Machine Learning
(the speaker's humble contributions)

MIROSLAV KUBAT

All men are boastful

- Youth:
  - “I will accomplish great things!”
- Man in his prime:
  - “I am soooo important!”
- Retiree:
  - “Once upon a time ….”

21 years of professional life

End of career

Primary fields of interest

- Machine learning
  - Mainly “class recognition”
- Artificial Intelligence

What is *class recognition* good for?

Recognition of objects that are difficult to describe precisely
- e.g., biological objects

Early applications: automated classification of critical events
- E.g., credit card fraud, email spam, …

Automated document categorization

Prediction of gene functions, analyses of DNA

…. all the way up to applications in computer vision (see next slide)

Learnable tasks in computer vision

Where is his nose?

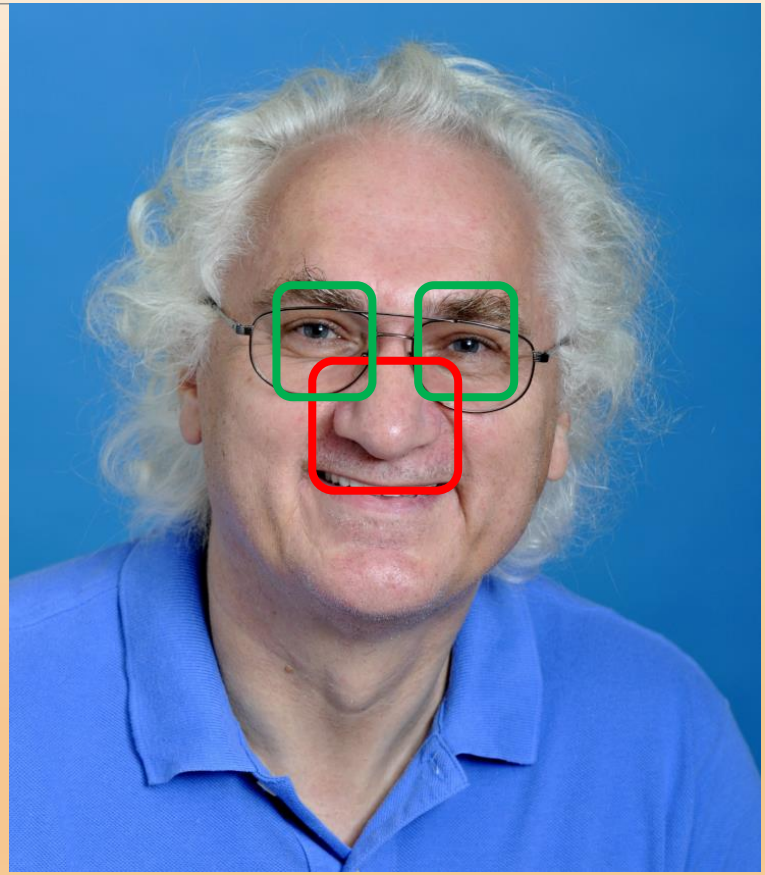
How many eyes does he have?

TWO

Is there an airplane in the picture?

NOOOOO.

These recognition tasks can be learned, too!

