Practical Anomaly Detection At Scale via Self-Supervised Learning

Baris Coskun

Amazon Web Services

Observability with Cloud Computing

- Control Plane API Logs
- Network flow logs
- DNS logs
- Data Plane Logs
- Runtime Logs
- …

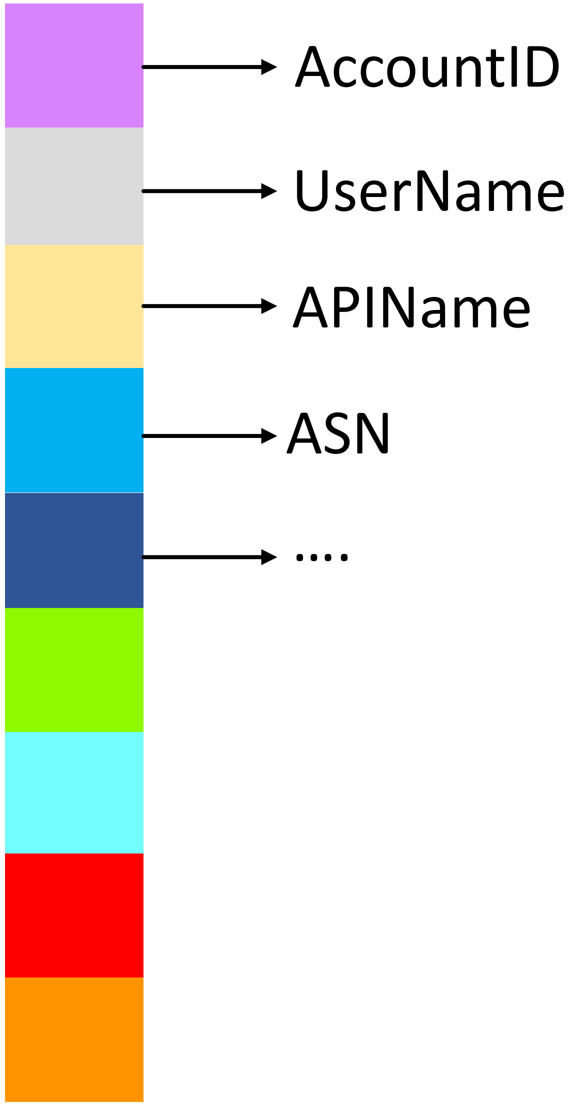
Lots of Data, No Labels

- Especially true for security use cases
- Need domain expertise for manual labeling
- Data drifts quickly
- Supervised learning often not viable

Self-supervised Learning → Anomaly Detection

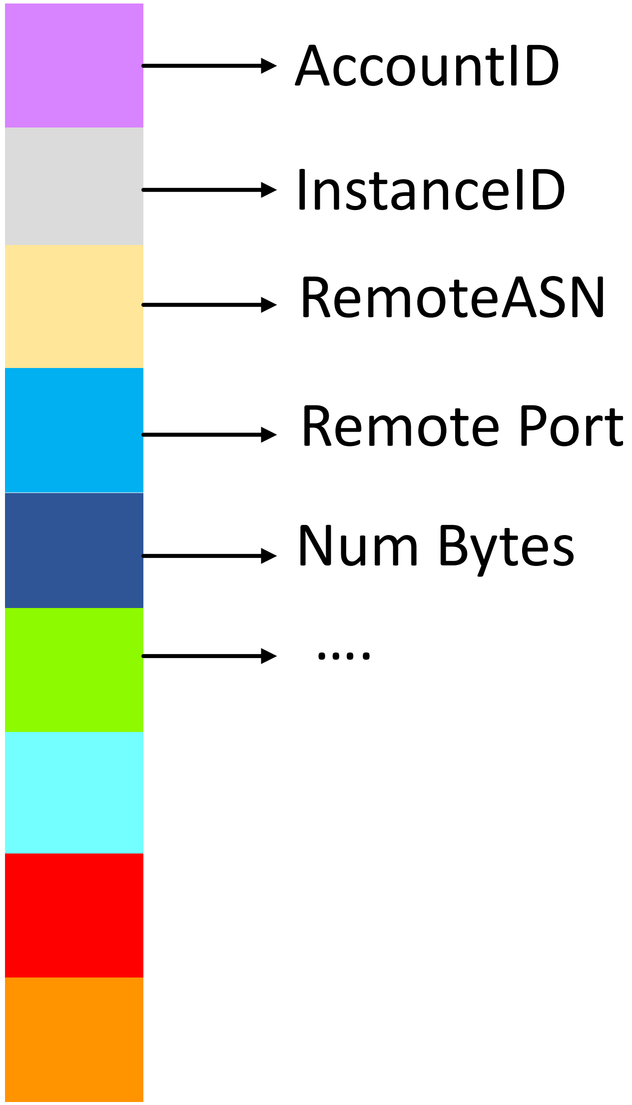
Closer Look at Data

- API Call Activity
- Network Activity
- Database Access Activity
- Compute Runtime Activity
- Container Cluster Activity
- etc.

