CYBER ELITE

BCG
BENCHACHINDA GROUP
SINCE 1960

CYBER ELITE

AI Adoption in Cyber Defense

10 June 2023

Agenda

- CBE Service Portfolio
- Global AI Adoption
- Defensive Cybersecurity and Key Issues
- AI Adoption in Cyber Defense Use Cases
  - UEBA (User and Entity Behavior Analytics)
  - NLP (Natural Language Processing)
- Key Takeaways

กว่า 60 ปี ที่ผ่านมา กลุ่มเบญจจินดา (Benchachinda Group “BCG”) มุ่งมั่นที่จะพัฒนาโครงสร้างพื้นฐานด้านโทรคมนาคม และดิจิทัลเทคโนโลยี ของประเทศไทยให้เป็นระดับแนวหน้าของภูมิภาคอาเซียน เพื่อยกระดับ คุณภาพ และเพิ่มขีดความสามารถในการแข่งขันขององค์กรไทย

ปัจจุบัน บีซีจี ดำเนินธุรกิจแบ่งออกเป็น 4 กลุ่มธุรกิจ ได้แก่
(1) กลุ่มธุรกิจ Digital Infrastructure and Solution Business
(2) กลุ่มธุรกิจ Distribution and Fulfillment Business
(3) กลุ่มธุรกิจ Content Business
(4) กลุ่มธุรกิจ Investment Business

BCG
BENCHACHINDA GROUP

DIGITAL INFRASTRUCTURE AND SOLUTION BUSINESS

UIH · CLOUD by UIH · UTEL · BB TEC · UTE · CYBER ELITE · brainerey

DISTRIBUTION AND FULFILLMENT BUSINESS

YAS · UD

CONTENT BUSINESS

inn · Rakbankerd

INVESTMENT BUSINESS

BCH

CYBER ELITE

บริษัท ไซเบอร์ อีลีท จำกัด (CYBER ELITE)

มุ่งมั่นในการเป็นผู้นำบริการด้านการรักษาความปลอดภัยไซเบอร์ที่ครบครันในระดับภูมิภาคอาเซียน มีความน่าเชื่อถือได้มาตรฐานระดับโลก มีผู้เชี่ยวชาญร่วมเฝ้าระวัง และรับมือภัยไซเบอร์ตลอด 24 ชั่วโมง CYBER ELITE ให้บริการครอบคลุม System Integration, Security Advisory, Managed Security Services, Cybersecurity Platform และ Training & Awareness

"We simplify the way you build trust and resilience in cyberspace"

Contact

บริษัท ไซเบอร์ อีลีท จำกัด
499 Benchachinda Bldg., Kamphaeng Phet 6 Rd.,
Ladyao, Chatuchak, Bangkok 10900.
www.cyberelite.co
Cyber Elite
mkt@cyberelite.co
02 016 5555

OUR ELITE TEAM

Our team has over 20 years of experience in cybersecurity as consultants, implementers, advisors, instructors, researchers, and service providers in various industries.

Financial	Telecom	Insurance
Retail	Healthcare	Government
Energy	Defense	National CERT
Leasing	Manufacuring	Entertainment

OUR CERTIFICATIONS

CISSP | CSSLP | CISA | CISM | CDPSE | CDPO | CCISO | GIAC GWAPT | ECSA | CEH | CHFI | ECES | ENSA | CSCU | CEI | CSIE | CSAE | CASP+ | CySA+ | Security+ | Pentest+ | Network+ | CTT+ | CNVP | CSAP | CNSP | IRCA ISO27001 PA | PECB ISO27001 SLI | PECB ISO27001 LI | PECB ISO22301 PI | PECB ISO31000 RM | ITILv3 Foundation

Our Security Operations Center (SOC) is 27001 and 27701 certified

CYBER ELITE

Stay in touch !

CYBER ELITE

"The Most Trusted and Supportive Cybersecurity Company"

We offer full range of end-to-end cybersecurity products and services, designed and tailer-made to fit each organization cybersecurity context and exposure

No "*one size fits all*" in cybersecurity

01 SECURITY SYSTEM INTEGRATION

02 MANAGED SECURITY SERVICES

03 SECURITY ADVISORY

04 TRAINING & AWARENESS

05 CYBERSECURITY PLATFORMS

CYBER ELITE

CYBER ELITE

Products, Solutions, and Services Highlights

Pre-defined custom solutions and services for each sector

Security Solutions

- Cloud Security
- PDPA
- Cyber Incident Response
- E-Insurance
- CRAF
- Cyber Hygiene
- Cyber Act
- Zero Trust
- Cyber Threat Intelligence
- IT Third Party Security
- Network Security
- OT Security

Security Advisory

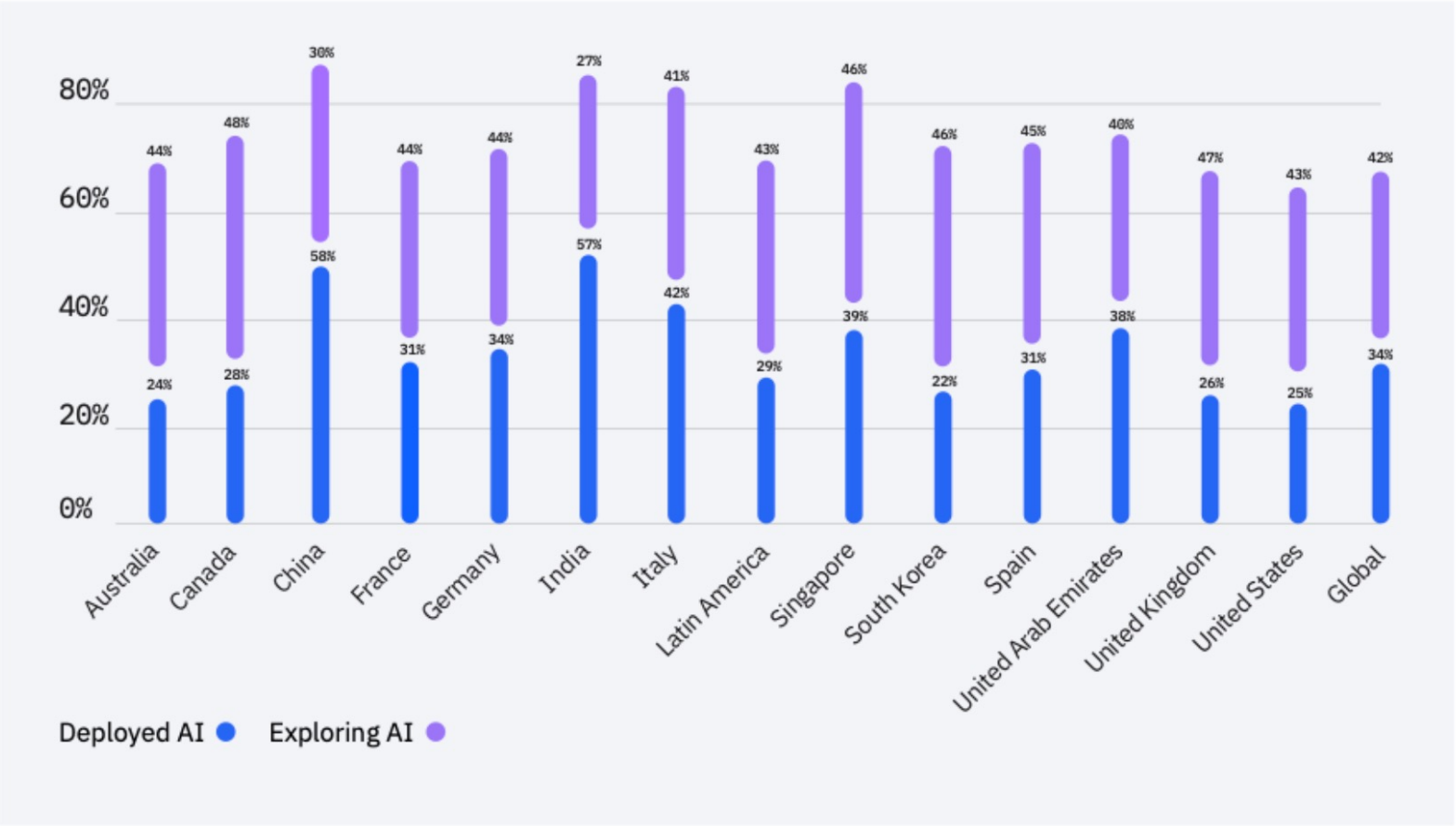
- Cyber Maturity Assessment
- Security Readiness Advisory
- Regulatory Compliance Advisory
- Security Assessment
- Security Infrastructure Review
- Security Hardening
- Cloud Security
- ISO27001 Consulting
- Cyber Drill and IRP Advisory

Managed Security Services

- Managed Cybersecurity Program (vCISO)
- Managed Network Security
- Managed CSOC
- Managed Vulnerability
- Managed Cyber Threat Intelligence
- Managed Cloud Gateway Security
- Managed Data Loss Prevention (DLP)
- Managed Data Tokenization

CYBER ELITE
Global AI Adoption

Global AI Adoption Index 2022



AI adoption rates around the world

https://www.ibm.com/downloads/cas/GVAGA3JP

Top 3 Benefits

1. Automation, Cost saving (54%)
2. Improvement in IT performance (53%)
3. Better experiences for customers (48%)

Top 5 Barriers to AI adoption

1. AI skills, expertise or knowledge (34%)
2. Price is too high (29%)
3. Lack of tools or platforms to develop models (25%)
4. Projects are too complex or difficult to integrate and scale (24%)
5. Too much data complexity (24%)

CYBER ELITE

Defensive cybersecurity

- Proactively attempting to prevent cyber attacks
- Reactively attempting to identify, block, and mitigate ongoing attacks

Key Issues	Use Cases	ML Algorithms	Proactive/ Reactive
Threat detection (False Positive, False Negative, MTTD)	?	?	?
Threat response time (MTTR)	?	?	?
New threat identification (Zero-day Attack)			
Staffing capacity and expertise	?	?	?
Large volume of cyber alerts			
How to manage?	?	?	?

CYBER ELITE

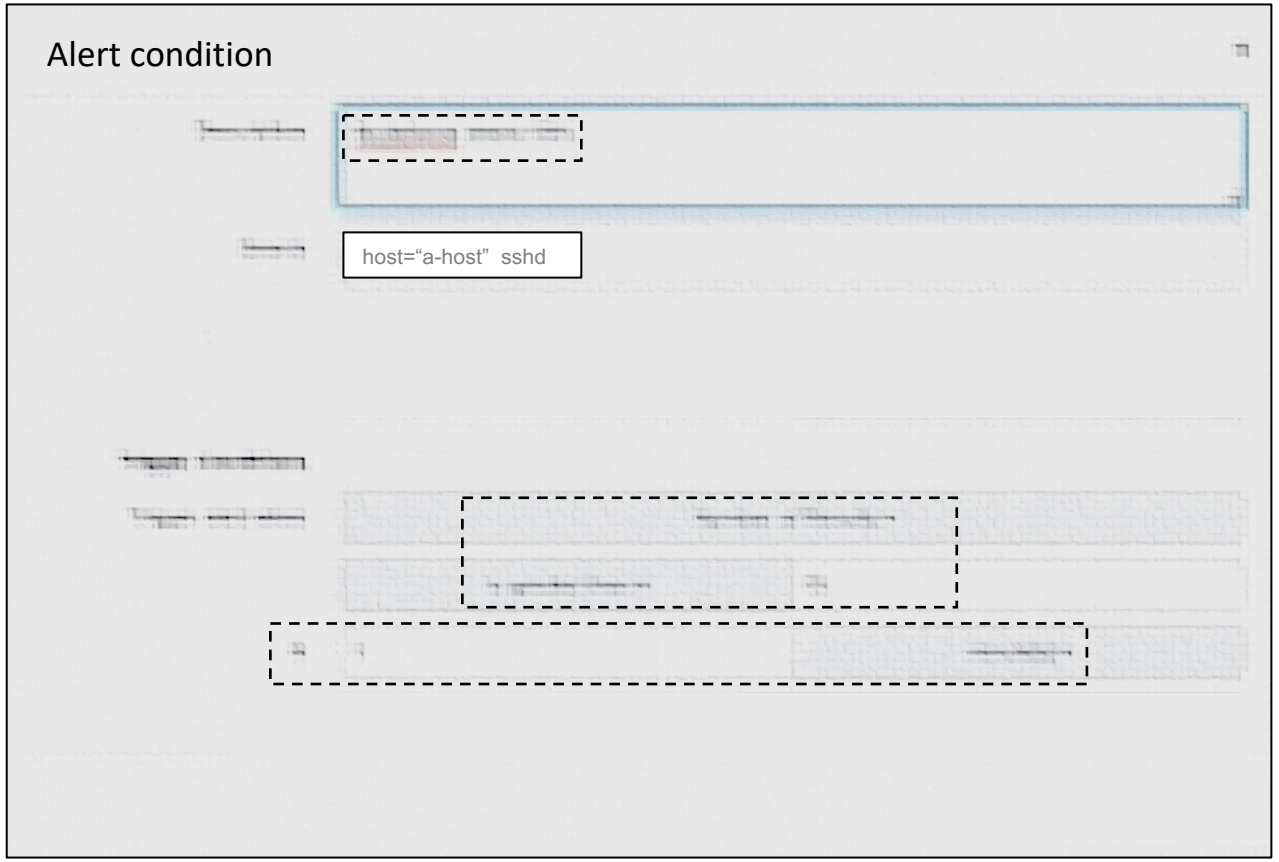
AI Adoption in Cyber Defense Use Cases

UEBA with Isolation Forest

CYBER ELITE

Use Cases : UEBA

Rule based alert (Example) (just for study)



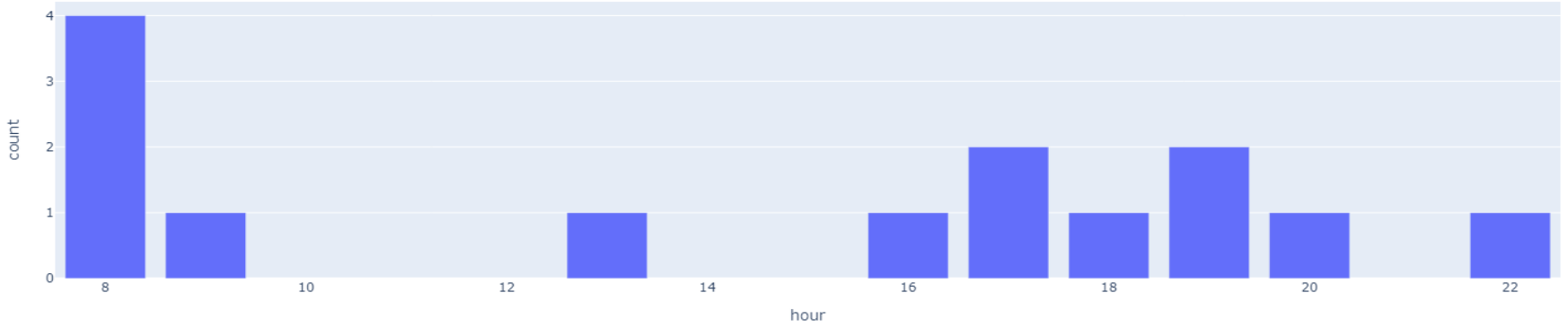
CYBER ELITE

Use Cases : UEBA

Anomaly Detection: Isolation forest
Example on actual login data

	_source.reason	_source.RemoteIPGeo.city_name	_source.RemoteIPGeo.country_name	_source.RemoteIPGeo.location.lon	_source.RemoteIPGeo.location.lat	_source.@timestamp	_source.user
399	sslvpn_login_permission_denied	Bangkok	Thailand	10	13	2023-03-09 20:52:36.088000+00:00	
227	sslvpn_login_permission_denied	Bangkok	Thailand	10	13	2023-04-10 09:13:31.008000+00:00	
427	sslvpn_login_permission_denied	Bangkok	Thailand	10	13	2023-04-03 22:24:28.045000+00:00	
42	sslvpn_login_permission_denied	Nonthaburi	Thailand	10	13	2023-03-10 08:26:03.372000+00:00	
30	sslvpn_login_permission_denied	Nonthaburi	Thailand	10	13	2023-03-27 08:37:08.362000+00:00	
375	sslvpn_login_permission_denied	Nonthaburi	Thailand	10	13	2023-03-30 17:58:09.062000+00:00	
333	sslvpn_login_permission_denied	Nonthaburi	Thailand	10	13	2023-03-31 16:01:15.544000+00:00	

