Google

Interpretable, Multidimensional, Multimodal Anomaly Detection with Negative Sampling

ICML 2020

John Sipple sipple@google.com

July 2020

Motivation

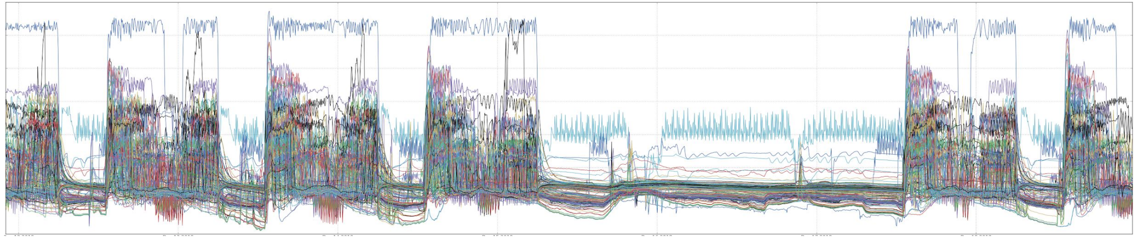
Complex systems generate vast streams of telemetry

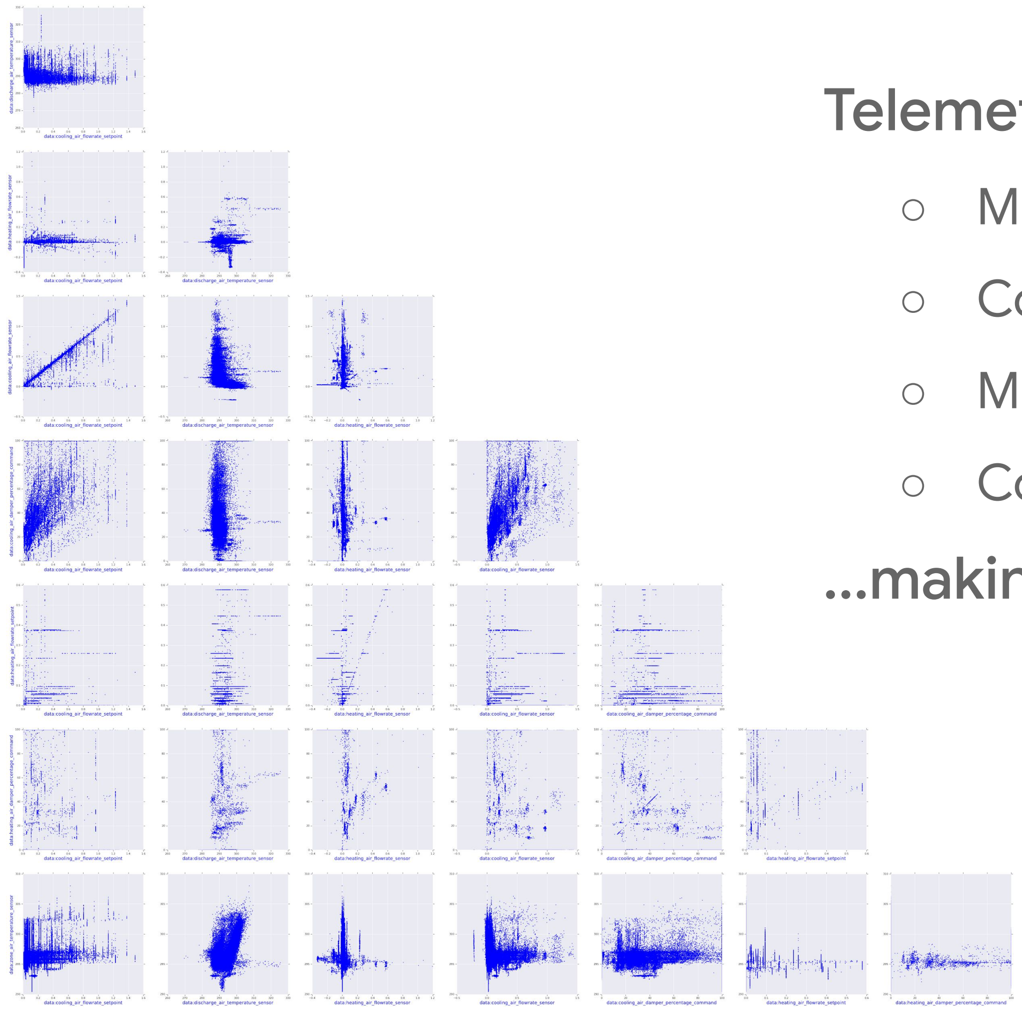
Malfunctions often can be predicted

- Outside range
- Correlations lost

Fixed rules or supervised approaches often ineffective

- Complex patterns
- Novel failure modes
- Few failure examples





Telemetry data streams are often...

