The Story of TL/Prevision

Léon Bottou

Joint work with Xavier Driancourt
and the Neuristique team.

Timeline.

1986 Finish X. Meet Yann Le Cun. Write SN.
1988 Start Ph.D. with F. Fogelman Soulié.
1989 Create Neuristique to sell SN.
1991 M.T.S. at AT&T Bell Labs. Meet Vapnik.
1992 Back to Neuristique.
1993 Start TL/Prevision project.
1995 Revise goals in life. Back to research.
AT&T Labs, then NEC Research Institute.

Neuristique 1992: Status.

- A hobby company.

  (12 associates, 11 board members.)

- One product: SN

  Neural network software. About \$10K.

  Customers: corporate R&D.

  PSA, Dassault, CEA, Thomson,

  AT&T, Intel, Sony, Taiwanese Telecoms, . . .

- Not a consulting shop.

Neuristique 1992: Problems.

SN is a vehicle for technology transfer.

- Little recurrent business.

Corporate R&D is sensitive to business conditions.

- Income variations from one year to the next.

Competition from Mimetics.

- F. Fogelman Soulié, Erik Marcadé and others.
  EasyReader, Mimenice, Consulting.

Advances in Machine Learning.

1991-1992, AT&T Bell Labs:

Conceptual advances in machine learning.

- capacity control
- structural risk minimization
- statistical regularisation
- support vector machines
- transductive inference

Practically usable.

[Guyon, Vapnik, Boser, Bottou & Solla, NIPS, 1992]

[Vapnik & Bottou, Neural Computation 4(6), 1992]

Better Statistics?

Yellow page statisticians:
- focus on linear models.
- low dimensionailty, few data points.
- fear multi-correlations.

Advanced statistics still very useful.
- robust regression.
- ridge regression.
- projection pursuit.
- missing and aberrant data.
- bootstrap and cross-validation.
- some bayesian ideas.

First Attempt.

Try solving high profile applications.

Example: Stock market forecast.
- Very difficult indeed.
- Returns is not a robust performance measure.
- Need expensive trading infrastructure to reliably leverage small statistical effects.

Hard problems require specific solutions.

$\Longrightarrow$ Consulting business.

$\Longrightarrow$ Not a product business.

We want generic problems and generic solutions!

Second Attempt.

<< *Instead of attacking difficult tasks,*
we should address simple (but generic) problems. >>

