Data Science As a Service

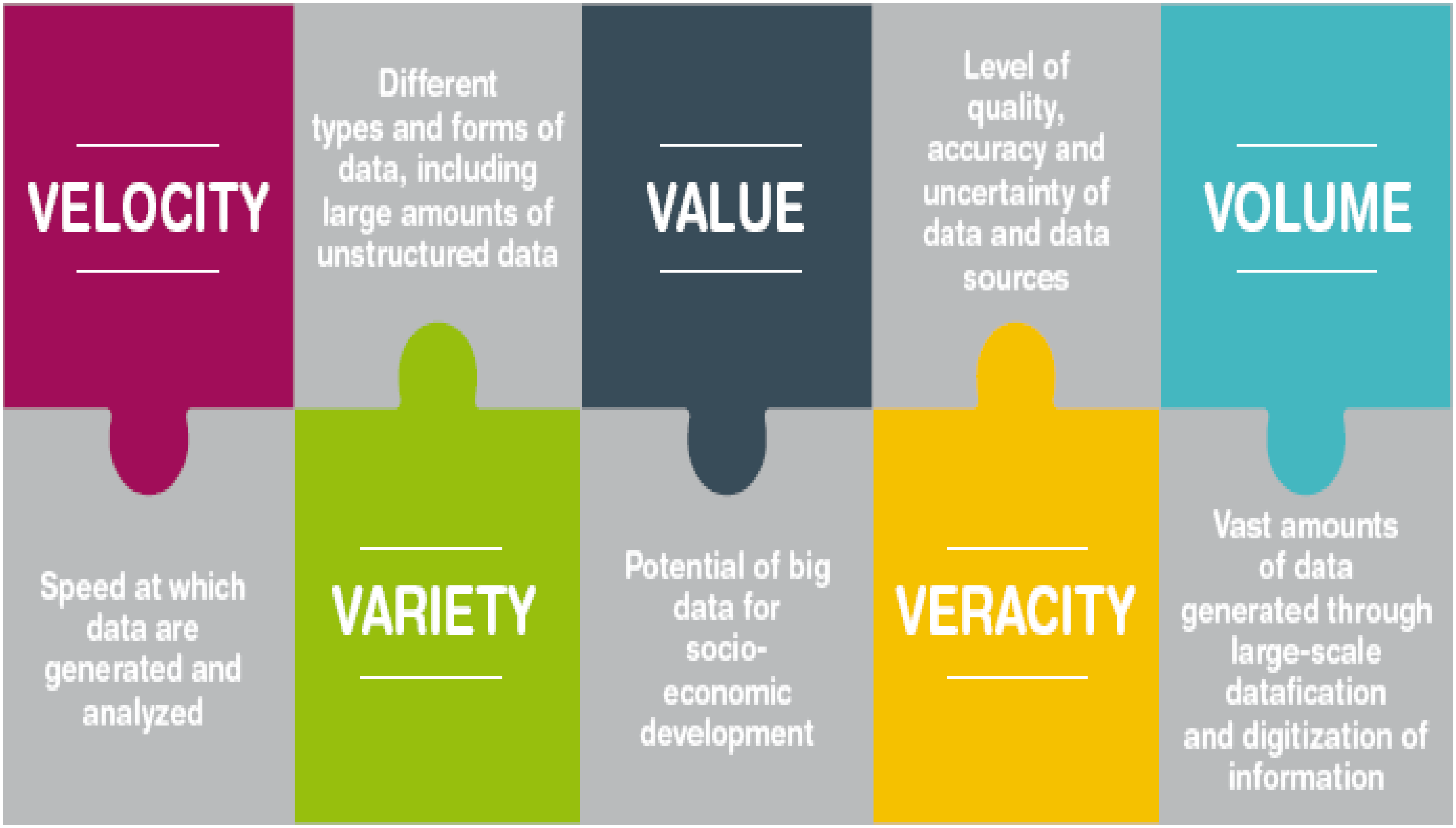
Morteza Saberi, PhD

UTS

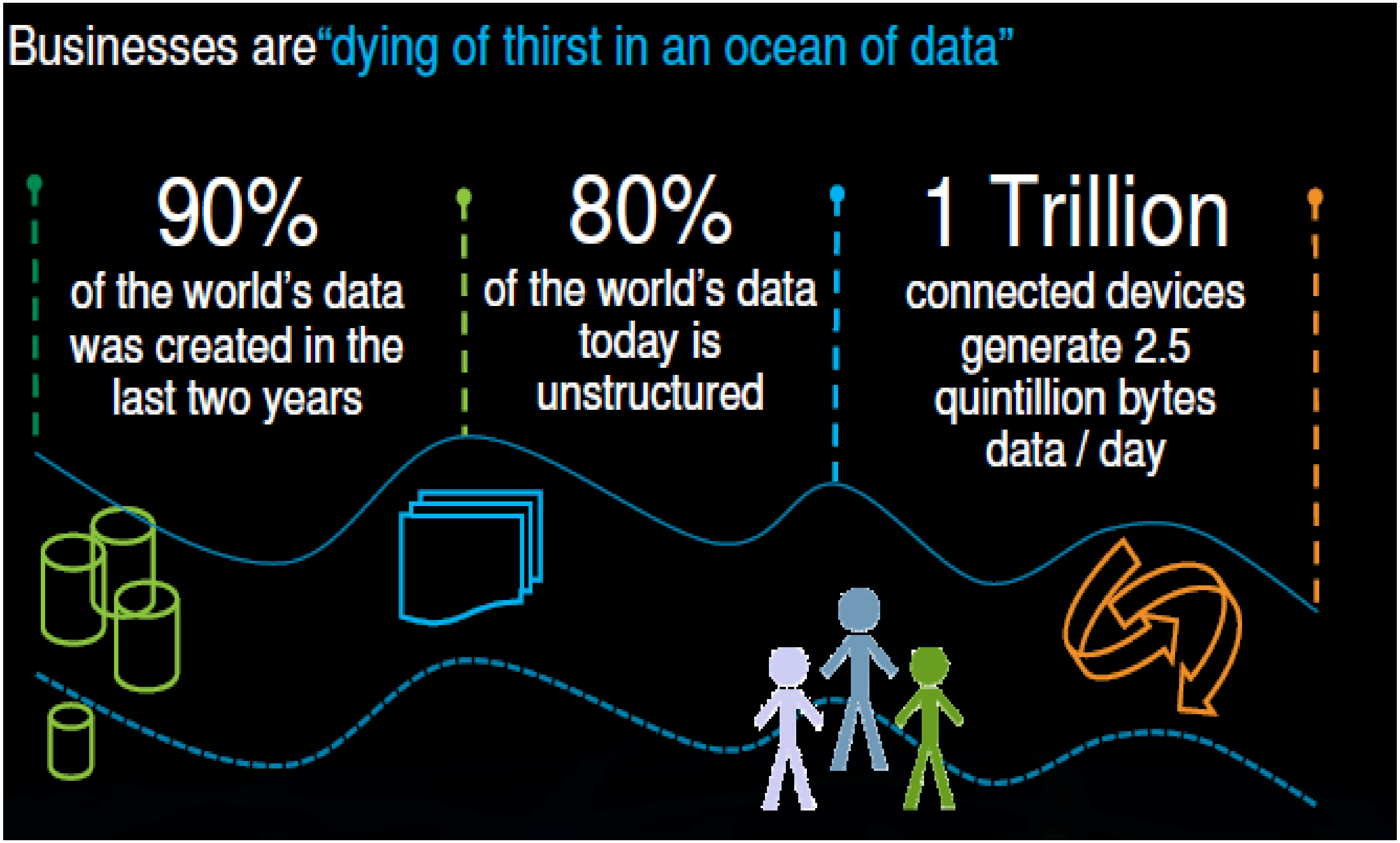
BIG Data

Big data is a term for data sets that **are so large** or complex that **traditional data processing** applications are inadequate to deal with them.

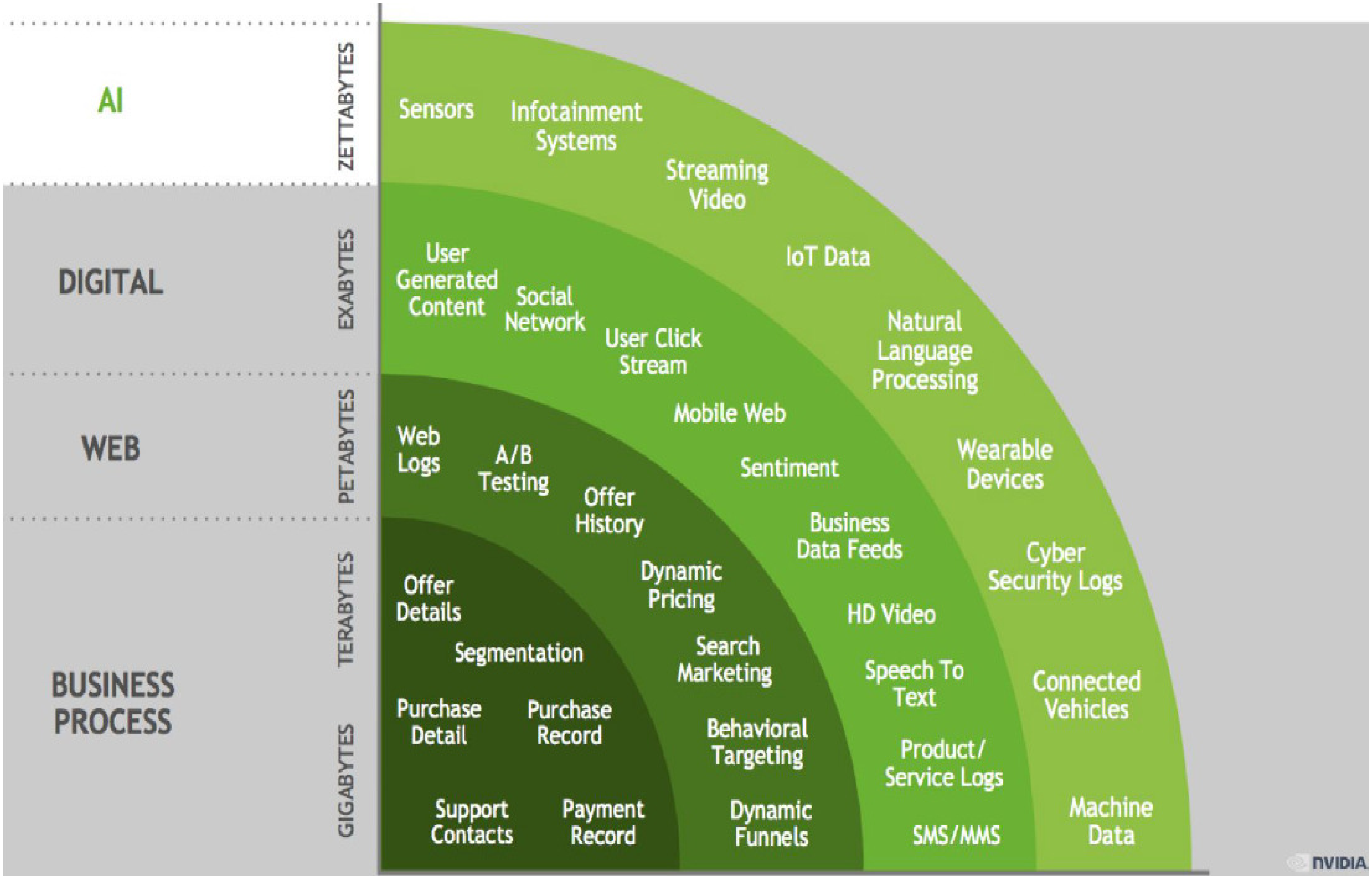
BIG Data


VELOCITY
Different types and forms of data, including large amounts of unstructured data
VALUE
Level of quality, accuracy and uncertainty of data and data sources
VOLUME
Speed at which data are generated and analyzed
VARIETY
Potential of big data for socio-economic development
VERACITY
Vast amounts of data generated through large-scale datafication and digitization of information

BIG Data

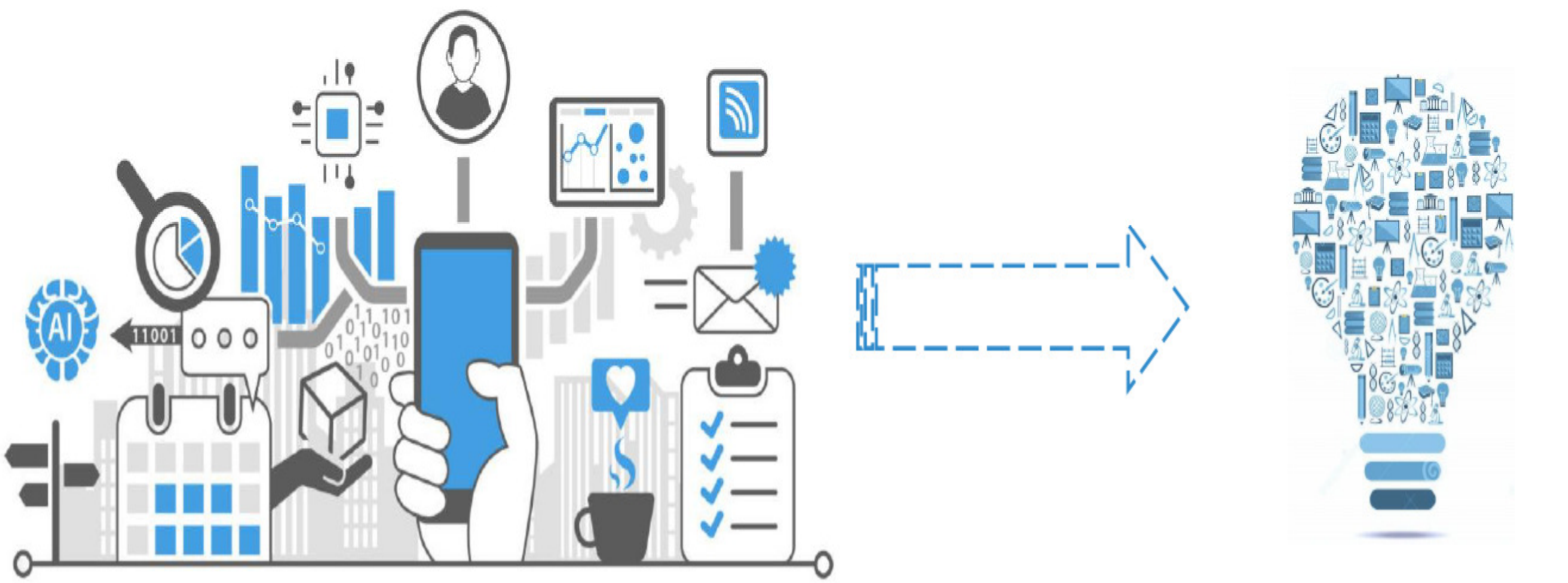

Businesses are"dying of thirst in an ocean of data"
90%
of the world's data was created in the last two years
80%
of the world's data today is unstructured
1 Trillion
connected devices generate 2.5 quintillion bytes data / day

