AI & Data Science, Friend or Foe?*

SDP 2024 Annual Conference

John Mark Agosta

San Jose State U

* With acknowledgements to Robert Horton, and John Mount at WV labs.

Outline

1. Why automate decisions - examples

2. What AI is – the unreasonable effectiveness of data

3. Merging value and predictive modeling

4. An “AI” as an opponent

5 million $1 decisions or 1 $million decision?

- Software automation means doing a simple thing with a computer many many times
- An automated decision process v/s analyzing a one-time ″strategic″ choice.
- Creating an AI calls for a “A Decision to Decide”
- Should these be automated?
  - Classifying insurance claims to refer them to legal staff?
  - Offshoring Xray classification?
  - Alerting failures for software roll-outs?

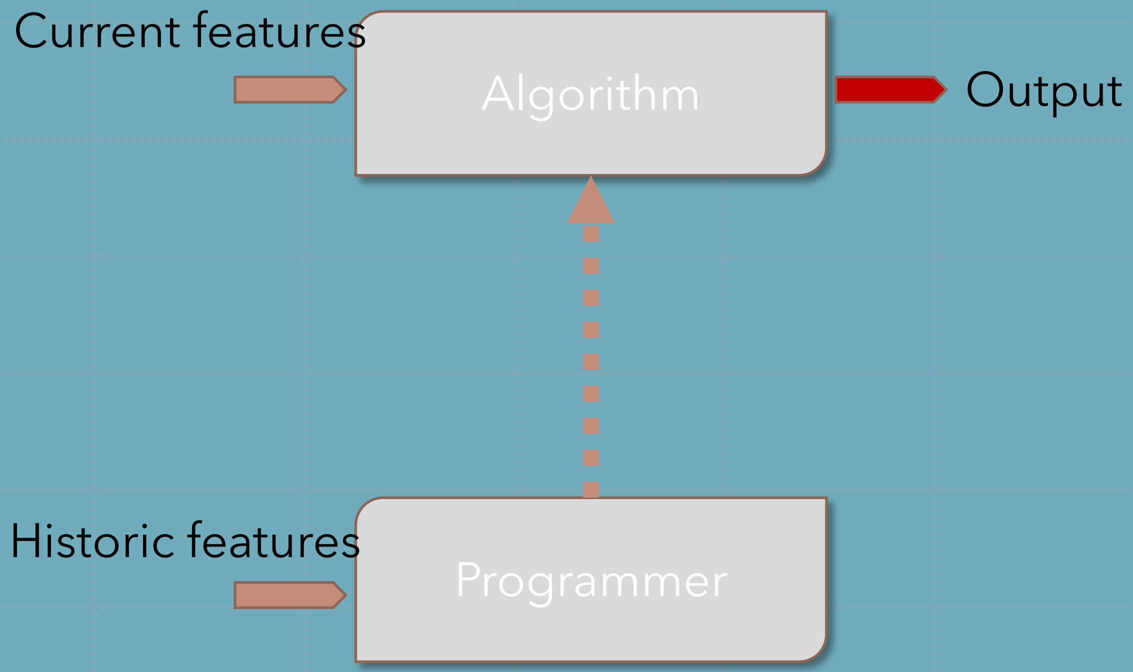
"AI" – rational choice or just "common sense"

- The goal of the discipline has changed, as techniques have evolved.
- AI began as the quest to create *rational agents*.
- A basis in rationality that is shared with Decision Analysis
- Language models embody "common sense" but not sound reasoning. They are design accelerators, but cannot be deciders.
- Keep eyes on the prize – rational action by building compound models.

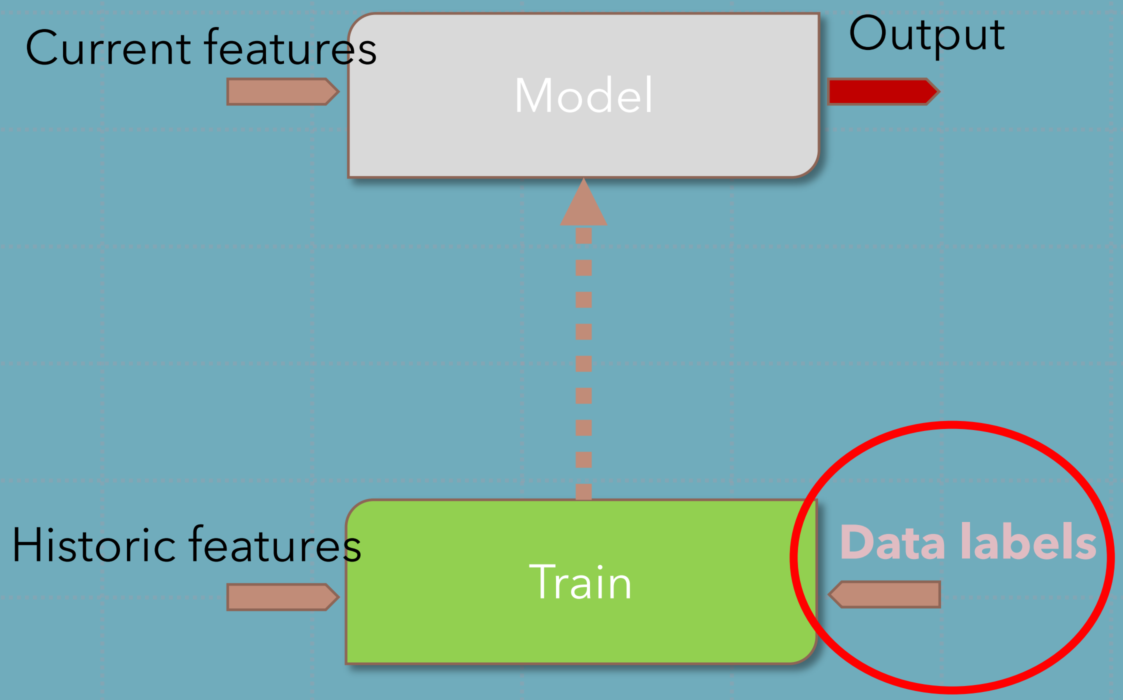
The unreasonable effectiveness of big data.

