Uncertainty versus Decisions

Some (false) dichotomies between Astrophysics and Machine Learning

Roban Hultman Kramer

Hightable

http://hightable.com

a part of Gerson Lehrman Group

http://gersonlehrmangroup.com

roban@astro.columbia.edu
http://roban.github.com/

Astrophysics

vs.

Machine Learning

Uncertainty is everything

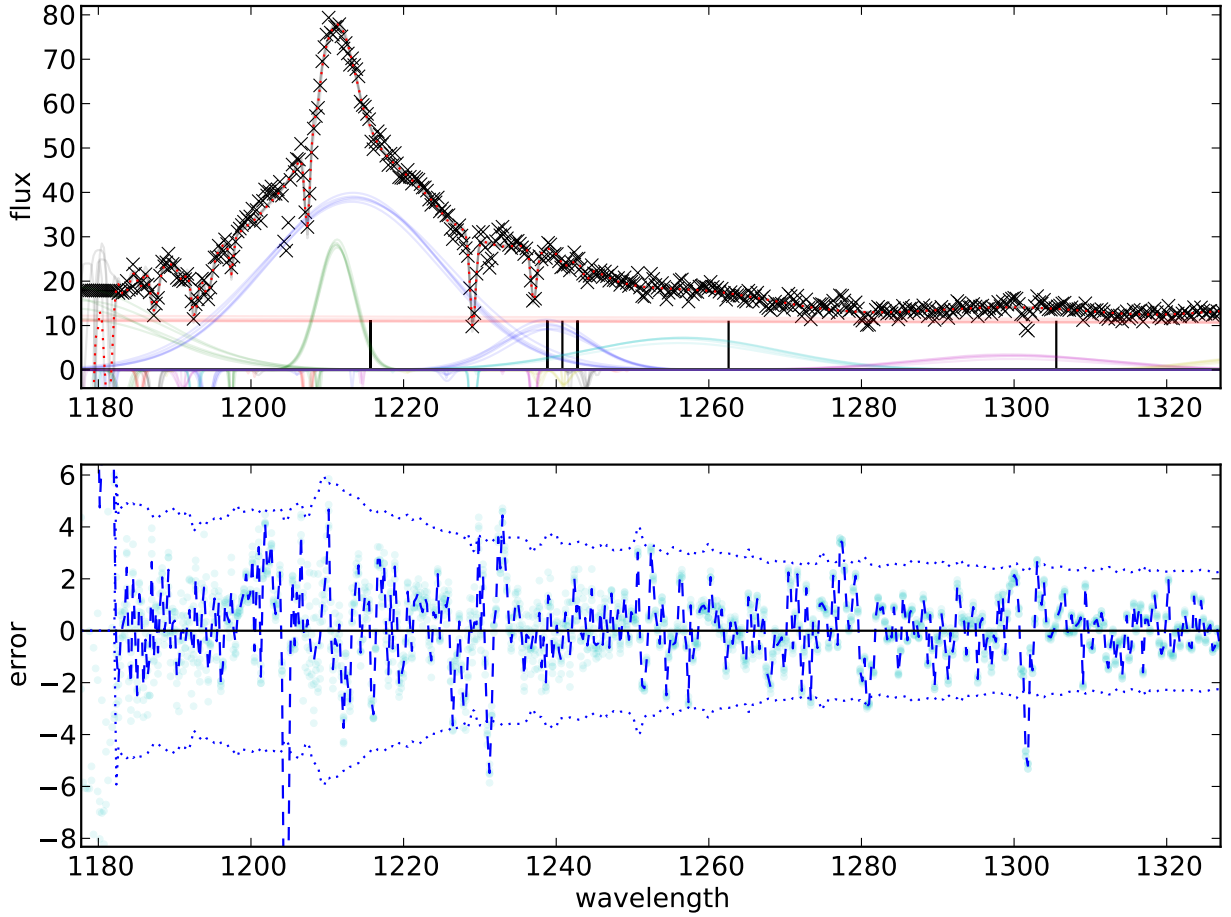
Decisions are everything

Constraining Parameters

Making Predictions

Astrophysics

Uncertainty is everything



