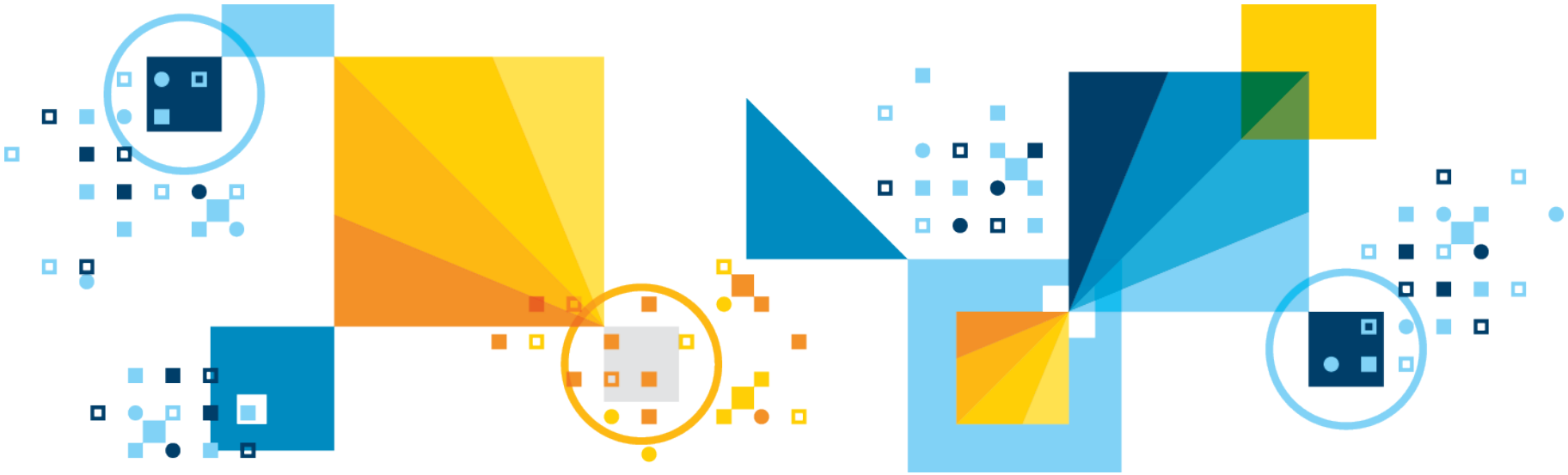
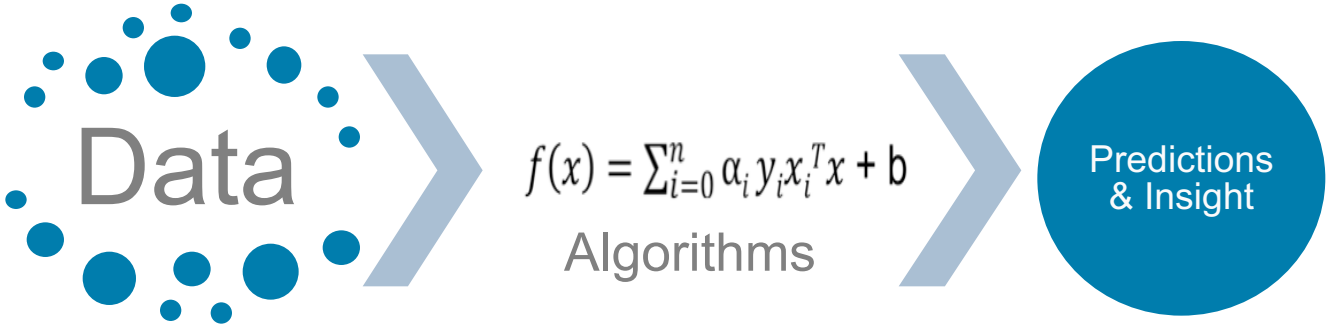
Introduction to Machine Learning

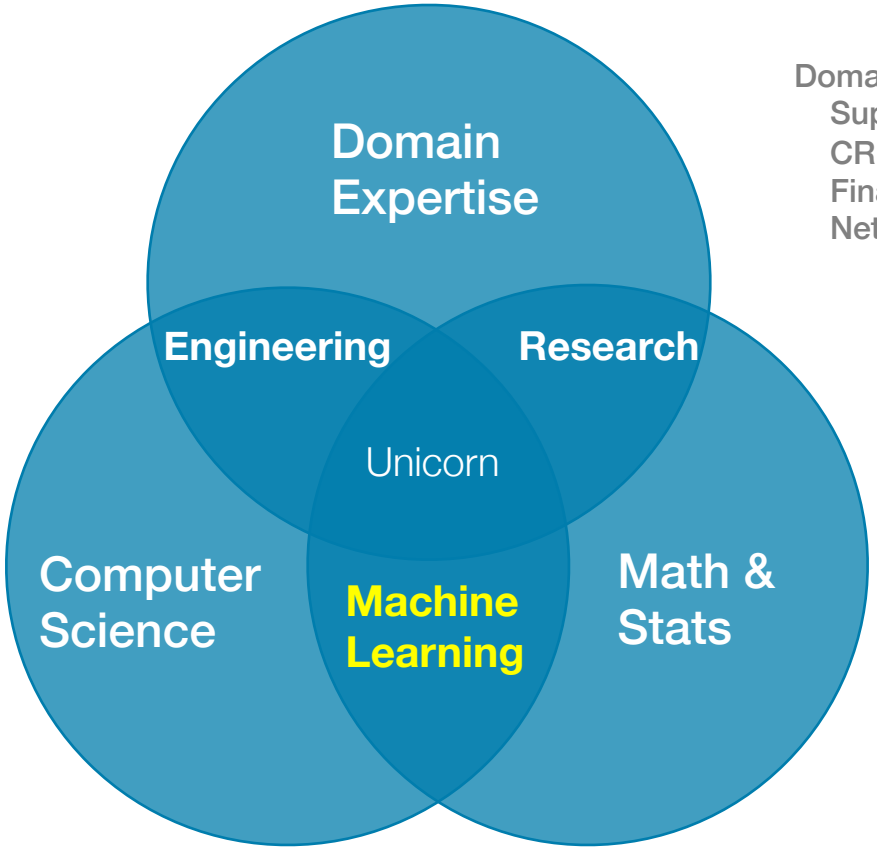


What is Machine Learning?

