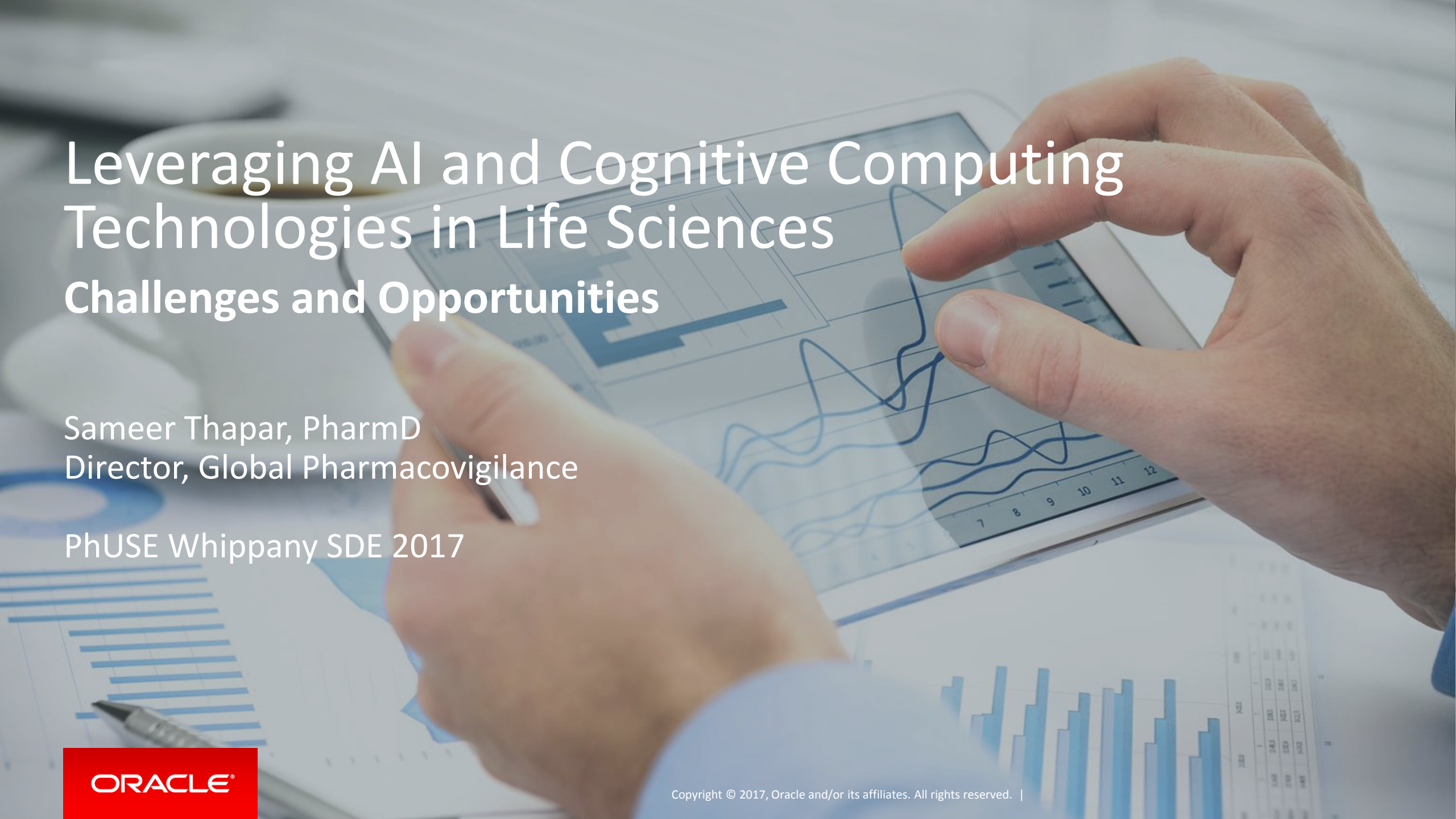


Leveraging AI and Cognitive Computing Technologies in Life Sciences

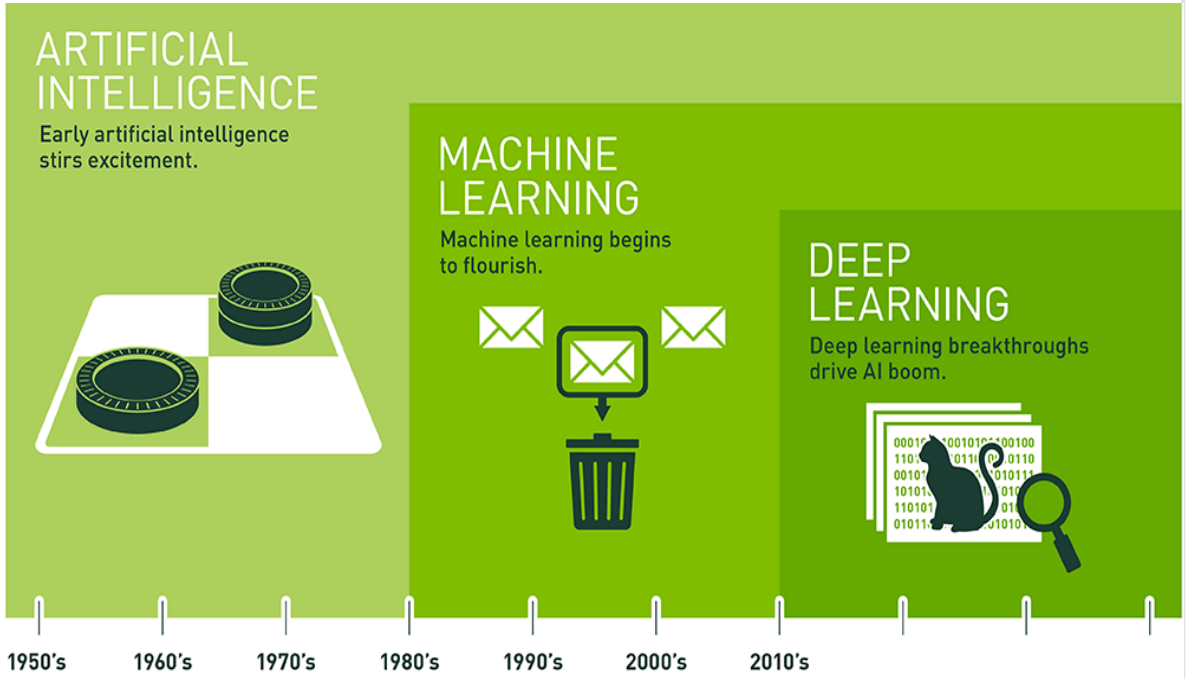
Challenges and Opportunities

Sameer Thapar, PharmD
Director, Global Pharmacovigilance

PhUSE Whippany SDE 2017

ORACLE

Basics: AI versus ML versus DL



- **Artificial Intelligence:** computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.
- **Machine learning**: subset of AI having the ability to automatically learn and improve from experience without being explicitly programmed.
- **Deep Learning**: subfield of machine learning concerned with algorithms called artificial neural networks.

Overview Only

This presentation will not go into detail on the extensive AI/CC/ML platforms and toolkits

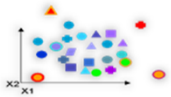
Classification

- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine
- Gaussian Mixture Models

Clustering

- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)

Anomaly Detection



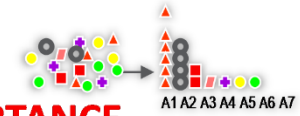
- One-Class SVM

Time Series

- Single, Double Exp Smoothing

Regression

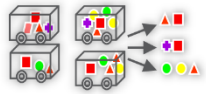
- Linear Model
- Generalized Linear Model
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- LASSO



Attribute Importance

- Minimum Description Length
- Principal Comp Analysis (PCA)
- Unsupervised Pair-wise KL Div

Association Rules



- A priori/ market basket

Predictive Queries

- Predict, cluster, detect, features

SQL Analytics

- SQL Windows, SQL Patterns, SQL Aggregates

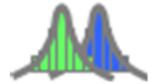
Feature Extraction

- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

Text Mining Support

- Algorithms support text type
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA) for document similarity

Statistical Functions



- Basic statistics: min, max, median, stdev, t-test, F-test, Pearson's, Chi-Sq, ANOVA, etc.

