RSAConference2018
San Francisco | April 16 – 20 | Moscone Center

MATTERS NOW

SESSION ID: AIR-W02

PREDICTING EXPLOITABILITY - FORECASTS FOR VULNERABILITY MANAGEMENT

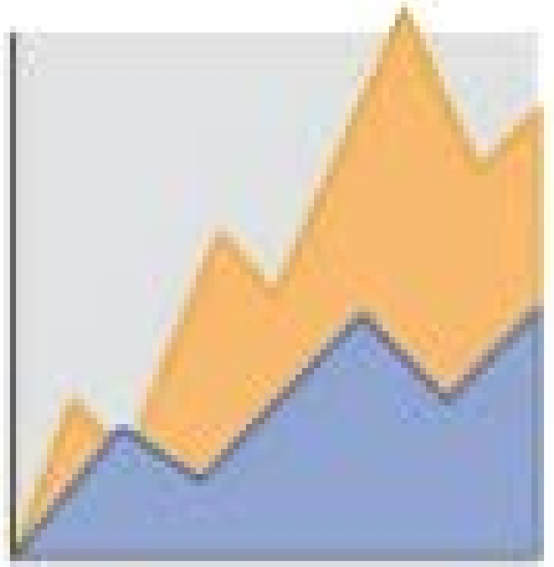
Michael Roytman

Chief Data Scientist
Kenna Security
@mroytman

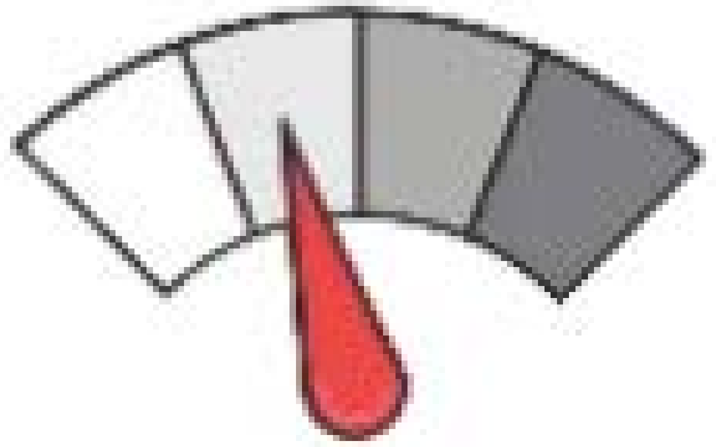
“Prediction is very difficult, especially about the future.”

-Niels Bohr

3 Types of Data-Driven



Retrospective
analysis and
reporting



Here-and-now
real-time processing
and dashboards

Predictions
to enable smart
applications

Too many vulnerabilities. How do we derive risk from vulnerability in a data-driven manner?

EXPLOITABILITY

1. RETROSPECTIVE
2. REAL-TIME
3. PREDICTIVE

EXPLOITABILITY

1. RETROSPECTIVE
2. REAL-TIME
3. PREDICTIVE

Retrospective Model: CVSS

Analyst Input

Temporal Score Estimation

Vulnerability Management Programs Augmenting Data

Vulnerability Researchers

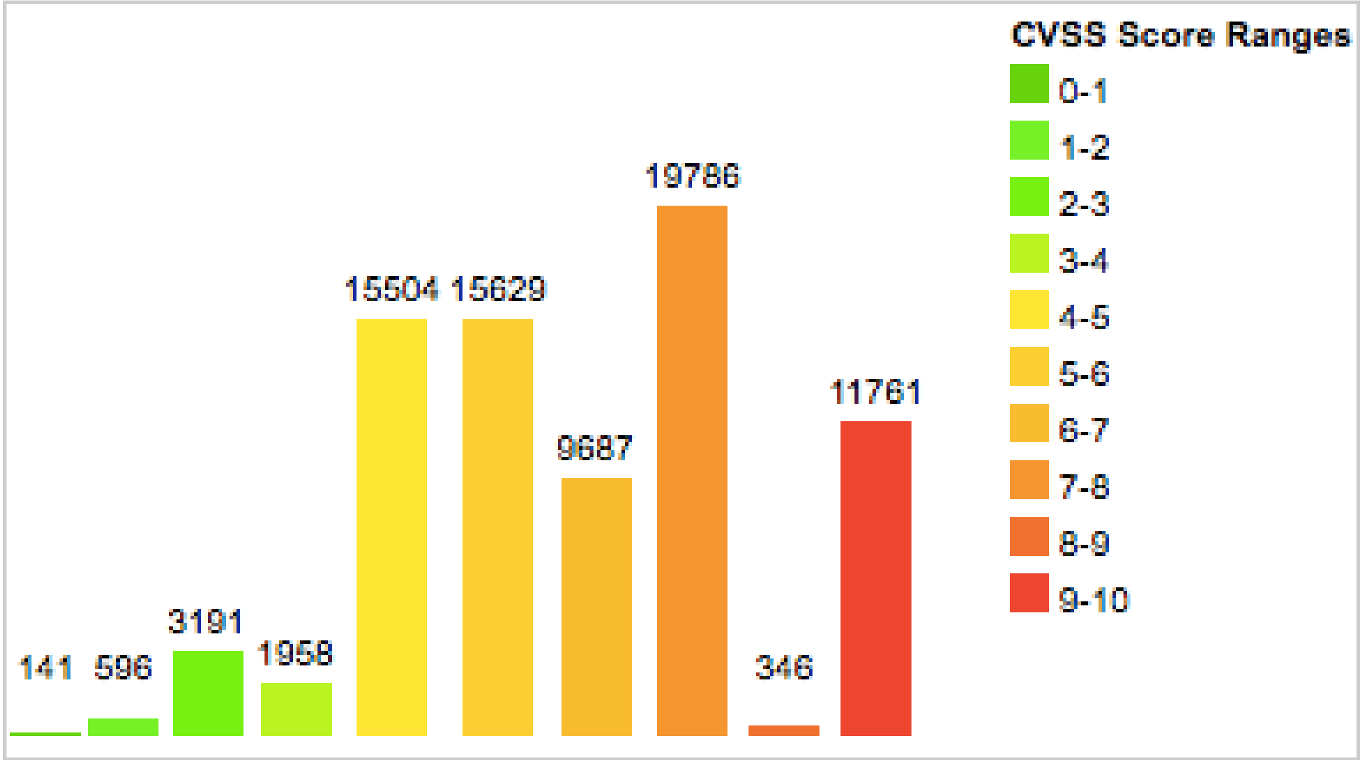
Current CVSS Score Distribution For All Vulnerabilities