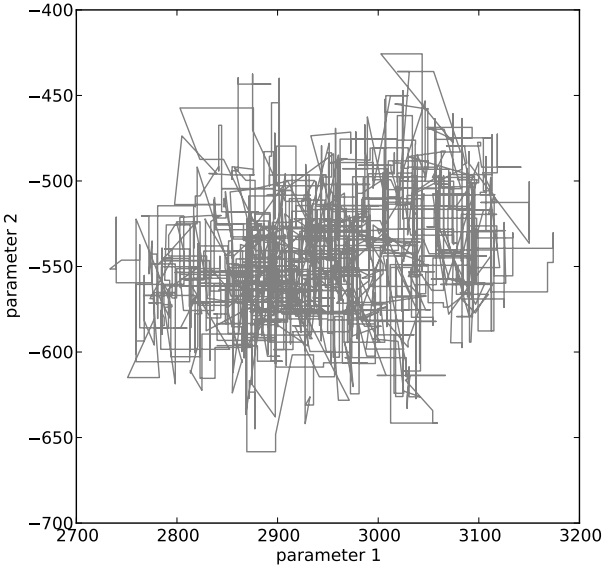
The error is as important as the measurement!

Astrophysics

Uncertainty

Example: MCMC

exploring parameter space



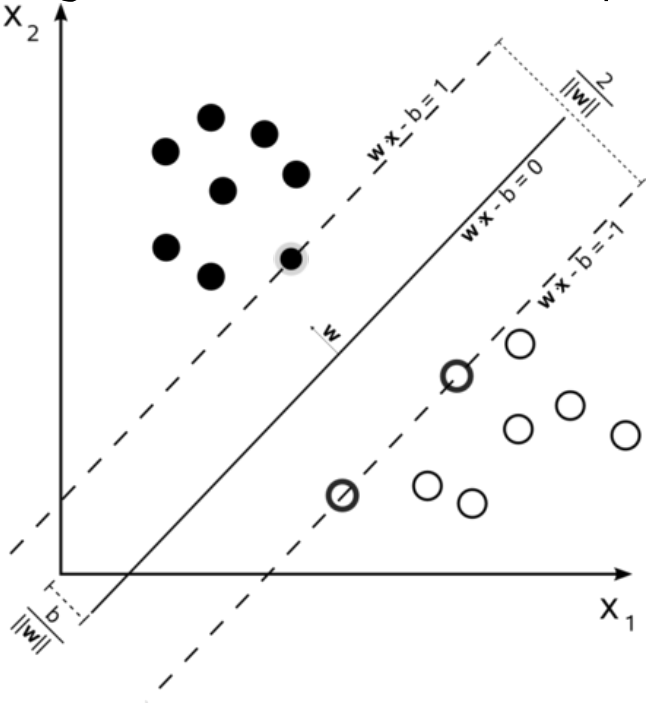
vs.

Machine Learning

Decisions

Example: SVM

finding boundaries in feature space



Credit: Wikimedia Commons
http://en.wikipedia.org/wiki/File:Svm_max_sep_hyperplane_with_margin.png

Astrophysics

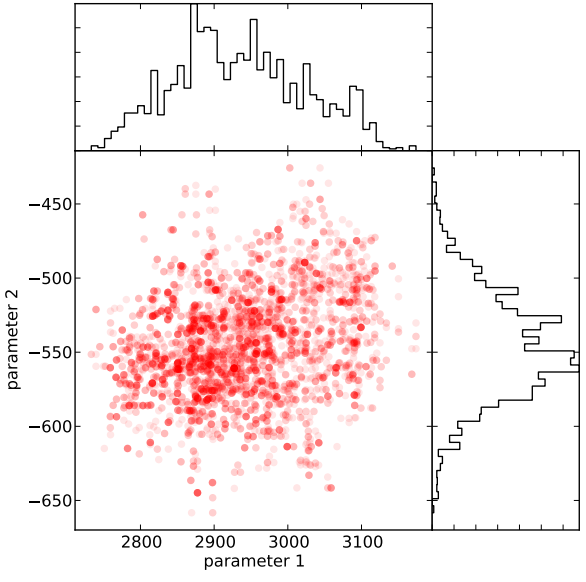
Uncertainty

Examples:

error bars

p-values

posterior distributions



vs.