R Packages

- CRAN R Algorithm Packages through Embedded R Execution
- Spark MLlib algorithm integration

Case Study: Automotive Industry Social Text Mining

Using Social Listening Algorithm to Keep Customer Centric Focus

Automotive Industry

- A Large Automotive Company wanted to know when its customers talk about car problems
- Attempt to gauge context and sentiment
- Social Data Listening Platform helps companies monitor social media conversations around their brands
- A simple, supervised, document classification example

"Yooo the bakes on the Camaro stopped working, we ran like 4 stop signs & almost died lmaooo"

"Well the Cadillac wont start"

"I realized that money doesn't buy happiness but I'd rather do my crying in a Corvette."

Automotive Case Applied to Life Sciences

Challenges

- Regulations on company branded communications to public
- Potential of adverse event or inadvertent medical communication and handling
- Once you know of disinformation, what are next steps?
- How to overcome the nuances of specialized lexicons, medical dictionaries, and coding?

Opportunities

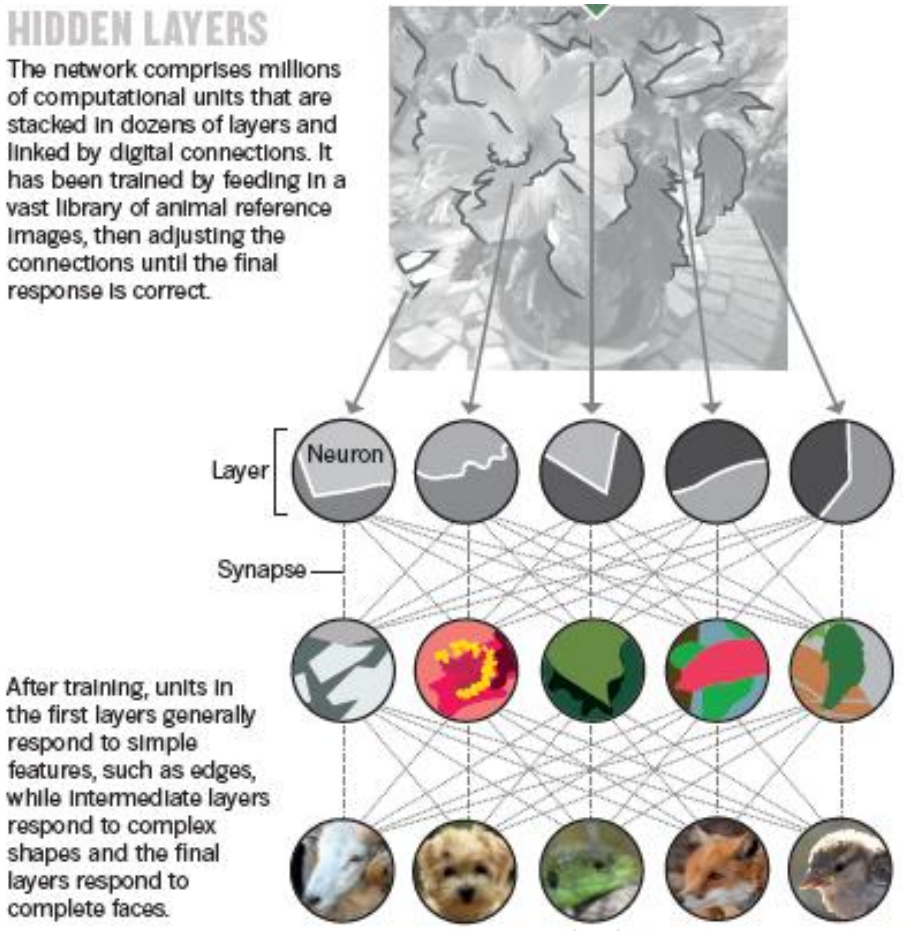
- Ability to gauge and prevent negative sentiments from proliferating
- Trend analysis insight
- Identification of brand advocates