Distribution of all vulnerabilities by CVSS Scores

CVSS Score	Number Of Vulnerabilities	Percentage
0-1	141	0.20
1-2	596	0.80
2-3	3191	4.10
3-4	1958	2.50
4-5	15504	19.70
5-6	15629	19.90
6-7	9687	12.30
7-8	19786	25.20
8-9	346	0.40
9-10	11761	15.00
Total	78599	

Weighted Average CVSS Score: **6.8**

Vulnerability Distribution By CVSS Scores



EXPLOITABILITY

1. RETROSPECTIVE
2. REAL-TIME
3. PREDICTIVE

Real-Time - The Data

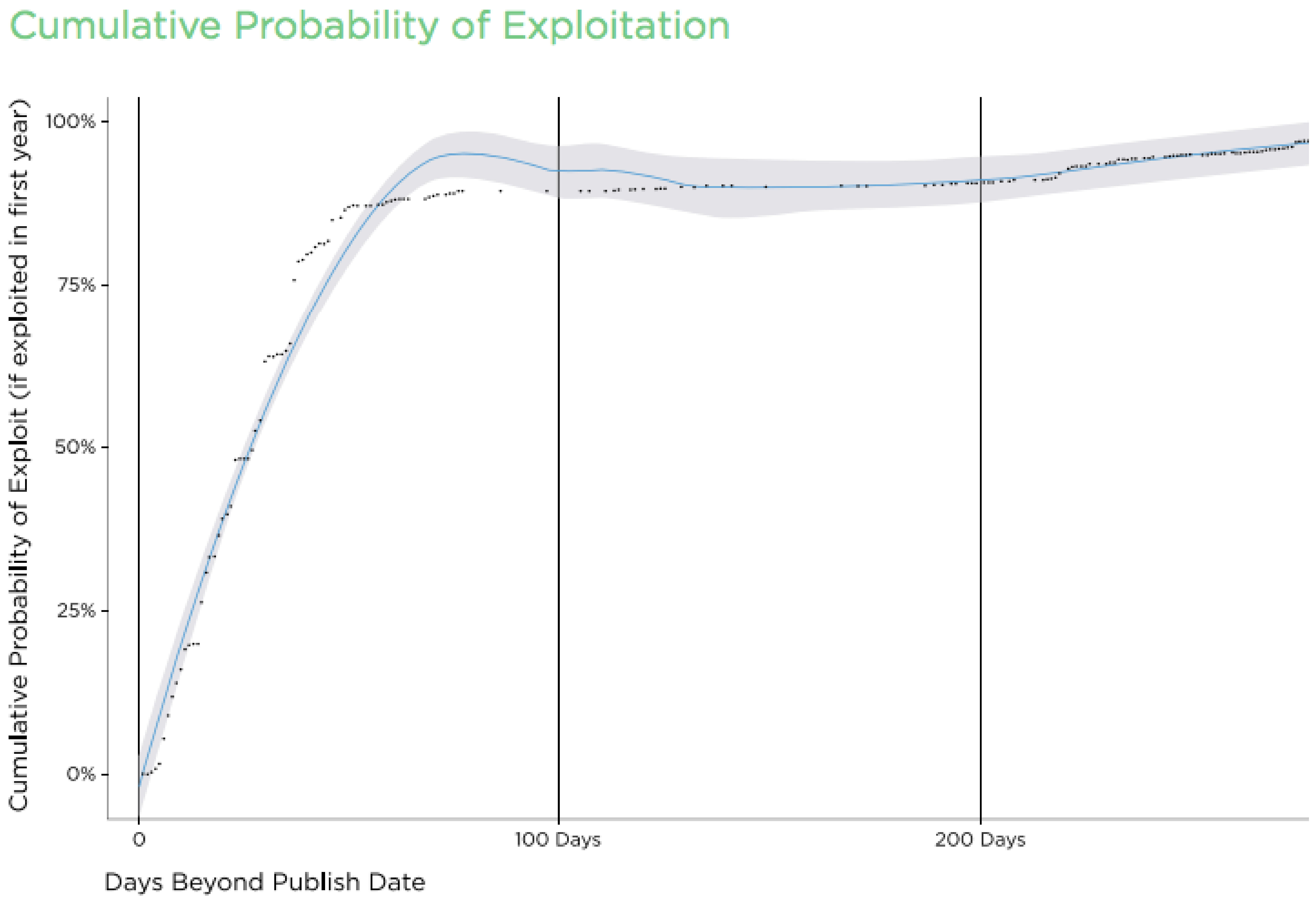
Vulnerability Scans (Qualys, Rapid7, Nessus, etc):

- 7,000,000 Assets (desktops, servers, urls, ips, macaddresses)
- 1,400,000,000 Vulnerabilities (unique asset/CVE pairs)