- Combining AI and Decision Analysis benefits from
- the power of data-driven approaches (and the flourishing of "Data Science" methods)
- with the rigor of Decision Quality
- The result is *alignment* between data models and organization's goals.

Manual coding v/s Machine Learning

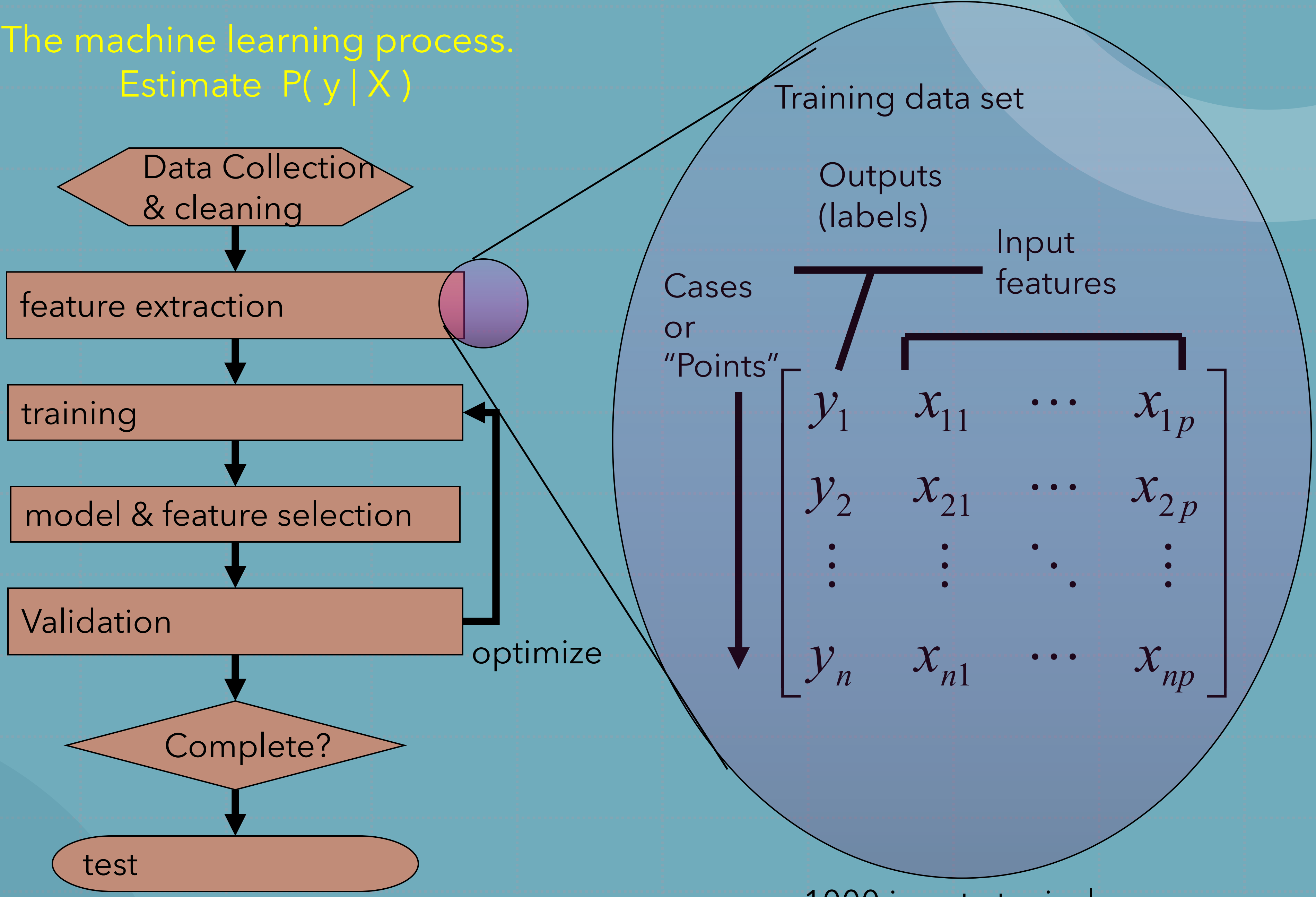


In traditional programming the programmer derives the algorithm



In Supervised Machine Learning, the program is “learned” from labels applied to historic data

The machine learning process.
Estimate P(y | X)



$$\begin{bmatrix} y_1 & x_{11} & \cdots & x_{1p} \\ y_2 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_n & x_{n1} & \cdots & x_{np} \end{bmatrix}$$

