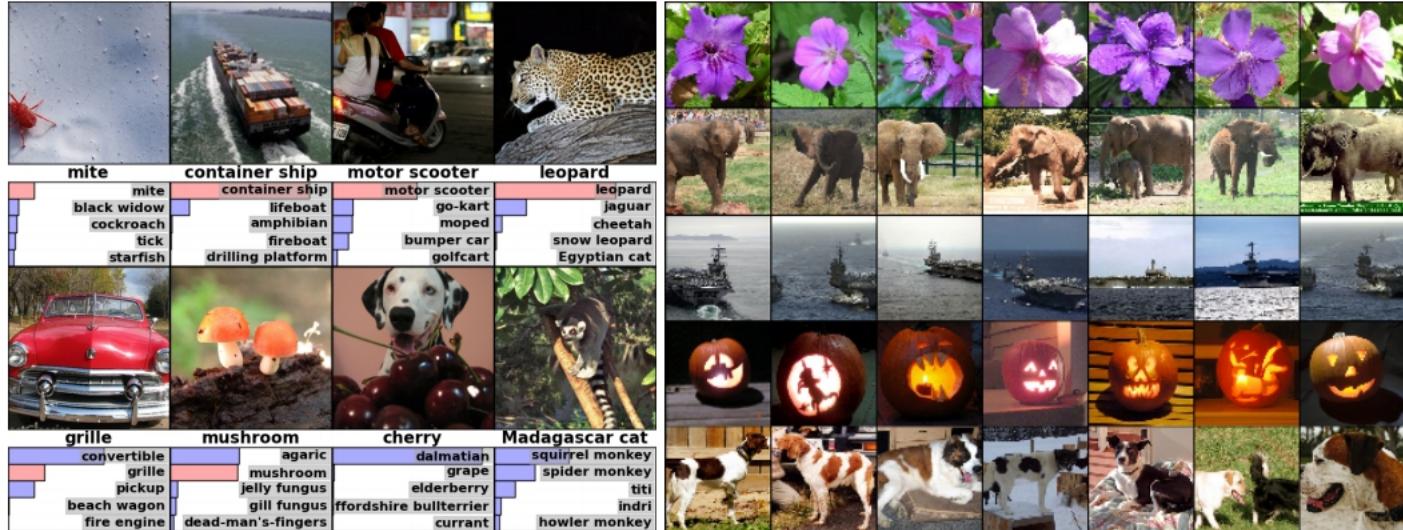
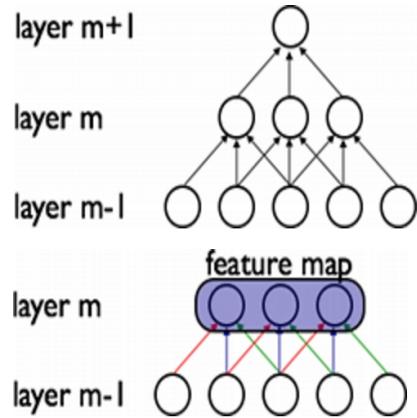
[1] 0.6615385

\$crand

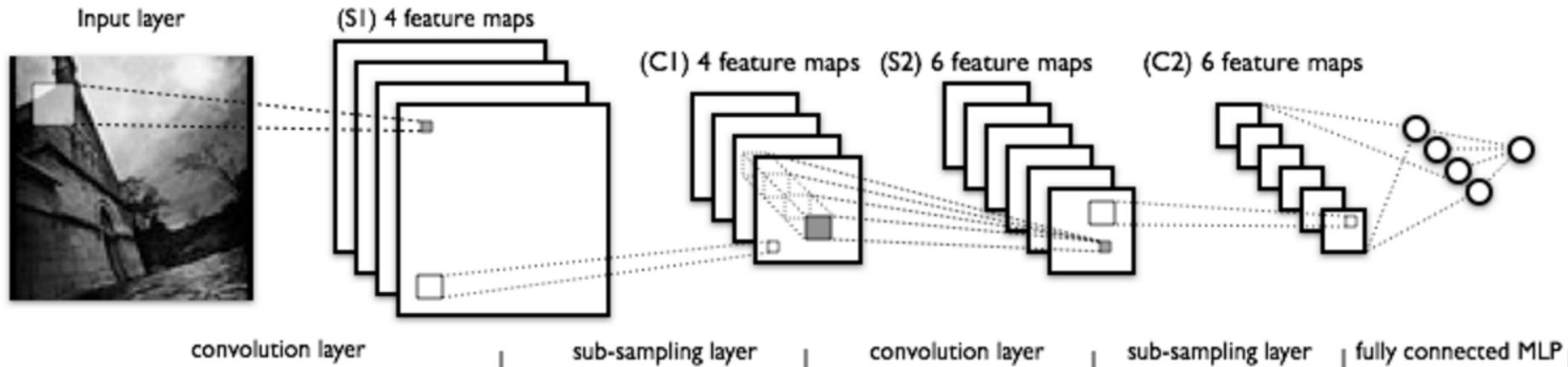
[1] 0.3241301

Overfitting? Underfitting?