	hour	count
0	8	4
1	9	1
2	13	1
3	16	1
4	17	2
5	18	1
6	19	2
7	20	1
8	22	1

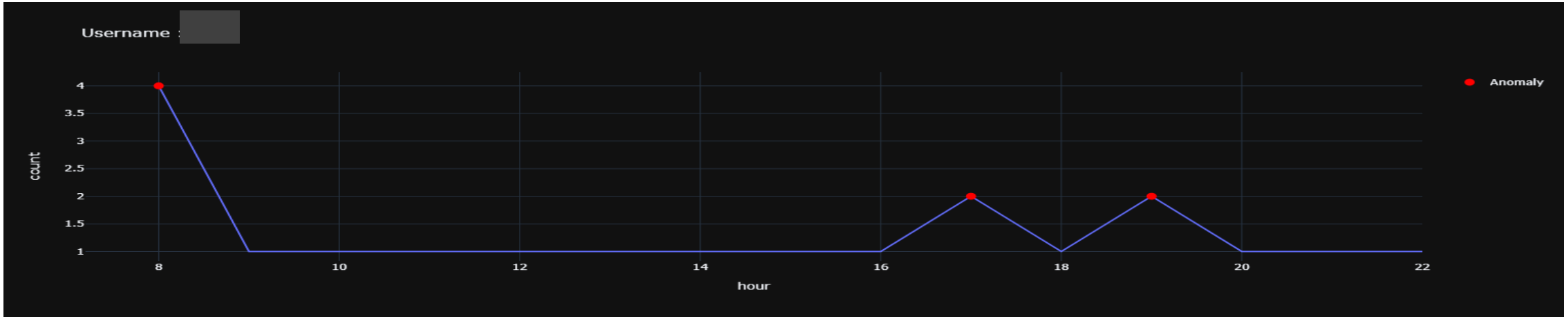


CYBER ELITE

CYBER ELITE

Use Cases : UEBA

Anomaly Detection: Isolation forest
Result



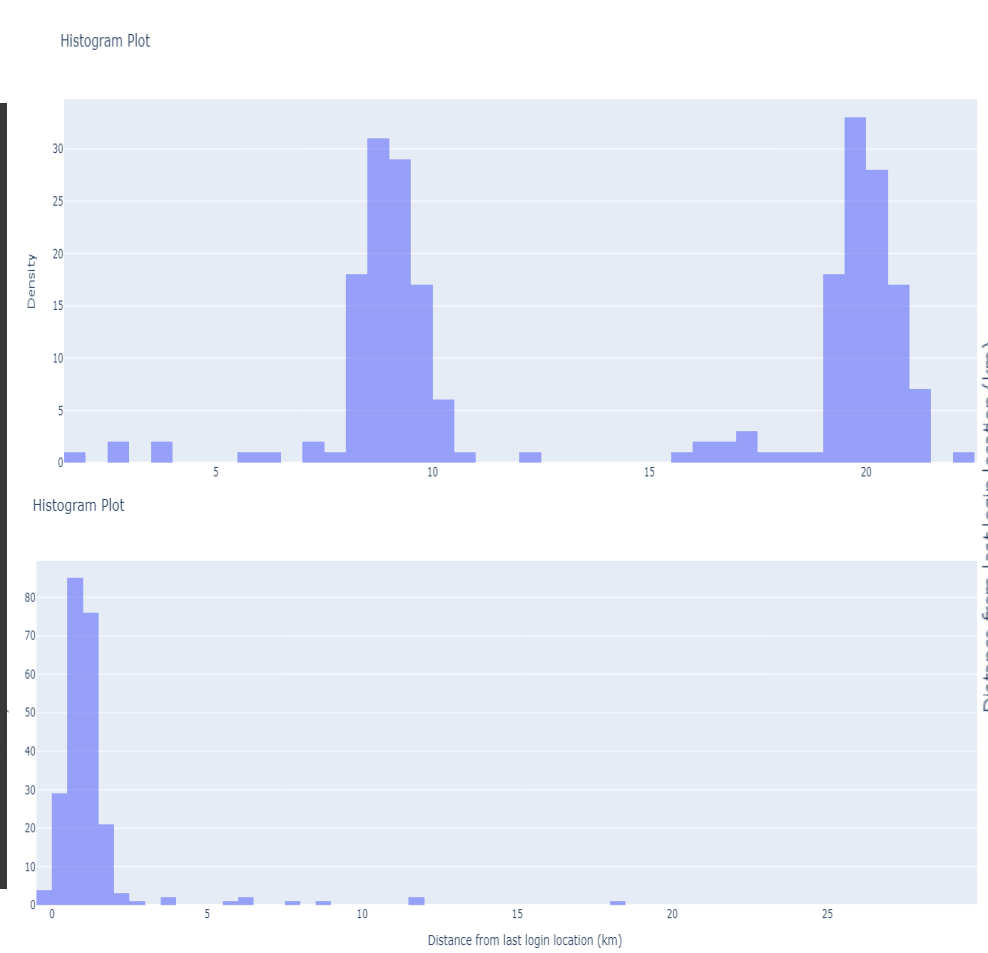
CYBER ELITE

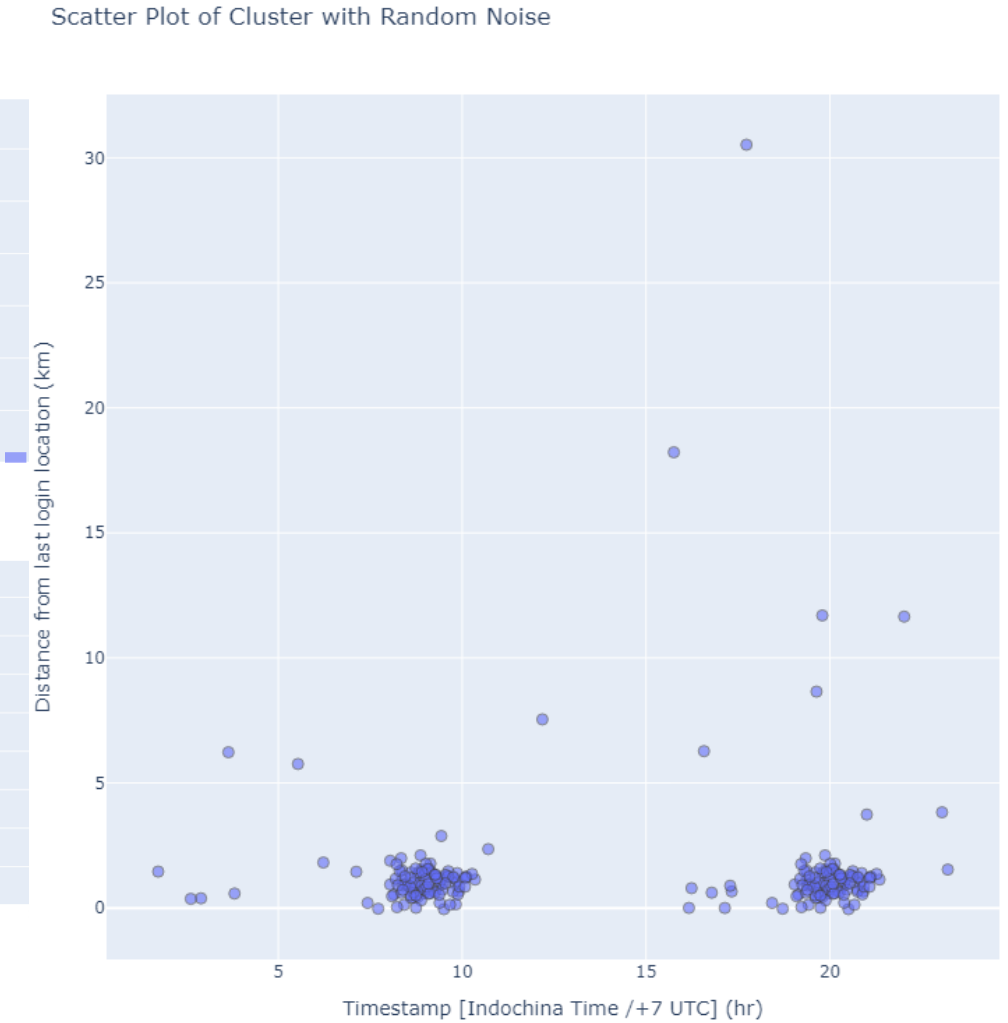
Use Cases : UEBA

- **Anomaly Detection: Isolation forest (2 dimensions)**

Let’s try to use **Isolation forest** use to on Login log to detect anomalies.

	Timestamp	Distance_diff
0	15.756032	18.223893
1	12.181968	7.542349
2	17.142040	0.000713
3	2.898547	0.387280
4	5.534569	5.758747
...	...	...
225	20.305775	1.180890
226	19.743112	0.496461
227	20.272761	1.334658
228	19.216450	1.754423
229	19.227055	0.033989



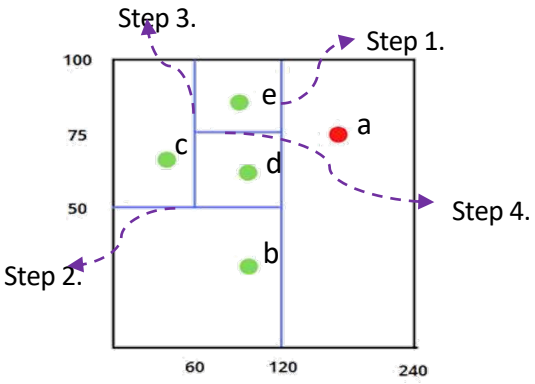


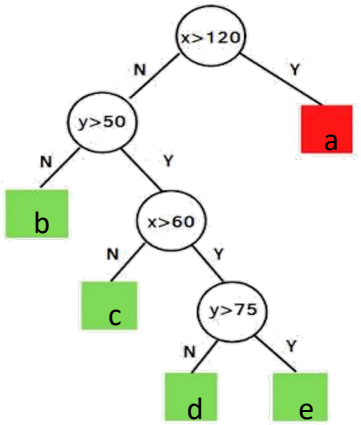
Use Cases : UEBA

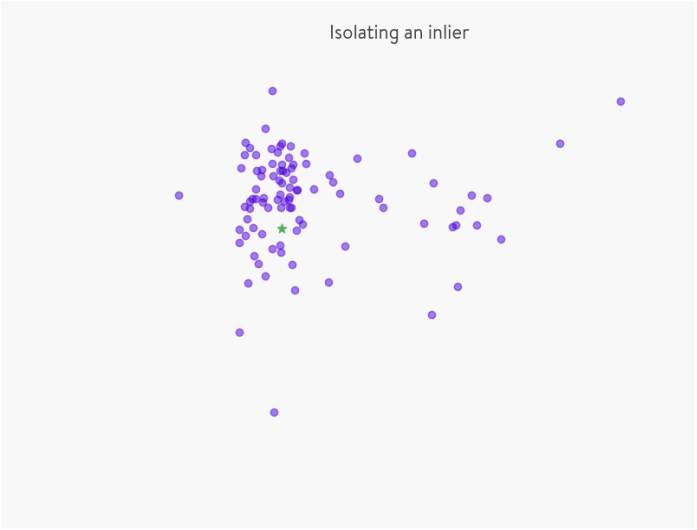
- **Anomaly Detection: Isolation forest**

The premise behind Isolation forest is that Anomalies easier to separate from the rest of the sample.

In other words, when construct a decision tree using randomly selects a feature and a split value, on average an anomalies will be isolated much sooner than normal data points.





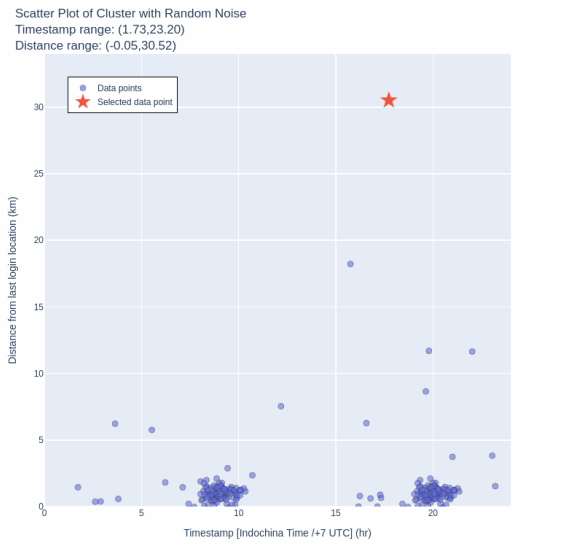


Use Cases : UEBA

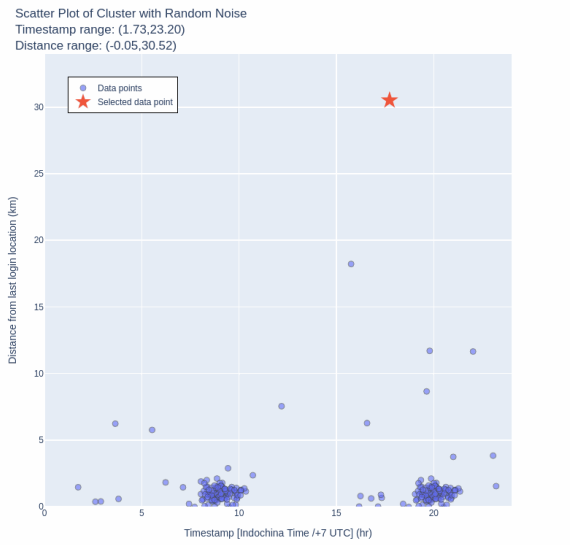
- **Anomaly Detection: Isolation forest**

Let’s try to use **Isolation forest** use to on Login log to detect anomalies.

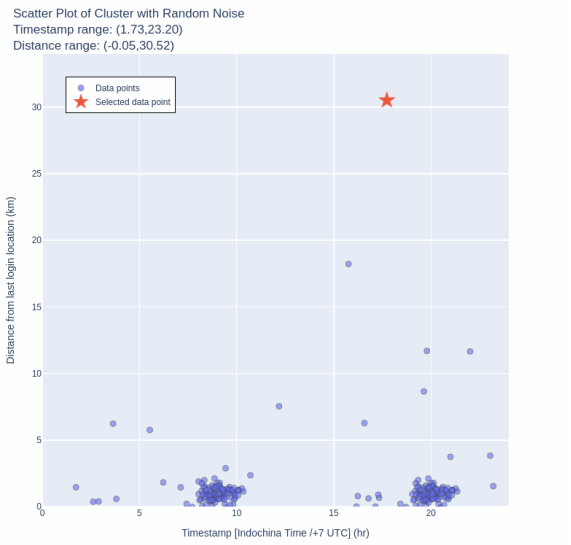
Do those steps multiple time then get average isolation number



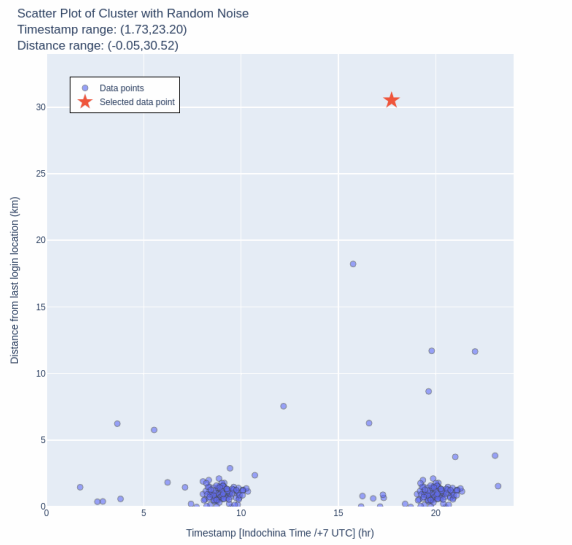
isolation number: 3



isolation number: 3



isolation number: 4



isolation number: 5

average isolation number = $\frac{3+3+4+5}{4} = 3.75$

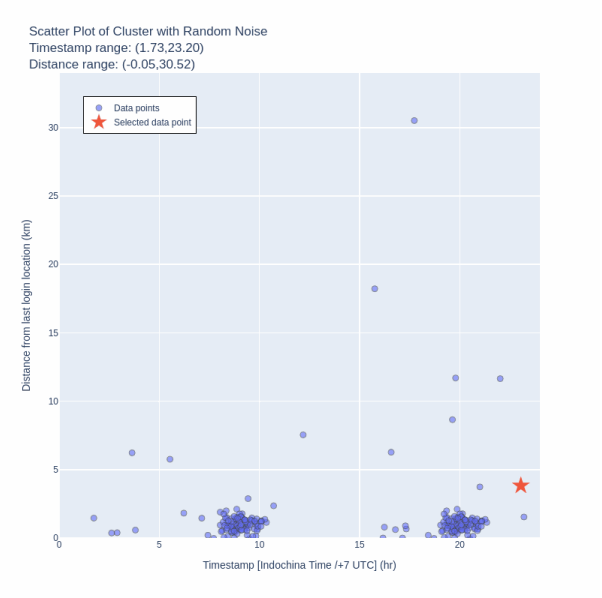
CYBER ELITE

Use Cases : UEBA

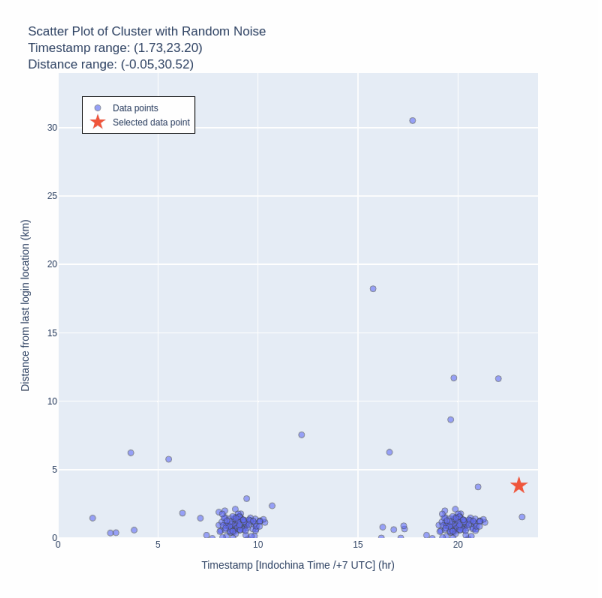
- **Anomaly Detection: Isolation forest**

Let’s try to use **Isolation forest** use to on Login log to detect anomalies.