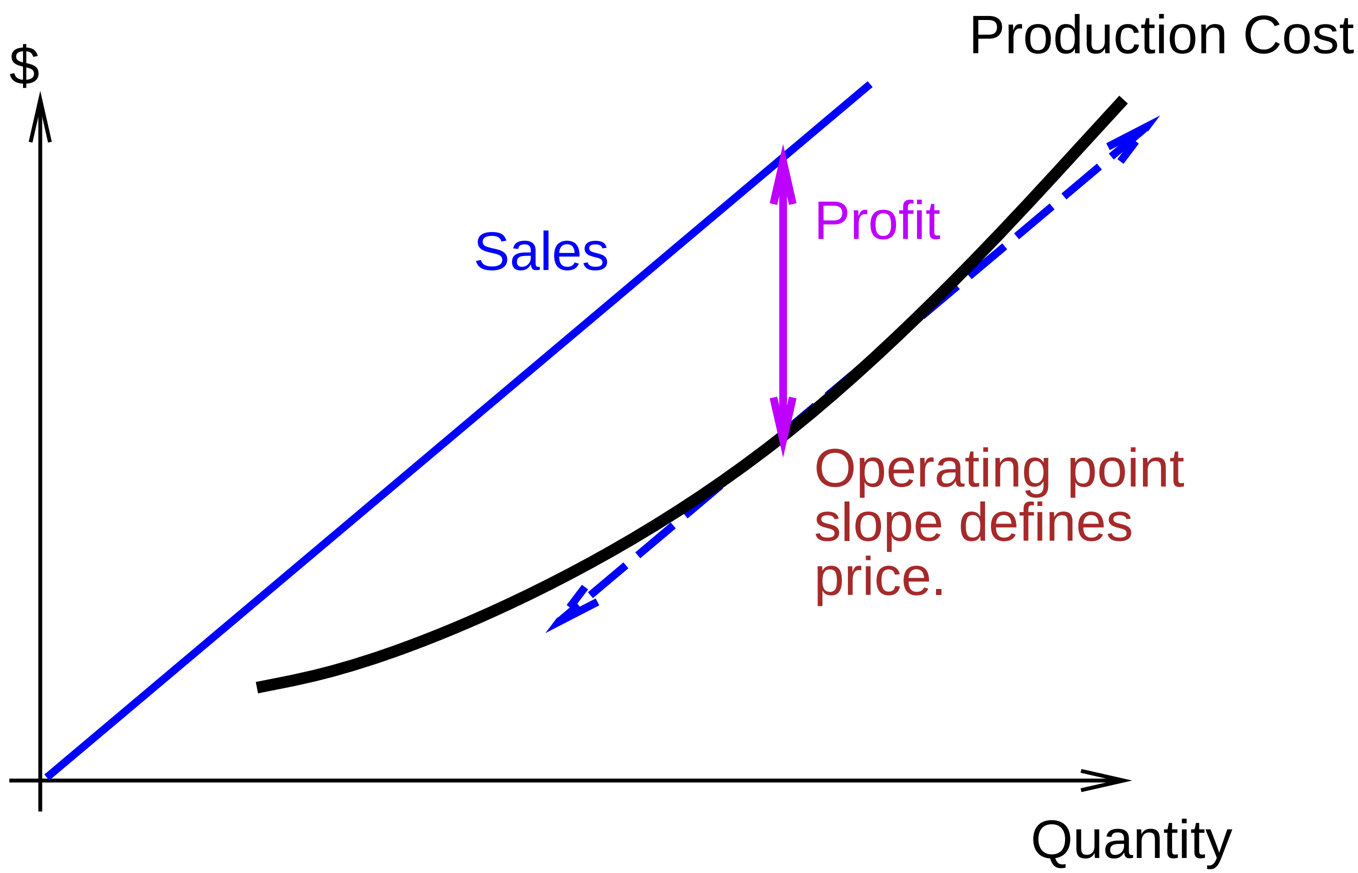
Where to find them ?

- Data mining was not a buzzword yet.

