The Rise of Engineering-Driven Analytics

Sameer M. Prabhu, Ph. D.
Industry Marketing Director
MathWorks
Sameer.Prabhu@mathworks.com
The Rise of Engineering-Driven Analytics

Images sourced from ASML, DAF Trucks, and Philips Healthcare websites
The Rise of Engineering-Driven Analytics
Analytics are now pervasive

- Engineering
- Business
- Transactional
- Desktop -
  - Multicore, GPU
  - Clusters
  - Cloud computing
  - Hadoop
- Neural Networks
- Classification
- Clustering
- Regression
- ...and much more...
Analytics are now pervasive

Apply robust, statistically-motivated methods to data produced from complex systems to understand what has happened and predict what will happen.
Analytics in e-commerce

Use **Image Processing** to add image data to the model, improving performance

- Images
- Social profile
- Geolocation
- Keystroke logs
- Transactions

**IMPROVED Predictive Model**

**Offer to Customer**
Source: Gartner Big Data Industry Insights, March 2016
The Rise of Engineering-Driven Analytics
Architecture of an analytics system

Data from instruments and connected systems

Data from business systems

Predictive Model deployed in smart systems using Model-Based Design

Predictive Model deployed on cloud and business systems

MATLAB Integrates in Embedded System and Enterprise IT Workflows
Example – BuildingIQ

Adaptive building energy management
Real-time, closed-loop optimization algorithms

Predictive Model
deployed on cloud with client system
and real-time data feeds

DATA - Billions
Physics, energy costs, ambient temperature, operation schedule

Analytics and
Machine Learning
plus system identification, control theory & more

Current energy costs & demand
Weather Feeds

MATLAB Toolboxes Just Work –
and work together!
We could rapidly translate our prototypes into production algorithms that deal reliably with real-world noise and uncertainty.

Borislav Savkovic, BuildingiQ

Why MATLAB?

- Robust numerical algorithms
- Extensive visualization and analytics tools
- Industry-robust and reliable mathematical optimization routines
- Good object-oriented framework
- Ability to interface with Java (for backend work)
- Running MATLAB in the cloud in production
- Unit-testing framework

MATLAB Impeccable Numerics for Trusted Results
Example – Scania
Automatic emergency braking using sensor fusion and analytics
50 km/h - sudden brake
Using Model-Based Design to build and deploy the analytics in an embedded control system

MATLAB Integrates Analytics and Model-Based Design
Implementing Sensor Fusion at Scania

Vehicle logs of video and radar data

Predictive Model deployed on vehicle

Machine learning to develop fusion algorithms for situation detection
The Rise of Engineering-Driven Analytics

- Automotive
- Off-highway vehicles
- Aeronautics

- Retail
- Finance
- Healthcare management
- Internet

- Industrial Automation
- Oil & Gas
- Medical Devices
- Clean Energy
Predictive Maintenance for polymer-based production machines

Sensor Data (~1 minute)
10-100 sensors/machine

Quality State (~40 minutes)

Classification using Statistics, Machine Learning, and Neural Networks
Deployment – a MATLAB App used by machine operators
The need for data scientists

- Domain expertise
- Coding and integration skills
- Statistical and mathematical knowledge
Shortage of data scientists, big data pros vexes IoT efforts

CRUNCH NETWORK
How To Stem The Global Shortage Of Data Scientists
Posted Dec 31, 2015 by Amy Gershkoff (@amygershkoff)

Big data talent shortage: How to bridge the gap?
By Abhishek Raval on May 29, 2015

What they say
• Expand university programs
• Train existing analysts
**Core Concepts in Data Analysis**

Learn both theory and application for basic methods that have been invented either for developing new concepts – principal components or clusters, or for finding interesting correlations – regression and classification. This is preceded by a thorough analysis of 1D and 2D data.

