Microsoft Azure

Azure Cognitive Services - Anomaly Detector

Swarnalata Tamanam
Swarnalata.tamanam@valuemomentum.com

Sheldon Dsouza
Sheldon.dsouza@valuemomentum.com

#GlobalAzure

CORPORATE FACTS

Established in 2000

90+
Customers Served

2300+
Employees

23%
CAGR since inception

Top 10
NA P/C IT Svs Provider
by # of customers

OUR DIGITAL & CLOUD SERVICES

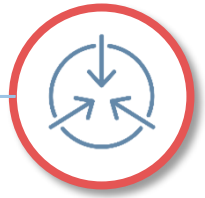
VALUEMOMENTUM
Deliver Value. Drive Momentum.

Customers trust ValueMomentum to rapidly deliver new experiences and stay competitive in today's digital-centric market.*

App Dev

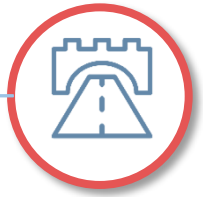
- Rapid App Development
- Cloud App Development
- Mobile App Development

Integration



- Integration Architecture
- API Development
- Service Modernization

Cloud & Infrastructure



- Cloud Platform Adoption
- Cloud Migration
- DevOps & CloudOps

ITSM

- IT Service Management
- IT Asset Management
- IT Operations Management

*To learn more, please log on to ValueMomentum – Digital & Cloud Services

VALUEMOMENTUM

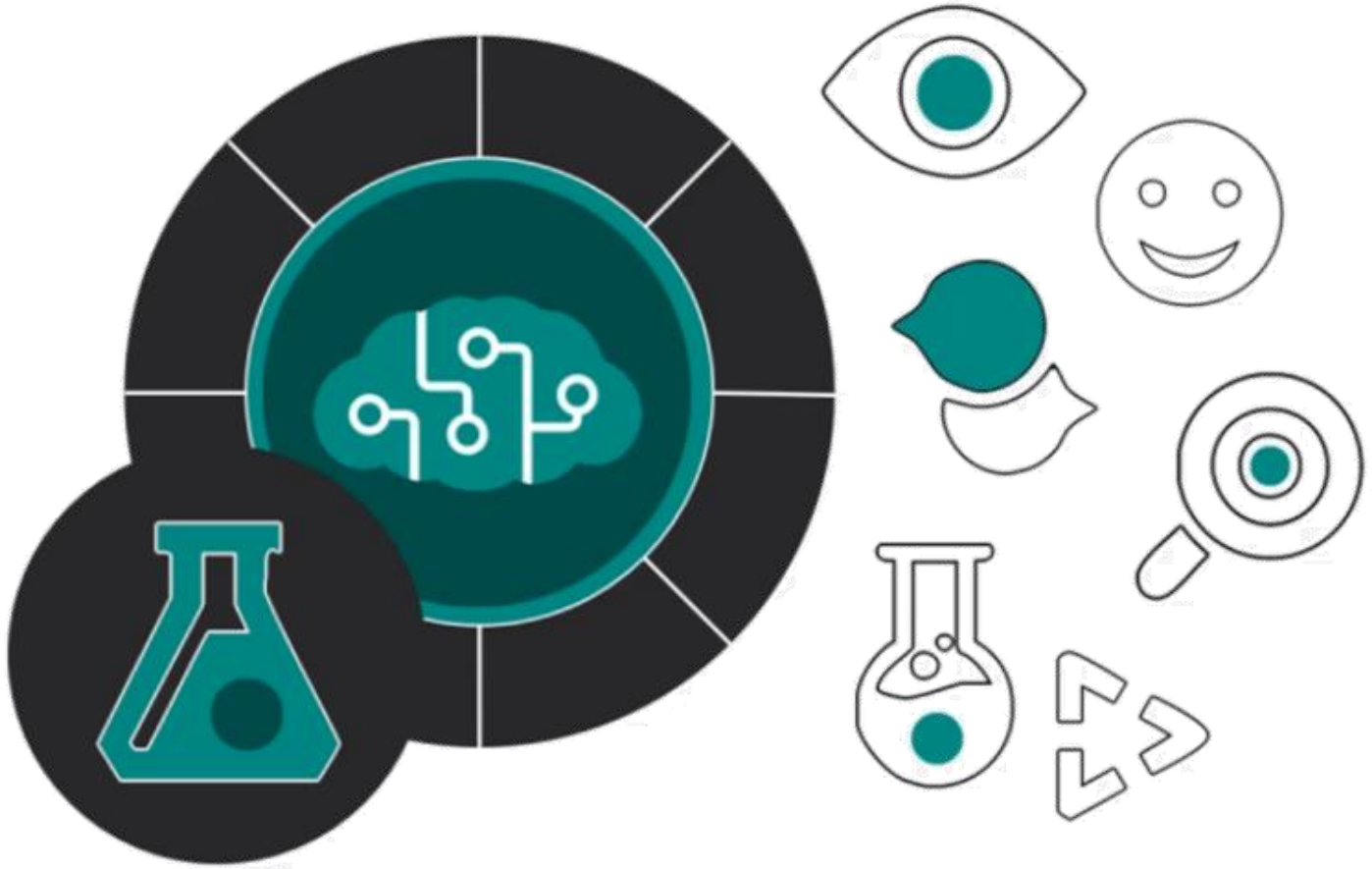
Deliver Value. Drive Momentum

Agenda

- Cognitive Services
- What is an Anomaly
- Types of Anomalies
- Key Challenges
- Use cases
- Anomaly Detector API

Cognitive Services

