



Developing Spidey Senses

Anomaly Detection for Javascript

Ron Dagdag

Spidey Sense?

- tingling sensation on the back of Peter Parker's skull
- ability to sense / react to danger

Uses

- Increases his ability to detect evil (and even clones)
- Helps him navigate if he is impaired (disoriented or unable to see/hear)
- Aids him in discovering secret passageways and find hidden/lost objects
- Helps fire his Web Shooters and swing instinctively



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)

Real Spider Sense

- **“hyper-awareness”**
- long, thin hairs, *trichobothria*
- low-level vibrations through their web
- can detect the vibrations of faint sounds
- small insects moving up to 3 meters away





Any new web developers here?

Spidey Sense?

Gut feeling

Vibe

Feeling

Intuition

Discover Blind Spots

Learning from the past



Agenda

What is Anomaly Detection?

Time Series Anomaly Detection

Demo

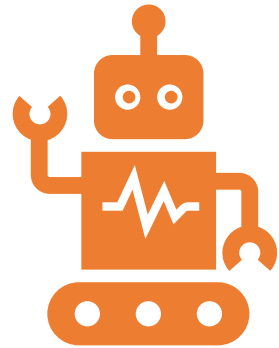
Takeaways

Anomaly Detection

- Identifying unexpected items or events in data sets, which differ from the norm
- An Outlier
- Assumptions:
 - Anomalies only occur very rarely in the data.
 - Their features differ from the normal instances significantly.



Causes of Outliers



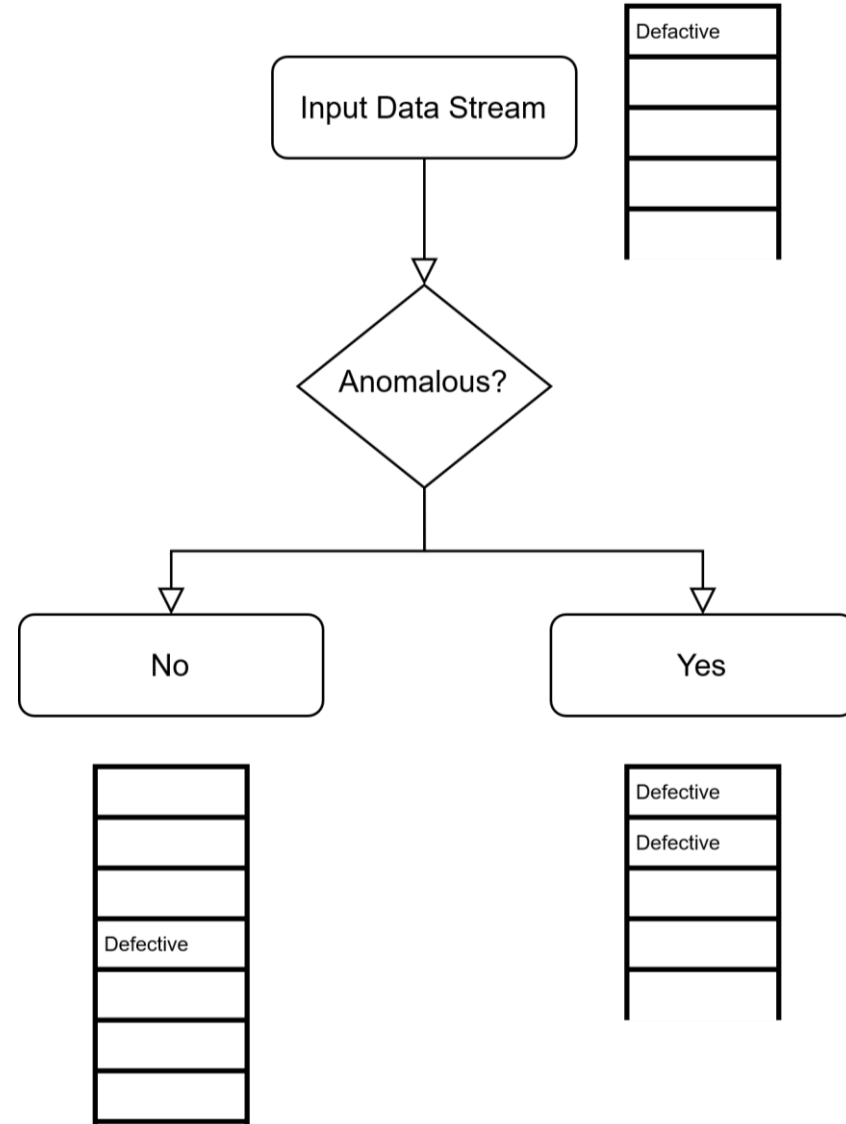
Artificial (Error) / Non-natural



Natural

Causes of Outliers

- Data Entry Errors: 100,000 vs 1,000,000 - fat fingered
- Measurement Error: common
- Experimental Error: start late in sprint
- Intentional Outlier: underreporting alcohol consumption
- Data Processing Error: extraction errors
- Sampling Error: reporting height for all athletes and included most basketball players
- Natural Outlier: When it's not artificial





Needle in a haystack

Methods



Rule-based Systems



Statistical Techniques

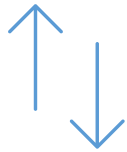


Machine Learning

Rule-based Systems



Specific Rules



Assign Threshold and
limits



Experience of Industry
Experts to detect
“known anomalies”



Doesn't Adapt as
patterns change



Data Labeling

Statistical Techniques



flags the data points => deviate from common statistical properties
(mean, median, mode, quantiles)



a rolling average or a moving average



n-period simple moving average
"low pass filter." e.g. Kalman Filters



Histogram-based Outlier Detection (HBOS)



More Interpretable and sometimes more useful than ML methods



Supervised
(e.g. Decision Tree, SVM, LSTM
Forecasting)



Unsupervised
(e.g. K-Means, Hierarchical Clustering,
DBSCAN)



Self-Supervised
(e.g. LSTM Autoencoder)

Machine Learning Methods

Anomaly Detection

- Very small number of positive examples
- Large number of negative examples
- Many different “types” of anomalies. Hard to learn from positive examples
- Future anomalies may not be discovered yet.