"Computers that learn without being explicitly programmed"
"Using algorithms to understand patterns in data"



Machine Learning in Data Science Echosystem….



Domain Knowledge
Supply Chain
CRM
Financials
Networking

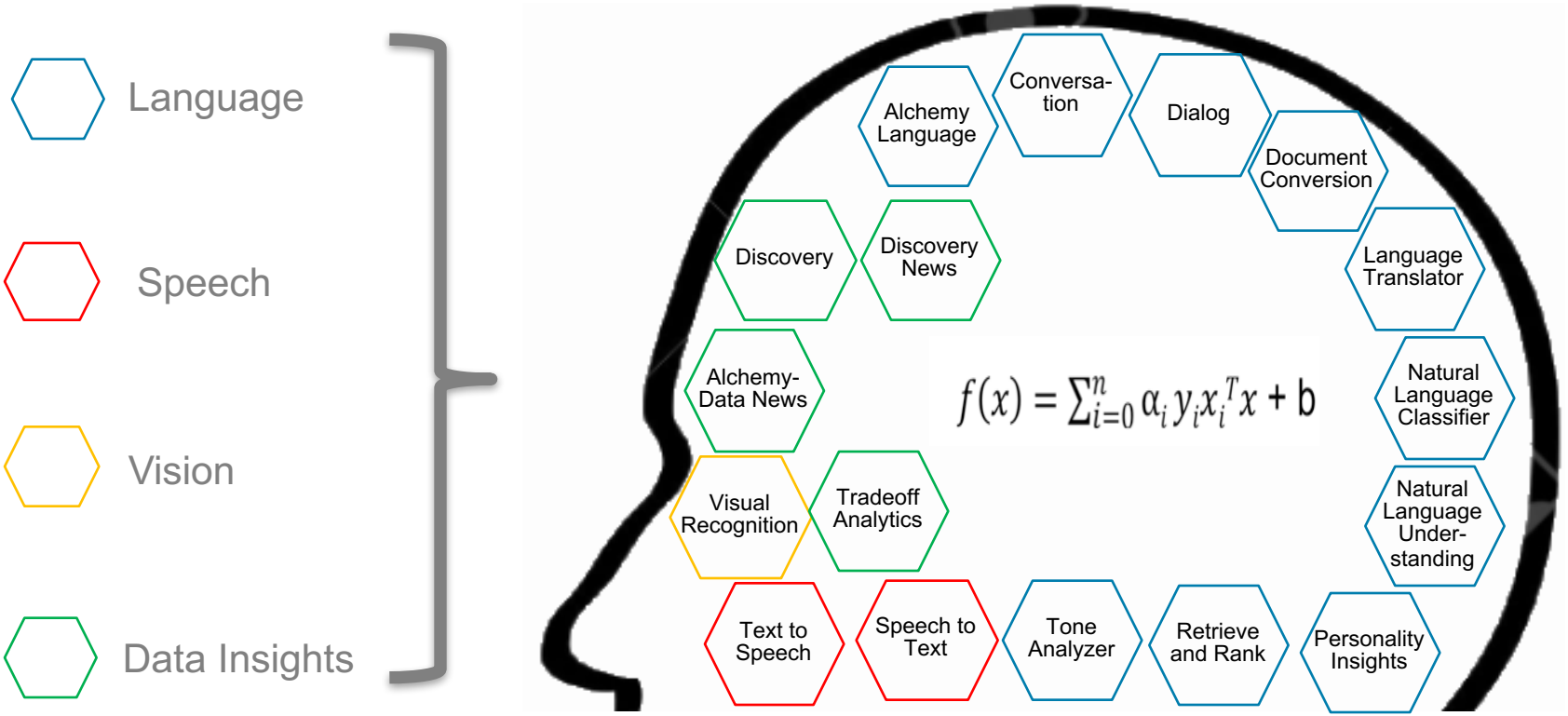
Scripting, SQL
Python, R Scala
Data Pipelines
Big Data/
Apache Spark

