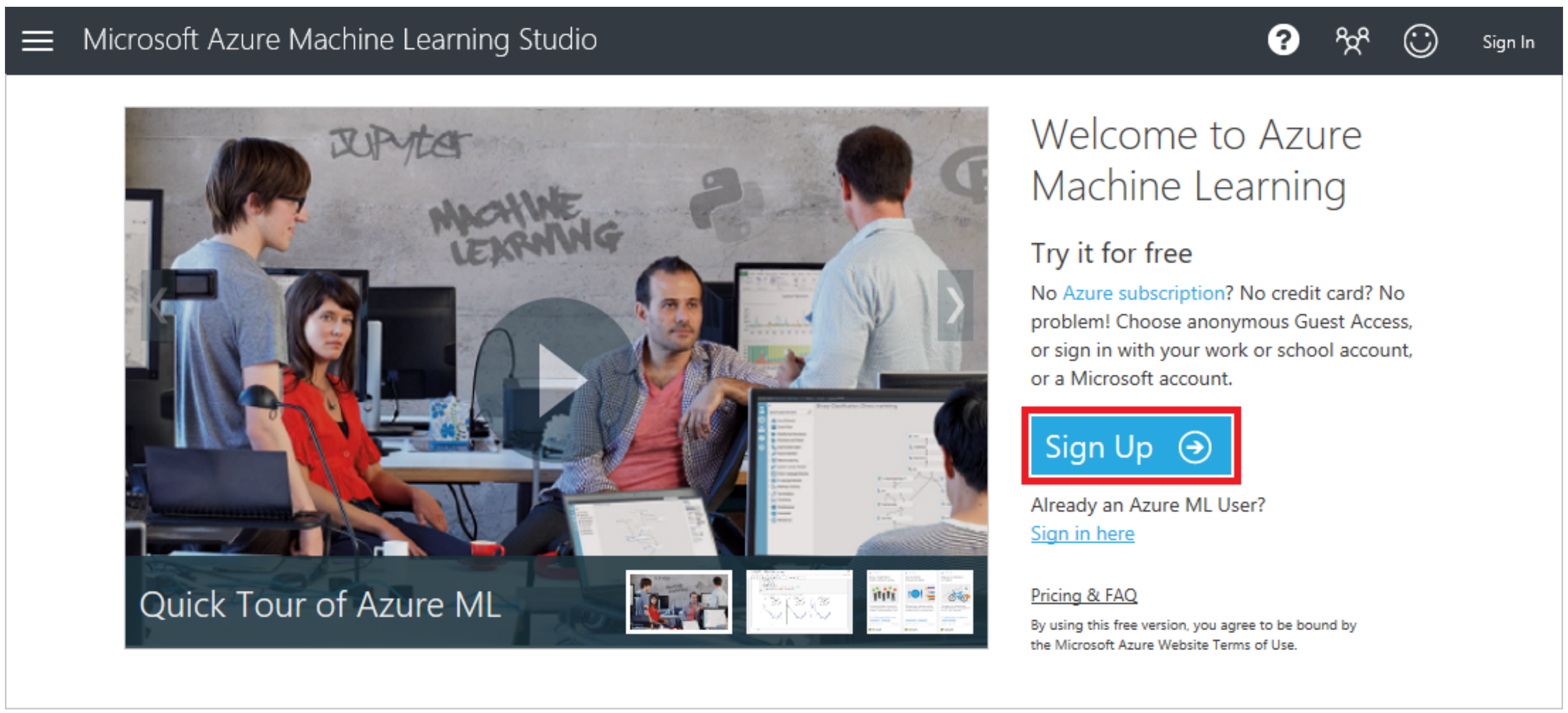
(OPTIONAL) TO PRACTICE: USE AZURE ML STUDIO

Using Azure ML Studio

1. In your Web browser, navigate to http://studio.azureml.net and click the **Sign Up** button.
2. Click **Sign In under Free Workspace.** Then sign in using your Microsoft account.



GETTING STARTED

MACHINE LEARNING

KISHAU ROGERS

- Background: Computer Science, Entrepreneur, 24yrs delivering software solutions
- Blog: www.bigthinking.io
- Email: kishau@bigthinking.io
- TwitterS: @kishau, @bigthinkingio
- Current Focus: Machine Learning @ time:study

Weill Cornell Medicine
NewYork-Presbyterian
Stony Brook Medicine
University of Colorado Hospital
UNIVERSITY OF COLORADO HEALTH
Columbia University Medical Center
Southampton Hospital
MedStar Washington Hospital Center

LEARNING OBJECTIVES

- Overview of Machine Learning vs. Artificial Intelligence
- How to leverage Systems Thinking to avoid common pitfalls
- Roadmap for Machine Learning projects
- Machine Learning algorithms by project goal
- Overview of Model Building Workflow

ORIENTATION

A.I. vs MACHINE LEARNING

ARTIFICIAL INTELLIGENCE : (ONE) DEFINITION

Artificial intelligence is the ability of a machine or a computer program to think and learn.

"The concept of AI is based on the idea of building machines capable of thinking, acting, and learning like humans."

ARTIFICIAL INTELLIGENCE : CHARACTERISTICS

"NOT A.I."	✓ Pre-programmed rules ✓ Pre-programmed responses
PRE-PROGRAMMED SYSTEMS	Difference between AI & preprogrammed systems is the # of Layers removed from the program (the degree that the programmers was NOT involved in the decision process for simulating the intellectual process)

"REAL A.I."	✓ Can Learn & Gain Knowledge Over Time ✓ Can Act on What It has Learned
ARTIFICIAL NARROW INTELLIGENCE	AI that specializes in one area. (ex: AI that can play chess but not make a medical diagnosis)
ARTIFICIAL GENERAL INTELLIGENCE	AI that can perform any intellectual task that a human being can. AI can accumulate knowledge and use it to solve different kinds of problems.
ARTIFICIAL SUPER INTELLIGENCE	"an intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills."

ARTIFICIAL INTELLIGENCE : DISCIPLINE BY GOAL

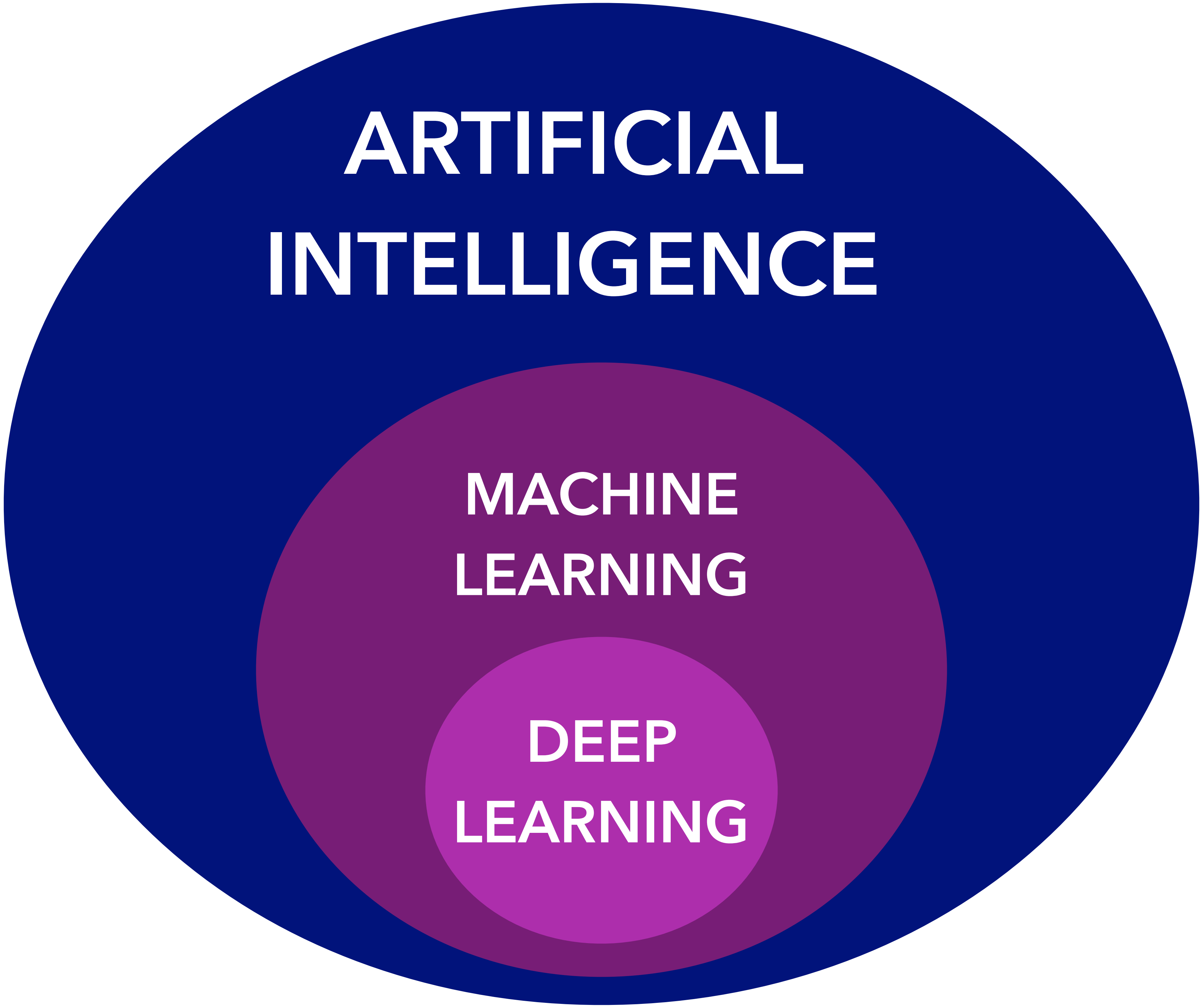
Machine Learning Learn from experience	**Searching & Planning** Synthesize a plan to reach a desired goal	**Knowledge Representation** Knowledge Engineering
Perception Reproduce human vision, touch, hearing	**Natural Language Processing** Understand human language	**Robotics** Move & Manipulate Objects

MACHINE LEARNING : DEFINITION

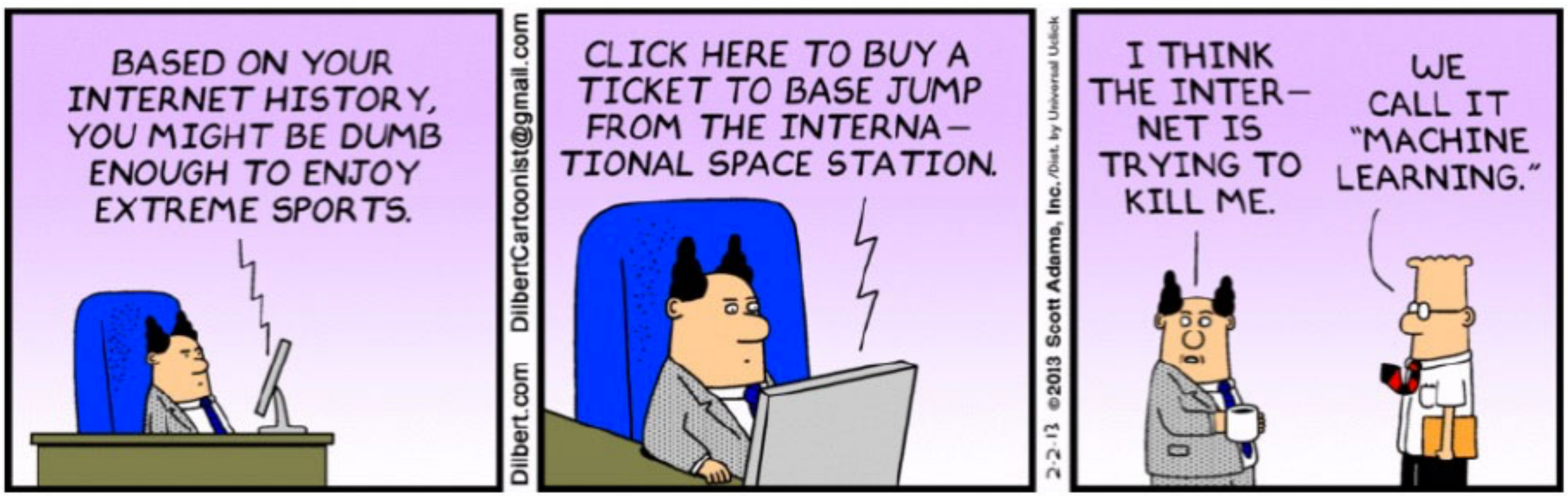
ML is an application of AI. A computer is learning if its performance of a certain task, as measured by computable score, improves with experience (Tom Mitchell). Tom Mitchell, Machine Learning (McGraw Hill, 1997).