- Multidimensional
- Correlated
- Multimodal
- Complex

...making Anomaly Detection hard!

Google

Anomaly Detection Problem

Why is it "anomalous"?

What is "normal"?

$$P(x \in Normal) \approx 0$$

How do we test?

x: observed point in $\mathbb{R}^D$

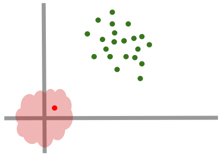
$Normal$: region in $\mathbb{R}^D$ representing expected behavior

Detect the Anomaly

- What is “normal”?
- How do we test?

Anomaly Detection

Few/no failure labels challenge supervised approaches



One-class Classifiers

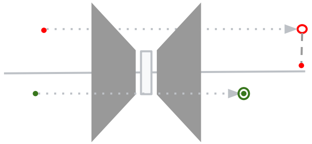
Learn a transformation to separate the observed points from the origin.

- One-Class SVM (2001)
- Deep SVDD (2018)

Density-Based

Anomalous points occur in low-density regions

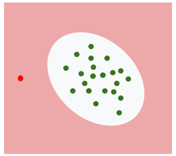
- Local Outlier Factor (2000)
- Isolation Forest (2009) and Ext. Isolation Forest (2018)



Autoencoders and Generative Models

Anomalies have larger reconstruction errors than Normal points

- AnoGAN (2017)
- GANomaly (2018)
- DAE-DBC (2018)



Negative Sampling Methods

Explicitly define negative space for anomalies.

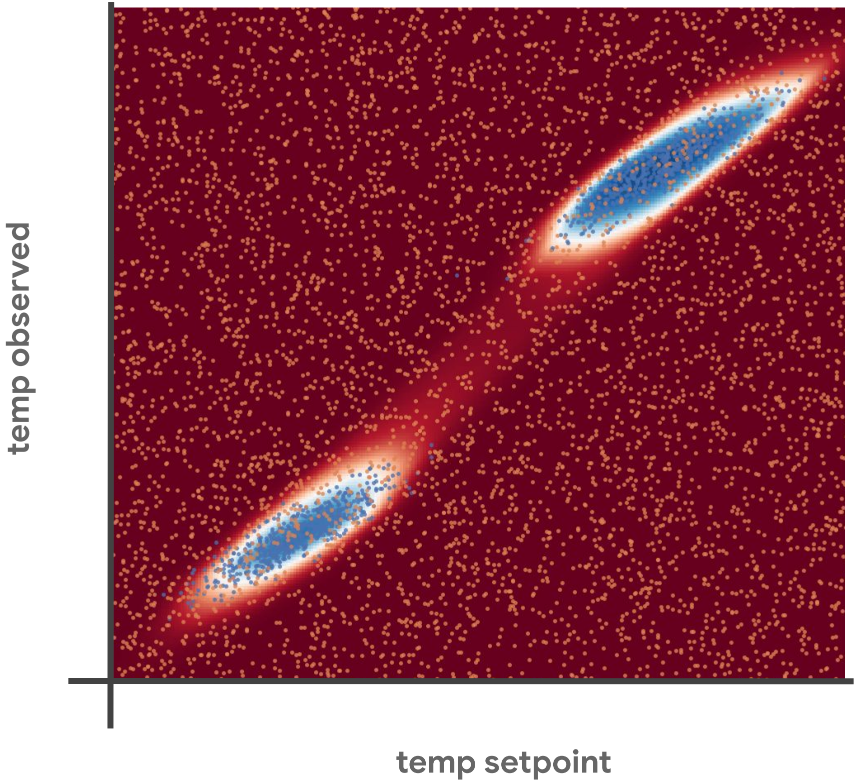
- Neg Selection Algorithms (NSA) (2002)
- Neg Sampling Classifiers (**this work**)

Negative Sampling Anomaly Detection

Positive Region = Observed ≈ Normal

Negative Region = Complement of Positive ≈ Anomalous

Train DNNs and Random Forests to predict P(x∈*Normal*)