BIG Data



AI
DIGITAL
WEB
BUSINESS PROCESS
ZETTABYTES
EXABYTES
PETABYTES
TERABYTES
GIGABYTES
Sensors
Infotainment Systems
Streaming Video
IoT Data
Natural Language Processing
Wearable Devices
Cyber Security Logs
Connected Vehicles
Machine Data
User Generated Content
Social Network
User Click Stream
Mobile Web
Sentiment
Business Data Feeds
HD Video
Speech To Text
Product/ Service Logs
SMS/MMS
Web Logs
A/B Testing
Offer History
Dynamic Pricing
Search Marketing
Behavioral Targeting
Dynamic Funnels
Offer Details
Segmentation
Purchase Detail
Purchase Record
Support Contacts
Payment Record
nVIDIA

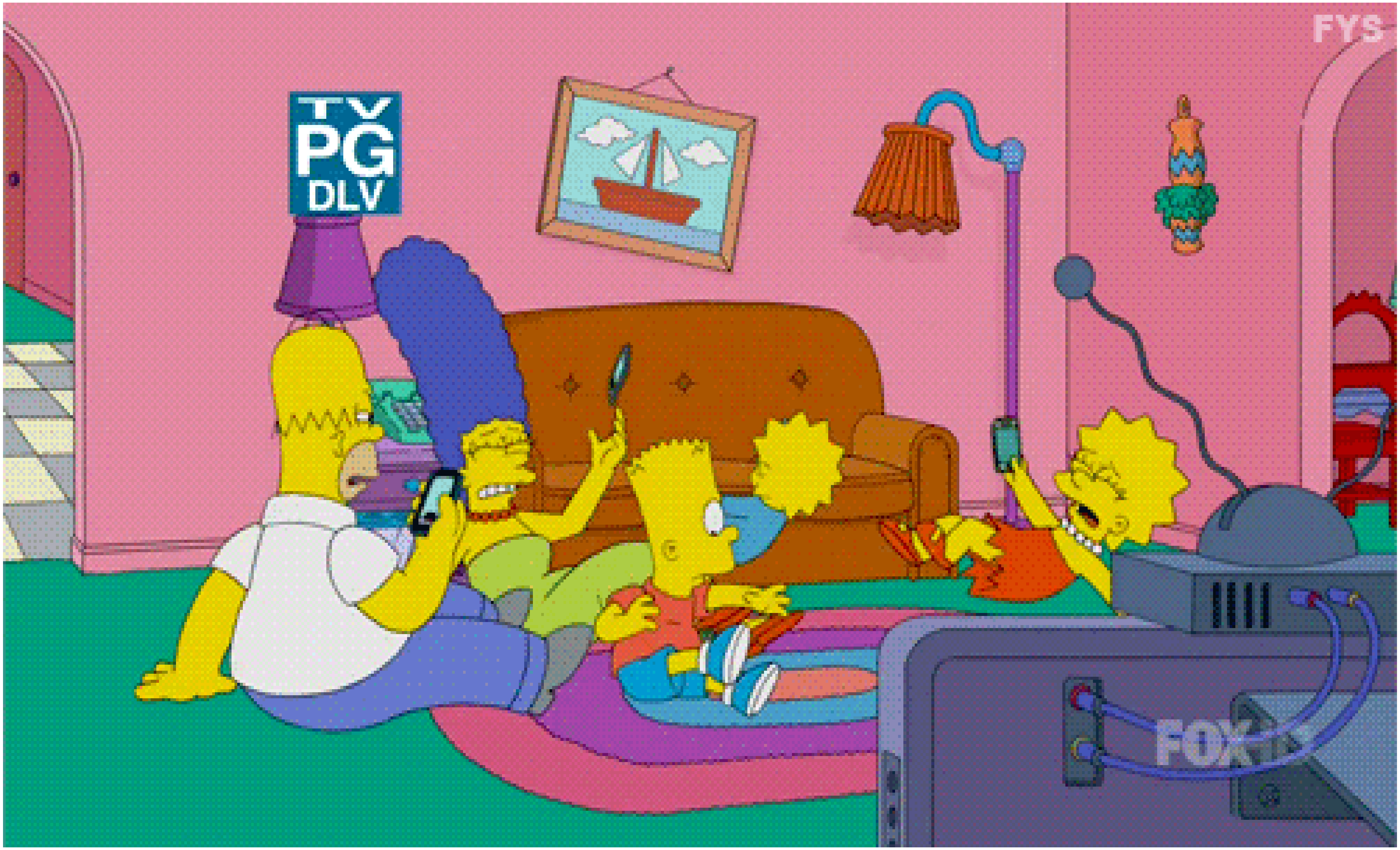
BIG Data



Big Data

Structured Knowledge & Insights

Big Data!!


TV
PG
DLV
FYS
FOX

Present of Big Data

Too big to handle

Why do we need data science?

- "The data is the computer"
  - Large amounts of data can be more powerful than complex algorithms and models
    - Google has solved many Natural Language Processing problems, simply by looking at the data
    - Example: misspellings, synonyms

Why do we need data science?

- Data is power!
  - Today, the collected data is one of the biggest assets of an online company
    - Query logs of Google
    - The friendship and updates of Facebook
    - Tweets and follows of Twitter
    - Amazon transactions
- We need a way to harness the collective intelligence

Evolution of Database Technology

- 1960s:
  - Data collection, database creation, IMS and network DBMS
- 1970s:
  - Relational data model, relational DBMS implementation
- 1980s:
  - RDBMS, advanced data models (extended-relational, OO, deductive, etc.)
  - Application-oriented DBMS (spatial, scientific, engineering, etc.)
- 1990s:
  - Data mining, data warehousing, multimedia databases, and Web databases
- 2000s
  - Stream data management and mining
  - Data mining and its applications
  - Web technology (XML, data integration) and global information systems

The data is also very complex

- Multiple types of data: tables, text, time series, images, graphs, etc

- Spatial and temporal aspects

- Interconnected data of different types:
  - From the mobile phone we can collect, location of the user, friendship information, check-ins to venues, opinions through twitter, status updates in FB, images though cameras, queries to search engines

Example: transaction data

- Billions of real-life customers:

  - WALMART: 20M transactions per day
  - AT&T 300 M calls per day
  - Credit card companies: billions of transactions per day.

- The point cards allow companies to collect information about specific users

Example: document data

- Web as a document repository: estimated 50 billions of web pages
- Wikipedia: 4.5 million articles (and counting)
- Online news portals: steady stream of 100's of new articles every day
- Twitter: ~500 million tweets every day

Example: network data

- Web: 50 billion pages linked via hyperlinks
- Facebook: 1.23 billion users
- Twitter: 270 million users
- Blogs: 250 million blogs worldwide, presidential candidates run blogs

Behavioral data

- Mobile phones today record a large amount of information about the user behavior
  - GPS records position
  - Camera produces images
  - Communication via phone and SMS
  - Text via facebook updates
  - Association with entities via check-ins

Behavioral data

- Amazon collects all the items that you browsed, placed into your basket, read reviews about, purchased.

- Google and Bing record all your browsing activity via toolbar plugins. They also record the queries you asked, the pages you saw and the clicks you did.

- Data collected for millions of users on a daily basis

So, what is Data?

- Collection of data objects and their attributes
- An attribute is a property or characteristic of an object
  - Examples: eye color of a person, temperature, etc.
  - Attribute is also known as variable, field, characteristic, or feature
- A collection of attributes describe an object
  - Object is also known as record, point, case, sample, entity, or instance

Attributes

Objects

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	**No**
2	No	Married	100K	**No**
3	No	Single	70K	**No**
4	Yes	Married	120K	**No**
5	No	Divorced	95K	**Yes**
6	No	Married	60K	**No**
7	Yes	Divorced	220K	**No**
8	No	Single	85K	**Yes**
9	No	Married	75K	**No**
10	No	Single	90K	**Yes**

Size: Number of objects
Dimensionality: Number of attributes

Data mining?

Data mining is the use of efficient techniques for the analysis of very large collections of data and the extraction of useful and possibly unexpected patterns in data.

Examples: What is (not) Data Mining?

- **What is not Data Mining?**
  - Look up phone number in phone directory
  - Query a Web search engine for information about "Amazon"

- **What is Data Mining?**
  - Certain names are more prevalent in certain US locations (O'Brien, O'Rurke, O'Reilly… in Boston area)
  - Group together similar documents returned by search engine according to their context (e.g. Amazon rainforest, Amazon.com,)

knowledge?

- **Valid**: generalize to the future
- **Novel**: what we don't know
- **Useful**: be able to take some action
- **Understandable**: leading to insight
- **Iterative**: takes multiple passes
- **Interactive**: human in the loop

Types of Attributes

- There are different types of attributes
  - Nominal: eye color, sex
  - Ordinal: taste of potato chips on a scale from 1-10), grades, height in {tall, medium, short}
  - Numeric
    - Examples: dates, temperature, time, length, value, count.
    - Discrete (counts) vs Continuous (temperature)
    - Special case: Binary attributes (yes/no, exists/not exists)

Numeric Record Data

- Such data set can be represented by an n-by-d data matrix, where there are n rows, one for each object, and d columns, one for each attribute

Projection of x Load	Projection of y load	Distance	Load	Thickness
10.23	5.27	15.22	2.7	1.2
12.65	6.25	16.22	2.2	1.1

Categorical Data

- Data that consists of a collection of records, each of which consists of a fixed set of nominal attributes

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	High	**No**
2	No	Married	Medium	**No**
3	No	Single	Low	**No**
4	Yes	Married	High	**No**
5	No	Divorced	Medium	**Yes**
6	No	Married	Low	**No**
7	Yes	Divorced	High	**No**
8	No	Single	Medium	**Yes**
9	No	Married	Medium	**No**
10	No	Single	Medium	**Yes**

Document Data

- Each document becomes a `term' vector,
  - each term is a component (attribute) of the vector,
  - the value of each component is the number of times the corresponding term occurs in the document.
  - Bag-of-words representation – no ordering

	team	coach	pla y	ball	score	game	wi n	lost	timeout	season
Document 1	3	0	5	0	2	6	0	2	0	2
Document 2	0	7	0	2	1	0	0	3	0	0
Document 3	0	1	0	0	1	2	2	0	3	0

Transaction Data

- Each record (transaction) is a set of items.

TID	*Items*
1	Bread, Coke, Milk
2	Beer, Bread
3	Beer, Coke, Diaper, Milk
4	Beer, Bread, Diaper, Milk
5	Coke, Diaper, Milk

- A set of items can also be represented as a binary vector, where each attribute is an item.
- A document can also be represented as a set of words (no counts)

Sparsity: average number of products bought by a customer

Ordered Data

- Genomic sequence data

```
GGTTCCGCCTTCAGCCCCGCGCC
CGCAGGGCCCGCCCCGCGCCGTC
GAGAAGGGCCCGCCTGGCGGGCG
GGGGGAGGCGGGGCCGCCCGAGC
CCAACCGAGTCCGACCAGGTGCC
CCCTCTGCTCGGCCTAGACCTGA
GCTCATTAGGCGGCAGCGGACAG
GCCAAGTAGAACACGCGAAGCGC
TGGGCTGCCTGCTGCGACCAGGG
```

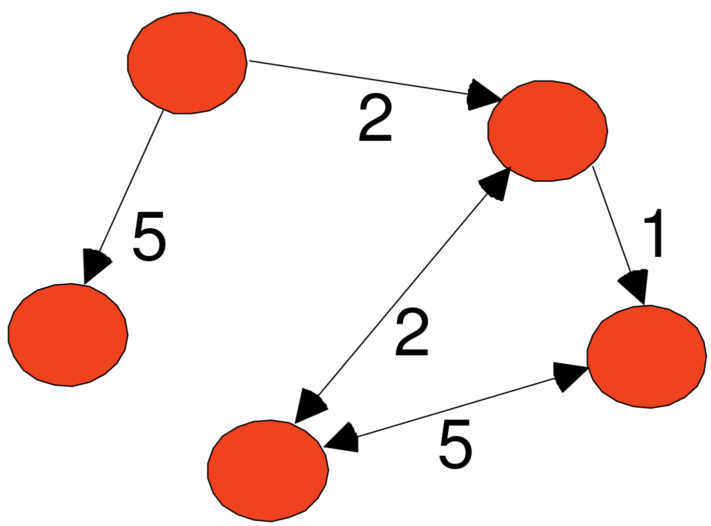
- Data is a long ordered string

Ordered Data

- Time series
  - Sequence of ordered (over "time") numeric values.

Graph Data

- Examples: Web graph and HTML Links
- Facebook graph of Friendships
- Twitter follow graph
- The connections between brain neurons



In this case the data consists of pairs:

Who links to whom

What can you do with the data?

- Suppose that you are the owner of a supermarket and you have collected billions of market basket data. What information would you extract from it and how would you use it?

TID	*Items*
1	Bread, Coke, Milk
2	Beer, Bread
3	Beer, Coke, Diaper, Milk
4	Beer, Bread, Diaper, Milk
5	Coke, Diaper, Milk

Product placement

Catalog creation

Recommendations

- What if this was an online store?

What can you do with the data?