Exploit Events - Successful Exploitations

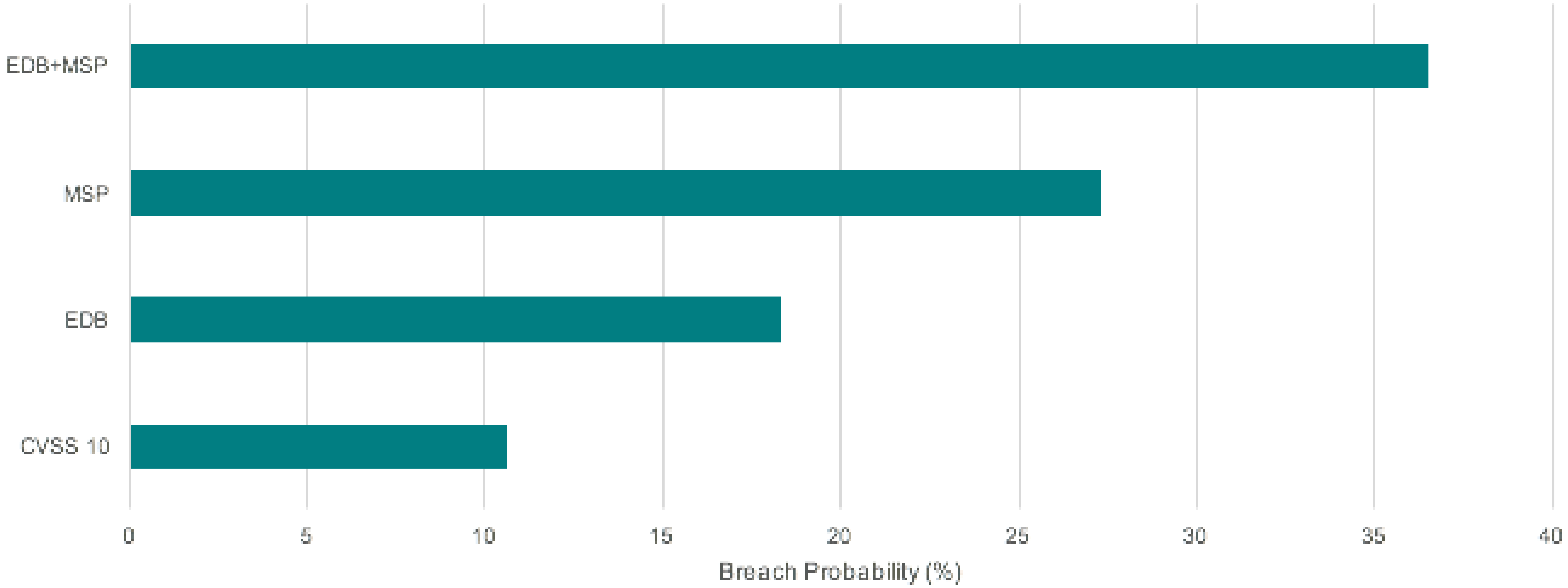
- ReversingLabs' backend metadata
  - Hashes for each CVE
  - Number of found pieces of malware corresponding to each hash
- Alienvault Backdoor
- "attempted exploits" correlated with open vulnerabilities

Attackers Are Fast



KENNA Security

Positive Predictive Value of Remediating:



KENNA Security

Q: "Of my current vulnerabilities, which ones should I remediate?"

A: Old ones with stable, weaponized exploits

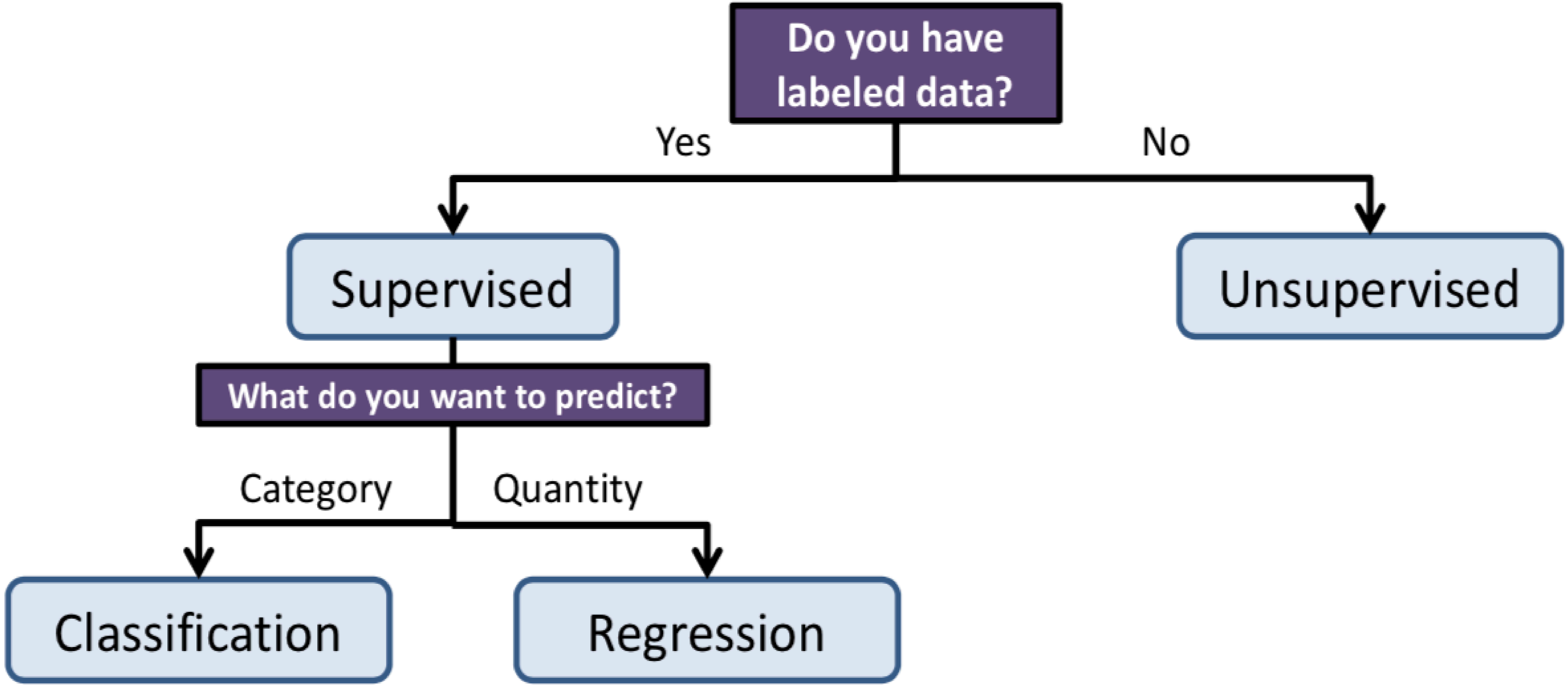
Q: "A new vulnerability was just released. Do we scramble?"