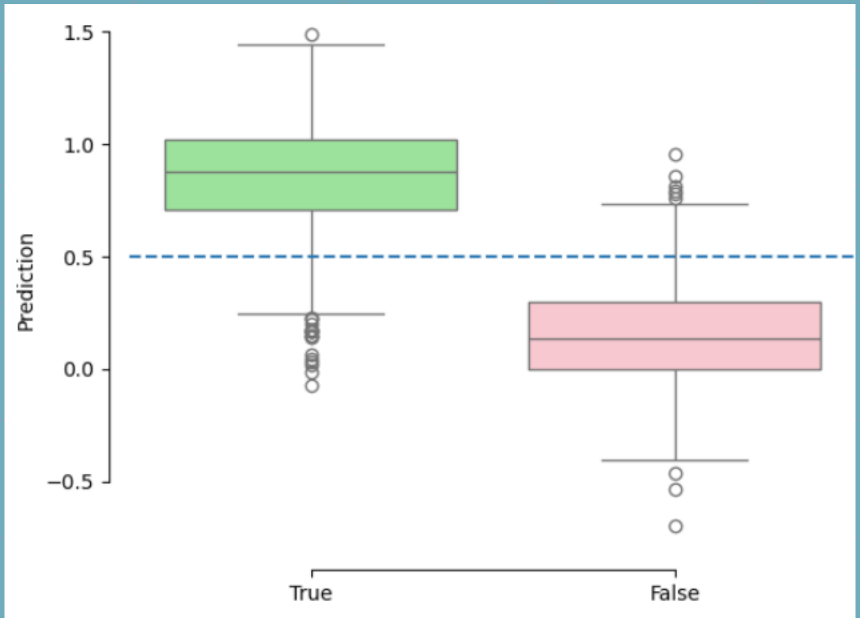
p = 1000 is not atypical, sometimes with p > n

Machine Learning: AI that generalizes from data.

- Fit a function, *f(X)*, to data.

	review	sentiment	vector
0	So there is no way for me to plug it in here i...	0	[0.08027008920907974, -0.04396028444170952, -0...
1	Good case, Excellent value.	1	[-0.009648566134274006, 0.10622689127922058, 0...
2	Great for the jawbone.	1	[-0.07081733644008636, 0.07361650466918945, 0....
3	Tied to charger for conversations lasting more...	0	[-0.0739610344171524, 0.06734045594930649, 0.0...
4	The mic is great.	1	[-0.09819574654102325, 0.010798277333378792, 0...
...	...	...	...
743	I just got bored watching Jessice Lange take h...	0	[-0.02032049186527729, -0.07333985716104507, 0...
744	Unfortunately, any virtue in this film's produ...	0	[-0.025788182392716408, 0.007497682701796293, ...
745	In a word, it is embarrassing.	0	[0.026193976402282715, 0.022175997495651245, 0...
746	Exceptionally bad!	0	[-0.027648691087961197, -0.004298456944525242,...
747	All in all its an insult to one's intelligence...	0	[0.00193118117749691, 0.08036555349826813, 0.0...

F(x)



Merging value and predictive modelling

- There is a challenge getting business people on the same page with their Data Scientists.
- Blind pursuit of model accuracy doesn't cut it
- Bayes networks / Influence Diagrams are known to both communities.
- Start with a decision model!