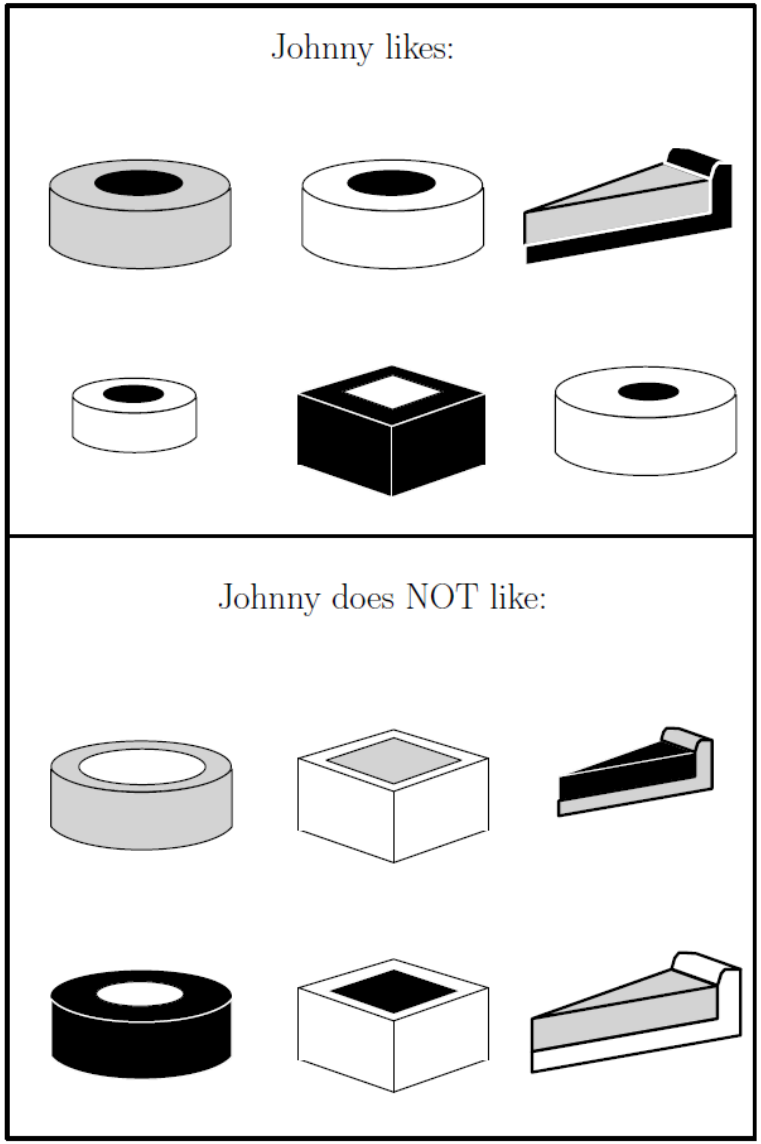


General framework for concept learning

Input: training examples, usually described by attribute vectors

Output: a classifier (logical expression, Bayesian probabilities, prototypes, linear classifiers, decision trees, neural nets, etc.)

What is needed: *machine learning* software for the induction of the classifier from training data

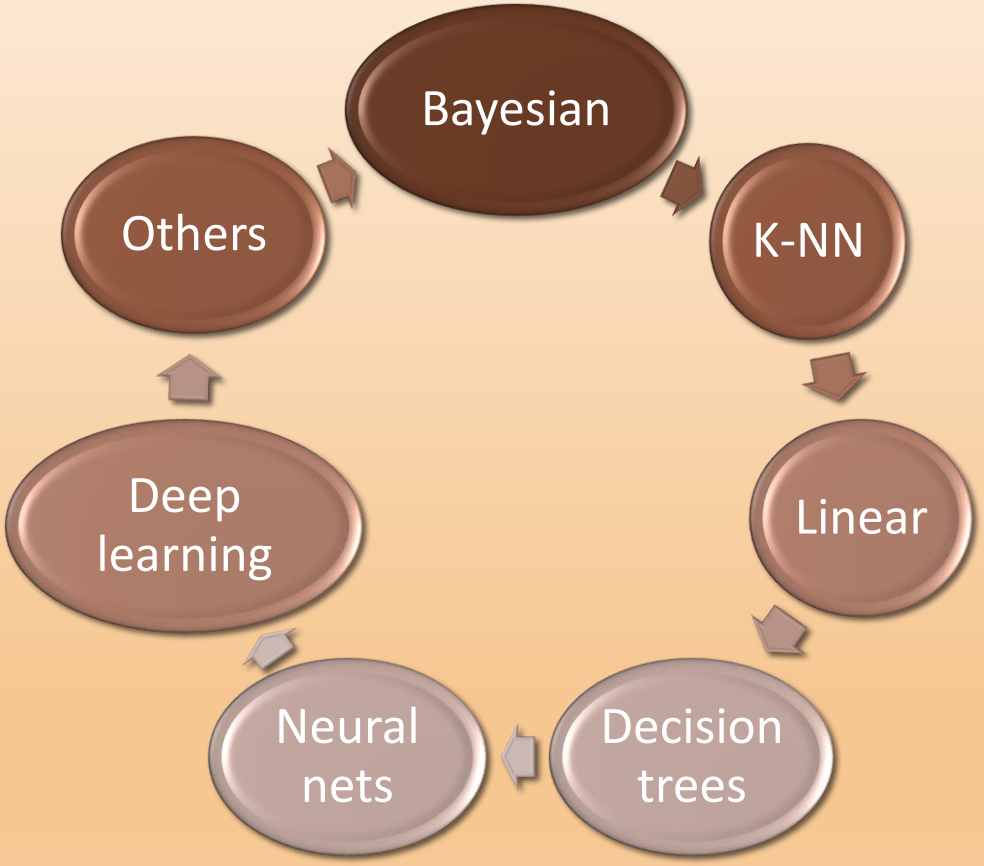


What kind of pie does Johnny like?