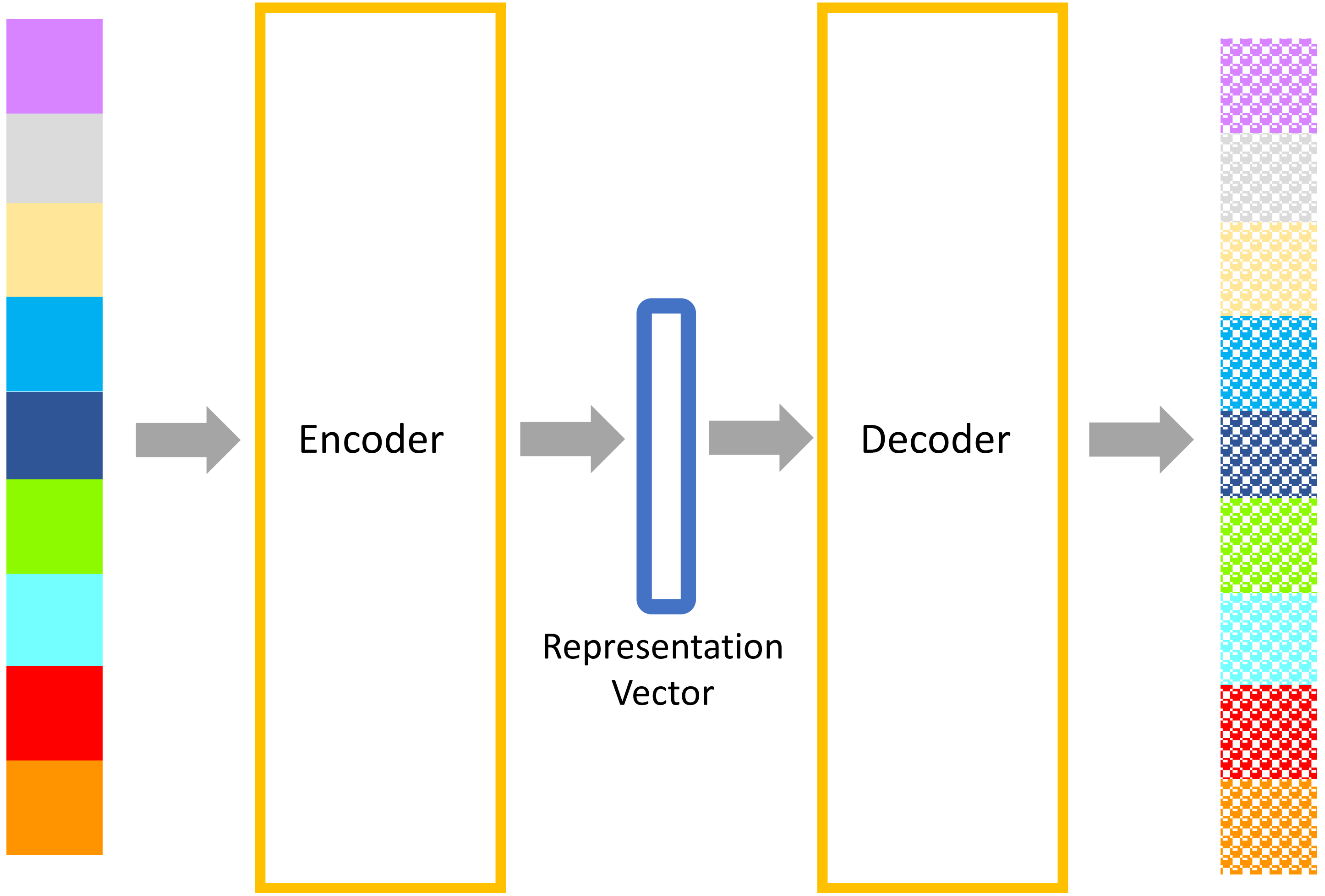


Closer Look at Data

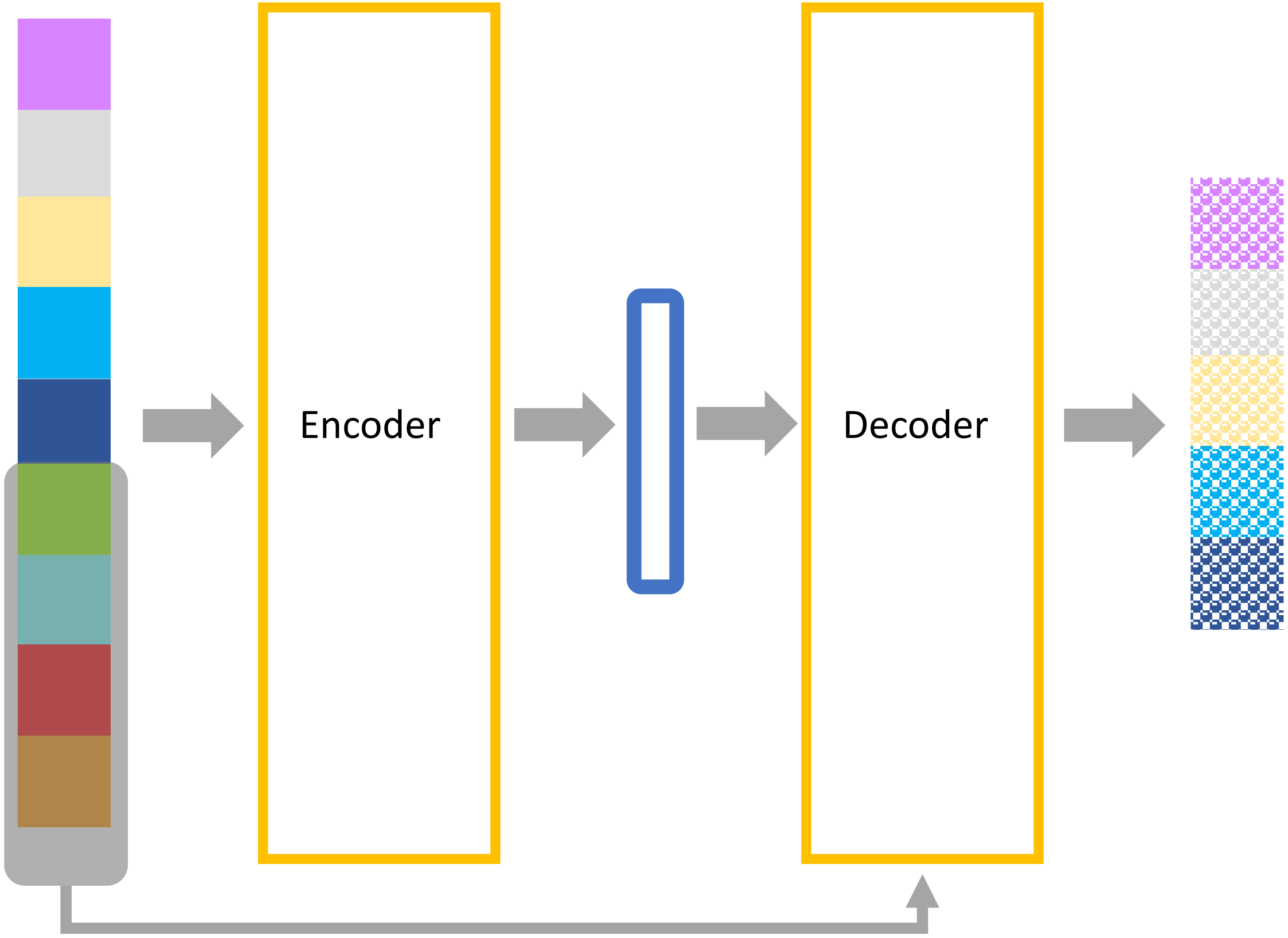
- API Call Activity
- Network Activity
- Database Access Activity
- Compute Runtime Activity
- Container Cluster Activity
- etc.