A combined value and prediction model

This influence diagram integrates

1. Value model *U(d,y) of*

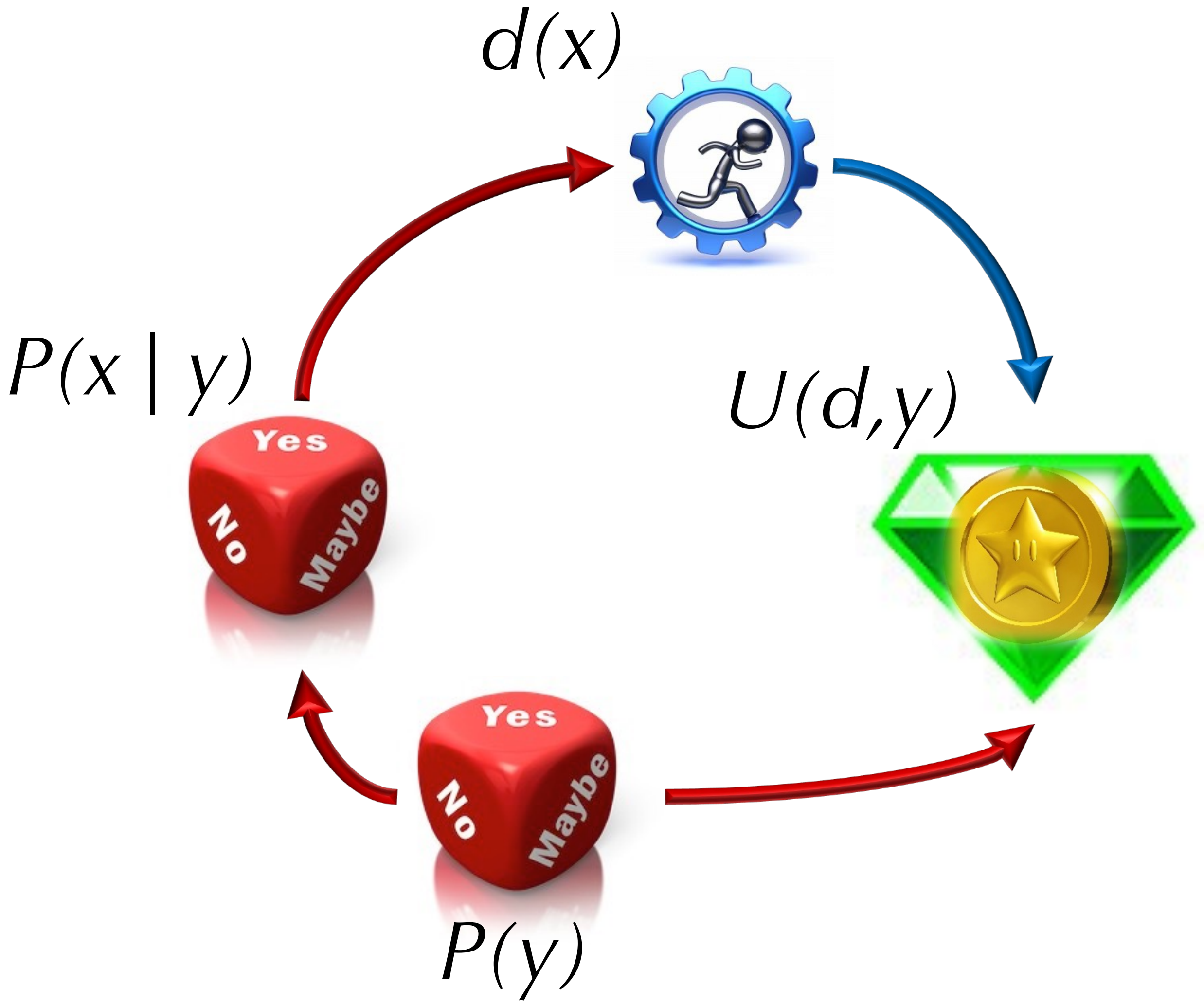
- decision *d*

- state uncertainty *y*

2. Predictive model *P (x | y)*

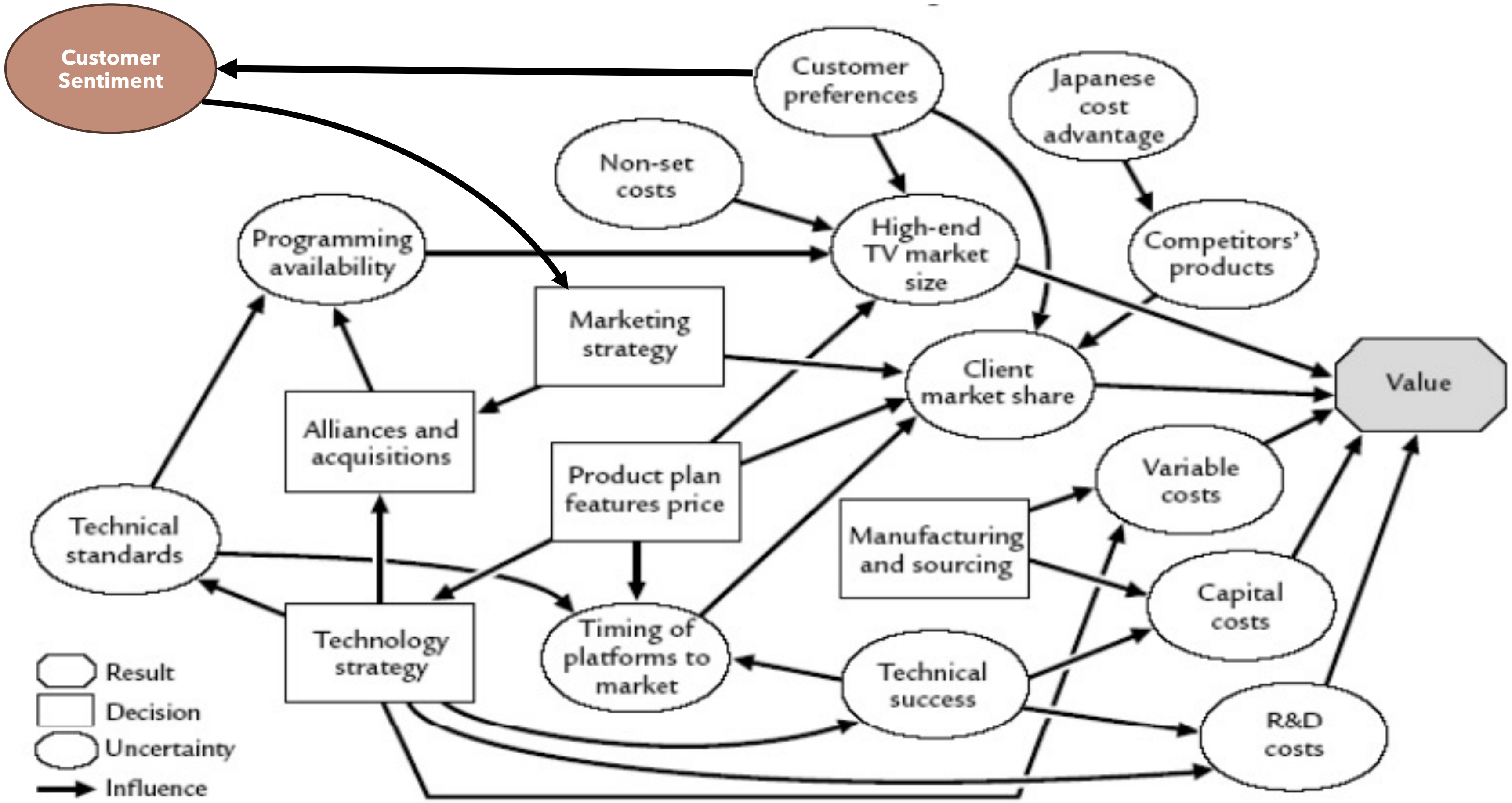
- observation x

- state uncertainty y



Extending a value model with observed data

The observable variable "customer sentiment" is learned from data about customer preferences and informs "marketing strategy" choice.



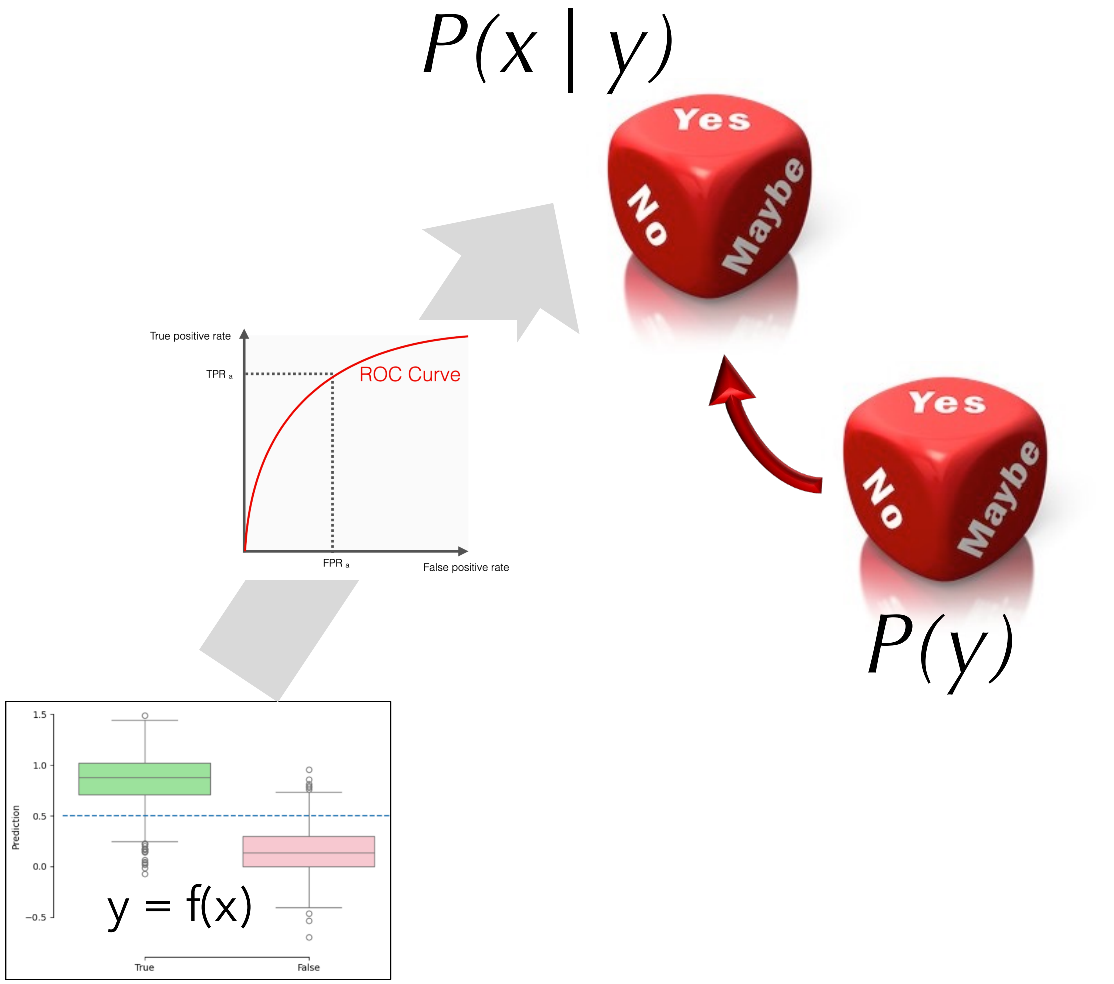
In Matheson: Decision Analysis = Decision Engineering 206 Tutorials in Operations Research, © 2005 INFORMS

But how do you get probabilities from an ML model?

- The influence diagram needs these 2 likelihoods:
- Sensitivity: P(positive | true) = TPR, and
- Specificity: P(negative | false) = 1 - FPR
- *To convert the model score to a "calibrated" likelihood use an ROC curve.*

"Flipping" the predictive model distribution.

- The score generated by the model, y = f(x) needs –
- * to be converted to a calibrated probability
- * to have the "data imbalance", an implicit prior, removed
- * for a binary observation, to be expressed as a likelihood.
- These likelihoods fill in the conditional probability table for the observation variable.
- With this one can calculate a value of information for deploying the model.