The training set from the previous slide

	shape	crust size	crust shade	filling size	filling shade	class
Ex.1	circle	thick	gray	thick	dark	**pos**
Ex.2	circle	thick	white	thick	dark	**pos**
Ex.3	triangle	thick	dark	thick	gray	**pos**
Ex.4	circle	thin	white	thin	dark	**pos**
Ex.5	square	thick	dark	thin	white	**pos**
Ex.6	circle	thick	white	thin	dark	**pos**
Ex.7	circle	thick	gray	thick	white	**neg**
Ex.8	square	thick	white	thick	gray	**neg**
Ex.9	triangle	thin	gray	thin	dark	**neg**
Ex.10	circle	thick	dark	thick	white	**neg**
Ex.11	square	thick	white	thick	dark	**neg**
Ex.12	triangle	thick	white	thick	gray	**neg**

Many paradigms



Other machine learning tasks

Reinforcement learning

- Learn to choose the right action in a concrete situation

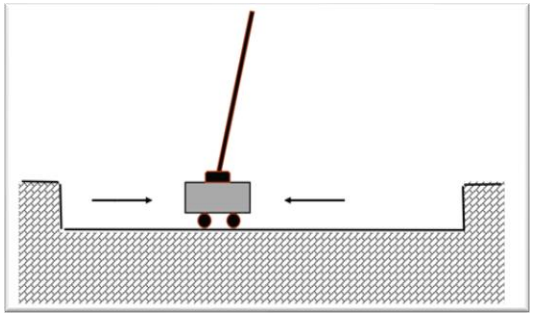
Unsupervised learning

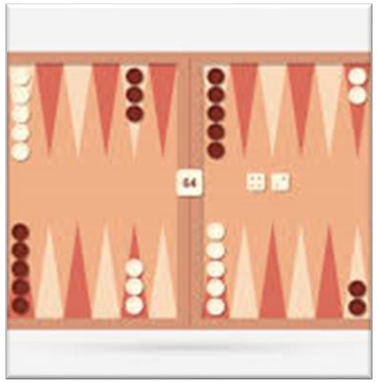
- Find underlying properties of the data, visualize them

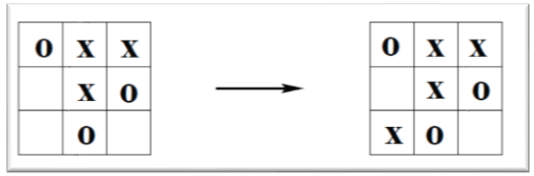
Hidden Markov Models

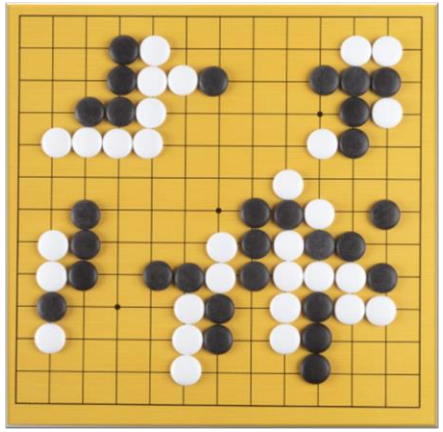
- Learn about sequences of events based on indirect clues