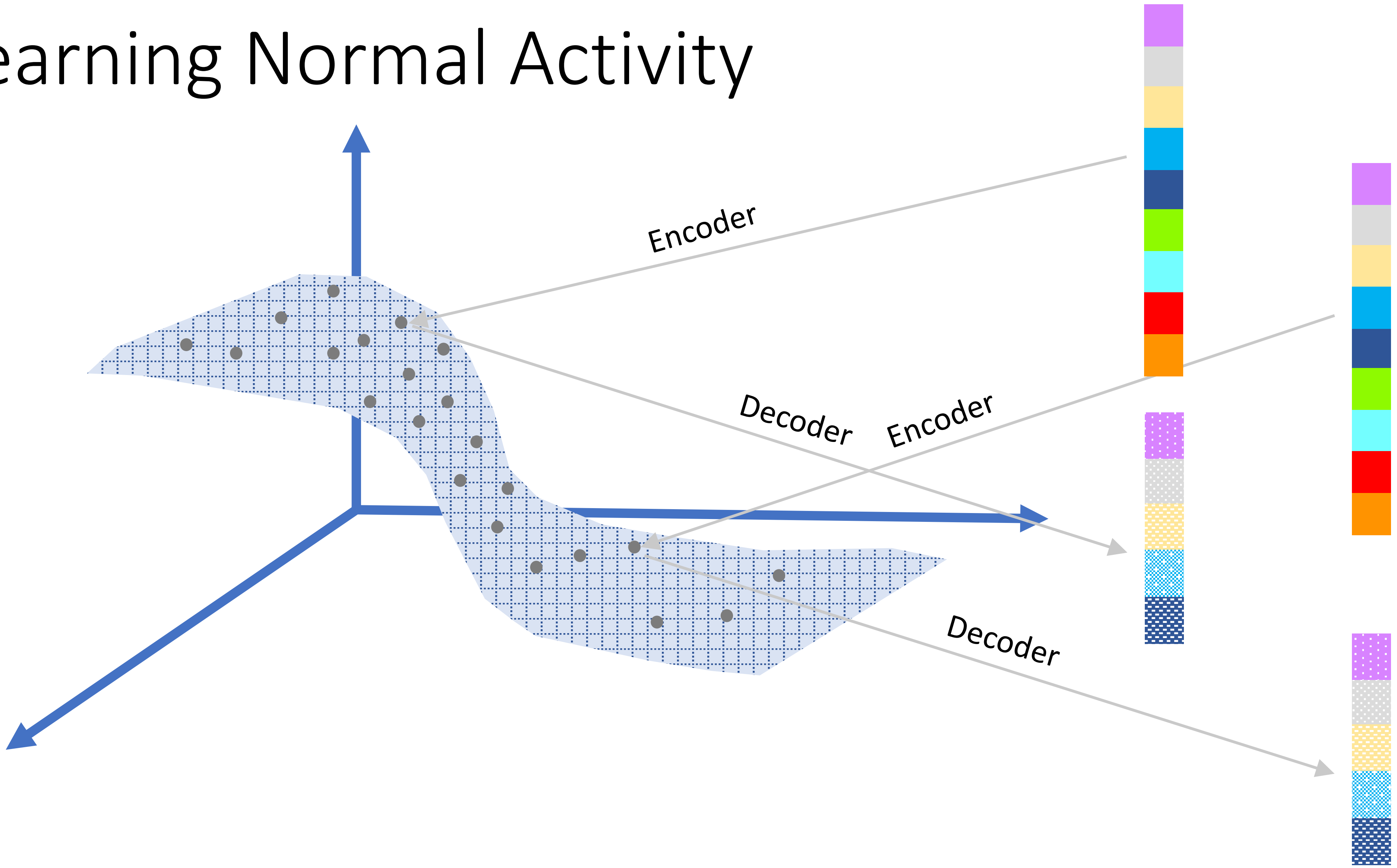
Autoencoder Model



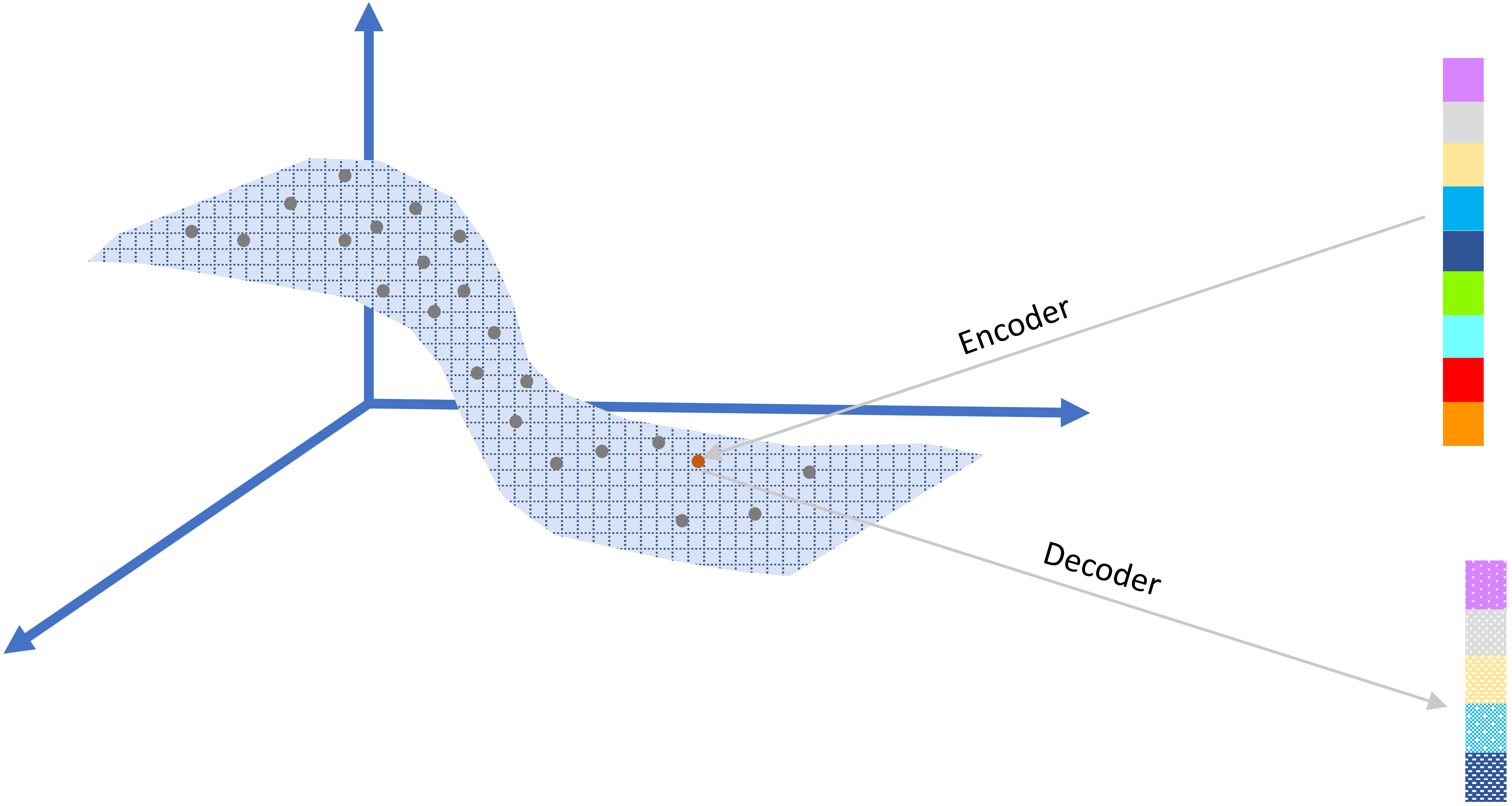
Conditional Autoencoder Model



