MATLAB EXPO 2016
Get More From Your Data with Data Analytics

Francesca Perino
What do we have to work with?
Buildings have thermodynamic properties

\[
\frac{\partial u}{\partial t} - \alpha \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = 0
\]
Temperatures change
Electricity demand varies
Humans have comfort bounds
\[ \frac{\partial u}{\partial t} - \alpha \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = 0 \]
BuildingIQ Develops Proactive Algorithms for HVAC Energy Optimization in Large-Scale Buildings

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 visualising 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 visualise energy performance data.
Traits of Data Analytics applications

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

1. Diverse and/or Big Data

Office buildings, hospital facilities, and large-scale commercial buildings often have heating, ventilation, and air conditioning (HVAC) systems that are about 30% of the energy consumed worldwide. The heating, ventilation, and air conditioning (HVAC) systems in these buildings often do not account for changing weather conditions, leading to energy waste. A single building often works independently, not sharing energy cost data. A single building often works independently, not sharing energy cost data. A single building often works independently, not sharing energy cost data.

To run their optimization algorithms, the BuildingIQ team created a model of a building to predict energy costs. The algorithm uses a cloud-based platform that analyzes historical energy data, energy forecasts, and planned maintenance schedules. This allows BuildingIQ to adjust HVAC systems based on historical and forecasted energy costs. BuildingIQ has developed Predictive Energy Optimization™ (PEO), a cloud-based software platform that reduces HVAC energy costs by 10–25% with BuildingIQ's energy optimization system.

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

Embedded Systems

Business Systems

Data Scientist
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

MATLAB EXPO 2016
Business and Engineering Data

Business and Transactional Data

<table>
<thead>
<tr>
<th>Repositories</th>
<th>Web Sources</th>
<th>File I/O</th>
<th>Communication Protocols</th>
<th>Real-Time Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Databases (SQL)</td>
<td>HTML, Mapping,</td>
<td>Text, Spreadsheet, XML, CDF/HDF, Image, Audio, Video, Geospatial</td>
<td>CAN (Controller Area Network), DDS (Data Distribution Service), OPC (OLE for Process Control), XCP (eXplicit Control Protocol)</td>
<td>Sensors, GPS, Instrumentation, Cameras, Communication systems, Machines (embedded systems)</td>
</tr>
<tr>
<td>NoSQL</td>
<td>Financial datafeeds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hadoop</td>
<td>RESTful, JSON</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Engineering, Scientific, and Field Data

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

“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
Accessing Data
Datastore

HDFS

Node

Data

Map

Reduce

Node

Data

Map

Reduce

Node

Data

Map

Reduce

Node

Data

Map

Reduce
Table
Preprocessing Data

Read in engine data
The data used is the sensor readings taken off of the equipment. Maintenance was done after 125 cycles of use, regardless of whether the equipment seemed to need it or not, so we only have the first 125 cycles off of each engine. Our maintenance staff tells us that while some of them were in need of maintenance, many were fine and could have run longer before being serviced. So far no failures have occurred prior to maintenance.

```matlab
sensorData = readtable('train_F0001_Unit_1.csv','ReadVariableNames',true);
```

Select relevant variable names
```matlab
variableNames = {'Unit' 'Time' 'LPCOutletTemp' 'HPCOutletTemp' ...};
```
Signal Processing

- cheby2
- filter
- rms
- pwelch
- periodogram
- xcov
- findpeaks
- movmean
- movstd
- ...

\[ H(j\omega) = \frac{e^{2} C_{N}(\omega_{1}/\omega)}{1 + e^{2} C_{N}(\omega_{2}/\omega)} \]

\[ C_{N}(\omega_{1}/\omega) = \begin{cases} \cos[N \cos^{-1}(\omega_{1}/\omega)] & , \quad |\omega| \geq \omega_{2} \\
\cosh[N \cosh^{-1}(\omega_{1}/\omega)] & , \quad |\omega| \leq \omega_{2} \end{cases} \]

\[ F_{\alpha}(f) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \int_{-\alpha}^{\alpha} \int_{-\alpha}^{\alpha} \text{rect}(\alpha - \beta) \cdot \xi_{\alpha}(\beta) \cdot \text{rect}(\beta - f) \cdot \xi_{\beta}(f) \, d\beta \, df \]

\[ c_{n}(m) = \begin{cases} \sum_{k=0}^{N-m-1} (x_{i\pi} + m) \cdot \frac{1}{N} \sum_{i=0}^{N-3} \xi_{k} \cdot \left( x_{i\pi} - \frac{1}{N} \sum_{i=0}^{N-3} \xi_{j} \right) & , \quad m \geq 0 \\
\xi_{\pi}(-m) & , \quad m < 0 \end{cases} \]
Image Processing

Perform image processing, analysis, visualization, and algorithm development

- Image display and exploration
- Image enhancement
- Image analysis
- Morphological operations
- Image registration
- Geometric transformation
- ROI-based processing
# Feature Engineering – Extracting Information from Data