Even: learn to play at master level





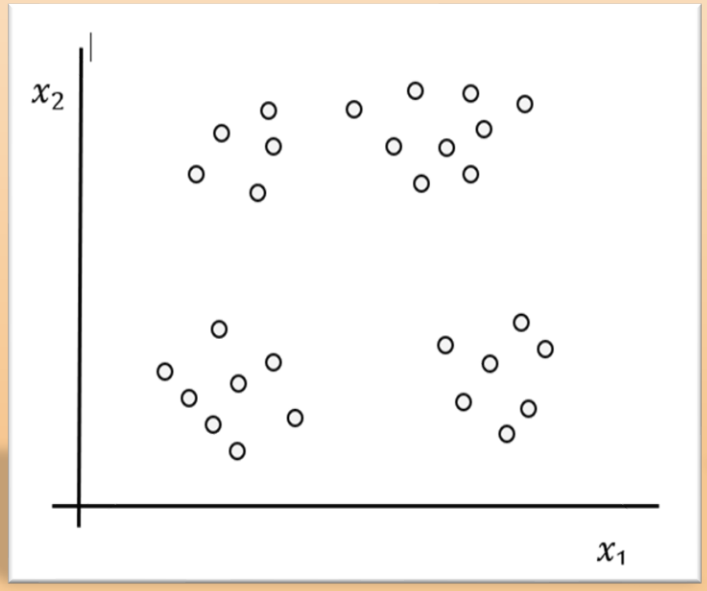


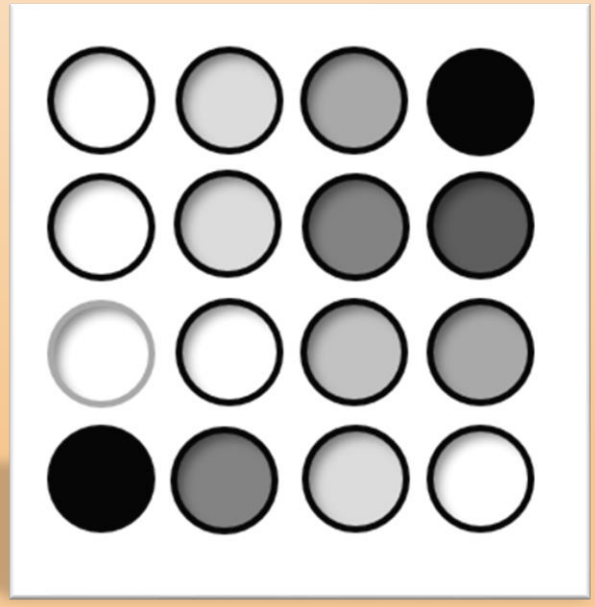


Another playground for ML: find properties of data

Identify clusters of similar examples.

Visualize distances among data by mapping them to two dimensions.





History of ML (outline)

- Single layer of neurons: Rosenblatt (1961)

- Early death: Minski and Papert (1969)

- Michalski, Carbonell, and Mitchell (1983)

- Resurrection of NN: Rummelhard and McLelland (1986)

- Early paradigms: nearest neighbor, Bayesian classification, decision trees, graph induction, explanation-based learning, inductive logic programming

- Major watershed: Mitchell's textbook (1992)

- Maturity (around 2000): reinforcement learning, hidden Markov models, then deep learning (especially the last 10-15 years).

- Shift from symbol manipulation to numeric approaches (including deep learning), then abandonment of symbol manipulation (e.g., Prolog)

Stages of success

Stage one: benefited from innocence (mistakenly sent a paper to a prestigious journal and it was accepted)

- M Kubat: "Conceptual inductive learning: the case of unreliable teacher," *Artificial Intelligence Journal, 1991*
  - Two theorems and a bit more math