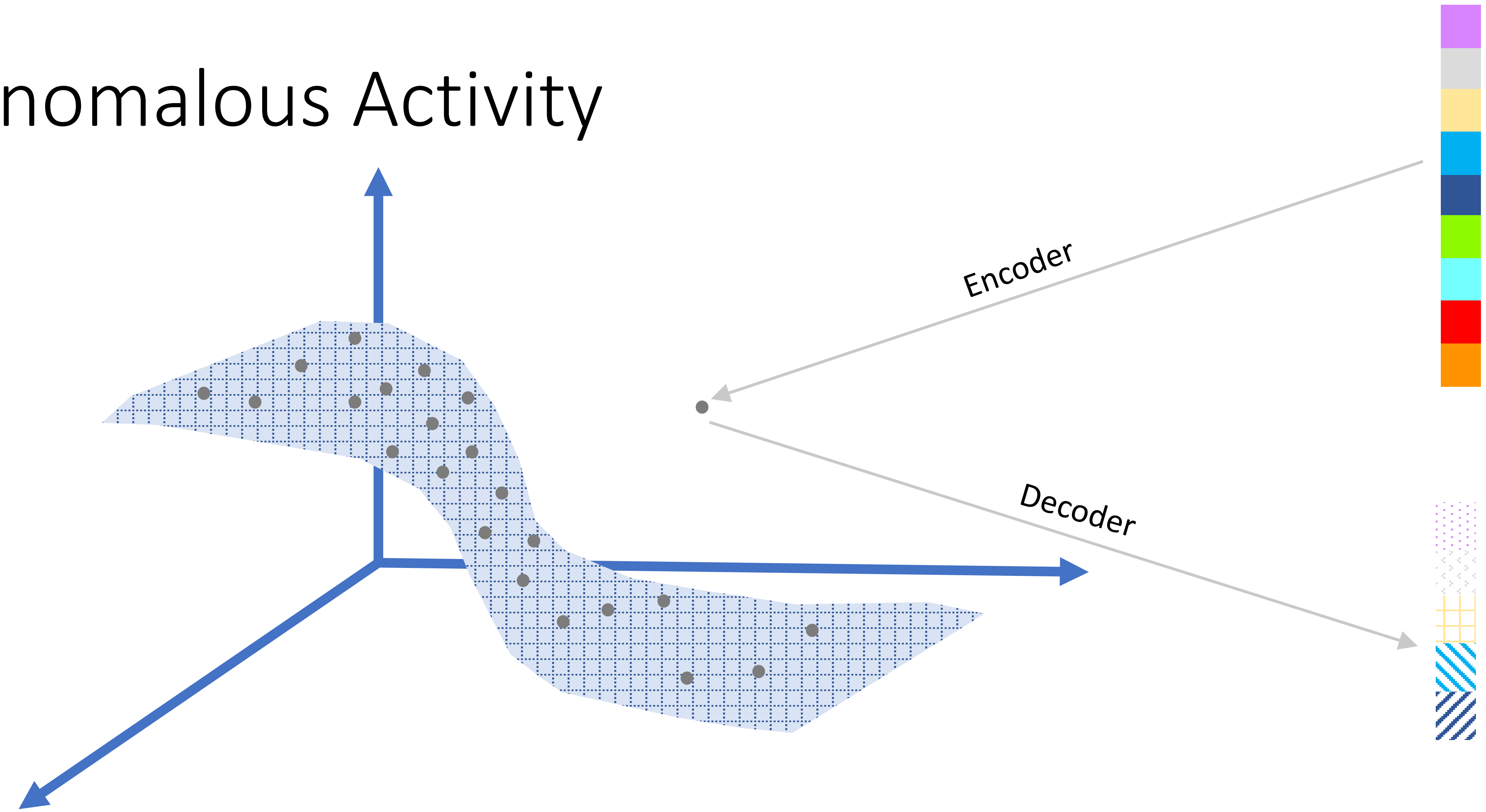
Learning Normal Activity



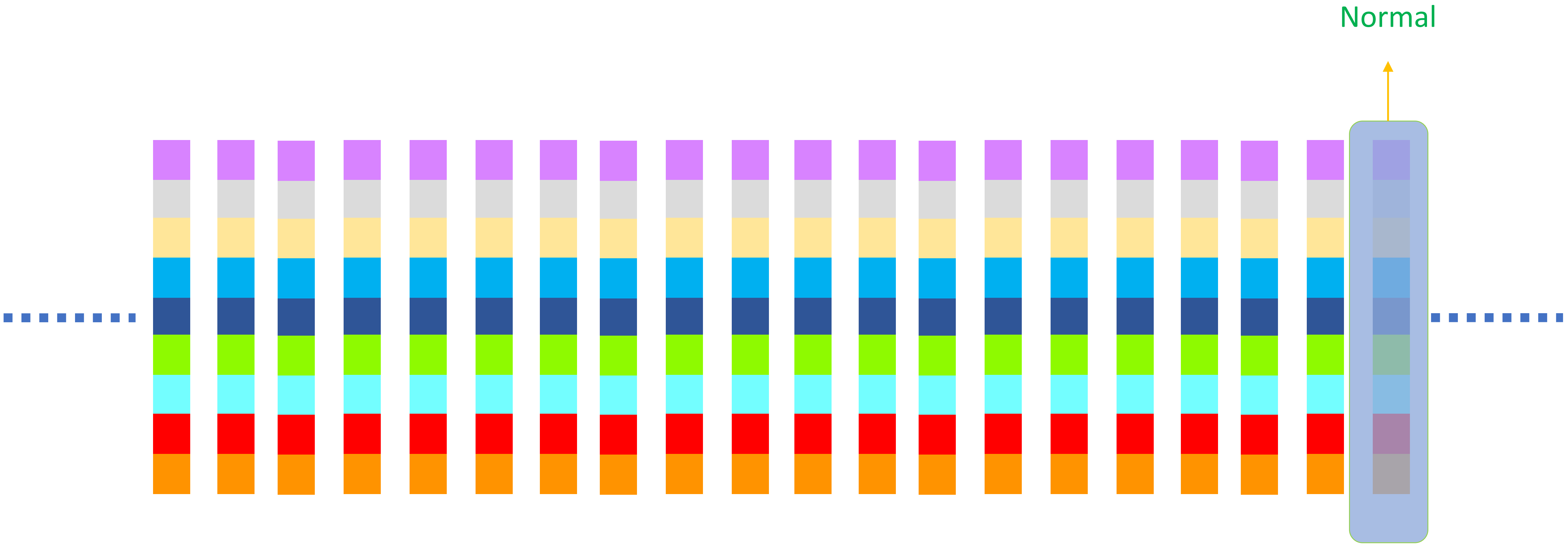
Normal Activity