Microsoft Cognitive Services are a **set of APIs, SDKs and services** available to developers, exposing a suite of machine learning APIs that enable developers to easily add intelligent features—such as emotion and video detection, facial, speech and vision recognition, and speech and language understanding—into their applications.



2019 Global Azure BOOTCAMP

Cognitive Services Stack

VALUE MOMENTUM
Deliver Value. Drive Momentum

Solutions	Extensible applications
Cognitive services / Bot framework	Easy to consume Artificial Intelligence
Data Science tools Data preparation, modeling, and operationalization	Most comprehensive data science capabilities
Algorithms	Best of Microsoft research and open source
Unique data assets	Rich data sources to enrich predictions
Analytics in Big Data Stores (cloud + on premise)	Flexible infrastructure support for analytics

2019 Global Azure BOOTCAMP

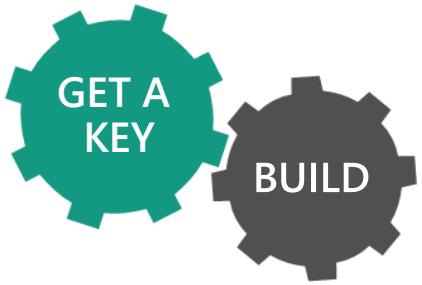
Why Azure Cognitive Services

VALUEMOMENTUM
Deliver Value. Drive Momentum

Easy

Roll your own with REST APIs

Simple to add: just a few lines of code required



Flexible

Integrate into the language and platform of your choice

Breadth of offerings helps you find the right API for your app

Tested

Built by experts in their field from Microsoft Research, Bing, and Azure Machine Learning

