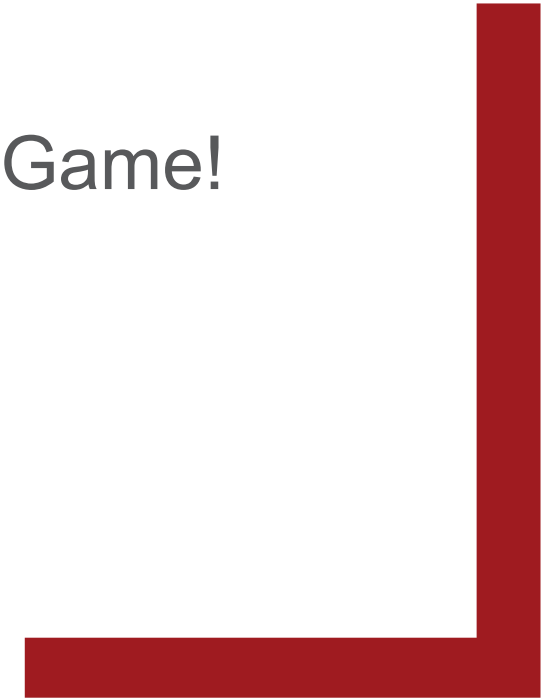
DATA SCIENCE
The Key to Upping Your Analytics Game!

2predict

Advanced Analytics Solutions



”AI is the New Electricity”

- Andrew Ng

Chief Scientist, Baidu
Founder, Google Brain
Co-Founder, Coursera
Adjunct Professor, Stanford

Agenda

- Quick Review – Key Concepts
- The Data Science Process
- Sample Use Cases
- Developing Your Data Science Strategy (and Know-How)
- Conclusion
- Q&A

Quick Review – Key Concepts

Rethinking The Role of Your Data

Question

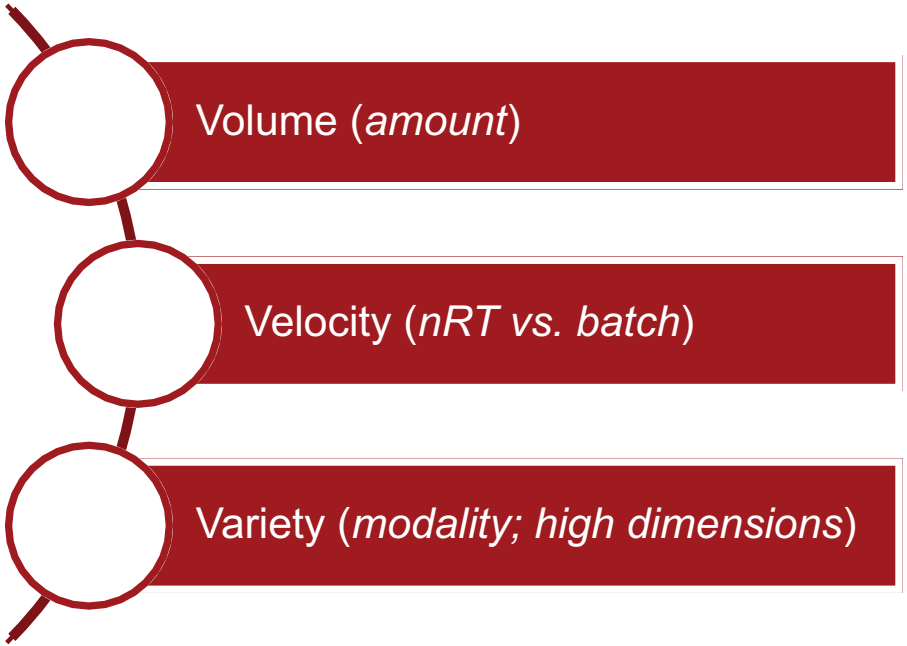
How many of you think the data you're collecting is valuable?

How do you know?

