

Workshop on Data Science and Industrial IoT

The logo for H2O.ai, featuring the text "H₂O.ai" in a bold, black, sans-serif font. The "H₂O" is in a larger font size than ".ai". The logo is set against a solid yellow square background.

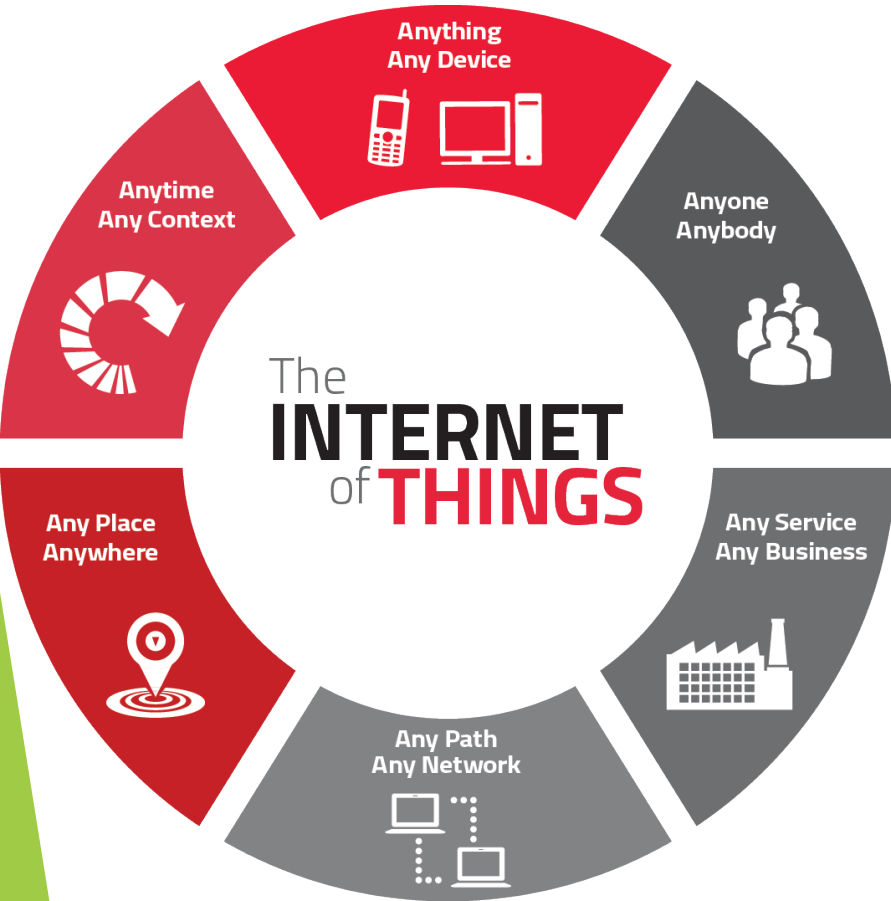
H₂O.ai

Sibanjan Das

Ajit Jaokar

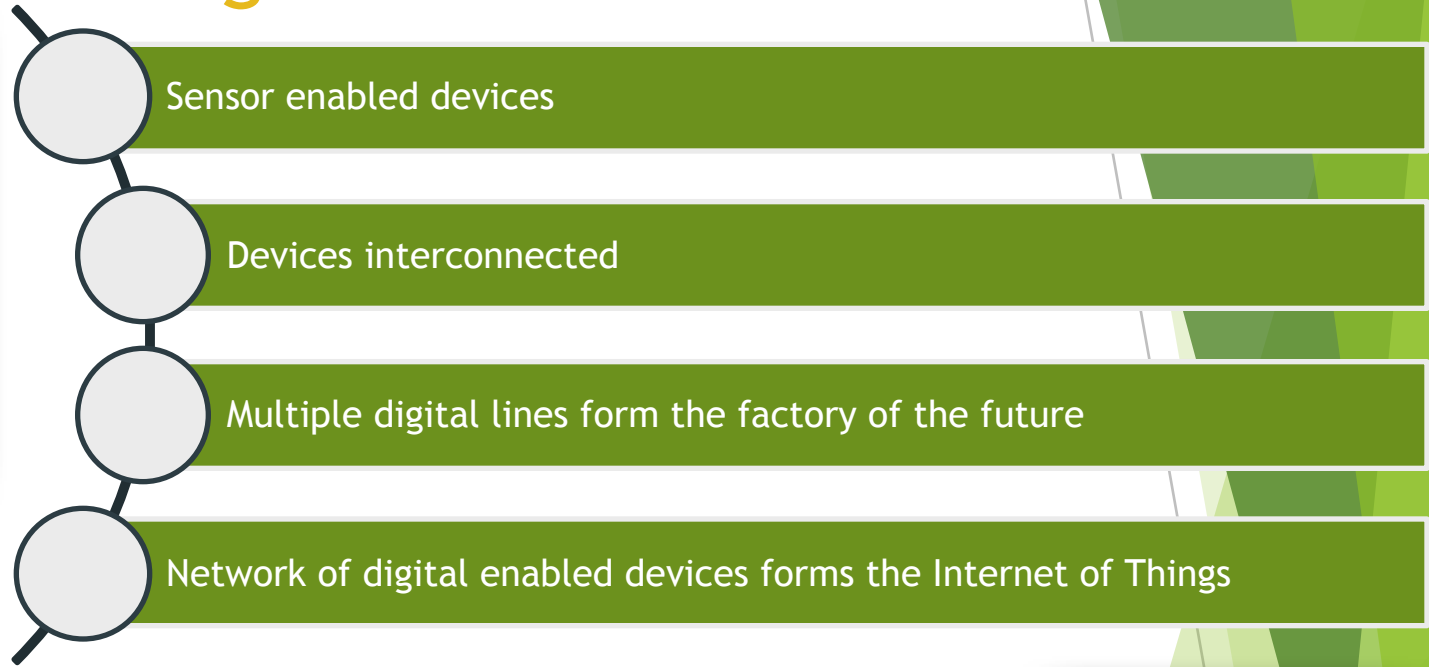
Jo-Fai Chow, H2O.ai

Internet of Things



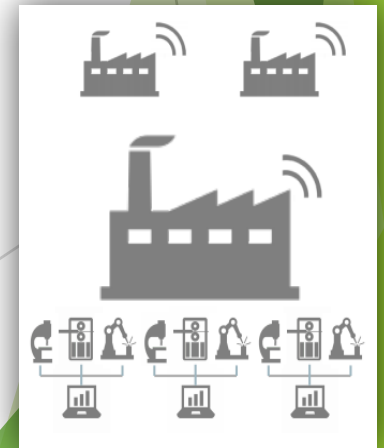
- The Internet of Things (IoT) is a futuristic technology trend
- It is the first step in becoming a truly digital business.
- By connecting people, assets, devices and systems, IoT delivers an Omni channel experience between the physical and digital world.

Future of Manufacturing and Industrial Internet of Things



The Biggest Challenge.... Data Explosion!!!!

Manufacturers are looking for ways to leverage value from this data



Manufacturing IoT: Big Data

Manufacturers are deploying instruments such as sensors and controllers or even smart, networked cameras or RFID readers to measure a wide range of operational processes



1 Storage data by sector derived from IDC.

As the amount of data used by business grows, there are new opportunities for analyzing it, which stands to change how we make day-to-day business decisions.

Contribution of Data Scientists in IIoT

As Michael Rothschild discusses manufacturers seeking to extract maximum value from investing in big data projects should work backwards and ask a few fundamental questions:

- ✓ What business processes or decisions do you want to improve?
- ✓ How will these decisions improve the business?
 - Customer profitability?
 - Product rationalization?
 - Capacity planning?
- ✓ What are you trying to maximize?
 - Profits? Asset utilization? ROI? Revenues?
- ✓ What are the most meaningful metrics to measure progress toward those goals?
 - Unit margins? Profit-per-hour of machine time?
- ✓ What types of analysis do you need to perform to expose the data,
 - Explore 'what if' scenarios and iterate through alternatives to maximize profitability?
- ✓ And then, finally, what types of data do you need to collect in
 - Order to feed the above analysis and decision-making?

Contribution of Data Scientists in IIoT

- ▶ Predictive Maintenance Models
- ▶ Anomaly Detection