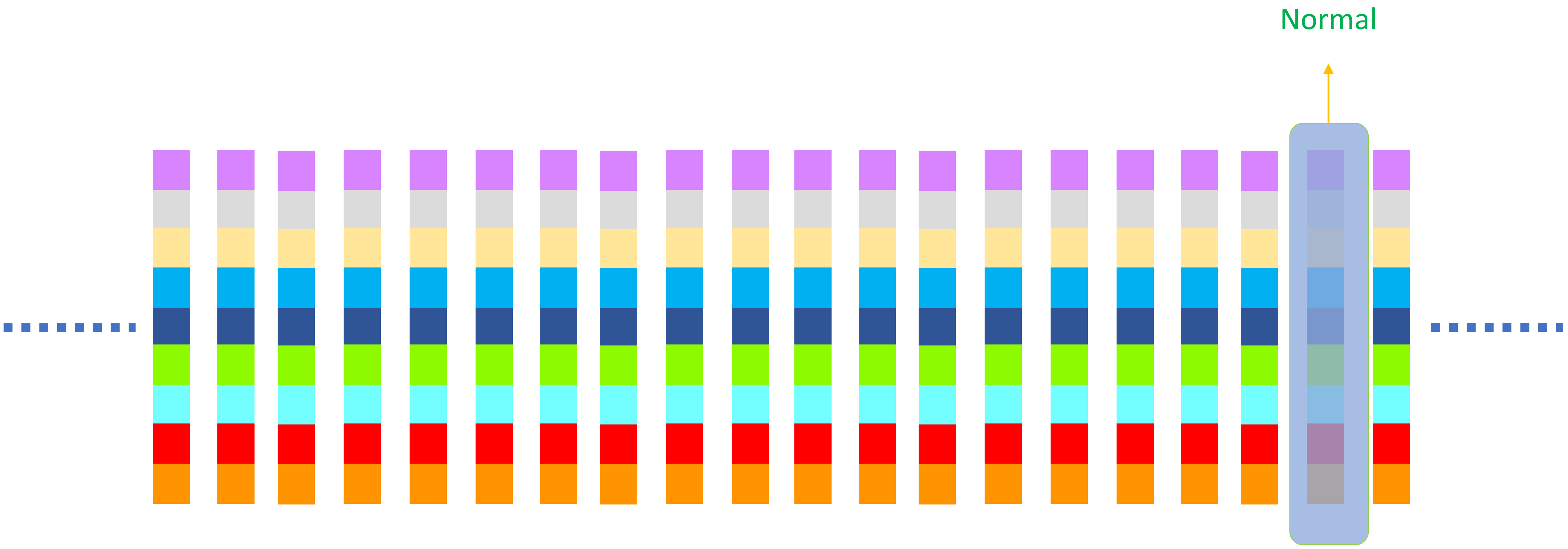
Anomalous Activity



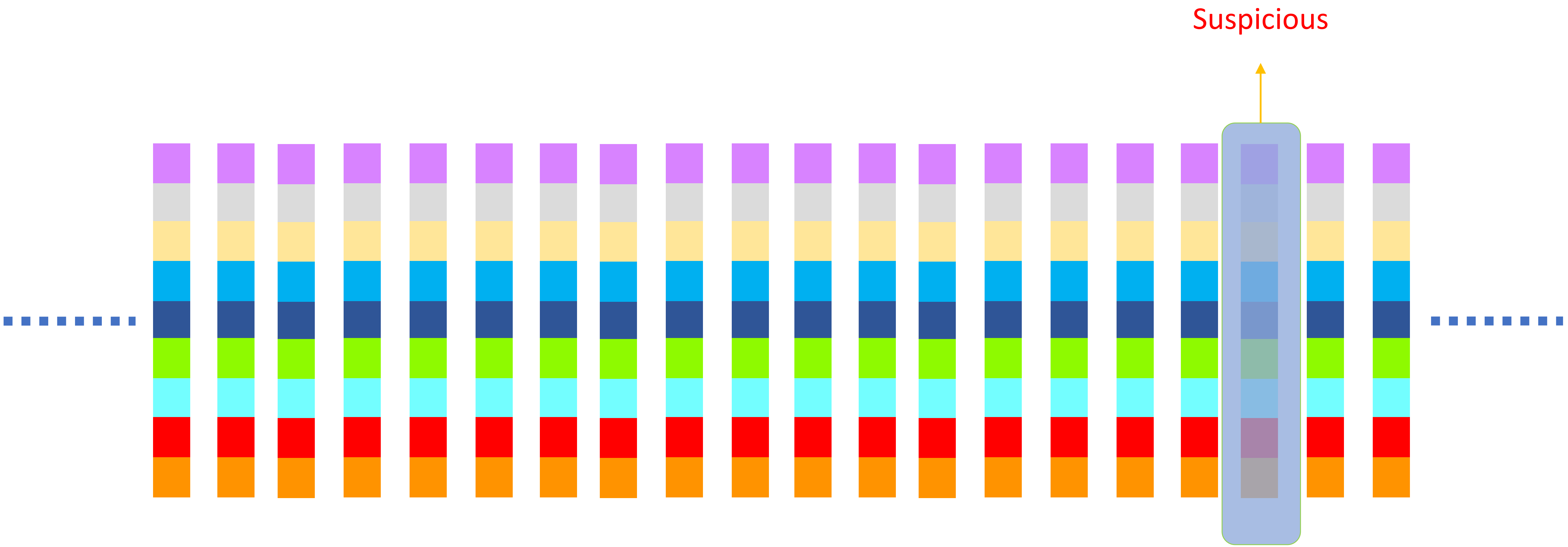
Inference



Inference



Inference



Evaluation

- Curate labeled test dataset
- Measure TPR and FPR
- P(alert|bad)~99% and P (alert|good)~0.1%
- Almost guaranteed not to work in production!