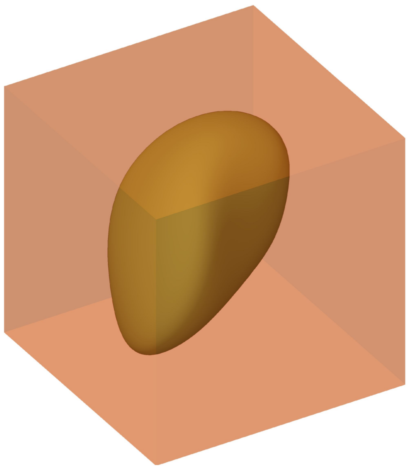
Sampling the Training Set

Positive Sample: Most observed points are normal, and anomalies are <u>rare</u>.

Negative Sample: Computationally hard to define a tight hull of an arbitrary shape in $\mathbb{R}^D$

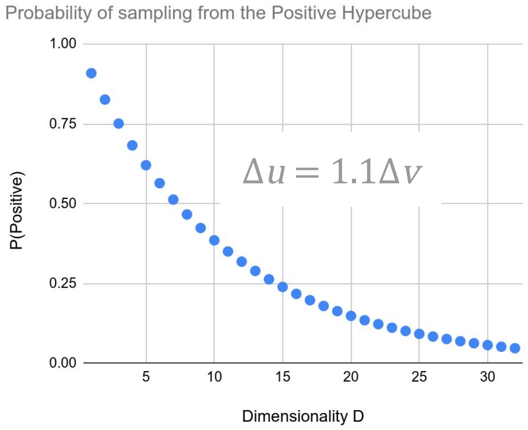
Alternatively, sample uniformly

Concentration Phenomenon: Volume increases exponentially with *D*

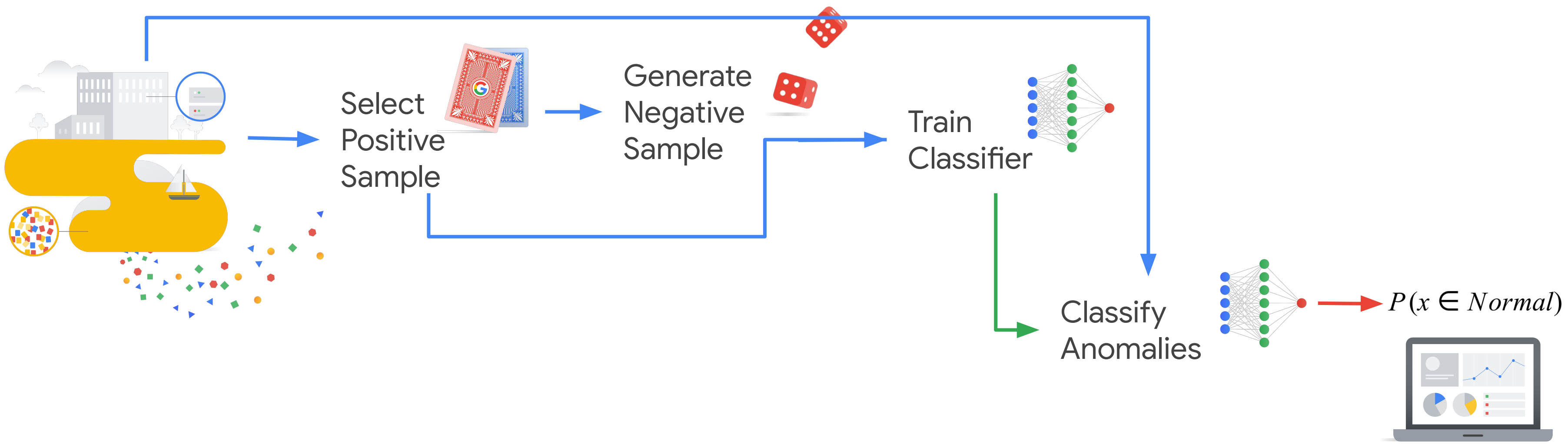


$$\lim_{D\to\infty} \prod_{d}^{D} \frac{\Delta v_d}{\Delta u_d} = 0$$

$$\Delta u_d > \Delta v_d$$



Anomaly Detection Pipeline



Anomaly Detection Results

ROC-AUC %	OC-SVM	Deep SVDD	Iso Forest	Extended Iso Forest	**NegSampleRnd Forest**	**NegSample Neural Net**
Forest Cover*	53 ±20	69 ±7	85 ±4	**93 ±1**	80 ±2	86 ±4
Shuttle*	93 ±0	88 ±9	**96 ±1**	91 ±1	**93 ±7**	**96 ±5**
Mammography*	71 ±7	78 ±6	77 ±2	**86 ±2**	**85 ±4**	84 ±2
Mulcross*	90 ±0	54 ±4	88 ±0	66 ±4	94 ±1	**99 ±1**