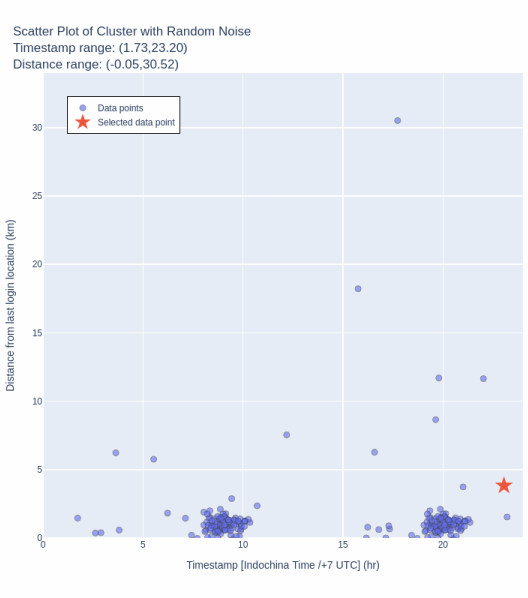
Do those step on all data points



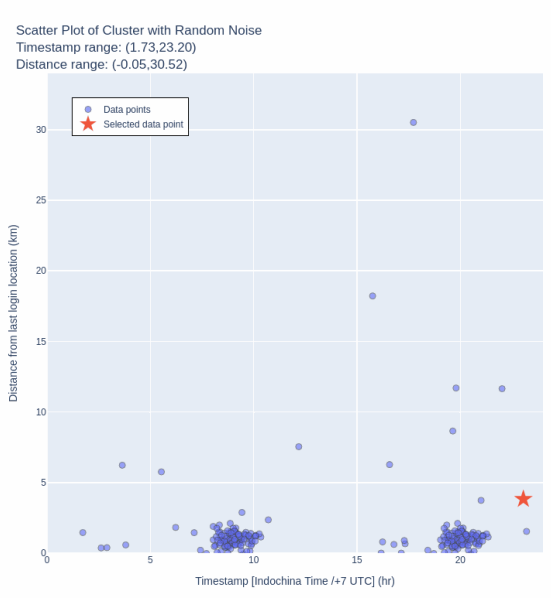
isolation number: 11



isolation number: 8



isolation number: 4



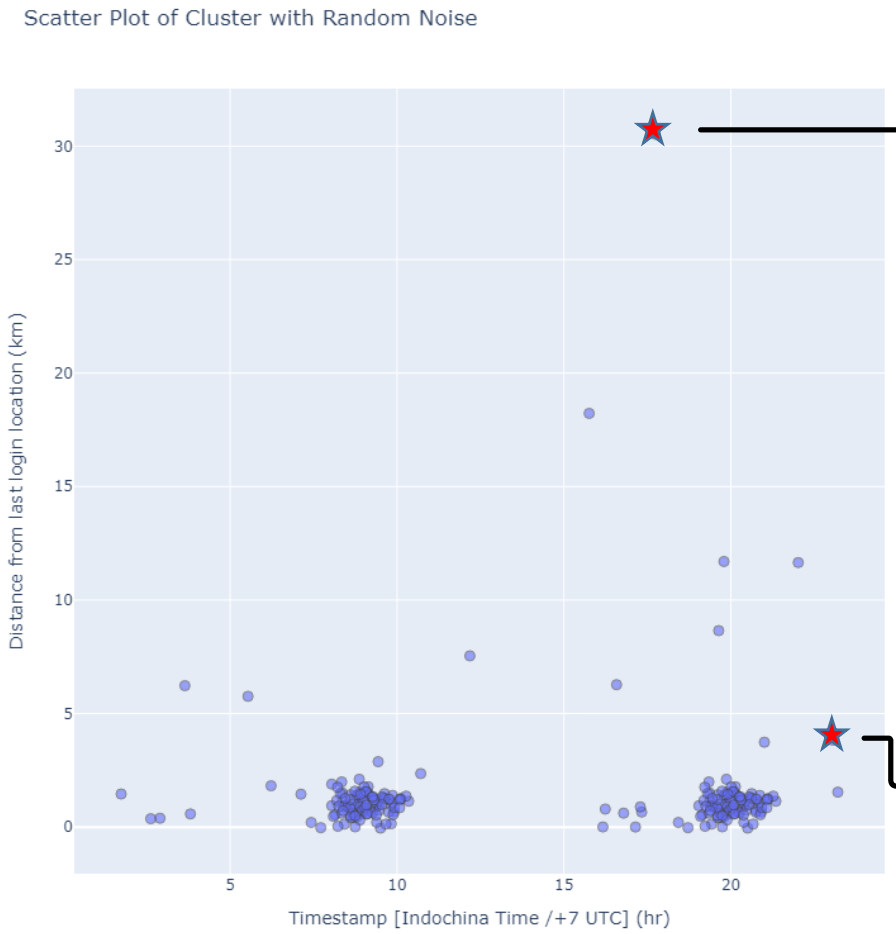
isolation number: 14

$$\text{average isolation number} = \frac{11+8+4+14}{4} = 9.25$$

CYBER ELITE

Use Cases : UEBA

- **Anomaly Detection: Isolation forest**



average isolation number $= 3.75$

** Anomalies data will have average isolation number lower than regular data

average isolation number $= 9.25$

CYBER ELITE

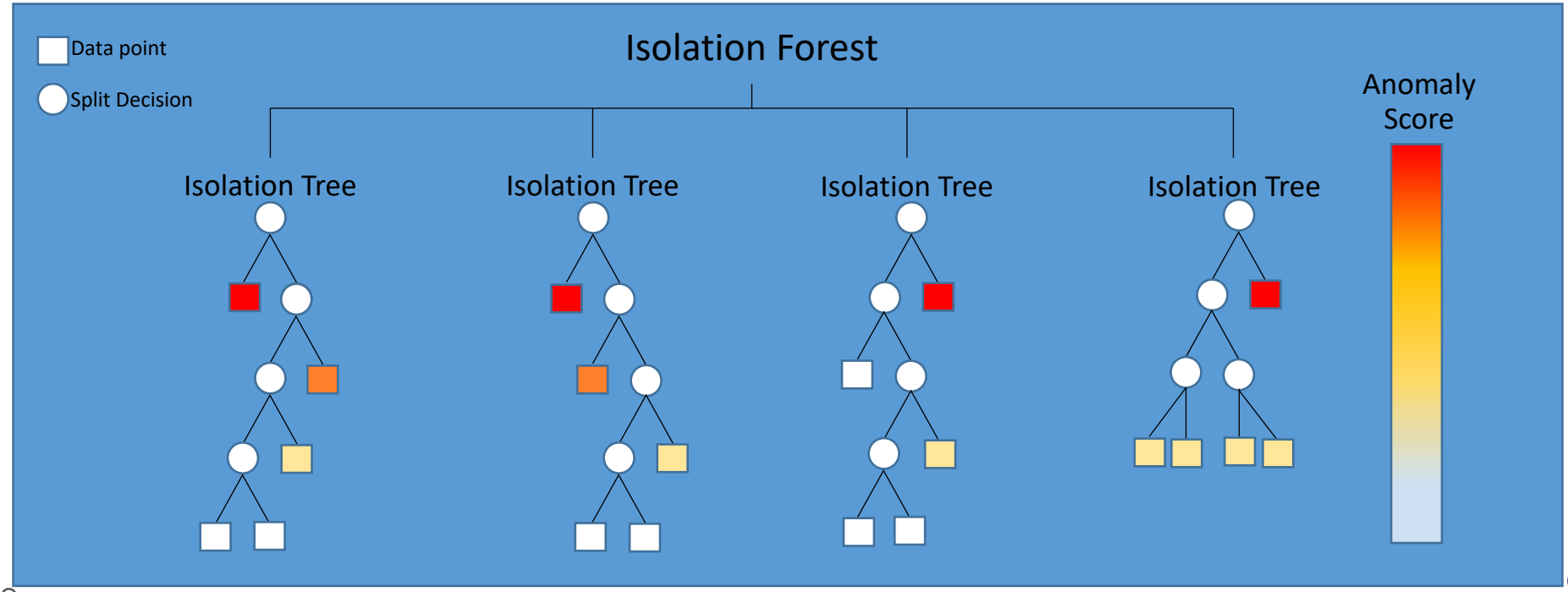
CYBER ELITE

Use Cases : UEBA

- **Anomaly Detection: Isolation forest**

What is Isolation forest?

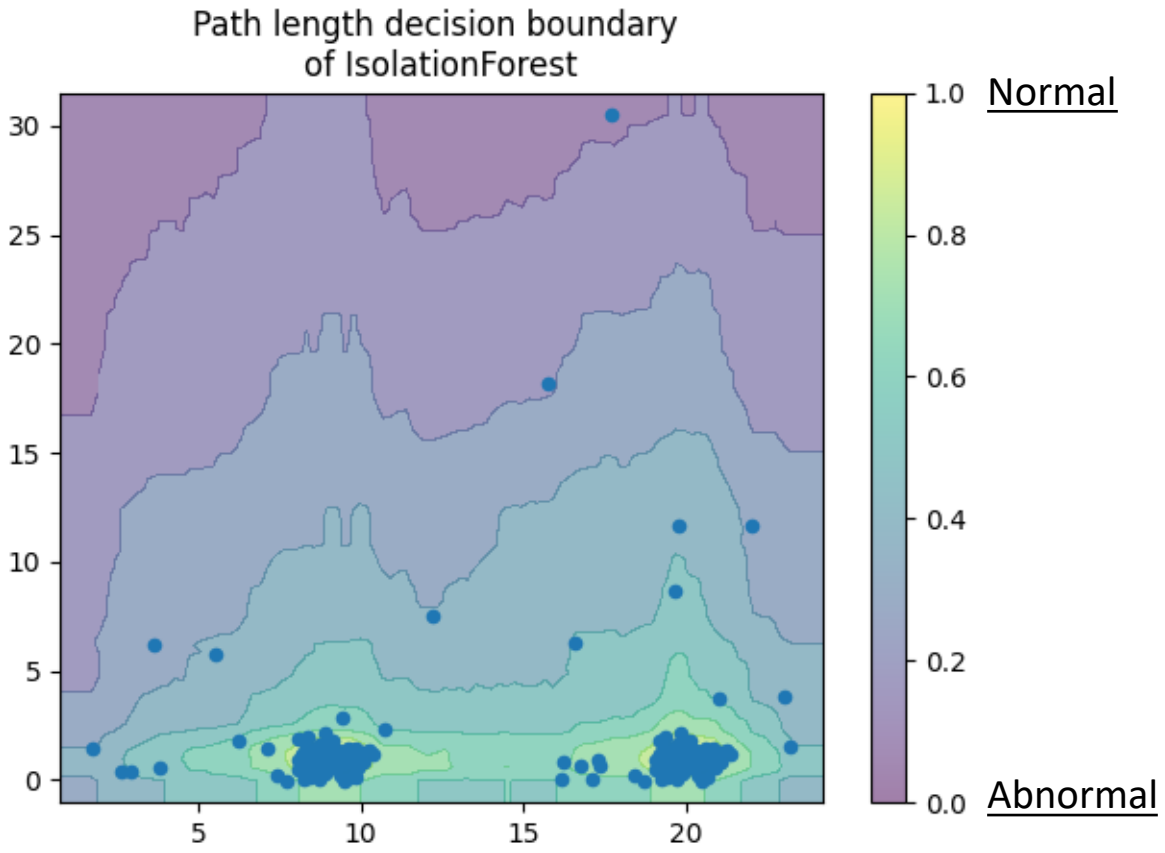
Isolation Forest is a machine learning algorithm that efficiently identifies anomalies or outliers in datasets by leveraging random partitioning and isolation techniques.



Use Cases : UEBA

Anomaly Detection: Isolation forest

Score interpretation



CYBER ELITE

CYBER ELITE

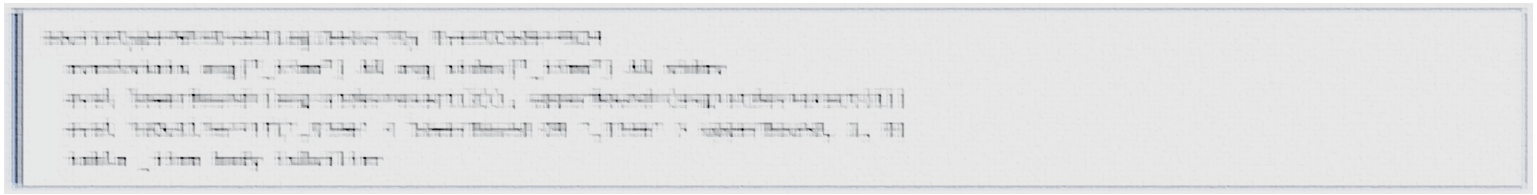
AI Adoption in Cyber Defense Use Cases

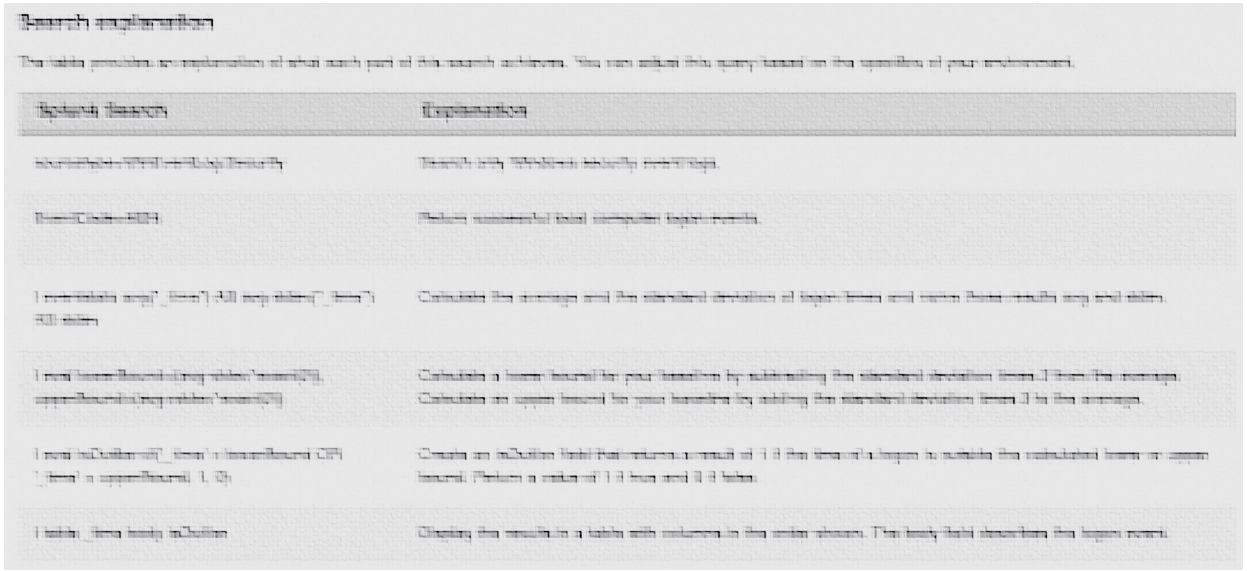
UEBA with Neural Network

CYBER ELITE

Use Cases : UEBA with Neural Network

Rule based (Example) (just for study)