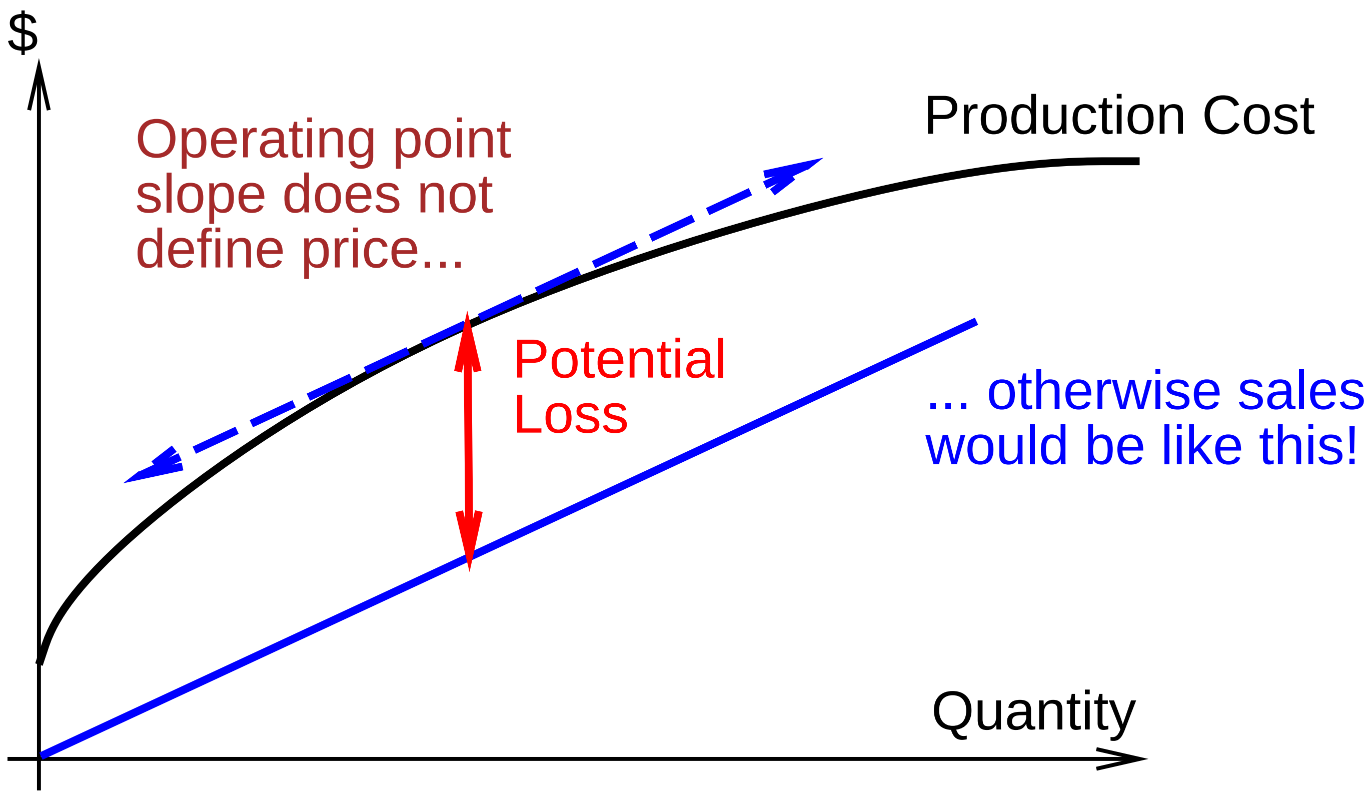
Non-convexity in micro-economy !

- How to pay for a bridge? How to tax?
- Mathematical foundations.

Convexity in Micro-Economy.



A more realistic Picture.



A more realistic Picture.

Price according to customers $\Longrightarrow$ C.R.M.

- Yield management
- Market segmentation

Fixed costs management $\Longrightarrow$ Provisioning

- Stock management.
- Demand forecast.

Economical justification for data intelligence.

Transition from industrial to information society ?