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

MACHINE LEARNING VS. ARTIFICIAL INTELLIGENCE



Pitfall : Not knowing when to use Machine Learning



CHARACTERISTICS OF ML APPLICATIONS

	PROBLEM CHARACTERISTICS	ML
INSIGHTS	How well do you understand this problem?	Absence of insights, ML becomes a Black box
COMPLEXITY	Can you code the rules (using a simple deterministic rule-based solution)? Is this a simple problem to solve? How many factors are involved?	Solution not easily coded & Involves Many Variables
ACCURACY	What accuracy rate is required? How quickly does your process need to adjust & learn from mistakes?	Enables Automated learning Accuracy is based on quality of data & model
SCALABILITY	Can a human perform this in a series of repeatable steps? Are you able to scale their efforts?	Offers speed & convenience Limited transparency
DATA ASSETS	Do you have the “right” data to “learn from”? Is the data clean (i.e. not “noisy”)	The “Right” Data is Required Requires data transformation & noise reduction
RESOURCES	Do you have resources to maintain your ML solution?	Data Science, IT, SME
BOTTOM LINE	What is the business impact for solving this problem?	Works well for solving problems with significant bottom-line impact

MACHINE LEARNING : USE CASES

Sample Use Cases

- Fraud detection
- Customer targeting
- Product recommendations
- Sentiment analysis
- Medical Diagnosis

EXERCISE: IDENTIFY A ML PROJECT

COMPLETE THE ML PROJECT EXERCISE

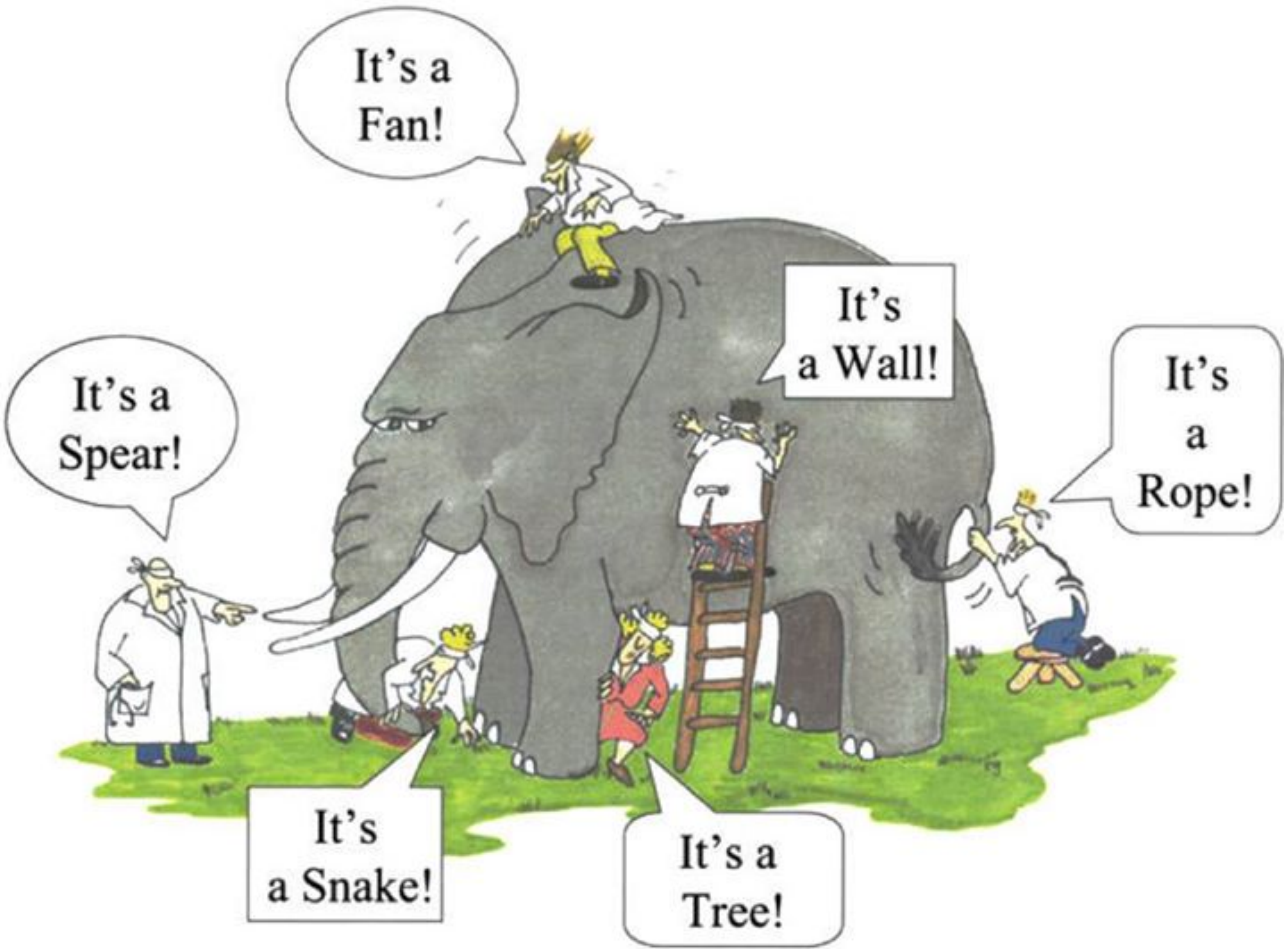
http://bigthinking.io/btFiles/training/MLIntroWorkBook.pdf

ML?	PROJECT DESCRIPTION	INSIGHTS	COMPLEXITY	ACCURACY	SCALE	DATA ASSETS	RESOURCES	BOTTOM-LINE IMPACT

ORIENTATION

SYSTEMS THINKING & ML

SYSTEMS THINKING : PURPOSEFUL MACHINE LEARNING



THINKING FRAMEWORKS

FRAMEWORK	DEFINITION	FOCUS	WHEN TO USE
Systems Thinking	To improve understanding of systems by thinking about the wider context of a situation, to uncover patterns and structures, to consider past and current state of systems.	Focusing on understanding complex systems and identifying opportunities for solving "root causes" to problems.	‣ To understand interconnected, complex problems ‣ To address divergent perspectives & stakeholders ‣ To avoid unintended consequences
Design Thinking	To innovate and transform systems by addressing the question "What can be?" The design thinking process has the following stages: Empathize, Define, Ideate, Prototype, Test.	Focusing on discovering solutions with customer-centered problem solving.	‣ To uncover needs ‣ To design solutions
Lean	A business methodology used to create more value for customers by using fewer resources and eliminating waste.	Focusing on improving the business model by relentlessly eliminating waste and delivering more value	‣ To eliminate waste ‣ To improve quality ‣ To reduce time & costs ‣ To preserve value with less work
Agile Methods	An approach to software development which emphasizes a continual planning, execution and delivery of working software.	Focusing on rapidly delivering working software solutions and continuously adapting to changing environments.	‣ To deliver working software ‣ To continuously adapt and respond to change, feedback and unpredictability ‣ Effective collaboration to ensure continuous value

SYSTEMS THINKING FRAMEWORK FOR MACHINE LEARNING

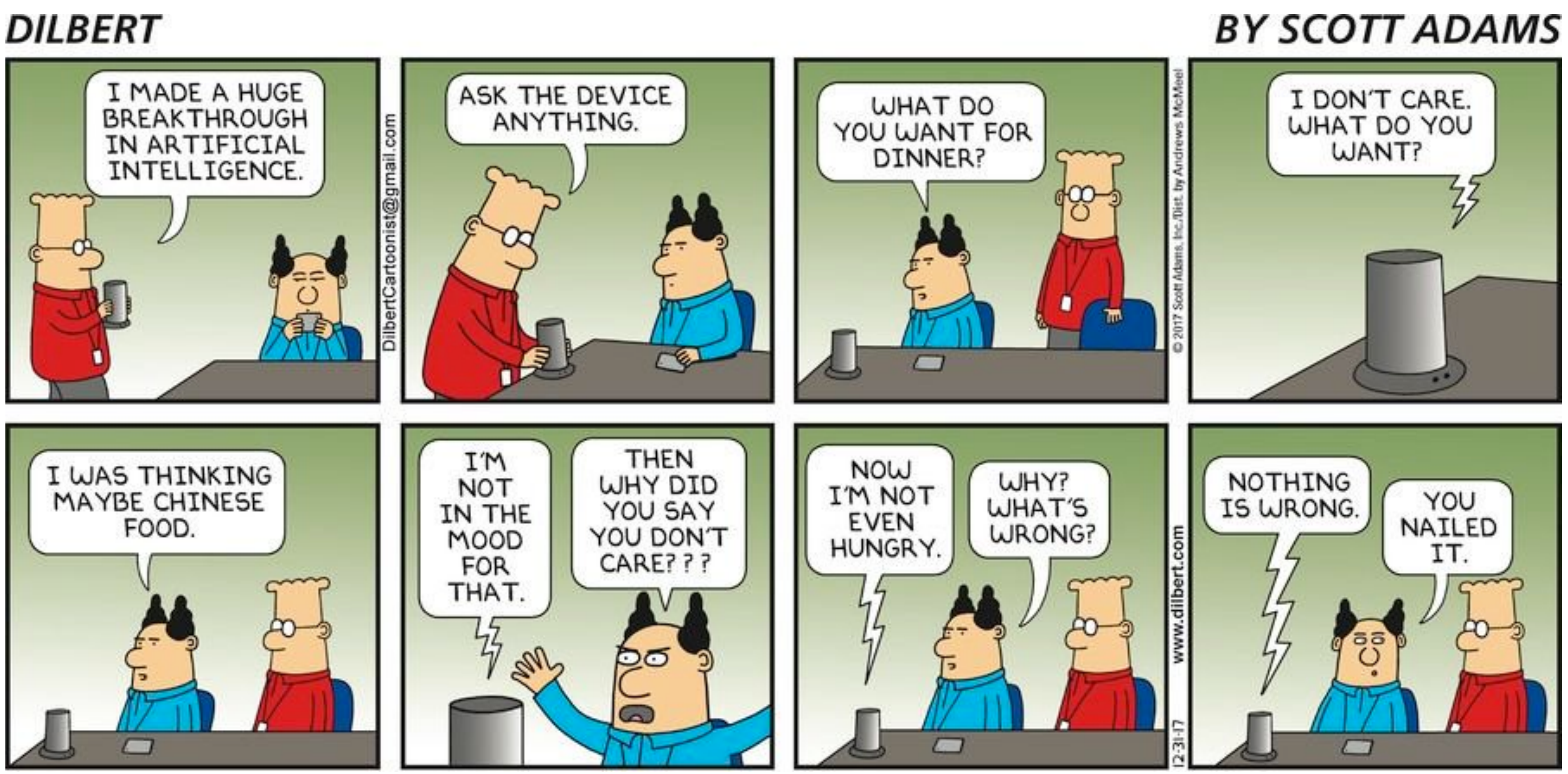
ML ROADMAP	CLASSIFY	ACQUIRE	PREPARE	BUILD	VALIDATE	DEPLOY	MONITOR
GOAL	Identify hypothesis	Acquire data assets & establishing context	Improve data quality & identify bias	Develop an appropriate learning system	Identify & Reduce error	Present results	Monitor change
PRINCIPLE	Purposeful	Openness	Multi-dimensional	Patterns & Trends	Counter-intuitive	Emergence	Adaptability
TOOLS	Archetypes Ladder of Inference	Iceberg Model	Stocks and Flows	Modeling & Simulation	Feedback Loops	Highest Leverage	Behavior Over Time
METRICS	Questions That Data Can Answer	Data Boundaries	Transparent open datasets	Experiments & Algorithms	Model Scores & Results	Predictions	Performance
INSIGHTS	Stakeholders	Data Owners	Data Managers	Engineers & Scientists	Engineers & Stakeholders	IT	Stakeholders
TOOLS & ARCHITECTURE	USE CASES	DATA LAKE	DATA WAREHOUSE	SAFE LEARNING SPACE	QA/QC	PREDICTION ENGINE	DATA AUDITS

STEP 1

CLASSIFYING THE PROBLEM

CLASSIFY THE PROBLEM

Pitfall : Not forming a question that can be answered with data



ASKING A WELL-FORMED QUESTION

QUESTION TYPE	SAMPLE QUESTIONS
Categorizing or Classifying Things	Will this customer renew their subscription? Is this a cat or a dog? Is this transaction fraudulent?
Predicting Values	What will the temperature be next Friday? How many new customers will we get next week?
Detecting Anomalies	Is this internet usage typical? Is this an unusual pressure reading?
Grouping Related Data	Which shoppers have similar tastes in footwear? What is the best way to break these items into 4 groups?
Determining the next best step	Where should this ad be placed on the page so that the viewer is most likely to click it? Should I adjust the temperature higher, lower or leave it? Should I accept this job offer?