Recap: Big Data Definition

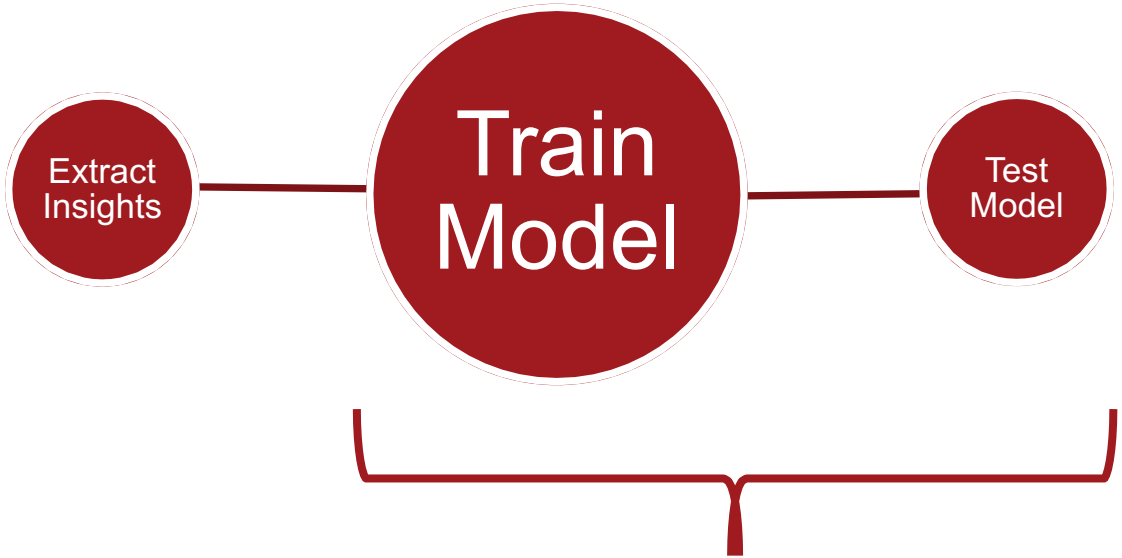
Data Generators

- **User Generated Content**
- **Sensors (IoT)**
- **Algorithmic Outputs**

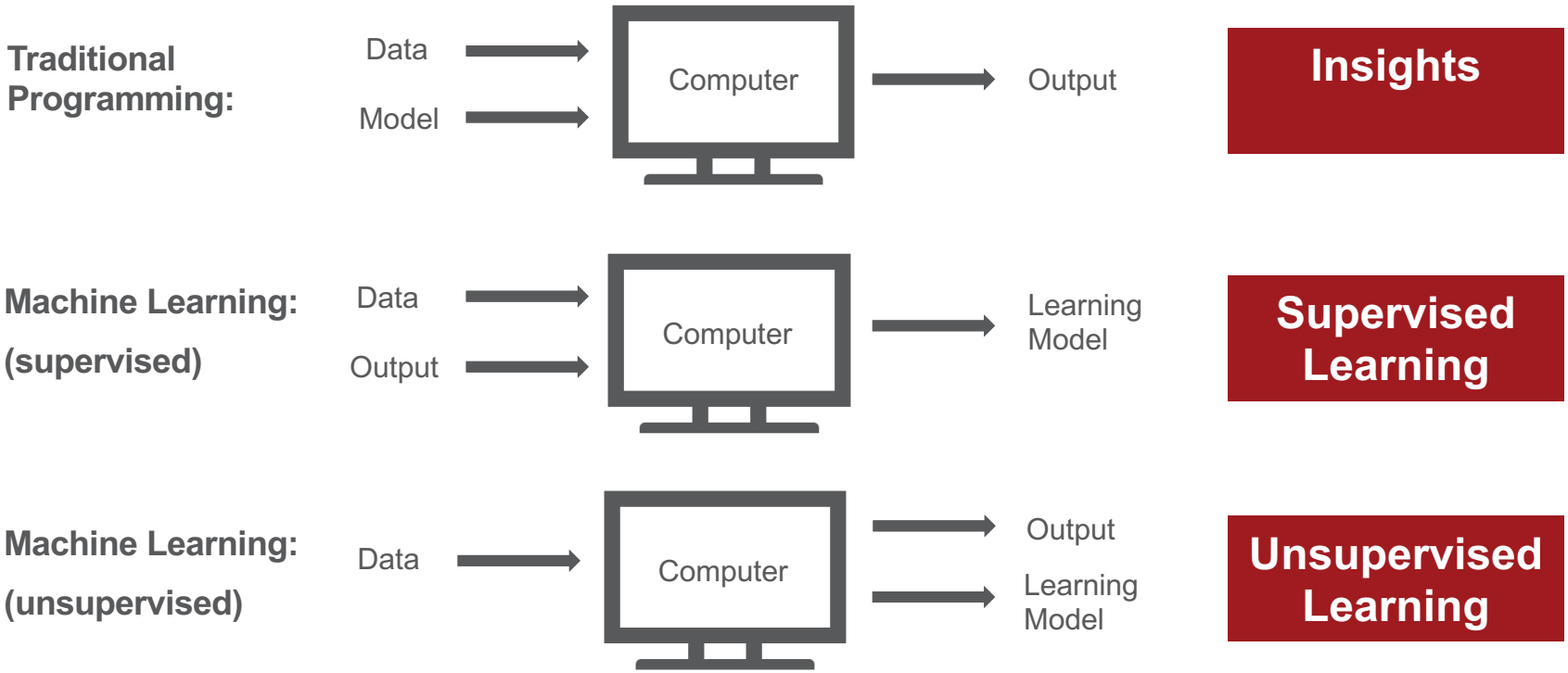


Problem: In the world of Big Data, how is it even possible to hypothesize?

So, What is Data's Role?



Learning - The New Compute Paradigm



In ML/AI, the Model is the “Output”

