Patterns in Time Series Data Are the New Gold Nuggets in the Digital Economy

TRENDALYZE

Companies have been walking along the digital transformation path for the past few years. COVID-19 forsed all of them to sprint.

Remote

Contactless

Reliable

The new powerlines that connect and power operations are data pipes.

!

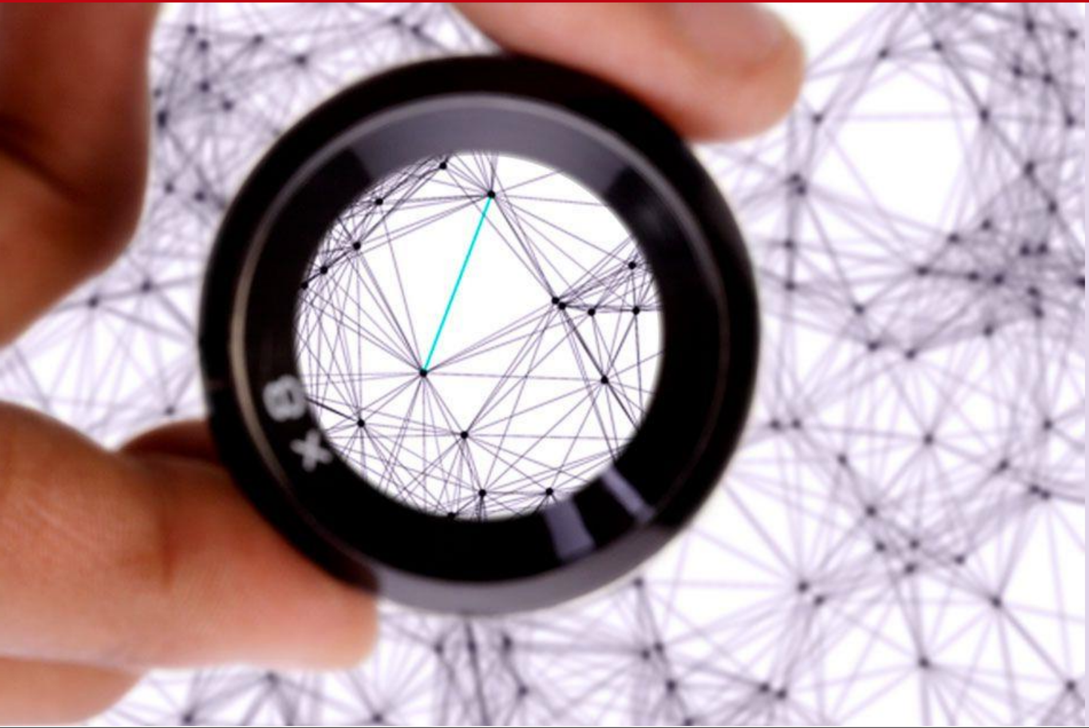
Not the data that you know and you are used to analyze.

Sensors

Data Pipes

People

Introducing granular time series data

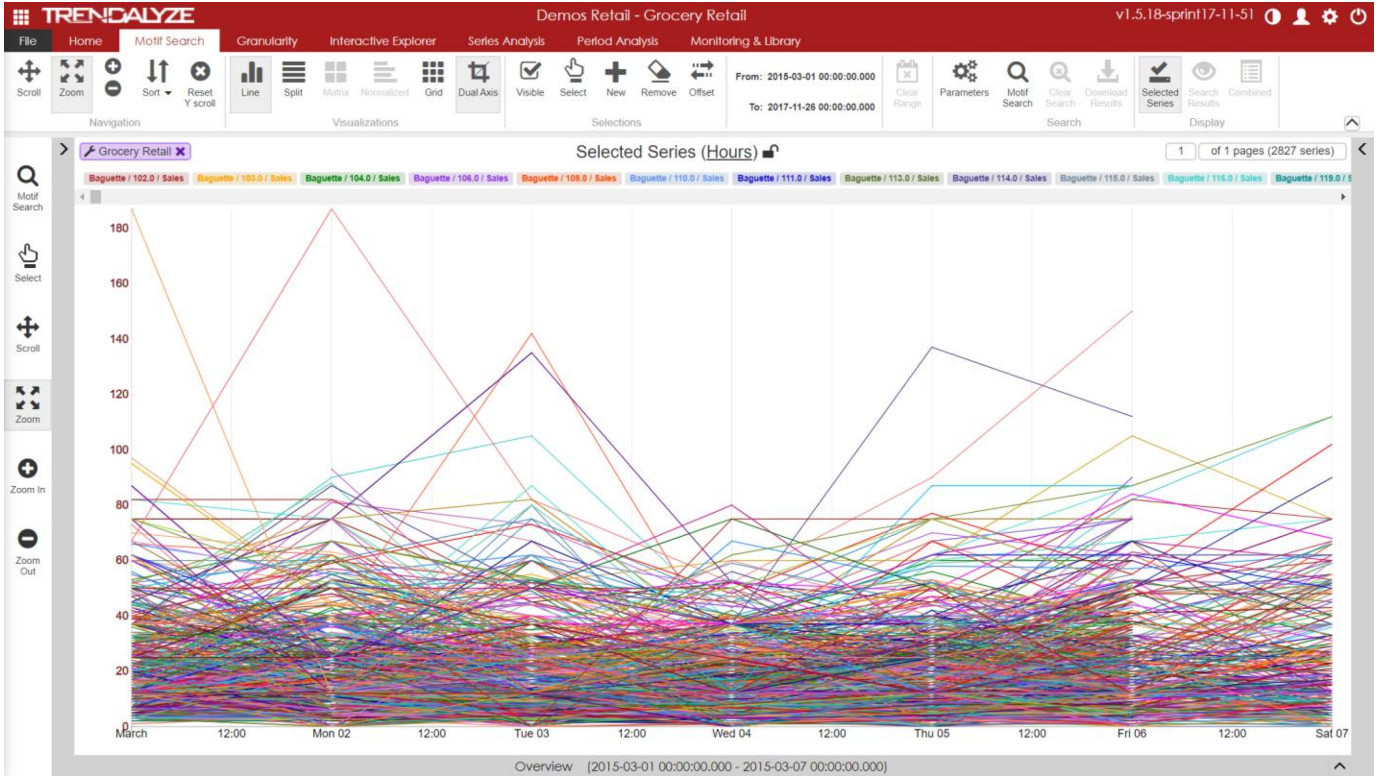


- Granularity – this is the sampling rate, i.e., the rate at which measurements are taken. Typically minutes, seconds, milliseconds, nanoseconds.

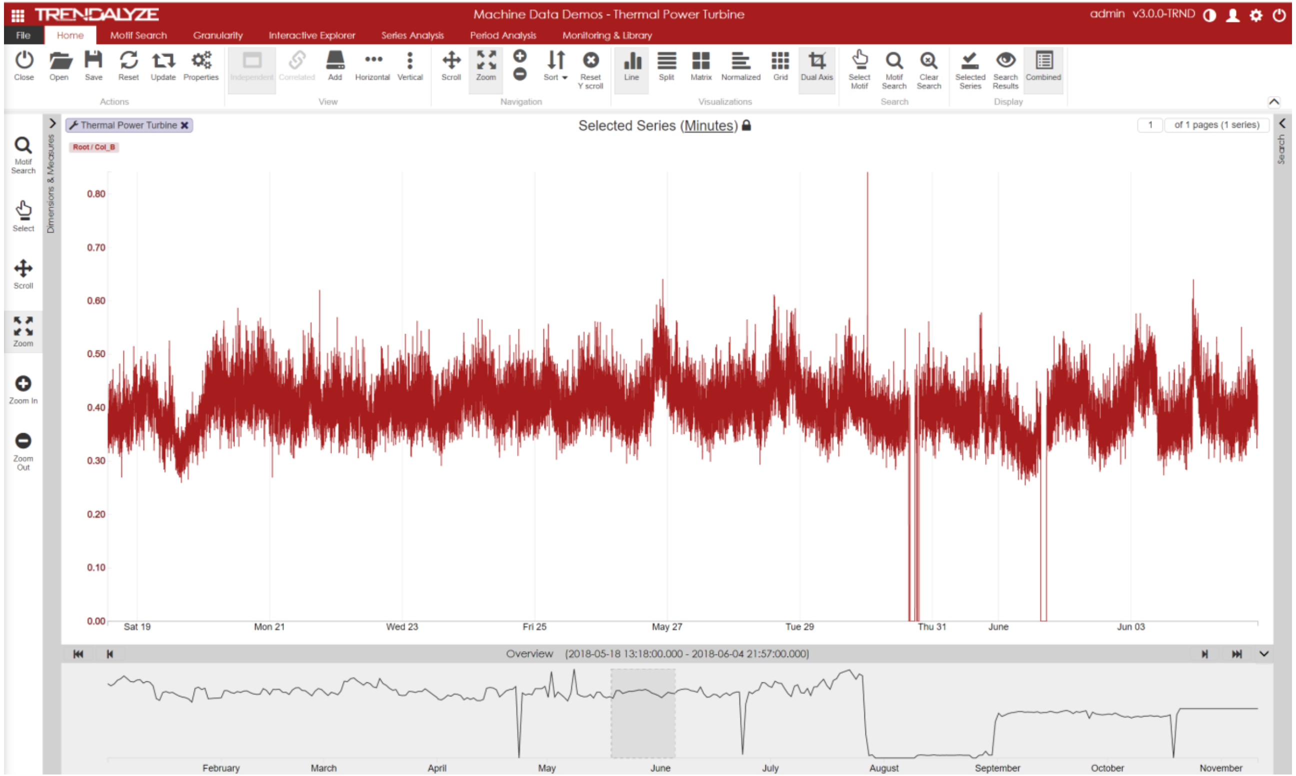
- Dimensionality – the descriptors and metadata for the individual unit of analysis, i.e., the SKU within a particular retail store, or the sensor at a particular car engine, etc.