Classificação via Deep Learning



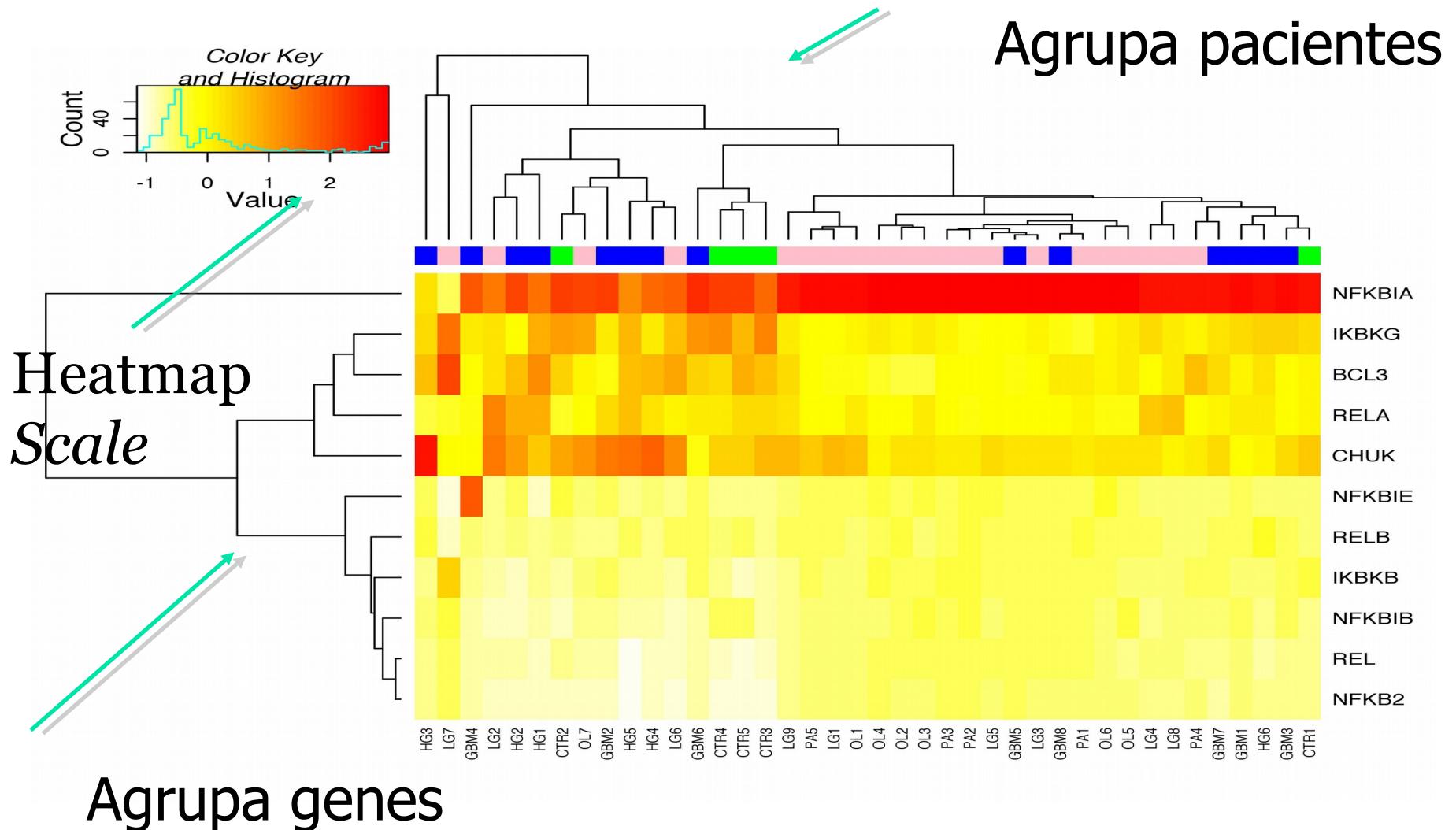
Convolutional Neural Networks (LeNet)