MACHINE LEARNING CATEGORIES

LEARNING CATEGORY	DESCRIPTION	USED FOR
SUPERVISED	Machine learns “by example.” You provide a dataset with known inputs/outputs and the algorithm finds the method for arriving at those inputs/outputs.	Categorizing or Classifying Things Predicting Values Detecting Anomalies
UNSUPERVISED	The machine studies the available data to identify patterns, then the algorithm tries to organize the data in a way that describes its structure.	Grouping Related Data Dimension Reduction
REINFORCEMENT LEARNING	Given a set of actions and end-values, the machine learns how to achieve the best possible result using trial and error and learning from past experiences.	Determining the next best step

TYPES OF ML SYSTEMS BY GOAL

YOUR GOAL	ML LEARNING TYPE	ALGORITHM TYPE
Is <this> A or B?	Supervised	Classification
Is <this> A, B, C or D?	Supervised	Multi-Class Classification
Is <this> weird or unexpected?	Unsupervised	Anomaly detection
How many or How much is <this>?	Supervised	Regression
How is <this> organized?	Unsupervised	Clustering
What should <this> do next?	Reinforcement Learning	Reward maximization

SELECTING AN ALGORITHM

Considerations:

- Your goal. What type of problem do you want to solve?
- The size & nature of your dataset
- Data quality
- Training time
- Accuracy

SELECTING AN ALGORITHM

LEARNING CATEGORY	QUESTION TYPE	ALGORITHM TYPE	ALGORITHMS
SUPERVISED	Categorizing or Classifying Things	Classification	Two-class SVM, Two-class Bayes point machine, Decision Forest, Logistic regression, Two-class Boosted decision tree, Decision Jungle, Locally deep SVM, Neural network, One-v-all multi class
	Predicting Values	Regression	Linear Regression, Ordinal Regression, Poisson Regression, Fast forest quantile, Bayesian Linear, Decision Forest, Boosted Decision Tree , Neural Network
	Detecting Anomalies	Anomaly Detection	One Class SVM PCA-Based
UNSUPERVISED	Grouping Related Data	Clustering	K-Means
REINFORCEMENT LEARNING	Determining the next best step	Reinforcement Learning	Neural Network

WHAT TYPE OF ALGORITHM SHOULD YOU USE TO ANSWER THIS QUESTION:

"HOW LIKELY IS THIS CUSTOMER TO RENEW THEIR SUBSCRIPTION?"

- Classification
- Regression
- Clustering
- Anomaly Detection

BREAK

10 MINUTES

PRACTICE

USE CASES

EXERCISE : PROJECT INITIATION BLUEPRINT

COMPLETE PROJECT INITIATION EXERCISE

bigthinking.io/btFiles/training/MLIntroWorkBook.pdf

TASK:

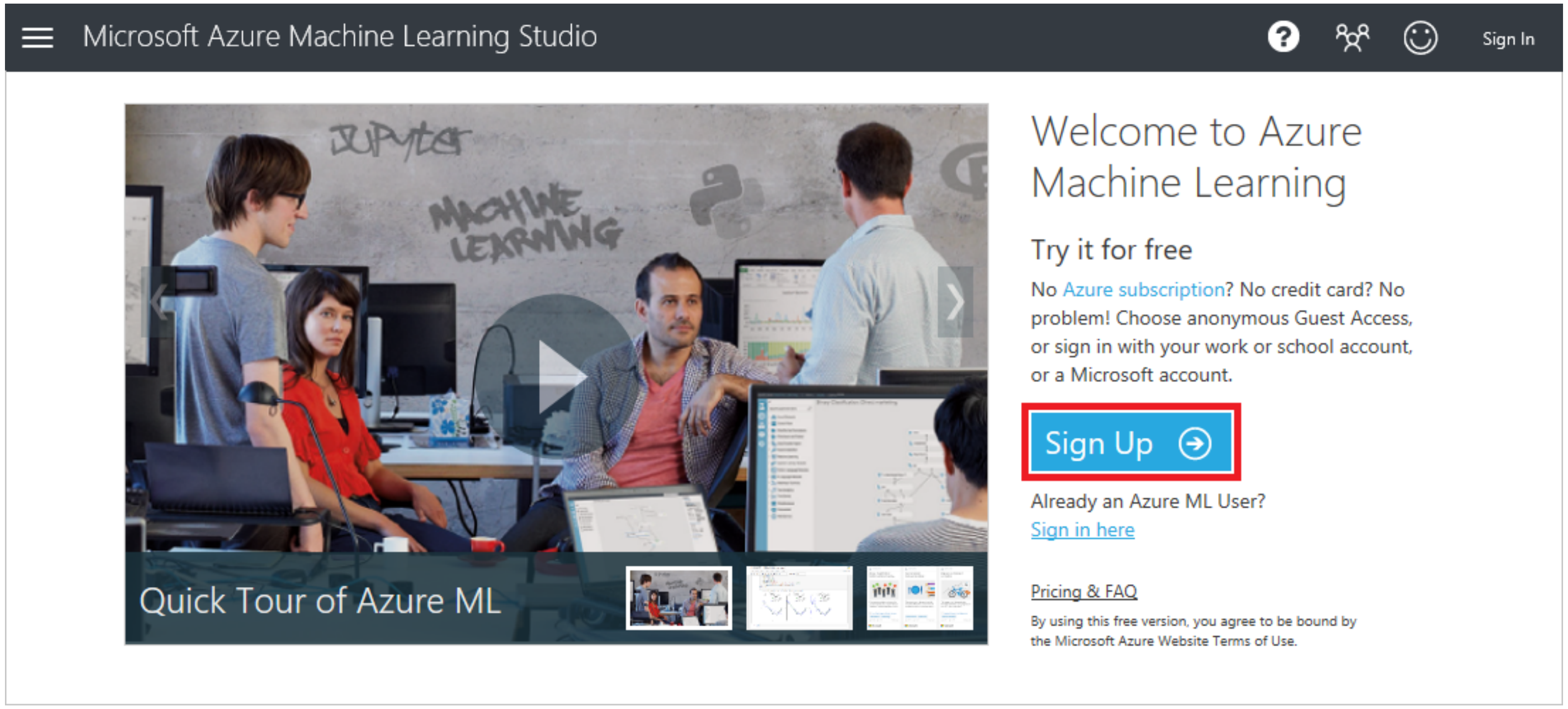
- ☐ Overview of Today's Use Case
- ☐ Review the "Data DataSheet"
- ☐ Identify a well-formed question that can be answered with the dataset(s) provided.
- ☐ Select the appropriate machine learning type

QUESTION TO BE ANSWERED			
LEARNING TYPE	SUPERVISED	UNSUPERVISED	REINFORCED

TO PRACTICE: USE AZURE ML STUDIO

Using Azure ML Studio

1. In your Web browser, navigate to http://studio.azureml.net and click the **Sign Up** button.
2. Click **Sign In under Free Workspace**. Then sign in using your Microsoft account.



BREAK

10 MINUTES

EXERCISE : SHARE YOUR WELL FORMED QUESTION

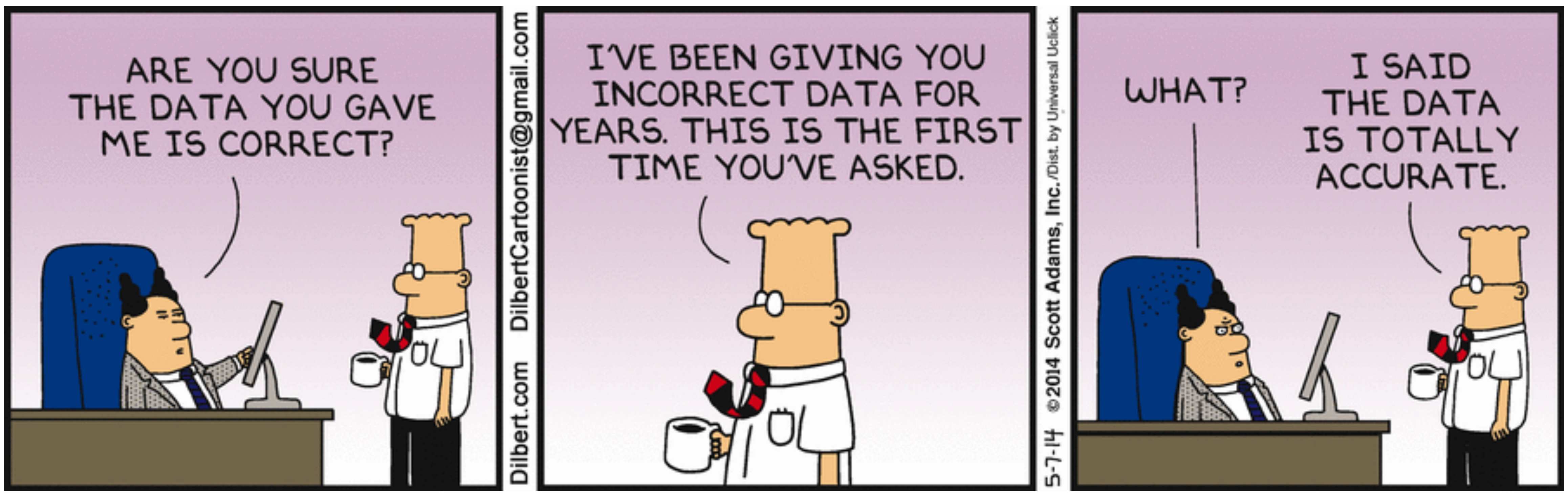
WHAT QUESTIONS CAN YOU ANSWER WITH THE SELECTED USE CASE?

STEP 2

ACQUIRING DATA

DATA ACQUISITION

Pitfall : Lack of appropriate data assets to answer the question



IMPACT OF DATA

- the task you are trying to perform,
- the performance you want to achieve,
- the input features you have,
- the noise in the training data,
- the noise in your extracted features,
- the complexity of your model
- …. and so on.

DATA ACQUISITION CONSIDERATIONS

SELECTION BIAS

Occurs when the samples used to produce the model are not fully representative of cases that the model may be used for in the future, particularly with new and unseen data.