2predict
Advanced Analytics Solutions

The Data Science Process

From Data to Decisions

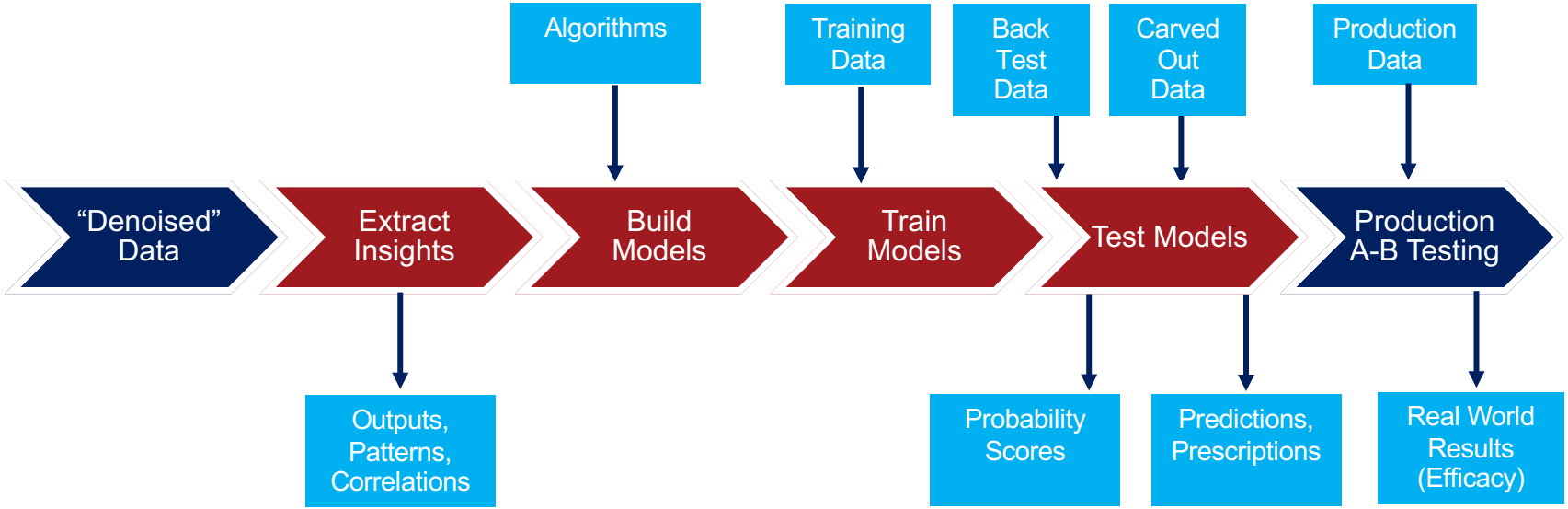
Question

Remember this?

ACTIONABLE **INSIGHTS**

How many of you are actually achieving this?
On a regular basis?

Typical Model-Based Systems Process



No overfitting!
Milestone based quick wins!

2predict
Advanced Analytics Solutions

Turnkey or Bespoke or Hybrid?

	Turnkey	Bespoke
Data Set Requirement	Common across vendor customer's datasets	Fairly unique to you
Insights	Good	Excellent
Industry Benchmarking	Good	Poor
Predictive Modeling	Poor (Limited)	Excellent - highly optimized (and thorough)

NOTE: For "turnkey", the longer the tool takes to train the models, the less turnkey it is.

Sample Use Cases

Disrupting Old Methods

Question

How many of you are seeing an ROI from your current analytics strategy?

Why Not?

Case Studies: AI Led Impacts

Old Methods	*New Methods*
Network Security: ”White” Lists + “Black” Lists	Model “normal” behavior; flag security anomalies
Credit Scoring: Based on past payment history; “to have credit, you need credit”	Model risk based on social physics datasets found on smartphone; geospatial and temporal
Call Center: Static call scripts; poor customer experience; poor NPS survey participation	Intonation (empathy) modeling to have dynamic conversations based on real time feedback of how conversation is going
Mental Health Detection: Q&A interview w/ a specialist - error prone; drug efficacy also error prone because of Q&A	Capture and diagnose “honest signals” via smartphone: GPS, call logs, phone logs, etc.

Impact

- Detect new types of intrusions; stay ahead of the curve