Quality documentation, sample code, and community support

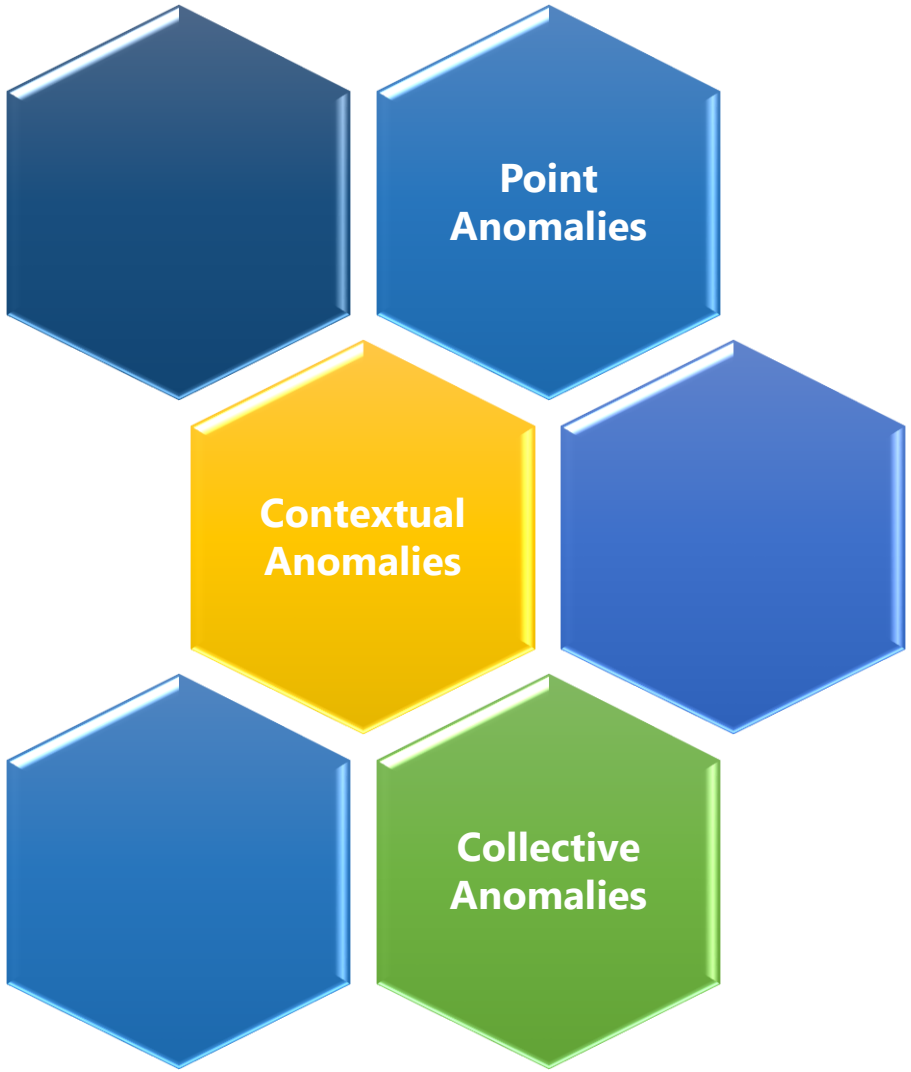
GitHub stackoverflow msdn uservoice

What is an anomaly?

- Something that deviates from what is standard, normal, or expected.
- Anomaly is a pattern in the data that does not conform to the expected behaviour
- Also referred to as outliers, exceptions, peculiarities, surprise, etc.

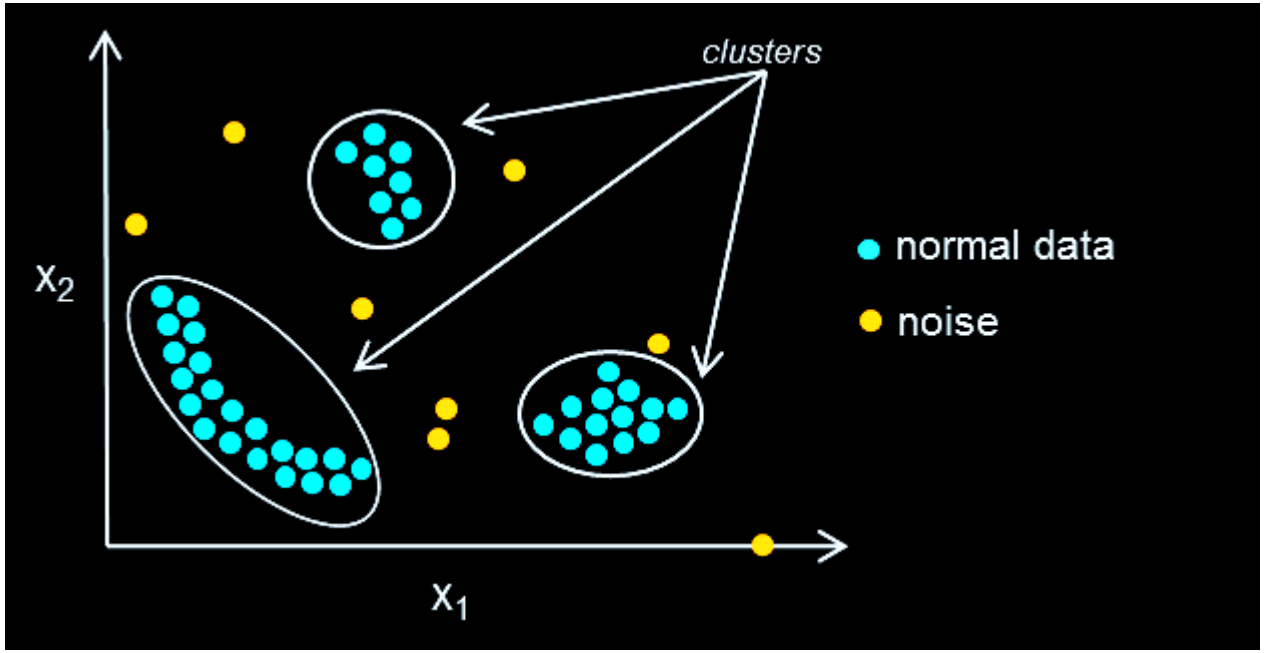
Needle in a haystack?