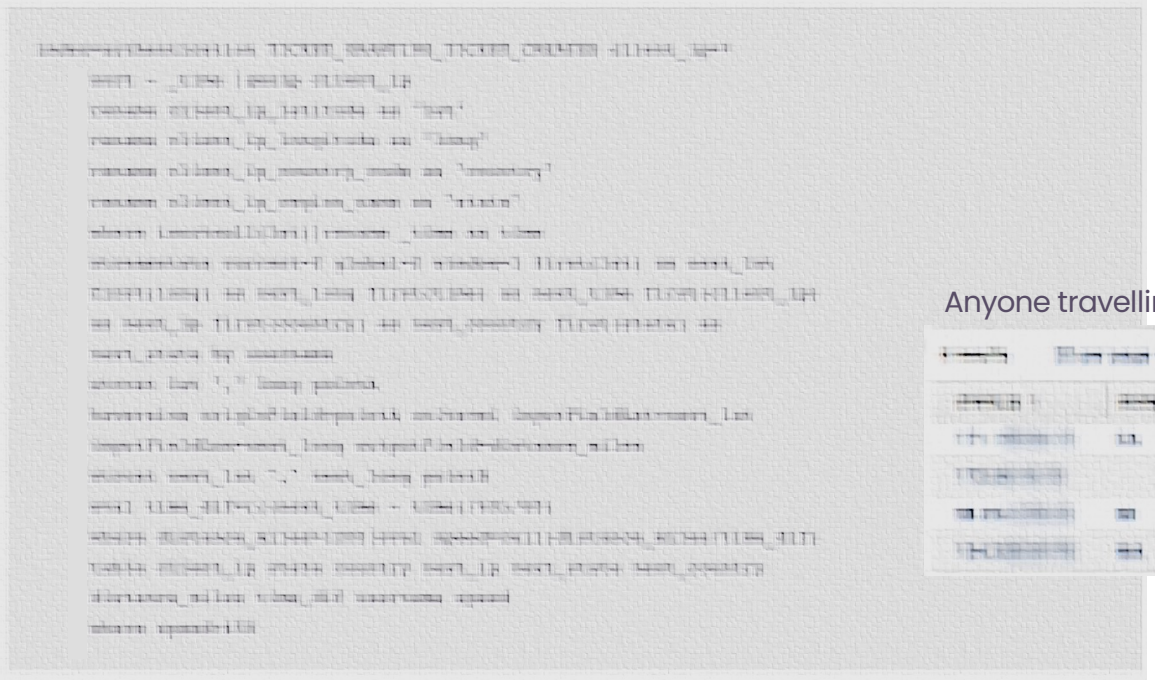


CYBER ELITE

Use Cases : UEBA with Neural Network

Rule based (Example) (just for study)

- Detecting Credential Theft Using Geographic Information



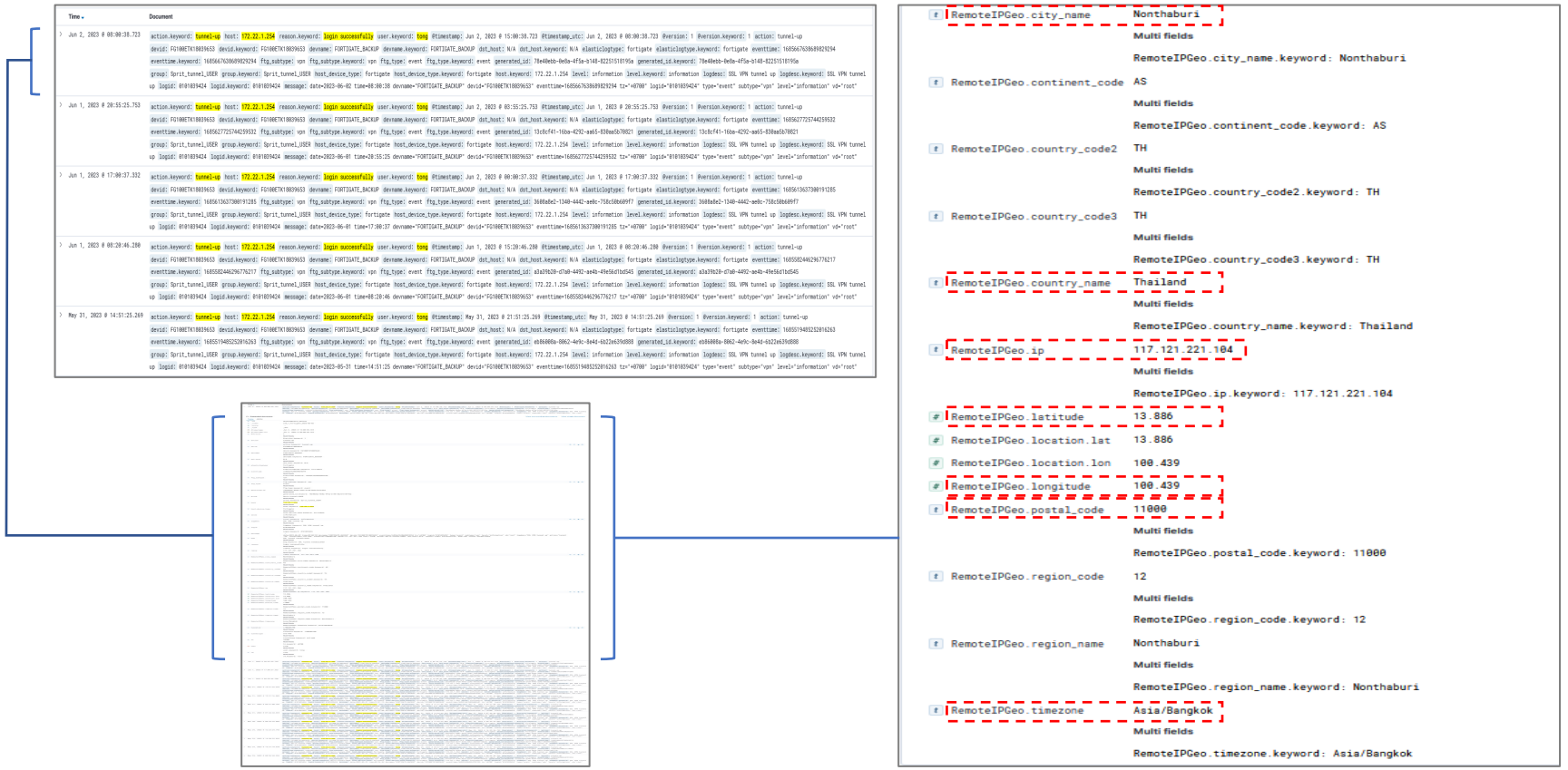
Anyone travelling faster than 450 miles/hour

CYBER ELITE

CYBER ELITE

Use Cases : UEBA with Neural Network

Data Source (Raw Log)



CYBER ELITE

Use Cases : UEBA

Dataset Preparation

user	login_date	login_time	country_name	city_name
tong	1/5/2023	08:59:23	Thailand	Bangkok
tong	1/5/2023	14:19:10	Thailand	Bangkok
tong	1/5/2023	20:04:46	USA	Chicago
tong	2/5/2023	09:29:35	Thailand	Bangkok
tong	2/5/2023	15:31:53	Thailand	Bangkok
tong	2/5/2023	20:39:33	USA	Chicago
tong	3/5/2023	10:22:57	Thailand	Bangkok
tong	3/5/2023	13:56:04	Thailand	Bangkok
tong	3/5/2023	20:13:11	USA	Chicago
tong	4/5/2023	08:23:27	Thailand	Bangkok
tong	4/5/2023	14:23:50	Thailand	Bangkok
tong	4/5/2023	19:43:13	USA	Chicago
tong	5/5/2023	10:02:00	Thailand	Bangkok
tong	5/5/2023	14:33:07	Thailand	Bangkok
tong	5/5/2023	19:14:47	USA	Chicago
tong	6/5/2023	10:33:43	Thailand	Bangkok
tong	6/5/2023	15:57:44	Thailand	Bangkok
tong	6/5/2023	21:34:16	USA	Chicago
tong	7/5/2023	09:40:30	Thailand	Bangkok
tong	7/5/2023	13:46:40	Thailand	Bangkok
tong	7/5/2023	21:06:50	USA	Chicago
tong	8/5/2023	10:16:01	Thailand	Bangkok
tong	8/5/2023	13:46:46	Thailand	Bangkok
tong	8/5/2023	20:06:08	USA	Chicago
. . .				
tong	27/5/2023	13:05:50	Thailand	Bangkok
tong	27/5/2023	21:39:03	USA	Chicago
tong	28/5/2023	09:57:07	Thailand	Bangkok
tong	28/5/2023	15:26:58	Thailand	Bangkok
tong	28/5/2023	20:33:49	USA	Chicago
tong	29/5/2023	09:39:22	Thailand	Bangkok
tong	29/5/2023	15:40:12	Thailand	Bangkok
tong	29/5/2023	21:01:29	USA	Chicago
tong	30/5/2023	09:47:32	Thailand	Bangkok
tong	30/5/2023	15:47:54	Thailand	Bangkok
tong	30/5/2023	21:14:08	USA	Chicago
tong	31/5/2023	09:25:29	Thailand	Bangkok
tong	31/5/2023	14:02:45	Thailand	Bangkok
tong	31/5/2023	20:24:16	USA	Chicago

encoding

login_time	encoding hour	country_name	encoding country_code
8:59:23	8	Thailand	66
14:19:10	14	Thailand	66
20:04:46	20	USA	1
09:29:35	9	Thailand	66
15:31:53	15	Thailand	66
20:39:33	20	USA	1
10:22:57	10	Thailand	66
13:56:04	13	Thailand	66
20:13:11	20	USA	1
08:23:27	8	Thailand	66
14:23:50	14	Thailand	66
19:43:13	19	USA	1
10:02:00	10	Thailand	66
14:33:07	14	Thailand	66
19:14:47	19	USA	1
10:33:43	10	Thailand	66
15:57:44	15	Thailand	66
21:34:16	21	USA	1
09:40:30	9	Thailand	66
13:46:40	13	Thailand	66
21:06:50	21	USA	1
10:16:01	10	Thailand	66
13:46:46	13	Thailand	66
20:06:08	20	USA	1
. . .			
13:05:50	13	Thailand	66
21:39:03	21	USA	1
09:57:07	9	Thailand	66
15:26:58	15	Thailand	66
20:33:49	20	USA	1
09:39:22	9	Thailand	66
15:40:12	15	Thailand	66
21:01:29	21	USA	1
09:47:32	9	Thailand	66
15:47:54	15	Thailand	66
21:14:08	21	USA	1
09:25:29	9	Thailand	66
14:02:45	14	Thailand	66
20:24:16	20	USA	1

Datetime Properties

`Series.dt.date`	Returns numpy array of python datetime.date objects (namely, the date part of Timestamps without timezone information).
`Series.dt.time`	Returns numpy array of datetime.time.
`Series.dt.year`	The year of the datetime
`Series.dt.month`	The month as January=1, December=12
`Series.dt.day`	The days of the datetime
`Series.dt.hour`	The hours of the datetime
`Series.dt.minute`	The minutes of the datetime
`Series.dt.second`	The seconds of the datetime
`Series.dt.microsecond`	The microseconds of the datetime
`Series.dt.nanosecond`	The nanoseconds of the datetime
`Series.dt.week`	The week ordinal of the year
`Series.dt.weekofyear`	The week ordinal of the year
`Series.dt.dayofweek`	The day of the week with Monday=0, Sunday=6
`Series.dt.weekday`	The day of the week with Monday=0, Sunday=6
`Series.dt.dayofyear`	The ordinal day of the year
`Series.dt.quarter`	The quarter of the date
`Series.dt.is_month_start`	Logical indicating if first day of month (defined by frequency)
`Series.dt.is_month_end`	Indicator for whether the date is the last day of the month.
`Series.dt.is_quarter_start`	Indicator for whether the date is the first day of a quarter.
`Series.dt.is_quarter_end`	Indicator for whether the date is the last day of a quarter.
`Series.dt.is_year_start`	Indicate whether the date is the first day of a year.
`Series.dt.is_year_end`	Indicate whether the date is the last day of the year.
`Series.dt.is_leap_year`	Boolean indicator if the date belongs to a leap year.
`Series.dt.daysinmonth`	The number of days in the month
`Series.dt.days_in_month`	The number of days in the month

CYBER ELITE

Use Cases : UEBA

Data Labeling (for Supervised Learning Algorithm)

Normal Bahavior

hour	country_code	label (1=nomal, 0=anomaly)
8	66	1
14	66	1
20	1	1
9	66	1
15	66	1
20	1	1
10	66	1
13	66	1
20	1	1
8	66	1
14	66	1
19	1	1
10	66	1
14	66	1
19	1	1
10	66	1
15	66	1
21	1	1
9	66	1
13	66	1
21	1	1
10	66	1
13	66	1
20	1	1

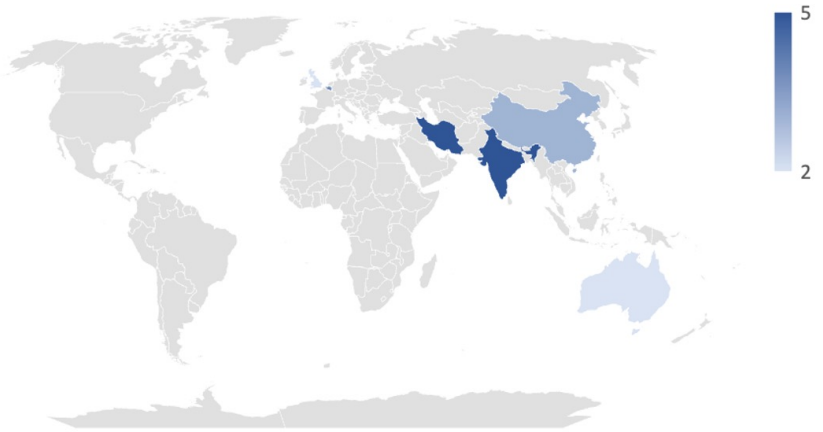
Abnomal Behavior

hour	country_code	label(1=nomal, 0=anomaly)
1	44	0
2	32	0
5	86	0
6	32	0
23	44	0
23	86	0
1	91	0
3	32	0
4	44	0
5	91	0
22	44	0
23	91	0
0	98	0
1	44	0
2	86	0
5	91	0
6	91	0
23	61	0
23	91	0
1	32	0
3	91	0
4	61	0
5	44	0
22	98	0

Normal Hour vs. Anomaly Hour

label (1=nomal, 0=anomaly)	0	1	2	3	4	5	6	8	9	10	13	14	15	19	20	21	22	23	Grand Total
0	1	4	2	2	2	4	2										2	5	24
1								7	14	10	9	13	9	10	10	11			93
Grand Total	1	4	2	2	2	4	2	7	14	10	9	13	9	10	10	11	2	5	117