Supervised Learning

- Large number of positive and negative examples
- Enough positive examples for algorithm to learn.
- Future positive examples likely to be similar to training set

Anomaly Detection

- Fraud Detection
- Manufacturing
(engines/machineries)
- Monitoring Data Center
- Internet of Things

Supervised Learning

- Email spam classification
- Weather prediction
- Cancer classification

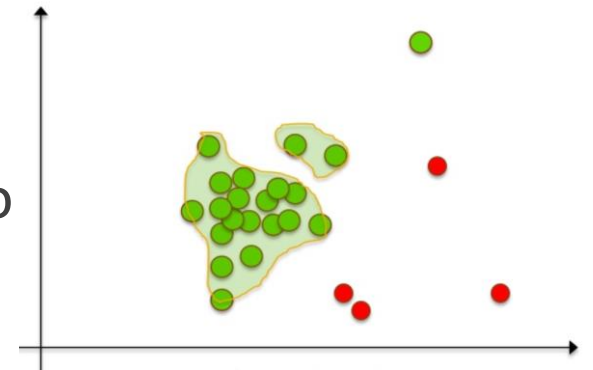
Machine Learning

Density-Based Anomaly Detection

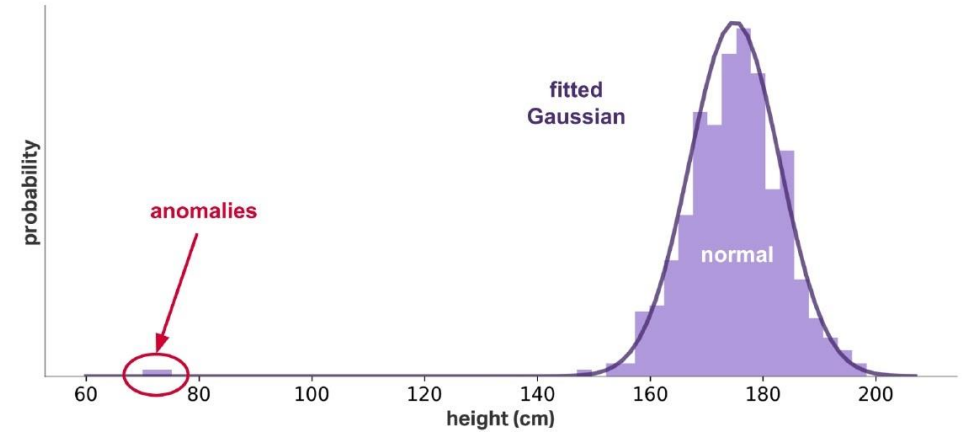
- based on the k-nearest neighbors' algorithm.
- *Assumption:* Normal data points occur around a dense neighborhood and abnormalities are far away.

Clustering-Based Anomaly Detection

- *Assumption:* Data points that are similar tend to belong to clusters --> distance from local centroids.
- K-means



Machine Learning



Gaussian Distribution

- Gaussian Distribution and given a new data-point,
- Compute the probability of the data-point
- If the probability is below a threshold => outlier or anomalous.

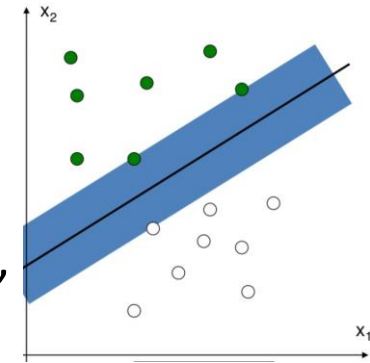
Machine Learning

Support Vector Machine-Based Anomaly Detection

- *OneClassSVM*
- *>100 features, aggressive boundary*
- find a function that is positive for regions with high density of points, and negative for small densities

PCA-Based Anomaly Detection

- analyzing available features to determine what constitutes a "normal" class
- applying distance metrics
- Fast training





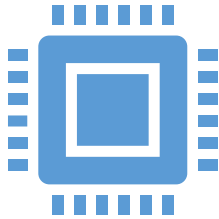
Simple Anomaly Detection DEMO

Time Series Data

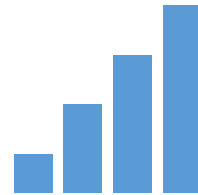
- Series of data points indexed in time order
- Examples:
 - Logs
 - Stock Market
 - Sales Data
 - Sensors
 - Any data captured with Time Stamp



Internet of Things



Increasing Data Volume
(sensors are cheaper)



Increased Data Speed
(improved networking)



Risk environment that are
moving very fast but failures are
not tolerated.