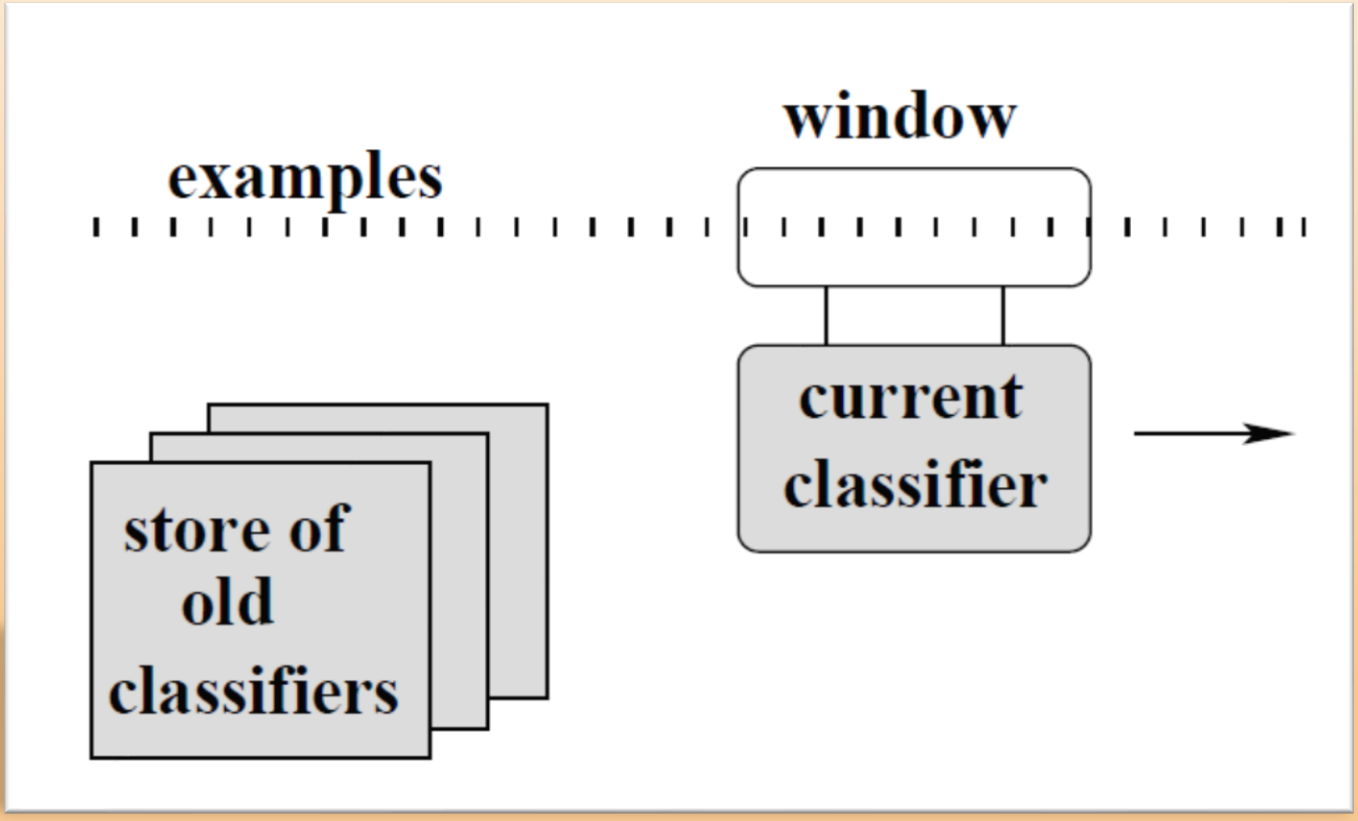
- This paper was largely ignored (as it deserved): 13 citations
- But: *Artificial Intelligence*– the most prestigious computer-science journal of the date
- thanks to the journal's prestige, invitations to editorial boards, program committees, etc.

Stages of success

Stage two: benefited from ignorance (succeeded thanks to a more experienced colleague)

- Did not know how to handle induction in time-varying domains (the meaning of some concepts varies in time)
- 1992 – 1996: series of papers on FLORA by M. Kubat and G. Widmer
- Culmination: G. Widmer and M Kubat: "Learning in the presence of concept drift and hidden contexts," *Machine Learning Journal, 1996*
- *2735 citations (Google's Scholar)*
- Consequences:
  - At least two special issues on the topic (including the MLJ)
  - Specialized conference workshops

One solution for concept-drift domains: FLORA



The FLORA paradigm

Examples arrive one at a time, or in batches, added to a window.

The classifier maximizes performance on examples in the window.