Mathematics
Computational

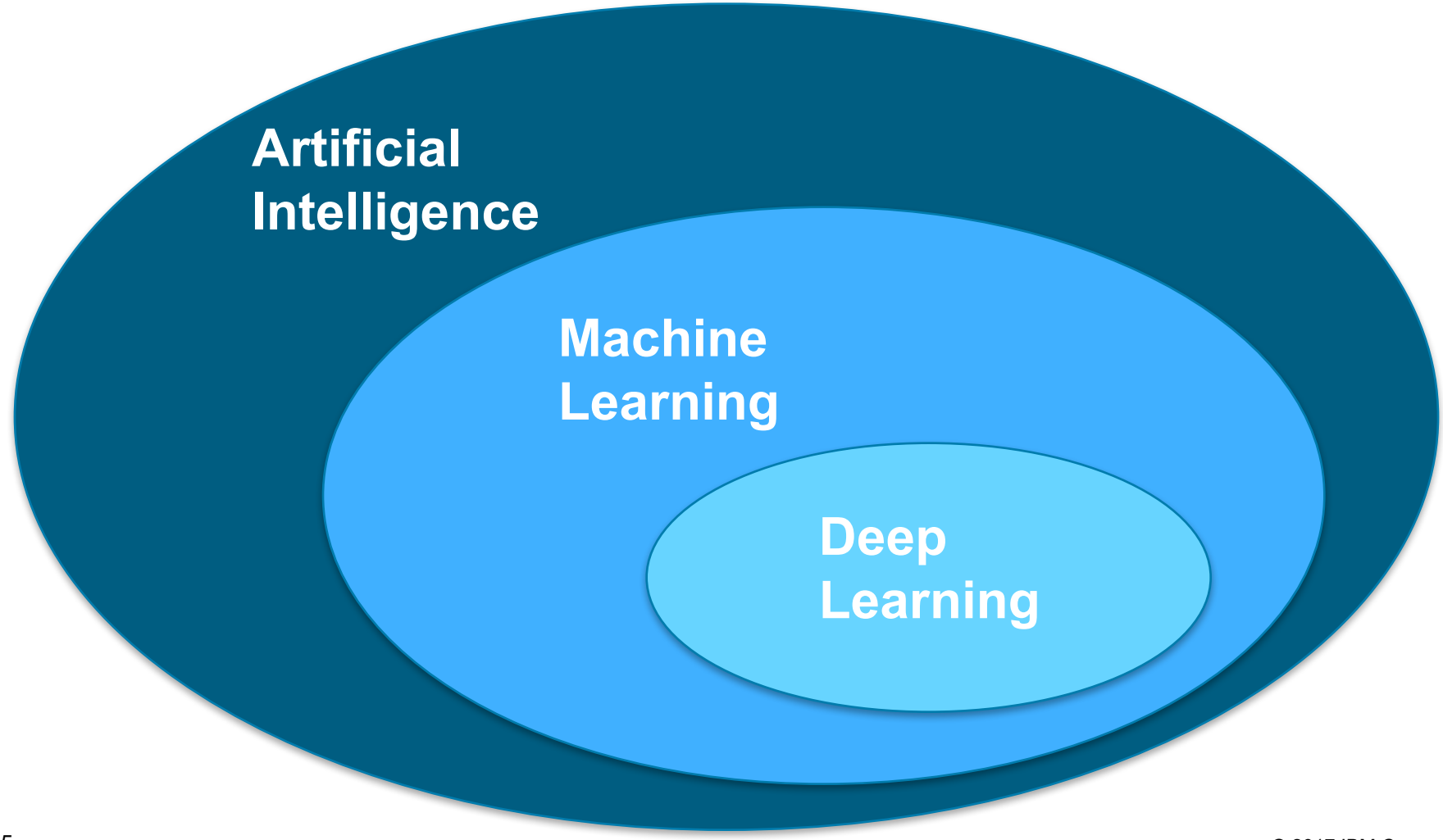
Data Science Projects Require multiple Skills

Machine Learning = Artificial Intelligence???