Types of Anomalies



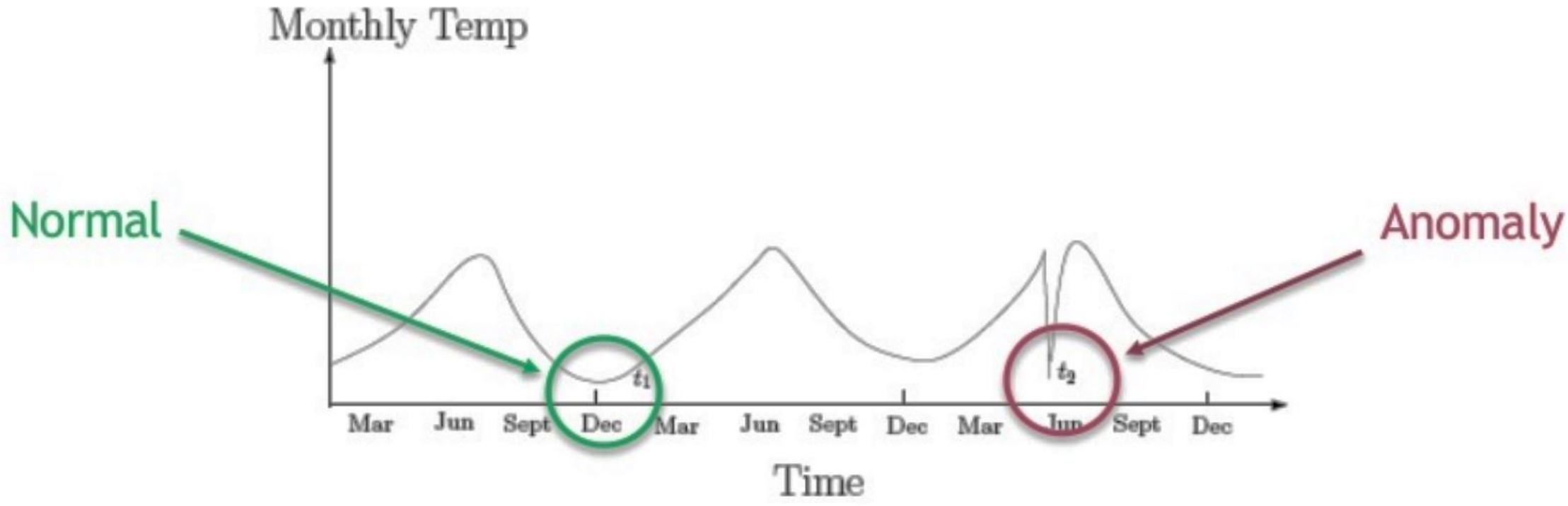
Point Anomalies

- An individual data instance is anomalous w.r.t. the data



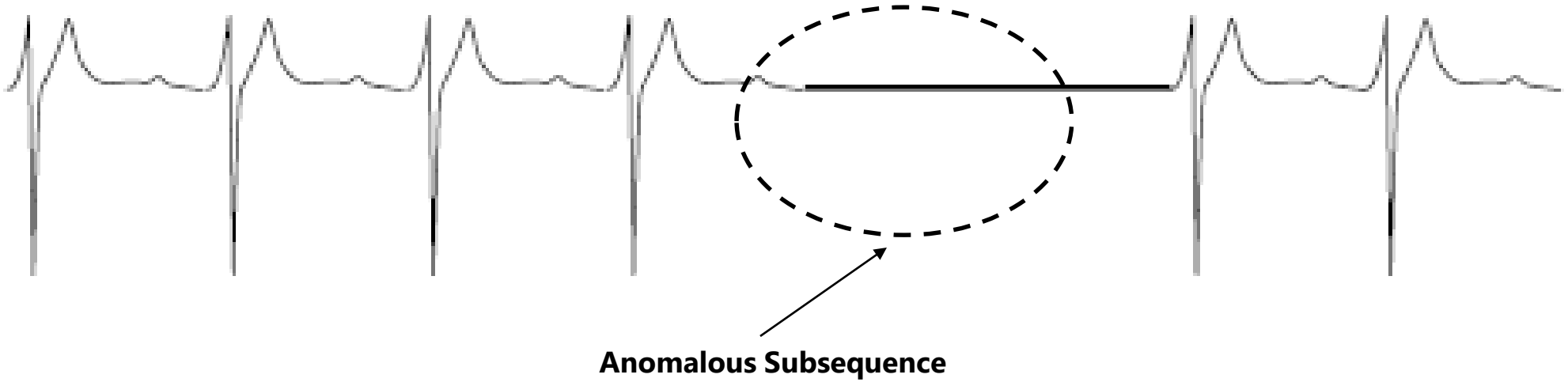
Contextual Anomalies

- An individual data instance is anomalous within a context
- Requires a notion of context
- Also referred to as conditional anomalies