TL/Prevision

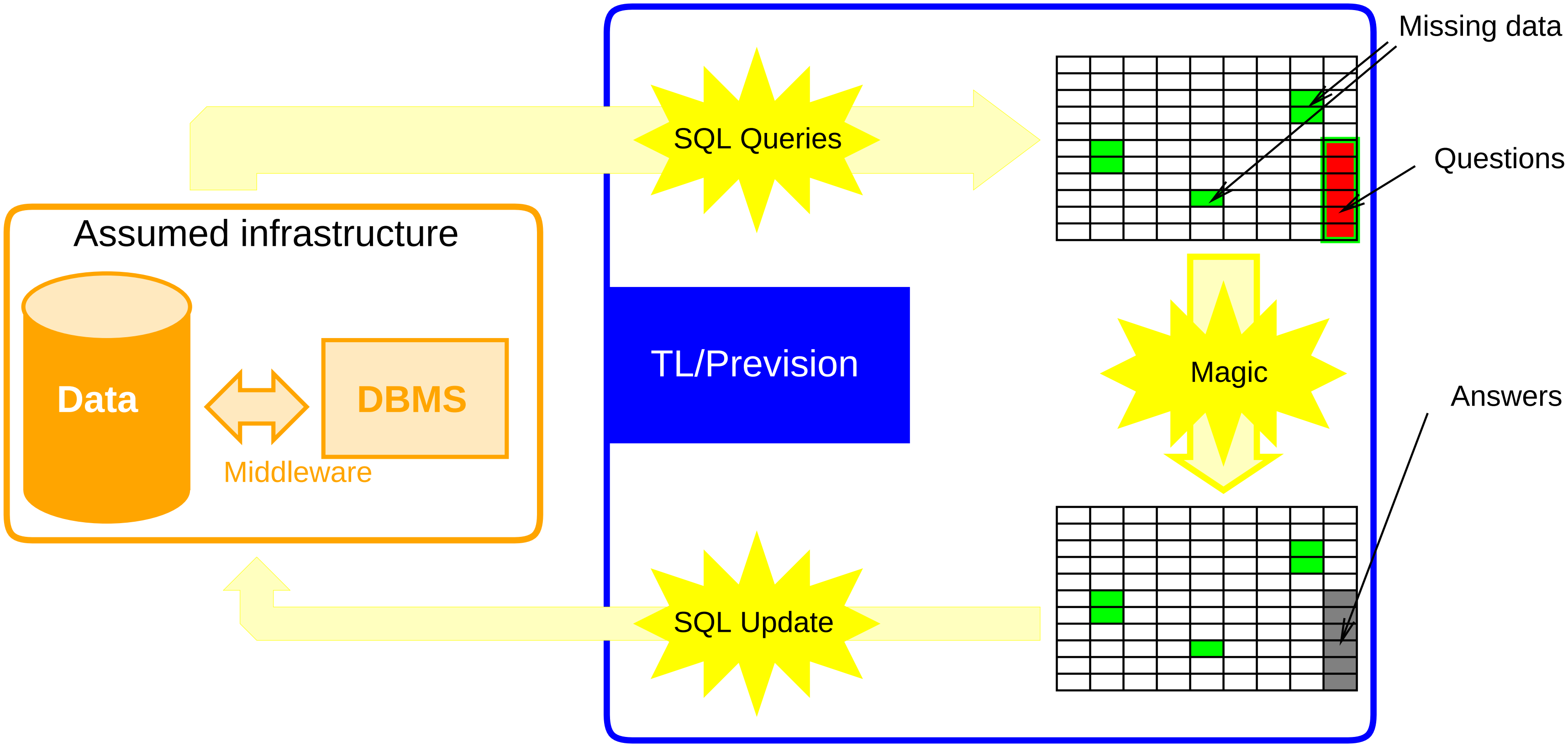
- Interface with popular DBMS.

- Create model, update model, use model.
  Handle data quality issues.

- Focus on simplicity, not accuracy!
  User must know about his data, not about statistics.
  Example: Adding an input variable to a model
  should not drastically reduce accuracy.

- ANVAR.

TL/Prevision.



Also: GUI, query management, visualisation, reports, . . .

TL/Prevision: Spot the mistakes

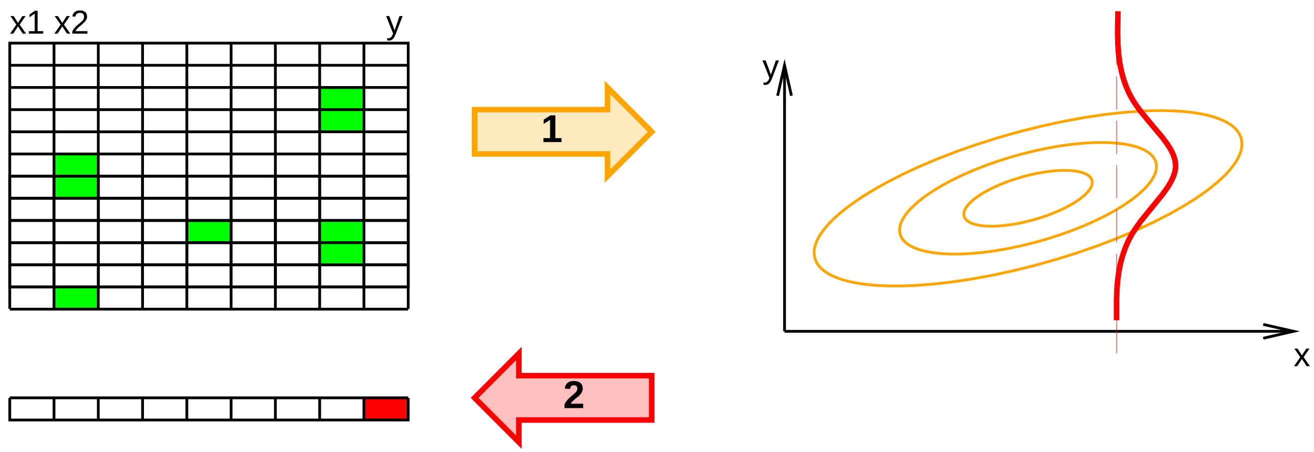
- DBMS infrastructure was not ready.
  (now "data warehouses".)