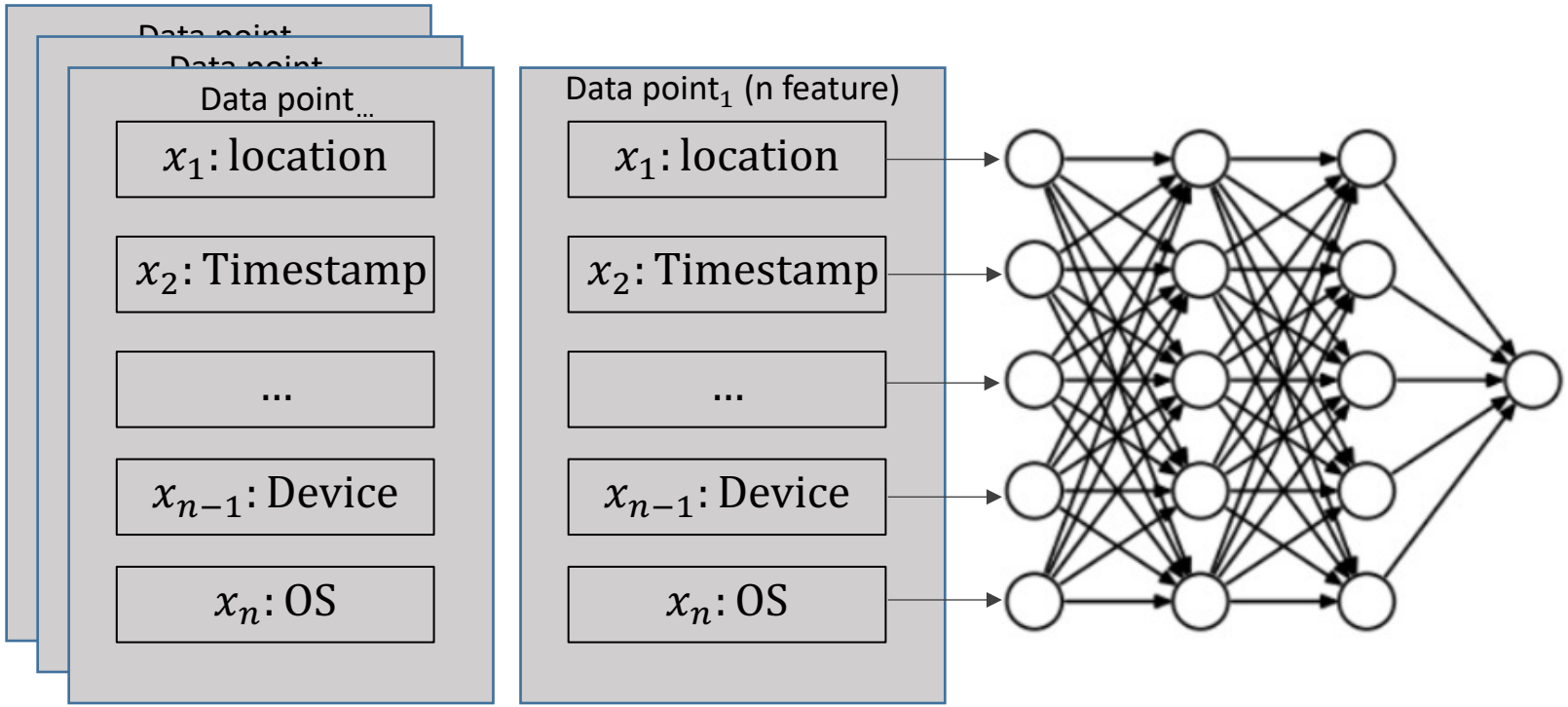
Anomaly Country



CYBER ELITE

Use Cases : UEBA

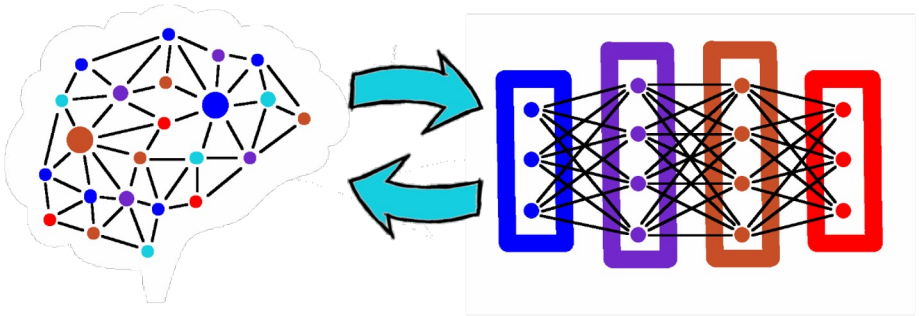
Use data to train Neural Network

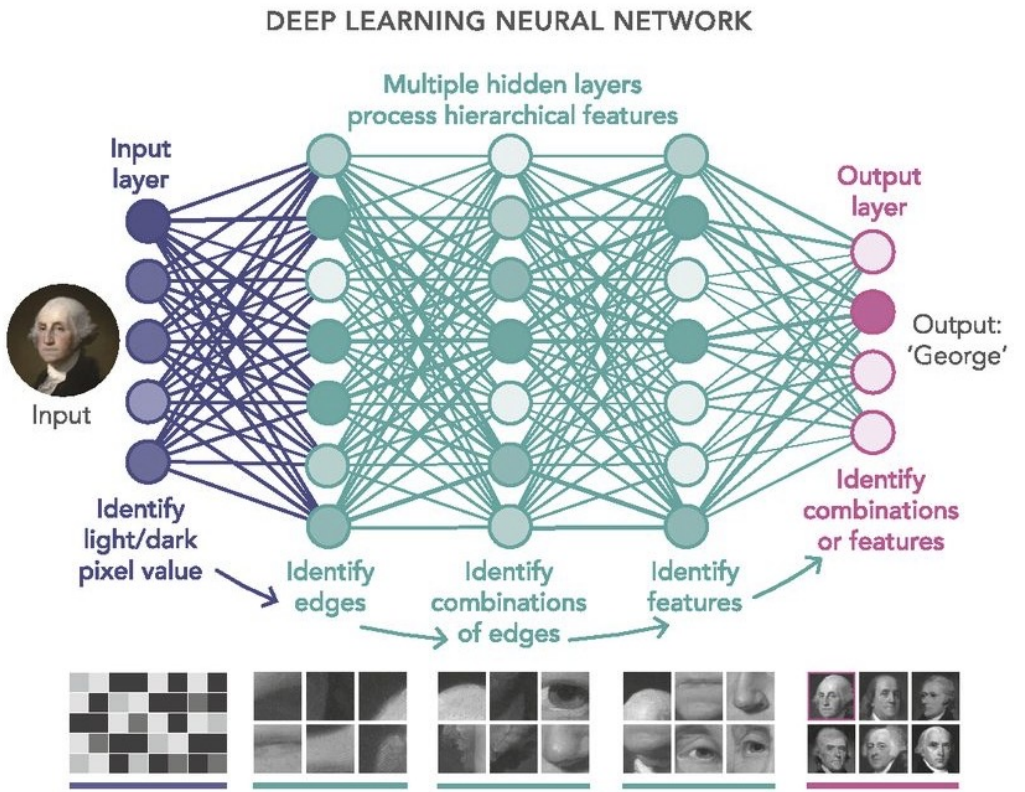


CYBER ELITE

Use Cases : UEBA with Neural Network

Instead of relying on those rule to handles case by case, we can use machine learning to find those rule for us.
In this specific we can use **Neural Network!**





CYBER ELITE

CYBER ELITE

Use Cases : UEBA with Neural Network

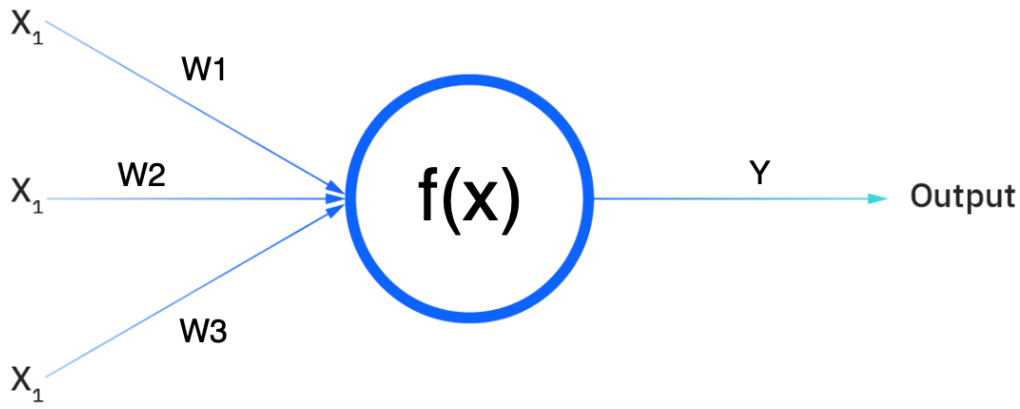
Neural Network. How is it compared to the working of brain.

You should go surfing (Yes: 1, No: 0).
Let's assume that there are three factors influencing your decision-making:

1. Are the waves good? (Yes: 1, No: 0)
2. Is the line-up empty? (Yes: 1, No: 0)
3. Has there been a recent shark attack? (Yes: 0, No: 1)

Then, let's assume the following, giving us the following inputs:

- X1 = 1, since the waves are pumping
- X2 = 0, since the crowds are out
- X3 = 1, since there hasn't been a recent shark attack
- W1 = 5, since large swells don't come around often
- W2 = 2, since you're used to the crowds
- W3 = 4, since you have a fear of sharks



$\sum w_i x_i + bias = w1x1 + w2x2 + w3x3 + bias$

output = f(x) = 1 if $\sum w1x1 + b >= 0$; 0 if $\sum w1x1 + b < 0$

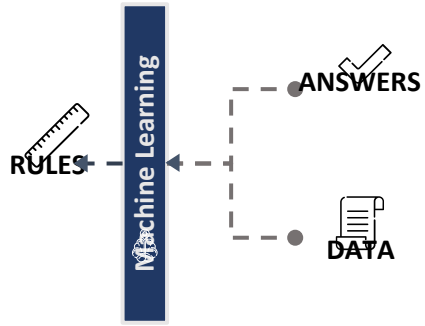
$Y = (1*5) + (0*2) + (1*4) - 3 = 6$

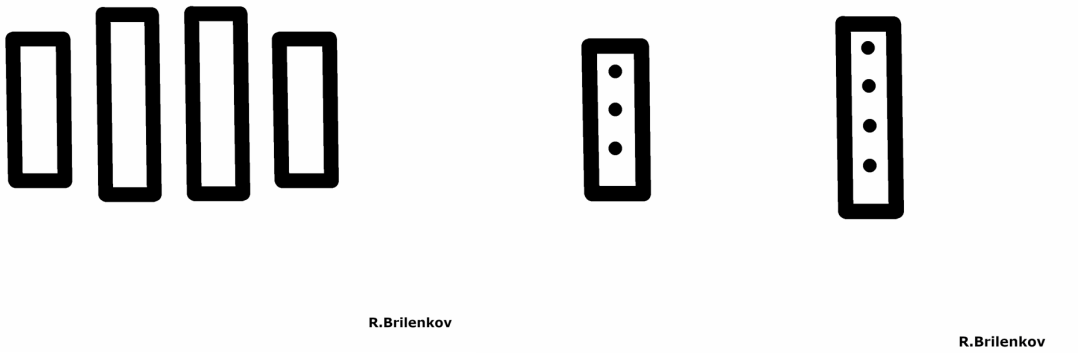
CYBER ELITE

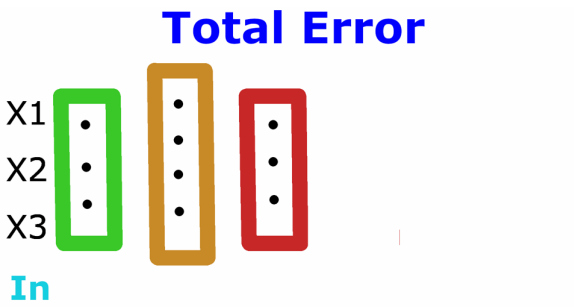
CYBER ELITE

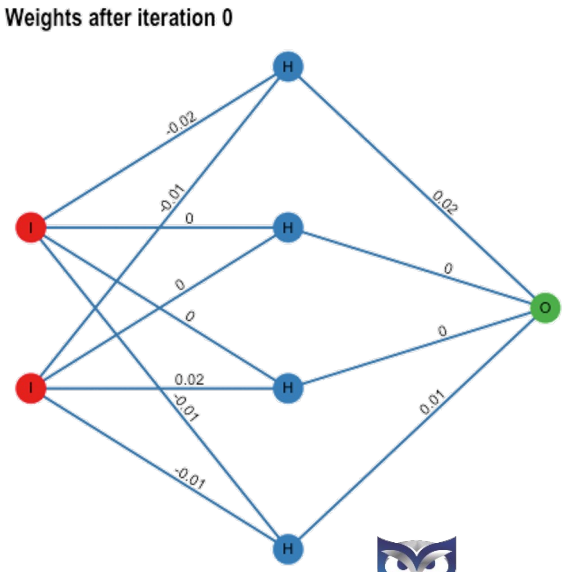
Use Cases : UEBA with Neural Network

Neural Network. How dose it work?







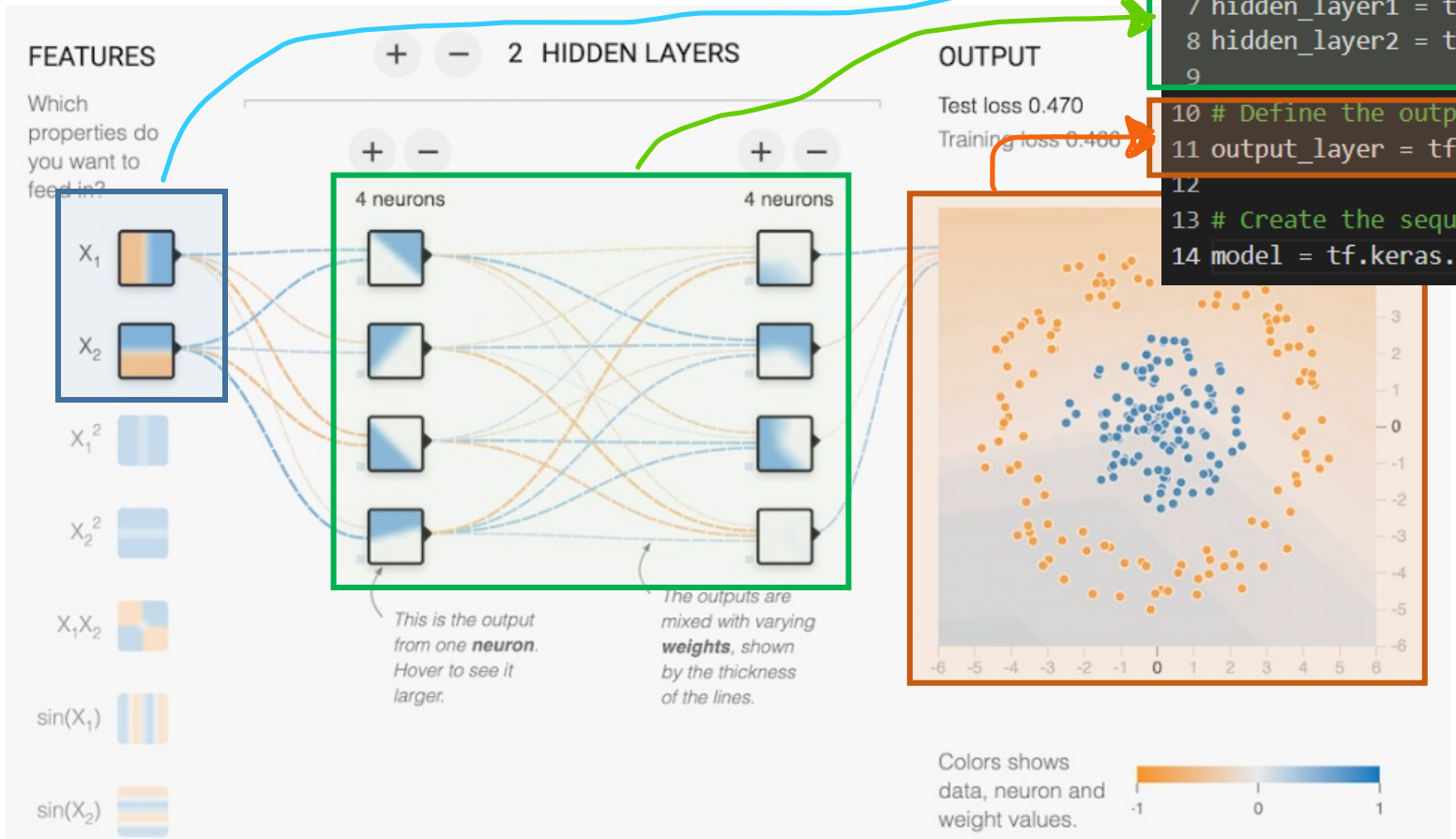


CYBER ELITE

Use Cases : UEBA

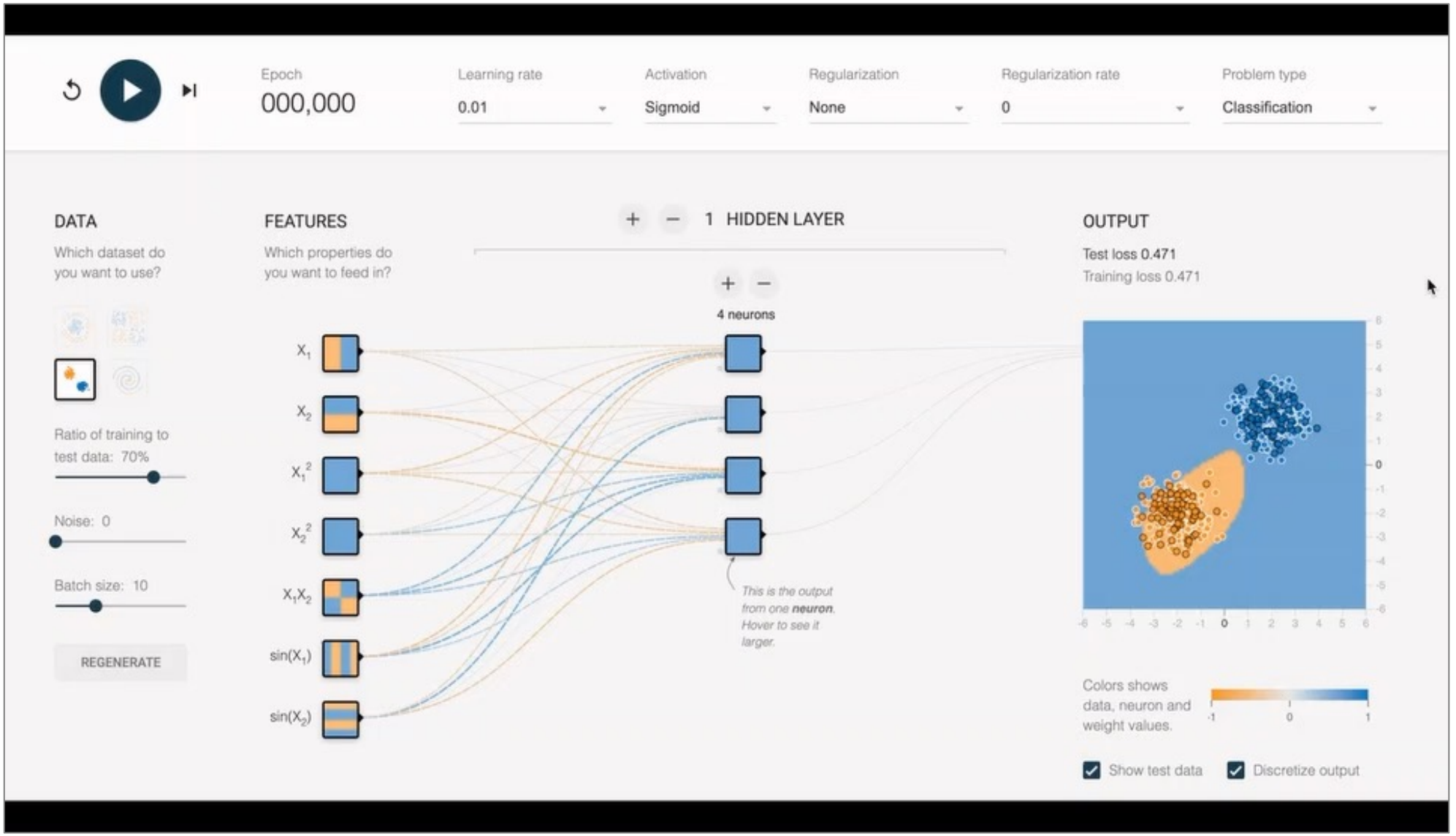
- **Neural Network Implementation**

```python
import tensorflow as tf

# Define input layer
input_layer = tf.keras.layers.Input(shape=(2,))

# Define the hidden layers
hidden_layer1 = tf.keras.layers.Dense(4, activation='relu')(input_layer)
hidden_layer2 = tf.keras.layers.Dense(4, activation='relu')(hidden_layer1)

# Define the output layer
output_layer = tf.keras.layers.Dense(1)(hidden_layer2)

# Create the sequential model
model = tf.keras.models.Model(inputs=input_layer, outputs=output_layer)
```