Internet of Broken Things



Time Series Anomaly Types



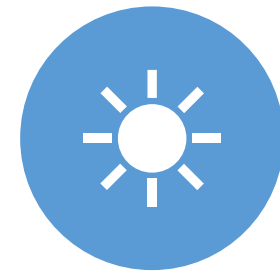
OUTLIER



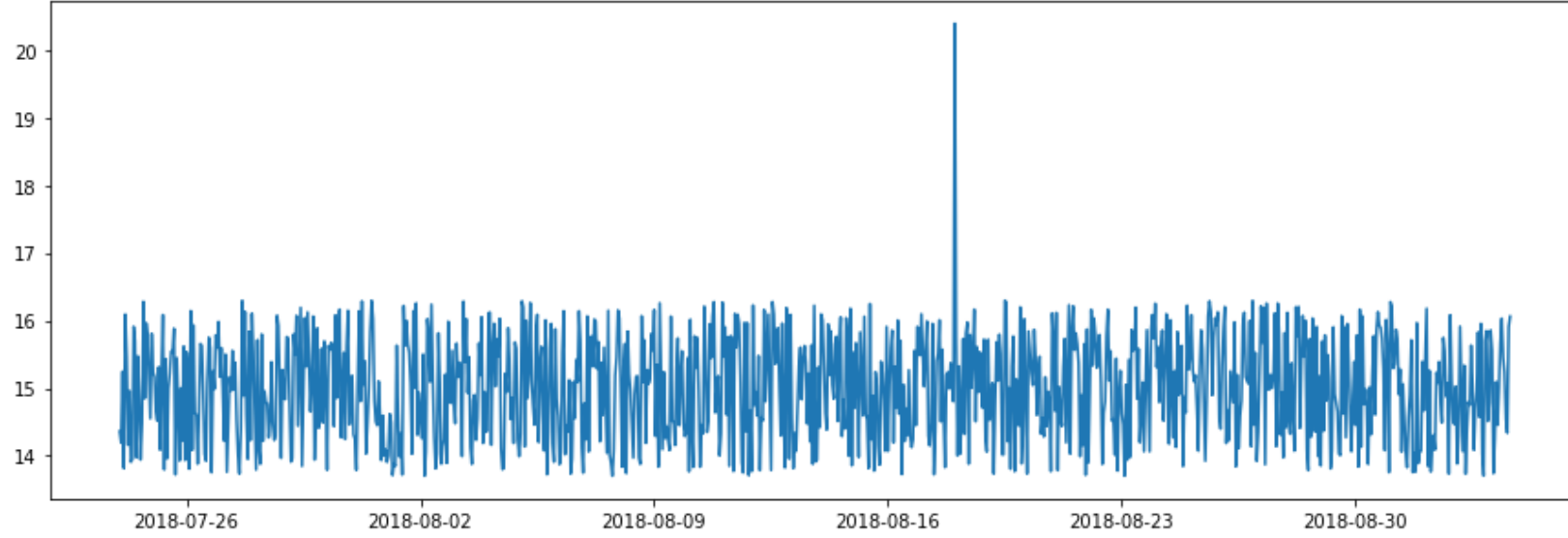
SPIKE AND
LEVEL SHIFT



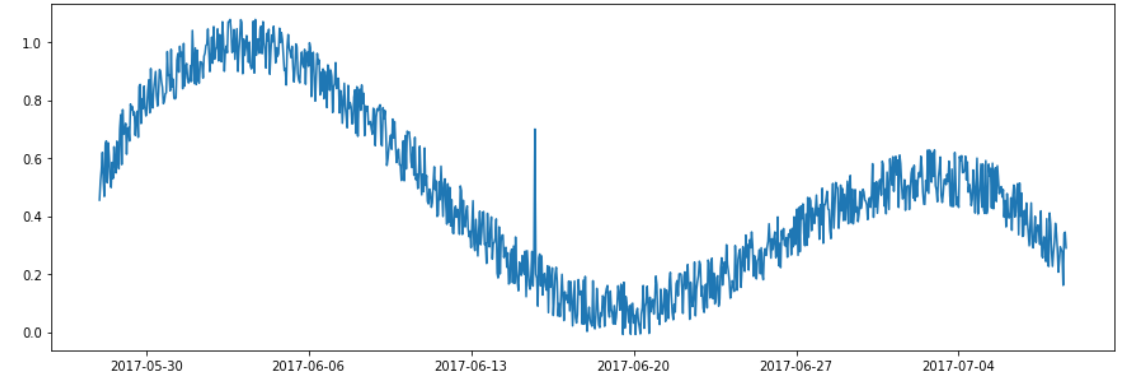
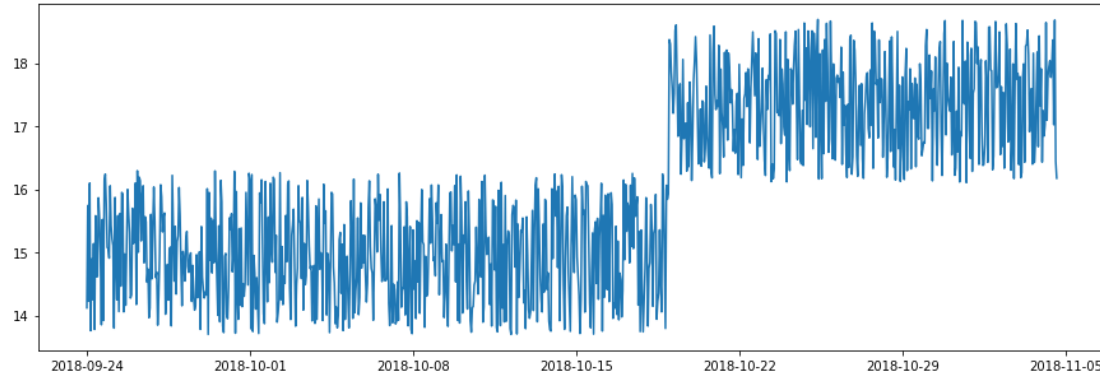
PATTERN
CHANGE