- Suppose you are a search engine and you have a toolbar log consisting of
  - pages browsed,
  - queries,
  - pages clicked,
  - ads clicked

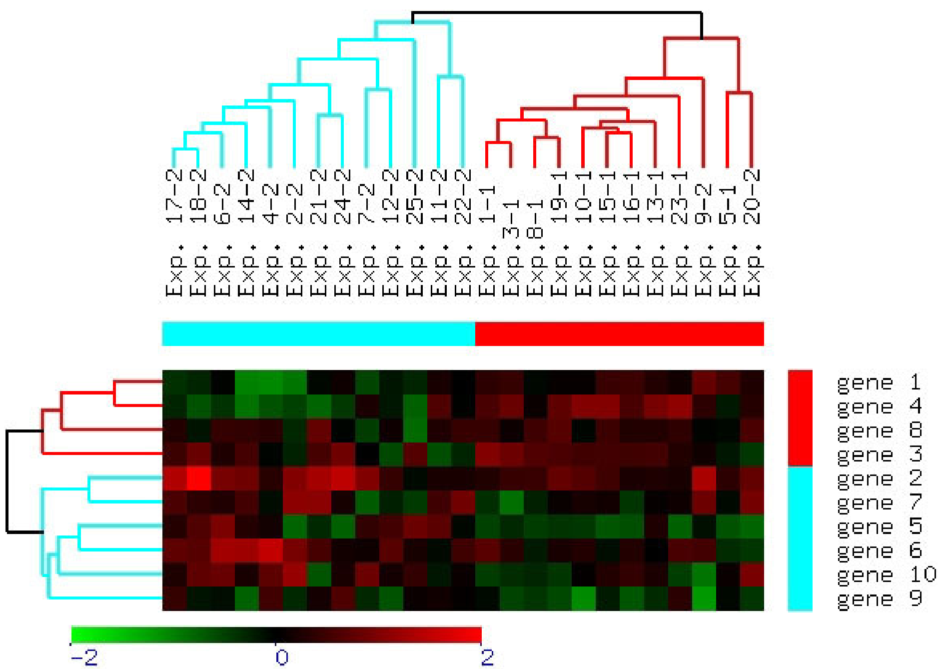
each with a user id and a timestamp. What information would you like to get our of the data?

Ad click prediction

Query reformulations

What can you do with the data?

- Suppose you are biologist who has microarray expression data: thousands of genes, and their expression values over thousands of different settings (e.g. tissues). What information would you like to get out of your data?



Groups of genes and tissues

What can you do with the data?

- Suppose you are a stock broker and you observe the fluctuations of multiple stocks over time. What information would you like to get our of your data?

Clustering of stocks

Correlation of stocks

Stock Value prediction

What can you do with the data?

- You are the owner of a social network, and you have full access to the social graph, what kind of information do you want to get out of your graph?

  - Who is the most important node in the graph?
  - What is the shortest path between two nodes?
  - How many friends two nodes have in common?
  - How does information spread on the network?

What is data mining again?

- The industry point of view: The analysis of huge amounts of data for extracting useful and actionable information, which is then integrated into production systems in the form of new features of products

  - Data Scientists should be good at data analysis, math, statistics, but also be able to code with huge amounts of data and use the extracted information to build products.

Tasks

- Classification
- Clustering
- Estimation
- Affinity groups

Two Main Types of Machine Learning

- Supervised learning: learn by examples
- Unsupervised learning: find structure w/o examples

	Supervised Learning	Unsupervised Learning
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

Classification: Definition

- Given a collection of records (*training set*)
  - Each record contains a set of *attributes*, one of the attributes is the *class*.
- Find a *model* for class attribute as a function of the values of other attributes.
- Goal: <u>previously unseen</u> records should be assigned a class as accurately as possible.
  - A *test set* is used to determine the accuracy of the model. Usually, the given data set is divided into training and test sets, with training set used to build the model and test set used to validate it.

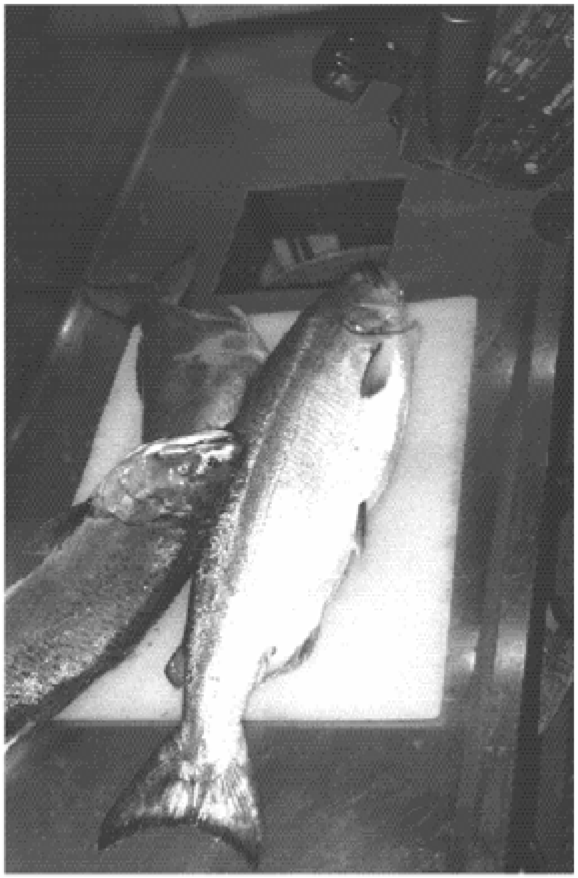
An Example

(from *Pattern Classification by* Duda & Hart & Stork – Second Edition, 2001)

- A fish-packing plant wants to automate the process of sorting incoming fish according to species

- As a pilot project, it is decided to try to separate **sea bass** from salmon using optical sensing

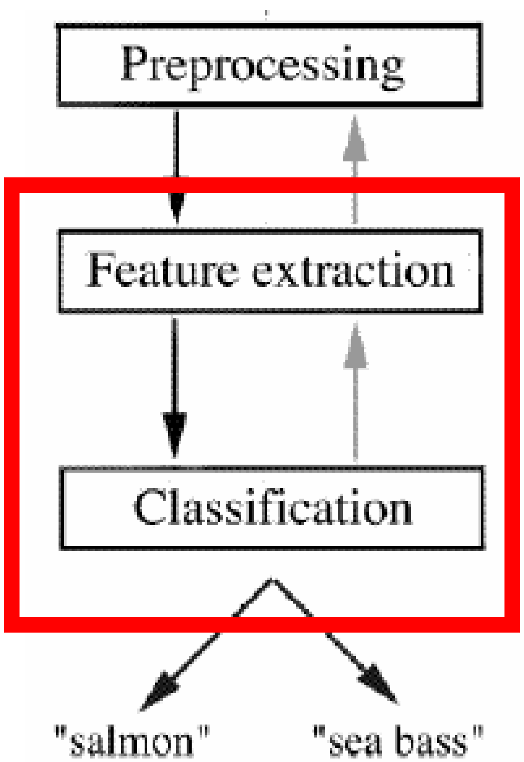
An Example (continued) Classification



Features (to distinguish):

- Length
- Lightness
- Width
- Position of mouth

An Example (continued) Classification



- **Preprocessing:** Images of different fishes are isolated from one another and from background;
- **Feature extraction:** The information of a single fish is then sent to a feature extractor, that measure certain "features" or "properties";
- **Classification:** The values of these features are passed to a classifier that evaluates the evidence presented, and build a model to discriminate between the two species

An Example (continued)

- Domain knowledge:
  - A sea bass is generally longer than a salmon
- Related feature*:* (or attribute)
  - Length

- Training the classifier:

  - Some examples are provided to the classifier in this form: <fish_length, fish_name>

An Example (continued)

- These examples are called training examples
- The classifier *learns* itself from the training examples,
- how to distinguish Salmon from Bass based on the *fish_length*

An Example (continued)

- Classification model (hypothesis):
  - The classifier generates a model from the training data to classify future examples (test examples)
  - An example of the model is a rule like this:
  - If *Length >= l* then sea bass* otherwise *salmon*
  - Here the value of *l** determined by the classifier

An Example (continued)

- Testing the model
  - Once we get a model out of the classifier, we may use the classifier to test future examples
  - The test data is provided in the form <fish_length>
  - The classifier outputs <fish_type> by checking *fish_length* against the model

Classification: Application 1

- Direct Marketing
  - Goal: Reduce cost of mailing by *targeting* a set of consumers likely to buy a new cell-phone product.
  - Approach:
    - Use the data for a similar product introduced before.
    - We know which customers decided to buy and which decided otherwise. This *{buy, don't buy}* decision forms the *class attribute*.
    - Collect various demographic, lifestyle, and company-interaction related information about all such customers.
      - Type of business, where they stay, how much they earn, etc.
    - Use this information as input attributes to learn a classifier model.

Classification: Application 2

- Fraud Detection
  - Goal: **Predict fraudulent cases in credit card transactions**.
  - Approach:
    - Use credit card transactions and the information on its account-holder as attributes.
      - When does a customer buy, what does he buy, how often he pays on time, etc

Classification: Application 2

- Fraud Detection
  - Goal: **Predict fraudulent cases in credit card transactions**.
  - Approach:
    - Label past transactions as fraud or fair transactions. This forms the class attribute.
    - Learn a model for the class of the transactions.
    - Use this model to detect fraud by observing credit card transactions on an account.

Classification: Application 3

- Customer Attrition/Churn:
  - Goal: **To predict whether a customer is likely to be lost to a competitor**.
  - Approach:
    - Use detailed record of transactions with each of the past and present customers, to find attributes.
      - *How often the customer calls, where he calls, what time-of-the day he calls most, his financial status, marital status, etc.*
    - Label the customers as loyal or disloyal.
    - Find a model for loyalty.

Data Mining Function: (4) Cluster Analysis

- **Unsupervised learning** (i.e., Class label is unknown)
- Group data to form **new** categories (i.e., clusters), e.g., cluster houses to find distribution patterns
- Principle: **Maximizing intra-class similarity & minimizing interclass similarity**
- Many methods and applications

Clustering Definition

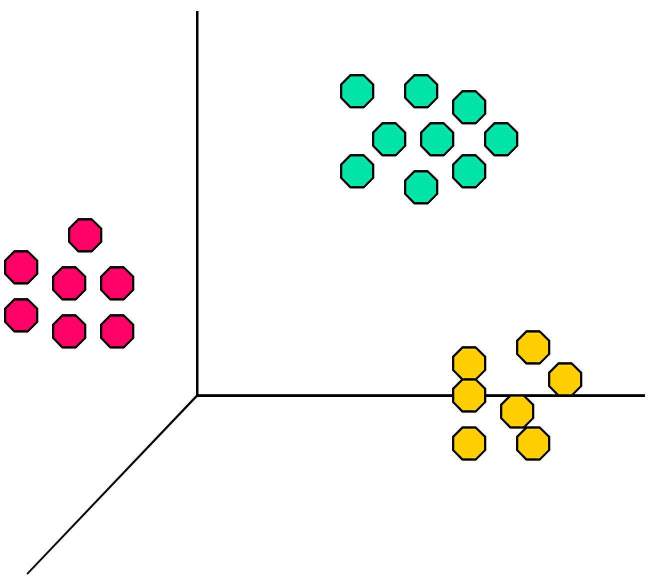
- Given a set of data points, each having a set of attributes, and a similarity measure among them, find clusters such that
  - Data points in one cluster are more similar to one another.
  - Data points in separate clusters are less similar to one another.
- Similarity Measures:
  - Euclidean Distance if attributes are continuous.
  - Other Problem-specific Measures.

Illustrating Clustering

- Euclidean Distance Based Clustering in 3-D space.

Intracluster distances are minimized

Intercluster distances are maximized



Clustering: Application 1

- **Market Segmentation**:
  - Goal: *subdivide a market into distinct subsets of customers where any subset may conceivably be selected as a market target to be reached with a distinct marketing mix.*
  - Approach:
    - Collect different attributes of customers based on their geographical and lifestyle related information.
    - Find clusters of similar customers.
    - Measure the clustering quality by observing buying patterns of customers in same cluster vs. those from different clusters.

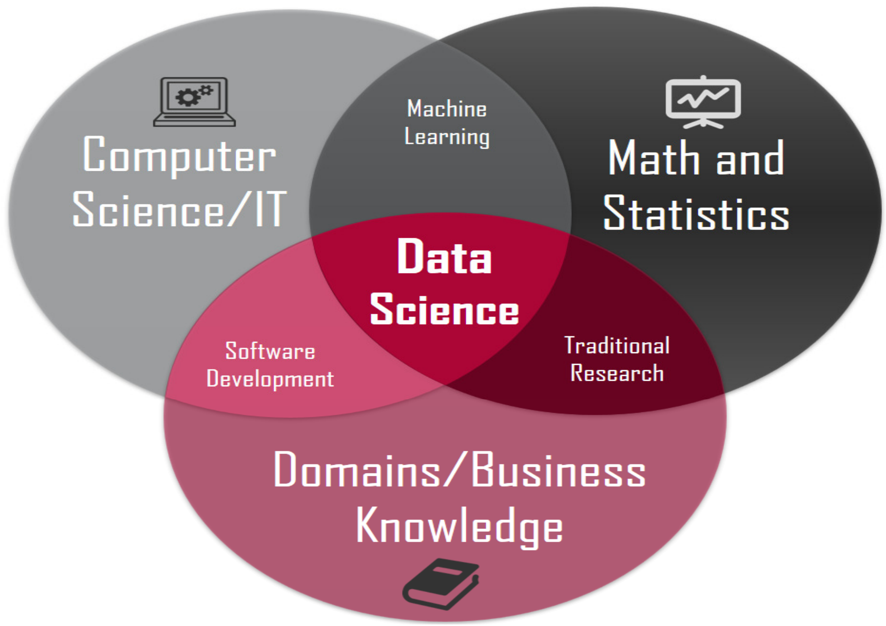
Clustering: Application 2

- Document Clustering:
  - Goal: To find groups of documents that are similar to each other based on the important terms appearing in them.
  - Approach: To identify frequently occurring terms in each document. Form a similarity measure based on the frequencies of different terms. Use it to cluster.
  - Gain: Information Retrieval can utilize the clusters to relate a new document or search term to clustered documents.

