Industrial Uses Cases of Big Data Analytics

Hideo Watanabe
Analytics & Optimization,
IBM Research - Tokyo
Big Data
Unifying Relevant Technologies Make Big Data Analysis Used in Real Issues

Age of AI

Increasing Data

- Traditional expert systems haven’t succeeded enough.

- Rule-based Reasoning
- Expert System
- Neural Networks
- Natural Language, Image, Speech Processing
- Operations Research (Optimization Theory)
- Statistics

Increasing Computing Power

- Web search and recommendation systems based on machine learning techniques are used in real.
- Machine learning technologies are mandatory in the areas of fraud detection of credit cards and credit decisions in banking.
- Machine learning is being tried at the statistical process controls.

- Machine Learning
- Agent Simulation
- Data Mining
- Text Mining
Research Areas in Analytics Team of IBM Research - Tokyo

Machine Learning
Estimate rules and models behind the big data

Optimization
Formulate an issue and solve it effectively

(Agent) Simulation
Predict and analyze issues based on the re-created virtual world

Cognitive Science
Human behavior modeling based on cognitive and psychological studies
Sensor Data
Examples of Sensor Data Analysis

**Battery Degradation Analysis**
- Predicting EV battery internal state changes non-destructively
- Higher precision than the state-of-the-art method

**Trajectory Analysis**
- Analyze trajectory of cars, humans, etc.
- Extracting valuable information for e.g. marketing

**Near-Miss Analysis**
- Classifying near-miss situation from drive-recorder data
- Developed a method to create feature vectors representing best the situation from time-series data.

**Anomaly Detection**
- Detecting anomalies from time-varying multiple sensor data
- A new approach which does not use thresholds set by human
Anomaly Detection: Manual threshold setting is very difficult to consider all dependencies among sensors

- **Accuracy of threshold-based method is poor.**
  - Many false alerts and miss alerts

- **In various dynamic systems, threshold is changing dynamically.**
  - i.e. high temp. under high pressure is not anomaly, but high temp. under low pressure is anomaly.
IBM Anomaly Analyzer for Correlational Data (AACD)

We have developed a unique approach for dependency-based anomaly detection for multivariate sensor data.

1. Sensor correlation graph structure was calculated from the past sensor data collected.

2. A current sensor-correlation graph is created from the real-time data.

3. Compare these two graphs.

- AACD learns hidden dependencies among the variables (sensors).
- The present dependencies are compared with the normal situation to compute the anomaly score of each variable.
### Advantages of AACD for Anomaly Detection

AACD increases maintenance efficiency and quality.

<table>
<thead>
<tr>
<th>Usual Approach</th>
<th>IBM Approach (AACD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Handcrafted rule-based approach</strong></td>
<td><strong>Data-driven approach using Big Data</strong></td>
</tr>
<tr>
<td>□ Growing shortage of skilled engineers</td>
<td>□ Increase maintenance efficiency and quality with high accurate anomaly detector</td>
</tr>
<tr>
<td>□ Many black-box electronic devices</td>
<td>□ enable whole devices’ anomaly detection with data-driven approach</td>
</tr>
<tr>
<td>□ Very high cost to create diagnosis rules</td>
<td>□ provide new knowledge to experts by statistical analysis</td>
</tr>
<tr>
<td>□ Limited accuracy due to unpredictable external disturbances (weather, water depth, see current, etc.)</td>
<td></td>
</tr>
</tbody>
</table>
Use Cases of Sensor Data Analytics

There are many use cases in various industries such as vessel, transportation, energy and vehicle.

(Note that the following examples include some of general sensor data analytics not using AACD.)

**Mistake Monitoring of Vehicle Plant Workers**
- Utilizing geomagnetic sensors attached to workers
- We have good accuracy for the real data.

**Remote Monitoring of Vessels**
- Automatic anomaly detection of vessel engines
- No manual threshold is required.

**Train Status Monitoring**
- Anomaly detection by monitoring temperature of wheel axels
- Better results than an existing method

**Quality Monitoring of Vehicle Production Line**
- A spark spectrum can decide the quality of the welding process.
- Almost 100% accuracy

**Sensor Monitoring**
- Checking the sensor data correctness of prototype cars
- Early detection of sensor anomalies led to the workload reduction.

**Power Generator Status Monitoring**
- Monitoring more than 100 sensors of power generator
- Succeeded early detection of anomalies which humans cannot
Anomaly Detection of Vessel Engines


<table>
<thead>
<tr>
<th>Business Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>To deliver vessel maintenance and management cloud system as the foundation of condition-based monitoring for the maritime industries.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technical Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>To compute anomaly score of vessel engines to realize a remote vessel engine maintenance system.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technical Challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vessel engine sensor data are largely varying on the sea condition such as weather.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Successfully created anomaly detection models and implemented into Diesel United’s Lifecycle Administrator (LC-A) system, a condition-based maintenance system for diesel engines.</td>
</tr>
</tbody>
</table>
Summary – Decision Making Supported by Big Data

Mathematical and scientific approaches are being applied to real problems in the big data analytics area.

These approaches realized mathematically-sound analysis and predictions based on the past data.

It is a time to change from an intuitive decision making to a decision making endorsed by the mathematical analytics.

THANK YOU