---

**Big Data Science with the BD2K-LINCS Data Coordination and Integration Center**

Learn various methods of analysis including unsupervised clustering, gene-set enrichment analyses, Bayesian Inference, network visualization, and supervised machine learning applications to LINCS data and other relevant Big Data from high content molecular and proteomic profiling of human cells.
Student Contest
use process control data
to improve semiconductor yields

- 21 teams competed
- Wafer Big Data in Hadoop
- MATLAB used by winning team and 2nd place team
MATLAB lets you be your own data scientist

MATLAB is Designed and Documented to be Easy for Engineers and Scientists to Use
Manufacturing Sustainability Hackathon

- Understand what drives power consumption
- Create an easy to use insightful dashboard
- Provide recommendations to reduce power consumption
  - Challenge: Too many sensors of unknown significance (1200+ sensors)
  - Prize overall best solution for one of S&P 10 companies
  - Competing with top of the industry companies and analytics platforms
In MATLAB

• Native support for engineering data
• Database interfaces
• Streaming

NEW for MATLAB

Audio System Toolbox R2016a
Vision HDL Toolbox R2015a

Limited users, scope, & technology

Big Data

• Engineering
• Business
• Transactional

Compute Power

• Desktop - Multicore, GPU
• Clusters
• Cloud computing
• Hadoop

Machine Learning

• Neural Networks
• Classification
• Clustering
• Regression
• ...and much more...

Pervasive users, scope, & technology
In MATLAB

- Native support for engineering data
- Database interfaces
- Streaming
- **Datastore** R2014b text, image, video, Excel files, ...
- **Mapreduce** R2014b

- Engineering
- Business
- Transactional

- Desktop - Multicore, GPU
- Clusters
- Cloud computing
- Hadoop

- Neural Networks
- Classification
- Clustering
- Regression
Big Data

- Engineering
- Business
- Transactional

In MATLAB

- Native support for engineering data
- Database interfaces
- Streaming
- Datastore R2014b text, image, video, Excel files, ...
- Mapreduce R2014b

Machine Learning

- Desktop
- Multicore, GPU
- Clusters
- Cloud computing
- Hadoop

MATLAB is fast:
- heavily optimized libraries
- JIT compiled
- takes advantage of the compute power you have

Limited users, scope, & technology

Pervasive users, scope, & technology
In MATLAB

- Native support for engineering data
- Database interfaces
- Streaming
- **Datastore** R2014b
text, image, video, Excel files, ...
- **Mapreduce** R2014b

- Multicore & GPU
- MATLAB Distributed Computing Server and EC2 Support
- **Hadoop support** R2014b
- MATLAB Production Server

NEW from MathWorks

The open data platform for the Internet of Things
Big Data

- Engineering
- Business
- Transactional

Compute Power

- Desktop
  - Multicore, GPU
- Clusters
- Cloud computing
- Hadoop

Machine Learning

- Neural Networks
- Classification
- Clustering
- Regression

In MATLAB

- Native support for engineering data
- Database interfaces
- Streaming
- **Datastore** R2014b
text, image, video, Excel files, ...
- **Mapreduce** R2014b
- Multicore & GPU
- MATLAB Distributed Computing Server and EC2 Support
- **Hadoop support** R2014b
- MATLAB Production Server
- Statistics and Machine Learning Toolbox
- **Classification Learner App** R2015a
- Neural Network Toolbox
- **CNNs for Deep learning** R2016a
- Machine learning with code generation
Classification Learner App
in Statistics and Machine Learning Toolbox
MATLAB Apps for Data Analytics

Distribution Fitting
System Identification
Signal Analysis
Wavelet Design and Analysis
Neural Net Fitting
Neural Net Pattern Recognition
Training Image Labeler

and many more…

With MATLAB Apps, you can complete data science tasks more quickly and easily than custom programming.
Using MATLAB R2016a

App Designer
Using MATLAB R2016a

App Designer
Deep Learning with Neural Network Toolbox - New in R2016a
Deep Learning with Neural Network Toolbox - New in R2016a
Example – cellscope®
First consumer otoscope in a mobile device machine learning and computer vision
The Rise of Engineering-Driven Analytics

Be your own Data Scientist!