- SQL aggregate queries are slow and clumsy.
  (now "OLAP".)

- Monolithic program
  (now "Components".)

Too much focus on the machine learning challenge.

Least Square Linear Regression.



1- Estimate Gaussian on training rows.

2- Marginalize on query row.

Gaussian model = linear least squares.

Regularized Linear Regression.

⚠ Inversion of a covariance matrix.

Stability

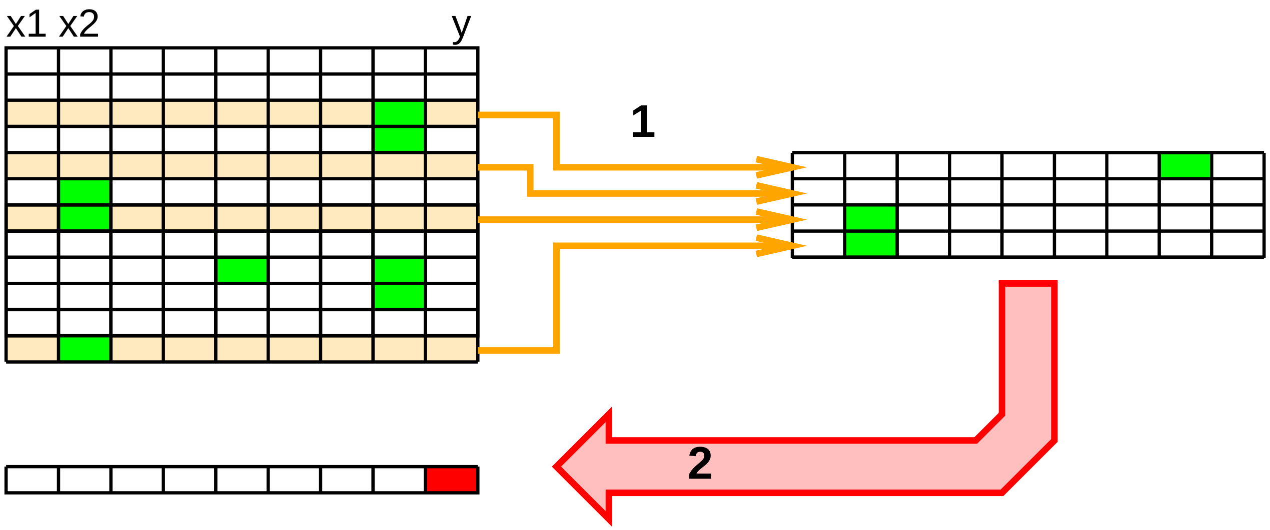
Limited computer accuracy causes numerical instability.
Add ε on the diagonal to enhance stability

Statistical Regularization

Covariance matrix are rough approximations.
Inversion is not stable.
Add sizeable positive values λ on the diagonal

This is a special case of
Structural Risk Minimization [Vapnik, 74].

Locally Regularized Regression.