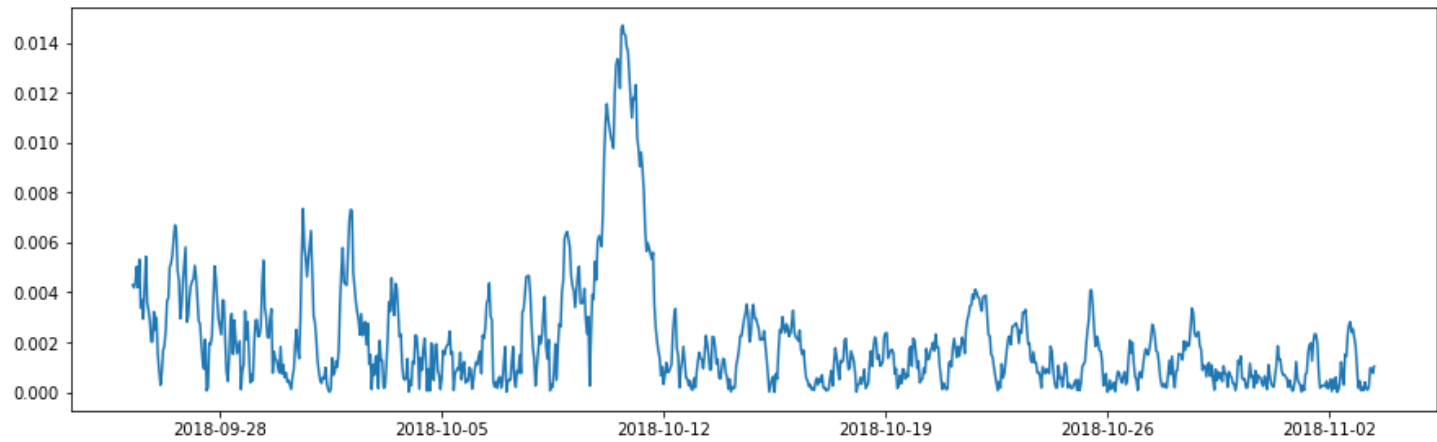
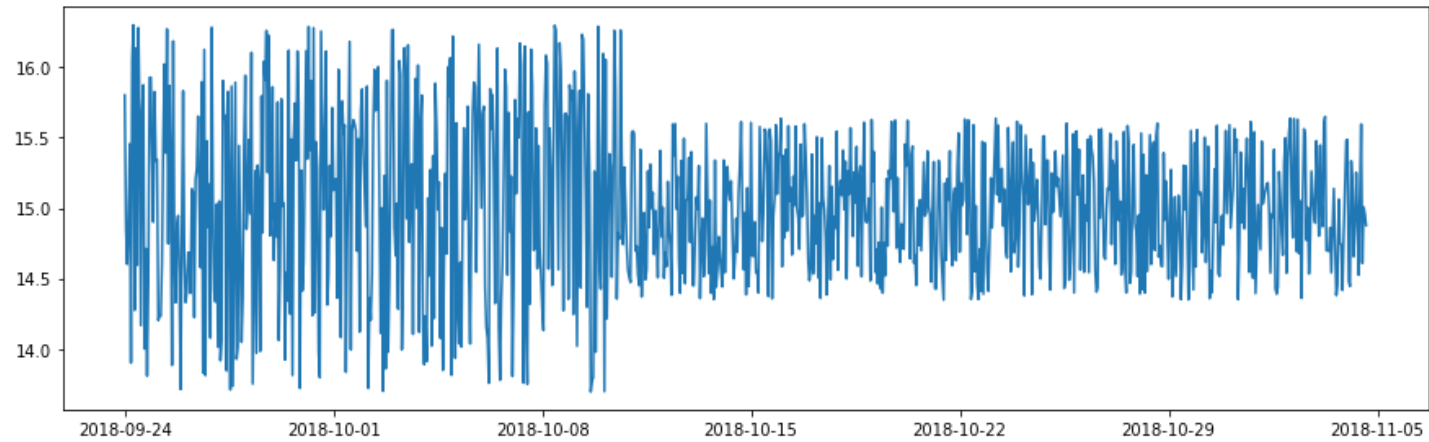
SEASONALITY