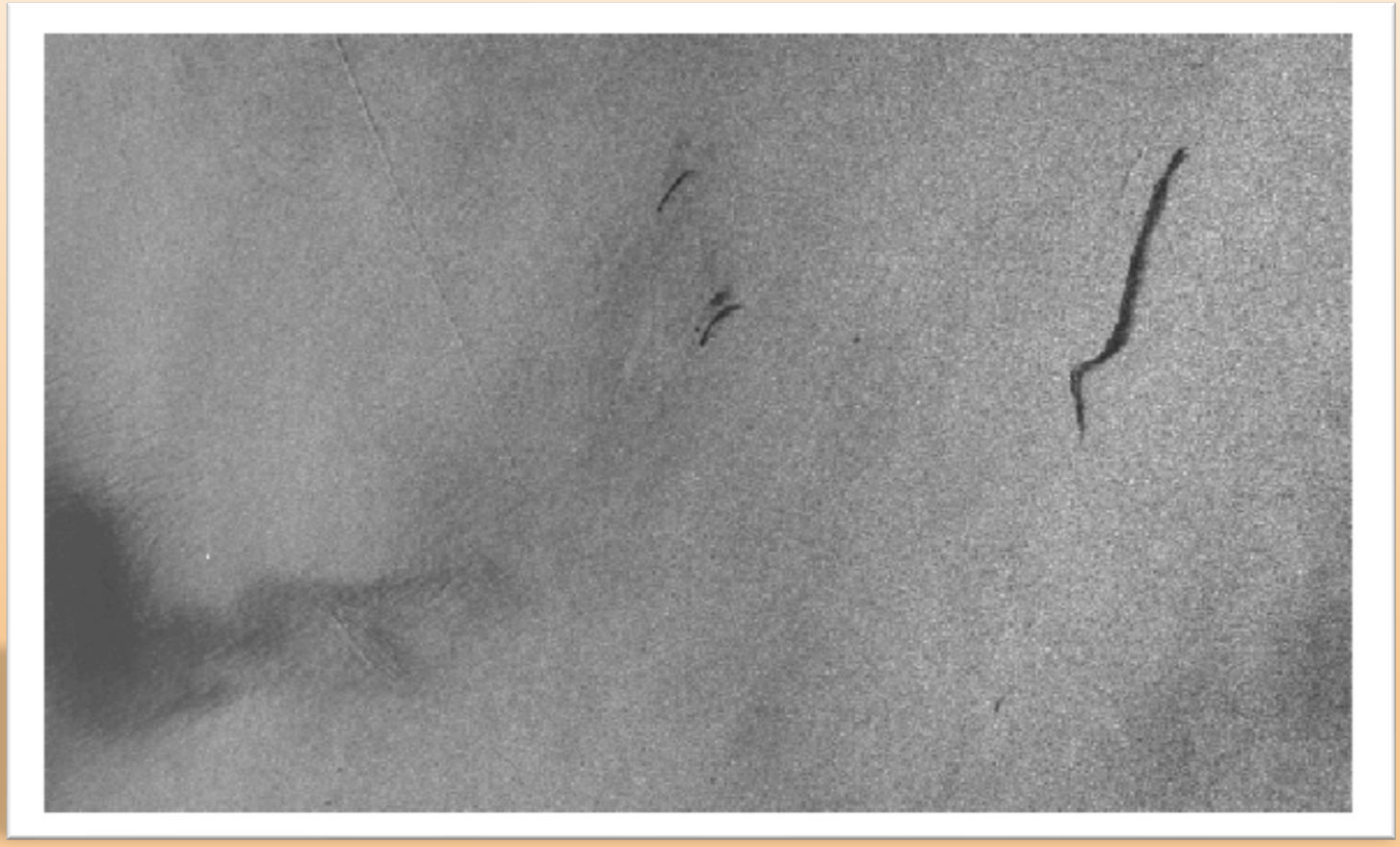
Each time an example is added to the window, the system decides:

- Should I delete old examples as obsolete?
- How many of the examples should I delete?
- Should I re-induce the classifier from scratch or just to modify it?

Sometimes, the classifier's older versions of the classifier are stored, to be retrieved when necessary.

- E.g., when the changes are periodical, such as in seasons.