<table>
<thead>
<tr>
<th>Data type</th>
<th>Common Techniques for Deriving Features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sensor data</strong></td>
<td>• Peak analysis</td>
</tr>
<tr>
<td></td>
<td>• Pulse and transition metrics</td>
</tr>
<tr>
<td></td>
<td>• Spectral measurements (power, bandwidth, mean frequency, median frequency)</td>
</tr>
<tr>
<td><strong>Image and video data</strong></td>
<td>• Bag of visual words</td>
</tr>
<tr>
<td></td>
<td>• HOG (Histogram of Oriented Gradients)</td>
</tr>
<tr>
<td></td>
<td>• Minimum Eigenvalue algorithm</td>
</tr>
<tr>
<td></td>
<td>• Local feature descriptors</td>
</tr>
<tr>
<td></td>
<td>• Edge detection</td>
</tr>
<tr>
<td><strong>Transactional data</strong></td>
<td>• Decomposing timestamps into components (day, month, day of week, etc.)</td>
</tr>
<tr>
<td></td>
<td>• Calculation of aggregate values</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
"MATLAB has helped accelerate our R&D and deployment with its robust numerical algorithms, extensive visualization and analytics tools, reliable optimization routines, support for object-oriented programming, and ability to run in the cloud with our production Java applications."

Borislav Savkovic, BuildingIQ
Apps - Classification Learner app
Language - MATLAB Execution Engine

- Redesigned execution engine runs MATLAB code faster
  - All MATLAB code can be JIT compiled
  - A platform for future improvements

The examples of all FSDA functions have been monitored under R2012a, R2014b, R2015a, R2015b. Results seem to indicate that, generally, on computationally intensive mathematical/statistical function R2015b yields remarkable time improvements.
Why MATLAB?

Analytics that increasingly require both business and engineering data

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

1. Smarter Embedded Systems
   - Develop embedded systems with analytics powered functionality

2. Enable Domain Experts to be Data Scientists
   - Data Scientist

3. Develop analytics to run on both enterprise and embedded platforms
   - Business Systems

4. DATA
   • Engineering, Scientific, and Field
   • Business and Transactional
Smarter Embedded Systems

MATLAB EXPO 2016

AIRBUS
Battery management

GM
Climate control

Festo
Industrial robots

Sonova
Hearing implants

Weinmann
Transport ventilator

manroland
Printing presses

FLIR
Thermal imaging

ABB
Smart Grid controller

Daimler
Cruise controller
MATLAB Code Generation
C Code Generation
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
Where does the processing happen?

<table>
<thead>
<tr>
<th>Devices</th>
<th>Business Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility</td>
<td>Entire System</td>
</tr>
<tr>
<td>Self, Maybe Neighbors</td>
<td>Grid (Unlimited)</td>
</tr>
<tr>
<td>Power</td>
<td>Battery (Low)</td>
</tr>
<tr>
<td>Processing</td>
<td>Microprocessor (Low)</td>
</tr>
<tr>
<td>Memory (Minimum Buffer)</td>
<td>CPU’s (Unlimited)</td>
</tr>
<tr>
<td>Storage</td>
<td>HD’s (Unlimited)</td>
</tr>
<tr>
<td>Updates</td>
<td>Difficult</td>
</tr>
<tr>
<td></td>
<td>Easy</td>
</tr>
</tbody>
</table>
Why not transfer all the data?

1. Data privacy concerns
2. Cost of network/storage
3. Power required to transmit data from device (for wireless)
4. Response time
Splitting computation
Smart Systems: Wearable Healthcare Technology

Wearables and Apps Are Shaping the Future of Medical Care

Motion analysis sensors and apps can precisely capture, measure, or record users’ movements. These devices provide doctors, researchers, and patients with biometric data that would otherwise be difficult to collect, such as minute changes in a patient’s breathing pattern, gait, and range of motion.

However, the proliferation of these devices poses a new problem: How do you analyze data collected with wearable technology and ensure that it is being put to good use?

MATLAB EXPO 2016
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

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

Data Scientist

Business Systems

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

MATLAB EXPO 2016
Learn More

Data Analytics

Download a Free Data Analytics Trial of MATLAB

Engineering and IT teams are using MATLAB to build today’s advanced Big Data Analytics systems ranging from predictive maintenance and telematics to advanced driver assistance systems and sensor analytics. Teams select MATLAB because it offers essential capabilities not found in business intelligence systems or open source languages.

Physical-world data: MATLAB has native support for sensor, image, video, telemetry, binary, and other real-time formats. Explore this data using MATLAB MapReduce functionality for Hadoop, and by connecting interfaces to ODBC/JDBC databases.

Machine learning, neural networks, statistics, and beyond: MATLAB offers a full set of statistics and machine learning functionality, plus advanced methods such as nonlinear optimization, system identification, and thousands of prebuilt algorithms for image and video processing, financial modeling, control system design.

High speed processing of large data sets. MATLAB’s numeric routines scale directly to parallel processing on clusters and cloud.

The Netflix Prize and Production Machine Learning Systems: An Insider Look

Download the white paper

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

— Dr. G. Subrahmanya VRK Rao
Congkent