The Time Series Data That We All Know

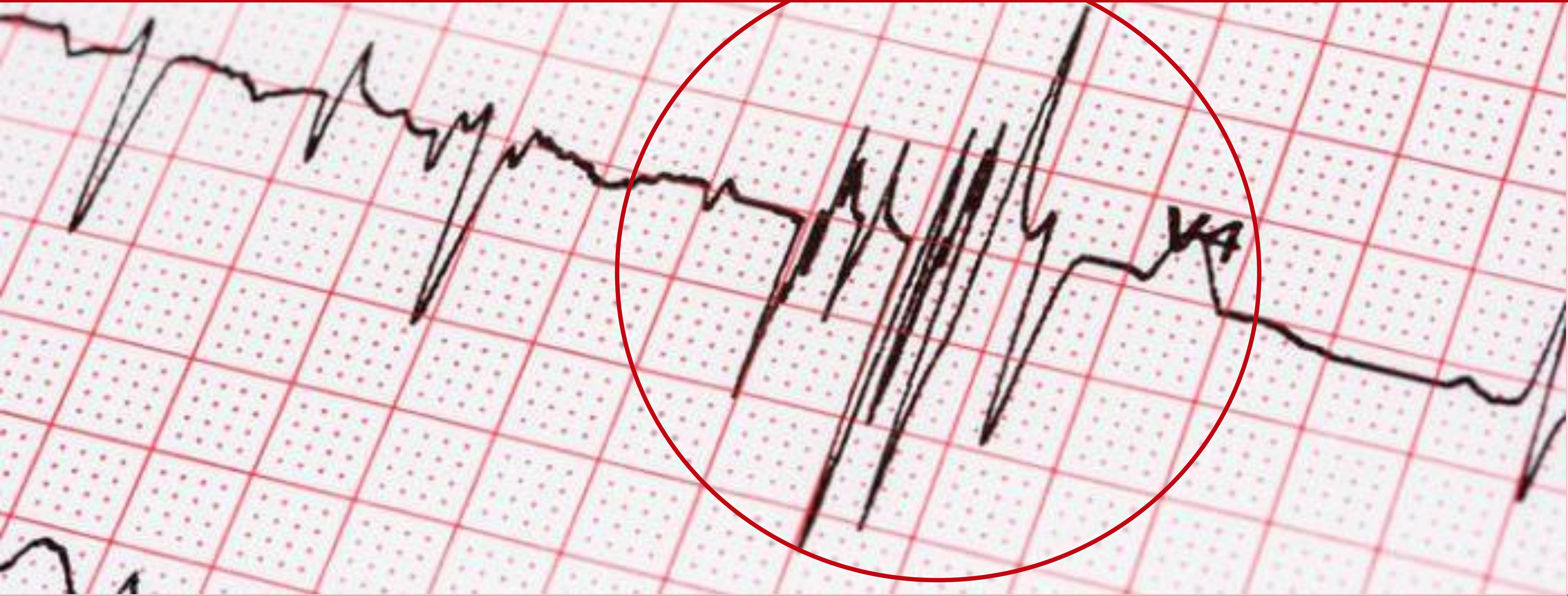
Highly Dimensional Time Series Data



Highly Frequency Time Series Data

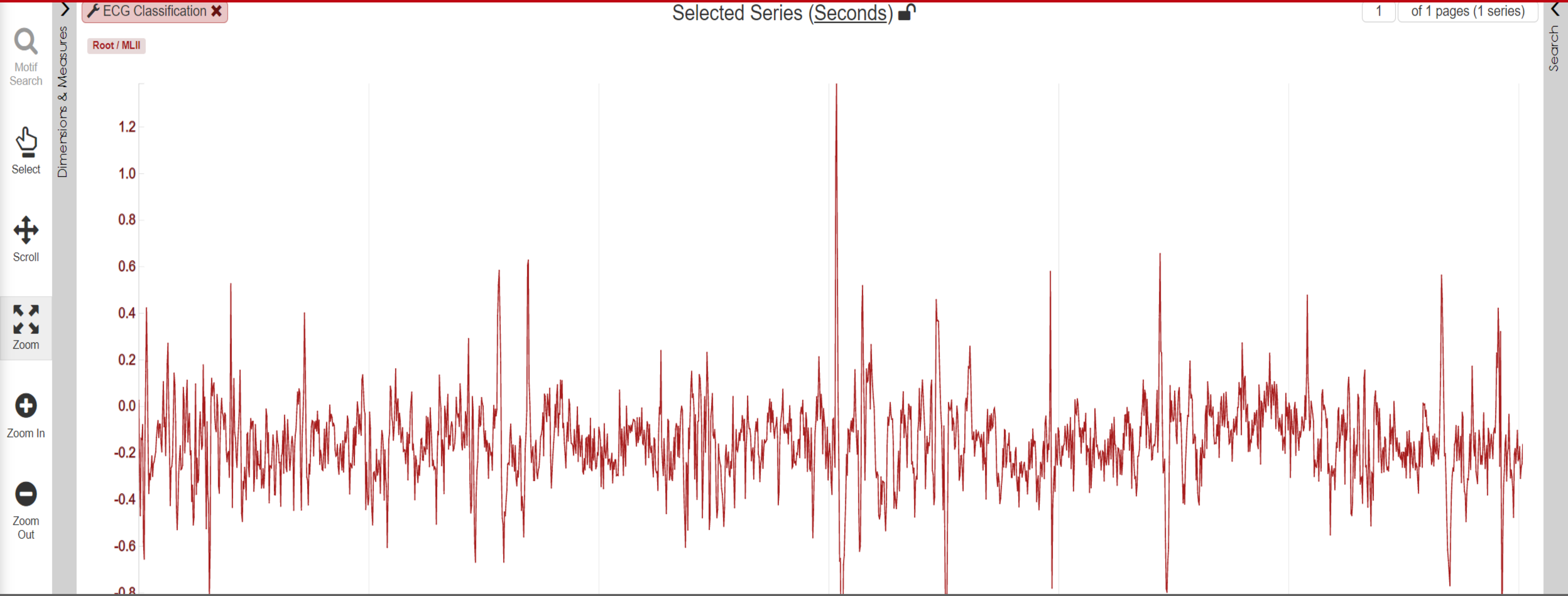


- Sensors and devices collect the “heartbeats” of everything
- Knowing the good and the bad “heartbeats” is gold
- By monitoring the “heartbeats” you can save or make money