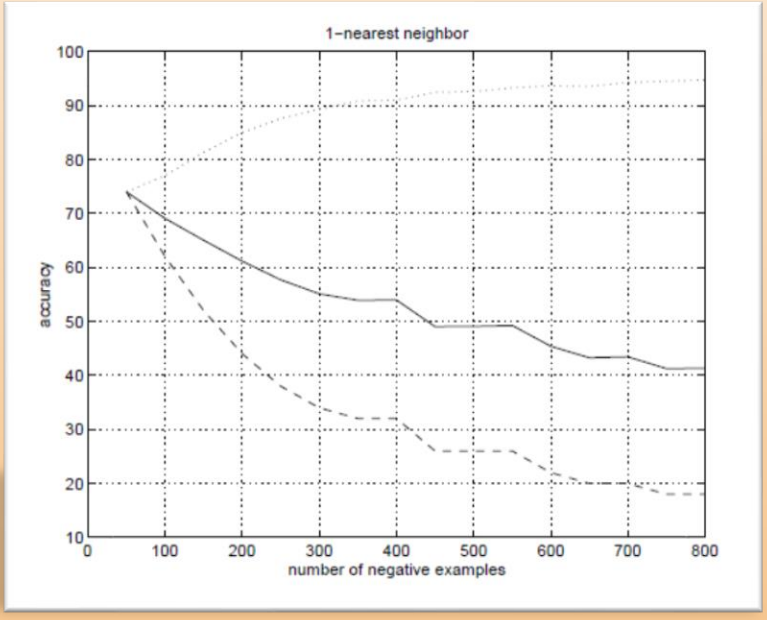
A talked-about application: oil-spills



Stages of success

- The oil-spill domain data (limited size):
  - 50 positive training examples
  - 850 negative training examples
- The speaker's complaint:
  - "None of my attempts succeeded!"
- His boss's request:
  - "Try every single learning algorithm you know!"
- Computers in 1990s: very slow, by today's standards
  - Therefore: used only 50 positive and 50 negative examples
  - Astonishment:
  - The results were much better!

The more imbalanced the training set, the worse the results



The set of positive examples is always the same.

The number of negative examples is in each experiment higher (horizontal axis).

Accuracy on the majority class keeps increasing (dotted line).

Accuracy on the minority class keeps decreasing (dashed line).

Geometric mean of the two goes down (solid line).

Imbalanced training sets: three methods to mitigate the harm

Modify the original induction algorithm:

- Linear classifiers: shift the classifier by changing the bias (threshold).
- K-NN classifiers: modify the voting rule (e.g., in 7-NN, the example may be labeled as positive even if only 3 neighbors are positive).
- Bayesian classifiers and ANNs: classify as positive whenever $P_{pos}(\boldsymbol{x}) > \theta$ where θ has been properly adjusted.

Majority-class under-sampling:

- Choose a random subset of the majority-class examples.
- Better: smart choice of this subset.

Minority-class over-sampling:

- E.g., duplicate (or triplicate, etc.) the minority-class examples.
- Better, but more difficult, solution: create new minority-class examples by adding artificial noise to existing examples.

Stages of success

- **Stage three**: success thanks to his laziness and his slow computer
- M Kubat and S Matwin: “Addressing the curse of imbalanced training sets.” *International conference on machine learning,* 1997
- 3680 citations
- Consequences:
  - Special issues of journals
  - Specialized conference workshops

Stages of the speaker's success

Stage four: success thanks to arrogance

1998: oil-spill recognition

M Kubat, RC Holte, S Matwin: "Machine learning for the detection of oil spills in satellite radar images," *Machine Learning Journal*, 1998

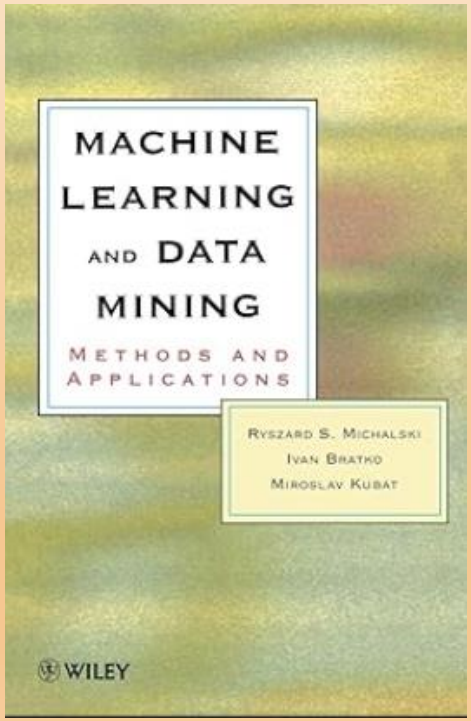
Contained suggestions for future research in the field

- Problems with data
- Problems with performance assessment
- Problems with algorithms