Função : (4) Agrupamentos

- Aprendizado não supervisionado (classe desconhecida)
- **Agrupamentos formam novas categorias (i.e., clusters)**, ex., agrupar clientes segundo perfil de consumo
- Princípio: Maximizar similariedade dos elementos no grupo & minimizar similariedade entre elementos de grupos distintos
- Uma gama diversificada de métodos
Clustering hierárquico, k-means, k-medoids, SOM, etc...

Função de Mineração: (4) Agrupamentos



Função: (5) Outliers / Anomalias

- Análise de padrões desviantes
 - Outlier: Uma observação que não segue o padrão geral de comportamento dentro do conjunto de dados observados
 - Ruído ou exceção? — **O que pode ser “lixo” para uma pessoa pode ser um “tesouro” para outra...**
 - Métodos: resultantes de aplicações de clustering ou análise de regressão, ...
 - Bastante usado para detectar casos de fraude, eventos raros

Função: (6) Tempo & Ordem

- Sequência, tendência e evolução
 - Tendência, série-temporal, e padrões desviantes: ex., regressão e predição
 - Mineração de padrões sequenciais
 - Ex., primeiro compra uma câmera, então compra um cartão de memória SD
 - Análise de periodicidade
 - Pesquisa de motivos em sequências
 - Aproximação, motivos consecutivos
 - Análise com base em similariedades
- Mineração de streams de dados
 - Ordenados, temporais, infinitos, data streams

Análise de Redes (Grafos)

- Mineração de Grafos
 - Buscar sub-grafos frequentes (ex, componentes químicos), árvores (XML), sub-estruturas (fragmentos web)
- Análise de Redes de Informação
 - Redes sociais: atores (vértices) e relacionamentos (arestas)
 - ex, autores em MD, redes terroristas
 - Múltiplas redes
 - Uma pessoa participa de diferentes redes: amigos, familiares, colegas de classes, ...
 - Links carregam a semântica da relação: *Link mining*
- Mineração da Web
 - **Web: do PageRank ao Google**
 - Análise da Web
 - Comunidades, opiniões, utilização, ...

Avaliação do Conhecimento

- **São todos os padrões realmente interessantes?**
 - Muitos dados >> Muitos Padrões >> ?? conhecimento
 - **Pode ser dimensionalmente relevante (tempo, espaço, ...)**
 - Pode não ser representativo, pode ser transiente, ...
- Avaliação do conhecido gerado → minerar apenas padrões interessantes?
 - Descritivo vs. preditivo
 - Cobertura
 - **Típico vs. novidade**
 - Precisão
 - Independente do tempo, espaço,...
 - **Utilidade**

Aplicações

- **Web: classificação de páginas, agrupamentos, rankings:
PageRank & HITS algorithms**
- Sistemas de Recomendação
- Análise do cesto de compra, perfil de consumo, marketing
- **Médica e Biológica: classificação, análise de agrupamentos
(microarray data analysis), sequências biológicas, biologia de
sistemas (genômica, proteômica, transcriptômica, etc...)**
- Mineração e engenharia de software (e.g., IEEE Computer, Aug. 2009 issue)
- Mineração de constelações
- **Mineração de caminhos, trajetórias...**

Desafios (1)

- Metodologia de Mineração
 - Minerar vários tipos de dados distintos, granularidades...
 - Conhecimento em espaço multidimensional
 - Mineração: Esforço constante e multidisciplinar
 - *Boosting the power of discovery*
 - Lidar com ruídos, incerteza, e dados incompletos
 - Avaliação de padrões, utilização de restrições
- Interatividade
 - Mineração interativa
 - Incorporação de conhecimento prévio
 - Apresentação e visualização

Desafios (2)

- Eficiência e escalabilidade dos algoritmos
 - Paralelo, distribuído, stream, e incrementais
- Diversidade
 - Complexidade na formação dos dados
- Impactos na sociedade
 - Questões sociais de comportamento
 - Mineração vs Privacidade
 - Mineração pervasiva

ICDM'10: DM Open Problems!!!