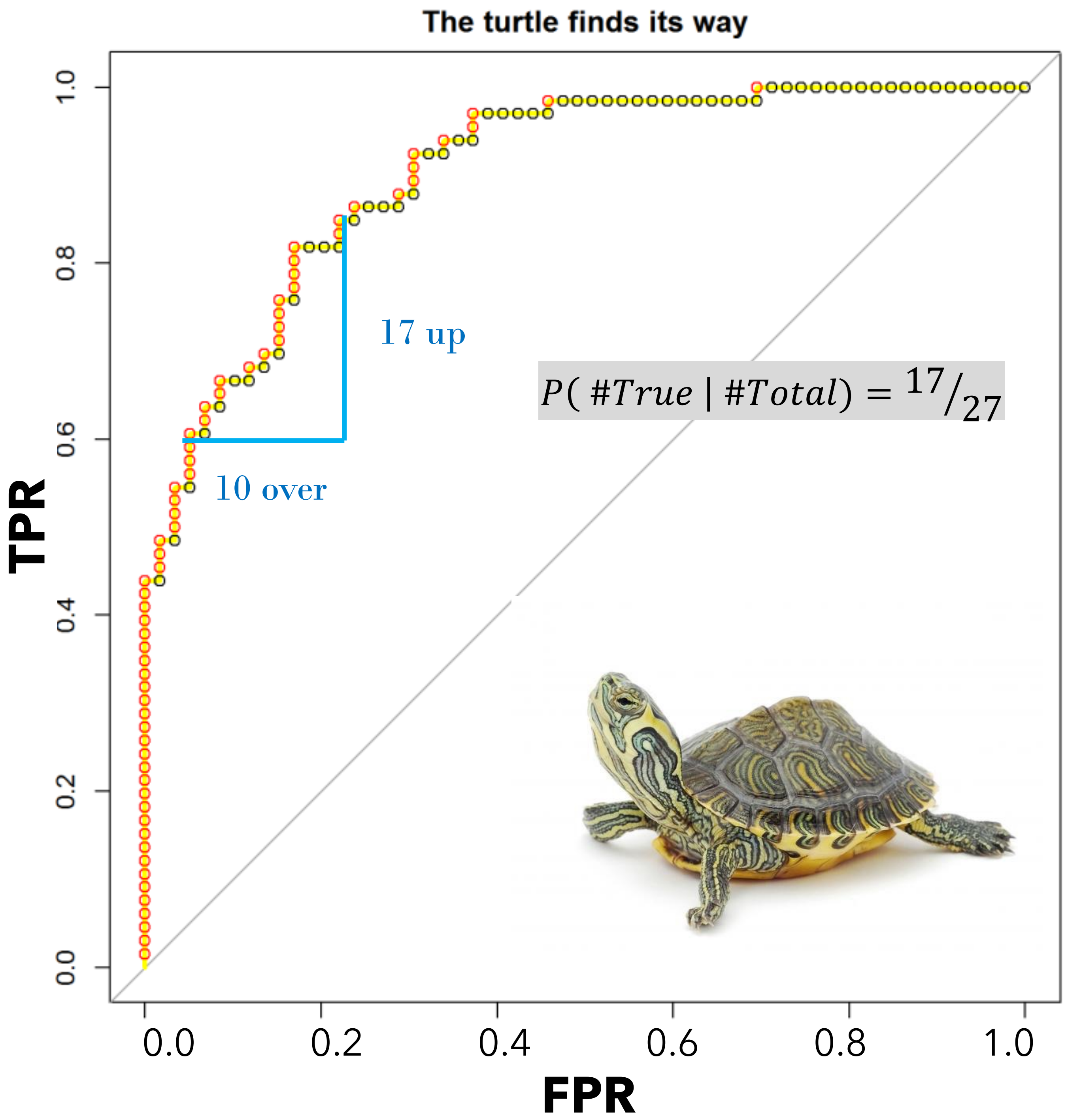


Calibration with the ROC Curve:

- sort test cases by their score from the model
- march along the sequence, stepping up for positives and right for negatives

This is the same as scanning across possible cutoff threshold values.

The slope of the curve shows the concentration of positives.



What is the *operating point*?
Setting an ROC threshold at the point that maximizes Utility