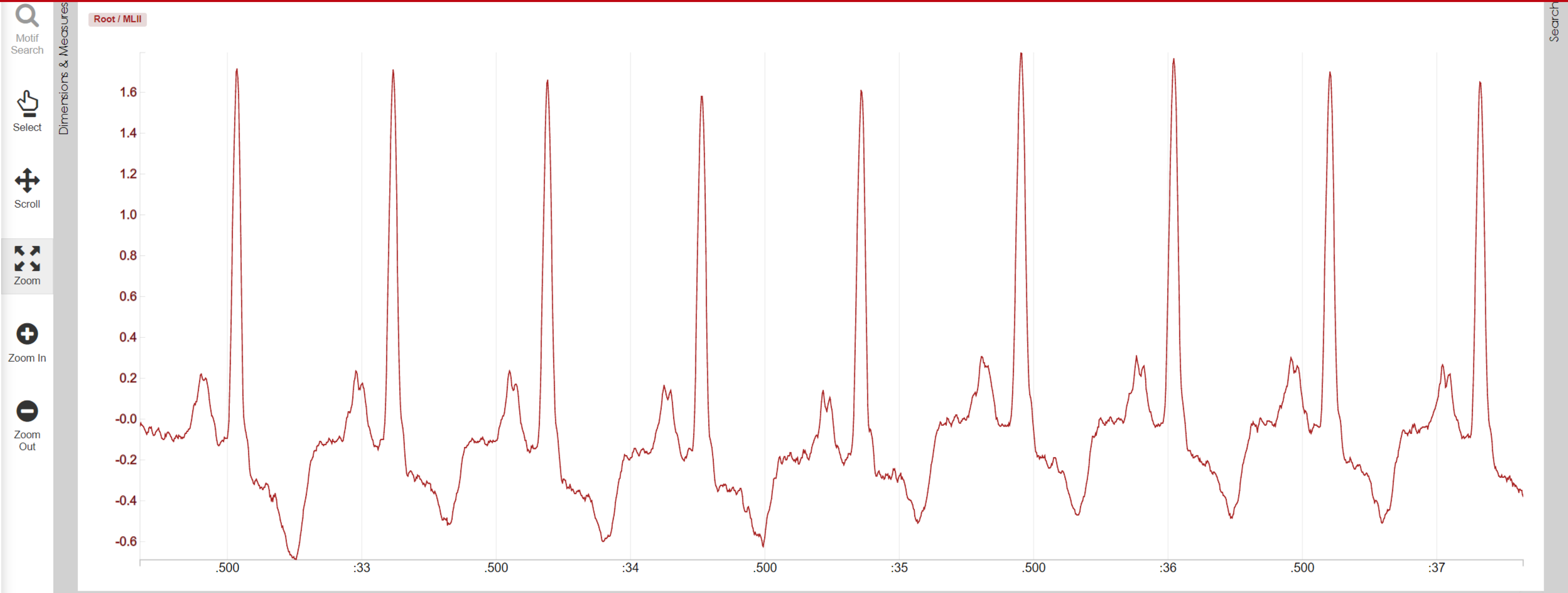


Why is granular time series data so important:

- What is this time series?
- You cannot recognize it.

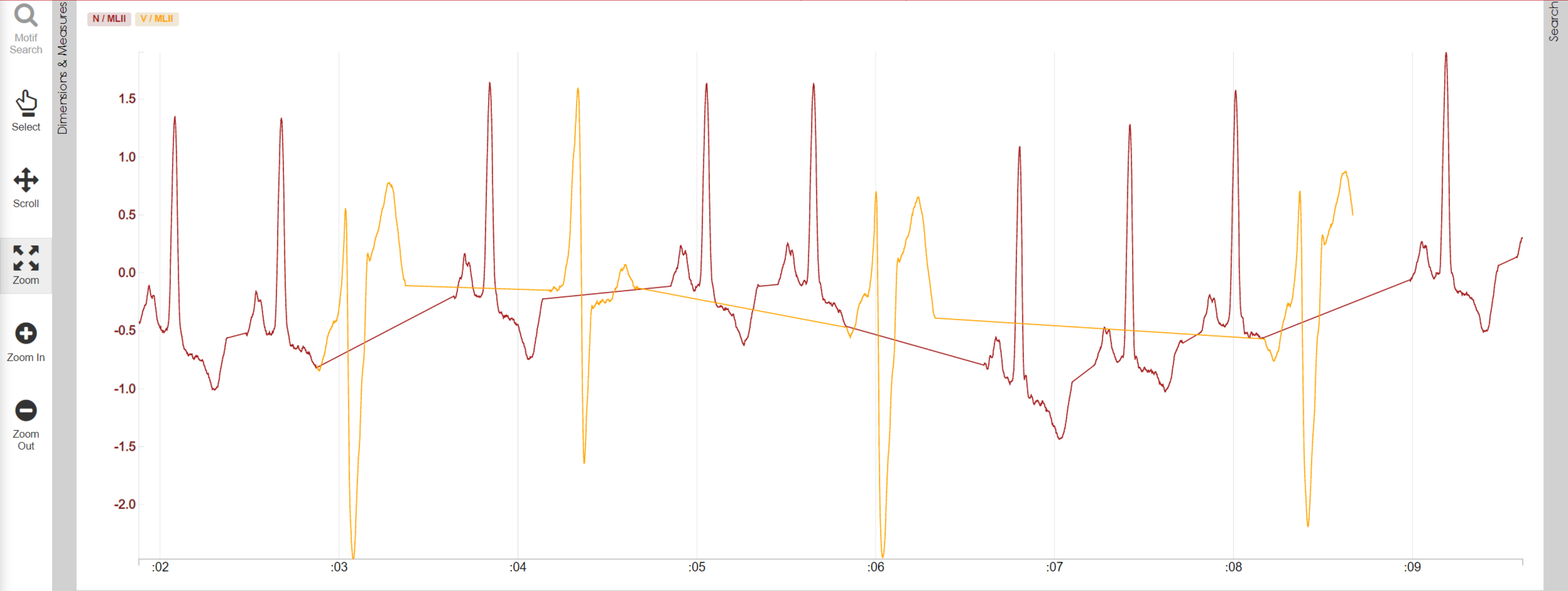


- This one you recognize instantly
- It is “heartbeats” from ECG
- The more detailed data reveals the shape

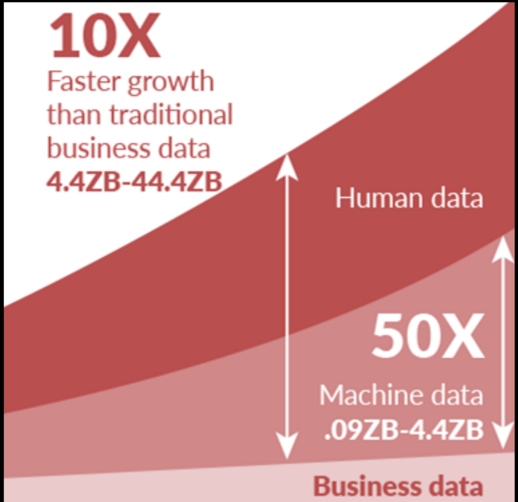


What are the golden nuggets in such data?

- The different "heartbeats" shapes
- Monitoring "heartbeat" shapes saves lives



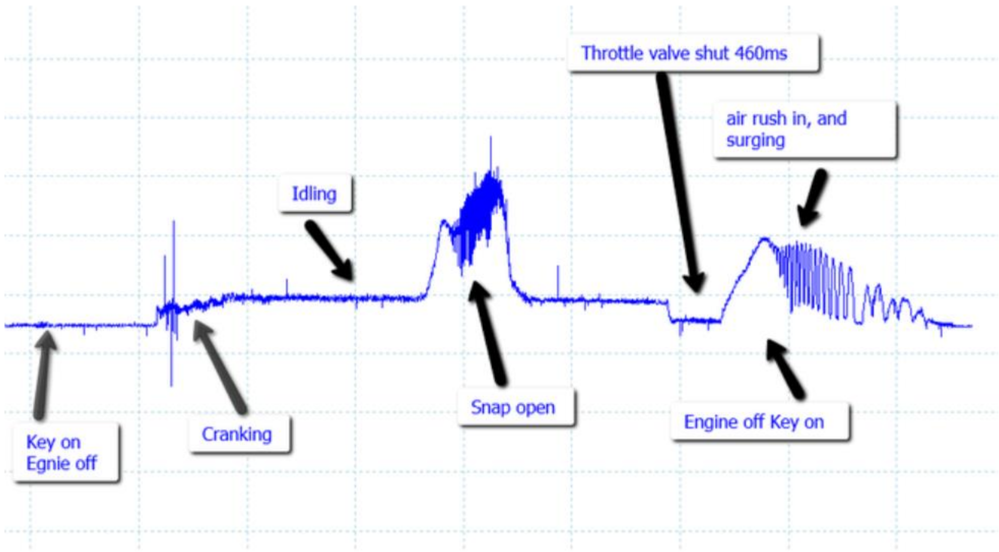
Today Every Industry Generates Loads to Granular Time Series Data

