

Apart from economic modeling (which is already widely integrated into decision-making), what fields, such as machine learning, network analysis, agent-based modeling, dynamical systems modeling, correlational big data mining, remote sensing, data fusion, simulation of physical systems, and so forth, are most likely to contribute to this as a goal, and how will they do so?

Tools to
manipulate
data

Tools to
describe the
past/present

Tools to
predict

It is not the method as much as
what it is used for

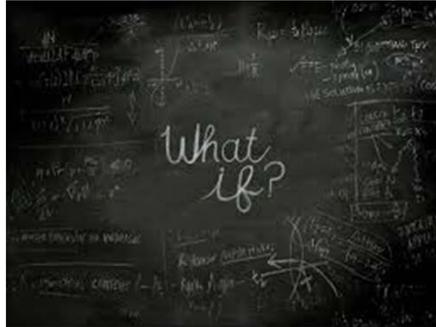
Prediction:

It is not only about the future

- 'predict' some quantity/information you need
 - SMS categories in Uganda "Youth Speaks"
 - Medical diagnosis
 - Spam & Fraud
- 'predict' alternative pasts
 - what 'would have happened ...'

counterfactuals

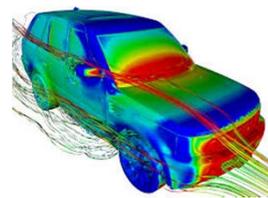
Policy and Prediction



Counterfactuals

Simulation

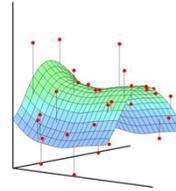
$$\begin{aligned}y(t) &= F^{-1}\left\{\frac{1}{1+4\pi^2 f^2}\right\} * F^{-1}\{G(f)\} \\ &= \frac{e^{-|t|}}{2} * g(t) \\ &= \frac{1}{2} \int_{-\infty}^{\infty} e^{-|t-\tau|} g(\tau) d\tau\end{aligned}$$



- Start with a set of equations
- Results are as good as your assumptions
- Data is used to maybe tune and validation

Predictive Modeling = Function Approximation

```
0.749380154021744 0.000590183986545028 779 20
0.798285720814701 0.00034631487664848 1962 20
0.540374430975371 0.001415737969704083 891 20
0.724142820274877 0.0011803594088993 1734 20
0.420231281165794 0.000683473608154817 1755 10
1.63577243467107 0.00229268808890358 1546 10
0.2501293492083 0.00843852618368423 1065 10
2.13146775632401 0.000785093522820787 779 10
0.815078913250702 0.00818205578664053 1473 20
0.77988940620191 0.0120380305675274 891 20
1.04225917097729 0.00314198703840882 1969 20
1.74645754395112 0.00165522594785467 703 10
0.67790040859134 0.00150284577448034 1236 10
1.05657287551612 0.0382022935590516 1912 20
0.74167007811735 0.0169101519222402 1830 20
0.83606256536671 0.000237500637072524 779 20
3.2219651310187 0.00956206776936697 916 10
0.47156025759991 0.000704493746174368 1065 10
0.851249341944925 0.05344262295082 1473 10
1.08403941502146 0.00160579161489708 1830 20
0.788523498700205 0.0432004773760777 1775 20
0.193584514913287 0.000495734011303755 586 10
```



- Start with data
- Results are as good as your data
- Little to no prior knowledge required

Potential & Limits of Predictive Modeling

Very effective in scenarios with

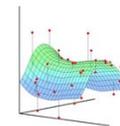
- very detailed level (individuals)
- get better and better with more data
- same setting for use as training
- probabilistic relationships

```
0.749380154021744 0.000590183986545028 779 20
0.798285720814701 0.00034631487664848 1962 20
0.540374430975371 0.001415737969704083 891 20
0.724142820274877 0.0011803594088993 1734 20
0.420231281165794 0.000683473608154817 1755 10
1.63577243467107 0.00229268808890358 1546 10
0.2501293492083 0.00843852618368423 1065 10
2.13146775632401 0.000785093522820787 779 10
0.815078913250702 0.00818205578664053 1473 20
0.77988940620191 0.0120380305675274 891 20
1.04225917097729 0.00314198703840882 1969 20
1.74645754395112 0.00165522594785467 703 10
0.67790040859134 0.00150284577448034 1236 10
1.05657287551612 0.0382022935590516 1912 20
0.74167007811735 0.0169101519222402 1830 20
0.83606256536671 0.000237500637072524 779 20
3.2219651310187 0.00956206776936697 916 10
0.47156025759991 0.000704493746174368 1065 10
0.851249341944925 0.05344262295082 1473 10
1.08403941502146 0.00160579161489708 1830 20
0.788523498700205 0.0432004773760777 1775 20
0.193584514913287 0.000495734011303755 586 10
```

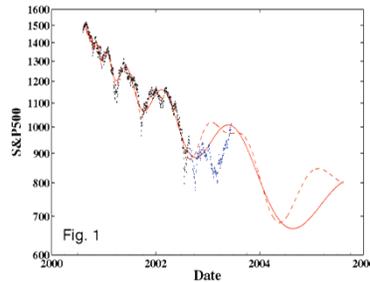


NOT Suitable for scenarios with

- rare events: Katrina
- few instances: Countries
- totally new circumstances: ObamaCare



Predictive Modeling & Policy



- Most of Policy is about Change: Extrapolation
 - with respect to time
 - with respect to circumstances

Will future decision makers in government and industry have better predictive tools when they consider policy alternatives?

- predictive tools have not changed much
- data has changed significantly
- use predictive modeling with caution
 - is PM even applicable?
 - do you have the right data?
 - quality control is extremely difficult
 - model is only as good as the person who build it

How will advances in the modeling and simulation of social and collective behaviors lead to better quantitative policy analysis?

Evaluating policy impact:

Causal Analysis

A/B Testing

- practical concerns
- amoral & unlawful with human subjects



Causal analysis on observational data

- Using targeted maximum likelihood to estimate causal impact (some form of PM)
- Can be done ex-post for different questions
- Use predictive modeling to control for **confounding**
- Data has to be 'rich' and cover all combinations of confounding and treatment

$$E[Y_{\text{policy}}] - E[Y_{\text{no policy}}]$$