DATA FRAGMENTATION

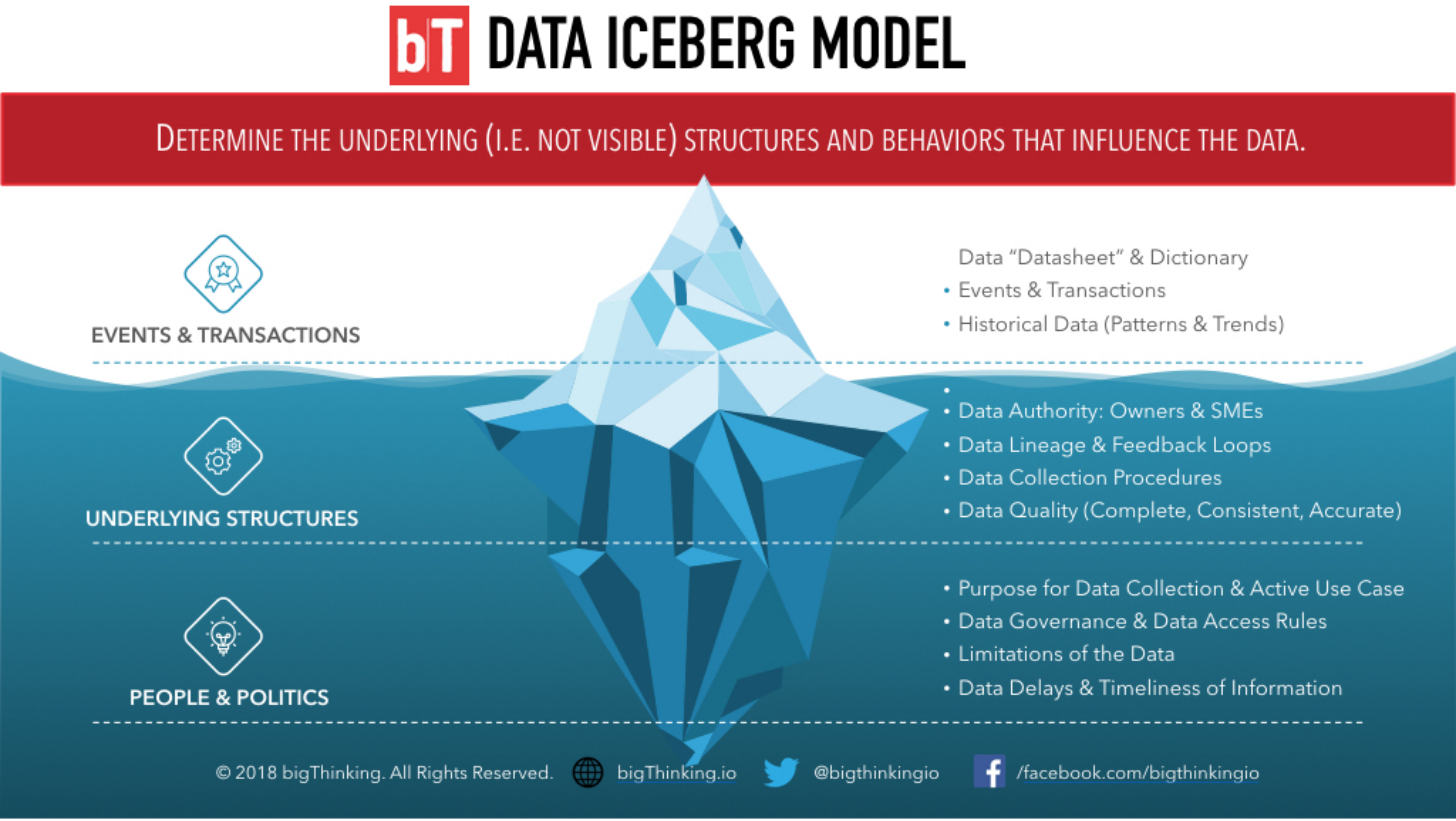
Data in disaggregated "data islands." This means that data stored in different systems and may be in diverse formats.

OTHER THINGS TO CONSIDER

1. What is the extent of the data you have available? Ensure you have a clear picture of everything that you can use.

2. What data is not available that you wish you had available? Identify data that is not recorded or cannot be recorded.

3. What data don't you need to address the problem? Note down which data you excluded and why.


bT DATA ICEBERG MODEL
DETERMINE THE UNDERLYING (I.E. NOT VISIBLE) STRUCTURES AND BEHAVIORS THAT INFLUENCE THE DATA.
EVENTS & TRANSACTIONS
Data "Datasheet" & Dictionary
• Events & Transactions
• Historical Data (Patterns & Trends)
UNDERLYING STRUCTURES
• Data Authority: Owners & SMEs
• Data Lineage & Feedback Loops
• Data Collection Procedures
• Data Quality (Complete, Consistent, Accurate)
PEOPLE & POLITICS
• Purpose for Data Collection & Active Use Case
• Data Governance & Data Access Rules
• Limitations of the Data
• Data Delays & Timeliness of Information
© 2018 bigThinking. All Rights Reserved.
bigThinking.io
@bigthinkingio
/facebook.com/bigthinkingio

DATA ACQUISITION : DOCUMENTING DATA ASSETS

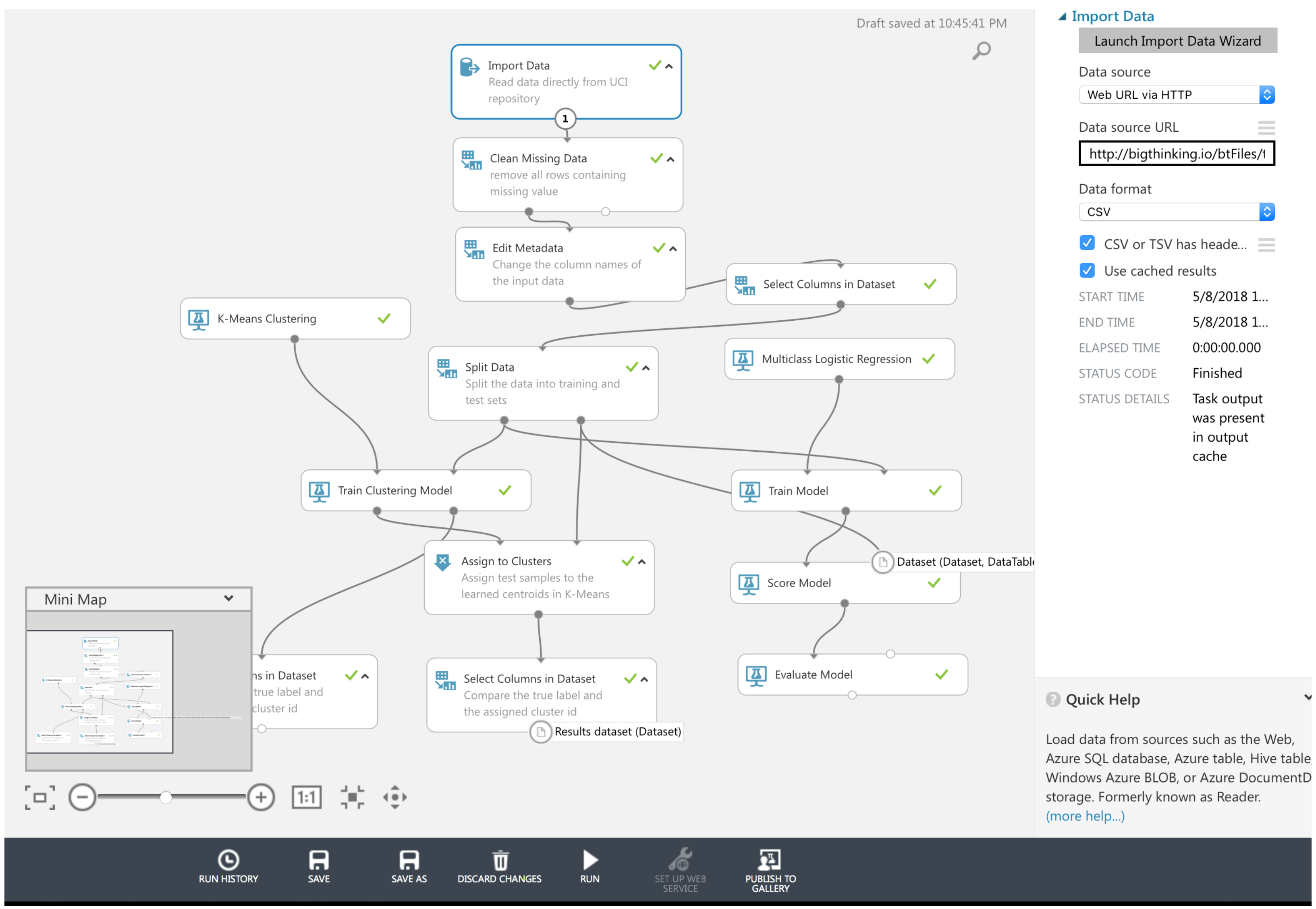
- IDENTIFIES DATA OWNERS AND SME
- DATA DEFINITIONS
- SOURCE, DATE & TIME OF ACQUISITION
- DATA LINEAGE
- DATA ACCESS POLICIES & PROCEDURES
- DATA CONNECTION PROTOCOL
- DATA REVIEW & ANNOTATIONS
- DATA FEEDBACK LOOP (USER REVIEW & FEEDBACK)

EXERCISE: DATA ACQUISITION

Using Azure ML Studio

COMPLETE DATA ACQUISITION EXERCISE

http://bigthinking.io/btFiles/training/MLIntroWorkBook.pdf



BREAK

10 MINUTES

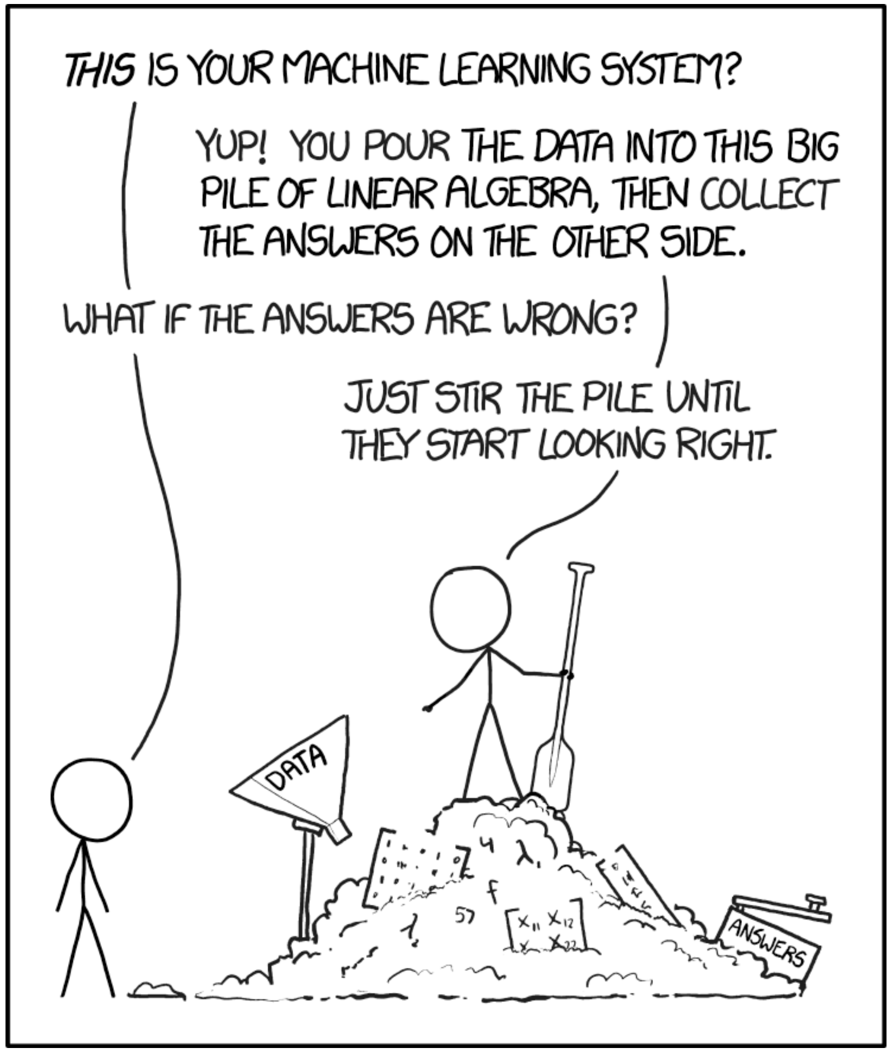
STEP 3

PREPARING & CLEANING DATA

DATA PREPARATION & CLEANING

Pitfall : Low quality data, inconsistent, duplicate or corrupt datasets

"Garbage (Data) In, Garbage (Results) Out"



DATA PREPARATION & CLEANING

POTENTIAL DATA ISSUES	TASKS
Duplicates	Duplicated rows of data
Irrelevant Data	Data which does not influence the solution
Structural Issues	Typos, misspellings
Missing Values	Determine if removal or (re)coding is required
Outliers (Maybe)	Outliers may be important to the solution space (ex: fraud detection), otherwise can also negatively impact training process.