Data Mining Function: (5) Outlier Analysis

- Outlier analysis
  - **Outlier**: A data object that **does not comply** with the general behavior of the data
  - **Noise or exception**? — One person's garbage could be another person's treasure
  - **Methods**: by product of clustering or regression analysis, …
  - Useful in fraud detection, rare events analysis

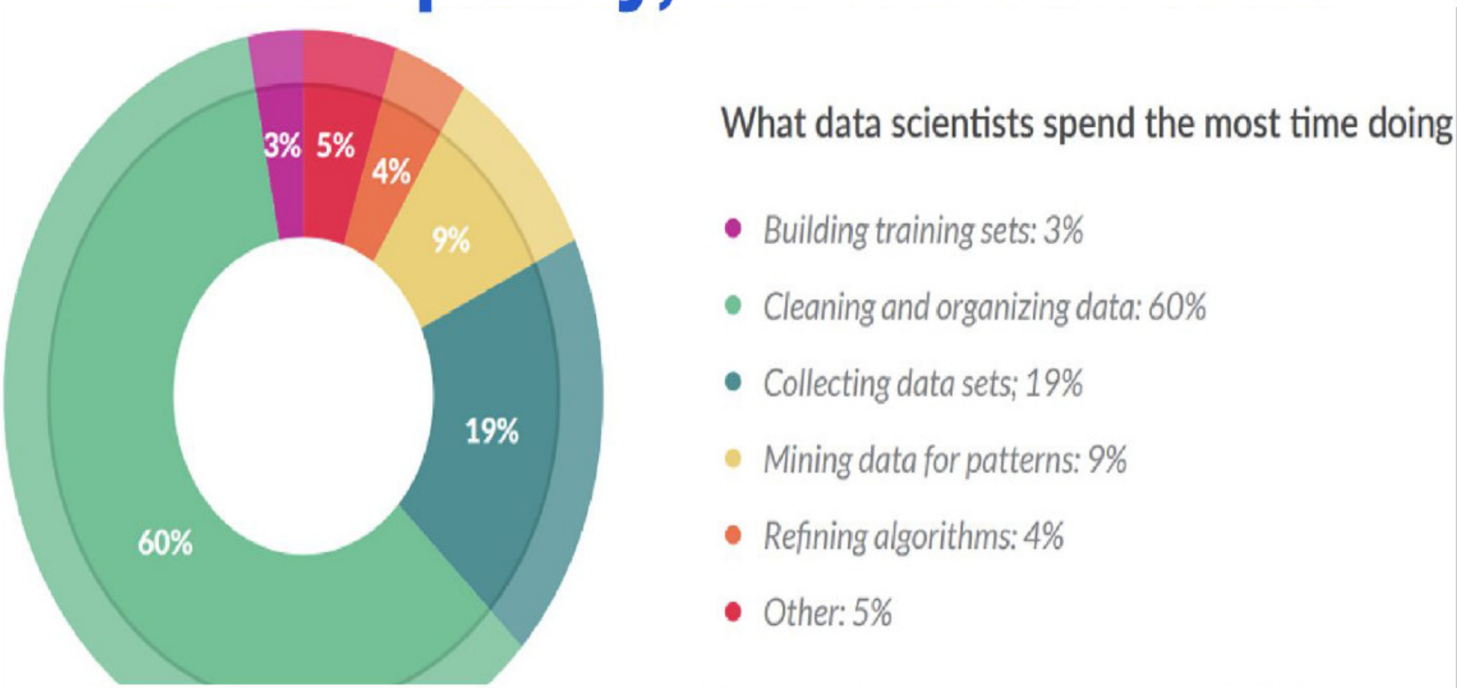
Data Science



Data Scientists should be good at data analysis, math, statistics, but also be able to code with huge amounts of data and use the extracted information to build products.

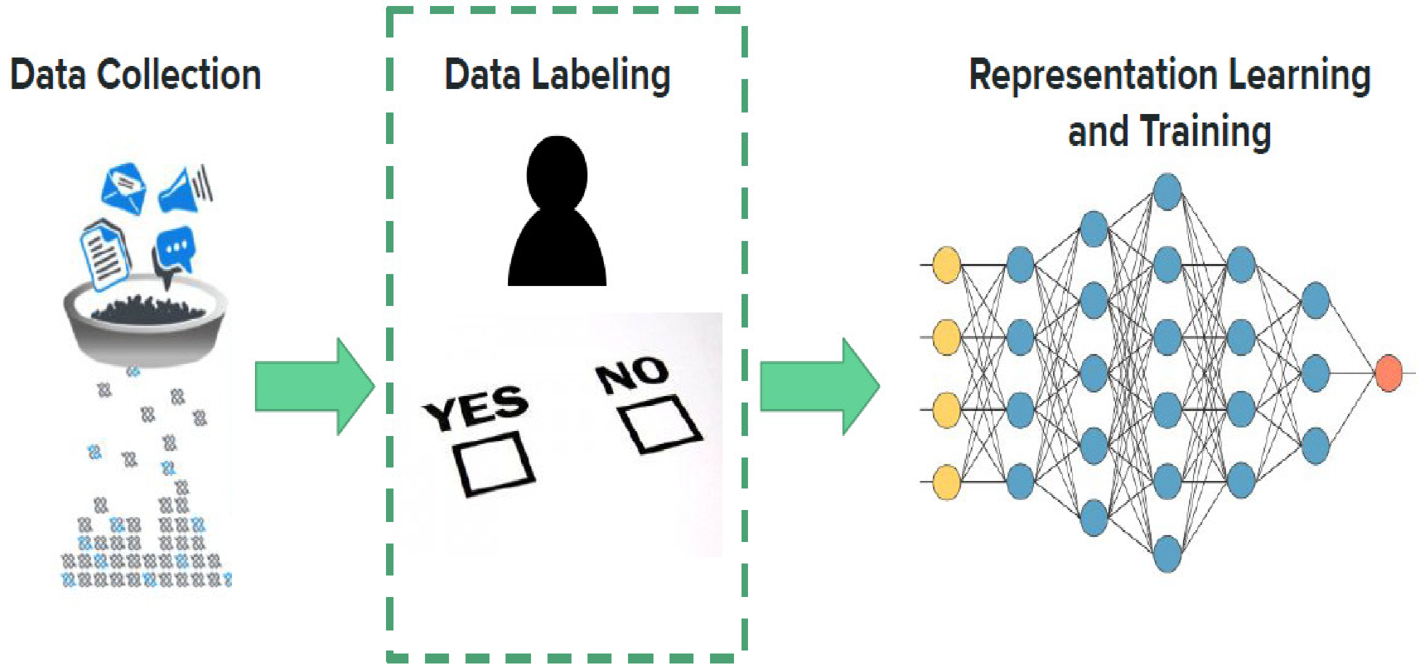
The Achilles' Heel of Modern Analytics

is low quality, erroneous data



Cleaning and organizing the data comprises 60% of the time spent on an analytics or AI project.

The ML Pipeline in the Deep Learning Era



A core pain point today, lots of time spent in labeling data.

Training Data: Challenges and Opportunities

- Collecting training data is **expensive** and **slow.**
- We are overfitting to our training data. [Recht et al., 2018]
  - Hand-labeled training data does not change
- Training data is the point to inject domain knowledge
  - Modern ML is too complex to hand-tune features and priors

The Rise of Weak Supervision

Definition: Supervision with noisy (much easier to collect) labels; prediction on a larger set, and then training of a model.

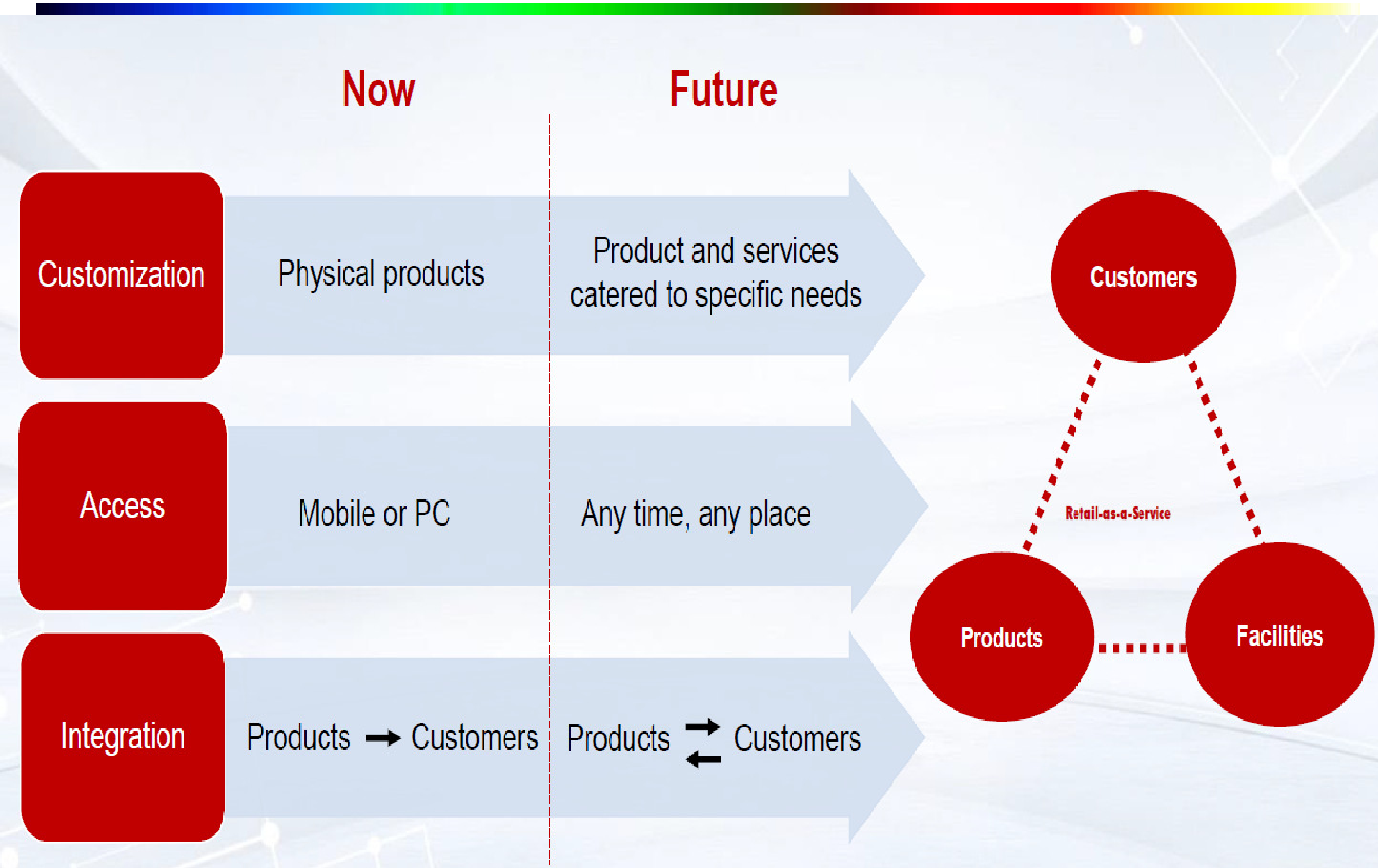
Semi-supervised learning and ensemble learning

Examples:

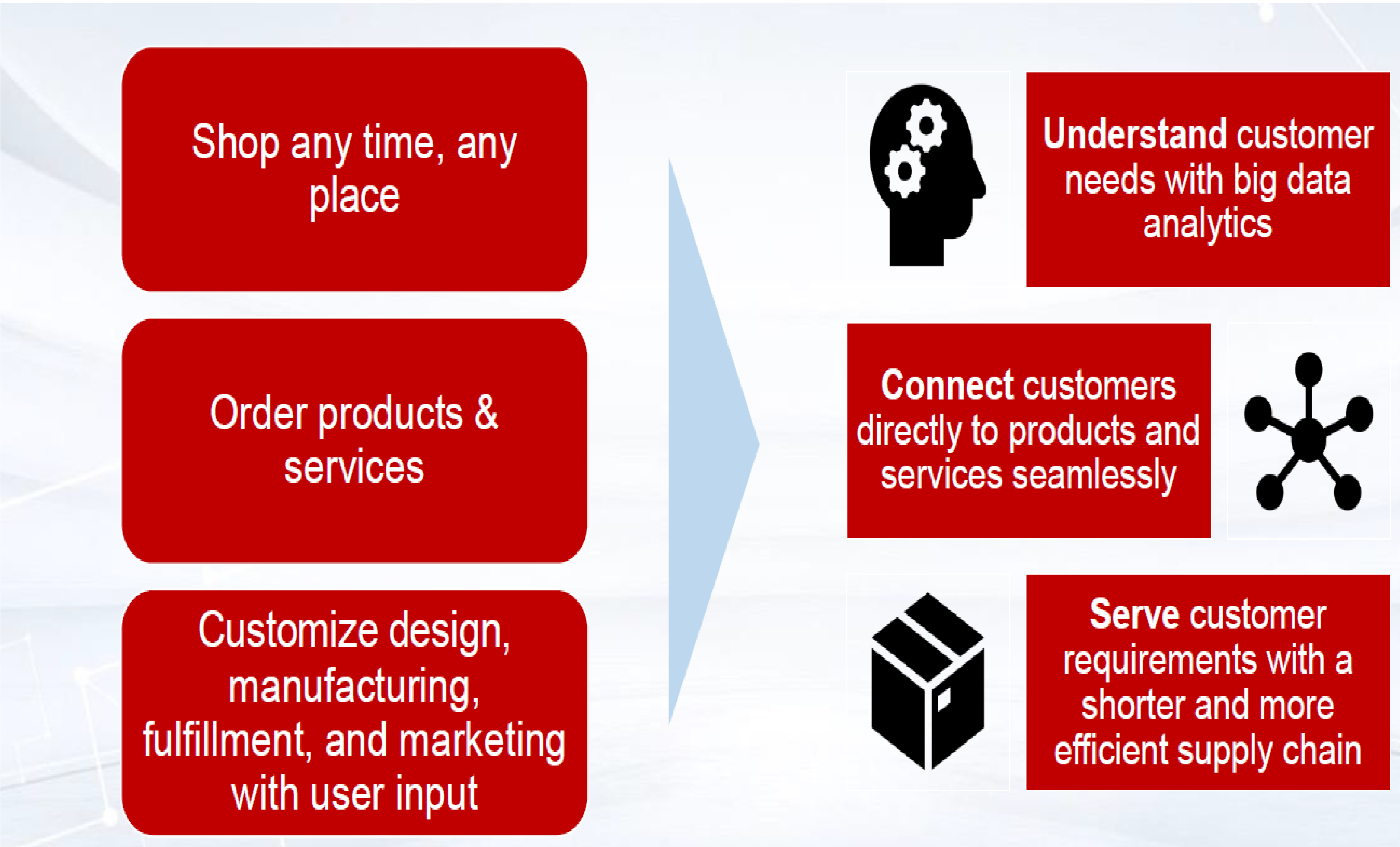
- use of non-expert labelers (crowdsourcing),
- use of curated catalogs (distant supervision)
- use of heuristic rules (labeling functions)