- ▶ Production Forecast
- ▶ Time Series Analysis
- ▶ 3D Printing





Predictive Maintenance

Reactive vs Proactive Maintenance



► Reactive Maintenance:

- Reactive approach; dealing with breakdowns or problems when they occur

► Proactive(Predictive) Maintenance

- Proactive approach
- An attempt to determine when best to perform preventive maintenance activities. Example: Reducing breakdowns through a program of lubrication, adjustment, cleaning, inspection, and replacement of worn parts

Advantages of Predictive Maintenance

- ▶ Reduces fuel/energy costs
- ▶ Reduces parts and labor
- ▶ Increased machine availability
- ▶ Increased throughput
- ▶ Fewer catastrophic failures
- ▶ Improved safety
- ▶ Optimizes manpower deployment





Workshop: Predictive Maintenance and Anomaly Detection using H2O.ai

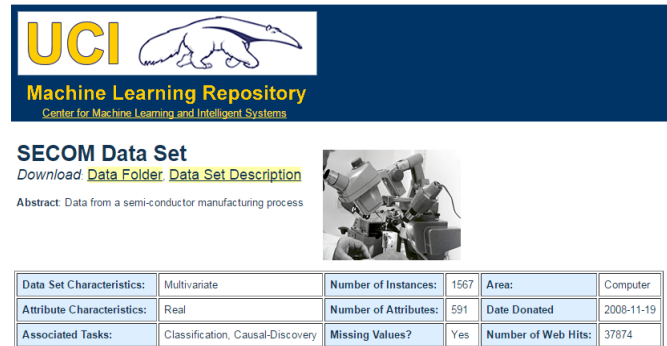
A Data Scientist's H2O.ai Toolbox



Some of the Smart
Applications using H2O.ai

Predictive Maintenance Use Case Data Set:

- SECOM dataset from UCI Machine Learning Repository



The screenshot shows the UCI Machine Learning Repository interface. At the top is the UCI logo and the text 'Machine Learning Repository Center for Machine Learning and Intelligent Systems'. Below this is the 'SECOM Data Set' title, followed by links for 'Download', 'Data Folder', and 'Data Set Description'. An abstract describes the data as 'Data from a semi-conductor manufacturing process' next to an image of a robotic arm. At the bottom is a table with data set characteristics.

Data Set Characteristics:	Multivariate	Number of Instances:	1567	Area:	Computer
Attribute Characteristics:	Real	Number of Attributes:	591	Date Donated	2008-11-19
Associated Tasks:	Classification, Causal-Discovery	Missing Values?	Yes	Number of Web Hits:	37874

- It consists of 1567 records taken from a wafer fabrication production line. Each record is a vector of 590 sensor measurements plus a label of pass/fail test.
- Among the 1567 records, there are 104 fail cases which are labeled as 1 and the rest are labeled as negative 0.

Anomaly Detection

- ▶ MTCARS Dataset
- ▶ The data was extracted from the 1974 *Motor Trend* US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973-74 models).



Results - Anomalies in Automobile Design

