MATLAB EXPO 2016
Modelle für die Zukunft dank prädiktiver Datenanalyse

Jérémy Huard, MathWorks
30%
\[
\frac{\partial u}{\partial t} - \alpha \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = 0
\]

Thermodynamic properties

Temperatures change

Humans have comfort bounds

Electricity demand varies
BuildingIQ Develops Proactive Algorithms for HVAC Energy Optimization in Large-Scale Buildings

Large-scale commercial buildings can reduce energy costs by 10–25% with BuildingIQ’s energy optimization system.

Office buildings, hospitals, and other large-scale commercial buildings account for about 30% of the energy consumed worldwide. The heating, ventilation, and air-conditioning (HVAC) systems in these buildings are often inefficient because they do not take into account changing weather patterns, variable energy costs, or the building's thermal properties.

BuildingIQ has developed Predictive Energy Optimization™ (PEO), a cloud-based software platform that reduces HVAC energy consumption by 10–25% during normal operation. PEO was developed in cooperation with the Commonwealth Scientific and Industrial Research Organisation (CSIRO), HVAC pressure sensors, as well as weather and energy cost data. A single building often produces billions of data points, and the scientists and engineers needed tools for efficiently filtering, processing, and visualizing this data.

To run their optimization algorithms, the scientists and engineers had to create an accurate mathematical model of a building's thermal and power dynamics. The algorithms would use this calculated model to run constrained optimizations that maintained occupant comfort while minimizing energy costs.

BuildingIQ needed a way to rapidly develop mathematical models, test optimization algorithms, and deploy them in the cloud.
Traits of Data Analytics applications

1. Diverse and/or Big Data
2. Advanced Algorithms
3. Deployment
Why MATLAB?

1. Analytics that increasingly require both business and engineering data

2. Enable Domain Experts to be Data Scientists

3. Develop embedded systems with analytics powered functionality

4. Develop analytics to run on both enterprise and embedded platforms

DATA
• Engineering, Scientific, and Field
• Business and Transactional

Embedded Systems

Business Systems

Data Scientist
Why MATLAB?

1. Analytics that increasingly require **both business and engineering data**

DATA
- Engineering, Scientific, and Field
- Business and Transactional

2. Enable **Domain Experts to be Data Scientists**

3. Develop **embedded systems** with analytics powered functionality

4. Develop analytics to run on **both enterprise and embedded platforms**

Smarter Embedded Systems

Business Systems

DATA
- Engineering, Scientific, and Field
- Business and Transactional
Business and Engineering Data

Business and Transactional Data

Repositories
- Databases
- Hadoop

File I/O
- Text
- Spreadsheet
- XML

Web Sources
- HTML
- Mapping
- Financial datafeeds
- RESTful
- JSON

Engineering, Scientific, and Field Data

File I/O
- Text
- Spreadsheet
- XML
- CDF/HDF
- Image
- Audio
- Video
- Geospatial

Communication Protocols
- CAN (Controller Area Network)
- DDS (Data Distribution Service)
- OPC (OLE for Process Control)
- XCP (eXplicit Control Protocol)

Real-Time Sources
- Sensors
- GPS
- Instrumentation
- Cameras
- Communication systems
- Machines (embedded systems)

“No matter what industry our client is in, and no matter what data they ask us to analyze—text, audio, images, or video—MATLAB enables us to provide clear results faster.”

Dr. G Subrahmanya VRK Rao, Cognizant
Data handling and visualization
# Feature Engineering

## Extracting Information from Data

### Sensor data
- Peak analysis
- Pulse and transition metrics
- Spectral measurements (power, bandwidth, mean frequency, median frequency)

### Image and video data
- Bag of visual words
- HOG (Histogram of Oriented Gradients)
- Minimum Eigenvalue algorithm
- Local feature descriptors
- Edge detection

### Transactional data
- Decomposing timestamps into components (day, month, day of week, etc.)
- Calculation of aggregate values
- Complexity reduction

---

## High-quality domain-specific libraries

<table>
<thead>
<tr>
<th>Data type</th>
<th>Common Techniques for Deriving Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor data</td>
<td><strong>Signal Processing</strong></td>
</tr>
<tr>
<td>Image and video data</td>
<td><strong>Image Processing</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Computer Vision</strong></td>
</tr>
<tr>
<td>Transactional data</td>
<td><strong>Statistics</strong></td>
</tr>
</tbody>
</table>
Why MATLAB?