Case Study: Facebook & Google Image Processing

Using Deep Learning toolkits to decipher and classify images

Facebook & Google Image Classification

- Facebook Research accomplishes an internet-size image classification in an hour with 90% accuracy
- Six years ago, same feat involved days of effort
- Google's open source Deep Dream software is designed to detect faces and other patterns in images, with the aim of automatically classifying images



NATURE | VOL 538 | 6 OCTOBER 2016

Facebook & Google Case Applied to Life Sciences

Challenges

- Current proof of concept involves understanding the “black box” of how the AI classifies
- Transparency in current deep learning algorithms is improving but not at confidence levels for science
- Trust in machine derived analysis over human inferred ones

Opportunities

- When used as an adjunct to data discovery processes, it can narrow the choices
  - Breast cancer imaging
  - Stroke radiographs
- Can be reverse engineered by feeding it a set of known artifacts and asking it to “find” similar sets

Case Study: Candidate Search and Talent Monitoring

Using AI for predictive analysis and employee hiring/retention

Candidate Search & Talent Monitoring

- AI service lets companies analyze workers' email to tell if they're feeling unhappy about their job
- AI will collect and crunch data from a variety of sources to find people with the right talents, with experience to match
- Use AI to find employees who may be likely to jump ship based on variables

HR Case Applied to Life Sciences

Challenges

- Privacy is an obvious concern when tracking employees just as HIPAA is when screening patients for trials
- Systems that sort people based on attributes also raise questions on the rigors of the parameters used
- Perceived bias in AI versus Human deduction

Opportunities

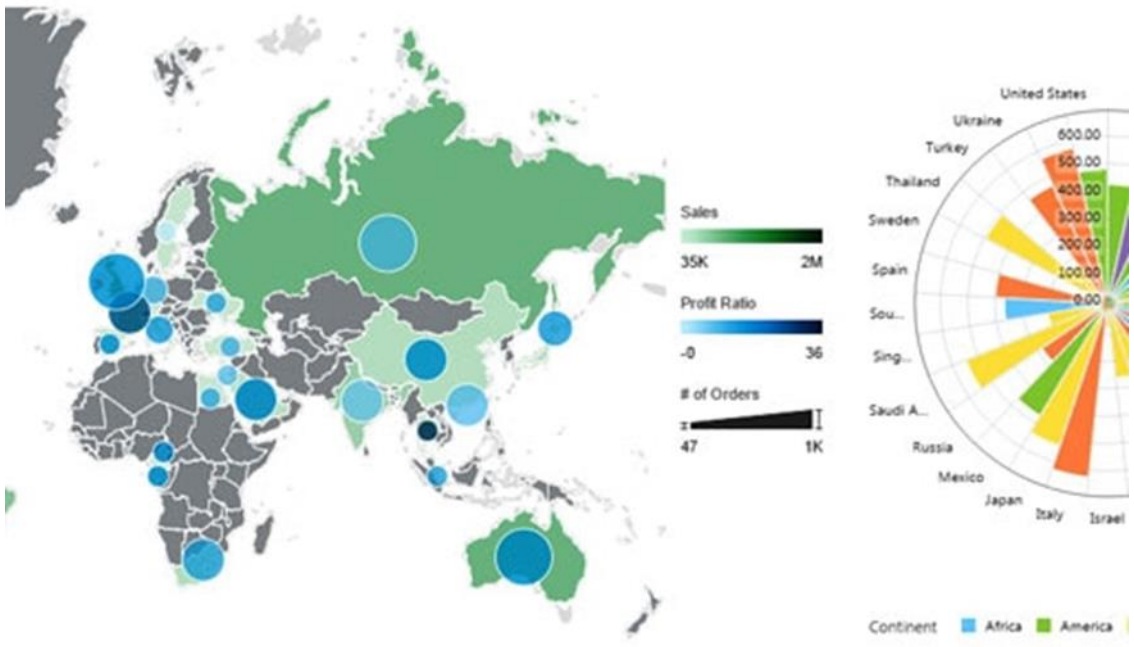
- AI systems as advisors by suggesting a set of options to human users, who ultimately make the final decisions (ie, PV)
- Can help organizations with decisions related to funding for clinical trials by offering new insights into the primary investigator, scientific topic and operational feasibility
- AI continually evolves based on new information, results and actions