- 1. Developing a Unifying Theory of Data Mining
- 2. Scaling Up for High Dimensional Data/High Speed Streams
- 3. Mining Sequence Data and Time Series Data
- 4. Mining Complex Knowledge from Complex Data
- 5. Data Mining in a Graph Structured Data
- 6. Distributed Data Mining and Mining Multi-agent Data
- 7. **Data Mining for Biological and Environmental Problems**
- 8. Data-Mining-Process Related Problems
- 9. Security, Privacy and Data Integrity
- 10. Dealing with Non-static, Unbalanced and Cost-sensitive Data

A Robot Took My Job – Was It a Robot or AI?

Automation Always Meant Shifting Job Needs!

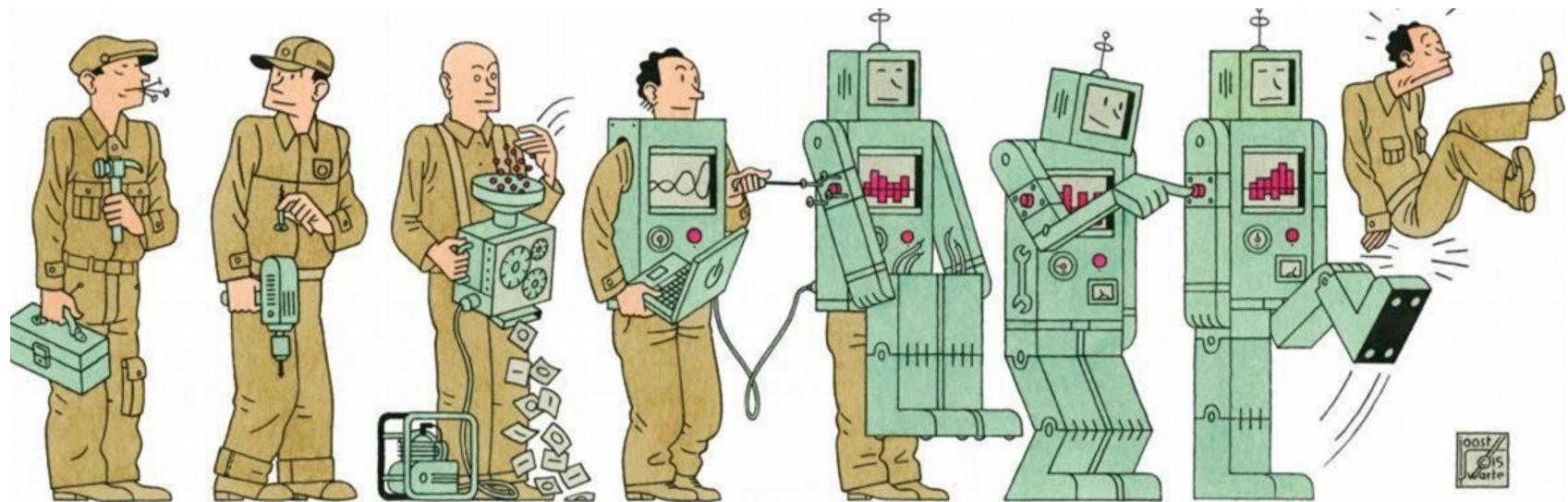
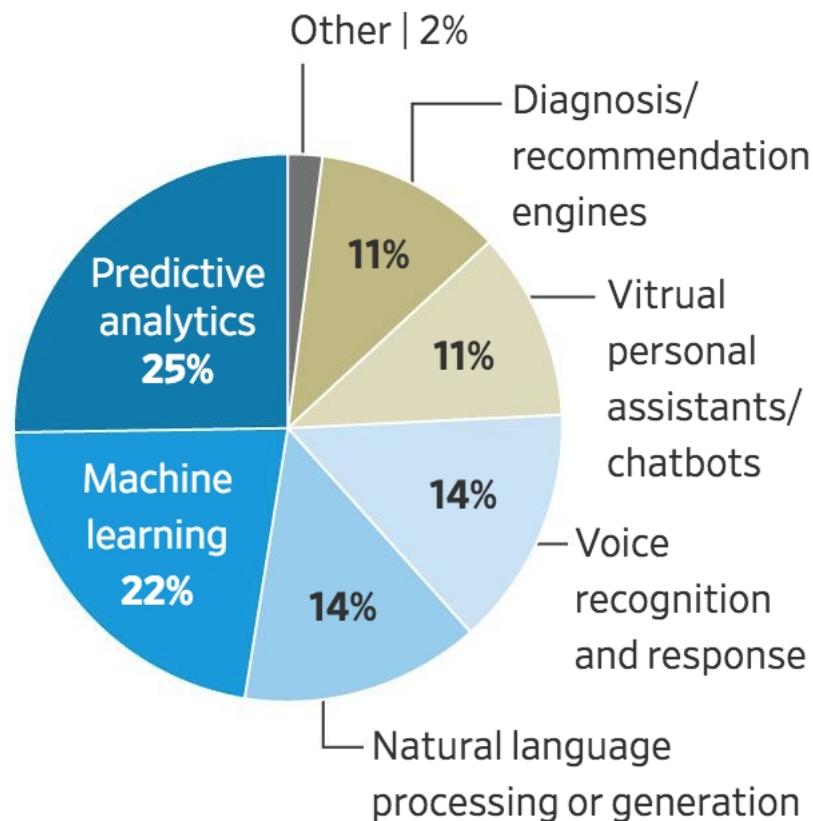