Data Science in Retailing

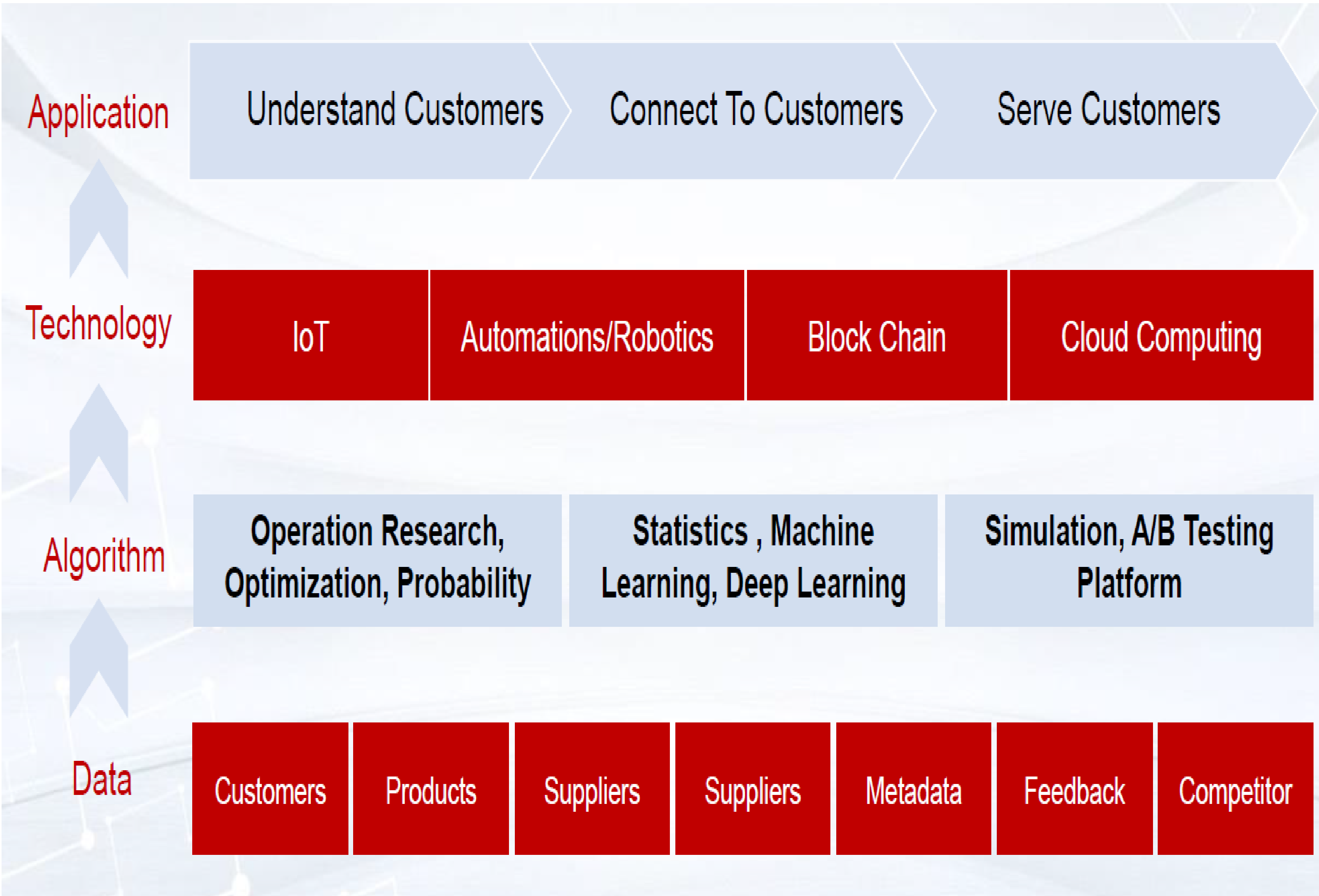
Example: Retail is Changing



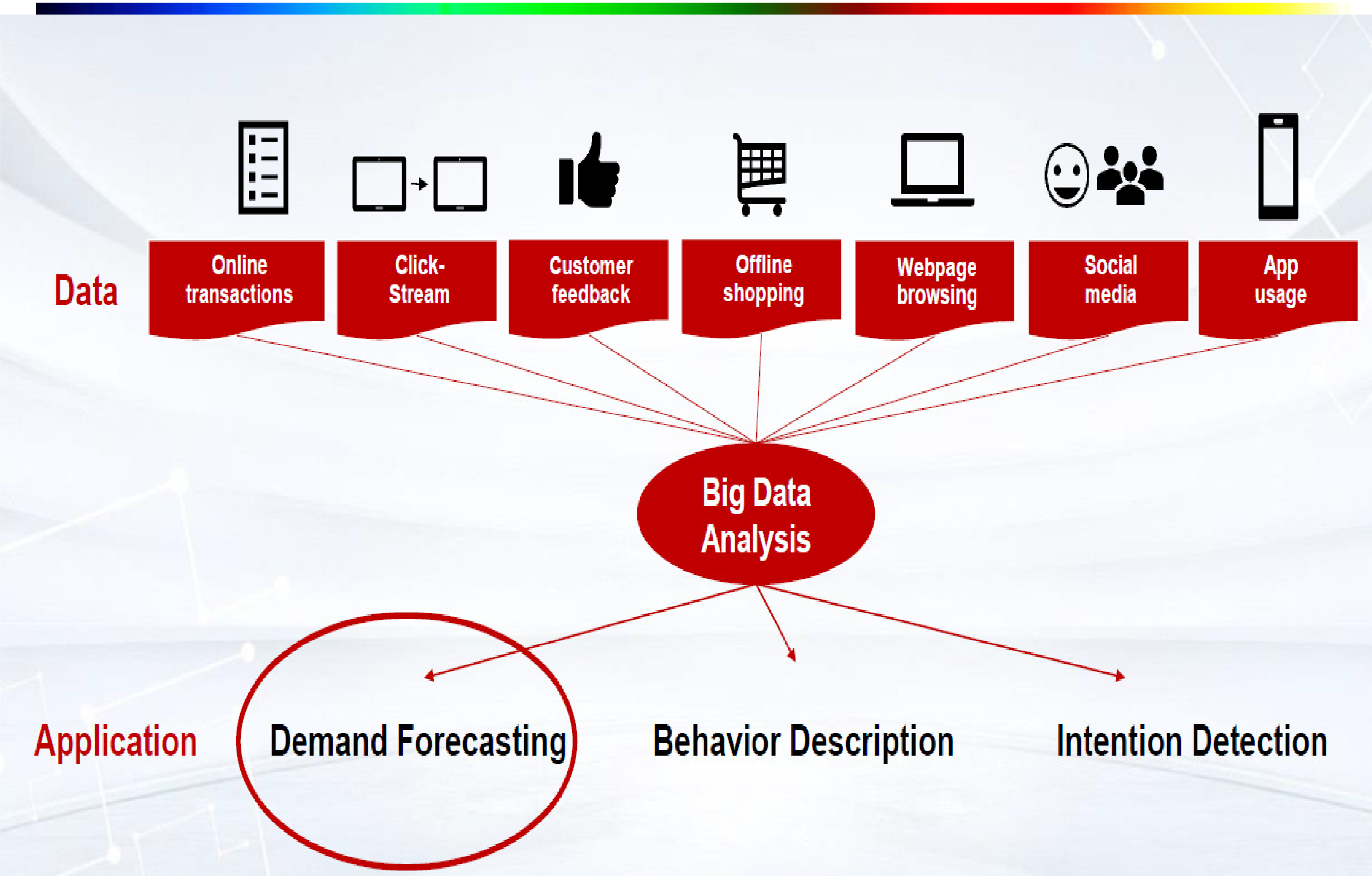
Example: Retail is Changing



Data-Driven Boundaryless Retail



Understand Customers with Big Data



Time Series Forecasting

Customer demand on a product forms a time series

Model-Driven	Data-Driven	Big Data Enrichment
Stochastic Time Series Models	**Machine Learning**	**Deep Learning**

Stochastic Time Series Models

- Linear Models:
  - ARIMA: Box-Jenkins methodology (1970)
  - AR,MA,ARMA,SARMA
  - VAR
- Non-linear Models:
  - ARCH (1982)
  - GARCH (1986)

Machine Learning

- Linear Regression
- Support Vector Regression (1996)
- Gaussian Process
- Tree-based Models
  - Regression Tress (1984)
  - Random Forest (1995)
  - Gradient Boosting
    - AdaBoost (1997)
    - XGBoost (2014)
    - LightGBM (2017)
    - CatBoost (2017)

Deep Learning

- MLP (<1965)
- RNN (1980s)
- LSTM (1997)
- Seq2seq (2014)

Time Series Forecasting

- Retailing is about getting the right products to the right people in the right place at the right time.
- Customers requirement vary by

Location
(e.g. stationery sales near a school)

Time
(e.g. ice-cream sales on sunny days)

Special Event
(e.g. toy sales after movie is released)

Personal Preference
(e.g. different fashion styles)

Demand Forecasting in E-Commerce

Highly variable customers needs

Stock inventory to provide buffer against demand variability

Millions of products (not to mention product-region pairs)

Supply chain issue like vendor lead time

Demand Forecasting in E-Commerce

Highly variable customers needs	Highly non-stationary demand time series	Stock inventory	Probabilistic forecast
Millions of products	Multiple time series	Vendor lead time	Multi-horizon forecast

ARIMA

- Auto-Regressive Integrated Moving Average
- George Box and Gwilym Jenkins developed in 1970s
- ARIMA(p,d,q)

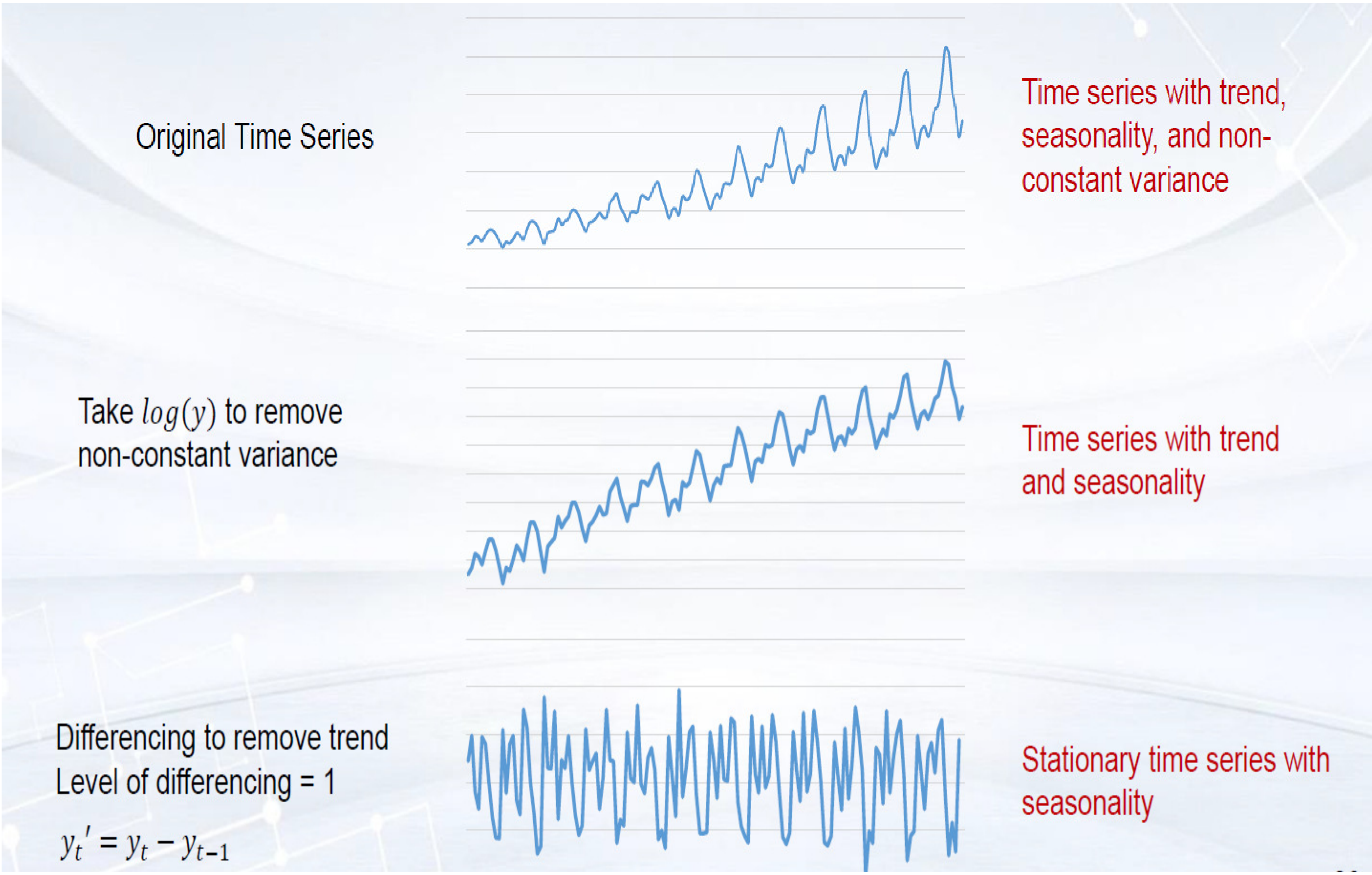
$$y_t = \delta + \underbrace{\phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p}}_{\text{AR(p) terms regress against past values}} + \underbrace{\theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q}}_{\text{MA(q) terms regress against past errors}}$$

- ARMA models can only be used for stationary time series
- Use finite differencing to 'stationarize' time series

$$y_t{}' = y_t - y_{t-d} \leftarrow \text{Level of differencing}$$

$$y_t' = \delta + \phi_1 y_{t-1}' + \phi_2 y_{t-2}' + \cdots + \phi_p y_{t-p}' + \theta_1 e_{t-1}' + \theta_2 e_{t-2}' + \cdots + \theta_q e_{t-q}'$$

