Overview of Machine Learning: Opportunities and Challenges

John D. Lee and Linda Ng Boyle
10/1/2018

HF Considerations for ML and ML Considerations for HF

– Human-centered ML represents a looming societal need

– ML deployment sometimes neglects design and systems thinking

– ML mindset for data analysis might save us from the p-value ritual

– ML and Human-centered tradeoffs in model development

Increasing Centrality of Machine Learning in System Design

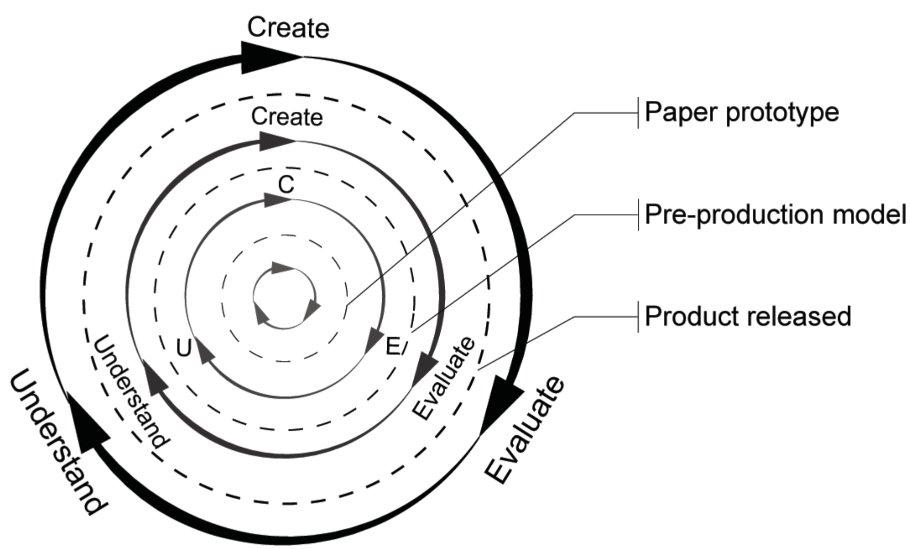
Human implications of increasingly powerful algorithms

Design Thinking and Systems Thinking for Machine Learning

– Design thinking: Empathy and understanding for person-centered solutions

– Systems thinking: 5 whys? 5 whats?

– Sea-level and C-level support for success



4/11

Machine Learning versus Inferential Data Analysis

– Linear regression and logistic regression as machine learning?

ML mindset	Traditional inferential mindset
Prediction	Inference
Cross validation and prediction error	Hypothesis testing and p values
Practical significance	Statistical significance