- Financial inclusion for the 2B+ unserved, including students
- Happier customer service experience; 100% automatic survey
- Less erroneous detection of mental health issues + data driven efficacy of prescriptions

2predict
Advanced Analytics Solutions

Predictive Modeling Use Cases
(Marketing & Supply Chain)

Marketing:

- ***Business Objectives:***
  - How do I increase MQL to Bookings ratio from 10% to 20%?
  - How do I increase MQL to Bookings velocity by 25%?
- ***Data Sources:***
  - Marketing Automation Platforms + CRM Platforms
- ***Predictions / Prescriptions:***
  - Create supervised and unsupervised learning models to predict campaign effectiveness forecasting including “nudges”

Predict campaign effectiveness, drive velocity and conversion rates

Supply Chain:

- ***Business Objectives:***
  - How do I increase my inventory turns ratio from current 4.0 to 4.5?
  - How do I reduce my expedited shipments by 50%?
- ***Data Sources:***
  - ERP and/or SCM Systems
- ***Predictions / Prescriptions:***
  - Create supervised learning models to perform SKU forecasting

Improve SKU forecasting, reduce costs, reduce working capital

2predict
Advanced Analytics Solutions

Predictive Modeling Use Cases
(Customer Service & Sensor Applications)

Customer Service:

- ***Business Objectives:***
  - How can I double my renewal rates?
- ***Data Sources:***
  - CRM + ticket handling systems
- ***Insights / Predictions:***
  - Extract meaningful correlations that can drive customers to renew; may be supervised or unsupervised learning

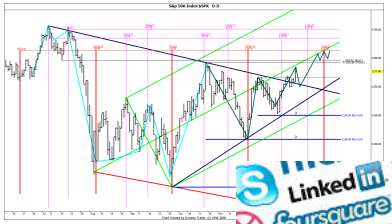
Understand what drives renewal rates and prescribe solutions to double such rates