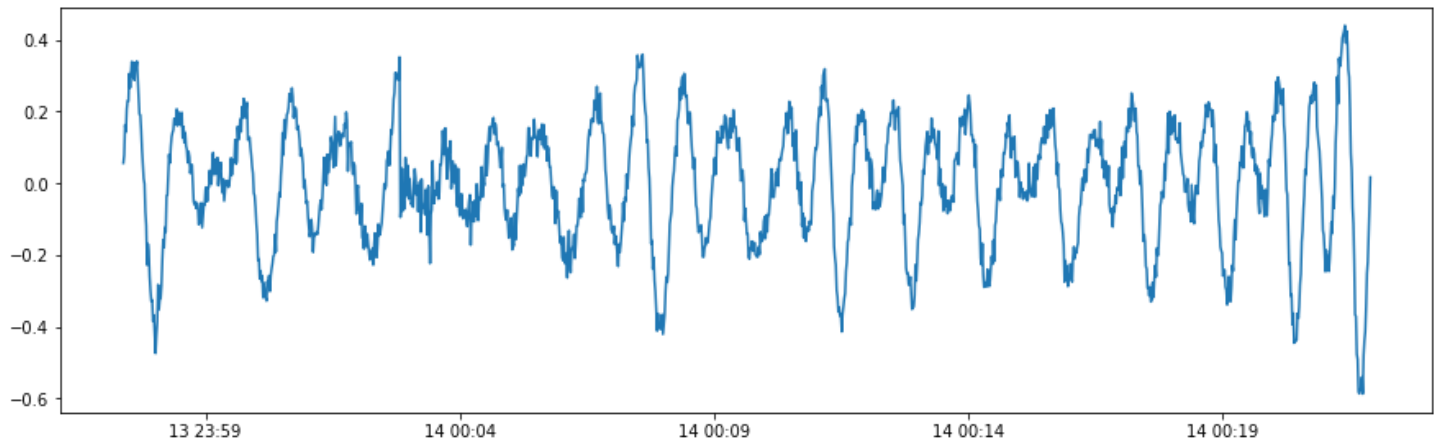
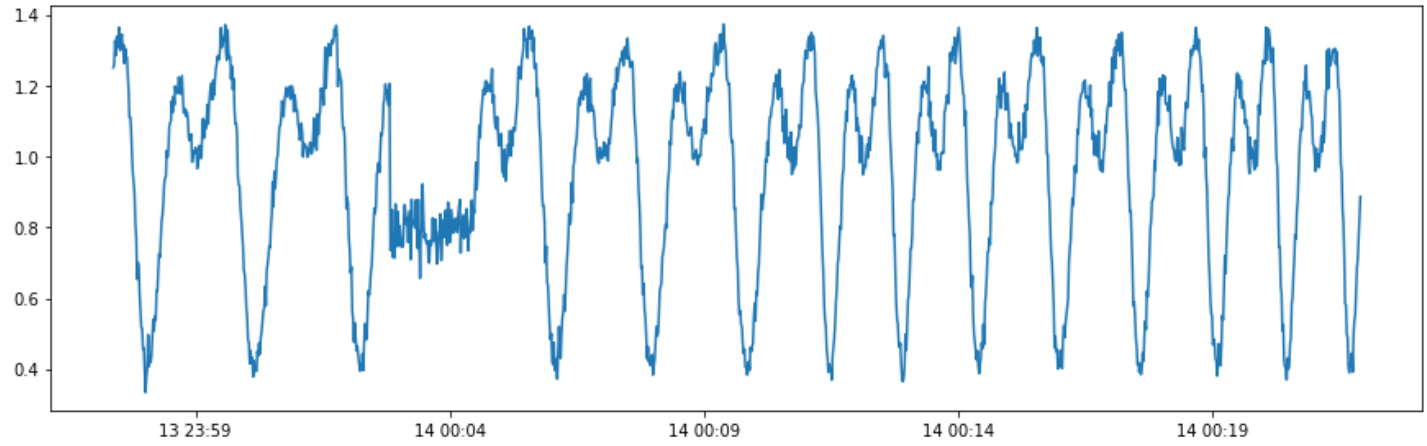
Outlier



Spike and Level Shift



Pattern Change



Seasonality

Anomaly Detector

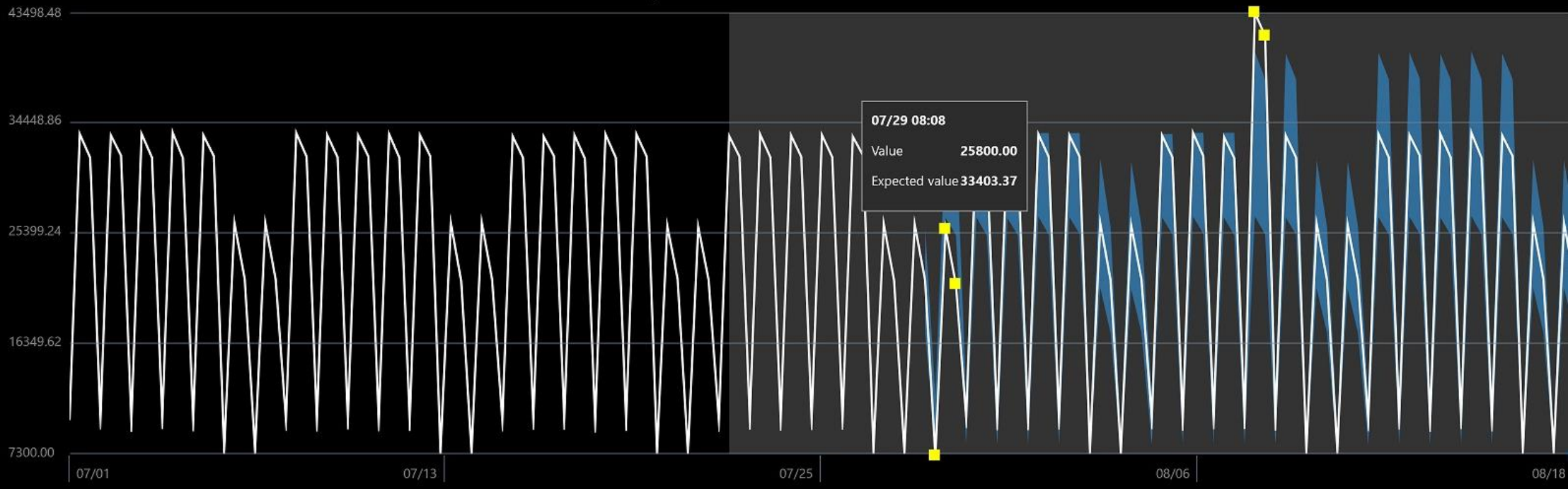
Bike rental Manufacturing Telecom Live sound



Adventure Works bike rentals

Adventure Works runs bike rental business in more than 60 cities which are divided into about 6,000 subareas. Hourly number of rides in each subarea forms a time series. Anomaly Detector monitors those thousands of metrics to help the company make critical operational decisions timely: for example, transporting hundreds of bikes to an area with demand spikes.