A: 

EXPLOITABILITY

1. RETROSPECTIVE
2. REAL-TIME
3. PREDICTIVE

Learning Machine Learning



KENNA Security

Classification!

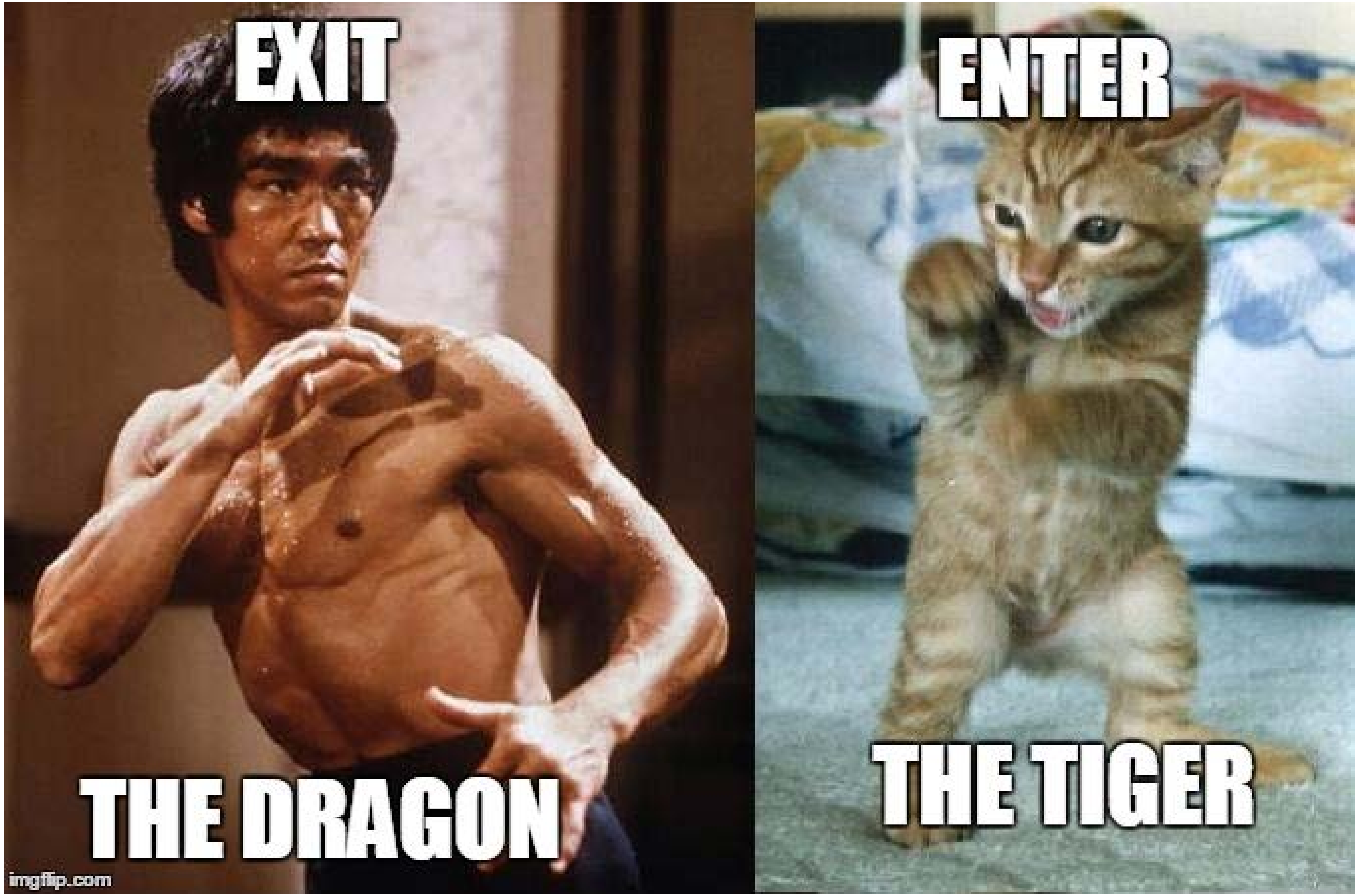
VS

The Future

- **Classification:** output is qualitative
- prediction:

“Will this vulnerability have an exploit written for it?”
(== cause more risk *later*)