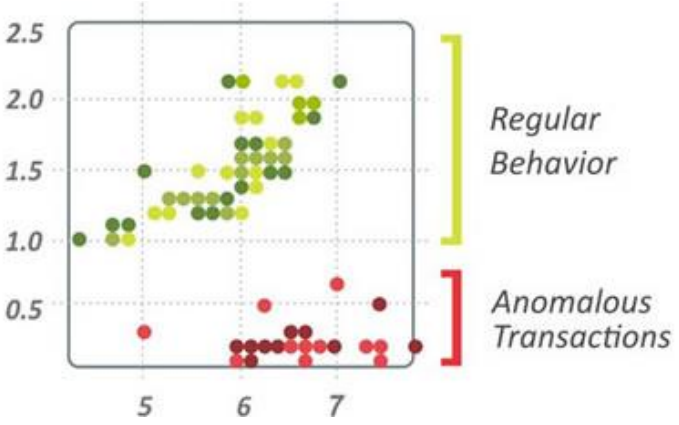
Collective Anomalies

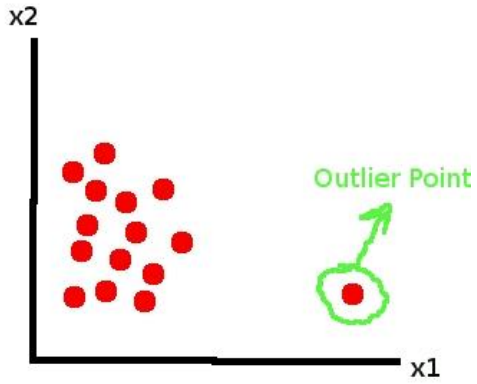
- A collection of related data instances is anomalous
- Requires a relationship among data instances
  - Sequential Data
  - Spatial Data
  - Graph Data
- The individual instances within a collective anomaly are not anomalous by themselves