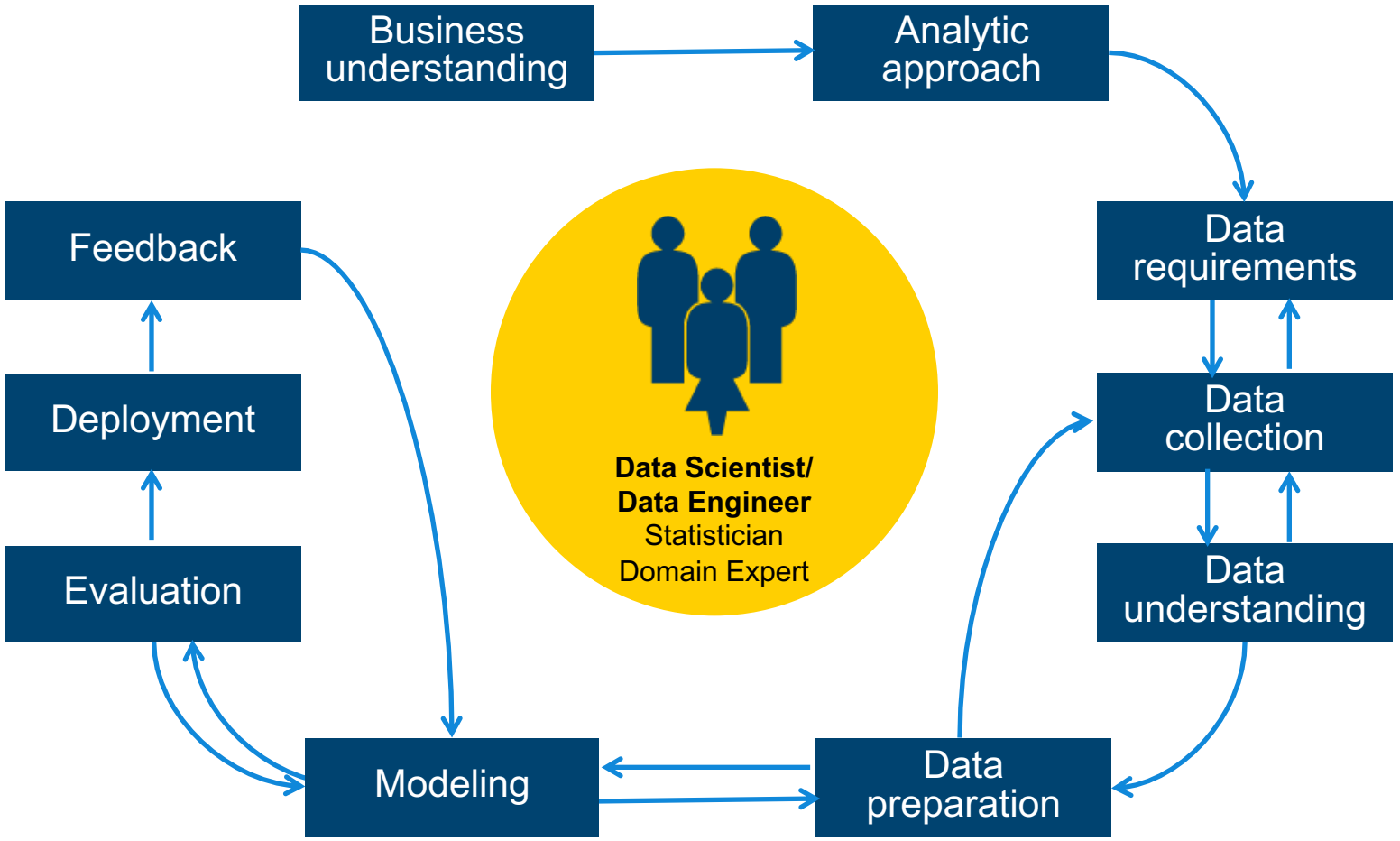
Data + Algorithms = Scored AI Models



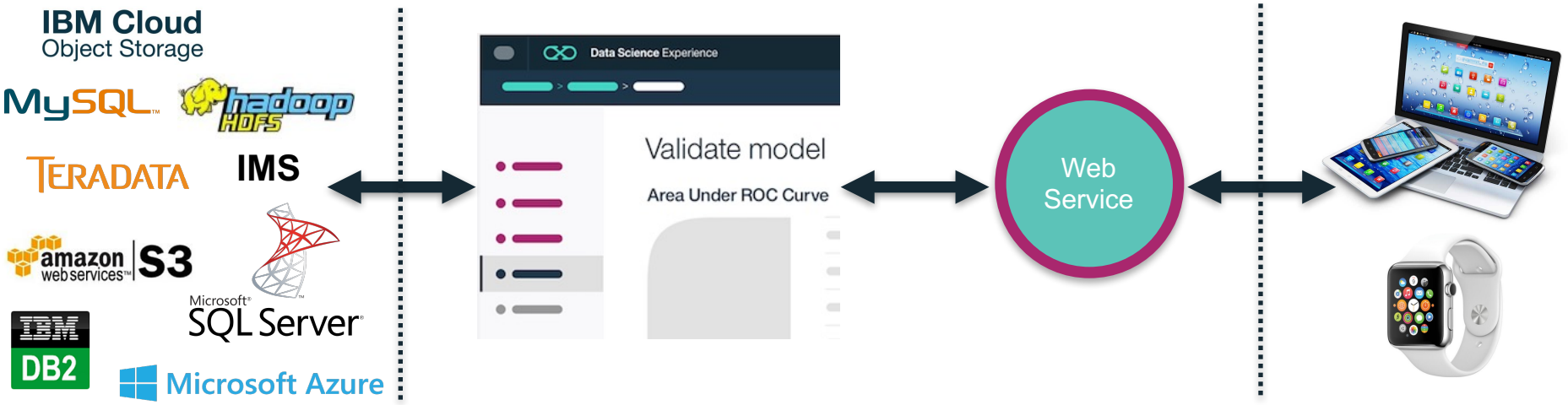
Understanding AI, ML & DL Relationship…



Methodology to create Machine Learning Model



Introducing IBM Watson Machine Learning Accelerating Digital Deployments in Cloud…



Data Access:

- Easily connect to Behind-the-Firewall and Public Cloud Data
- Catalogued and Governed Controls through Watson Data Platform

Creating Models:

- Single UI and API for creating ML Models on various Runtimes
- Auto-Modelling and Hyperparameter Optimization

Web Service:

- Real-time, Streaming, and Batch Deployment
- Continuous Monitoring and Feedback Loop

Intelligent Apps:

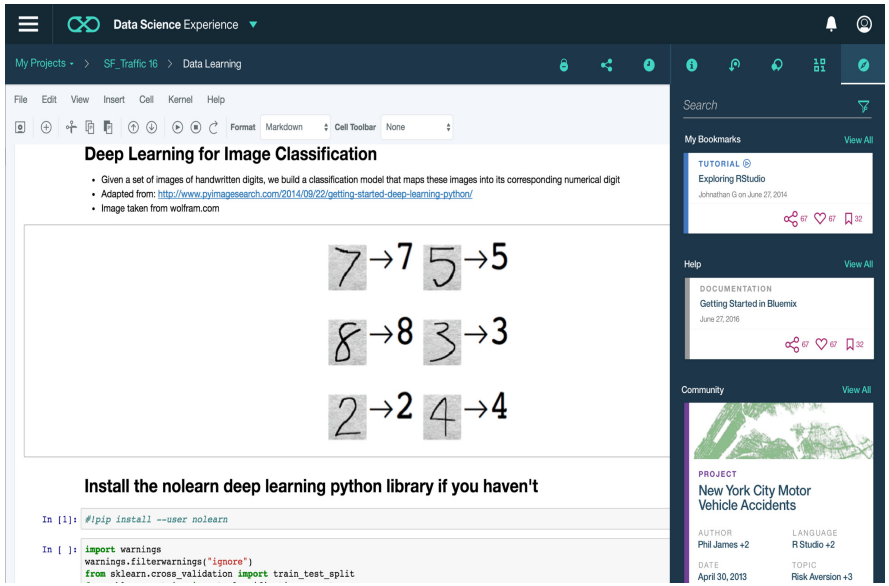
- Integrate ML models with apps, websites, etc.
- Continuously Improve and Adapt with Self-Learning