Enter: AWS ML



The Data

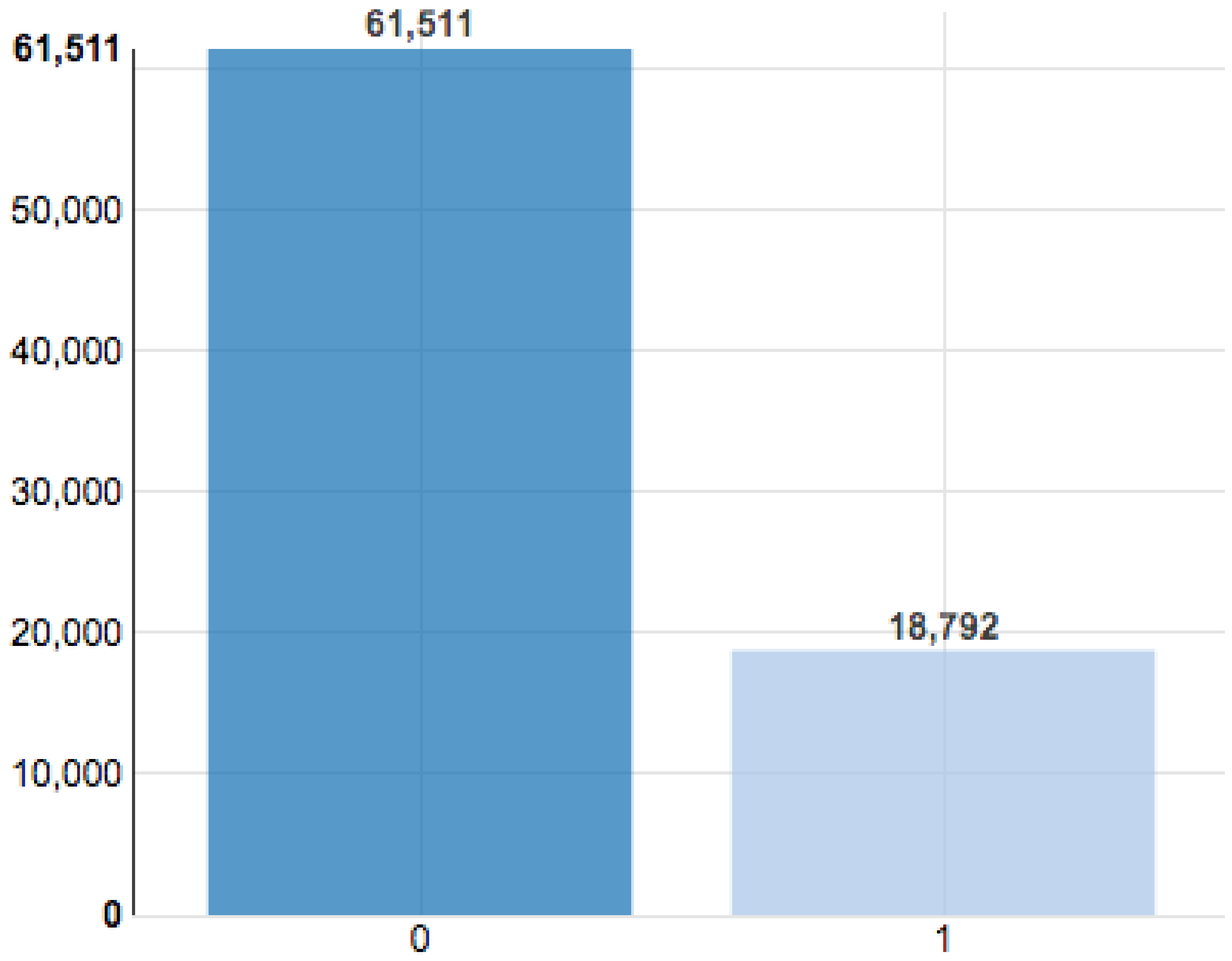
All CVE. Described By:

1. National Vulnerability Database
2. Common Platform Enumeration
3. Occurrences in Kenna Scan Data

Labelled as Exploit Available/Not:

1. Exploit DB
2. Metasploit
3. D2 Elliot/Canvas
4. Blackhat Exploit Kits

N = 81303



KENNA Security

All Models:

70% Training, 30% Evaluation Split

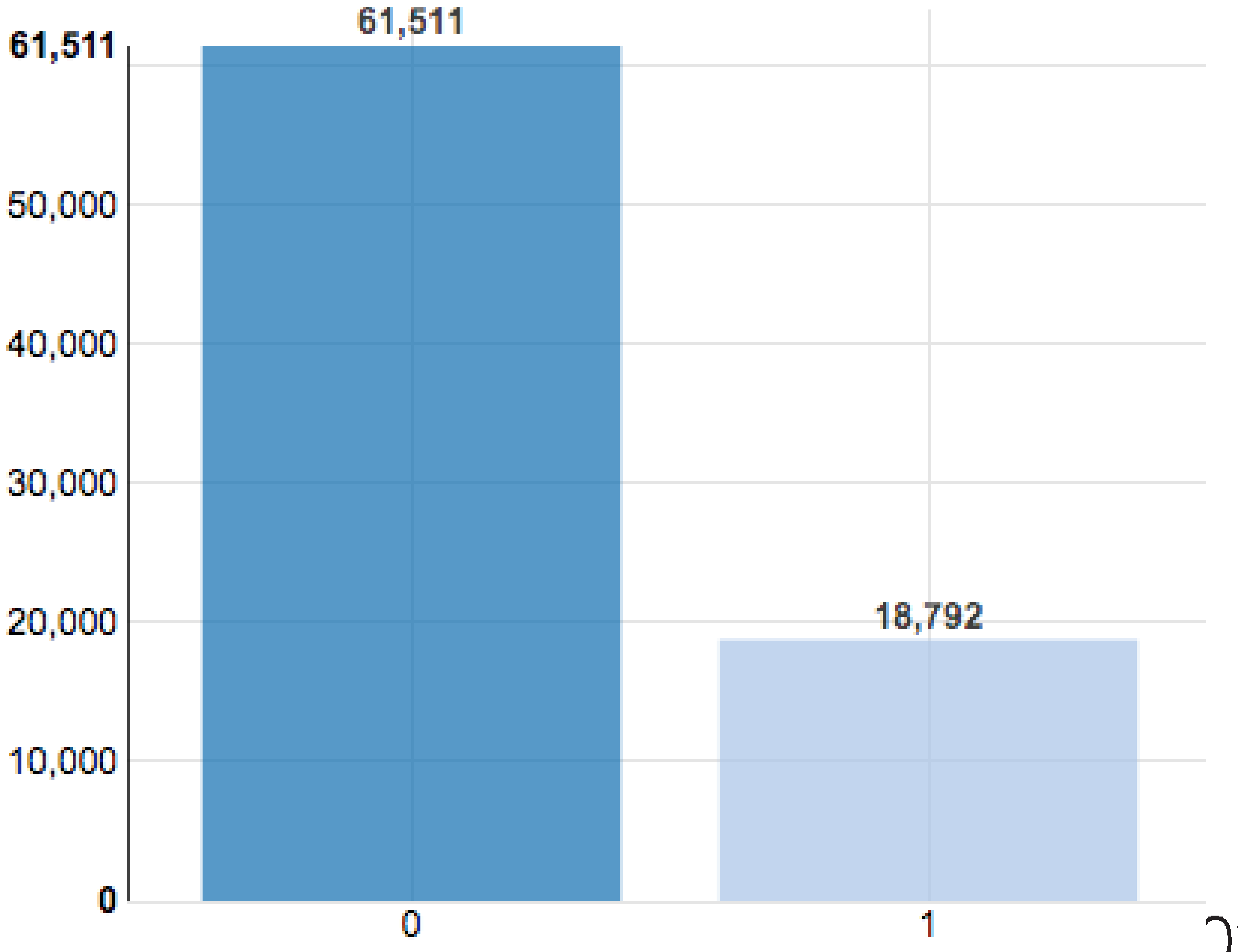
L2 regularizer

1 gb

100 passes over the data

Receiver operating characteristics for comparisons

N = 81303



KENNA Security

Predictive - The Expectations

Distribution is not uniform. 77% of dataset is not exploited

1. Accuracy of 77% would be bad

Precision matters more than Recall

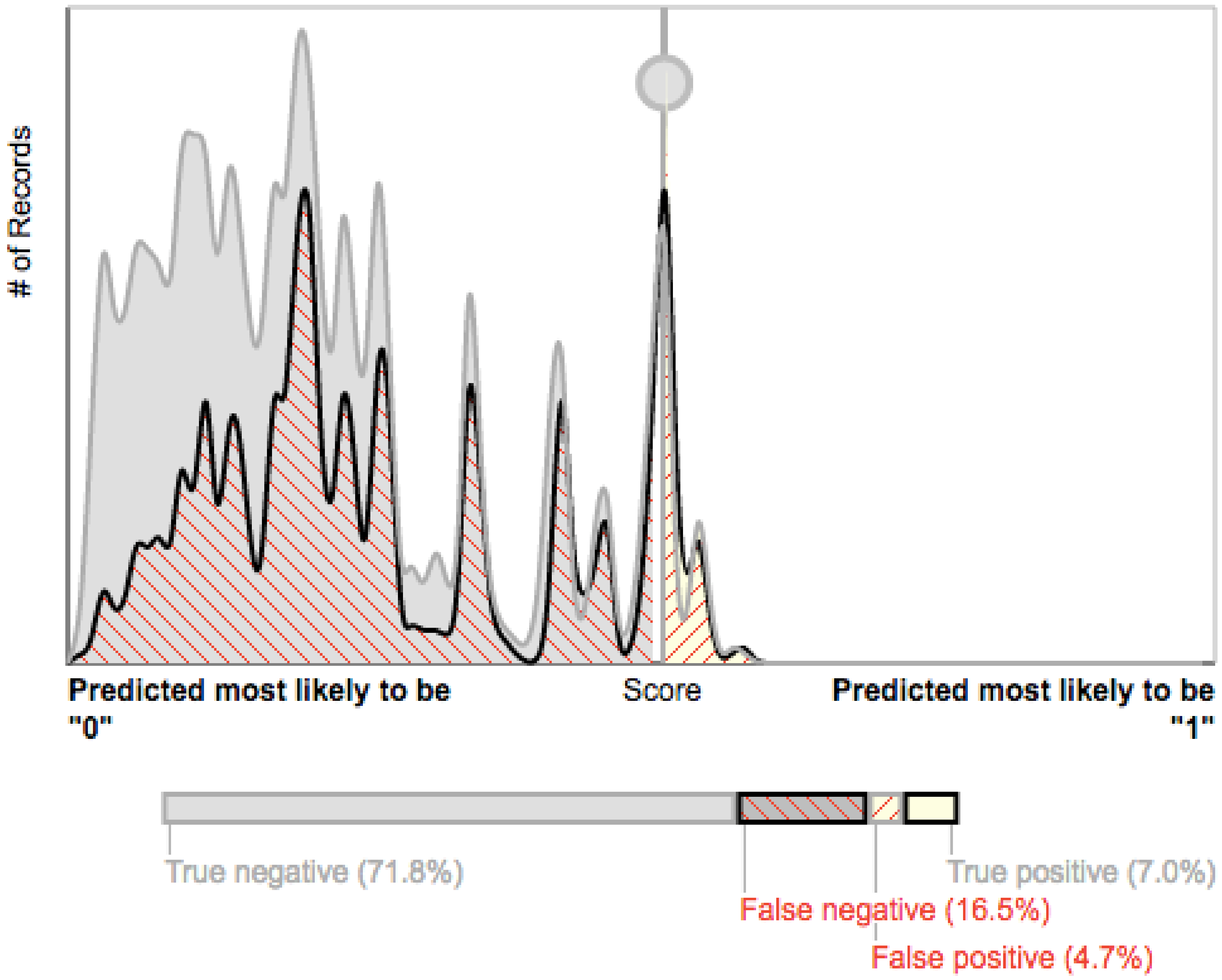
1. No one would use this model absent actual exploit available data.
2. False Negatives matter less than false positives - wasted effort

We are not modeling when something will be exploited, just IF

Could be tomorrow or in 6 months. Re-run the model every day

Model 1: Baseline

-CVSS Base
-CVSS Temporal
-Remote Code Execution
-Availability
-Integrity
-Confidentiality
-Authentication
-Access Complexity
-Access Vector
-Publication Date



- **79% are correct**
  1,699 true positive
  17,517 true negative
- **21% are errors**
  1,153 false positive
  4,020 false negative

- 12% of the records are predicted as "1"
- 88% of the records are predicted as "0"