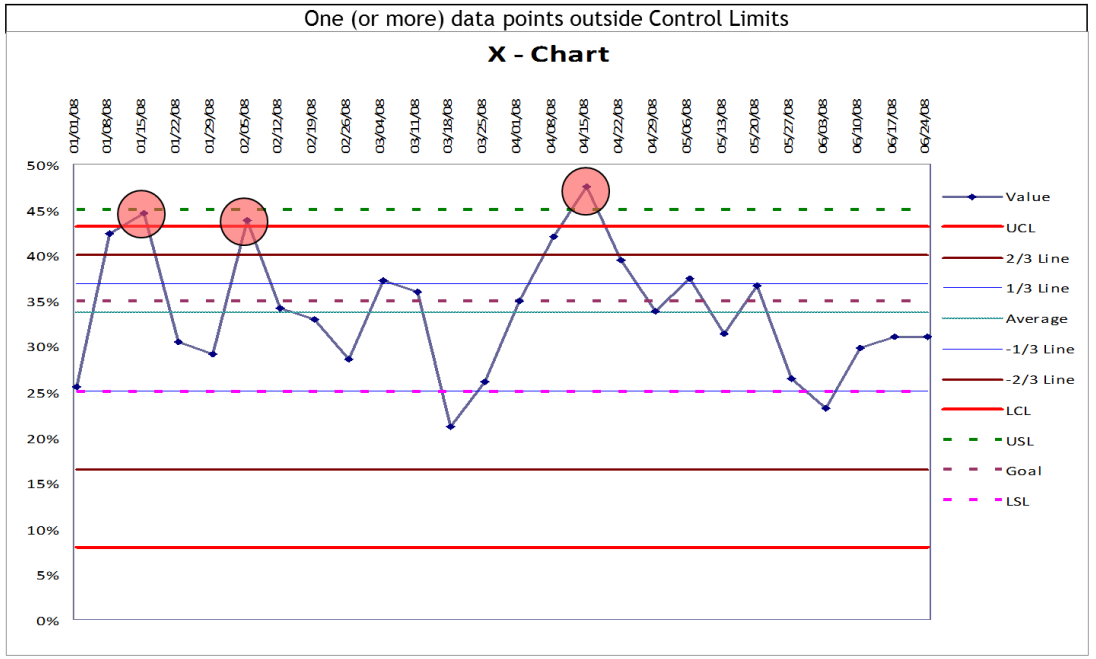


Key Challenges

- Defining a representative normal region is challenging
- The boundary between normal and outlying behaviour is often not precise
- The exact notion of an outlier is different for different application domains
- Availability of labelled data for training/validation
- Malicious adversaries
- Data might contain noise
- Normal behaviour keeps evolving





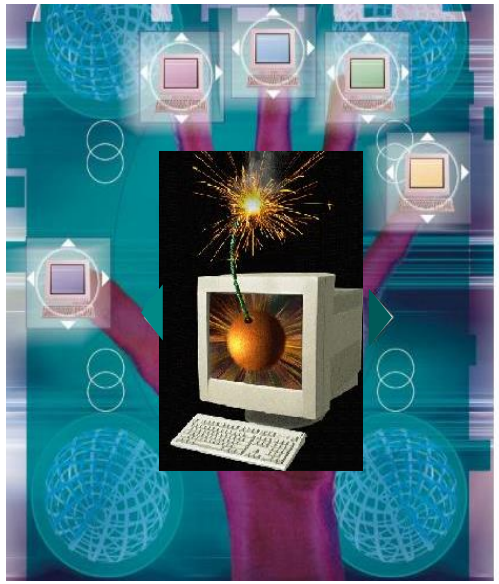


Use cases

Intrusion Detection

- Intrusion Detection:
  - Process of monitoring the events occurring in a computer system or network and analyzing them for intrusions
  - Intrusions are defined as attempts to bypass the security mechanisms of a computer or network

- Challenges
  - Traditional signature-based intrusion detection systems are based on signatures of known attacks and cannot detect emerging cyber threats
  - Substantial latency in deployment of newly created signatures across the computer system