IBM Watson Machine Learning in Data Science Experience

Data Scientist

APIs for Jupyter Notebooks

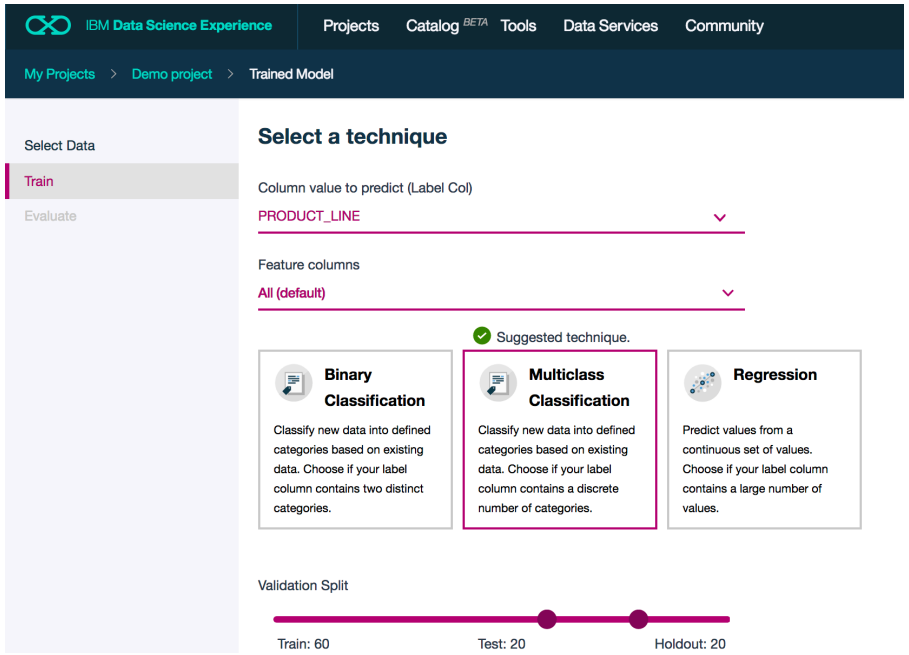
GA



Model Builder GUI

Lab 2

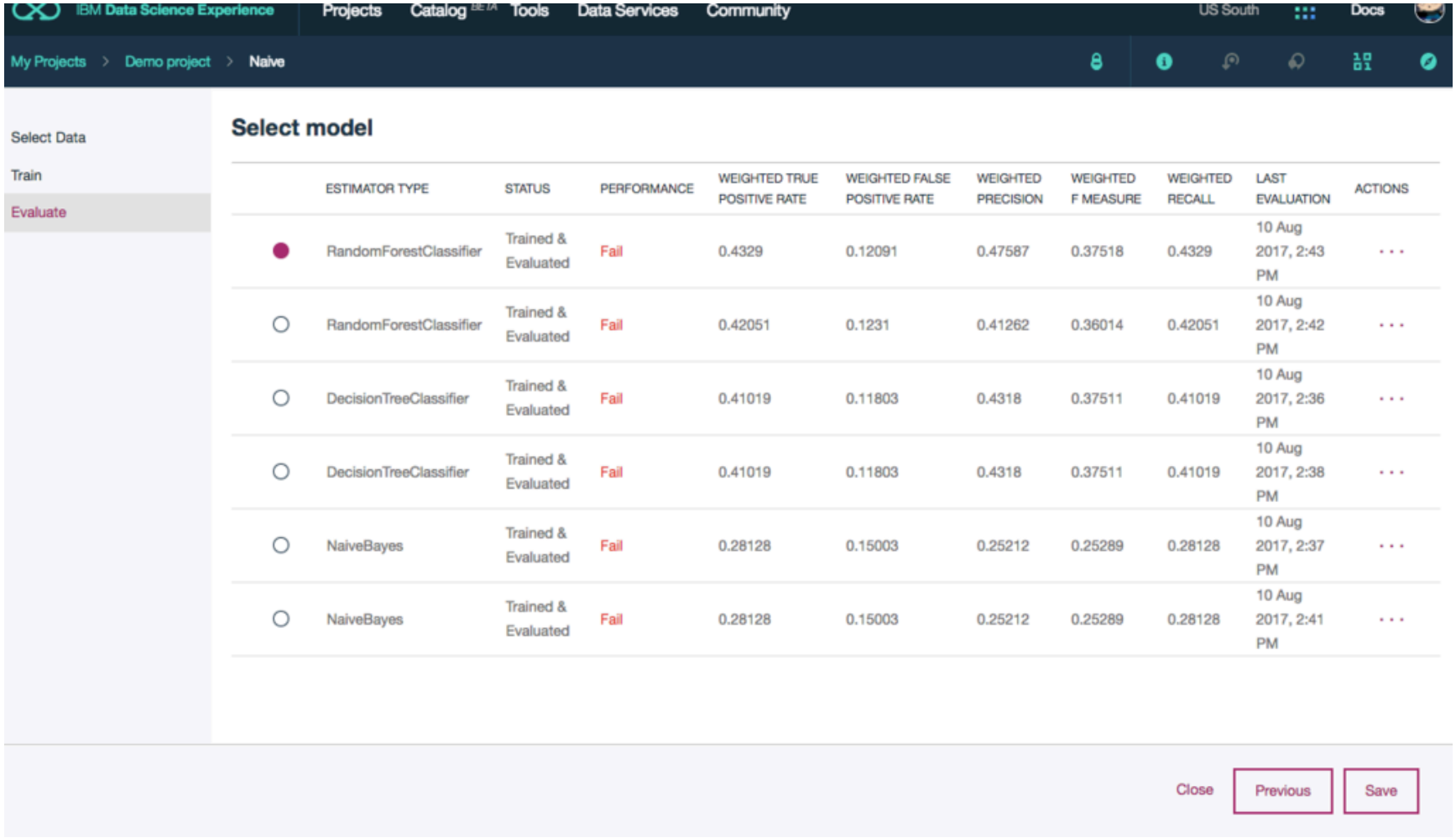
Open Beta



WML GUI - Train multiple models at same tim… System tells you best model…

Data Scientist