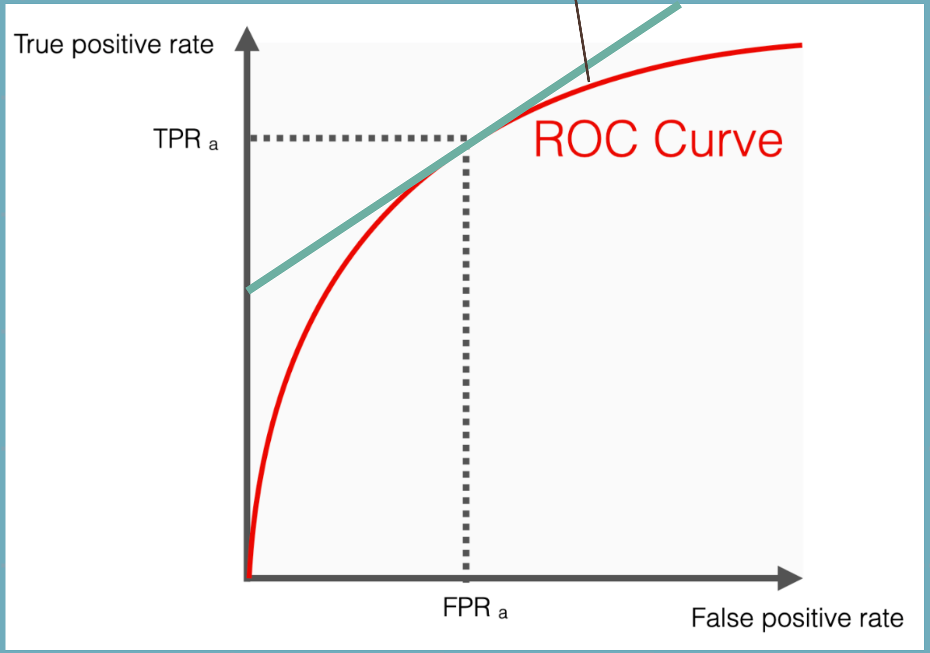
```
(TP_value * tpr * π) +               # accepted customers
(FP_value * fpr * (1 - π)) +         # un-deserving
(TN_value * (1 - fpr) * (1 - π)) + # rejected
(FN_value * (1 - tpr) * π)           # lost customers
```

π : prior probability (prevalence) for positive sentiment

`TP_value`, `FP_value`, `TN_value`, `FN_value`: values (or costs) assigned to TP, FP, TN and FN cases

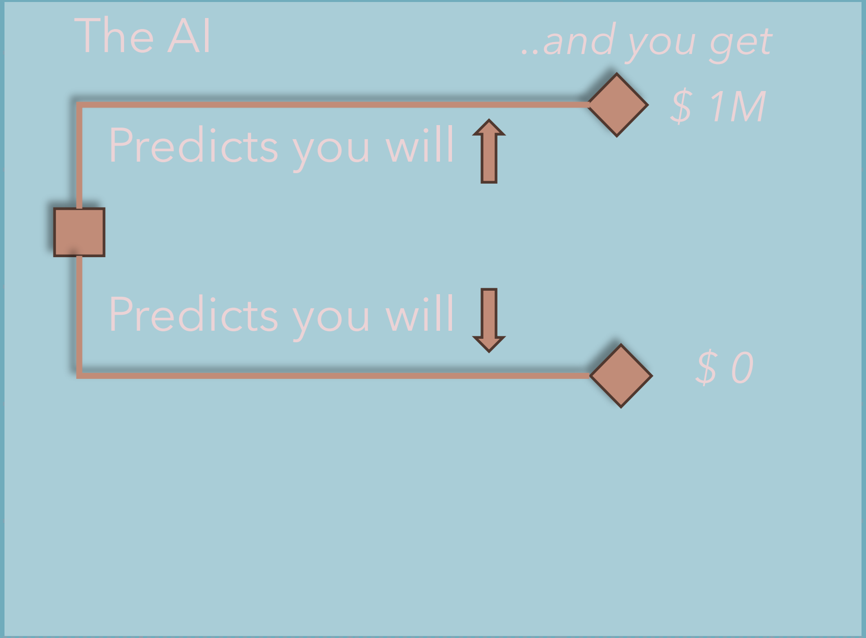
`tpr`, `fpr`: true positive and false positive rates from the operating point

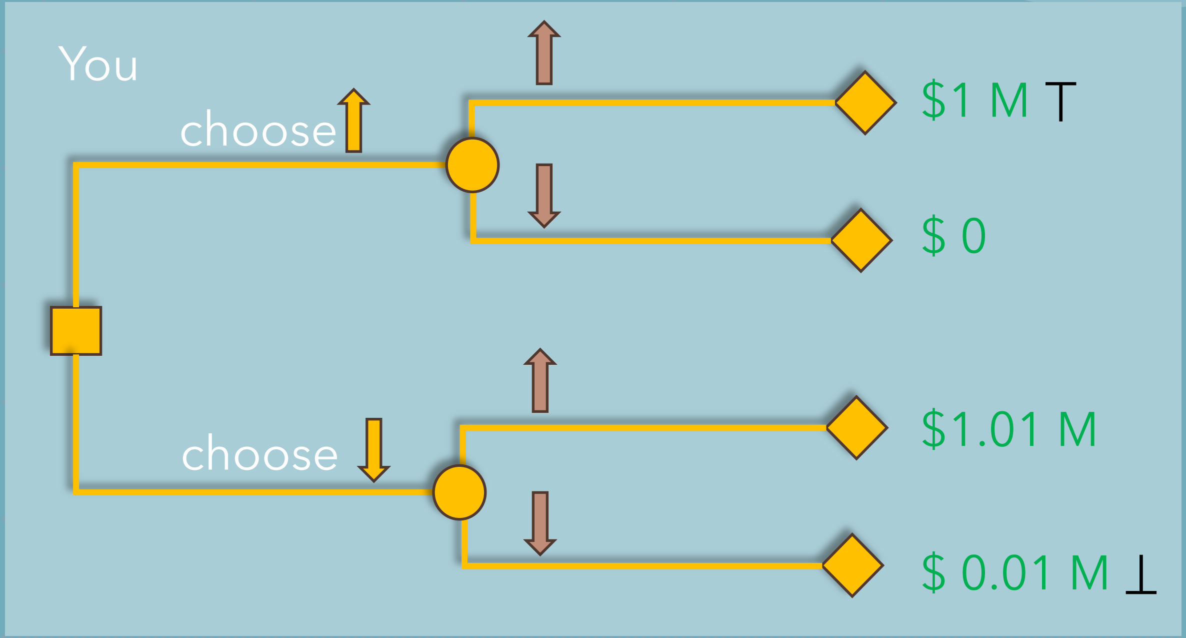
$\frac{TP\_value}{FP\_value}$



When an AI participates in a decision, it may be a friend or foe.
It pushes the boundaries of individual rationality.