Image Source: MIT Technology Review

A Robot Took My Job – Was It a Robot or AI?

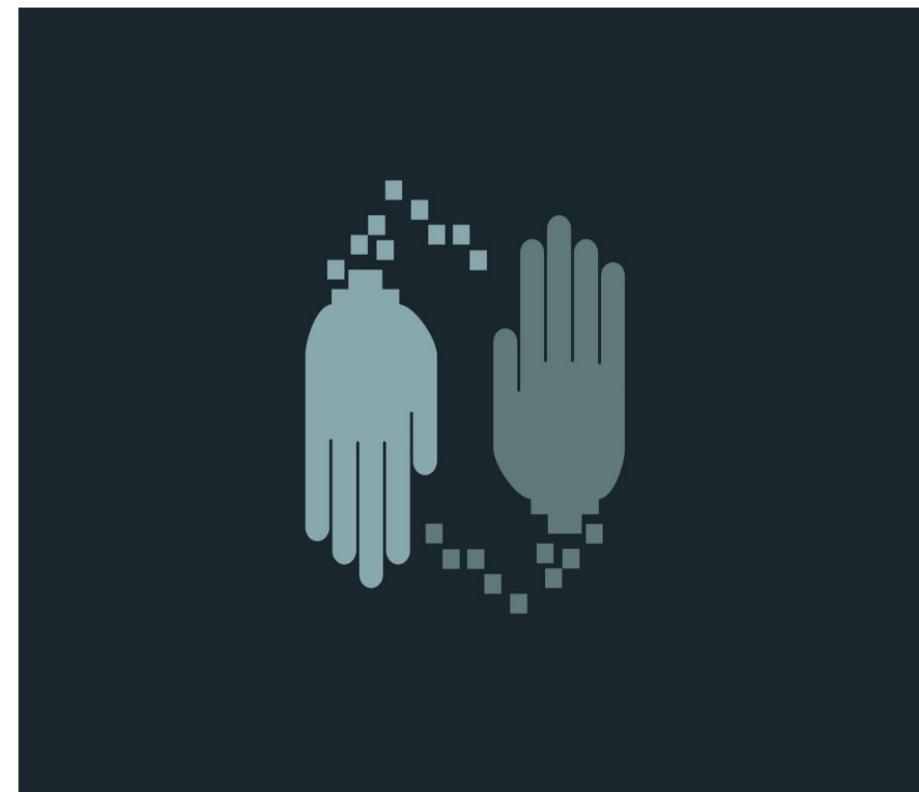
Putting AI to Work

The most common uses of AI at surveyed businesses



In a project called **AutoML**, Google's researchers have taught **machine-learning software to build machine-learning software**. In some instances, what it comes up with is more powerful and efficient than the best systems the researchers themselves can design.

GOOGLE'S LEARNING SOFTWARE LEARNS TO WRITE LEARNING SOFTWARE



[WORKFORCE PLANNING]

The Jobs That Artificial Intelligence Will Create

A global study finds several new categories of human jobs emerging, requiring skills and training that will take many companies by surprise.