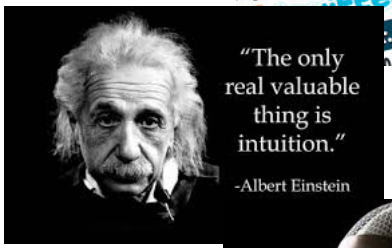
Sensor Applications (Horizontal Drilling):

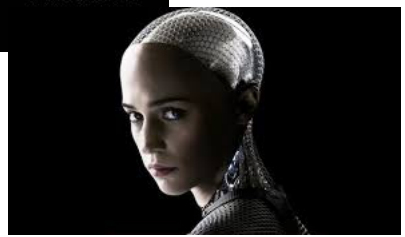
- ***Business Objectives:***
  - How do I reduce my drill bit failures by 50%?
- ***Data Sources:***
  - 100Hz time series drilling well data sets; topside drilling sensors
- ***Predictions / Prescriptions:***
  - Extract meaningful correlations and use deep learning model architecture to predict when drill bits will fail so that corrective action can be taken

Predict drilling failure, reduce costs, accelerate time to revenue

So What is Really, Really Difficult? (examples)







- Predicting Stock Market
  (HFT, sentiment, macro-, micro-fundamentals, geo-political, etc.)
- Behavior Modeling based on Social Media
  (no "honest signals")
- Intuition
- Singularity

2predict
Advanced Analytics Solutions

Creating your Data Science Strategy

Methods For Upping Your Decision Game

Question

How many of you are truly excited about achieving ACTIONABLE analytics?

Me Too!

Data Science Strategy Development Process

Key Inputs:
Vision/Goals, Metrics, SWOT Analysis

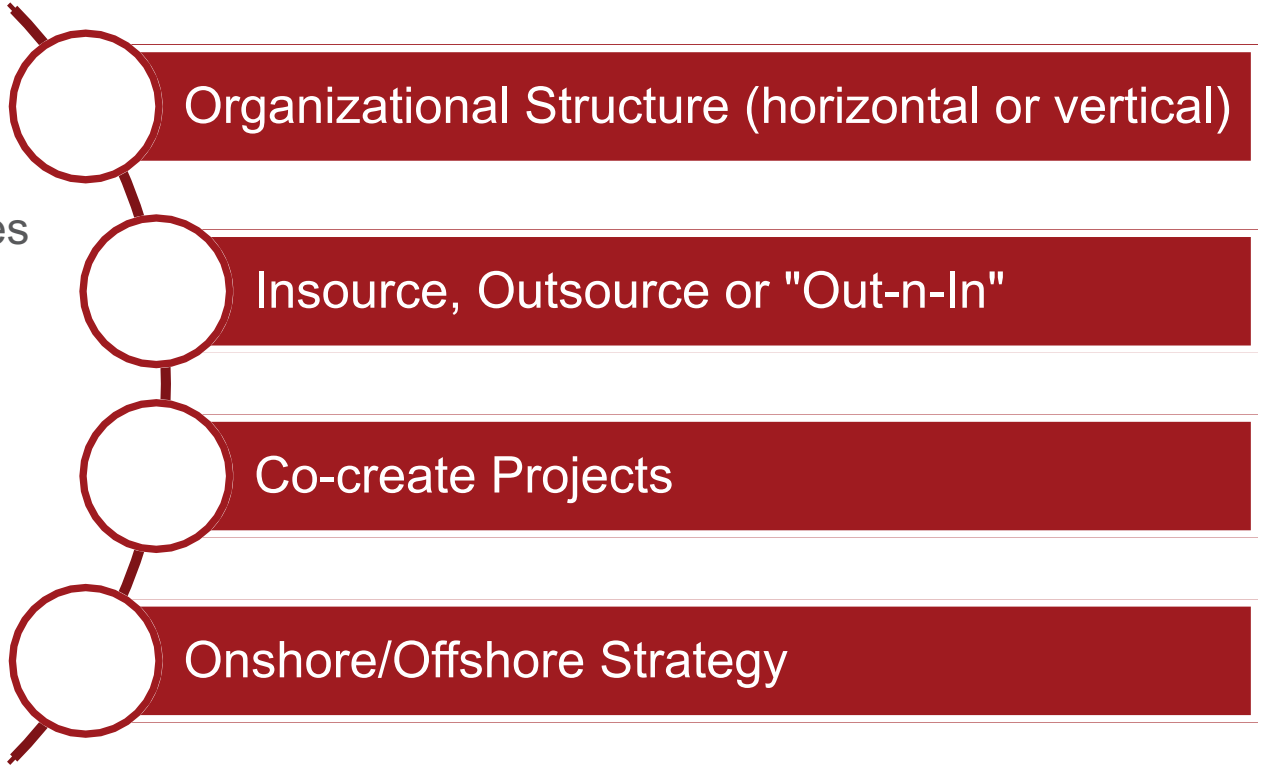
Key Outputs:
Strategy, Organizational Structure, Analytics Roadmap (inc. Quick Wins), Operating Plan