Satellite*	51 ±1	62 ±3	67 ±2	**71 ±3**	65 ±4	**73 ±3**
Smart Buildings	76 ±1	60 ±7	71 ±7	80 ±4	**95 ±1**	93 ±1

* Courtesy of ODDS Library [http://odds.cs.stonybrook.edu].
Stony Brook, NY: Stony Brook University, Department of Computer Science

Interpret the Anomaly

- Why is it "anomalous"?

Anomaly Interpretation

Attribute influence with differentiable classifier function $F(x)$, and **Integrated Gradients** (Sundararajan, 2017)

$$B_d(x) \equiv (u_d^* - x_d) \times \int_{\alpha=0}^{1} \frac{\partial F(x + \alpha \times (u^* - x))}{\partial x_d} d\alpha$$

Requires a neutral, baseline point, u*.

(1) Choose a **baseline set** U* from the positive sample U, where U* are Normal

$$U^* \subset U : \forall_{u \in U^*} F(x) \approx 1$$

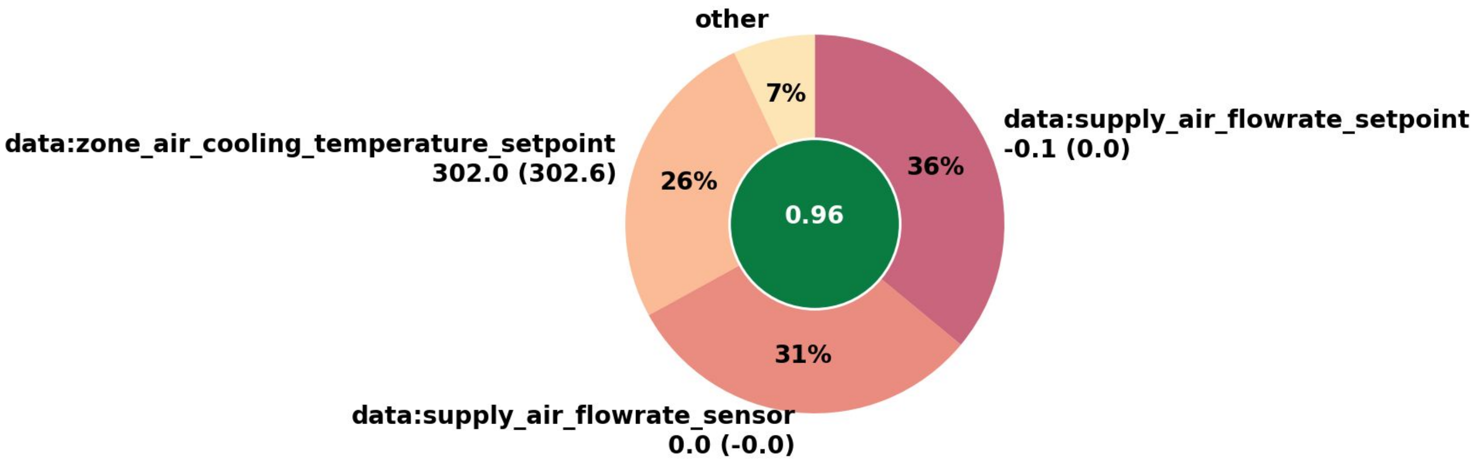
(2) Choose $u*$ from $U*$ with the minimum distance $dist(\cdot,\cdot)$ to Anomaly *x*