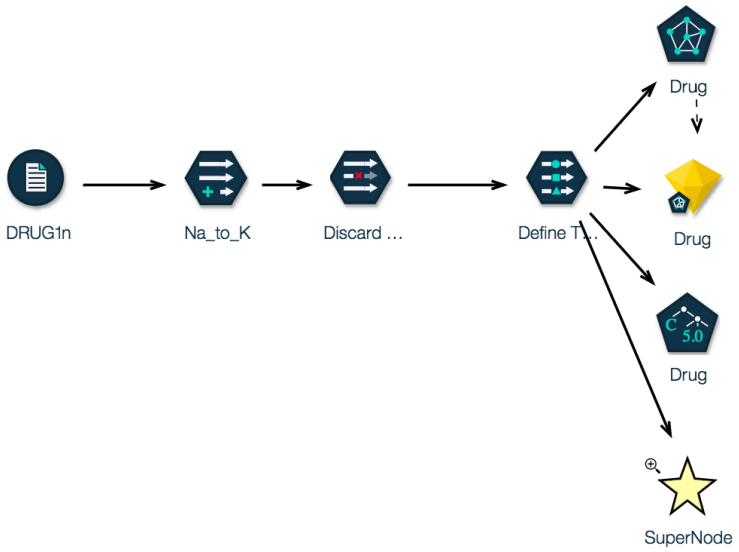
Open Beta

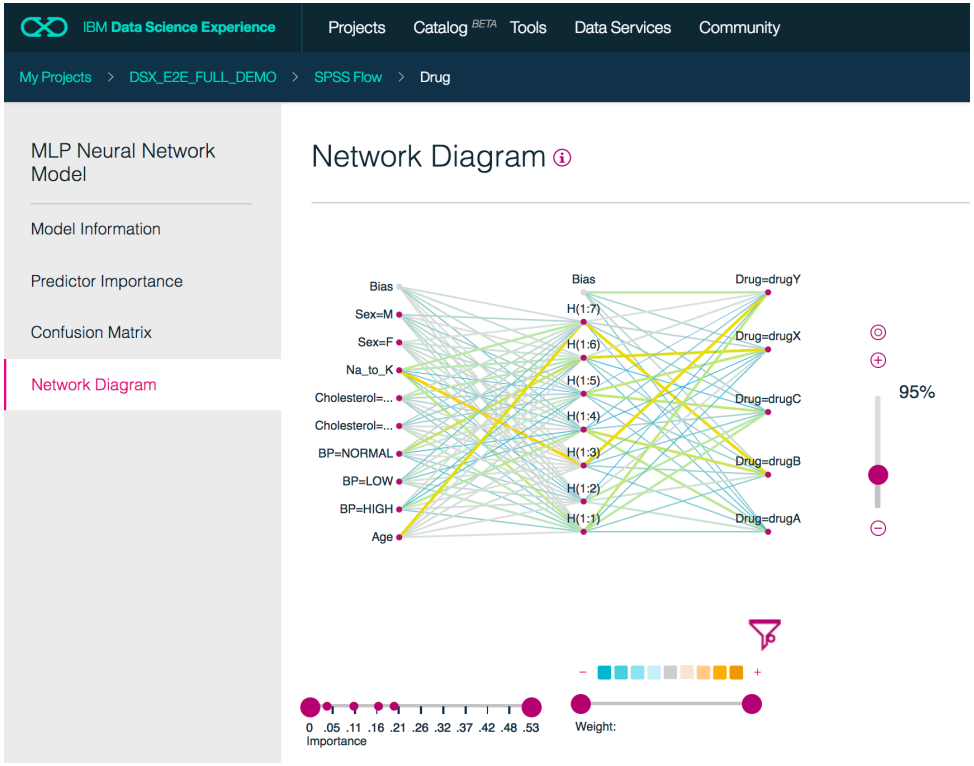


IBM Watson Machine Learning in Data Science Experience

Data Scientist

Flows *Open Beta*





Create Advanced Models without coding

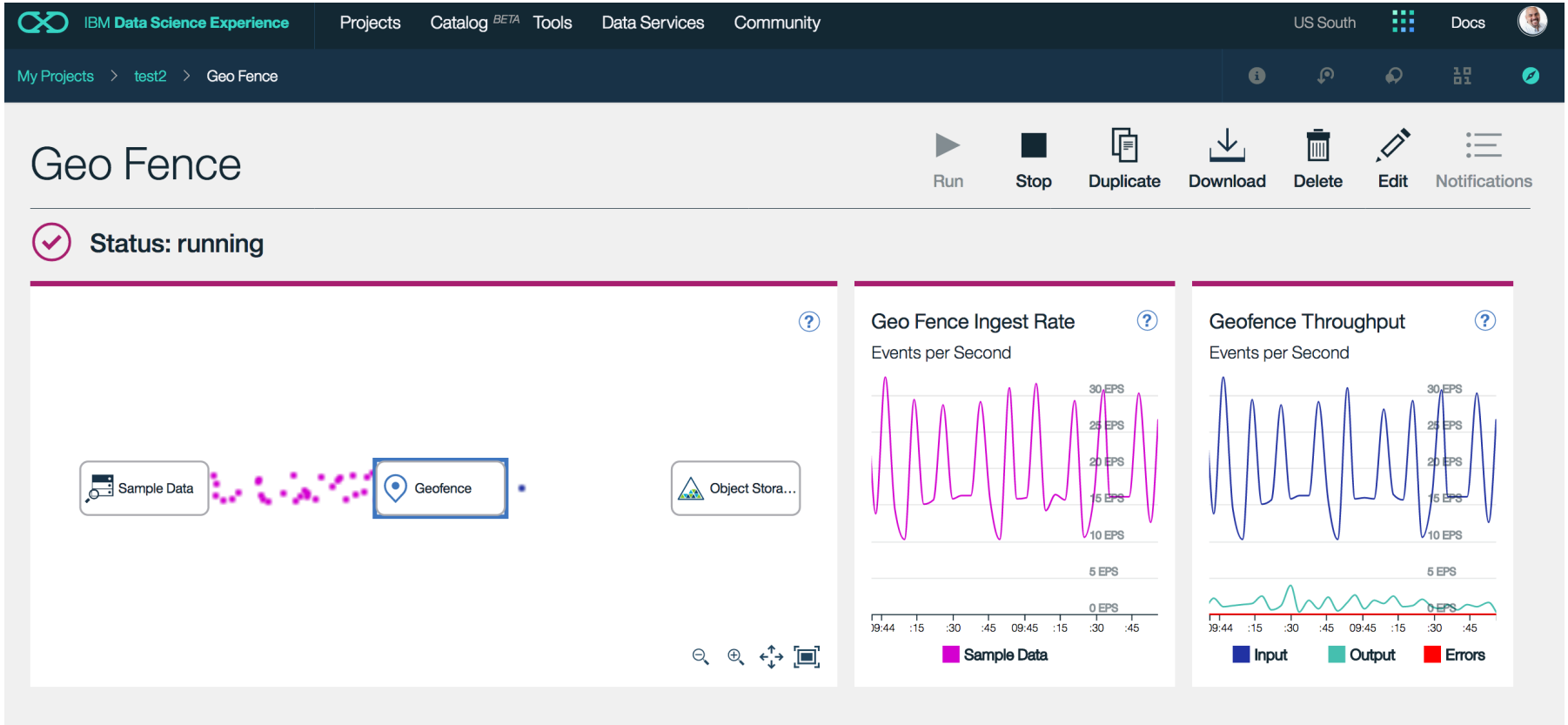
Advanced Model Visualization: Easy to understand their performance