Case Study: Financial Voice and Data Analysis

Using NLP to analyze structured and unstructured data

Financial Voice and Data Analysis

- Large global bank wanted its phone calls and voice records screened and analyzed into structured datasets
- Included multiple languages
- ML collapsed the *"raw data actionable insights/predictions"* process from 4-6+ weeks to mins/hours



Financial Case Applied to Life Sciences

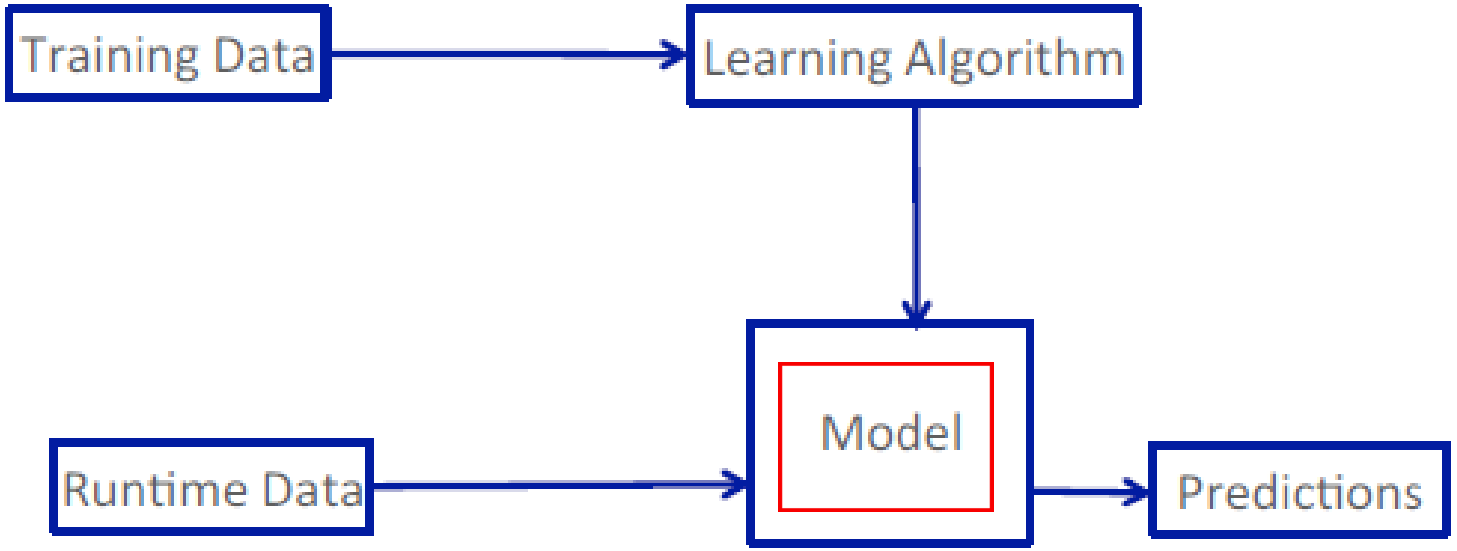
Challenges

- Multiple domains in Life Science industry, so data is in silos and bottlenecked
- Tapping into unstructured and structured datasets ideally requires a single source of truth to exist company-wide
- “Psychological embrace” – doubt that any AI/ML/DL can perform to the same accuracy and specificity as human who has learned the role via experience