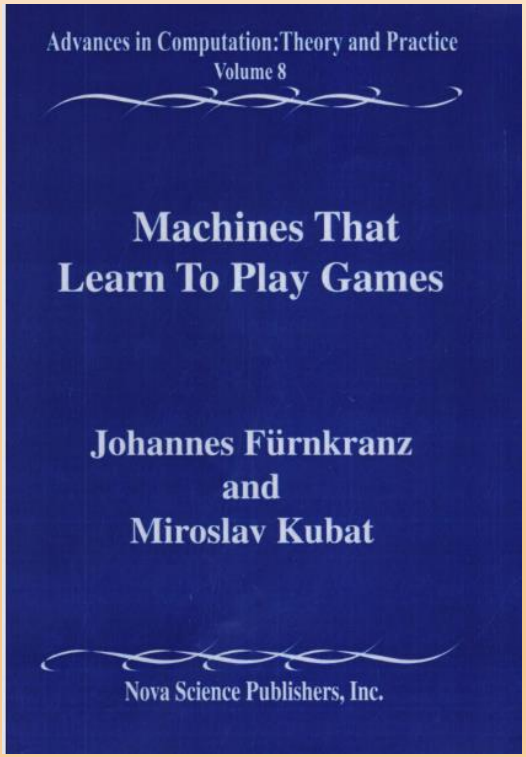
1890 citations

Edited books

MICHALSKI, BRATKO, KUBAT (1998)

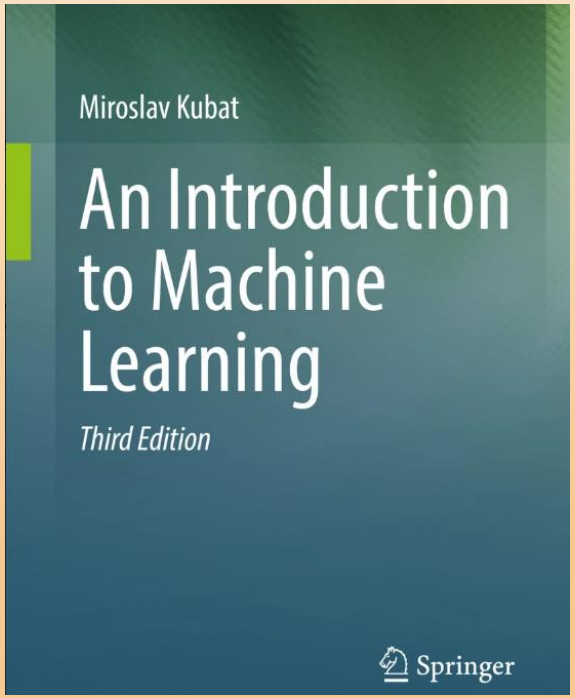


FURNKRANZ AND KUBAT (2001)

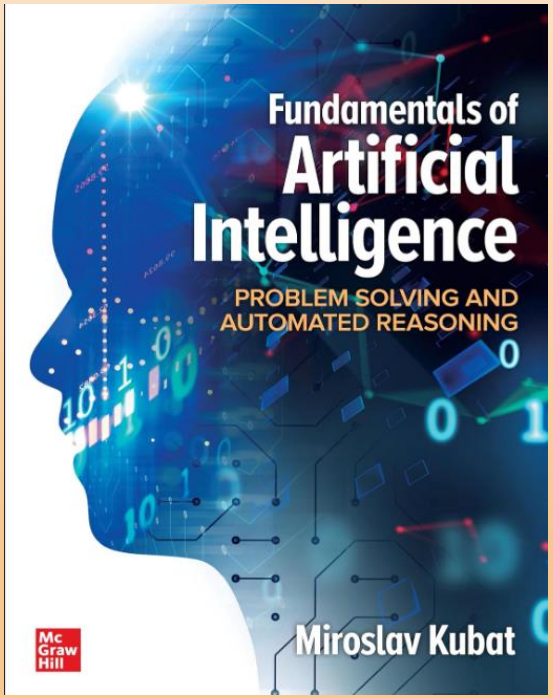


Textbooks

INTRO TO ML (2015, 2017,2021)



FUNDAMENTALS OF AI (2023)



Kubat's services to community

- Member of three editorial boards in scientific journals
- 60 PC memberships at conferences and workshops
- Co-chair of a major conference in Miami Beach (ICMLA, 2012)
- Hundreds of reviews of scientific papers

Kubat's ML curriculum

Some 100 papers, 13,000 citations in Google's Scholar

Four books

Later successes:

- Contribution to association mining (isomorphism between two data structures) (2003)
- tumor detection in MRI (2015) (thanks to a brilliant PhD student)
- Gene analysis (Nature Neuroscience 2019) (thanks to a friendly invitation)
- Some others … too boring to be listed here

A lesson to be learned?

... None (each career is different)

… A bit of luck can be more than helpful

… Most of all:

- Everybody has a talent for something
- If you find yours, you'll conquer the world

The most popular slide ….

Thank you !

☺