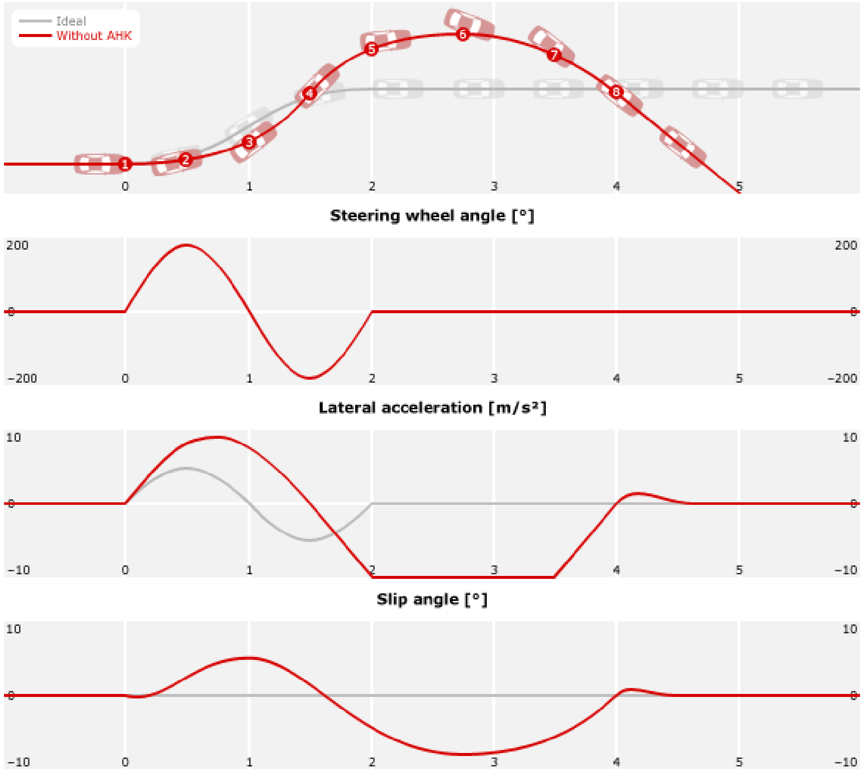


$16
Billion
Business Intelligence

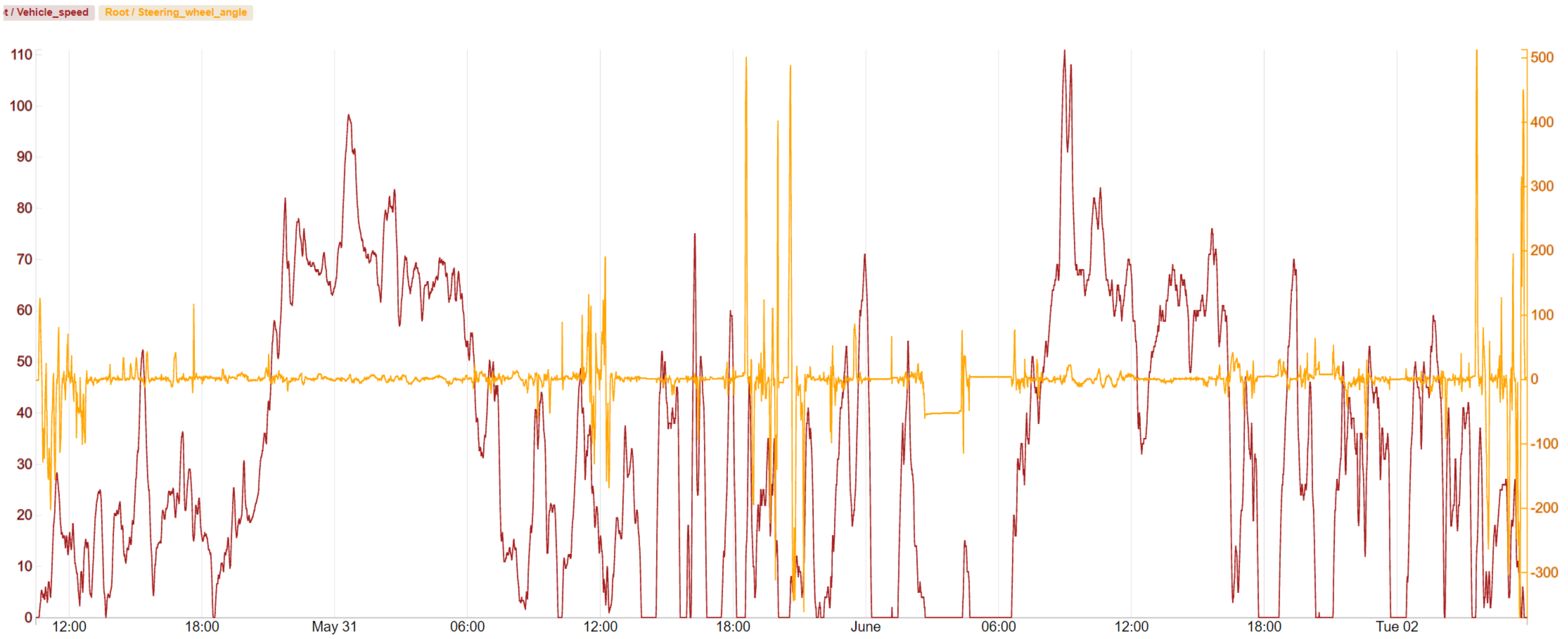
Time Series Data & Analytics Is Growing Fast

Engine Diagnostics & Fleet Management Examples





Driver Fingerprinting & Forensics Examples



Covid 19 Examples

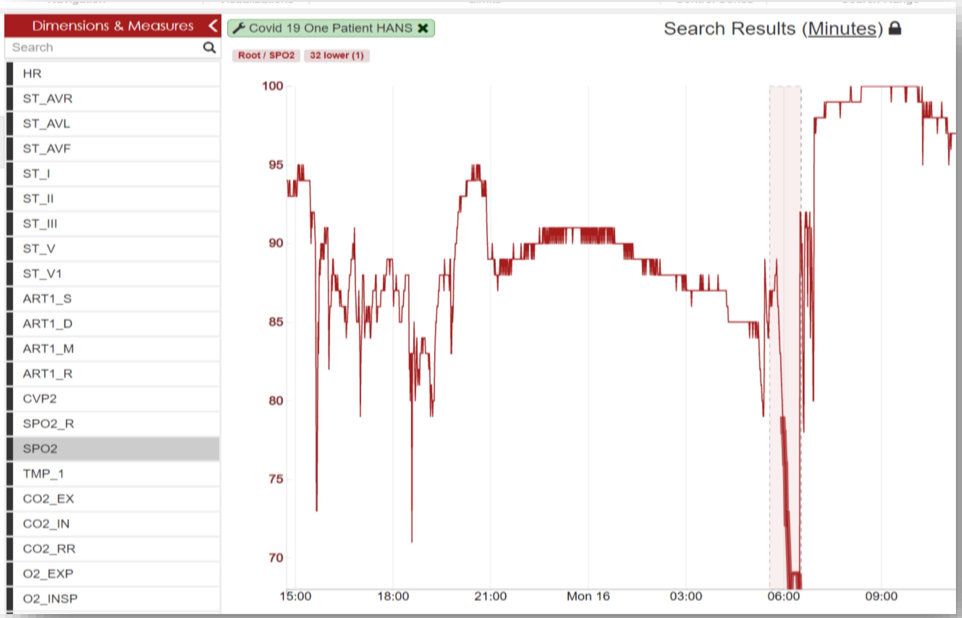
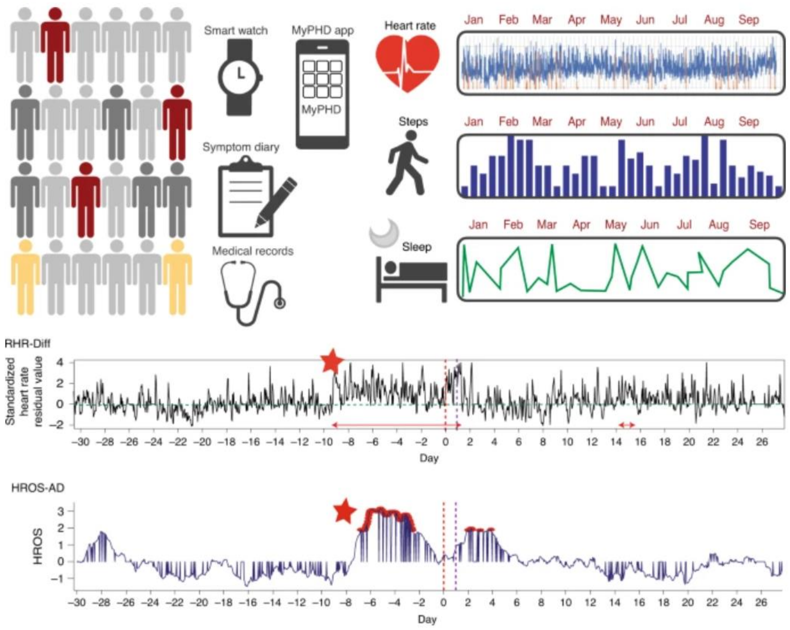