Hourly number of rides in area A76



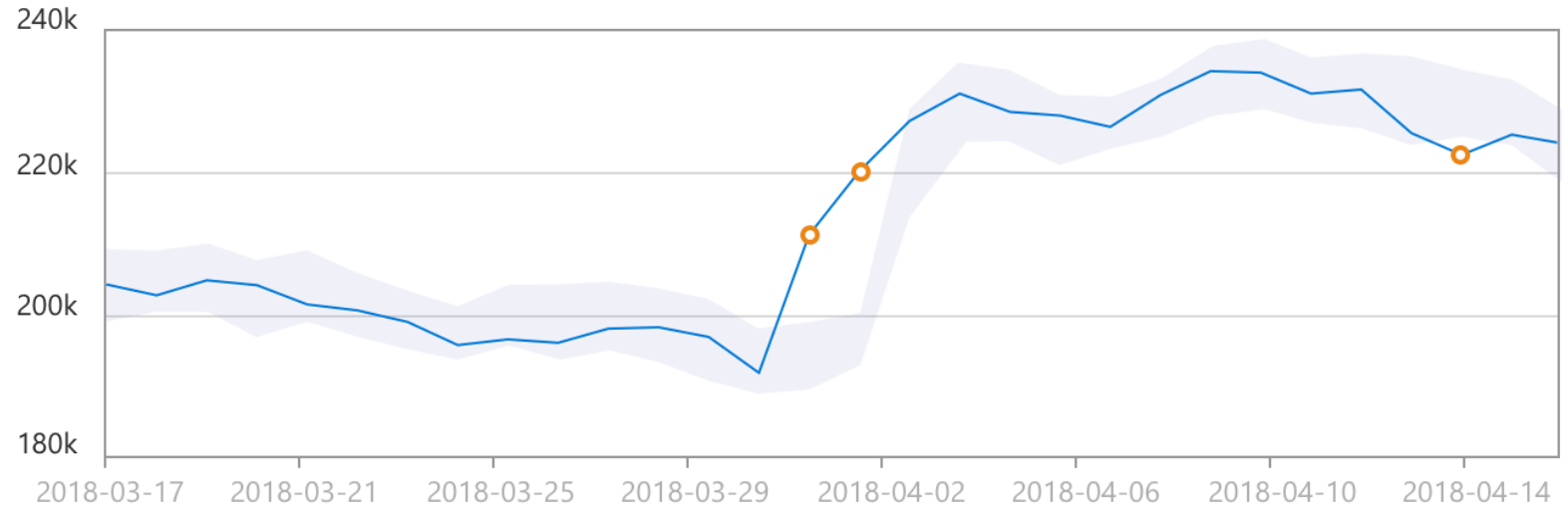
Detection Mode: Streaming ⓘ

Sensitivity

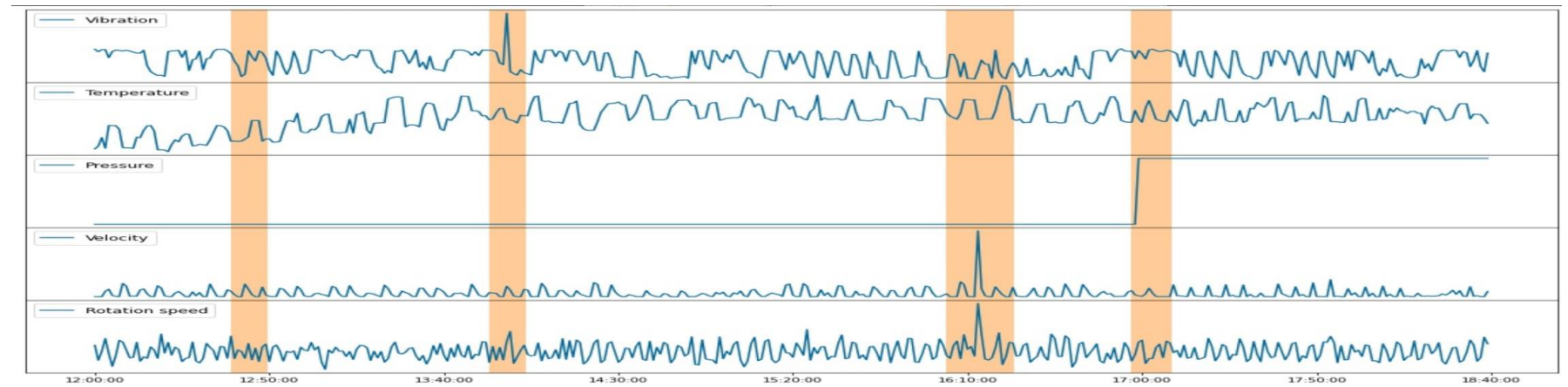
Detect Anomalies

Time Series Anomalies

- **Univariate**



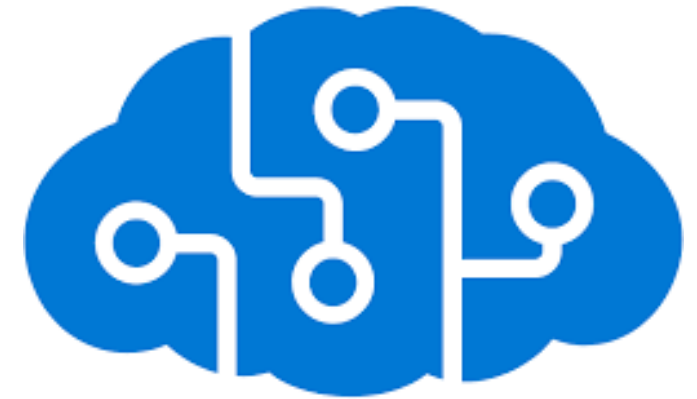
- **Multivariate**



Azure Cognitive Services

- AI for every developer— w/o requirement ML expertise.
- Just an API call

Decision	Make smarter decisions faster
Language	Anomaly Detector <small>PREVIEW</small> Identify potential problems early on.
Speech	Content Moderator Detect potentially offensive or unwanted content.
Vision	Personalizer Create rich, personalized experiences for every user.
Web search	