IBM Watson Machine Learning in Data Science Experience

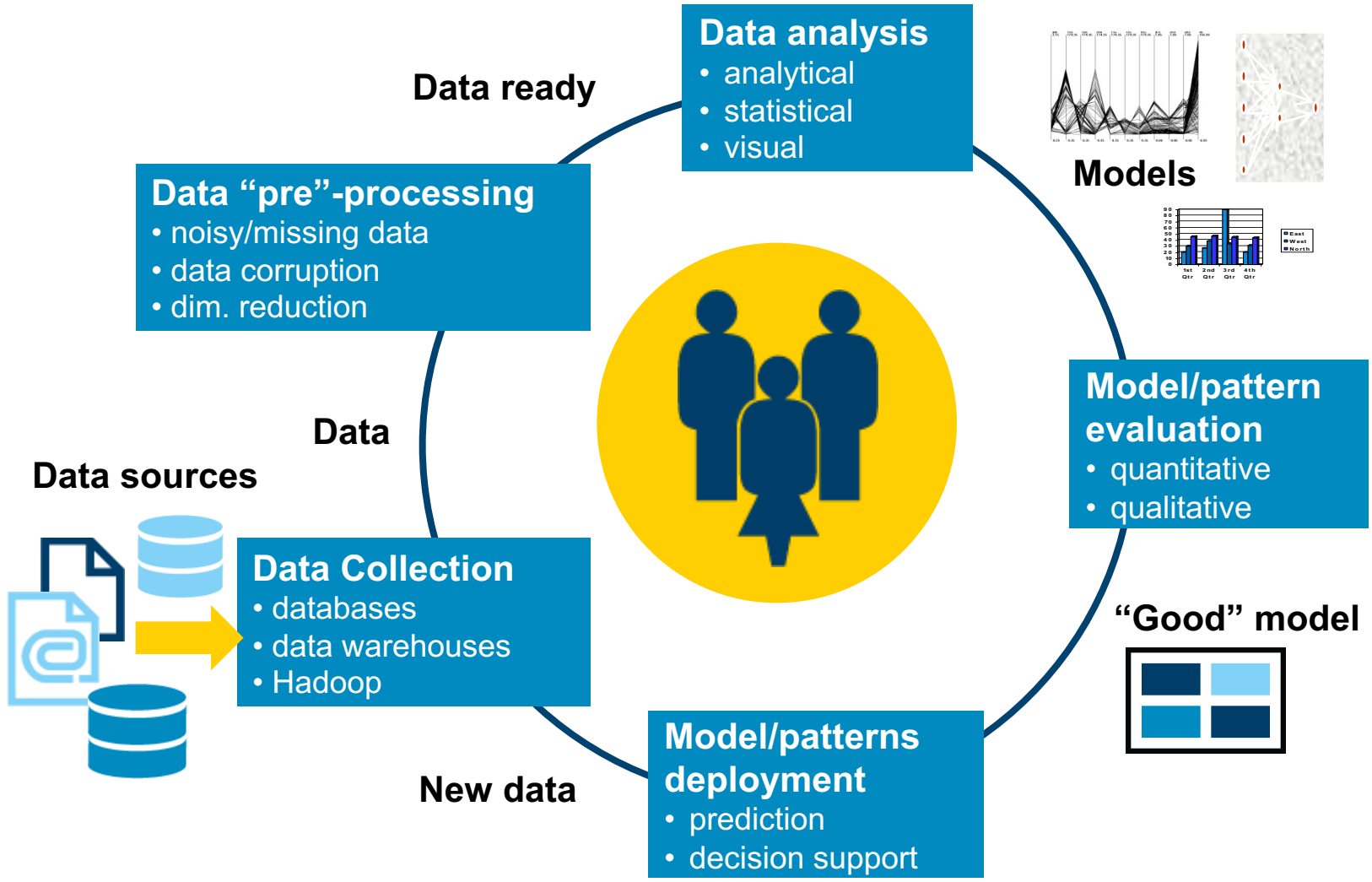
Data Scientist

Streaming Pipelines

Open Beta



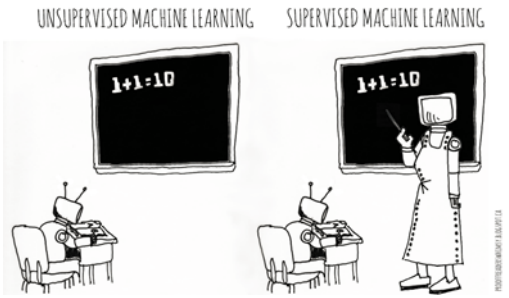
Data centric view of methodology



Categories of Machine Learning

- **Supervised learning**
  - The program is “trained” on a pre-defined set of “training examples”, which then facilitate its ability to reach an accurate conclusion when given new data
  - The algorithm is presented with example inputs and their desired outputs (correct results)
  - The goal is to learn a general rule that maps inputs to outputs