Fig. 1: Overview of the study design, cohort and data.



https://www.nature.com/articles/s41551-020-00640-6

Manufacturing Industry Examples

Quality Analysis Using Motif Discovery from Small Data

Amanjeet Singh Bhatia · Lianhua Chi · Rado Korotov

Abstract Modern manufacturing demands high standards of quality. New automated quality control solutions monitor quality using image classification with Deep Learning (DL). DL requires large data samples, expensive computational infrastructure, and time-consuming model training. As a result, many manufacturers cannot deploy DL-based quality control. The problem is worse in small and medium enterprises (SME) that lack the required resources and are thus put at a competitive disadvantage. In this paper we propose a new time series motif discovery method that could effectively learns from significantly smaller datasets compared to DL. The new approach uses one image sample as the "golden quality standard" and converts it into a time series motif. Then we apply motif-based similarity search to compare it with the converted time series of product images coming out of the production line. A distance score from the golden standard determines the quality status of the product item. Meanwhile, to enhance the performance, rather than comparing the entire image we compare separately the known segments in the image. We use image decomposition to extract core shapes and areas. By decomposing the image and performing multiple comparisons, we could achieve higher accuracy for each similarity search. The advantage of the present method is that it can be configured entirely by the engineers who are not trained in data science. Furthermore, it can be changed quickly when the production outputs change. In this paper we conduct a three-way comparison between DL, motif

Fig. 1 A casted plate labeled "ok".

Fig. 2 A casted plate labeled "def".

Fig. 11 A visualization of the converted to time series "Ok" images [A] and "Def" images [B] showing clear differences in shapes

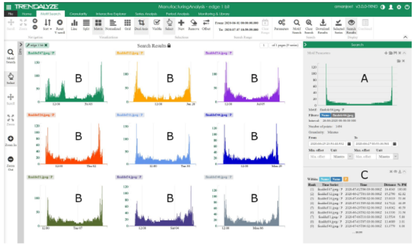
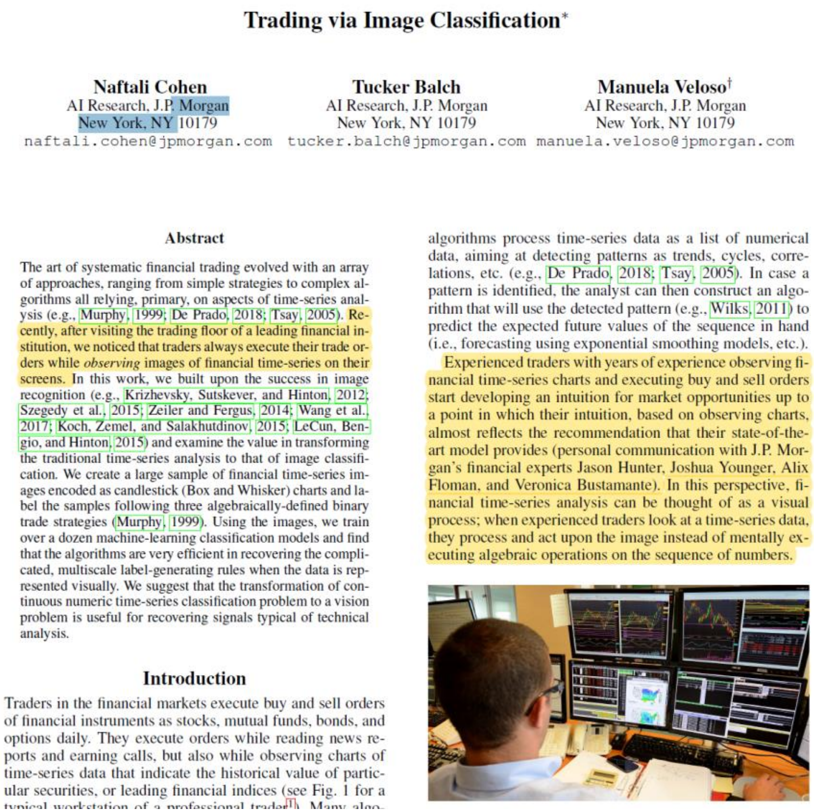


Fig. 12 A reference motif of an "Ok" image [A] used as a reference to search for most different patterns from it, matching correctly the "Def" images [B] and showing the difference in similarity [C].



Trading via Image Classification*

Naftali Cohen
AI Research, J.P. Morgan
New York, NY 10179
naftali.cohen@jpmorgan.com

Tucker Balch
AI Research, J.P. Morgan
New York, NY 10179
tucker.balch@jpmorgan.com

Manuela Veloso†
AI Research, J.P. Morgan
New York, NY 10179
manuela.veloso@jpmorgan.com