Playing against an algorithm:
What do you do if an AI could predict your actions?
("Newcomb's paradox. Nozick, 1969)





Act rationally, and you get ⊥, when you could have gotten ⊤!

How would an AI actually predict?

Newcomb's paradox (and the extensions to conventional decision theory it inspires) raises current questions when facing off with an AI.

- Although we do know the data the AI was trained on.
- We don't know on what basis (rationally or otherwise) it acts.

Thanks to machine-learning algorithms, the robot apocalypse was short-lived.

Conclusion:

- Join with Data Scientists to bring
- the principles of data quality
- for formulating data-driven automated decision-making.

Citations

C. Spetzler, H. W. (2016). *Decision Quality: Value Creation from Better Business Decisions .* Wiley.

Caruana, R. a.-M. (2004). Data Mining in Metric Space: An Empirical Analysis of Supervised Learning Performance Criteria. *KDD'04*. ACM.

Fienberg, S. E. (1982). The Comparison and Evaluation of Forecasters. *Proceedings of the 1982 I.O.S. Annual Conference on Practical Bayesian Statistics. 32 No. 1/2*, pp. 12-22. NYC: Journal of the Royal Statistical Society. Series D (The Statistician).

G. Allison, P. Z. (1999). *Essence of Decision 2nd Ed.* Pearson.

Kahneman, D. (2011). *Thinking Fast, Thinking Slow.* Farrar, Straus & Giroux.

Merkhofer, M. W. (2012). *Decision Science and Social Risk.* Springer.

Naeini, M. P. (2015). Binary Classifier Calibration Using a Bayesian Non-Parametric Approach. *Proceedings of the SIAM International Conference on Data Mining.*, 208–216.

Platt, J. C. (1999). *Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods* (Vol. Advances in Large Margin Classifiers). (P. B. A. Smola, Ed.) Cambridge, MA: MIT Press.

David H. Wolpert and Gregory Benford, *The lesson of Newcomb's paradox,* Synthese Vol. 190, No. 9 (June 2013), pp. 1637-1646 (10 pages)

Decision Quality applied to Automation

John Mark Agosta

Mapping The Decision Quality Framework to Data Science Applications

The principles developed for assuring a rational approach to strategic decisions can apply to effective applications of data science. Each Decision Quality principle has an analog that applies to autonomous systems.

DQ Principle	Decision Analysis of Strategic choices	Automated Decision Systems
Decision Basis	Find an alignment among decisions, outcomes and uncertainties within the problem framework, by building agreement on the problem and relaxing constraints where possible.	Identify the underlying domain causal model, drawing on techniques such as structural equation modeling, systems engineering, control block diagrams, and influence diagrams.
Creative Alternatives	Avoid reducing the choices just to advocacy for or against a position, by finding creative alternatives that meet objectives.	Identify those variables in the model that can be controlled, and the sequence they occur in. Note that the data's label set and the decision alternatives are not the same thing.
Relevant Information	Consider the sensitivity of choices to information, to remove irrelevant details, or to guide further information collection.	Apply feature selection driven by the loss (value) function, to compute a true value of information beyond the mutual information surrogates currently in use.
Comprehensive Objectives	Canvas stakeholders to make tradeoffs explicit and quantifiable.	Train the model against loss functions derived from business values, a, not just operational KPIs, and show robustness of results among different loss functions, not just measures of model accuracy.
Sound Reasoning	Use techniques to counter biases, flatten priors and ground subjective perceptions. Total analysis efforts should not compare to the value of the outcomes at risk. (Don't over analyze.)	Model risk reduction, not just accuracy, by calibrating probability outcomes. Engage in good model management, realizing that environments are rarely stationary. Include operational costs & latency if comparable to value achievable .
Commitment to Implement	"If the analysis is correct, would you take its recommendation?"	Anticipate deployment and production engineering challenges, and the strategic cultural change entailed. Match level of automation to value achievable.

AI & Data Science, Friend or Foe?

Abstract:

Is it more valuable to make millions of 1 dollar decisions or one 1 million dollar decision? Automating decisions in software opens a new frontier in Decision Science. It is another proving ground for Decision Quality (DQ).

Don't get confused by what "AI" was and what has it now become. AI's origins have common "credential" with DQ in the pursuit of rationality. The siren song of current generative AI language models and their appearance of "common sense" do not replicate rational decision making.

This talk explains how the value modelling familiar to Decision Analysts can be integrated to complement Data Science's predictive modelling, to solve the so-called "alignment" problem.

Then speculatively, I broach how this new world of automation raises an interesting moral question in light of Newcomb's paradox.