CYBER ELITE

Neural Network. How dose it work?



https://playground.tensorflow.org/

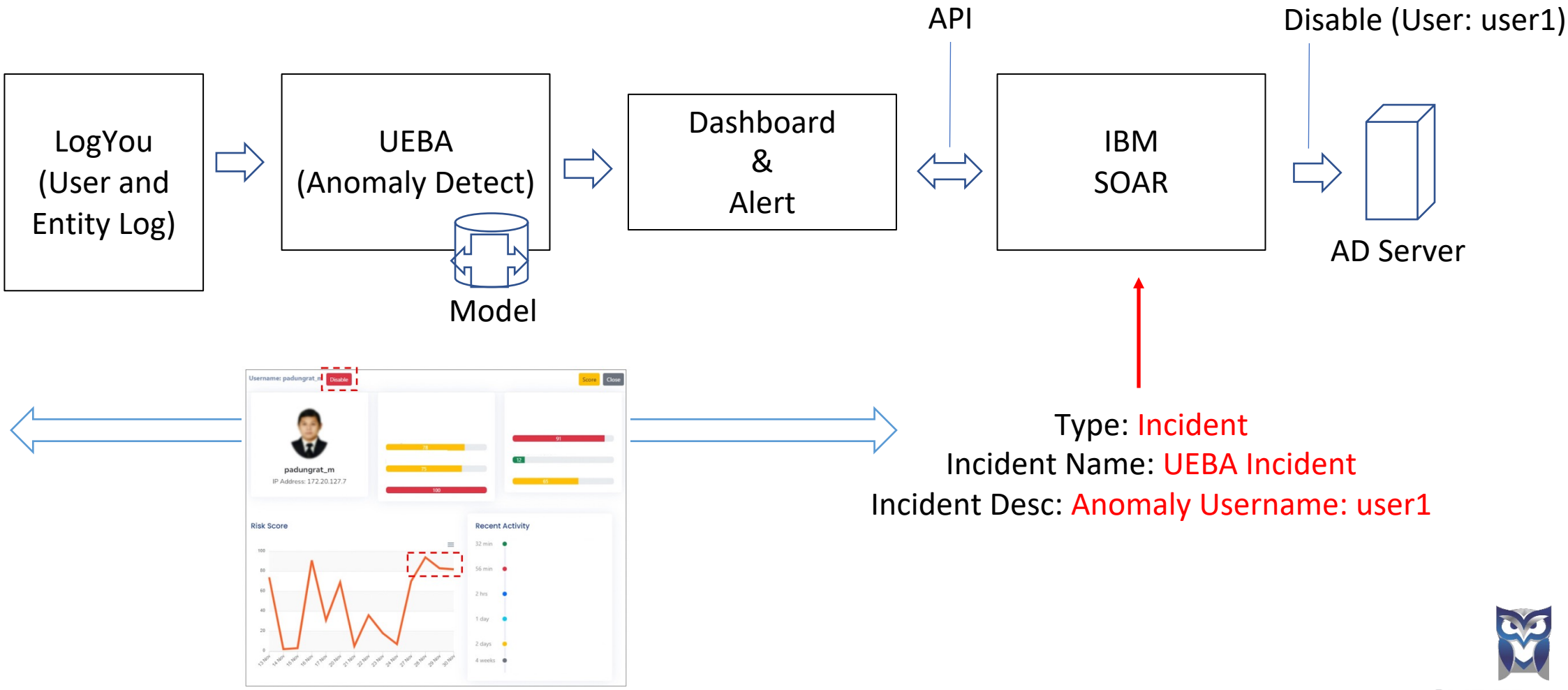
CYBER ELITE

UEBA, Other use cases (Example)

- Session Duration Analysis -> Ex: Web Application Session Duration, Database Session Duration
- Data Transfer Analysis -> Data Exfiltration
- Data Access Activity Analysis -> Data Protection(File, Database, Application)
- Fraud Detection -> Transaction Behavior Analysis

CYBER ELITE

Use Cases : UEBA + IBM SOAR



CYBER ELITE

AI Adoption in Cyber Defense (Conclusion)

Key Issues	Use Cases	ML Algorithms	Proactive/ Reactive
Threat detection (False Positive, False Negative, MTTD)	UEBA	Anomaly Detect (Isolation Forest (Unsupervised), Neural Network (Supervised))	Proactive/ Reactive
Threat response time (MTTR)			
New threat identification (Zero-day Attack)			
Staffing capacity and expertise			
Large volume of cyber alerts			
How to manage?			

CYBER ELITE

AI Adoption in Cyber Defense Use Cases

Natural Language Processing (NLP)

CYBER ELITE

AI Use cases: Natural Language Processing

- **Issues Recurring Detection**
- **Text Clustering**

CYBER ELITE

AI Adoption in Cyber Defense Use Cases

NLP : Issues Recurring Detection

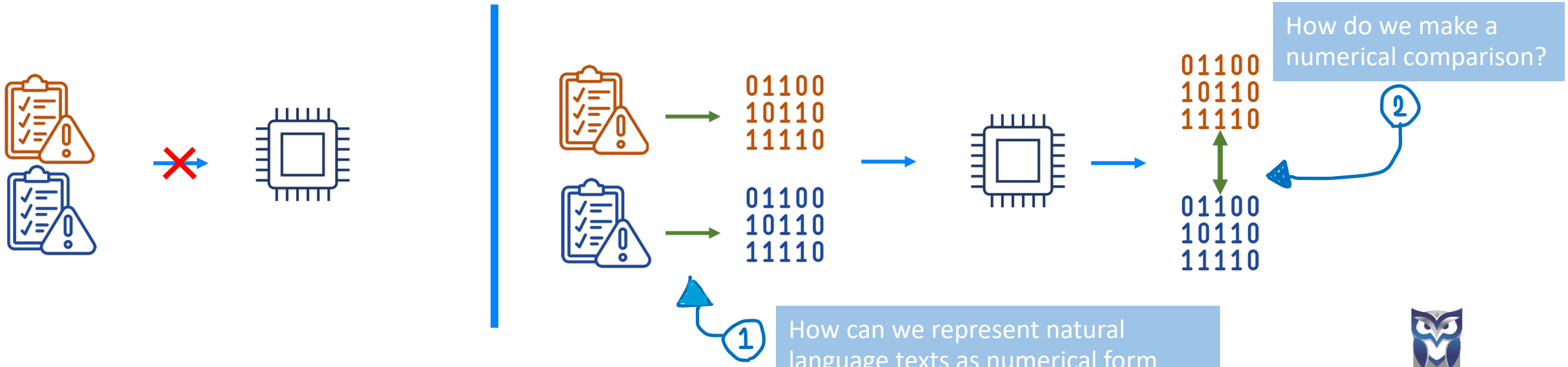
CYBER ELITE

AI Use cases: Natural Language Processing

- **Issues Recurring Detection**

How?

Intuitive way to approach is to reshape the ways we think about/ ask the question



CYBER ELITE

CYBER ELITE

AI Use cases: Natural Language Processing

- **Issues Recurring Detection** 

Classical non-contextual algorithms

Jaccard Similarity. The simplest way to compare two texts.

$$Jaccard\ Similarity = \frac{Number\ of\ common\ unique\ words}{Total\ Number\ of\ unique\ words} = \frac{AND\ operation\ then\ bit\ count}{OR\ operation\ then\ bit\ count} = \frac{3}{5} = 0.6$$

Example.

			Numerical representation				
No.	Sentence	Unique words	เรา	ชอบ	กิน	กาแฟ	ชา
1	เราชอบกินกาแฟ	[เรา,ชอบ,กิน,กาแฟ]	1	1	1	1	0
2	เราชอบกินชา	[เรา,ชอบ,กิน,ชา]	1	1	1	0	1

CYBER ELITE

AI Use cases: Natural Language Processing

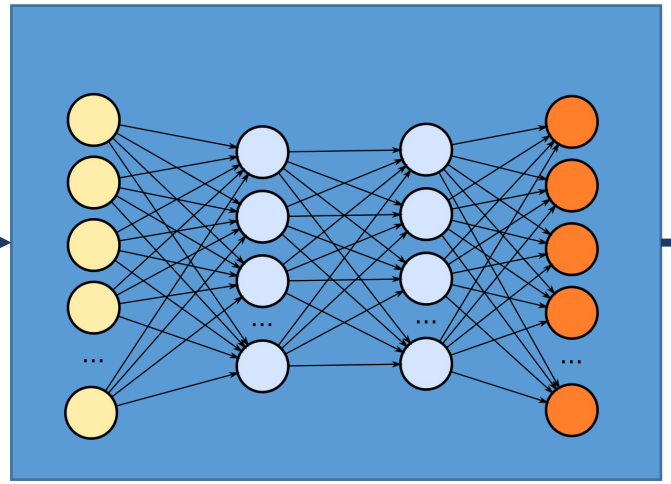
- **Issues Recurring Detection**

Modern contextual algorithms

Utilizing large language model (LLM) that is pre-trained on massive amount of text(over 1 billion sentence).

It capable of **encode sentences into vector representations that capture the meaning of the sentence**