BY H. JAMES WILSON, PAUL R. DAUGHERTY, AND NICOLA MORINI-BIANZINO

REPRESENTATIVE ROLES CREATED BY AI

Accenture's global study of more than 1,000 large companies identified the emergence of three new categories of uniquely human jobs.

| | | |
|------------|---|--|
| TRAINERS | Customer-language tone and meaning trainer | Teaches AI systems to look beyond the literal meaning of a communication by, for example, detecting sarcasm. |
| | Smart-machine interaction modeler | Models machine behavior after employee behavior so that, for example, an AI system can learn from an accountant's actions how to automatically match payments to invoices. |
| | Worldview trainer | Trains AI systems to develop a global perspective so that various cultural perspectives are considered when determining, for example, whether an algorithm is "fair." |
| EXPLAINERS | Context designer | Designs smart decisions based on business context, process task, and individual, professional, and cultural factors. |
| | Transparency analyst | Classifies the different types of opacity (and corresponding effects on the business) of the AI algorithms used and maintains an inventory of that information. |
| | AI usefulness strategist | Determines whether to deploy AI (versus traditional rules engines and scripts) for specific applications. |
| SUSTAINERS | Automation ethicist | Evaluates the noneconomic impact of smart machines, both the upside and downside. |
| | Automation economist | Evaluates the cost of poor machine performance. |
| | Machine relations manager | "Promotes" algorithms that perform well to greater scale in the business and "demotes" algorithms with poor performance. |

Onde buscar referências? DBLP, CiteSeer, Google

- Data mining and KDD (SIGKDD: CDROM)
 - Conferences: ACM-SIGKDD, IEEE-ICDM, SIAM-DM, PKDD, PAKDD, etc.
 - Journal: Data Mining and Knowledge Discovery, KDD Explorations, ACM TKDD
- Database systems (SIGMOD: ACM SIGMOD Anthology—CD ROM)
 - Conferences: ACM-SIGMOD, ACM-PODS, VLDB, IEEE-ICDE, EDBT, ICDT, DASFAA
 - Journals: IEEE-TKDE, ACM-TODS/TOIS, JIIS, J. ACM, VLDB J., Info. Sys., etc.
- AI & Machine Learning
 - Conferences: Machine learning (ML), AAAI, IJCAI, COLT (Learning Theory), CVPR, NIPS, etc.
 - Journals: Machine Learning, Artificial Intelligence, Knowledge and Information Systems, IEEE-PAMI, etc.
- Web and IR
 - Conferences: SIGIR, WWW, CIKM, etc.
 - Journals: WWW: Internet and Web Information Systems,
- Statistics
 - Conferences: Joint Stat. Meeting, etc.
 - Journals: Annals of statistics, etc.
- Visualization
 - Conference proceedings: CHI, ACM-SIGGraph, etc.
 - Journals: IEEE Trans. visualization and computer graphics, etc.

Resumo

- Descoberta de padrões interessantes e conhecimento a partir de grandes quantidades de dados
- Evolução natural das mais diversas tecnologias, grande demanda, com uma imensa abrangência
- Processo de KDD: limpeza, integração, seleção, transformação, **extração de padrões**, avaliação, e representação do conhecimento
- Funções: caracterização, discriminação, associação, classificação, agrupamentos, outliers, tendências, etc.
- Interpretabilidade x performance
- Automatização do aprendizado de máquinas
- Existem diversas questões em aberto...

Referências

- **S. Chakrabarti. Mining the Web: Statistical Analysis of Hypertext and Semi-Structured Data.** Morgan Kaufmann, 2002
- **R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification, 2ed., Wiley-Interscience,** 2000
- **T. Dasu and T. Johnson. Exploratory Data Mining and Data Cleaning.** John Wiley & Sons, 2003
- **U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. Advances in Knowledge Discovery and Data Mining.** AAAI/MIT Press, 1996
- **U. Fayyad, G. Grinstein, and A. Wierse, Information Visualization in Data Mining and Knowledge Discovery,** Morgan Kaufmann, 2001
- **J. Han and M. Kamber. Data Mining: Concepts and Techniques.** Morgan Kaufmann, 3rd ed., 2011
- **D. J. Hand, H. Mannila, and P. Smyth, Principles of Data Mining,** MIT Press, 2001
- **T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed., Springer-Verlag,** 2009
- **B. Liu, Web Data Mining,** Springer 2006.
- **T. M. Mitchell, Machine Learning,** McGraw Hill, 1997
- **G. Piatetsky-Shapiro and W. J. Frawley. Knowledge Discovery in Databases.** AAAI/MIT Press, 1991
- **P.-N. Tan, M. Steinbach and V. Kumar, Introduction to Data Mining,** Wiley, 2005
- **S. M. Weiss and N. Indurkhya, Predictive Data Mining,** Morgan Kaufmann, 1998
- **I. H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations,** Morgan Kaufmann, 2nd ed. 2005