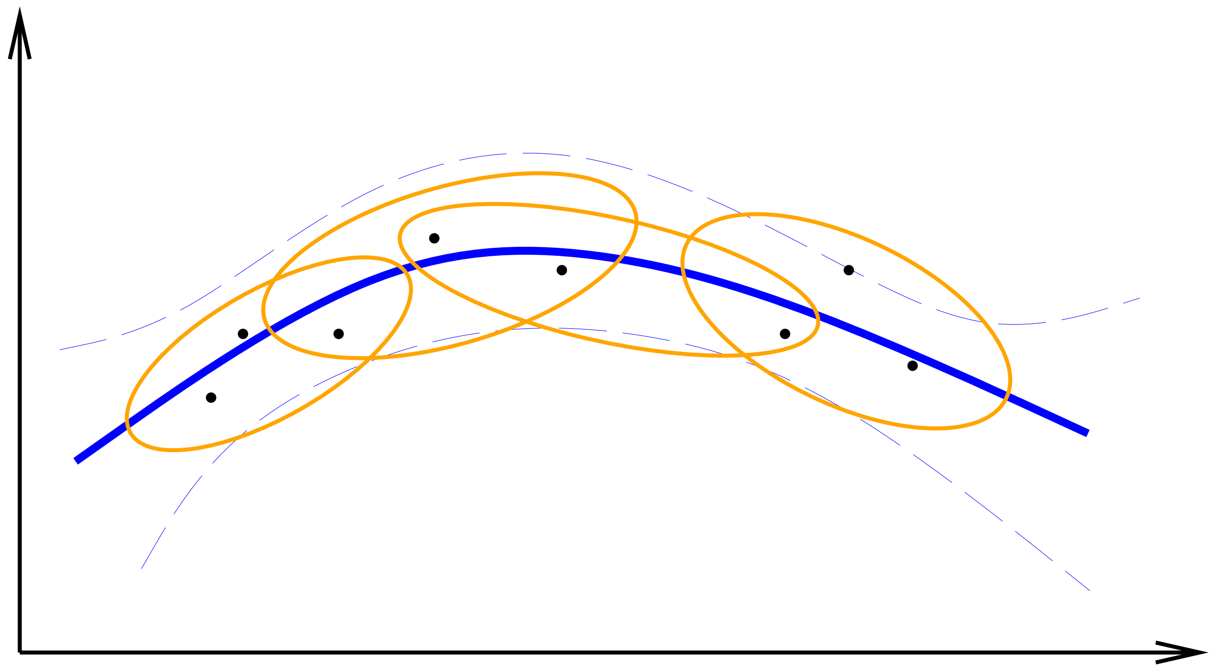


1- Select K training rows close to query

2- Apply Regularised Linear Regression.

⚠ K small $\Longrightarrow$ Larger λ.

Locally Regularized Regression.



⚠ Non linear model with a price in computing time.

TL/Prevision models.