FEATURES

Feature: Property of data being used; a column of data in your input set.

Feature Engineering: Process of determining the best representation of the data to learn from; process of turning your inputs into something that the algorithm can understand.

METHODS	DESCRIPTION
Feature Extraction	Automatically constructing new features from raw data.
Feature Construction	The manual construction of new features from raw data.
Feature Selection	Removing unnecessary features; selecting the few that are useful.

FEATURE SELECTION

Eliminating the unnecessary and identifying the appropriate features; also called "variable selection"

- Provides faster predictors
- Improves prediction performance
- Reduces complexity of the model

FEATURE SELECTION APPROACHES

APPROACH	DESCRIPTION
Filter	Selects variables regardless of the model; evaluates features independently
Wrapper	Selects variables according to their usefulness. A wrapper method evaluates a subset of features thus it takes the interactions between features into account.
Embedded	Selects variables as a part of the learning procedure.

FEATURE SELECTION

Getting Started: Machine Learning › Edit Metadata › Results dataset

rows 165324 | columns 7

view as

MaskedPersonID	ActivityDate	Administration	Teaching	ServiceDelivery	OtherActivity	Label
620	2017-08-15T00:00:00	0	0	0	8.25	Independent
1603	2016-03-25T00:00:00	0	0	0	0	Unclassified
398	2016-03-18T00:00:00	0	0	0	9	Independent
1384	2013-09-10T00:00:00	0	1.5	0	4.5	Independent
206	2015-02-17T00:00:00	0	0.5	4.5	0	Service
1319	2012-06-05T00:00:00	1	0	9.5	0	Service
1301	2014-09-19T00:00:00	4	3	0	0	Administrator
487	2015-06-05T00:00:00	1	0	10	0	Service
1586	2016-11-04T00:00:00	8	0	0	0	Administrator
375	2016-08-02T00:00:00	0	0	8	0	Service
1046	2017-10-12T00:00:00	0	0	0	10	Independent
113	2015-10-12T00:00:00	0	0	0	9	Independent

Statistics

Unique Values	5
Missing Values	0
Feature Type	String Feature

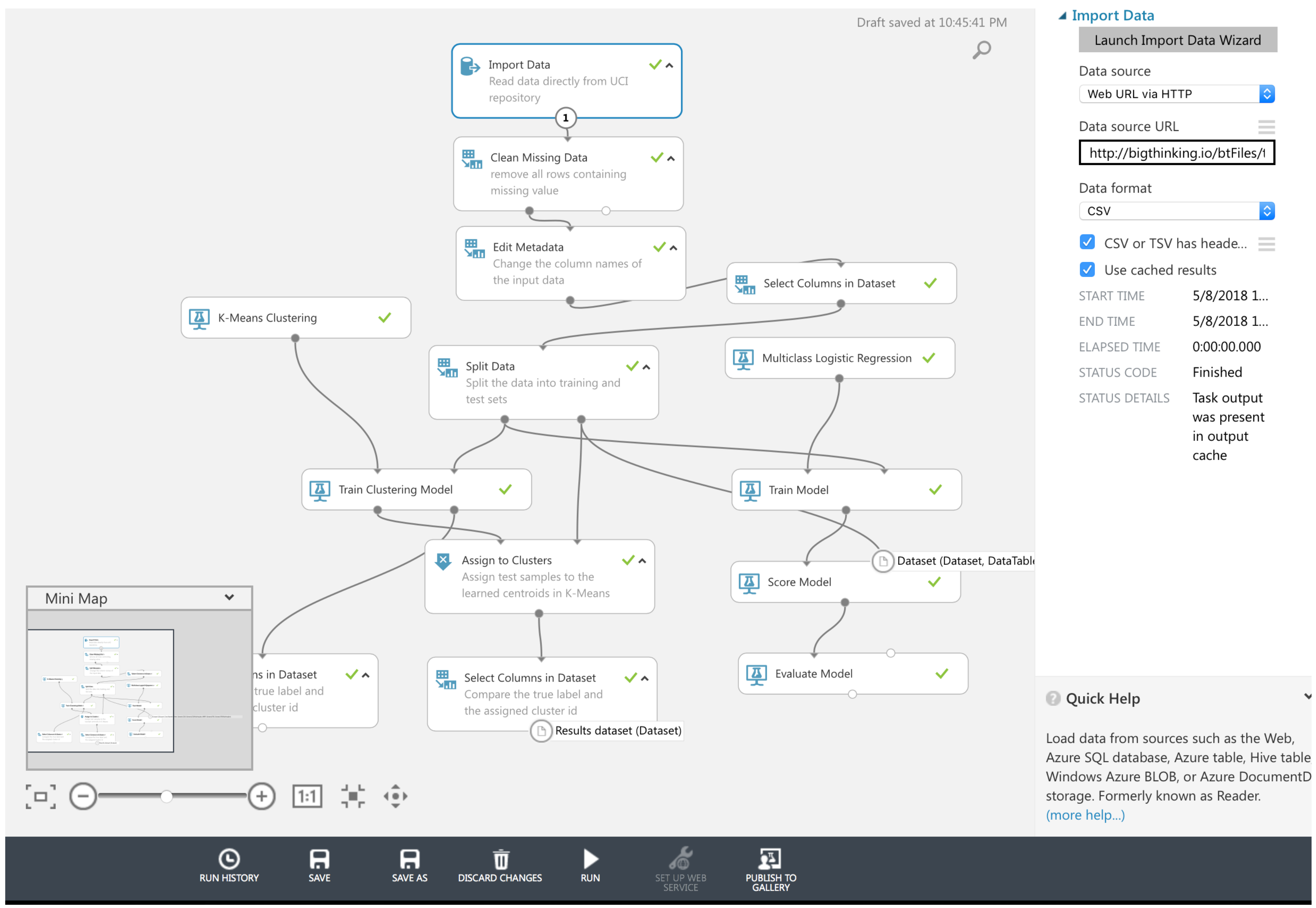
Visualizations

EXERCISE: DATA CLEANING

Using Azure ML Studio

COMPLETE DATA CLEANING EXERCISE

http://bigthinking.io/btFiles/training/MLIntroWorkBook.pdf



BREAK

10 MINUTES

FEATURE EXTRACTION IS THE THE MANUAL CONSTRUCTION OF NEW FEATURES FROM RAW DATA.

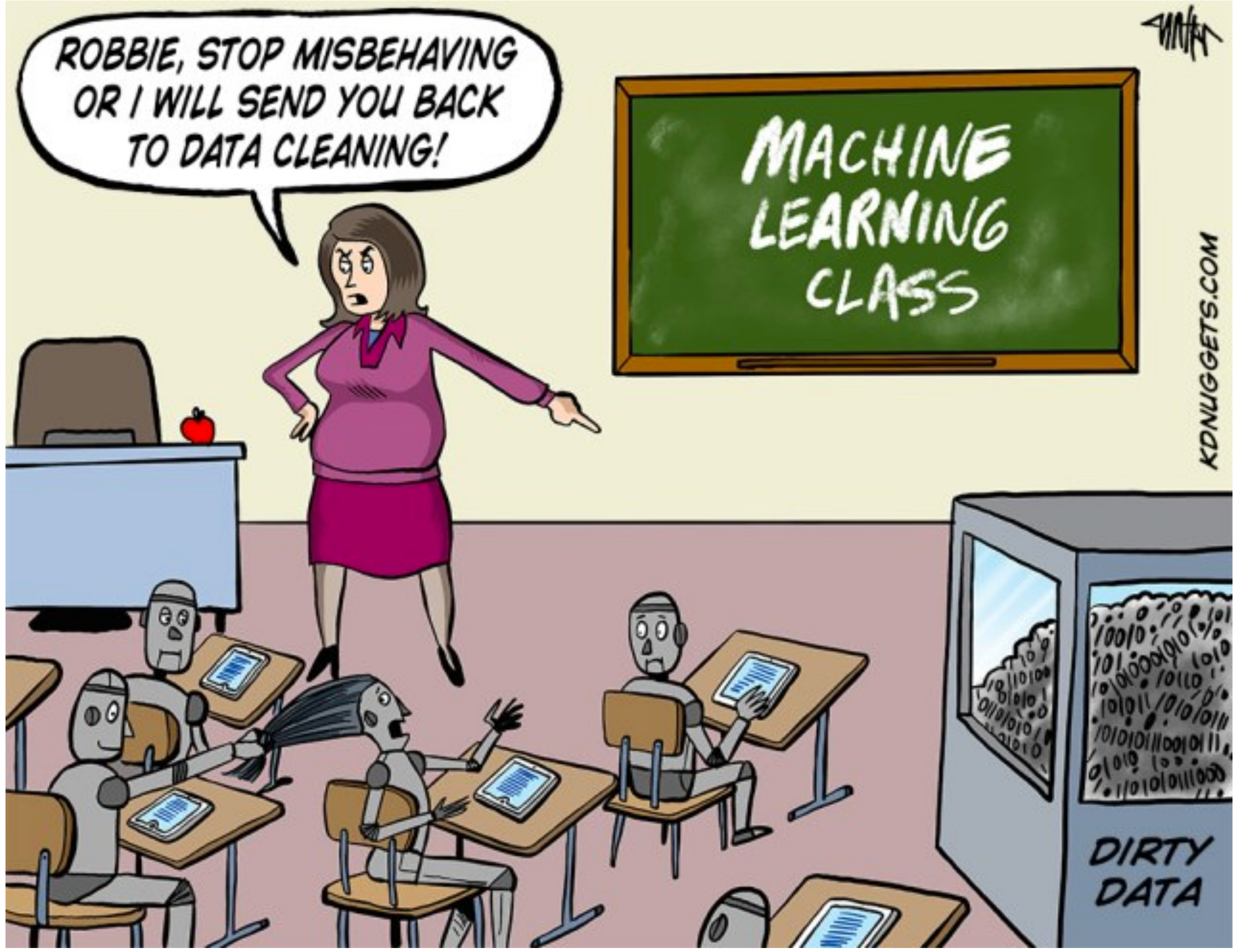
- Yes
- No

STEP 4

BUILDING MODELS

MODEL BUILDING

Pitfall : Building a model that fails to learn (fails to generalize for unseen data).