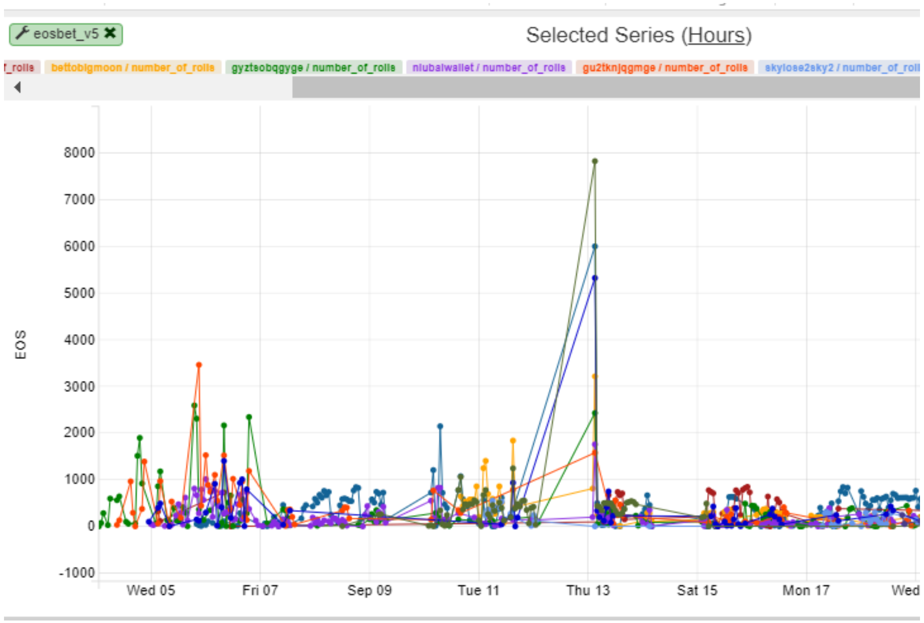
Abstract

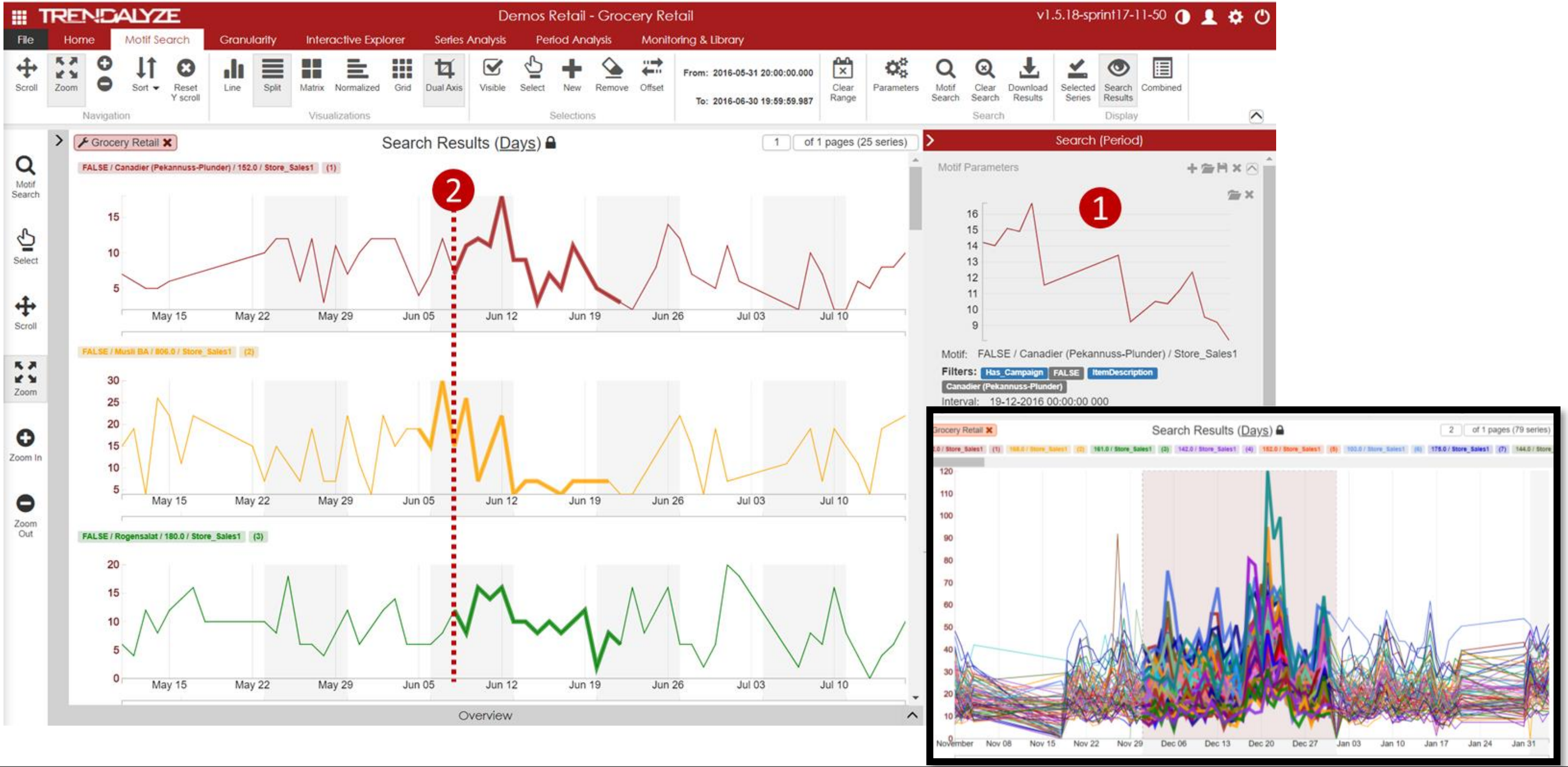
The art of systematic financial trading evolved with an array of approaches, ranging from simple strategies to complex algorithms all relying, primary, on aspects of time-series analysis (e.g., Murphy, 1999; De Prado, 2018; Tsay, 2005). Recently, after visiting the trading floor of a leading financial institution, we noticed that traders always execute their trade orders while *observing* images of financial time-series on their screens. In this work, we built upon the success in image recognition (e.g., Krizhevsky, Sutskever, and Hinton, 2012; Szegedy et al., 2015; Zeiler and Fergus, 2014; Wang et al., 2017; Koch, Zemel, and Salakhutdinov, 2015; LeCun, Bengio, and Hinton, 2015) and examine the value in transforming the traditional time-series analysis to that of image classification. We create a large sample of financial time-series images encoded as candlestick (Box and Whisker) charts and label the samples following three algebraically-defined binary trade strategies (Murphy, 1999). Using the images, we train over a dozen machine-learning classification models and find that the algorithms are very efficient in recovering the complicated, multiscale label-generating rules when the data is represented visually. We suggest that the transformation of continuous numeric time-series classification problem to a vision problem is useful for recovering signals typical of technical analysis.

Introduction

Traders in the financial markets execute buy and sell orders of financial instruments as stocks, mutual funds, bonds, and options daily. They execute orders while reading news reports and earning calls, but also while observing charts of time-series data that indicate the historical value of particular securities, or leading financial indices (see Fig. 1 for a typical workstation of a professional trader[1]). Many algorithms process time-series data as a list of numerical data, aiming at detecting patterns as trends, cycles, correlations, etc. (e.g., De Prado, 2018; Tsay, 2005). In case a pattern is identified, the analyst can then construct an algorithm that will use the detected pattern (e.g., Wilks, 2011) to predict the expected future values of the sequence in hand (i.e., forecasting using exponential smoothing models, etc.).

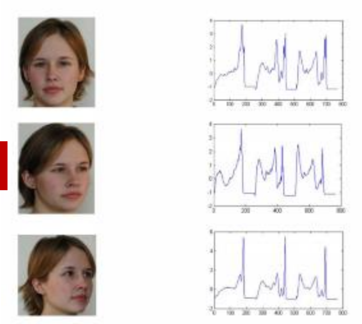
Experienced traders with years of experience observing financial time-series charts and executing buy and sell orders start developing an intuition for market opportunities up to a point in which their intuition, based on observing charts, almost reflects the recommendation that their state-of-the-art model provides (personal communication with J.P. Morgan's financial experts Jason Hunter, Joshua Younger, Alix Floman, and Veronica Bustamante). In this perspective, financial time-series analysis can be thought of as a visual process; when experienced traders look at a time-series data, they process and act upon the image instead of mentally executing algebraic operations on the sequence of numbers.

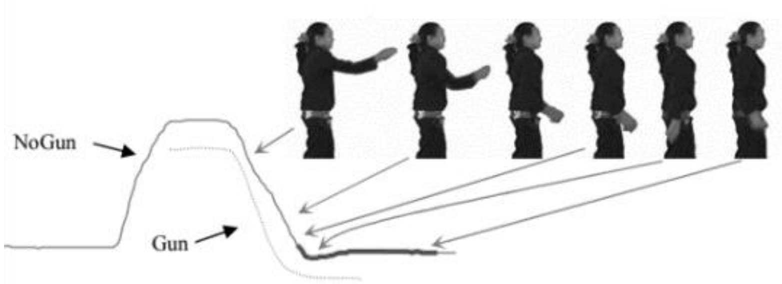



TRENDALYZE
Demos Retail - Grocery Retail
v1.5.18-sprint17-11-50
File
Home
Motif Search
Granularity
Interactive Explorer
Series Analysis
Period Analysis
Monitoring & Library
Scroll
Zoom
Sort
Reset Y scroll
Navigation
Line
Split
Matrix
Normalized
Grid
Dual Axis
Visualizations
Visible
Select
New
Remove
Offset
Selections
From: 2016-05-31 20:00:00.000
To: 2016-06-30 19:59:59.987
Clear Range
Parameters
Motif Search
Clear Search
Download Results
Search
Selected Series
Search Results
Combined
Display
Grocery Retail
Search Results (Days)
of 1 pages (25 series)
FALSE / Canadier (Pekannuss-Plunder) / 152.0 / Store_Sales1
FALSE / Musli BA / 806.0 / Store_Sales1
FALSE / Rogensalat / 180.0 / Store_Sales1
May 15
May 22
May 29
Jun 05
Jun 12
Jun 19
Jun 26
Jul 03
Jul 10
Overview
Zoom In
Zoom Out
2
Search (Period)
Motif Parameters
1
Motif: FALSE / Canadier (Pekannuss-Plunder) / Store_Sales1
Filters: Has_Campaign FALSE ItemDescription Canadier (Pekannuss-Plunder)
Interval: 19-12-2016 00:00:00 000
Search Results (Days)
of 1 pages (79 series)
November
Nov 08
Nov 15
Nov 22
Nov 29
Dec 06
Dec 13
Dec 20
Dec 27
Jan 03
Jan 10
Jan 17
Jan 24
Jan 31

- Nearly all data can be converted to time series data
- Text, images, voice, video
- Shapes can be detected in all of it

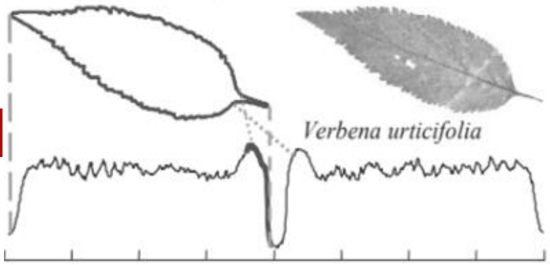
Face Detection

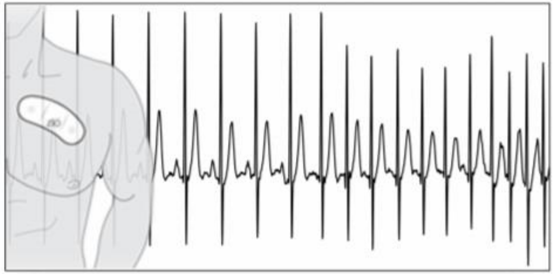




Gesture Detection

Object Detection





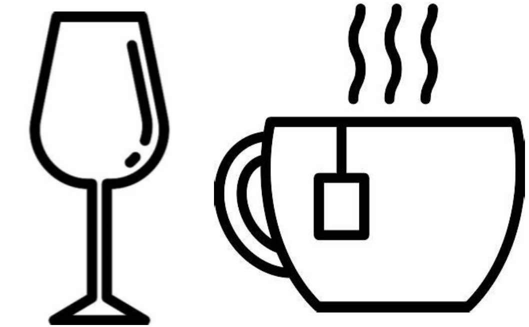
Process Monitoring

Types of Time Series Machine Learning

AI 1.0	AI 2.0	AI 3.0
ARIMA	Deep Learning	Motif Discovery
Statistical Analysis	Statistical Learning	Shape Based Learning

The Fundamental Difference Between Machine & Human Learning

How do humans and machines learn the differences between these two shapes?