Anomaly Detector Features



Detect anomalies as they occur in real-time.



Detect anomalies as a batch.



Automatically adapts and learns from new data



Fine Tune Sensitivity

Anomaly Detector Features



REST API



No machine learning expertise needed

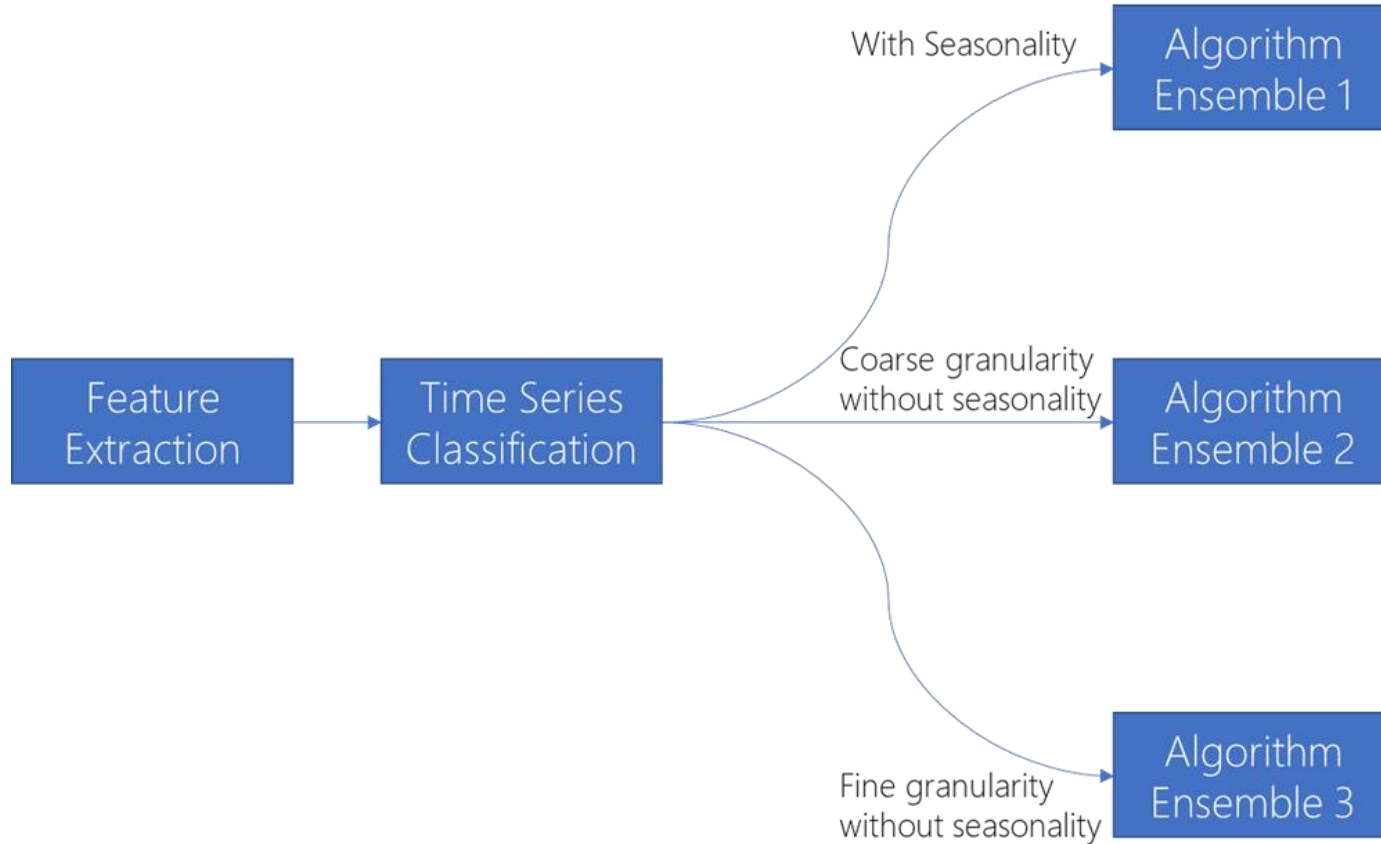


Eliminate need for labeled training data



Automatically identify and apply best-fitting model

Gallery of Algorithms



Fourier
Transformation
Extreme Studentized
Deviate (ESD)
[STL Decomposition](#)
Dynamic Threshold
Z-score detector
[SR-CNN](#)

Limitations

- Data Granularity – Daily, Hourly, Minutely, Monthly, Weekly, Yearly
- Series Data Points – 12 to 8640 entries

- ```
JSON
{
 "granularity": "daily",
 "series": [
 {
 "timestamp": "2018-03-01T00:00:00Z",
 "value": 32858923
 },
 {
 "timestamp": "2018-03-02T00:00:00Z",
 "value": 29615278
 },
]
}
```

```
JSON
{
 "granularity" : "minutely",
 "customInterval" : 5
}
```