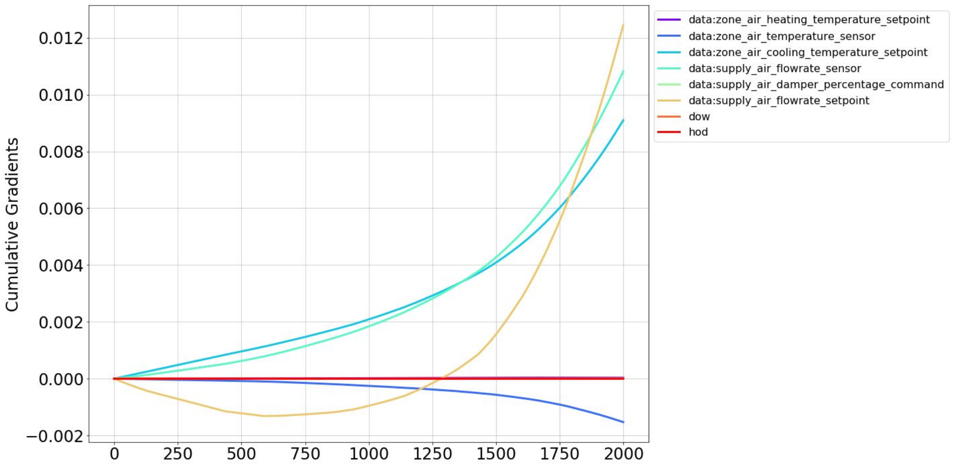
$$u^* = argmin_{u \in U^*} \{dist(x, u)\}$$

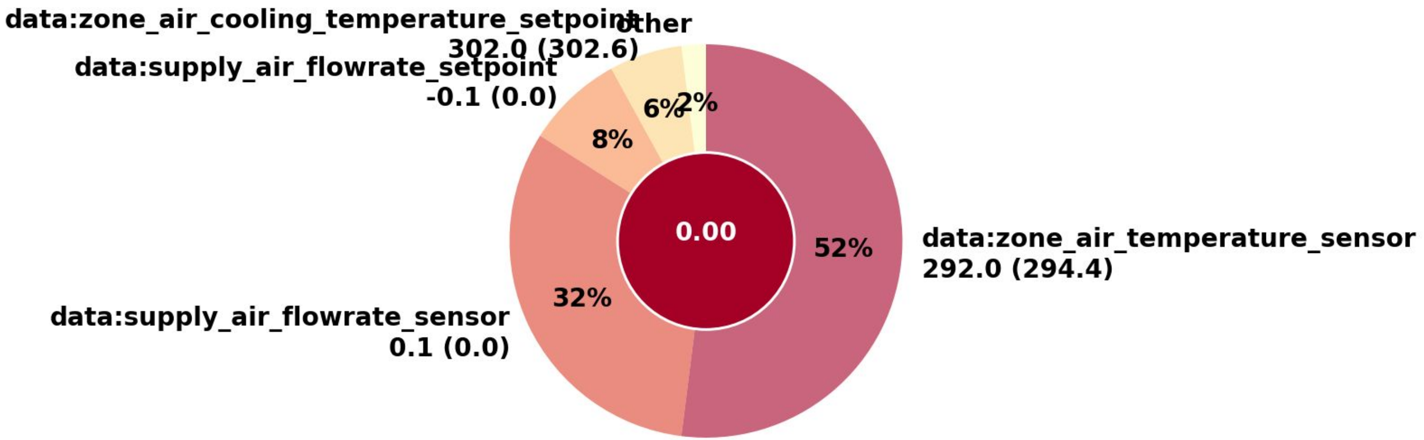
By the Completeness Axiom, the sum across all dimensions should be nearly 1