Human Learning:

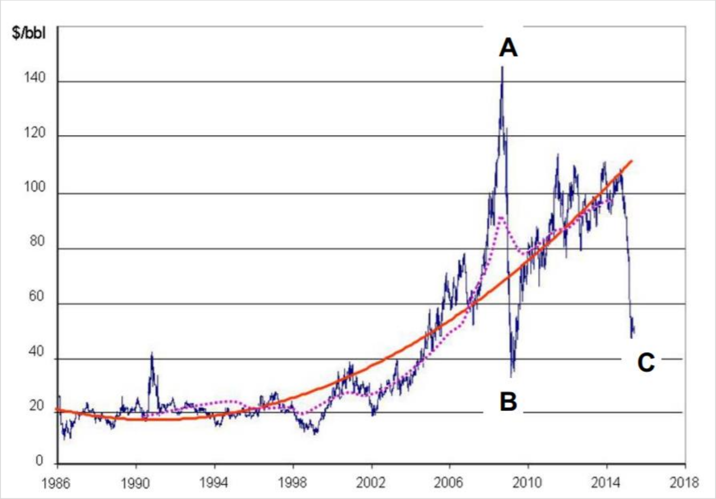
- Recognize shape differences immediately
- Learn from a few examples the differences
- Identify known shapes instantly

Machine Learning:

- Requires thousands of picture to train the model
- All pictures have to be labeled precisely by humans
- Small shape differences confuse machines

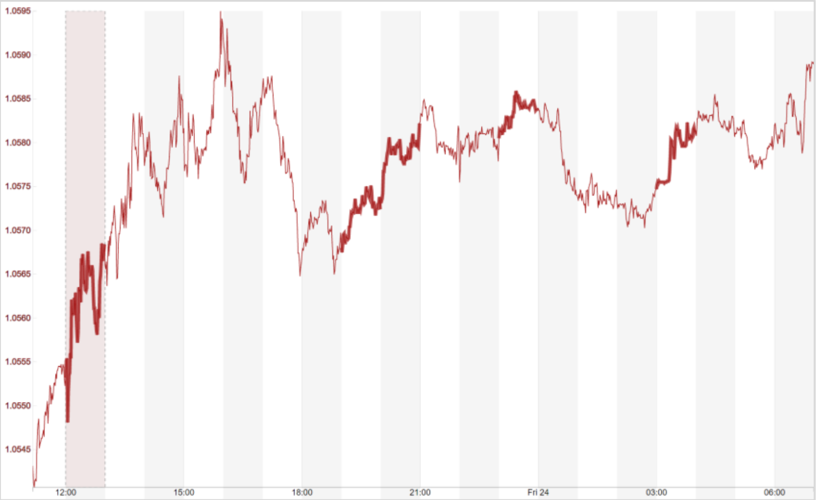
Statistical Modeling vs. Motif Discovery

Statistical Model Fit



Model fit (see red vs blue line) captures the macro trends but omits all micro trends (A, B and C segments). The predictions based on the model will be close to the red line.

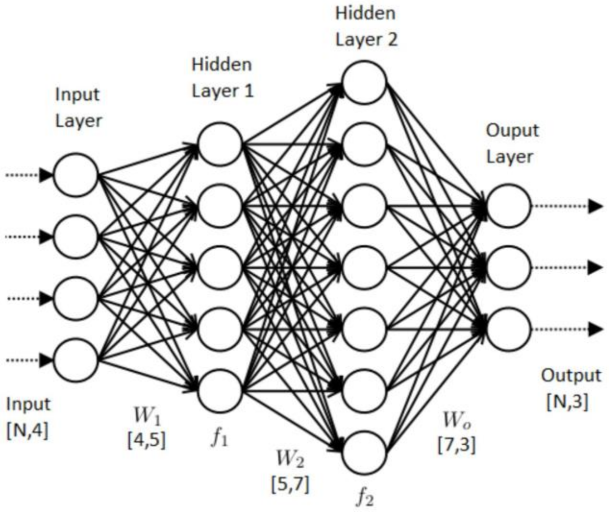
Motif Discovery & Search



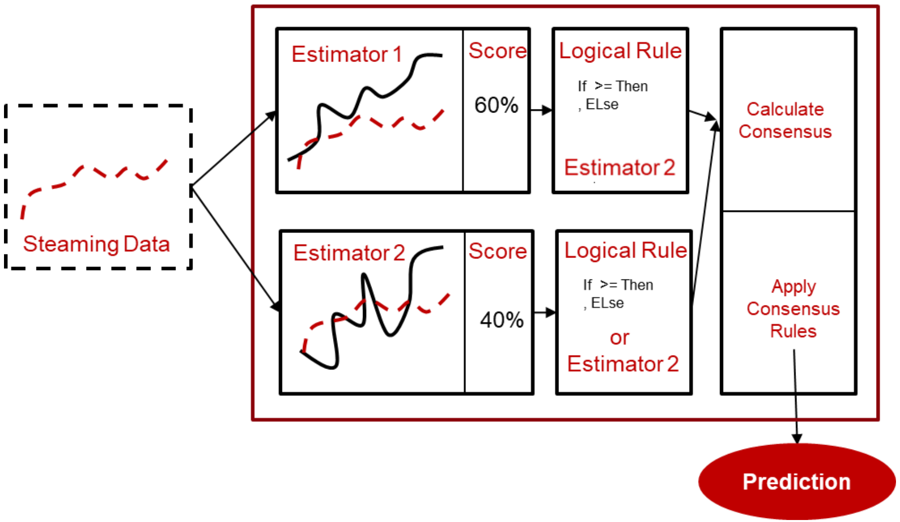
Motif search is like Google search and rank learning algorithms. Motif 1 is used to identify, and rank all matches 2. The process can be used for mining, anomaly detection and real time monitoring.