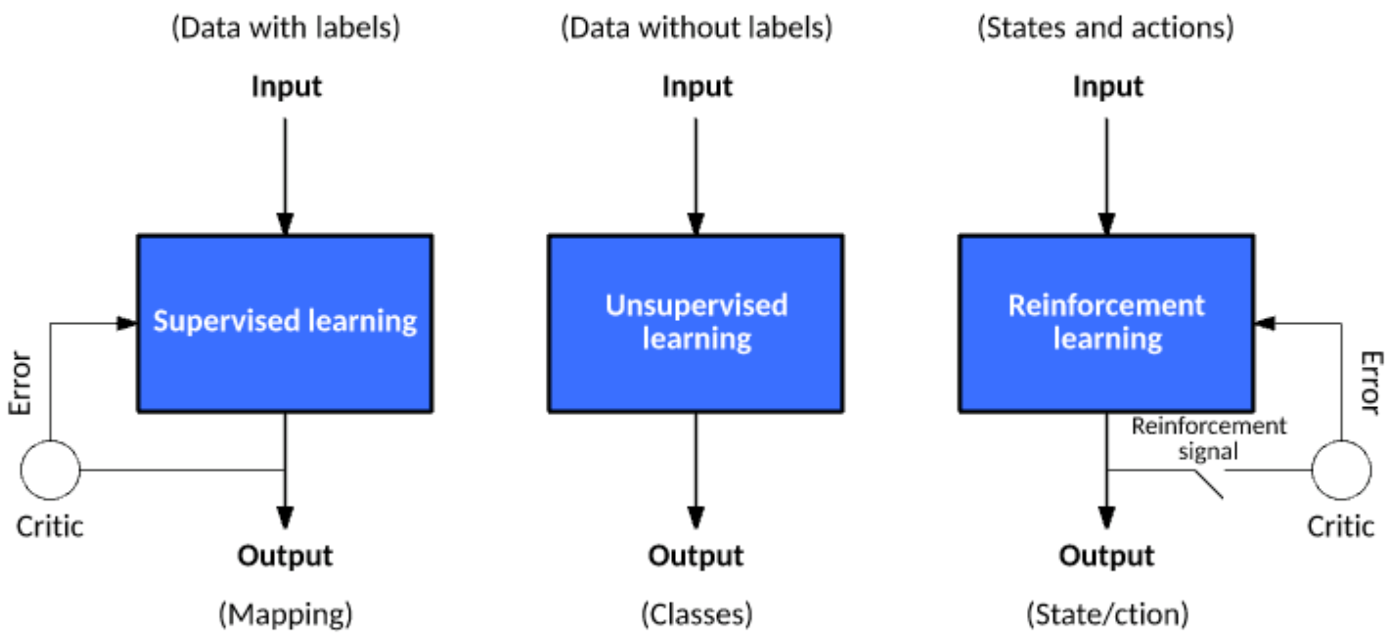
5/11

Types of Machine Learning

– Supervised learning: Predicting a known output

– Unsupervised learning: Identifying unknown patterns or clusters

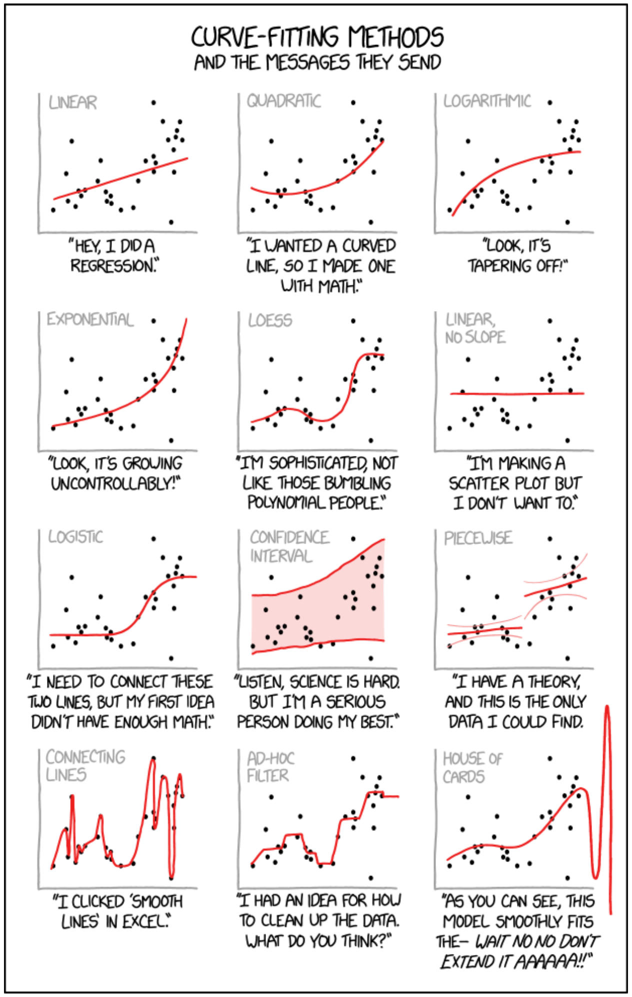
– Reinforcement learning: Learning through interactions with a system



https://www.ibm.com/developerworks/library/cc-models-machine-learning/index.html

6/11

Supervised Learning: More than curve fitting?



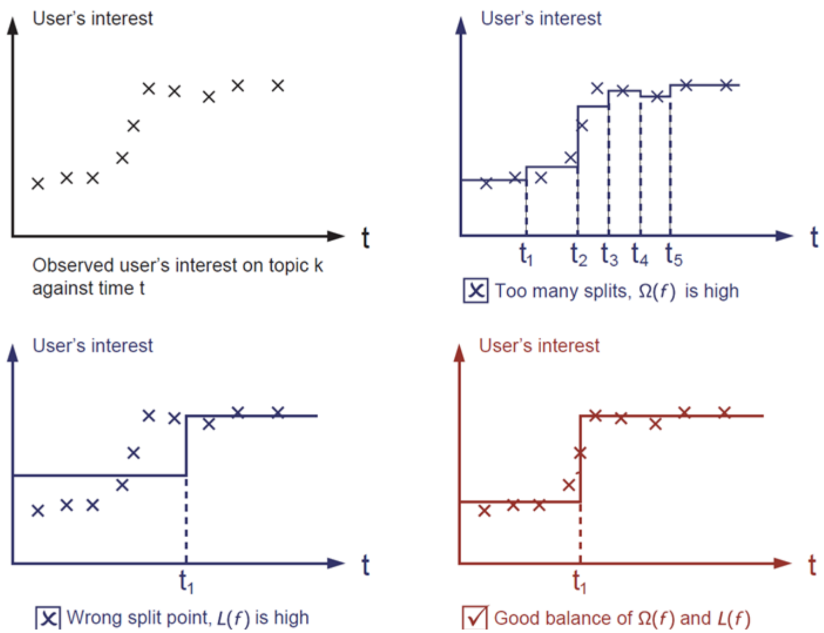
Essential Tradeoffs in ML Design

Variance-bias tradeoff of a model θ

$\Omega(\theta) =$ Model complexity, ideally small variance and good generalization

$L(\theta) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 =$ Model error, ideally small bias and precise predictions

Objective function$(\theta) = \Omega(\theta) + L(\theta)$



https://xgboost.readthedocs.io/en/latest/index.html

8/11

Essential Tradeoffs in ML Design

Human-centered tradeoffs

– Different errors: Cost of a miss may differ than a cost of a false alarm

$L(\theta) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ Positive errors equally problematic as negative errors??

– Trust depends on more than AUC (Area Under the Receiver-Operator Curve): Hard hits don't compensate for easy misses

– People might value understandable models more than precise models

$\Omega(\theta)$ Model complexity does not typically reflect perceived complexity

Essential Tradeoffs in ML Design

Hyperparameters

– Hyperparameters are parameters of the ML algorithms that govern how they fit the data – Some hyperparameters allow a better model fit – Some hyperparameters allow a more robust model – Some hyperparameters allow for a more understandable model

HF Considerations for ML and ML Considerations for HF

– Human-centered ML represents a looming societal need

– Design and systems thinking is sometimes neglected with ML deployment

– ML mindset for data analysis might save us from the p-value ritual

– ML and Human-centered tradeoffs in model development