2predict
Advanced Analytics Solutions

Organizational (+Resourcing) Playbook

Considerations

- Abilities / Capabilities
- Risk Tolerance
- Timelines
- Learning Curve
- Cost



2predict
Advanced Analytics Solutions

Data Scientists Skills Comparisons

Requirements	Sr. Data Scientists	Sr. Software Engineers	Sr. Data Engineers
Modus Operandus	Scientists = “research”/ exploratory work (POI) driven	engineering milestones (POR) driven	engineering milestones driven
Education	Must have: Advanced Degrees (PhD, MS in Mathematics / Comp Sci)	Nice to have: Advanced Degrees	Nice to have: Advanced Degrees
Domain Knowledge	Nice To Have	Yes	Nice To Have
Language/Libraries	Python; R; TensorFlow; Notebooks	C/C++; Java; JavaScript; Scala	Python; C/C++; Java; JavaScript
Skills	Modeling; ML/DL/NLP algorithm development as applicable	Commercialization – deployment of code in end customer environments	Data Lake Architecture / Design; SQL, Hadoop; Basic Modeling; Basic ML; Data Mining;
U.S. Salary	$400K - $500K / year	$175K/year	$150K/year

Question

Too Daunting?

Don’t worry – we’ve got a playbook for you!

A “Walk Before You Run” Strategy

- ***What You Do Know:***
  - What business problems matter to you? Where you would really like to make data driven decisions?
- ***What You Don’t Know:***
  - How good your data sets are for answering the questions that matter to you.
- ***Solution:*** **Learn** before committing to a strategy

Key Outputs:

- Actionable analytics (proof that it won’t be waste of time)
- Extremely valuable data gap analysis (i.e. data audit)
- Extremely valuable knowledge (to help develop strategy)

2predict
Advanced Analytics Solutions

A Quick Win Strategy

Focus on:

ACTIONABLE **INSIGHTS**

Start w/ Valued Questions

1. Does online ad spend in CIO magazine result in MQLs?
2. How do we reduce our expedited freight charges?

Not Actionable

1. Yes – it is correlated
2. Here is an expedited freight SKU report from last quarter

Actionable

1. It is weakly correlated w/ a correlation coefficient of 0.24 – so probably better to reallocate the dollars
2. If we take action A, our predictive model shows, with 78% probability, that we can reduce expedited freight charges by $X on top 3 SKUs

In Conclusion

- ✓ Big Data (potentially) is a store of significant value
- ✓ AI skills is the key to extrapolating value from your dataset
- ✓ Modeling allows you to predict future and make confident decisions
- ✓ Data scientists possess key skills to create predictive models
- ✓ Your data science strategy must be dictated by your needs, capabilities and constraints

Era of “data driven decisions” has finally arrived!

Thank you!

Praveen Mandal
praveen@2predict.co

2predict
Advanced Analytics Solutions