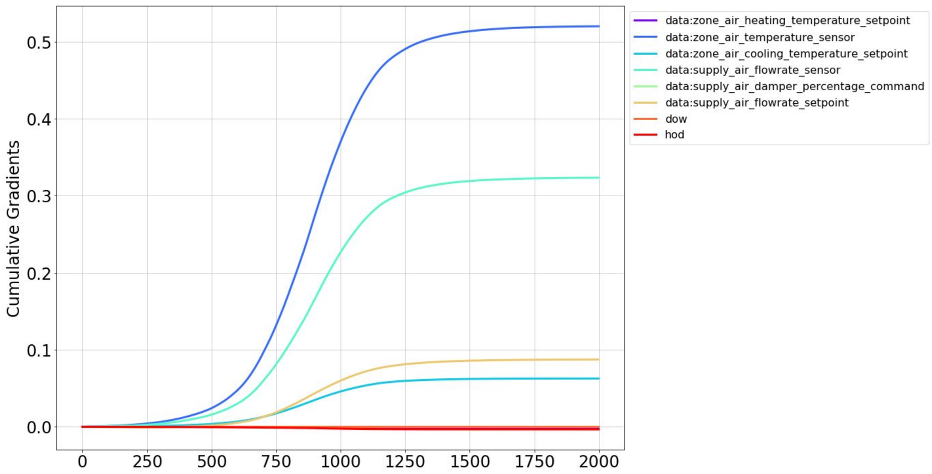
$$\sum_{d \in D} B_d(x) \approx 1$$

Each dimension d gets a proportional **blame** B_d

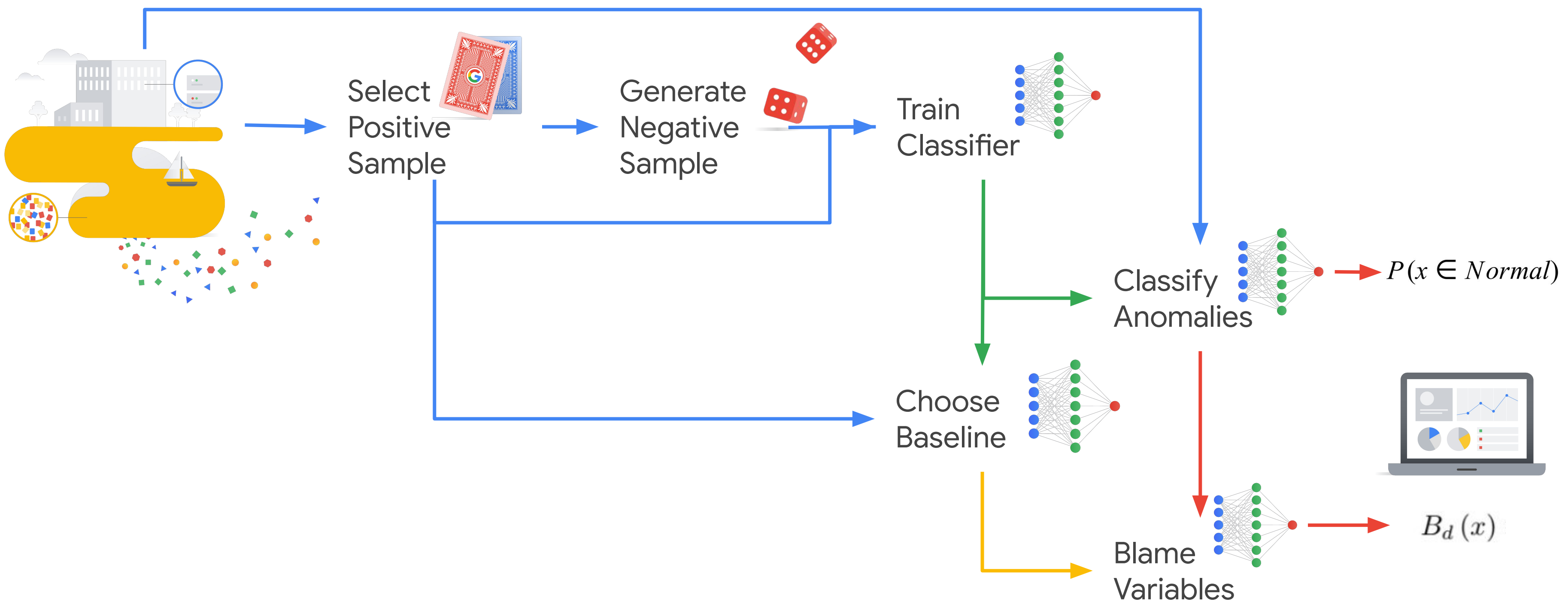

other
7%
data:zone_air_cooling_temperature_setpoint
302.0 (302.6)
26%
0.96
36%
data:supply_air_flowrate_setpoint
-0.1 (0.0)
31%
data:supply_air_flowrate_sensor
0.0 (-0.0)


Cumulative Gradients
0.012
0.010
0.008
0.006
0.004
0.002
0.000
−0.002
0
250
500
750
1000
1250
1500
1750
2000
data:zone_air_heating_temperature_setpoint
data:zone_air_temperature_sensor
data:zone_air_cooling_temperature_setpoint
data:supply_air_flowrate_sensor
data:supply_air_damper_percentage_command
data:supply_air_flowrate_setpoint
dow
hod


data:zone_air_cooling_temperature_setpoint
302.0 (302.6)
other
data:supply_air_flowrate_setpoint
-0.1 (0.0)
6%
2%
8%
0.00
52%
data:zone_air_temperature_sensor
292.0 (294.4)
32%
data:supply_air_flowrate_sensor
0.1 (0.0)