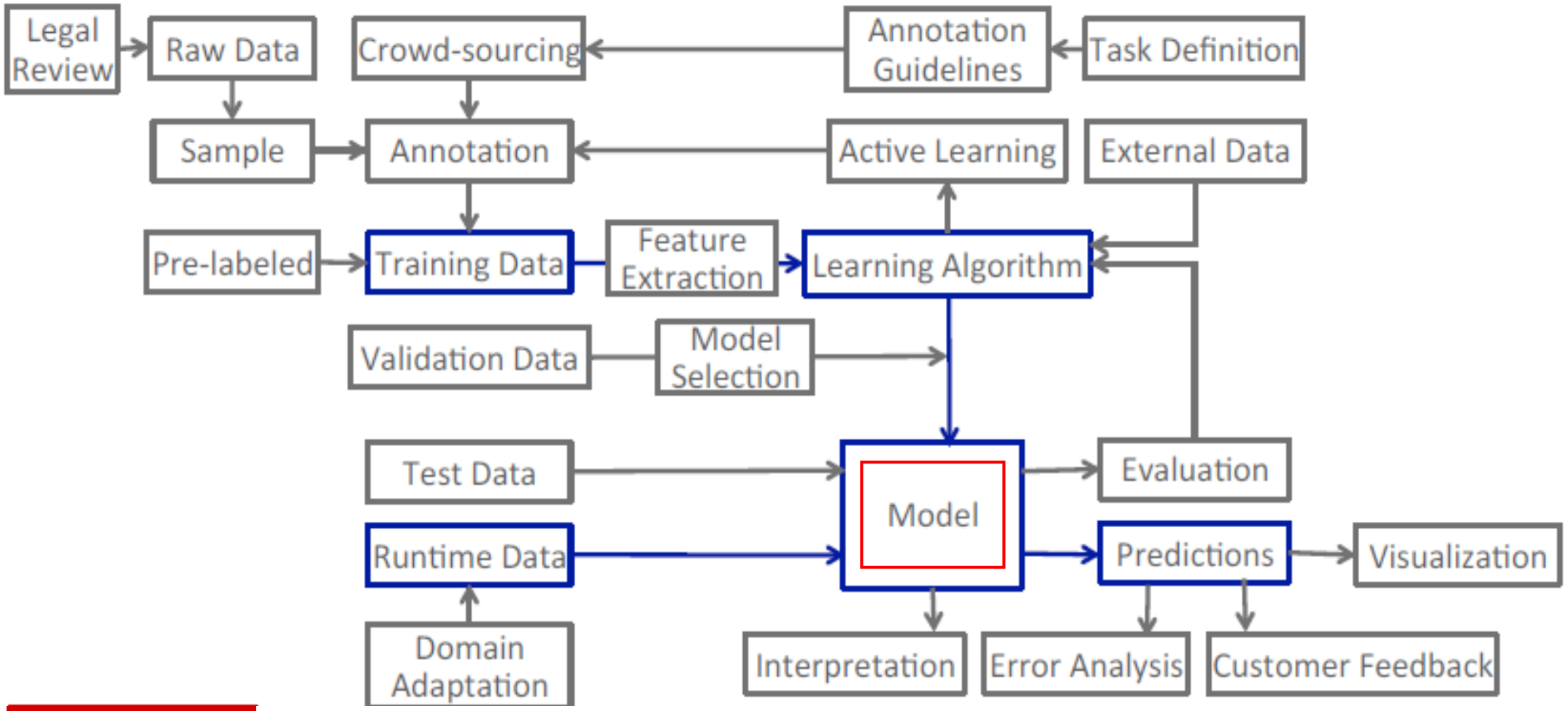
Opportunities

- Silo systems foster redundancies and breaking down these allows for AI to functional optimally
- Analysis of voice data has scientific basis for disease discovery (i.e., Parkinson disease)
- Some human interaction will always remain