# Calling the Anomaly Detector API



## Client SDK

C#, Python, Node



## REST API

Any language supporting HTTP calls

# Anomaly Detector Demo



# Where can you use this?

---

C#, Javascript, Python

---

Docker Containers

---

Power BI

---

Azure Databricks for streaming data

# Anomaly Detector Demo



@rondagdag

[This Photo](#) by Unknown Author is licensed under [CC BY-NC-ND](#)



# Where can you use this?

Javascript, Python, C#

<https://docs.microsoft.com/en-us/azure/cognitive-services/anomaly-detector/quickstarts/client-libraries?pivots=programming-language-csharp&tabs=linux>

Docker Containers

<https://docs.microsoft.com/en-us/azure/cognitive-services/anomaly-detector/anomaly-detector-container-howto>

- **Power BI**

- <https://docs.microsoft.com/en-us/azure/cognitive-services/anomaly-detector/tutorials/batch-anomaly-detection-powerbi>

- **Azure Databricks for streaming data**

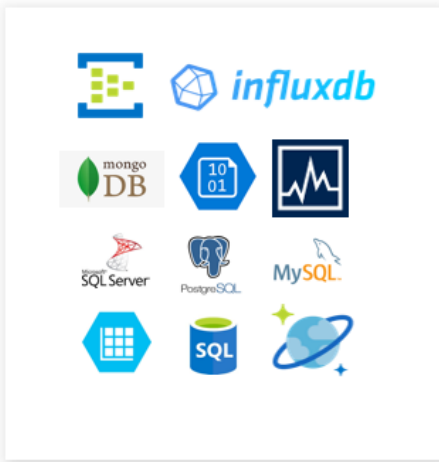
- <https://docs.microsoft.com/en-us/azure/cognitive-services/anomaly-detector/tutorials/anomaly-detection-streaming-databricks>

# Metrics Advisor

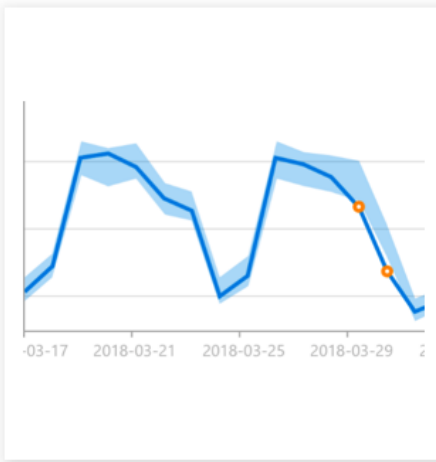
- Part of Azure Cognitive Services
- Performs data monitoring, anomaly detection in time series data
- Automates applying models
- Analyze multi-dimensional data from multiple data sources
- Identify and correlate anomalies
- Configure and fine-tune the anomaly detection model
- Diagnose anomalies and help with root cause analysis
- REST API and Web Portal
  
- Currently in preview



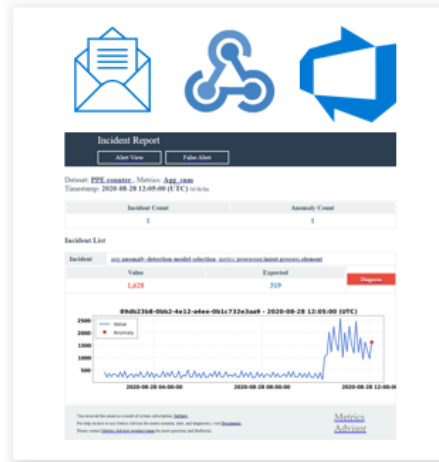
Collect time-series data



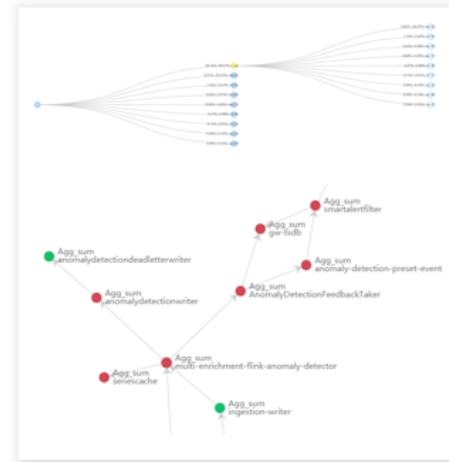
Detect anomalies



Send incident alerts



Analyze root cause



The best superpower you can give to your project is a “spidey-sense”.





<https://github.com/rondagdag/spidey-sense-js>

@rondagdag

# Recap

- ✓ What is Anomaly Detection?  
process of identifying unexpected items or events in data sets
- ✓ What is Time Series Data  
series of data points indexed in time order
- ✓ Anomaly Detector  
API to detect anomalies  
automatically adapts  
learns from new data

# About Me

Ron Dagdag



- Lead Software Engineer / AI Edge Specialist
- 5<sup>th</sup> year Microsoft MVP awardee
- Personal Projects  
[www.dagdag.net](http://www.dagdag.net)
- Email: [ron@dagdag.net](mailto:ron@dagdag.net)  
Twitter [@rondagdag](https://twitter.com/rondagdag)
- Connect me via Linked In  
[www.linkedin.com/in/rondagdag/](https://www.linkedin.com/in/rondagdag/)
- Thanks for geeking out with me about Spidey Senses and Anomaly Detection  
[@rondagdag](https://twitter.com/rondagdag)