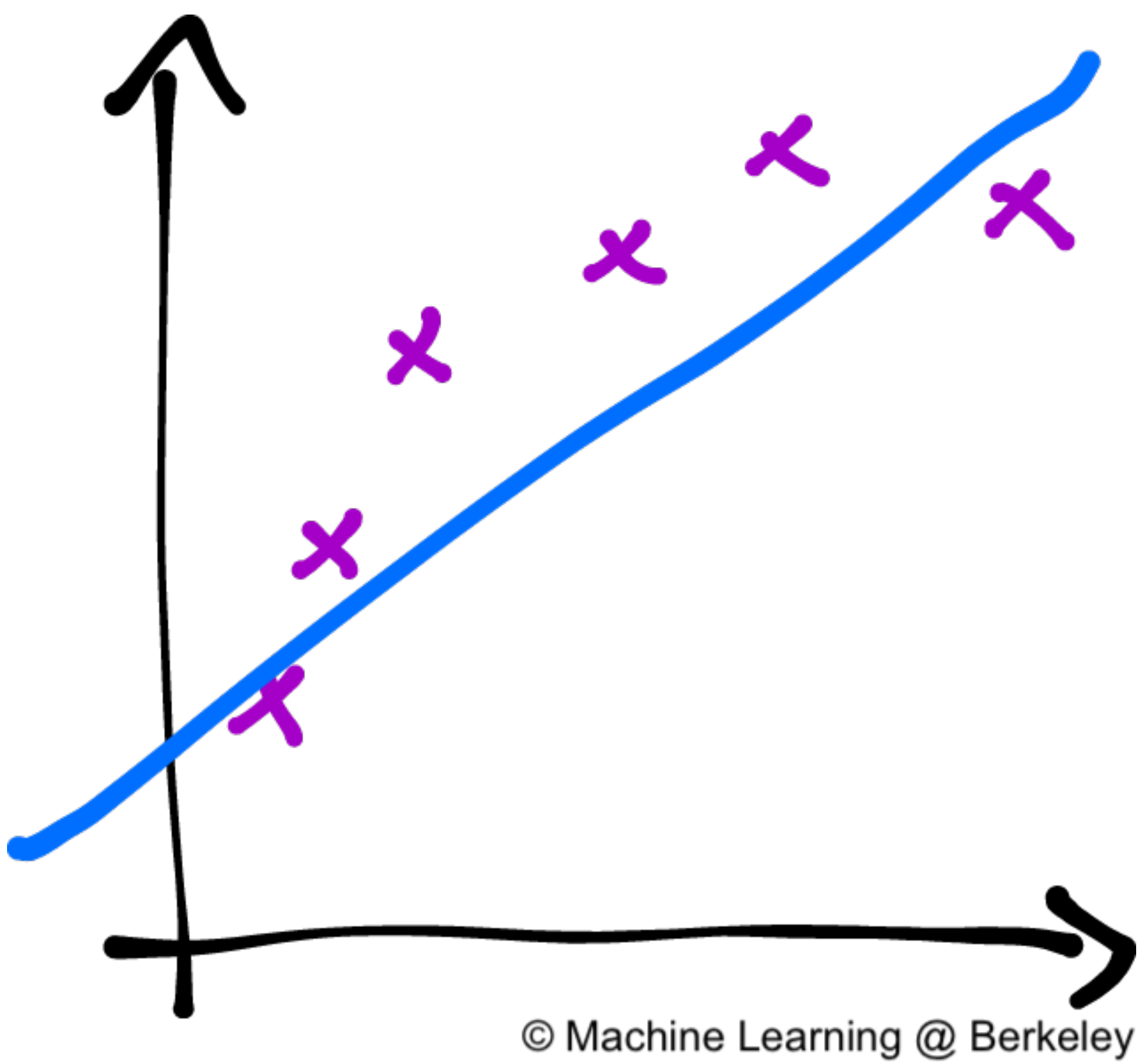


MODELS: OVERFITTING VS. UNDERFITTING

- Failure to address patterns & trends VS. structure (data may reflect patterns, but fail to address the "structure" and "behaviors" that create the patterns).
- A good model can separate signals from noise.
- Overfitting occurs when the model fits the training data TOO well and fails to generalize for unseen data.
- Under-fitting occurs when the model does not fit the data well. Cannot model the training data nor generalize for unseen data.

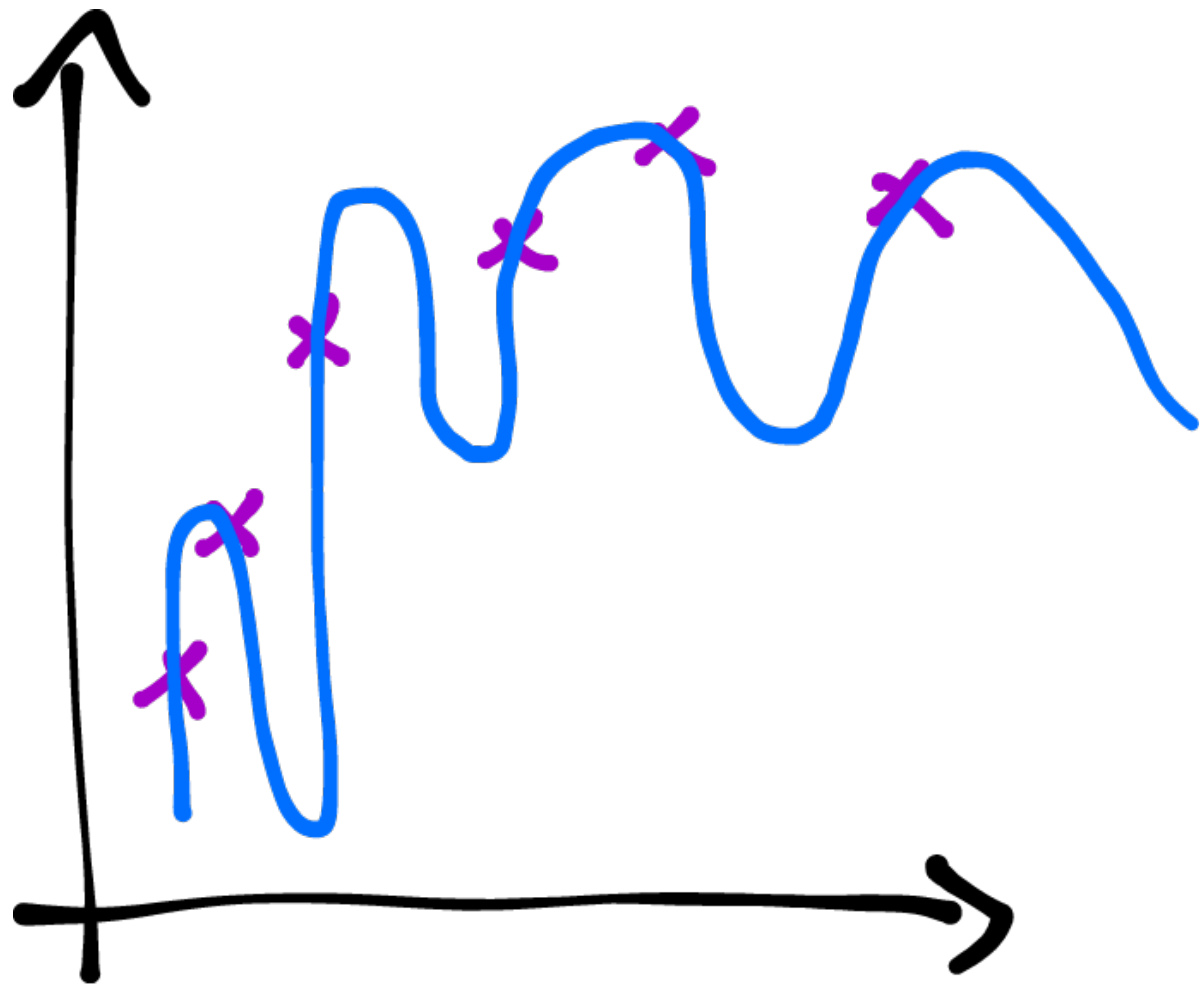
MODEL BUILDING : UNDER-FITTING

Model performs poorly on training data and poorly generalizes for unseen data.



MODEL BUILDING : OVERFITTING

Model performs well on training data, but performs poorly on unseen data.



RESPONDING TO OVERFITTING

Model performs well on training data, but performs poorly on unseen data.

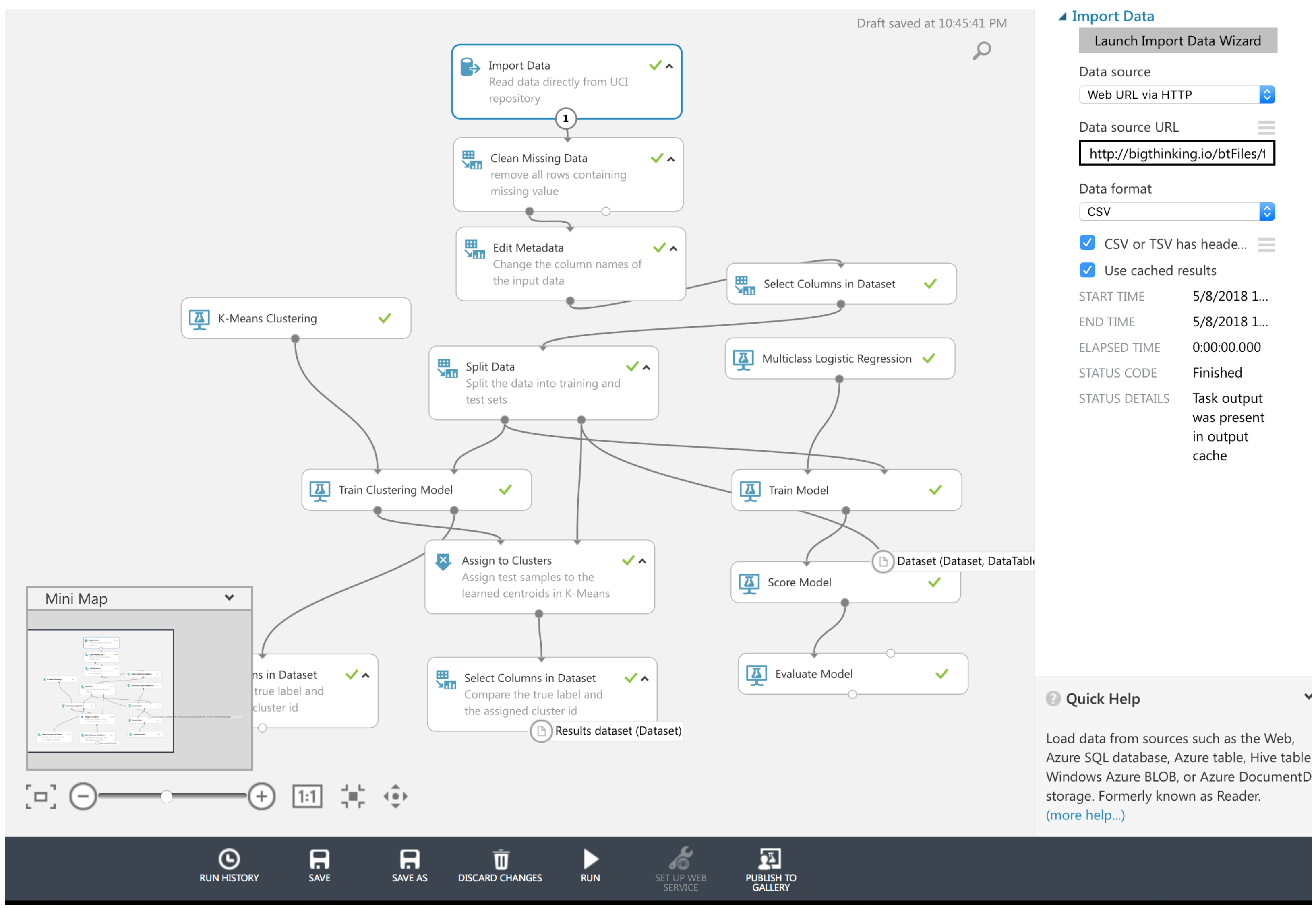
METHODS	DESCRIPTION
Cross Validation	To minimize generalization errors, train the model on complementary subsets of the input data (folds); keeping your test data separate.
Train with More Data	(Sometimes) adding more clean data will enable the algorithms to better find the signal.
Feature Selection	Removing unnecessary features; selecting the few that are useful.

EXERCISE: MODEL BUILDING

Using Azure ML Studio

COMPLETE MODEL BUILDING EXERCISE

http://bigthinking.io/btFiles/training/MLIntroWorkBook.pdf



BREAK

10 MINUTES

STEP 5

VALIDATING MODELS

Pitfall : Failure to identify biased datasets (the cure can be worse than the disease).

"THE FIRST TEST WAS FALSE-POSITIVE, THE SECOND TEST WAS FALSE-NEGATIVE. WHAT ARE YOU TRYING TO PULL?"

EVALUATING PERFORMANCE : CONFUSION MATRIX

Identifying Precision, Sensitivity, Specificity, Accuracy.

	POSITIVE actual	NEGATIVE actual
POSITIVE predicted	True Positive	False Positive
NEGATIVE predicted	False Negative	True Negative

EVALUATING PERFORMANCE : ACCURACY

Identifying Precision, Sensitivity, Specificity, Accuracy.

	POSITIVE actual	NEGATIVE actual
POSITIVE predicted	True Positive	False Positive
NEGATIVE predicted	False Negative	True Negative

EVALUATING PERFORMANCE : PRECISION

Identifying Precision, Sensitivity, Specificity, Accuracy.

	POSITIVE actual	NEGATIVE actual
POSITIVE predicted	True Positive	False Positive
NEGATIVE predicted	False Negative	True Negative

EVALUATING PERFORMANCE : SENSITIVITY

Identifying Precision, Sensitivity, Specificity, Accuracy.