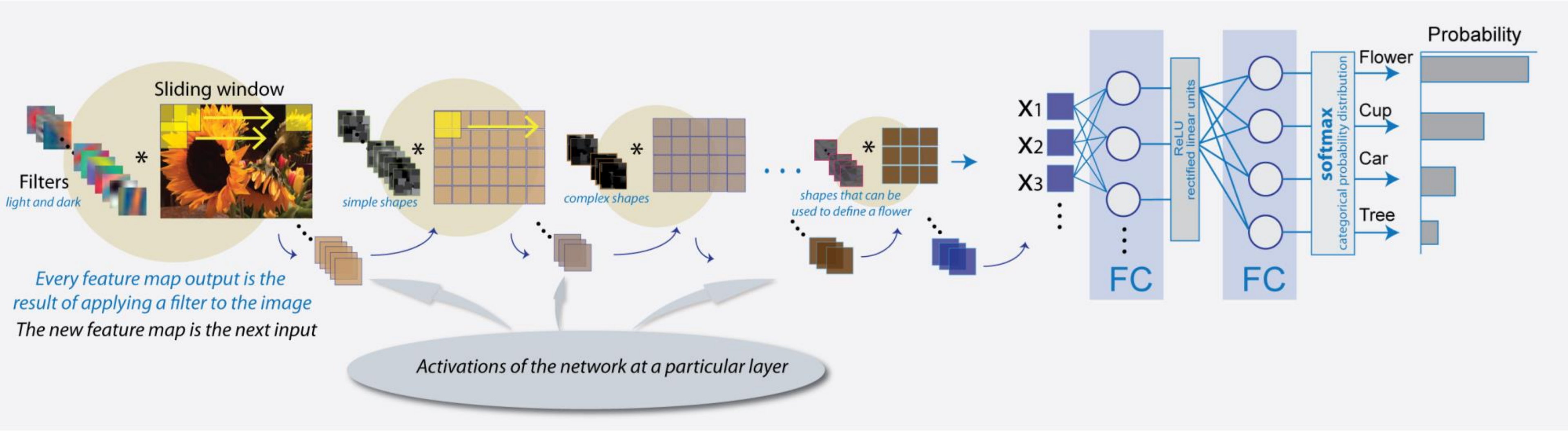
Deep Learning vs. Motif Discovery

Neural Network



Logical Network

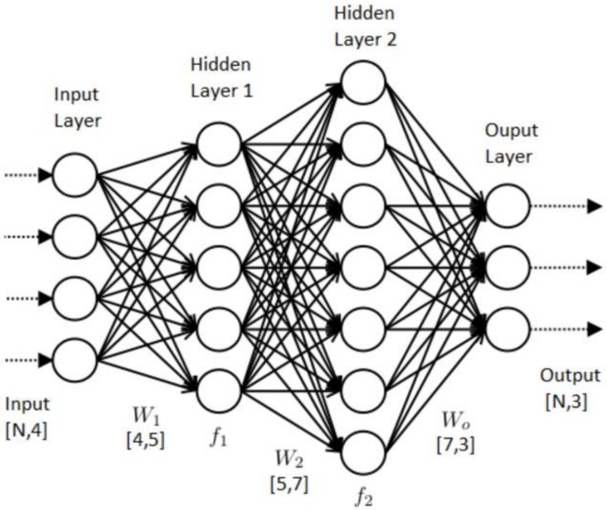



Sliding window
Filters
light and dark
*
simple shapes
complex shapes
shapes that can be used to define a flower
Every feature map output is the result of applying a filter to the image
The new feature map is the next input
Activations of the network at a particular layer
X1
X2
X3
FC
ReLU
rectified linear units
FC
softmax
categorical probability distribution
Probability
Flower
Cup
Car
Tree

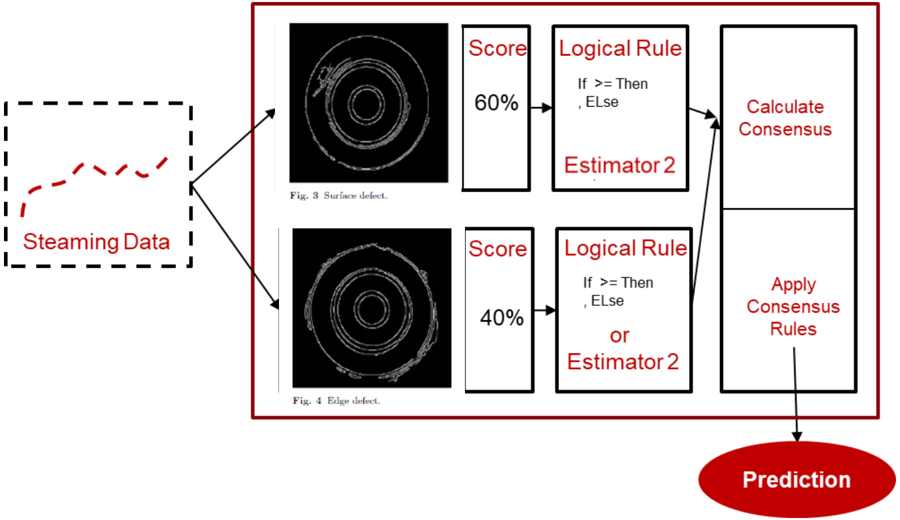
$$\star \qquad \text{kernel} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Deep Learning vs. Motif Discovery

Neural Network



Logical Network



The Choice of a Method Depends on Four Factors

Cost Time Resources Reliability

AI 1.0	AI 2.0	AI 3.0
ARIMA	Deep Learning	Motif Discovery
Statistical Analysis	Statistical Learning	Shape Based Learning

Hype, Magic, and Myths

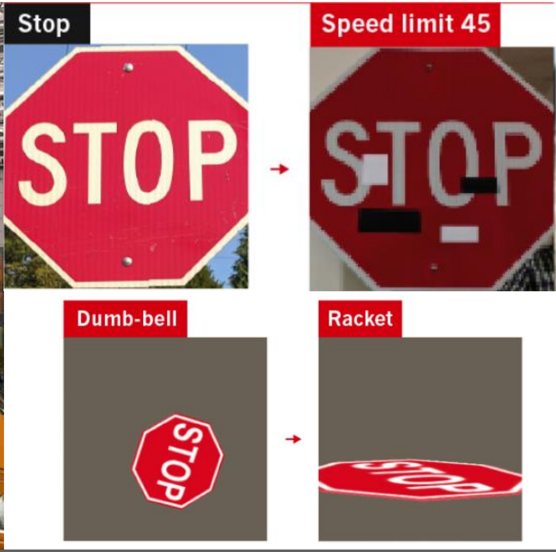
Machines learn on their own

Machines are Faster than humans

Machines are very accurate

Hype, Magic, and Myths

Cost Time Resources Reliability



What You Should Know About Machine Learning

A recent MITSloan survey found that only 10% of companies obtain significant financial benefits and achive ROI with AI. And they are not alone. Gartner has found that 85% of ML projects fail. Worse yet, the research company predicts that this trend will continue through 2022.

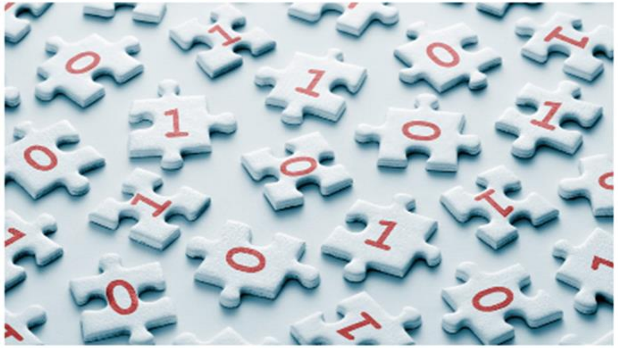
Harvard Business Review

Small Data Can Play a Big Role in AI

by H. James Wilson and Paul R. Daugherty

February 17, 2020

Summary Save Share Print $8.95 Buy Copies



The way we train AI is fundamentally flawed

The process used to build most of the machine-learning models we use today can't tell if they will work in the real world or not—and that's a problem.

MIT Technology Review

They trained 50 versions of an image recognition model on [ImageNet](), a dataset of images of everyday objects. The only difference between training runs were the random values assigned to the neural network at the start. Yet despite all 50 models scoring more or less the same in the training test—suggesting that they were equally accurate—their performance varied wildly in the stress test.

The Risk of Not Doing vs Doing AI Wrong

Sure Profit:

Doing AI right ensures effectiveness and competitiveness

Money Bleed:

Doing it wrong can lead to failures, litigation and other costly mistakes

Risk of Not Doing It

Safe Trials:

Without competitive pressure there is time to do it right

Money Waste:

Your experiments lead to dead ends and only increase future risks

Risk of Doing It Wrong

Decision Styles:	Reactive		Pro-Active		Data-Driven		Trend-Driven	
Methods:	Descriptive	Diagnostic	Forecasting	Predictive Modeling	Machine Learning	Deep Learning	Trend Learning	Pattern Prediction
Sales								
Questions:	How much did we sell?	Which SKUs not selling?	What is the sales forecast?	Who is likely to churn?	What is the right price?	What to recommend?	What are the items' trends?	If we know the rends ......
Techniques:	Report/Dashboard	Ad Hoc/ Discovery	Seasonality Analysis	Logistic Regression	Price Sensitivity	Images Classification	Motif Mining/Search	Motif Prediction
Platforms:	Business Intelligence		Advanced Analytics		AI 2.0 Data Modeling		AI 3.0 Shape Intelligence	

The Top 6 **Time Series Intelligence Software**

Check out this list of the top Time Series Intelligence Software products based on user satisfaction. A product's satisfaction score is calculated by a **proprietary algorithm** that factors in real-user satisfaction ratings from review data. Software buyers can compare products according to their satisfaction scores to streamline the buying process and quickly identify the best products based on the experiences of their peers.

Collapse All		G2 Satisfaction Score	Compare
#1	Azure Time Seri...	78	Compare
#2	Trendalyze	76	Compare
#3	Anodot	69	Compare
#4	Google Cloud Int...	60	Compare
#5	TrendMiner	59	Compare
#6	Avora	7	Compare

We help companies understand the pulse of their business.

Thank you