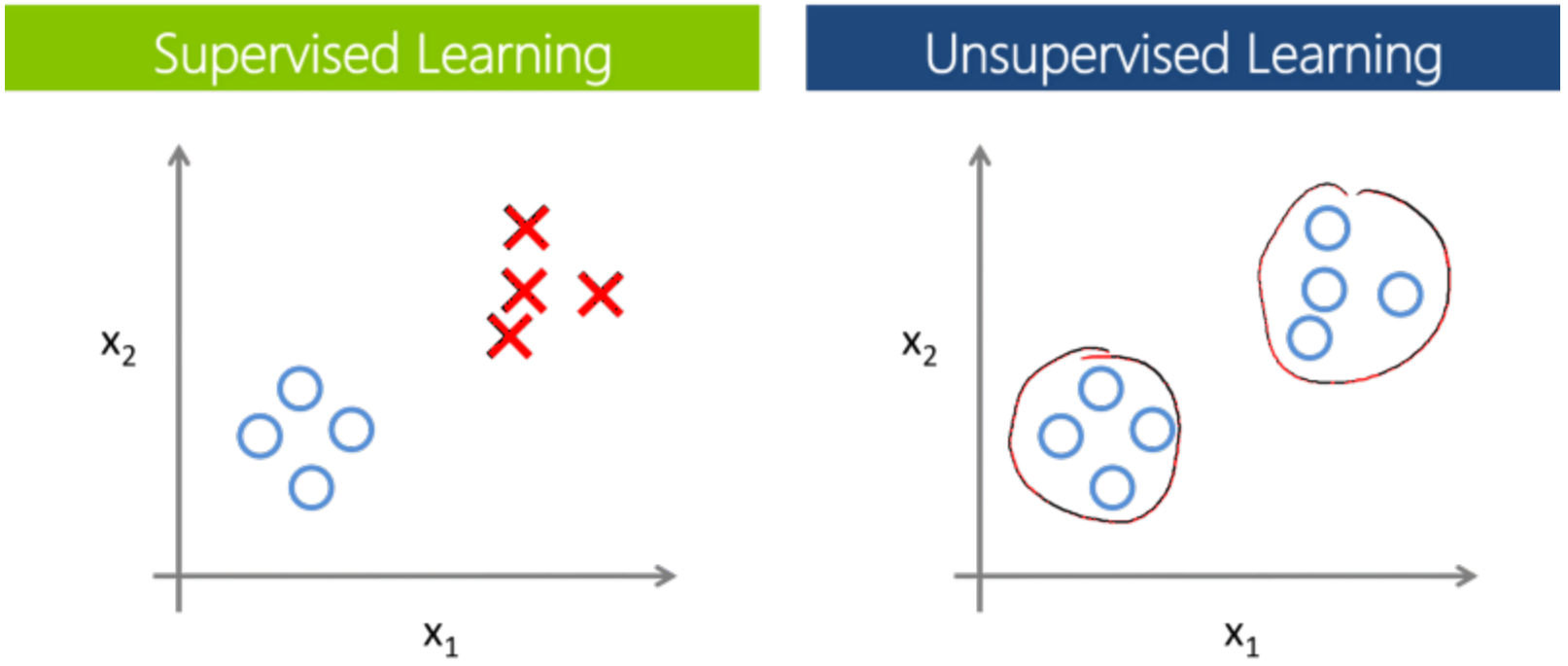
- **Unsupervised learning**
  - No labels are given to the learning algorithm, leaving it on its own to find structure (patterns and relationships) in its input
  - Unsupervised learning can be a goal in itself (discovering hidden patterns in data) or a means towards an end (feature learning)



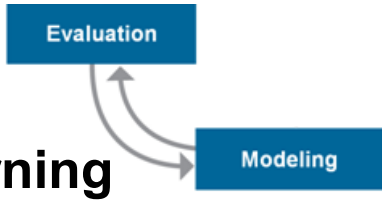
Categories of Machine Learning

	Discrete Output	Continuous Output
Supervised Learning (require Ground-Truth)	• **Classification** (outcome is discrete) • Binary Classification • Linear Models (Logistic Regression) • Decision Trees • Naïve Bayes • Multi class Classification • Decision Trees • Naïve Bayes • K-NN	• **Regression** - Linear - Ridge - Lasso • **Decision Trees** • Random Forest • Gradient Boosted Trees
Unsupervised Learning (no Ground-Truth data required)	• Clustering - k-means • FP-Growth	• Clustering - k-means - Gaussian Mixture • Dimensionality Reduction - PCA - SVD

Supervised vs. Unsupervised Learning



Training, testing, & validation sets



- **During the model development process, supervised learning techniques employ training and testing sets and sometimes a validation set.**
  - Historical data with known outcome (*target*, *class*, *response*, or *dependent variable*)
  - Source data randomly split or sampled… mutually exclusive records
- **Why?**
  - Training set ➔ build the model (**iterative**)
  - Testing set ➔ tune the parameters & variables during model building (**iterative**)
    - Assess model quality during training process
    - Avoid overfitting the model to the training set
  - Validation set ➔ estimate accuracy or error rate of model (**once**)
    - Assess model's expected performance when applied to new data

Machine Learning – A more formal definition

Tom Mitchell of Carnegie Mellon University provides a widely quoted, more formal definition of machine learning

"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"

Q & A