	POSITIVE actual	NEGATIVE actual
POSITIVE predicted	True Positive	False Positive
NEGATIVE predicted	False Negative	True Negative

EVALUATING PERFORMANCE : SPECIFICITY

Identifying Precision, Sensitivity, Specificity, Accuracy.

	POSITIVE actual	NEGATIVE actual
POSITIVE predicted	True Positive	False Positive
NEGATIVE predicted	False Negative	True Negative

CONFUSION MATRIX

Clustering: Work Style data › Evaluate Model › Evaluation results ×

Overall accuracy	0.843333
Average accuracy	0.937333
Micro-averaged precision	0.843333
Macro-averaged precision	0.836563
Micro-averaged recall	0.843333
Macro-averaged recall	0.811653

Confusion Matrix

Predicted Class

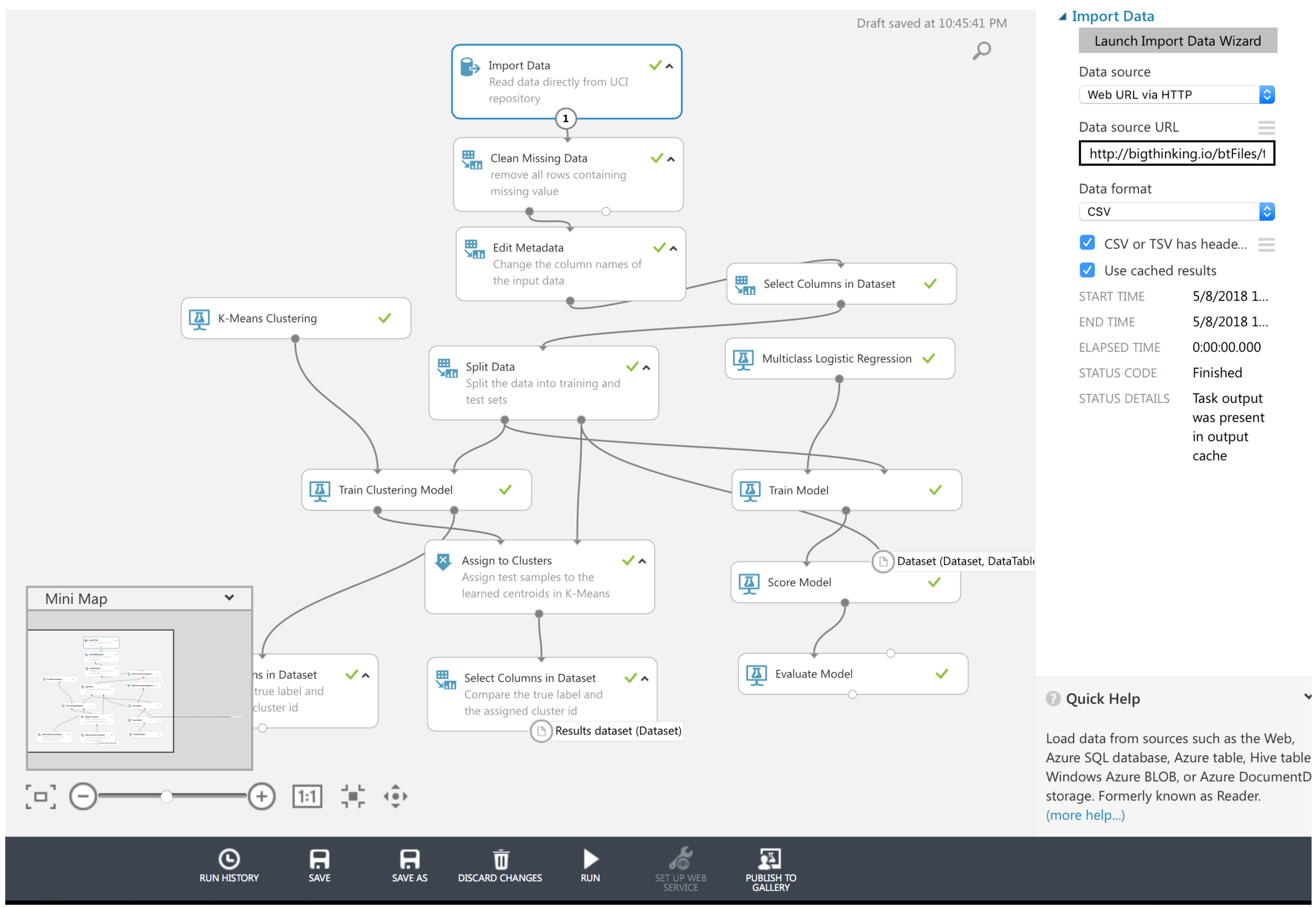
Actual Class	Administ...	Independ...	Service	Teacher	Unclassi...
Administ...	80.2%				19.8%
Independ...		97.4%			2.6%
Service			94.1%		5.9%
Teacher			1.2%	73.2%	25.6%
Unclassi...	5.5%	12.7%	12.7%	8.2%	60.9%

EXERCISE: MODEL EVALUATION

Using Azure ML Studio

COMPLETE MODEL EVALUATION EXERCISE

http://bigthinking.io/btFiles/training/MLIntroWorkBook.pdf



BREAK

10 MINUTES

STEP 6

DEPLOYING SOLUTIONS

DEPLOY : EXPLORATORY VS. PRODUCTION SYSTEMS

Pitfall : Building models that are unable to add real-world value (failure to identify the point of highest leverage).

"First thing Monday, we're gonna scale back the Machine Learning budget."

brianmooredraws.com

PRODUCTION STANDARDS

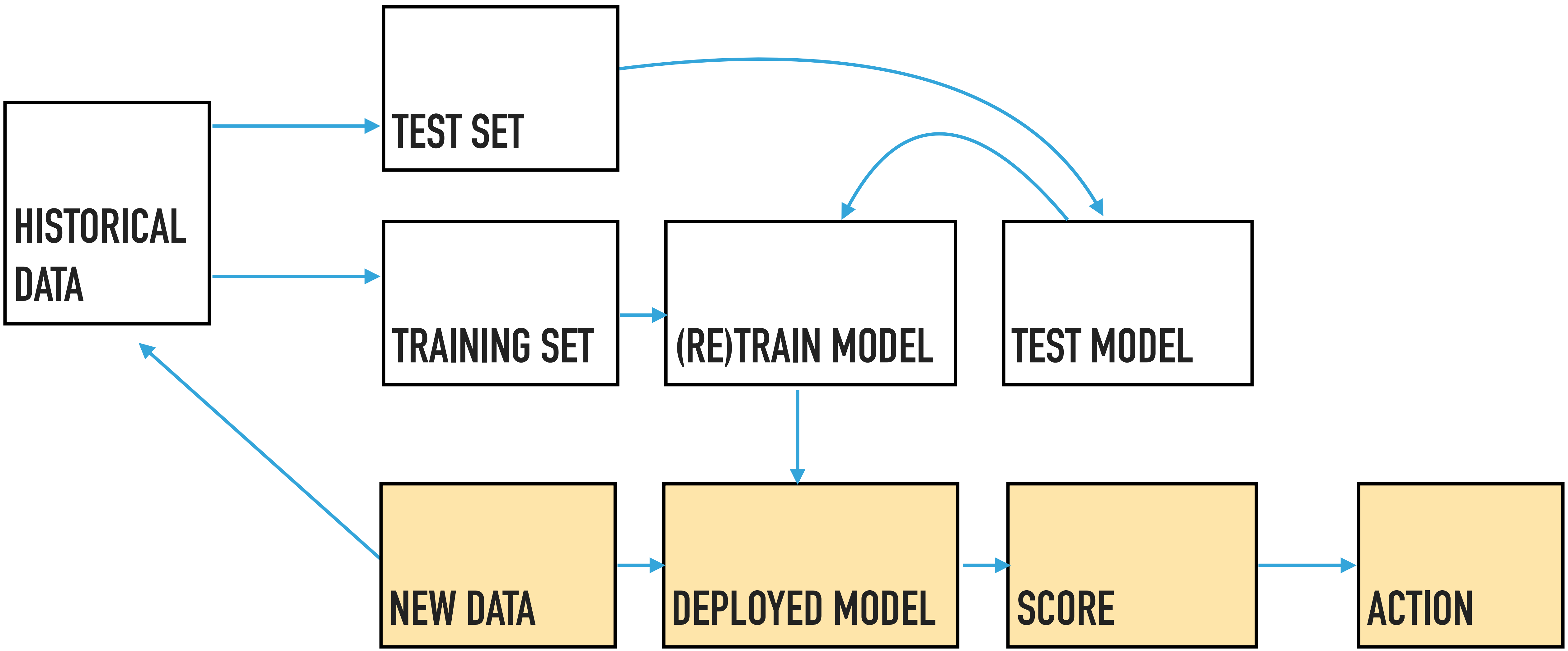
STANDARD	DESCRIPTION
Scope	An clear indication of the well-formed question to be answered & stakeholder impact
Data Quality	Access to the "right" data; automated process for cleaning data.
Model Quality	Model produces timely feedback, results are repeatable and explainable
Performance	Performance standards defined (accuracy, speed etc.). Tested for scalability. Automated data ingestion.
Maintainability	Well documented model versioning, deployment & configuration procedures.
Monitoring	Real-time alerts on performance and adaptability.
Security	Meets data privacy & info security standards.

PRODUCTION-READINESS TESTING

ML Concept (Exploratory vs. Production Prediction Engines) - Determining how this ML system fits in the larger system; where is the appropriate point of intervention?

METHODS	DESCRIPTION
Repeatable Training	Training steps can be reproduced.
ML Pipeline Integration	Use integration testing to validate the ML Pipeline; that you can move through each stage with a working model (data, feature generation, model training, verification, deployment)
Model Quality	Meets quality, accuracy & security thresholds. Comparing predictions on a validation set; compare predictions to prior model versions.
Debuggable	How difficult is it to determine why a model is not working as expected. Can the model can be debugged via step-by-step analysis?
Rollback	Is there a model rollback procedure? Can you quickly revert back to a known-good state?

FEEDBACK LOOPS : DISCOVERY—>QA—>PRODUCTION



DEFINE the point in the production workflow that an action can be taken on this score?

What decisions are made?

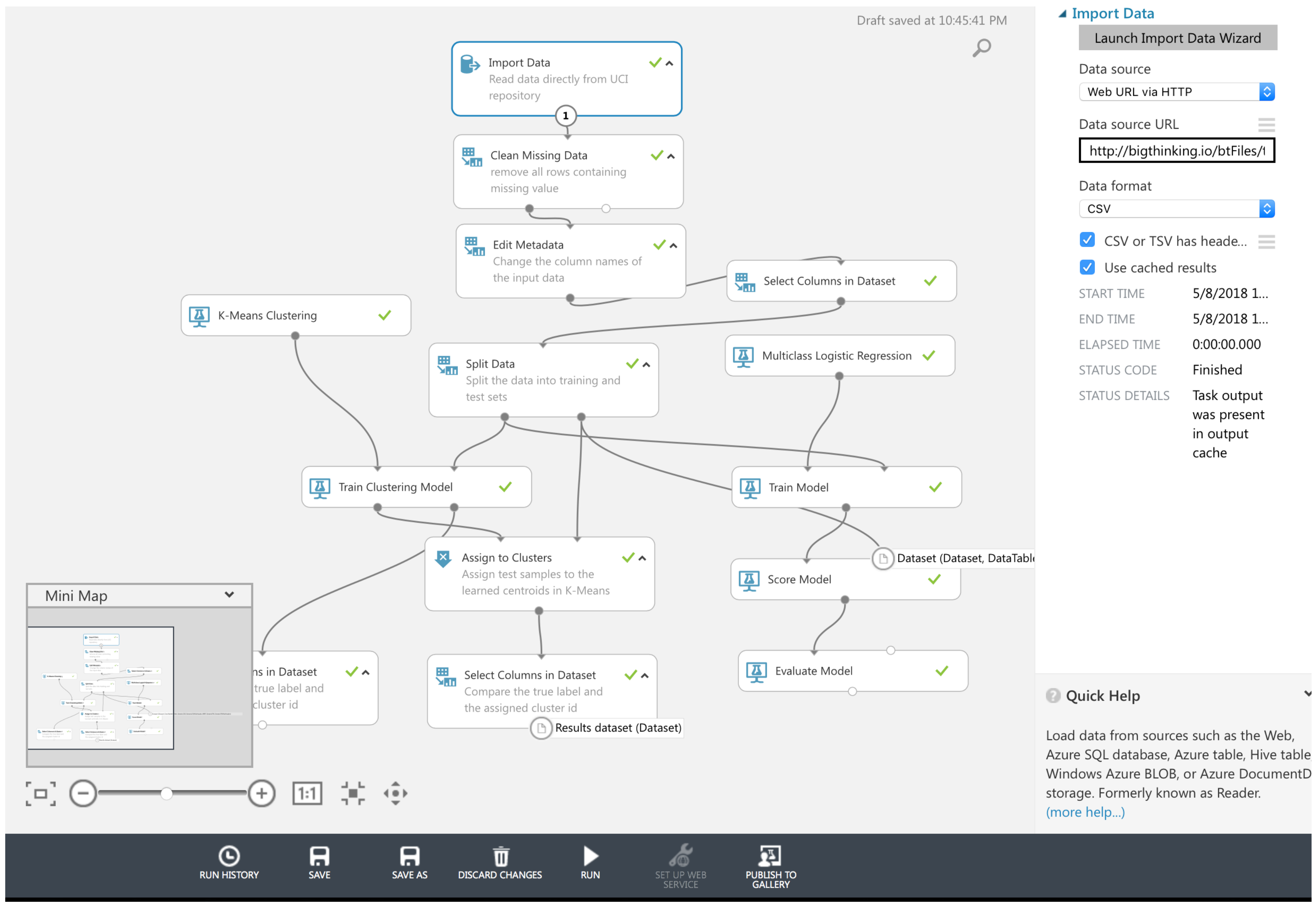
What is the business impact of the action?