Tokenize	Encode
I	4324
Like	2695
Apple	10682
<NULL>	0
…	…
<NULL>	0



Large Language Model

0.15	0.55	0. 08	…	0.89	0.45	0.93

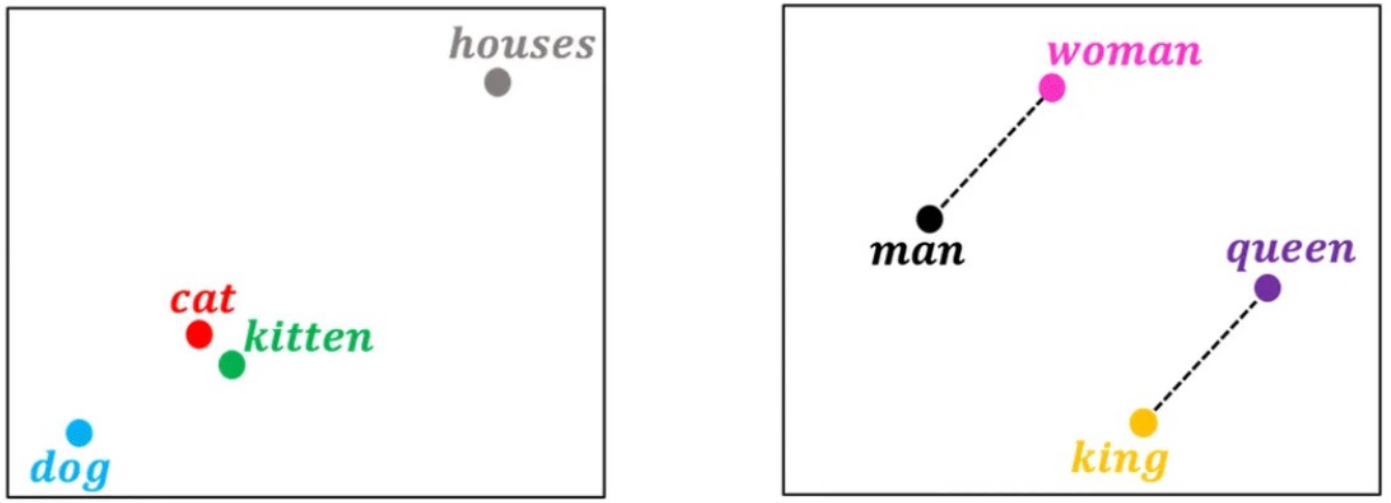
vector representation of the sentence

CYBER ELITE

AI Use cases: Natural Language Processing

- **Issues Recurring Detection**

  Modern contextual algorithms



Visualization of word embeddings in 2D

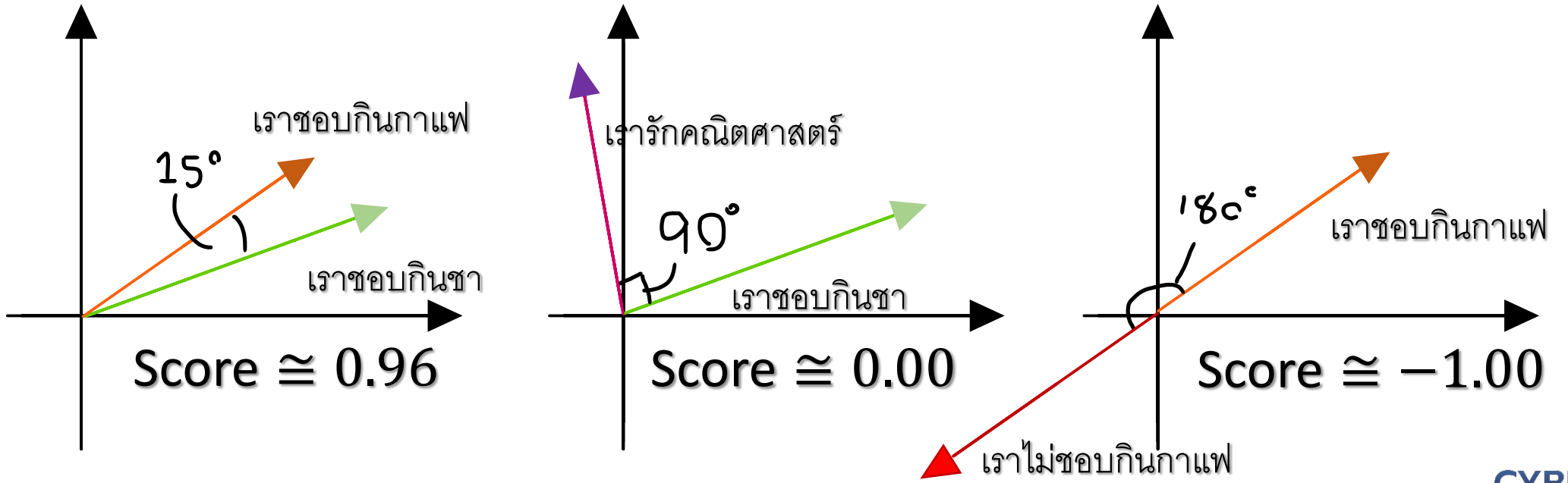
CYBER ELITE

AI Use cases: Natural Language Processing

- **Issues Recurring Detection**

Modern contextual algorithms

Cosine similarity example with 2d vector

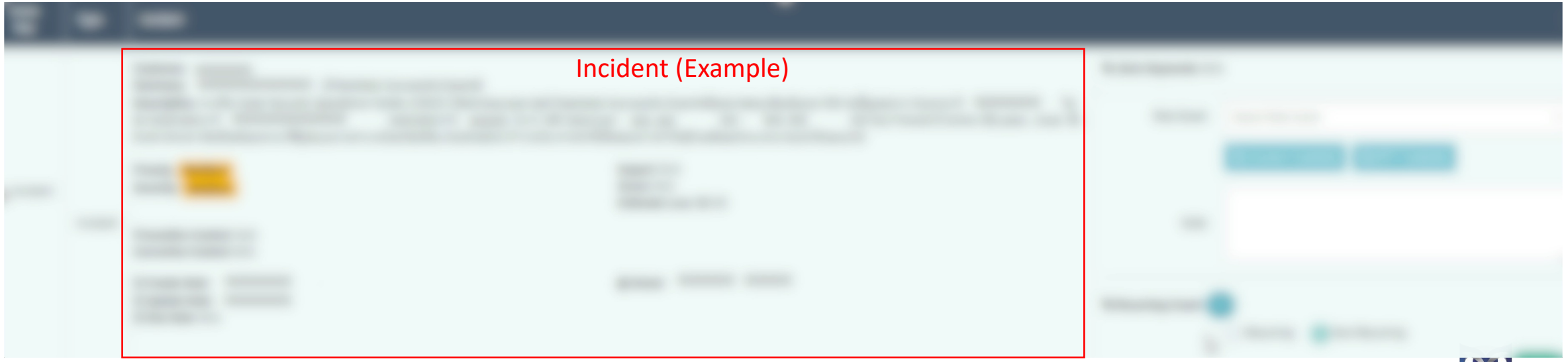


CYBER ELITE

AI Use cases: Natural Language Processing

- **Issues Recurring Detection**

  Showcase

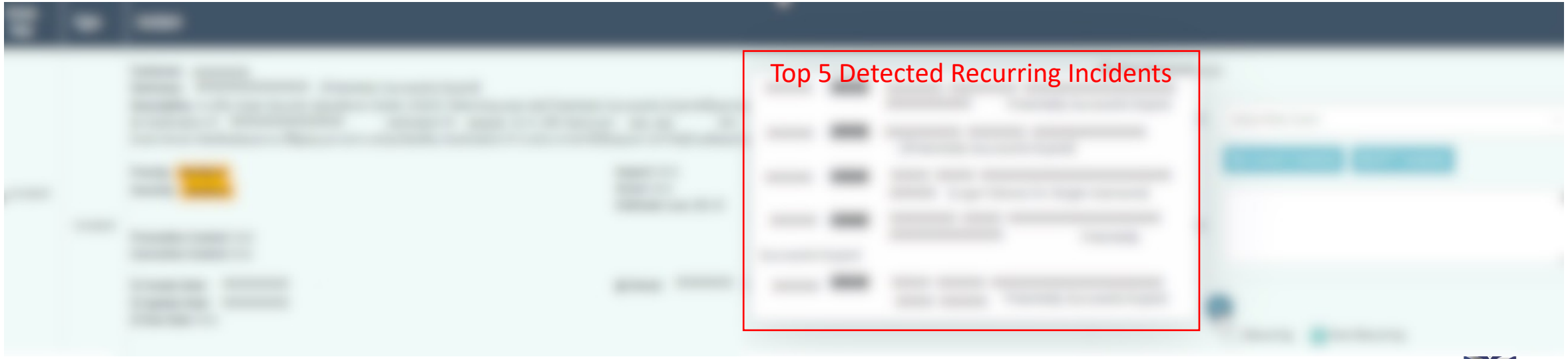


CYBER ELITE

AI Use cases: Natural Language Processing

- **Issues Recurring Detection**

  Showcase

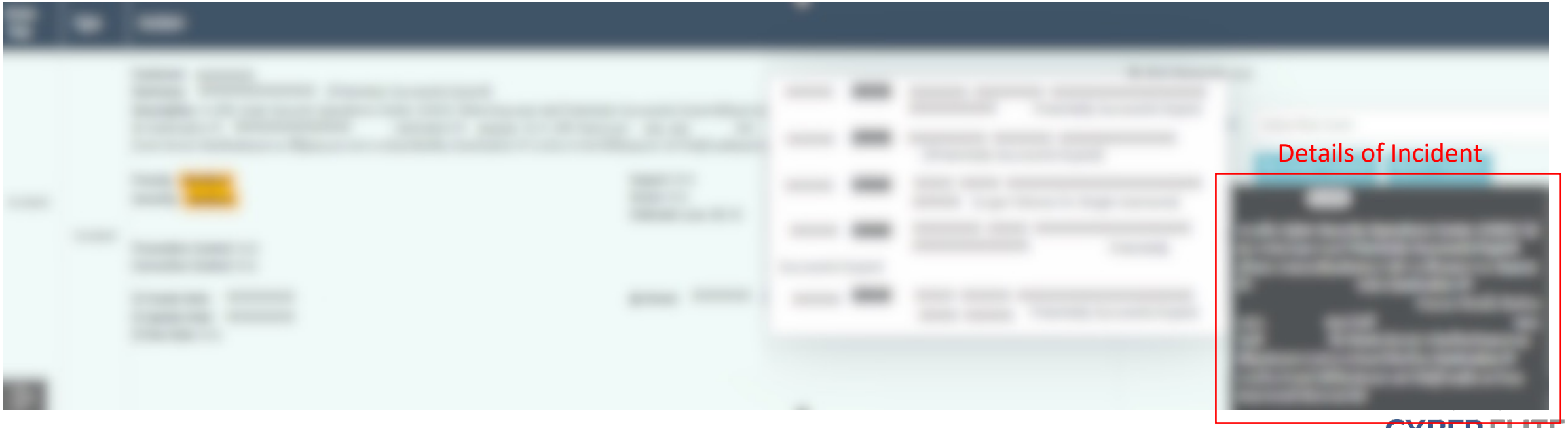


CYBER ELITE

AI Use cases: Natural Language Processing

- **Issues Recurring Detection**

  Showcase



CYBER ELITE

AI Adoption in Cyber Defense Use Cases

NPL : Text Clustering

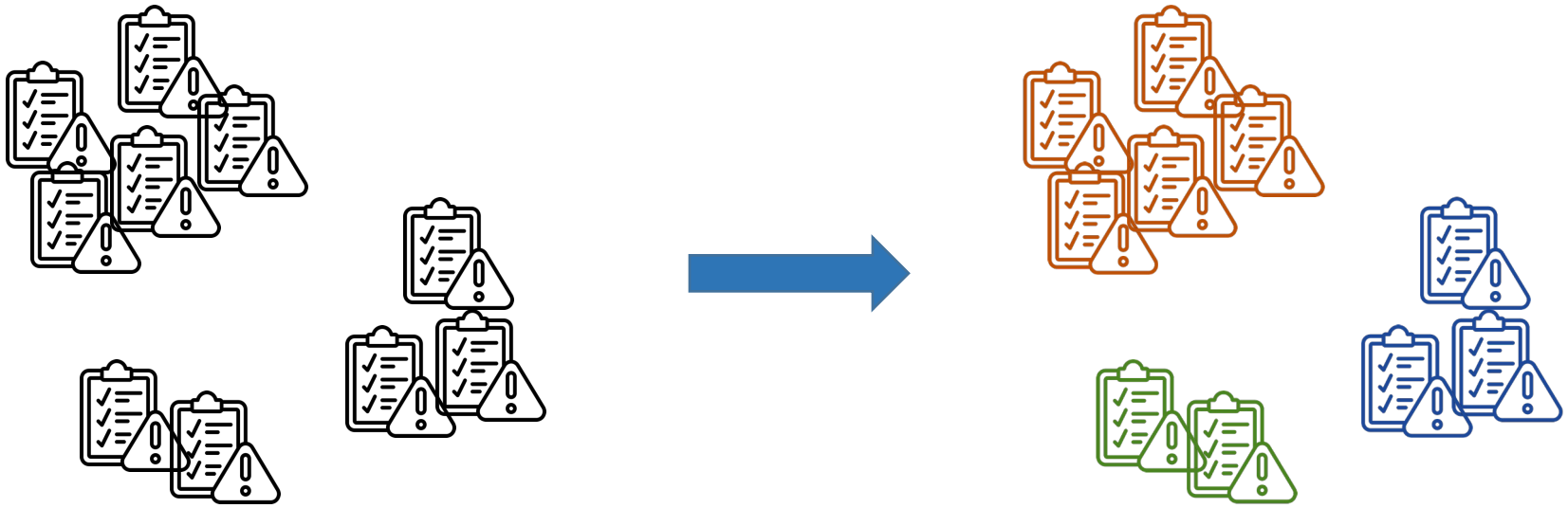
CYBER ELITE

AI Use cases: Natural Language Processing

- **Text Clustering**

What is Text Clustering?

Text clustering is a technique that **groups similar documents together** based on their content.



CYBER ELITE

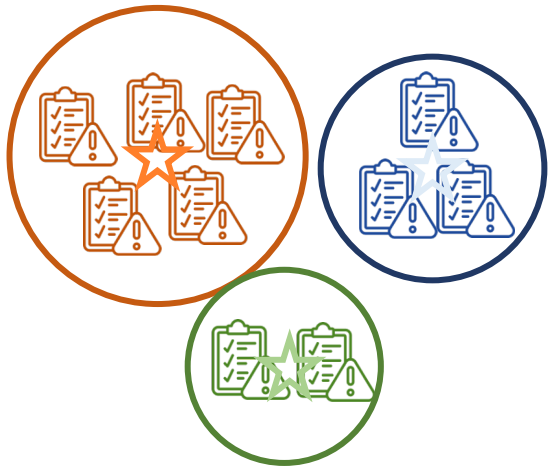
AI Use cases: Natural Language Processing

- **Text Clustering**

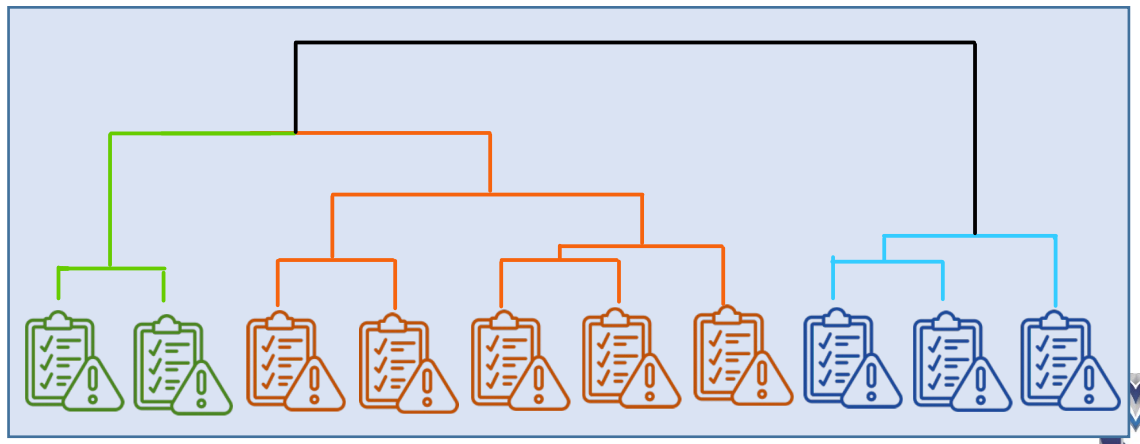
  How to utilizing Text Clustering technique?

  Clustering algorithms

K-means



Hierarchical Clustering



CYBER ELITE

AI Use cases: Natural Language Processing

- **Text Clustering**

  K-means algorithm

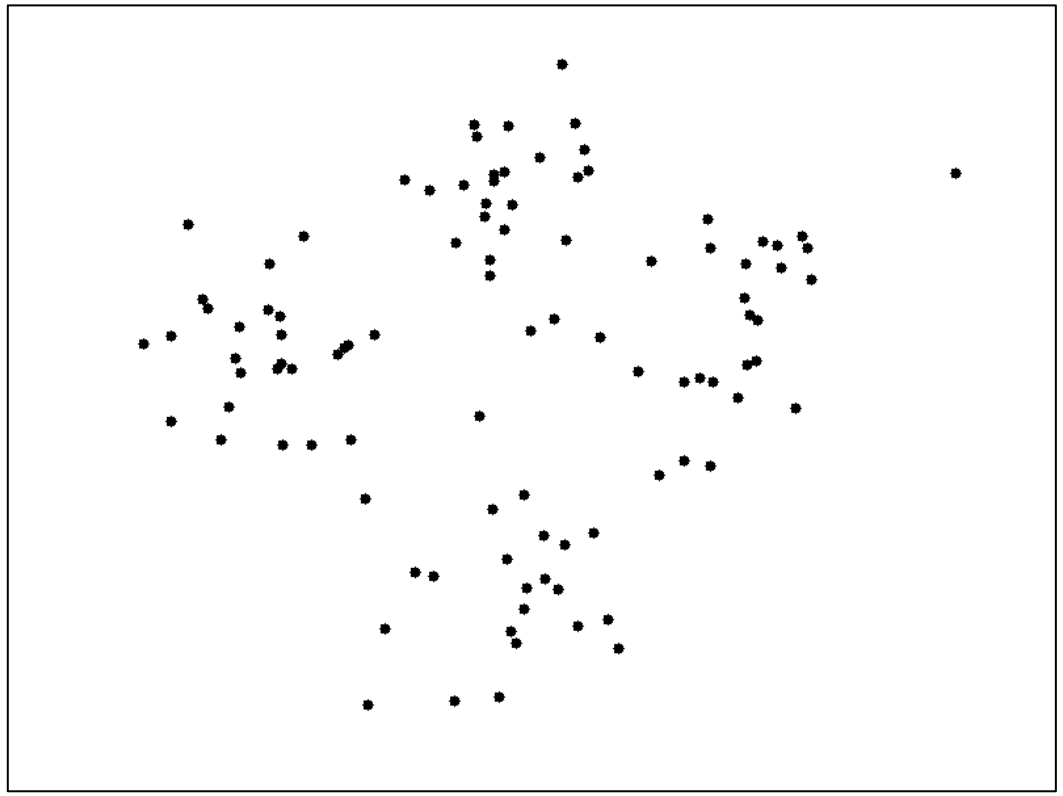
1. Select the number of clusters, **k**, and initialize the cluster centers randomly.
2. Repeat the following steps until convergence
   - Assign each data point to the nearest cluster center
   - Recalculate the cluster centers as the mean of the assigned data points
   - Check for convergence by comparing the new cluster centers with the previous ones.

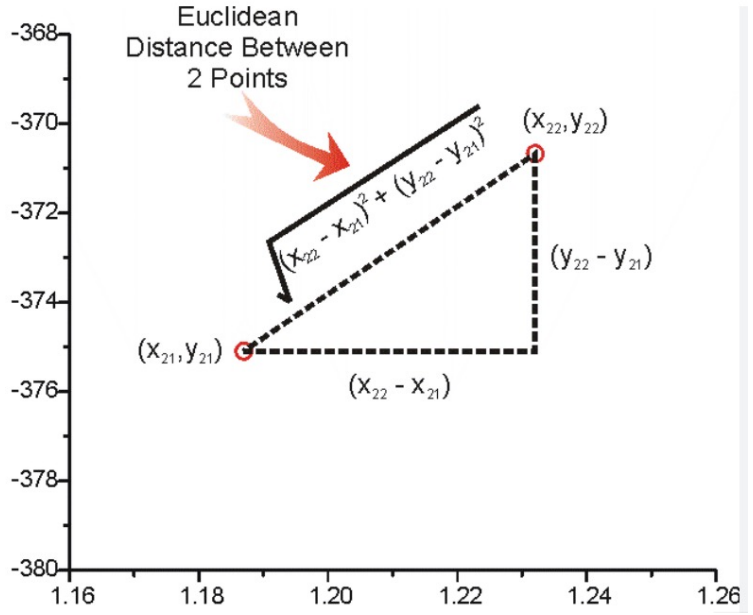
CYBER ELITE

AI Use cases: Natural Language Processing

- **Text Clustering**

  K-means algorithm , on sample data



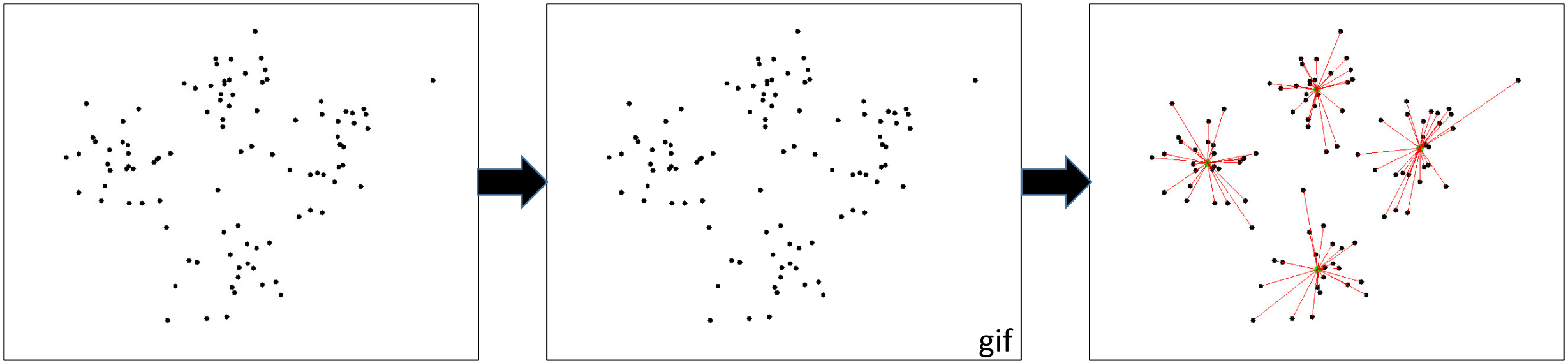


CYBER ELITE

AI Use cases: Natural Language Processing

- **Text Clustering**

  K-means algorithm



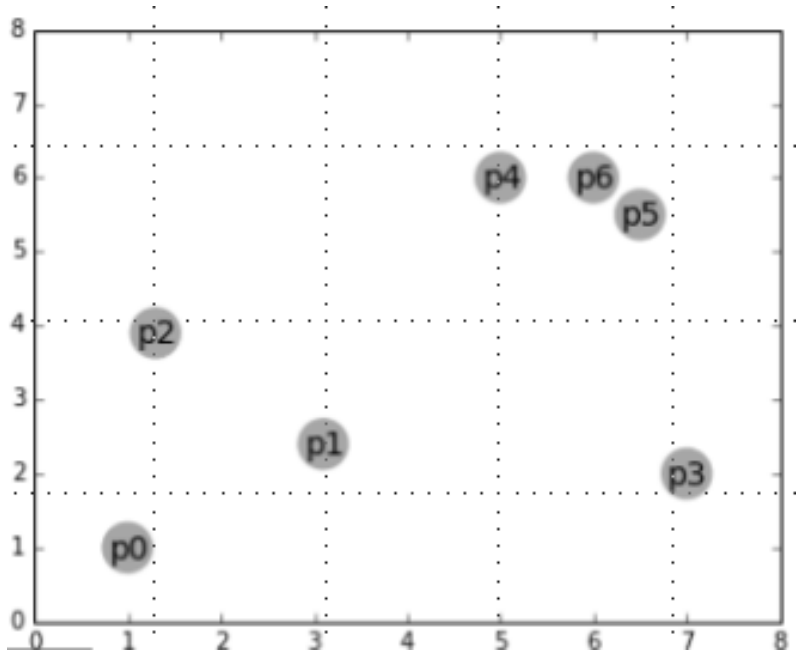
CYBER ELITE

AI Use cases: Natural Language Processing

- **Text Clustering**

  How to utilizing Text Clustering technique?