Evaluation

- Labeling is hard
- Test data distribution != production data distribution
- Production models score tens of billions of events per minute
- Searching for 1 in billions
- Need hundreds of billions of data points for accurate estimation

Security Value

Very low constant

$$P(bad|alert) = \frac{P(alert|bad)P(bad)}{P(alert)}$$

Security Value

Reducing Alert Rate

$$P(bad|alert) = \frac{P(alert|bad)P(bad)}{\boxed{P(alert)}}$$

- Run at production scale to estimate P(alert)
- Tune model hyperparameters
- Tune threshold
- Post-processing
  - Filtering
  - Aggregation

Keep Detection Rate High

$$P(bad|alert) = \frac{P(alert|bad)P(bad)}{P(alert)}$$

- Inject known attack traffic
- Inject synthetic attack
- Look for signs of TP on actual detections
  - Eyeballing
  - Automation

Firenze: Model Evaluation Using Weak Signals

Bhavna Soman, Ali Torkamani, Michael J. Morais, Jeffrey Bickford, Baris Coskun
Amazon Web Services
{bhsoman, alitor, moraismi, jbick, barisco}@amazon.com

If all goes well, launch!

Amazon GuardDuty introduces new machine learning capability to more accurately identify potentially malicious activity

Posted On: Mar 12, 2021

Amazon GuardDuty has incorporated new machine learning techniques that have proven highly effective at discerning potentially malicious user activity from anomalous, but benign operational behavior within AWS accounts. This new capability continuously models API invocations within an account, incorporating probabilistic predictions to more accurately isolate and alert on highly suspicious user behavior. This new approach has proven to identify malicious activity associated with known attack tactics, including discovery, initial access, persistence, privilege escalation, defense evasion, credential access, impact, and data exfiltration. The new threat detections are available for all existing Amazon GuardDuty customers with no action required and at no additional costs.

If all goes well, launch!

Amazon GuardDuty introduces new machine learning capability to more accurately identify potentially malicious activity

Posted On: Mar 12, 2021

Amazon GuardDuty has incorpo
malicious user activity from ano
API invocations within an accou
behavior. This new approach ha:
access, persistence, privilege esc
available for all existing Amazor

Amazon GuardDuty introduces new machine learning capabilities to more accurately detect potentially malicious access to data stored in S3 buckets

Posted On: Jul 6, 2022

Amazon GuardDuty has incorporated new machine learning techniques that are highly effective at detecting anomalous access to data stored in Amazon Simple Storage Service (Amazon S3) buckets. This new capability continuously models S3 data plane API invocations (e.g. GET, PUT, and DELETE) within an account, incorporating probabilistic predictions to more accurately alert on highly suspicious user access to data stored in S3 buckets, such as requests coming from an unusual geo-location, or unusually high volumes of API calls consistent with attempts to exfiltrate data. The new machine learning approach can more accurately identify malicious activity associated with known attack tactics, including data discovery, tampering, and exfiltration. The new threat detections are available for all existing Amazon GuardDuty customers that have GuardDuty S3 Protection enabled, with no action required and at no additional costs. If you are not using GuardDuty yet, S3 protection will be on by default when you enable the service. If you are using GuardDuty, and are yet to enable S3 Protection, you can enable this capability organization-wide with one-click in the GuardDuty console or through the API.

If all goes well, launch!

Amazon GuardDuty introduces new machine learning capability to more accurately identify potentially malicious activity

Posted On: Mar 12, 2021

Amazon GuardDuty has incorpo
malicious user activity from ano
API invocations within an accou
behavior. This new approach ha:
access, persistence, privilege esc
available for all existing Amazor

Amazon GuardDuty introduces new machine learning capabilities to more accurately detect potentially malicious access to data stored in S3 buckets

Posted On: Jul 6, 2022

Amazon GuardDuty has incorporated ne
stored in Amazon Simple Storage Servic
(e.g. GET, PUT, and DELETE) within an ac
access to data stored in S3 buckets, such
consistent with attempts to exfiltrate da
associated with known attack tactics, in
all existing Amazon GuardDuty custome
costs. If you are not using GuardDuty ye
and are yet to enable S3 Protection, you
through the API.

Amazon GuardDuty introduces new machine learning capability to enhance threat detection for Amazon EKS detections

Posted On: Nov 9, 2023

Amazon GuardDuty has incorporated new machine learning techniques to more accurately detect anomalous activities indicative of threats to your Amazon Elastic Kubernetes Service (Amazon EKS) clusters. This new capability continuously models Kubernetes audit log events from Amazon EKS to detect highly suspicious activity such as unusual user access to Kubernetes secrets that can be used to escalate privileges, and suspicious container deployments with images not commonly used in the cluster or account. The new threat detections are available for all GuardDuty customers that have GuardDuty EKS Audit Log Monitoring enabled.

Takeaways

- Security has become a data problem
- Lack of labels leads to self-supervised appraches
- Evaluation is as important as modeling
- Evaluation is as hard as modeling
- Need to rethink evaluation

Thank You!