ARIMA Example

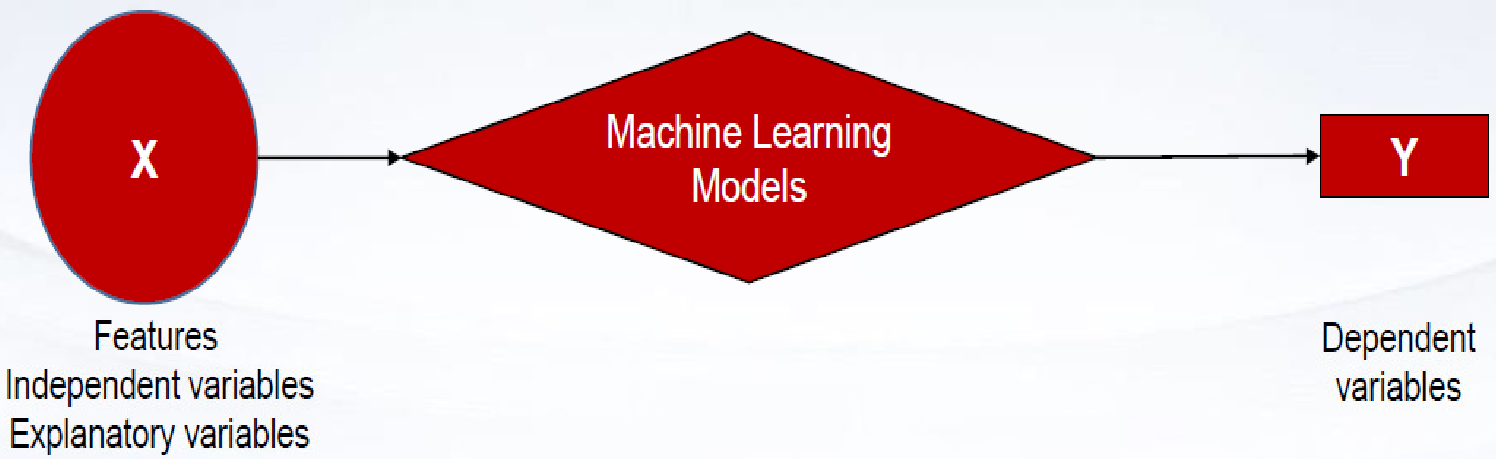


ARIMA Limitations

- ARIMA assumes the underlying time series is linear
- Difficult to fit highly non-stationary time series
- Cannot deal with multiple time series at the same time

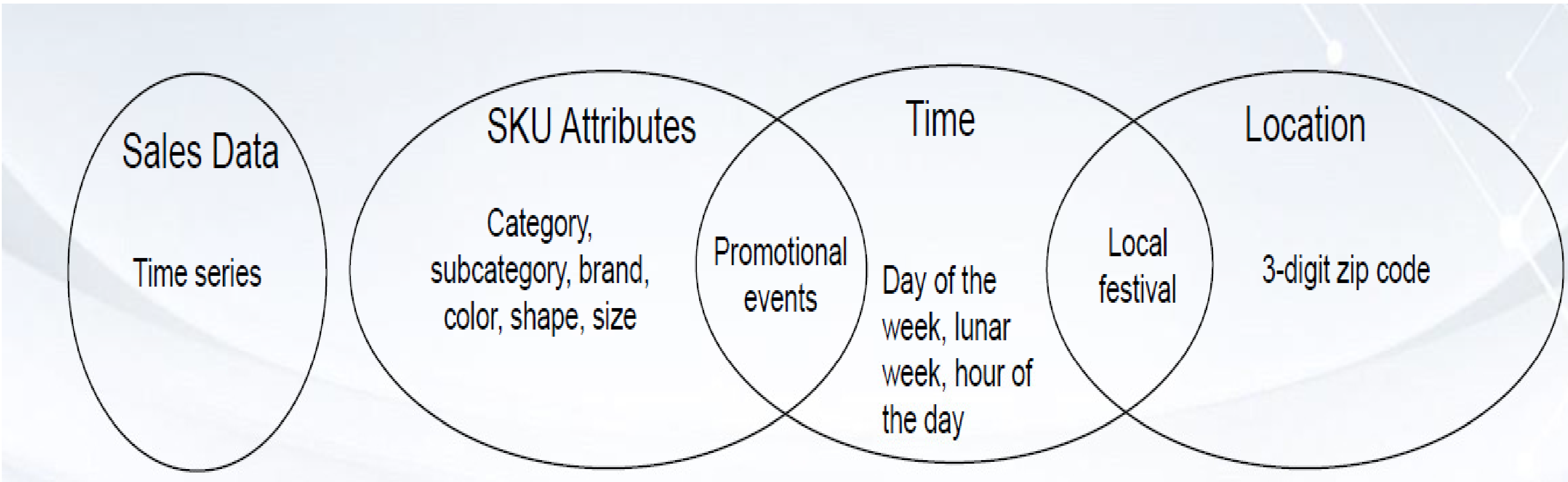
	Scorecard
Highly non-stationary	Limited
Multiple time series	Limited
Multi-horizon forecast	Yes
Probabilistic forecast	Yes

Machine Learning Model



- Flexible in having more features (X) in the model
- No assumption w.r.t the demand distribution
- One model for all time series
- Feature engineering is important

Feature Engineering



Mean of the past 7-day sales
Variance of the past 7-day sales
Max sales of the past 14 days
Sales of the 7th day in the past
90% sales quantile of last month

Festival encoding ([0,0,0,1,0,0])
Percentage of discount
Promotional type (hash id)
Category (hash id)
SKU Name (embedding vector)

ML Limitations

- Incorporating features requires manual work
  - Requires human expertise
  - Some features are difficult to capture
  - Time consuming

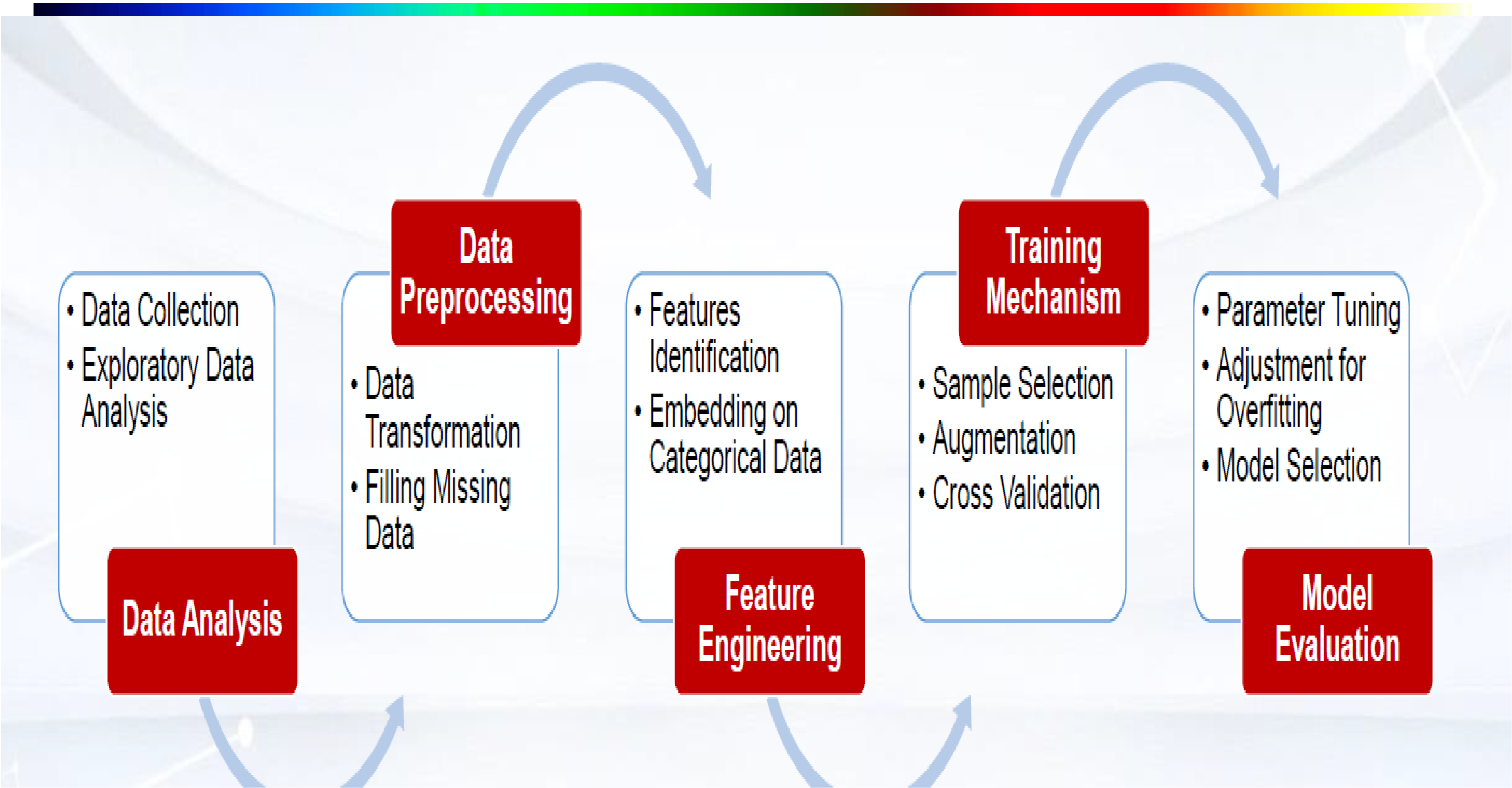
	Scorecard
Highly non-stationary	Yes
Multiple time series	Yes
Multi-horizon forecast	Yes
Probabilistic forecast	Yes

Methods Comparison

- Stochastic time-series models
  - Good model interpretability
  - Limited model complexity to handle non-linearity
  - Difficult to incorporate cross features among multiple time series
- Machine learning
  - Flexible and can incorporate any feature explicitly
  - Heavy workload in terms of feature engineering
- Deep learning
  - Very flexible and automated feature detection
  - Poor model interpretability

	Stochastic Time Series	Machine Learning	Deep Learning
Highly non-stationary	Limited	Yes	Yes
Multiple time series	Limited	Yes	Yes
Multi-horizon forecast	Yes	Yes	Yes
Probabilistic forecast	Yes	Yes	Yes

Model Framework



Connect To Customers

- Connecting products to customers seamlessly in all scenarios.
- People are different in many ways

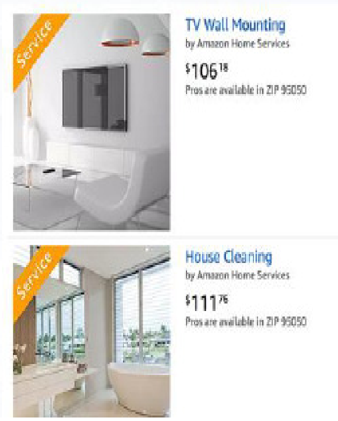
Background | Locations | Activities | Social Connections

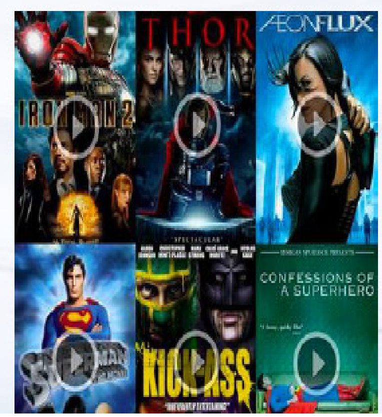



Connect To Customers

- Products are different in many ways

Physical Products	Digital Goods	Service	Content

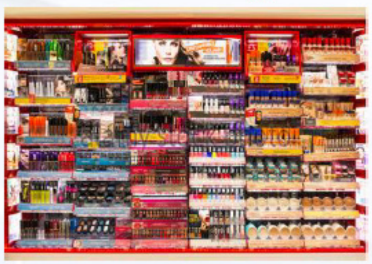




Connect To Customers

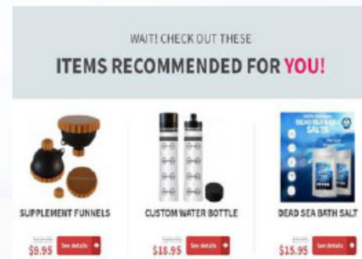
- Delivering the right products to the right customers at the right place and right time

Browsing

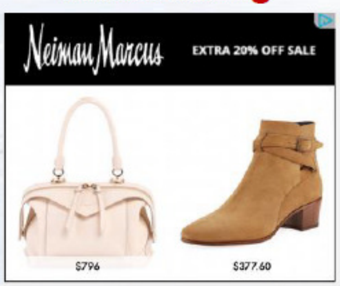


Search

Recommendation



Advertising



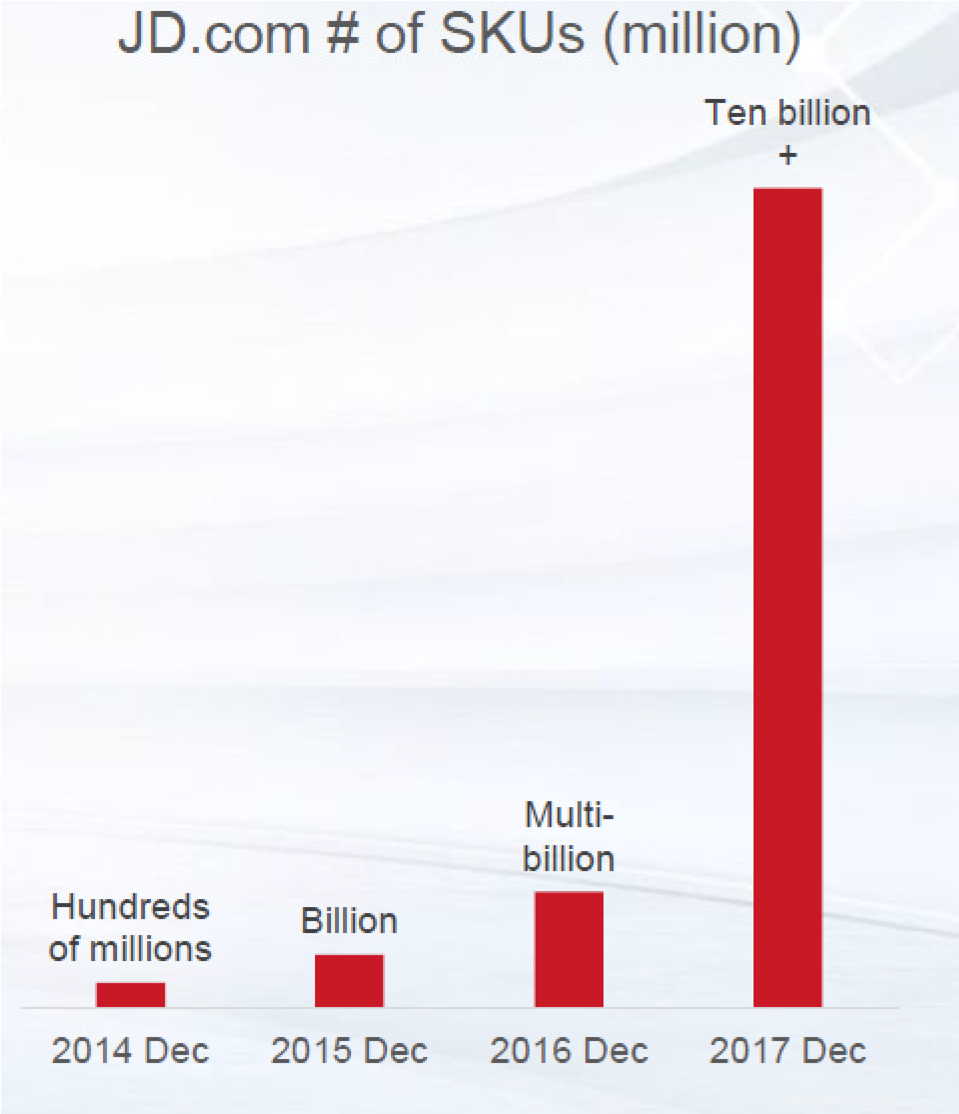
Social Network

C2M

Product Recommendation in E-Commerce

- With an ever increasing number of products available to customers, **delivering the most appropriate products to customers** has become a core functionality of retail platforms.
- Naturally, product recommendation has now become a centerpiece of e-commerce platforms.