1. Analytics that increasingly require **both business and engineering data**

2. **Enable Domain Experts to be Data Scientists**

3. Develop **embedded systems** with analytics powered functionality

4. Develop analytics to run on **both enterprise and embedded platforms**

DATA
- Engineering, Scientific, and Field
- Business and Transactional

Smarter Embedded Systems

Business Systems

Data Scientist
Enabling Domain Experts to be Data Scientists
Built-in algorithms

Clustering
- Hierarchical Clustering
  Produce nested sets of clusters
- k-Means and k-Medoids Clustering
  Cluster by minimizing mean or medoid distance, calculate Mahalanobis
- Gaussian Mixture Models
  Cluster based on Gaussian mixture models using the EM algorithm
- Nearest Neighbors
  Find nearest neighbors using exhaustive search or kd-tree search
- Hidden Markov Models
  Markov models for data generation
- Cluster Visualization and Evaluation
  Plot clusters of data and evaluate optimal number of clusters

Regression
- Linear Regression
  Multiple, stepwise, multivariate regression models, and more
- Generalized Linear Models
  Logistic regression, multinomial regression, Poisson regression
- Nonlinear Regression
  Nonlinear fixed- and mixed-effects regression models
- Support Vector Machine Regression
  Support vector machines for regression models
- Gaussian Process Regression
  Gaussian process regression models (kriging)
- Regression Trees
  Binary decision trees for regression
- Regression Tree Ensembles
  Random forests, boosted and bagged regression trees

Classification
- Classification Trees
  Binary decision trees for multiclass learning
- Discriminant Analysis
  Regularized linear and quadratic discriminant analysis
- Naive Bayes
  Naive Bayes model with Gaussian, multinomial, or kernel predictors
- Nearest Neighbors
  k nearest neighbors classification using Kd-tree search
- Support Vector Machine Classification
  Support vector machines for binary or multiclass classification
- Classification Ensembles
  Boosting, random forest, bagging, random subspace, and ECOC ensembles for multiclass learning
- Model Building and Assessment
  Feature selection, cross validation, predictive performance evaluation, classification accuracy comparison tests

Confusion Matrix

MATLAB EXPO 2016
Interactive Apps to focus on machine learning, not programing

Neural network Apps

Features
- Train models
- Assess results
- Export models to the MATLAB or generate MATLAB code
Why MATLAB?

1. Analytics that increasingly require **both business and engineering data**

DATA
- Engineering, Scientific, and Field
- Business and Transactional

2. Enable **Domain Experts to be Data Scientists**

3. Develop **embedded systems** with analytics powered functionality

4. Develop analytics to run on **both enterprise and embedded platforms**

Smarter Embedded Systems

Business Systems

Data Scientist
Smarter Embedded Systems

- RESEARCH
- REQUIREMENTS
- DESIGN
  - Environment Models
  - Physical Components
  - Algorithms
- IMPLEMENTATION
  - C, C++
  - VHDL, Verilog
  - Structured Text
    - MCU
    - DSP
    - FPGA
    - ASIC
    - PLC
- TEST AND VERIFICATION
- INTEGRATION

Airbus
Battery management
GM
Climate control
Festo
Industrial robots
Sonova
Hearing implants
Weinmann
Transport ventilator
manroland
Printing presses
FLIR
Thermal imaging
Daimler
Cruise controller
ABB
Smart Grid controller

MATLAB EXPO 2016
Why MATLAB?

1. Analytics that increasingly require both business and engineering data

2. Enable Domain Experts to be Data Scientists

3. Develop embedded systems with analytics powered functionality

4. Develop analytics to run on both enterprise and embedded platforms

DATA
- Engineering, Scientific, and Field
- Business and Transactional

Smarter Embedded Systems
Deploying Algorithms to Enterprise Systems

**MATLAB Compiler** enables sharing MATLAB programs without integration programming

**MATLAB Compiler SDK** provides implementation and platform flexibility for software developers

**MATLAB Production Server** provides the most efficient development path for secure and scalable web and enterprise applications
Enterprise Integration – Forecasting Model
MATLAB Differentiators

1. Analytics that increasingly require both business and engineering data

DATA
- Engineering, Scientific, and Field
- Business and Transactional

3. Develop embedded systems with analytics powered functionality

2. Enable Domain Experts to be Data Scientists

4. Develop analytics to run on both enterprise and embedded platforms

Smarter Embedded Systems

Business Systems
Learn More

Presentations

1.2. Data Science mit MATLAB
Session Chair: Dr. Alexander Diethert

15:15 Analyse von operationellen Flugdaten aus einem Hadoop System unter Verwendung von MapReduce und dem MATLAB Distributed Computing Server
Lukas Hohndorf, TU München

15:45 Mensch-Maschine-Interface zur multisensorischen Prozessüberwachung in der Polymerindustrie
Michael Kohlert, Mondi Gronau
Dr. Sarah Drewes und Elmar Tarajan, MathWorks

16:15 Algorithmen für Predictive Maintenance effizient entwickelt mit MATLAB
Dr. Sarah Drewes, MathWorks

MATLAB EXPO 2016

Machine Learning

mathworks.com/machine-learning

Consulting

Training