Machine Learning

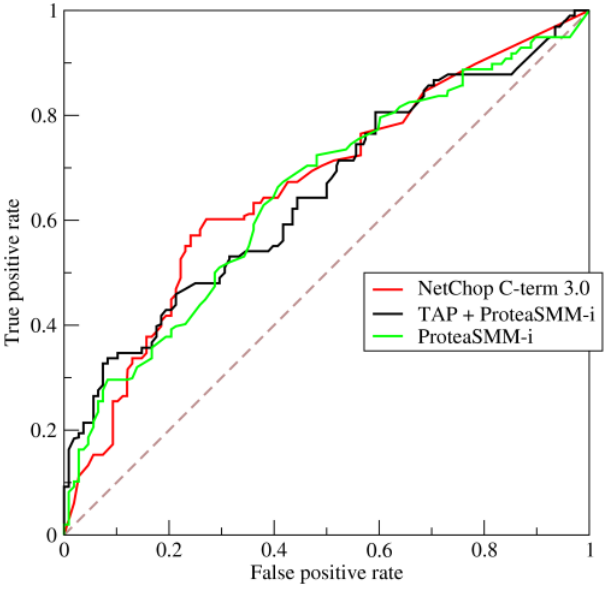
Decisions

Examples:

F_β-scores

lift

ROC curves



Wikimedia Commons `http://en.wikipedia.org/wiki/File:Roccurves.png`

Astrophysics

~~Uncertainty~~

Counter Example

Decisions

vs.

Machine Learning

Decisions

planning observations
target selection

Limited by:
telescope time
instrument budgets

(Hubble oversubscribed by $\approx 600\%$)

recommendation engines
targeted marketing

Limited by:
user attention span
marketing budgets

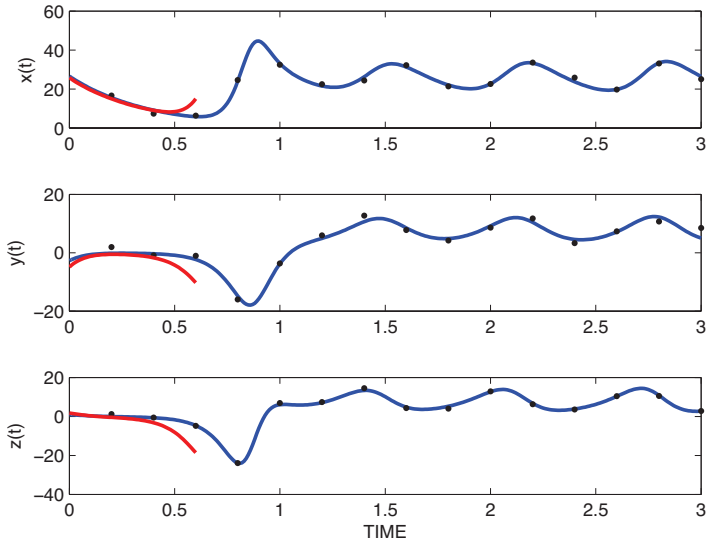
Astrophysics

Computational bottleneck: model complexity

VS.

Machine Learning

Computational bottleneck: data size



"Efficient MCMC for Climate Model Parameter Estimation: Parallel Adaptive Chains and Early Rejection" Solonen et al. *Bayesian Analysis* 7, 3 (2012), 715-736.

Astrophysics

Computational bottleneck:
model complexity

VS.

Machine Learning

Computational bottleneck:
data size

Counter Example

The Square Kilometer Array

Data Rate:
1 TB per second
after pre-processing

Computational bottleneck:
data size

"Artist's impression of the SKA dishes." Credit: SKA Organisation/TDP/DRAO/Swinburne Astronomy Productions