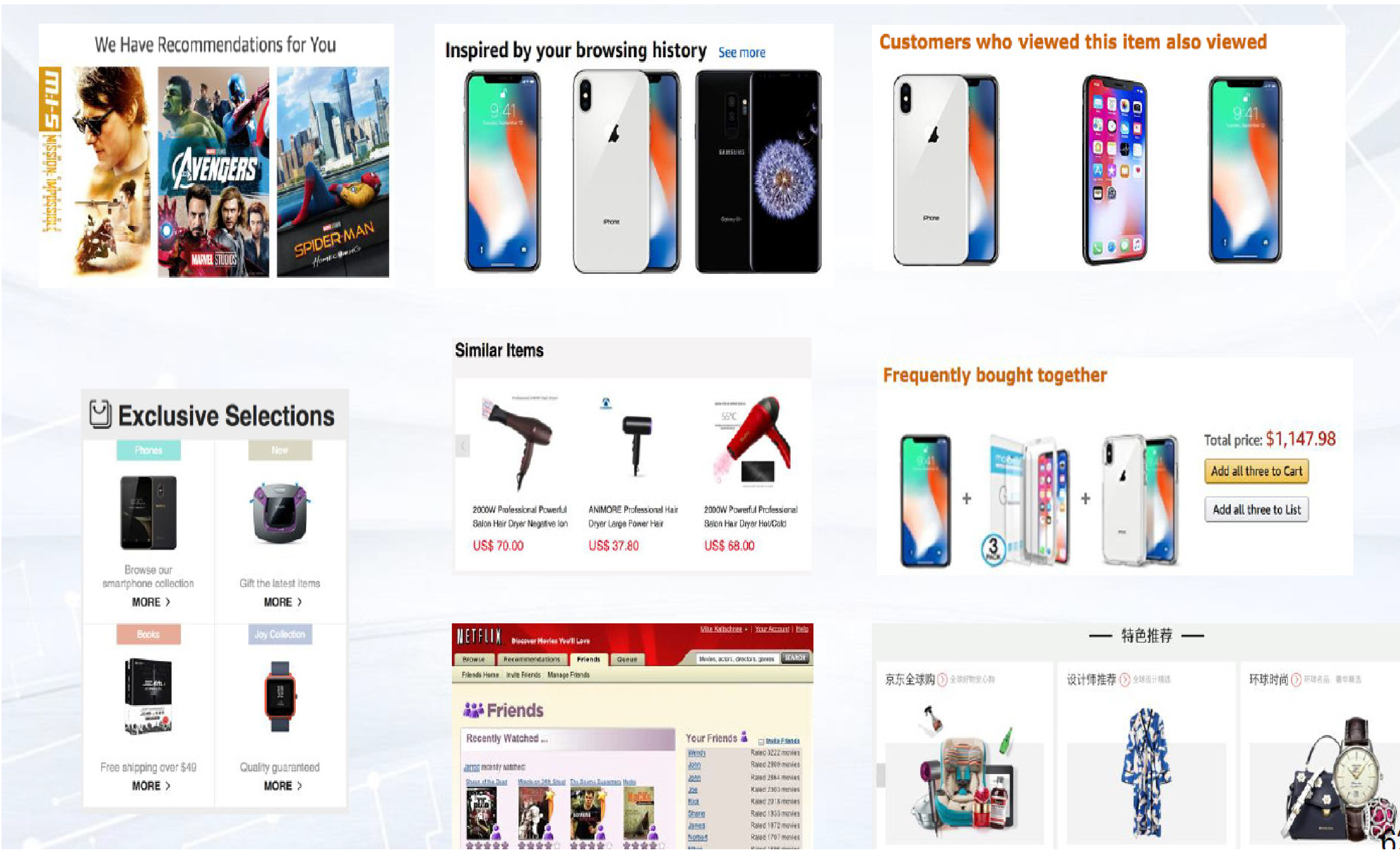
Product Recommendation in E-Commerce



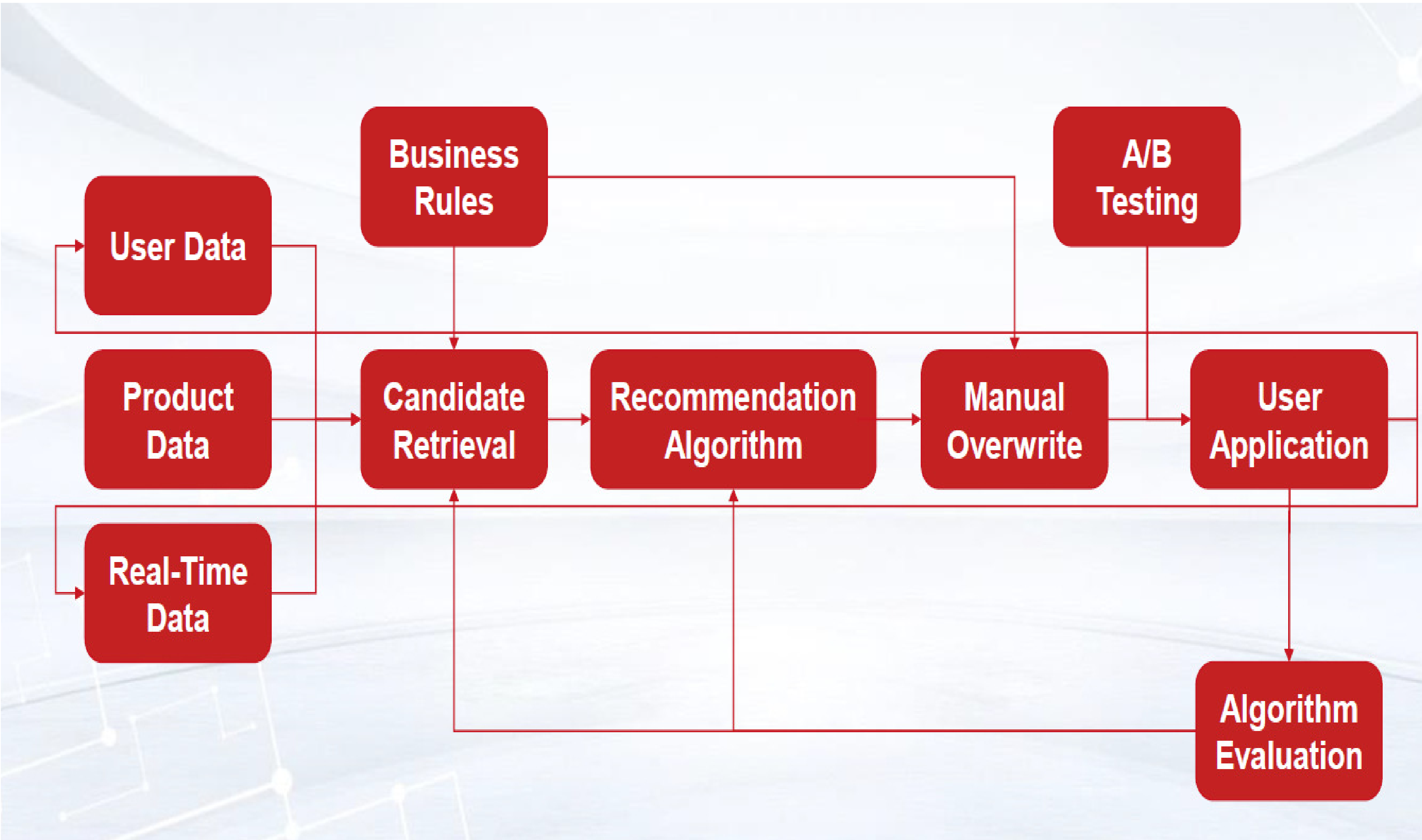
Product Recommendation in E-Commerce

- 35% of goods purchased on Amazon and 75% of content watched on Netflix come about as a result of product recommendations.

Product Recommendation in E-Commerce



Product Recommendation in E-Commerce



Product Recommendation: inputs

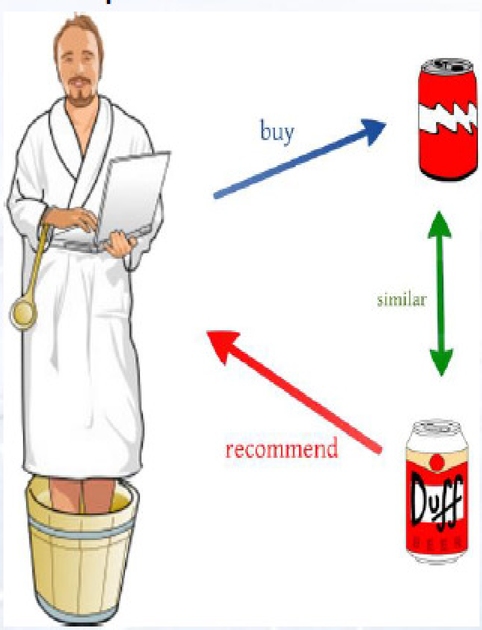
User Data	Product Data	Real-Time Data
User Identifier	Category	Location
Demographic Information	Brand/Manufacture	Time
Shopping Habit	Origin	Device
Shopping History	Rating	Session Information
Browse History	Product Price	Product Searches
Favorite/Disliked Items	Product Description	Product Impressions
Devices	Product Images	Product Browses
…	…	…
Describe users, their preferences, their histories, etc.	Describe the all things related to the products and all product-related user interactions.	Describe the shopping scenario and users' interaction with the shopping scenario

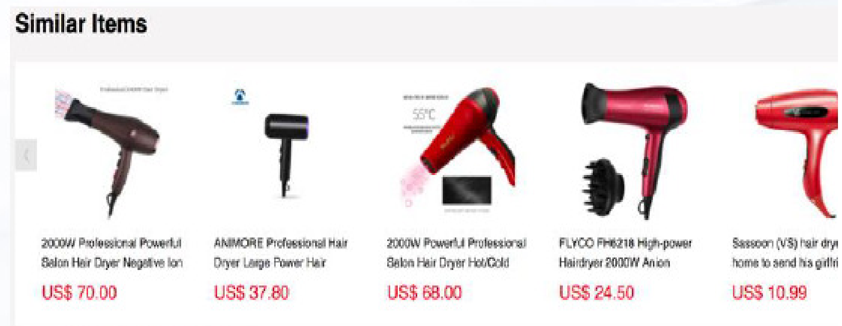
Types of Product Recommendation Algorithms

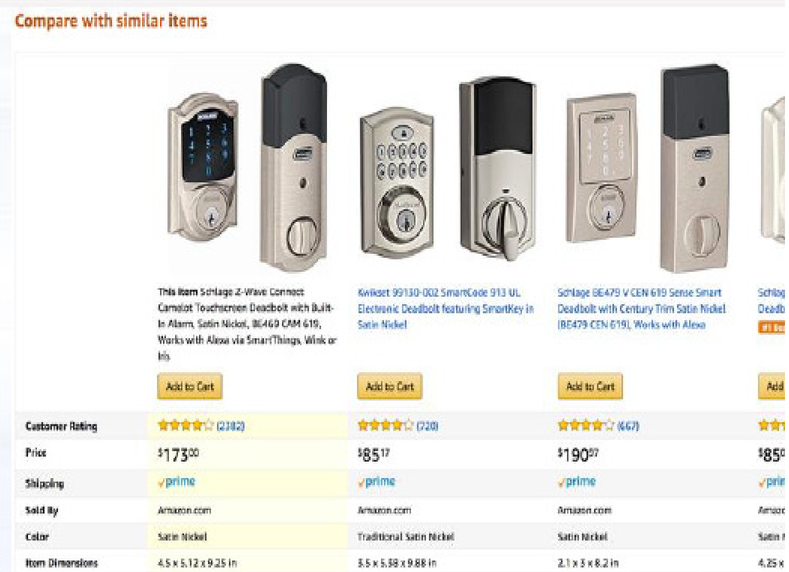
- Content Based Methods (Ricci et al., 2015; Pazzani and Billsus, 2007)
  - Recommends items similar to those liked/purchased by the customer in the past
  - Use attributes of items/customers
- Collaborative Filtering Based Methods (Goldberg et al., 1992; Linden et al., 2003; Schafer et al., 2007)
  - Recommends items liked or purchased by similar customers
  - Enable exploration of diverse content

Content Based Recommendation

- Based on similarity of item attributes
  - Item name, categorical information, price, description, technical specs, etc.
- Challenges:
  - Vague definition of similarity
  - Cannot provide diverse content



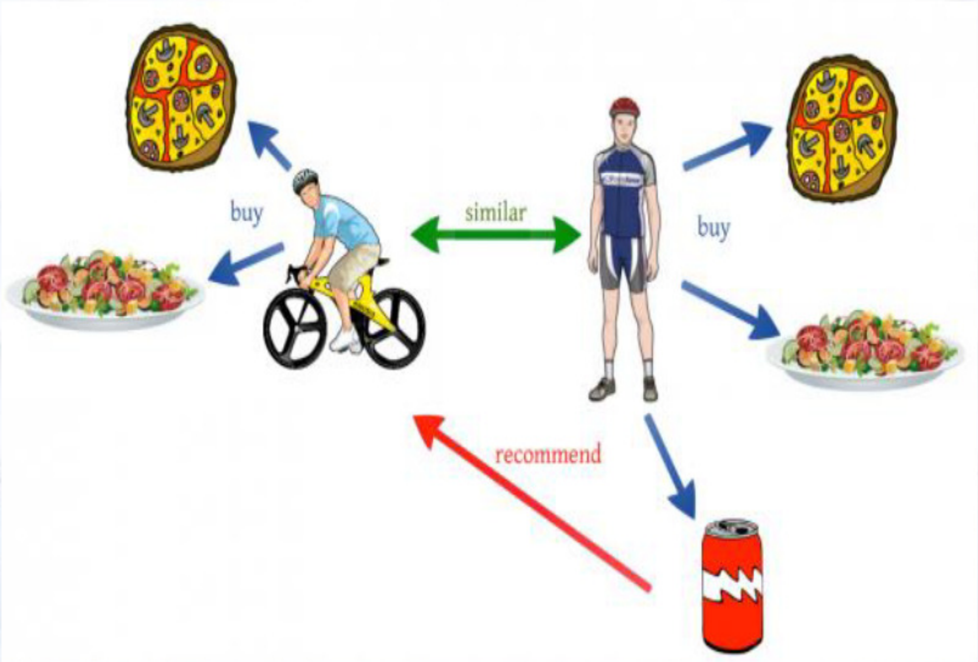




/d4datascience.wordpress.com/2016/07/22/recommender-systems-101/)

Collaborative Filtering

- Collaborative Filtering is the process of filtering or evaluating items using the opinions of other people.