Fraud Detection

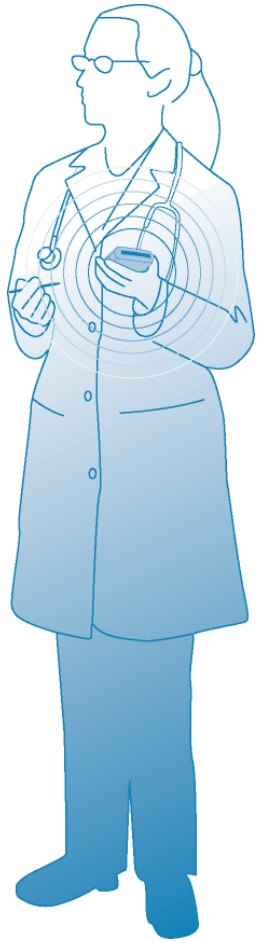
- Fraud detection refers to detection of criminal activities occurring in commercial organizations

- Challenges
  - Fast and accurate real-time detection
  - Misclassification cost is very high

Healthcare Informatics/Medical Diagnostics

- Detect anomalous patient records

- Key Challenges
  - Only normal labels available
  - Misclassification cost is very high
  - Data can be complex: spatio-temporal



Industrial Damage Detection

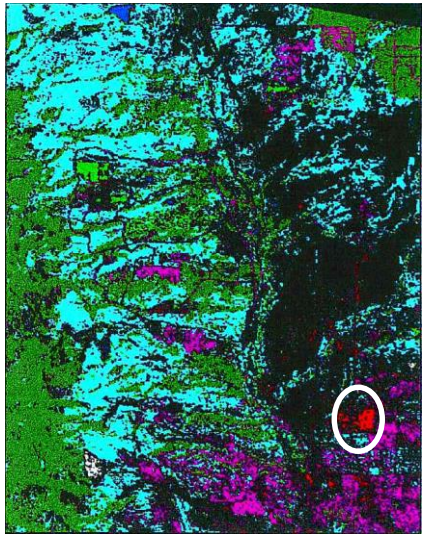
- Industrial damage detection refers to detection of different faults and failures in complex industrial systems, structural damages, intrusions in electronic security systems, suspicious events in video surveillance, abnormal energy consumption, etc.

- Key Challenges
  - Data is extremely huge, noisy and unlabelled
  - Most of applications exhibit temporal behaviour
  - Detecting anomalous events typically require immedia

Image Processing

- Detecting outliers in an image monitored over time
- Detecting anomalous regions within an image

- Key Challenges
  - Detecting collective anomalies
  - Data sets are very large



Anomaly Detector API

The Anomaly Detector API enables you to monitor and detect abnormalities in your **time series data** with machine learning. The Anomaly Detector API adapts by automatically identifying and **applying the best-fitting models** to your data, regardless of industry, scenario, or data volume. Using your time series data, the API determines **boundaries** for anomaly detection, **expected values**, and which **data points** are anomalies.

Features of Anomaly Detector API

- Detect anomalies as they occur in real-time.
- Detect anomalies throughout your data set as a batch.
- Get additional information about your data – expected values, anomaly boundaries and positions
- Adjust anomaly detection boundaries – increase or decrease API's sensitivity to data anomalies for better fit

Demo Use Case - credit card fraud

Credit Card Information

- User ID
- IP address
- Timestamp
- Email address
- Phone number
- Device ID / signature
- Amount
- Credit card number / payment method
- Billing address
- Shipping address

Instance 1


```
{
  "timestamp": "2017-12-08T12:00:00Z",
  "BillAddress": "Mumbai",
  "Amount": 8122.99
},
{
  "timestamp": "2017-12-08T13:00:00Z",
  "BillAddress": "Hyderabad",
  "Amount": 8145.49
},
{
  "timestamp": "2017-12-08T14:00:00Z",
  "BillAddress": "New York",
  "Amount": 7942
}
```

Instance 2

Timestamp	Amount
3/23/18 17:00	15000
3/24/18 12:00	38000
3/24/18 19:00	56000
3/25/18 9:00	350000
3/26/18 14:00	2500
3/27/18 18:00	9000
3/28/18 20:00	265
3/29/18 10:00	2590

Demo

© 2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries.
The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.