Machine Learning: Perception

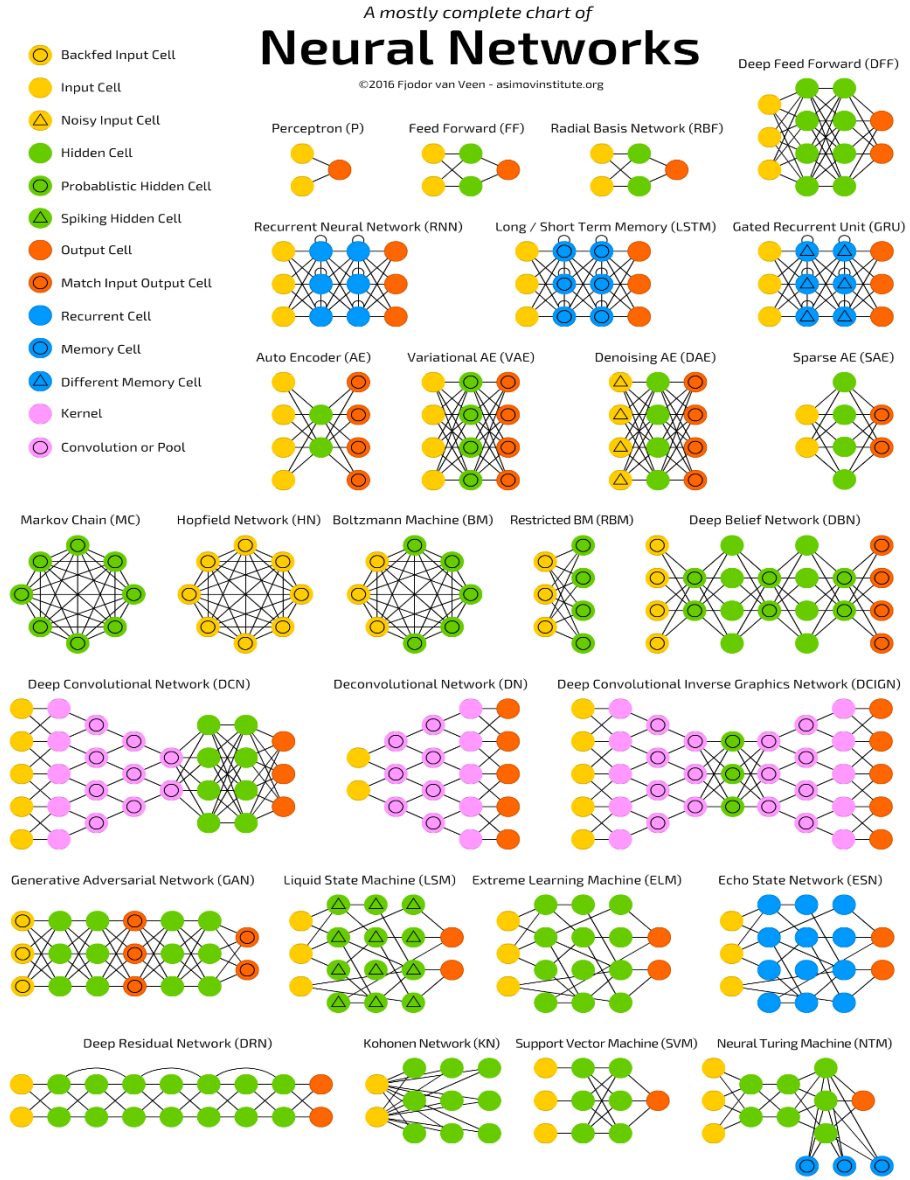


Machine Learning: Actuality



When Not To Use

- Low-budget or low-commitment problems
- Interpreting and communicating model parameters/feature importance to a general audience
- Establishing causal mechanisms
- Learning from “unstructured” features



Bibliography and Links

- nVidia blog: What's the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning, July 2016, accessed August 2017
- https://research.fb.com/publications/imagenet1kin1h/
- http://www.newyorker.com/magazine/2017/04/03/ai-versus-md
- https://www.wsj.com/articles/how-ai-is-transforming-the-workplace-1489371060
- https://www.quora.com/What-are-the-various-applications-of-machine-learning
- http://hyperparameter.space/blog/when-not-to-use-deep-learning

Q&A

Sameer Thapar, PharmD
Director, Global Pharmacovigilance
Oracle Health Sciences Consulting

LinkedIn.com/in/Thapar

Safe Harbor Statement

The preceding is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Hardware and Software
Engineered to Work Together