Serve Customers

A transformed shopping experience driven by cutting-edge technologies in big data and operations research

Unlimited choices available online

Convenience

- Anywhere
- Anytime

Fast delivery

Serve Customers

Big data introduces new opportunities to better serve customers, as well as challenges to traditional solution methods

	Unlimited choices available online	Convenience	Fast delivery
Challenges	• Limited capacity at local warehouses • Delivery speed • Inventory placement • …	• Local demand • Inventory replenishment • …	• Balance online and offline demand • Omni-channel fulfillment • …

Serve Customers

Inventory Placement

JD’s nationwide convenience stores

Expanding nationally, especially in rural areas

Expected to reach 1M stores by 2023

Cater to local needs and support fulfilling online demand

Inventory Placement

JD's nationwide convenience stores

Expanding nationally, especially in rural areas

Expected to reach 1M stores by 2023

Cater to local needs and support fulfilling online demand

Inventory Placement

Problem:

- How should inventory be allocated to JD's stores nationwide?

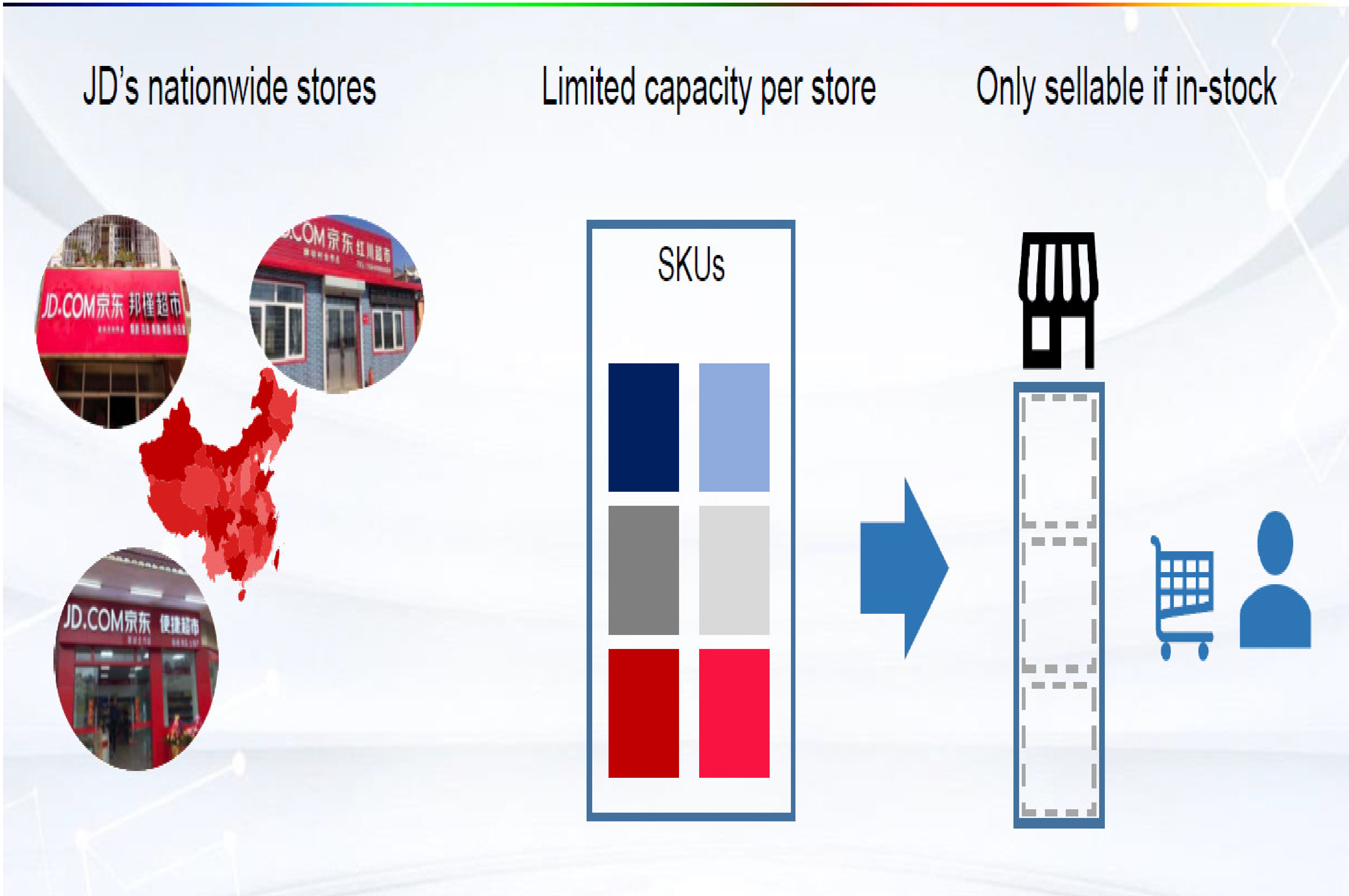
Goal:

- Delivery products to meet local needs
- Satisfactory fulfillment rate

Constraint:

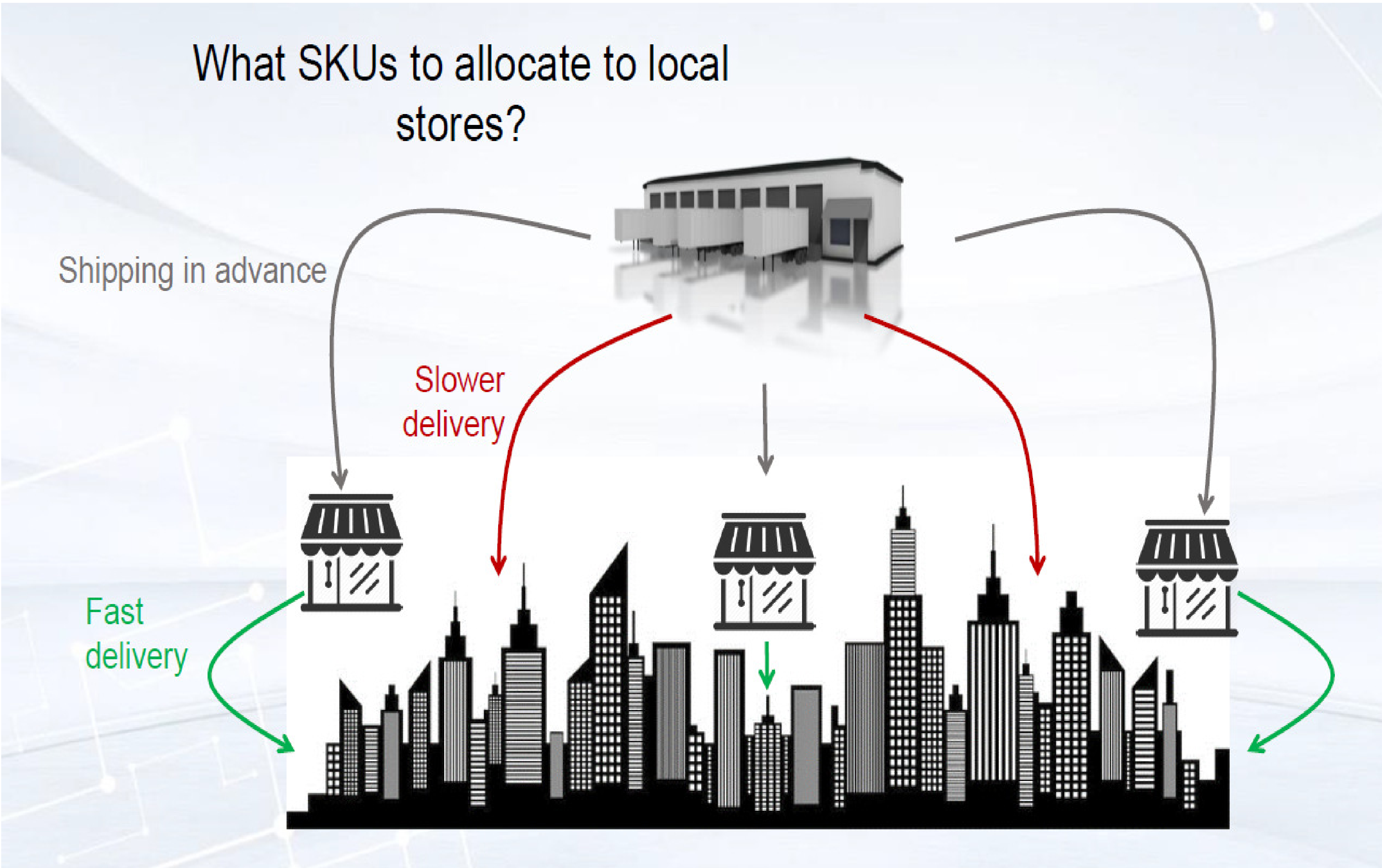
- Limited store capacity
- …

Inventory Placement-Offline Demand

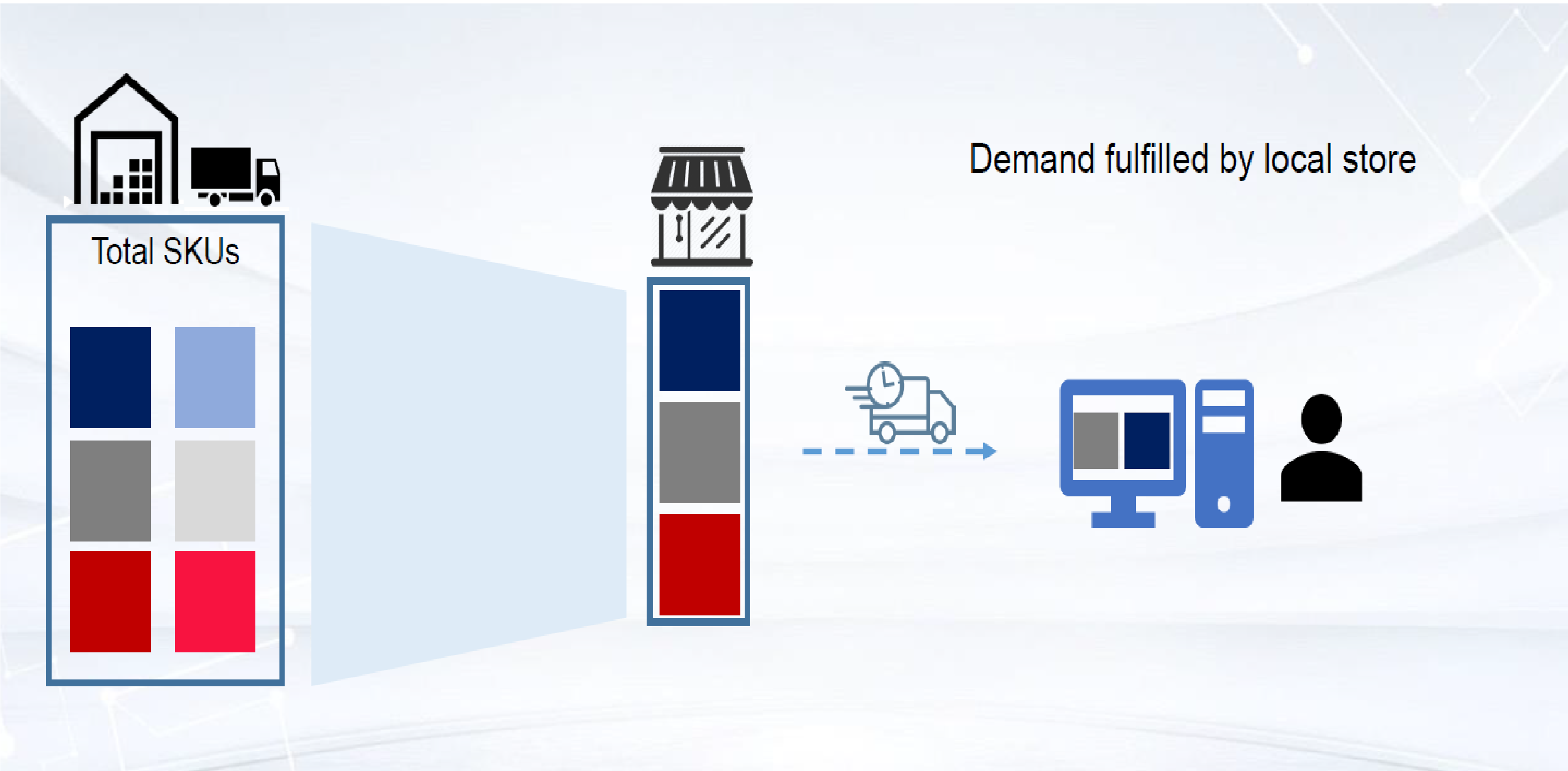


Inventory Placement-Online Demand

Inventory Placement



An Assortment Problem



Local fulfillment enables expediated delivery that delights customers