  Clustering algorithms : Hierarchical Clustering



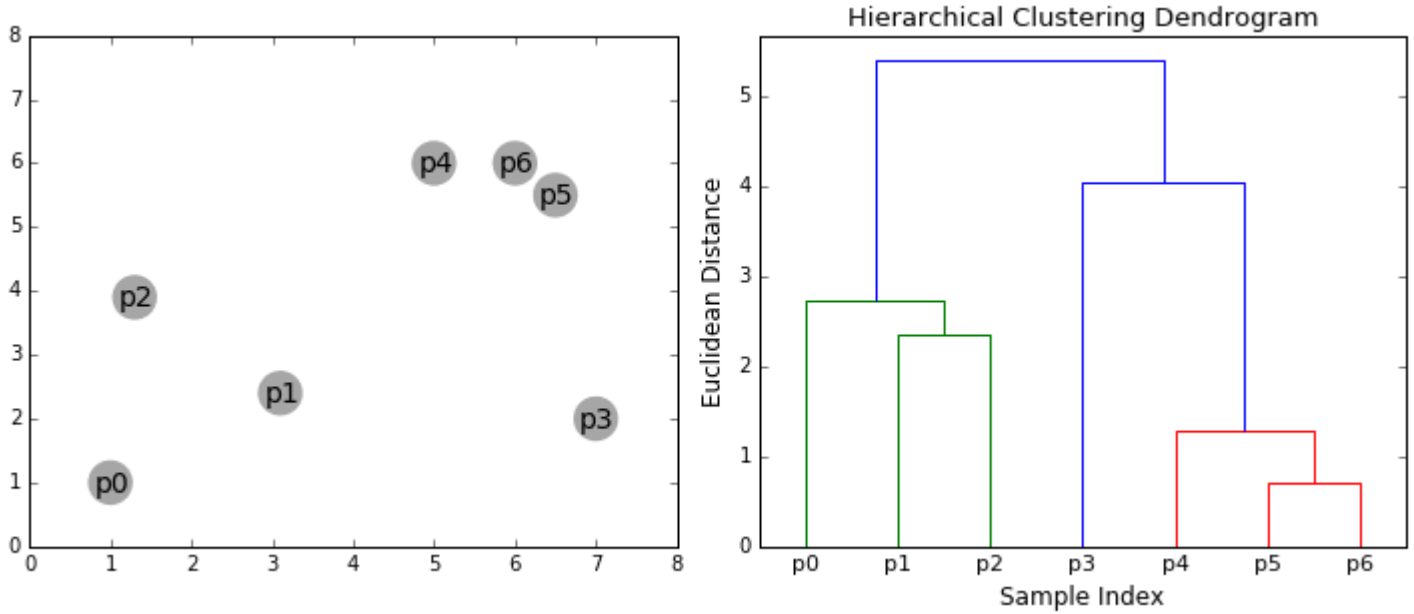
CYBER ELITE

AI Use cases: Natural Language Processing

- **Text Clustering**

  How to utilizing Text Clustering technique?

  Clustering algorithms : Hierarchical Clustering



CYBER ELITE

CYBER ELITE

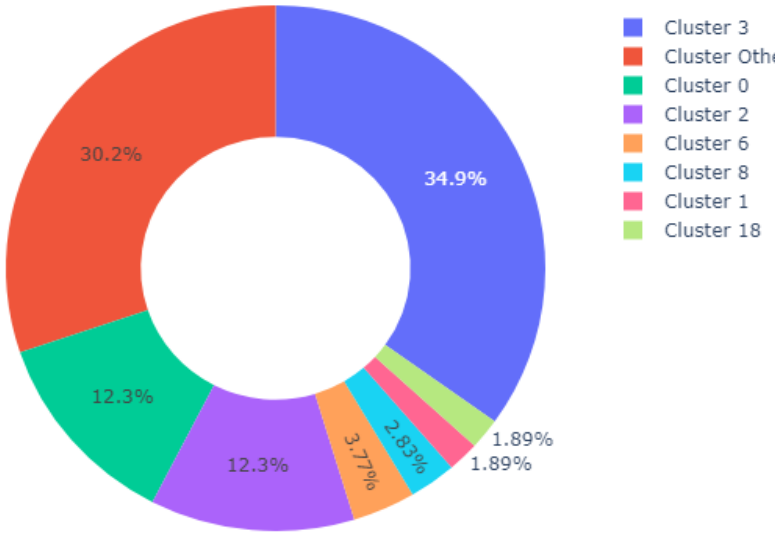
AI Use cases: Natural Language Processing

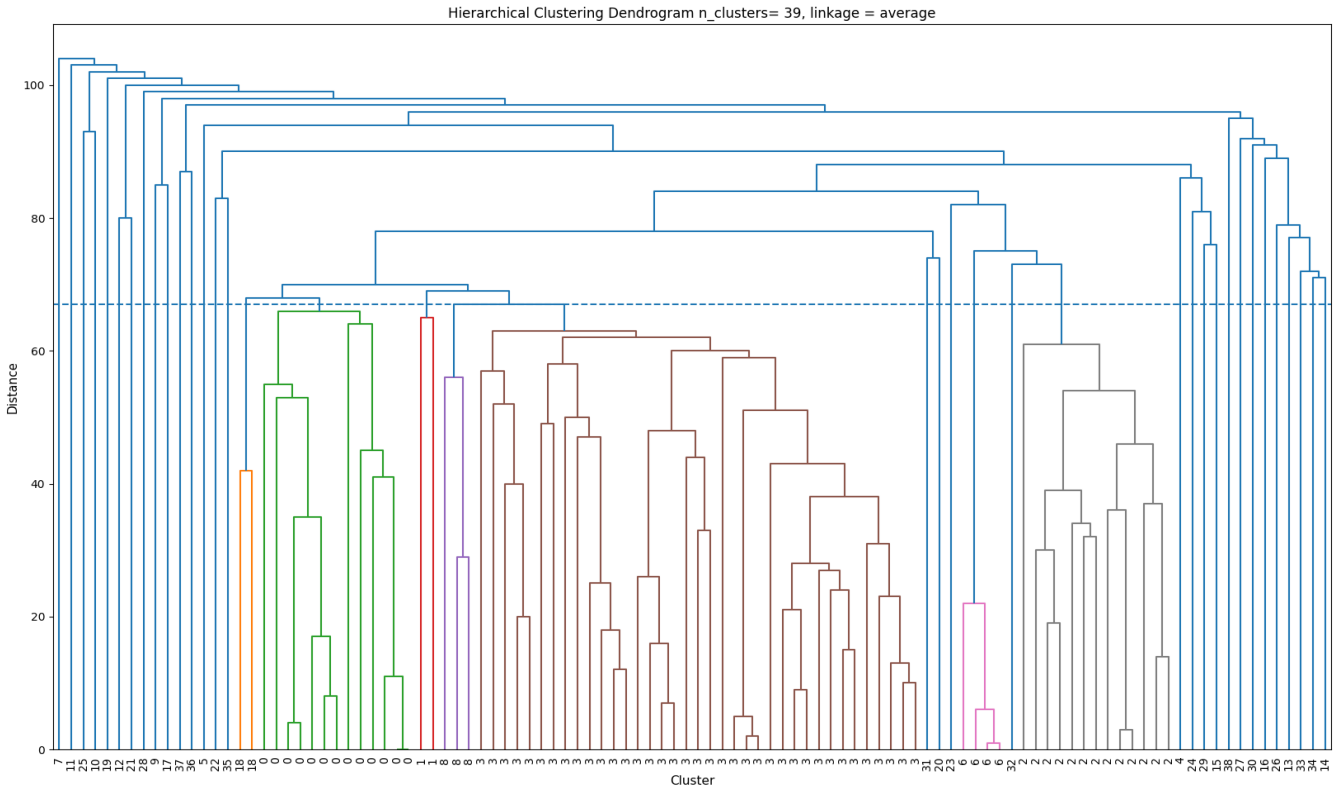
- **Text Clustering**

  Showcase: Hierarchical Clustering

  On our issues/incidents data (≅ 100 Samples

Donut Chart





CYBER ELITE

AI Use cases: Natural Language Processing

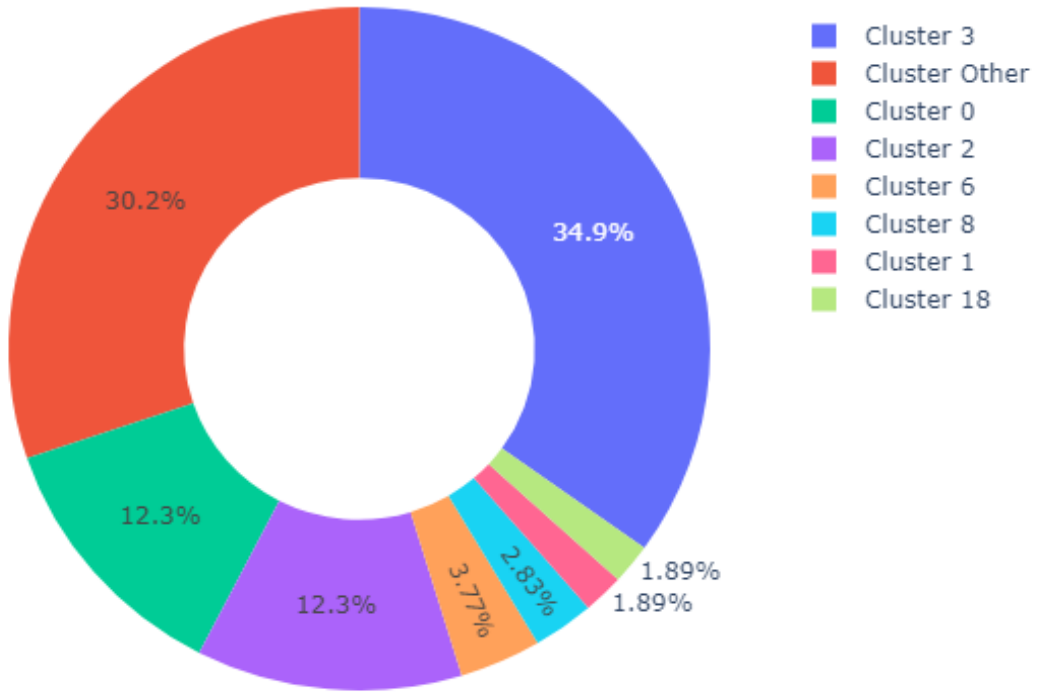
- **Text Clustering**

  Showcase: Hierarchical Clustering

  On our issues/incidents data (≅ 100 Samples)

Cluster Details (from manual review)	Count
Malicious code/ software/ activity	37
Unauthorized activities	13
Failure or disruption of communication links	13
Network Reconnaissance	4
Denial of service	3
Network outage	2
Brute force	2
Other	32

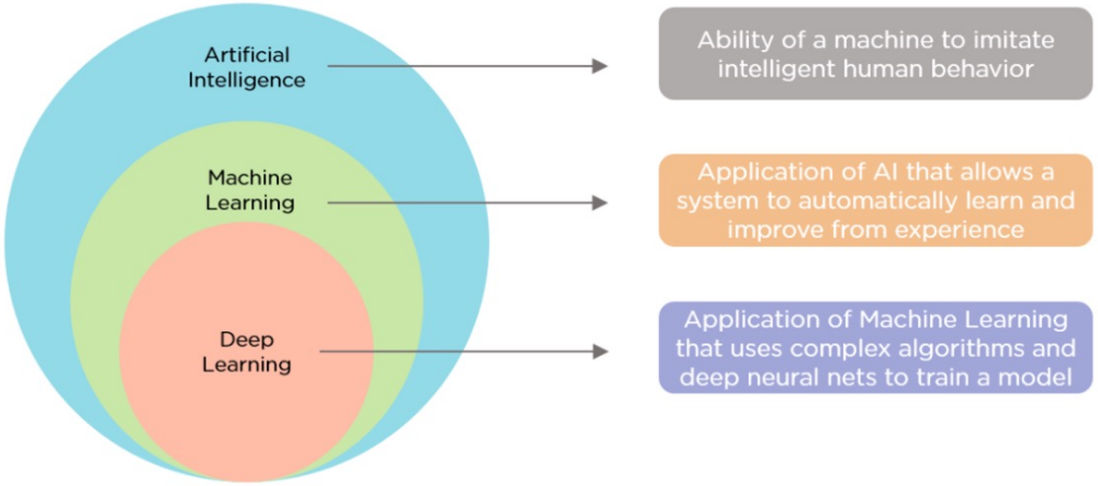
Donut Chart

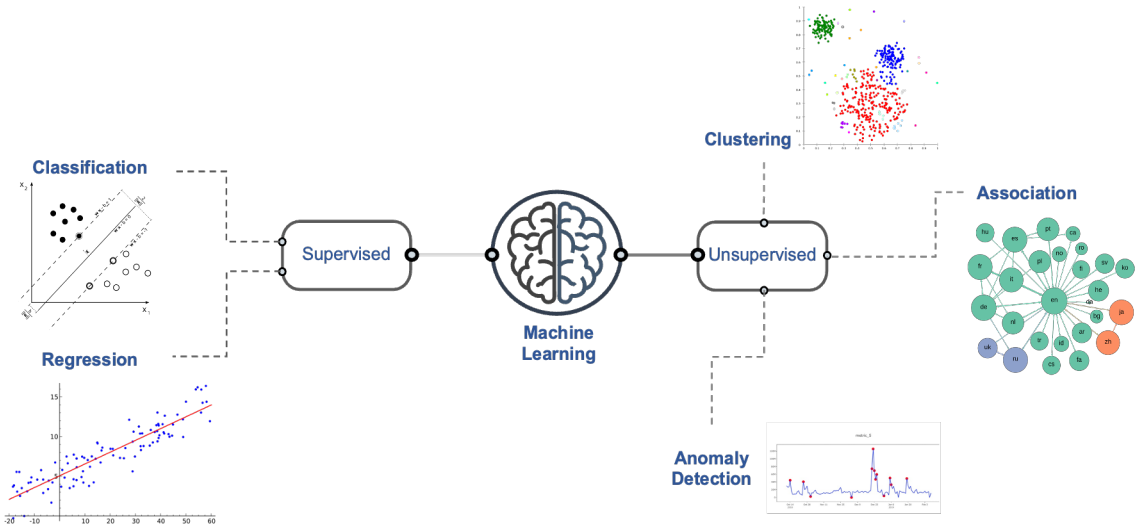


AI Adoption in Cyber Defense (Conclusion)

Key Issues	Use Cases	ML Algorithms	Proactive/ Reactive
Threat detection (False Positive, False Negative, MTTD)	UEBA	Anomaly Detect (Isolation Forest (Unsupervised), Neural Network (Supervised))	Proactive/ Reactive
Threat response time (MTTR)			
New threat identification (Zero-day Attack)			
Staffing capacity and expertise			
Large volume of cyber alerts			
How to manage?	Threat Category, Prioritization	Text Similarity, Text Clustering	Proactive

AI Adoption in Cyber Defense
Key Takeaways





Key factors to consider when implementing AI solutions

1. **Domain Expert (cybersecurity)**
2. **Data Quality**
3. **Model Selection**
4. **Hardware**
5. **Data Understanding (especially in sensitive fields like healthcare or finance)**
6. **Security and Privacy**
7. **Scalability**
8. **Ethical Implications (such as bias and fairness)**
9. **Integration**
10. **Maintenance (AI require regular maintenance and updates)**
11. **Monitoring**
12. **Bug in AI (vulnerable to attack)**

CYBER ELITE

AI Adoption in Cyber Defense (Conclusion)

Key Issues	Use Cases	ML Algorithms	Proactive/ Reactive
Threat detection (False Positive, False Negative, MTTD)	UEBA	Anomaly Detect (Isolation Forest (Unsupervised), Neural Network (Supervised))	Proactive/ Reactive
Threat response time (MTTR)	Root Cause Analysis	Time Series Anomaly Detect, Pattern Recognition, Supervised Attack Pattern (Neural Network)	Proactive/ Reactive
New threat identification (Zero-day Attack)			
Staffing capacity and expertise	AI Adoption	Unsupervised and Supervised Algorithms	Reactive
Large volume of cyber alerts			
How to manage?	Threat Category, Prioritization	Text Similarity, Text Clustering	Proactive

Get in Touch with CYBER ELITE

CYBER ELITE

BCG
BENCHACHINDA GROUP
SINCE 1960

094 480 4838

SALES@CYBERELITE.CO

WWW.CYBERELITE.CO