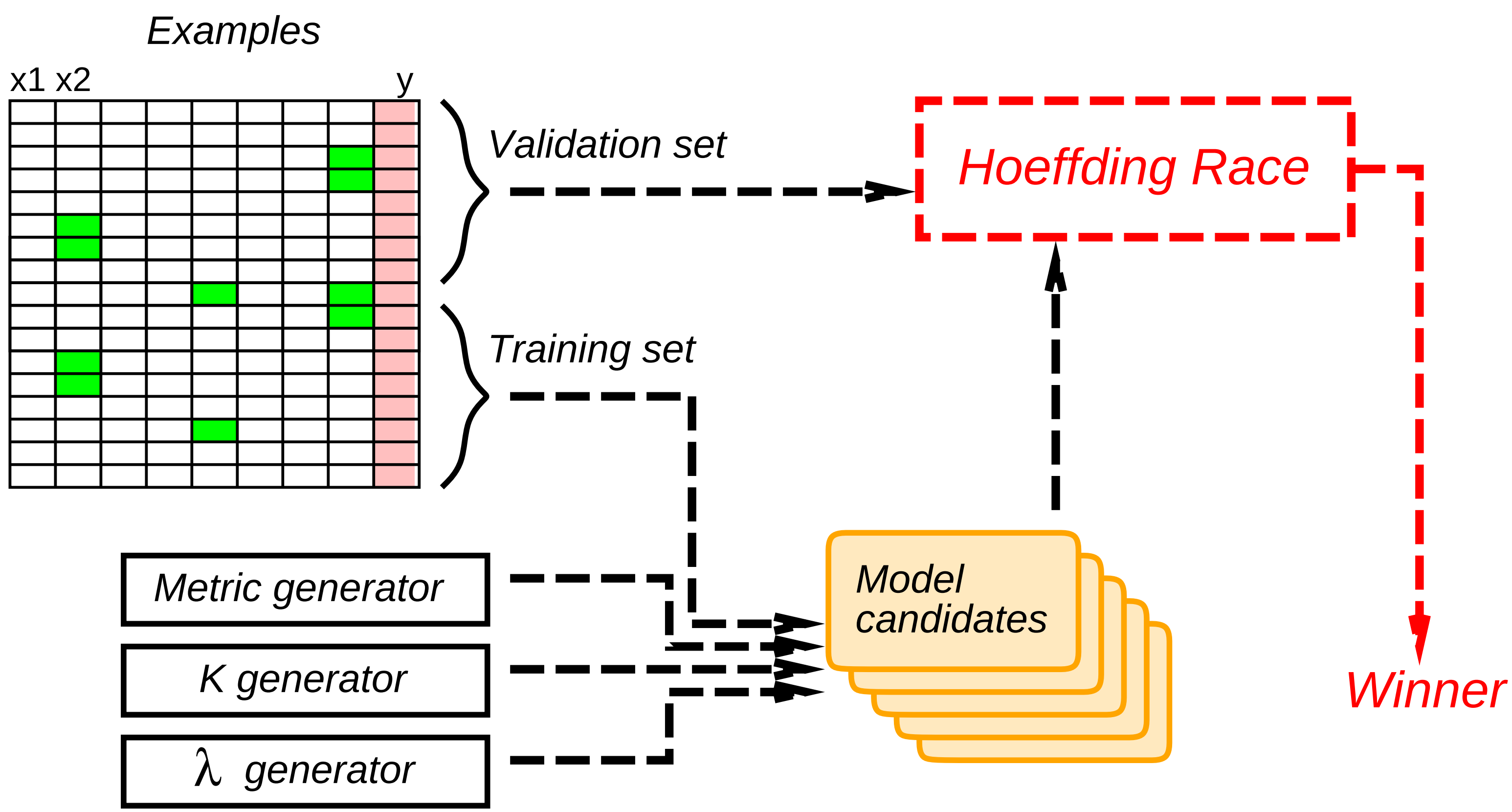
- SQL fragments.

- A set of training rows.

- A distance on rows (to select the K closest.)

- Value for K and λ.

TL/Prevision model creation.



Case study: Highway traffic forecast.

Daily and hourly traffic forecast.

Inputs:	traffic history,
	calendar and holidays,
	weather forecast,
	segments
Usage:	staffing. (tolls, security, . . .)
	maintainance.

Case study: Highway traffic forecast.

Where:	Autoroutes du Sud de La France (SOPHIA) Cofiroute (PTCT)
Results:	About 30% less error than conventional methods. Number of “large” errors divided by 3.

Case study: Highway traffic forecast.

Able to use hundreds of correlated inputs.

- holiday
- day of the week
- 1 day before/after holiday
- 2 days before/after holiday
- week-end markers
- boolean combinations of the above

Contrast with conventional methods

Smart data encoding is required.

- number of days since last “special” day.
- number of days until next “special” day.

Case study: Call center management.

Call volume $\Longrightarrow$ Staffing

[Erik Marcadé, Francoise Fogelman Soulié, Atos, 1995?]

Larger application reveals problems.

- Scalability of implementation.
- Monolithic software.
- Architecture too closed.
- SQL nightmares.

Consulting vs. Product.

Make a product, not a consulting business.

$\Longrightarrow$ Indirect road to market.
Get big solution companies to distribute.

We could only get small/medium companies.

- Lack of experience.
- Long decision cycles.
- Neuristique was too secretive.
- Lack of evangelism.
- Lack of resources.

Wrap-up.

Neuristique became a hobby company

$\Longrightarrow$ Lottery ticket company

$\Longrightarrow$ Endless arguments

$\Longrightarrow$ Counter-productive environment.

TL/Prevision technology has eventually been licensed to Kxen Inc. (www.kxen.com).