(OPTIONAL) EXERCISE: DEPLOY

Using Azure ML Studio

COMPLETE WEB SERVICE EXERCISE

http://bigthinking.io/btFiles/training/MLIntroWorkBook.pdf



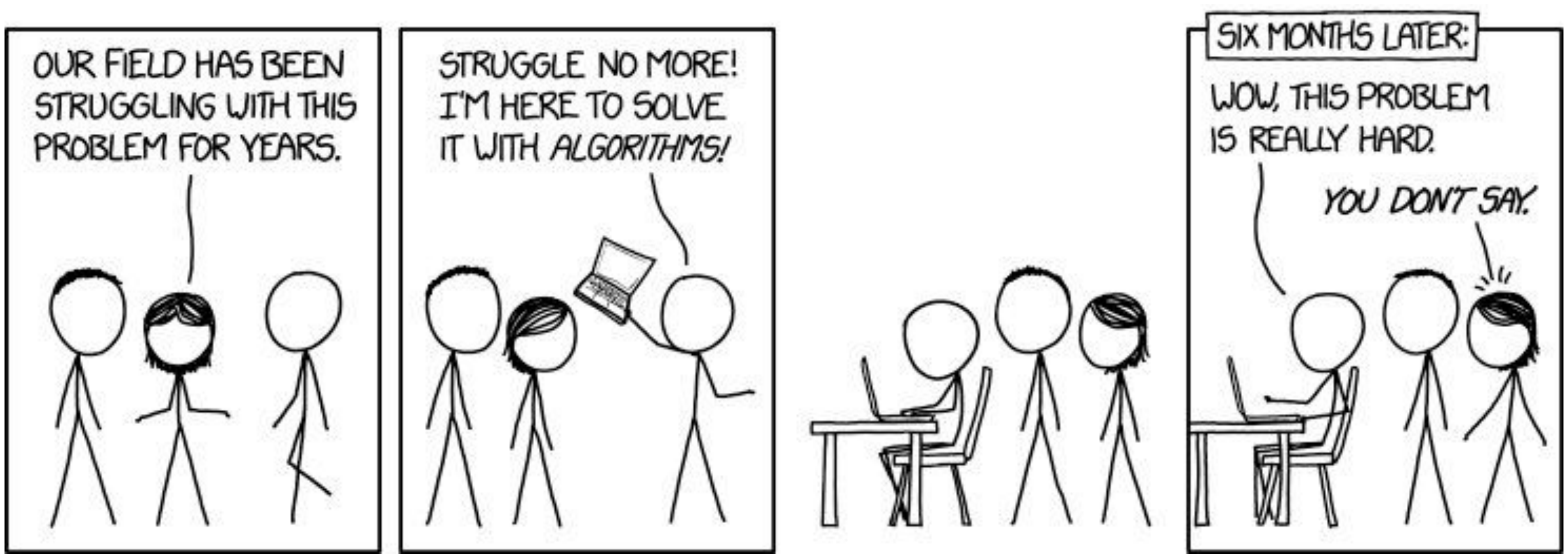
STEP 7

MONITORING OUTCOMES

MACHINE LEARNING : MONITORING OUTCOMES

Cause and Effect are not closely related in time and space

#MLPitfall : Failure to monitor models for adaptability problems & performance drift.



MONITORING CONSIDERATIONS

- Avoid “Drifting Goals” & react to subtle and gradual changes (see “The Boiling Frog” syndrome)
- Use a dashboard to visual model performance metrics
- Detect and act upon abnormal changes
- Track performance degradation over time
- Monitor process failures
- Use Cases for unintended consequences

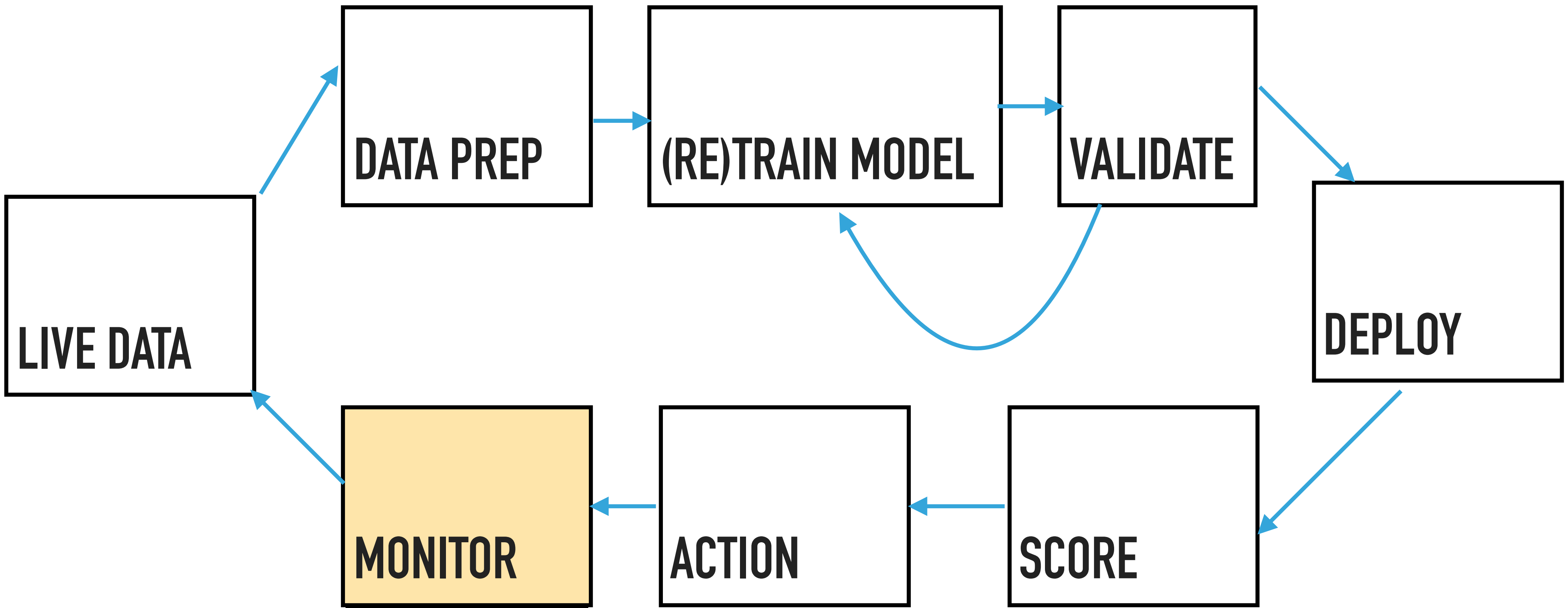
MONITORING : ANOMALY DETECTION

Classic Systems Archetypes

	Description	Mental Model	Key Strategy
Limits to Success	Success or growth is leveling off or declining.	"We'll get bigger and better by continuing to do more of what we are doing now."	Identify the limit that is causing the decline, then plan for that limit.
Success to the Successful	Decisions are being made in allocating resources, so that one party is getting attention and resources at the expense of another party.	"Because that person (department, project) is successful, they must be good and others are not."	Avoid win-lose situations in allocating resources. Find ways to make teams collaborators rather than competitors.
Tragedy of the Commons	Everyone is using a common resource that nobody owns. Overall usage goes up, but returns to individuals go down. Eventually, the resource may be destroyed.	"This resource belongs to me." Or "This resource is so vast that it'll never run out/collapse."	Identify the common resource and how people are drawing on it. Then work with users to plan how to allocate and/or limit access to the resource.
Growth and Underinvestment	We neglect or are unable to invest in the capacity to succeed.	"We don't need to invest in capacity; we can get through the present crunch by applying greater effort. We can invest down the line."	Identify the limited capacity that is causing the heroic efforts. Recognize the unintended consequences of the current course of action, then plan to invest in capacity or to deal with the consequences of choosing not to do so.
Fixes That Fail	All the quick fixes we have tried have worked at first but the problem keeps getting worse.	"Time is money, and neither time nor money should be wasted. Therefore, the first answer must be the right one."	Identify the quick fix and understand how it has undermined a long-term solution. Take robust actions that solve the problem once and for all.
Shifting the Burden	We know the fundamental solution, but are unwilling, or unable, to take it, so we implement a symptomatic solution and deal with the side effects.	"We know what we need to do, but it's too difficult to deal with, so let's put on a bandage instead."	Identify the addictive behavior to the symptomatic solution. Then commit to implementing the fundamental solution, no matter how difficult it may be.
Drifting Goals	We have lowered our standards to close the gap between the actual and desired performance.	"Our current level of activity is acceptable, even though it is below standard."	Identify the goal and how it has shifted. Recommit to or possibly redefine the goal. Then stay focused on the goal.

SYSTEM ARCHETYPES - Ideas for using Anomaly detection algorithms to flag unexpected weird changes in behaviors, over time.

MONITORING : MODEL PERFORMANCE



MONITOR & ANALYZE the models performance on production data, over time

DATA PREDICTION

ML TOOLS for using Anomaly detection algorithms to flag unexpected weird changes in behaviors, over time.

WHAT IS YOUR BIGGEST BARRIER TO DEPLOYING PRODUCTION-READY MACHINE LEARNING SOLUTIONS?

- Lack of Stakeholder Buy-in
- Unclear Use Cases for my industry/sector
- Data Acquisition - lack of access to the “right” data
- Data Science - limited understanding of ML algorithms & statistics
- Model Building Tools - limited experience using ML tools
- Insufficient IT/Architecture resources or personnel
- Other

DISCUSSION

REFLECTION & Q&A

DISCUSSION : REFLECTION AND Q&A

Reflection and Q&A

KISHAU ROGERS

- Background: Computer Science, Entrepreneur, 24yrs delivering software solutions
- Blog: www.bigthinking.io
- Email: kishau@bigthinking.io
- TwitterS: @kishau, @bigthinkingio
- Current Focus: Machine Learning @ time:study

Weill Cornell Medicine

NewYork-Presbyterian

Stony Brook Medicine

University of Colorado Hospital
UNIVERSITY OF COLORADO HEALTH

Columbia University Medical Center

Southampton Hospital

MedStar Washington Hospital Center

ADDENDUM

SUPPORTING RESOURCES

FOR SUPPLEMENTAL MACHINE LEARNING RESOURCES

http://MachineLearning.bigthinking.io