Cumulative Gradients
0.0
0.1
0.2
0.3
0.4
0.5
0
250
500
750
1000
1250
1500
1750
2000
data:zone_air_heating_temperature_setpoint
data:zone_air_temperature_sensor
data:zone_air_cooling_temperature_setpoint
data:supply_air_flowrate_sensor
data:supply_air_damper_percentage_command
data:supply_air_flowrate_setpoint
dow
hod

Anomaly Detection Pipeline with Interpretability

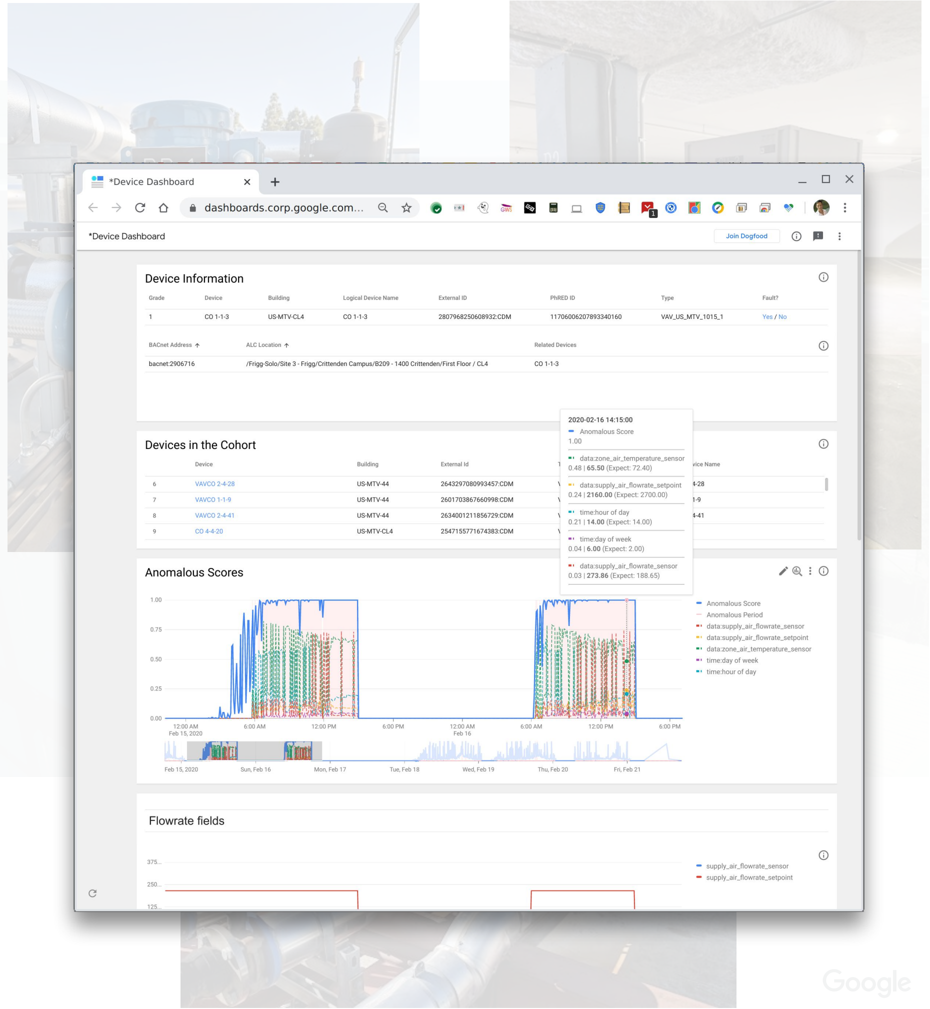


Case Study: Smart Buildings

Objective: *Make buildings smarter, secure and reduce energy use! Improve occupant comfort and productivity while also improving facilities' operation efficiencies.*

120 million measurements daily, generated by over **15,000 climate control devices, in 145 Google buildings**

Since going live in June 2019, FDD has created **458 facilities technician work orders, with a 44% True Positive rate**



Thank You

https://github.com/google/madi

Google