False positive rate **0.0618**

Precision **0.5957**

Recall **0.2971**

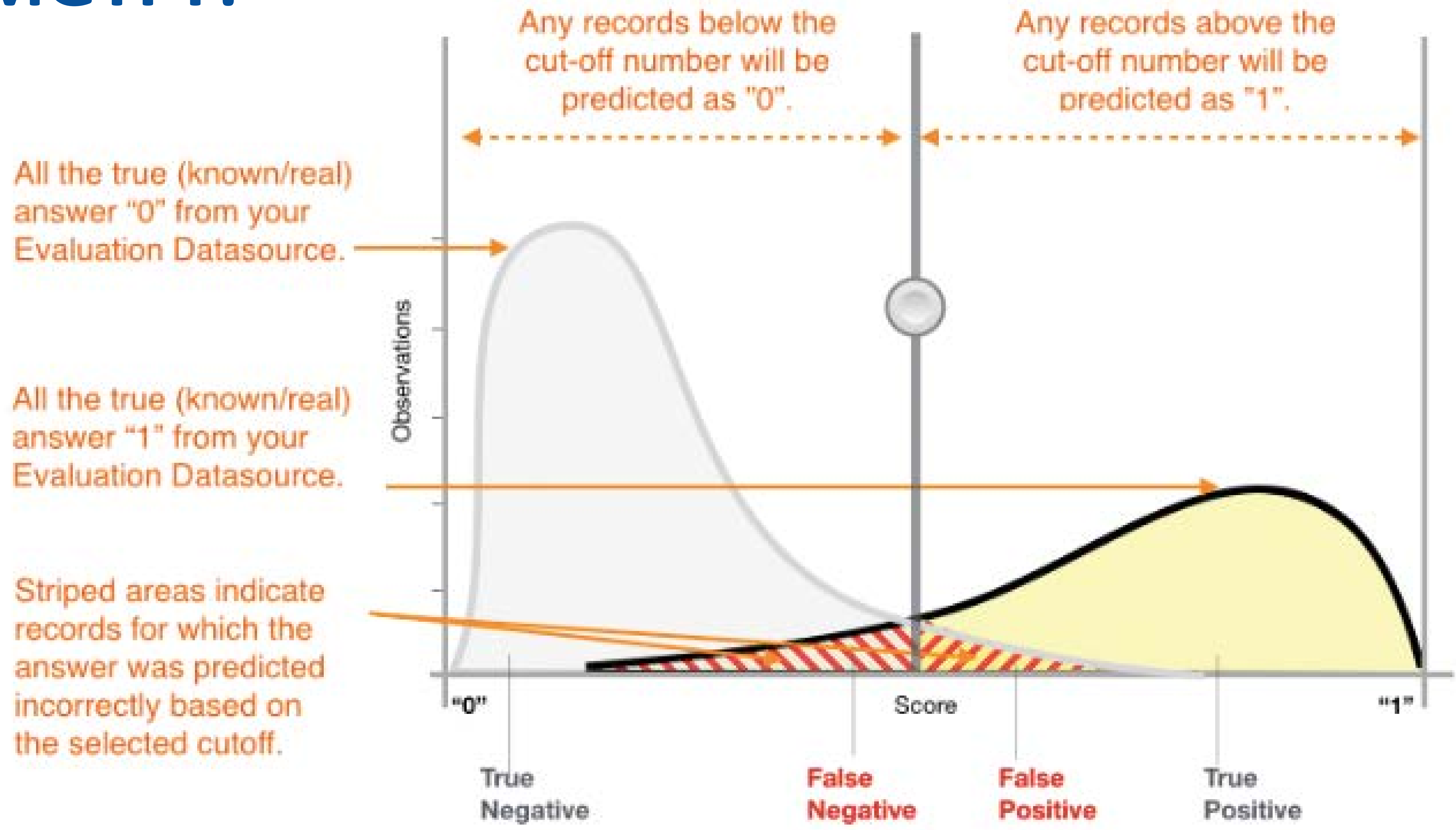
Accuracy **0.7879**

KENNA Security

RSAConference2018

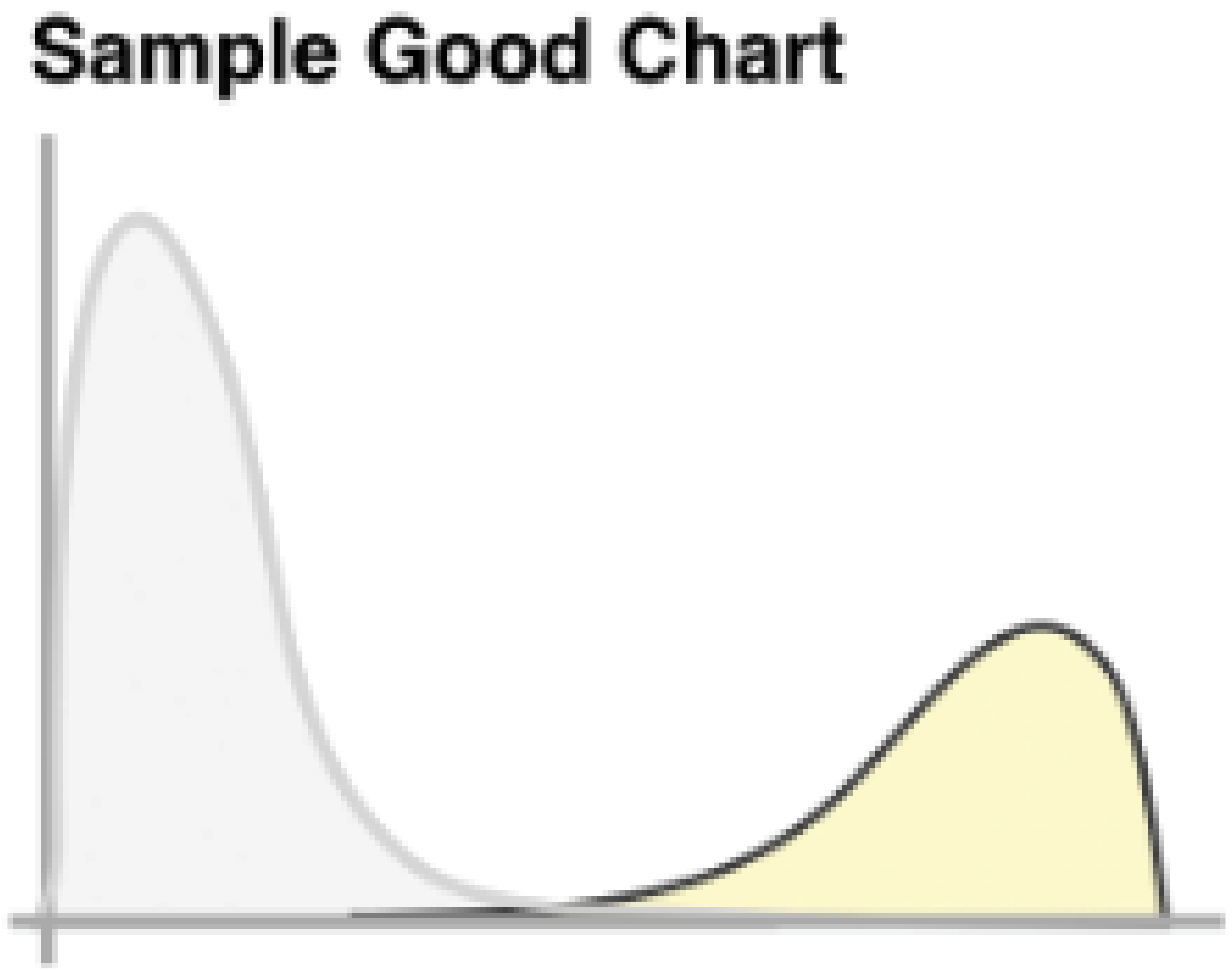
#RSAC

LMGTFY:

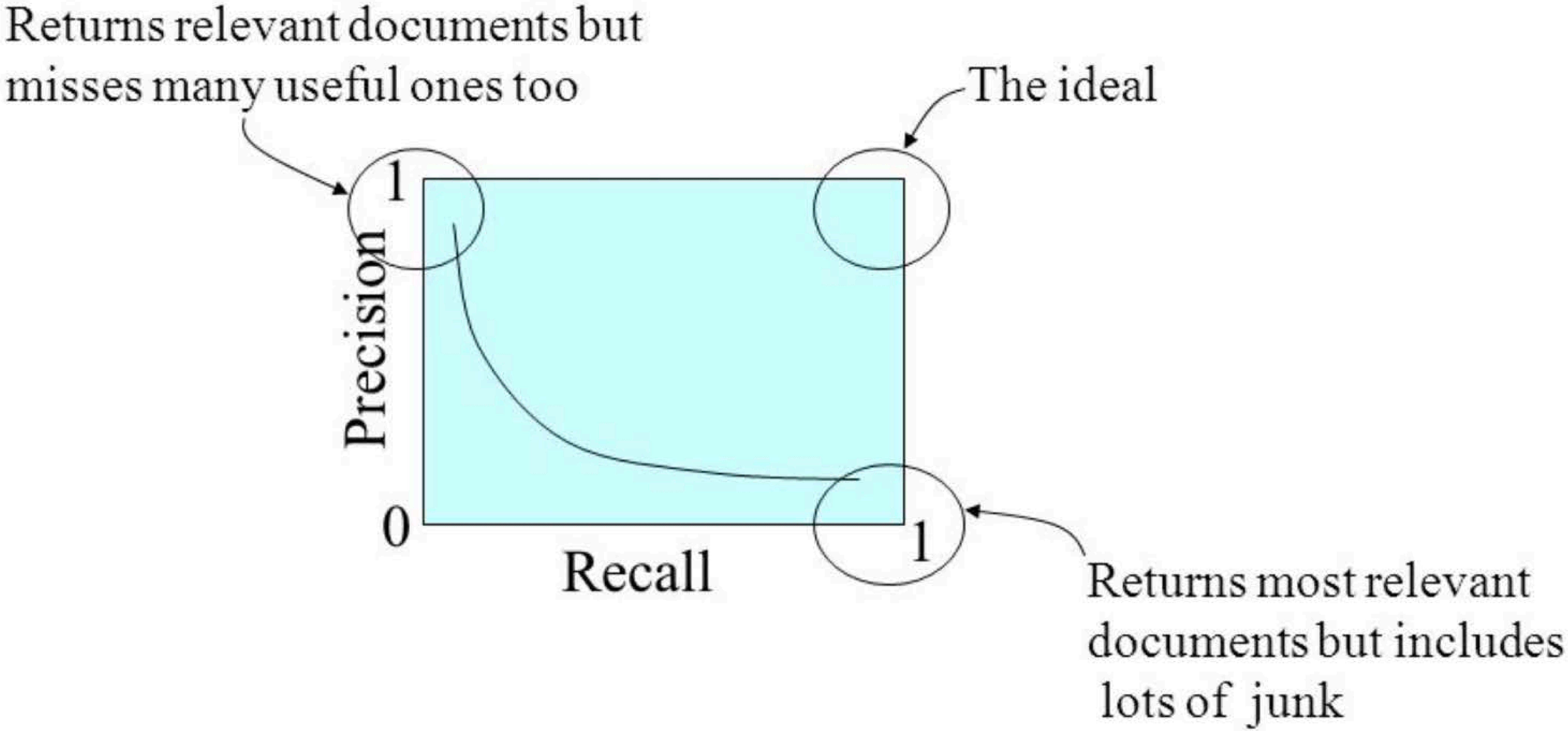


Any records below the cut-off number will be predicted as "0".
Any records above the cut-off number will be predicted as "1".
All the true (known/real) answer "0" from your Evaluation Datasource.
All the true (known/real) answer "1" from your Evaluation Datasource.
Striped areas indicate records for which the answer was predicted incorrectly based on the selected cutoff.
Observations
"0"
Score
"1"
True Negative
False Negative
False Positive
True Positive

Moar Simple?

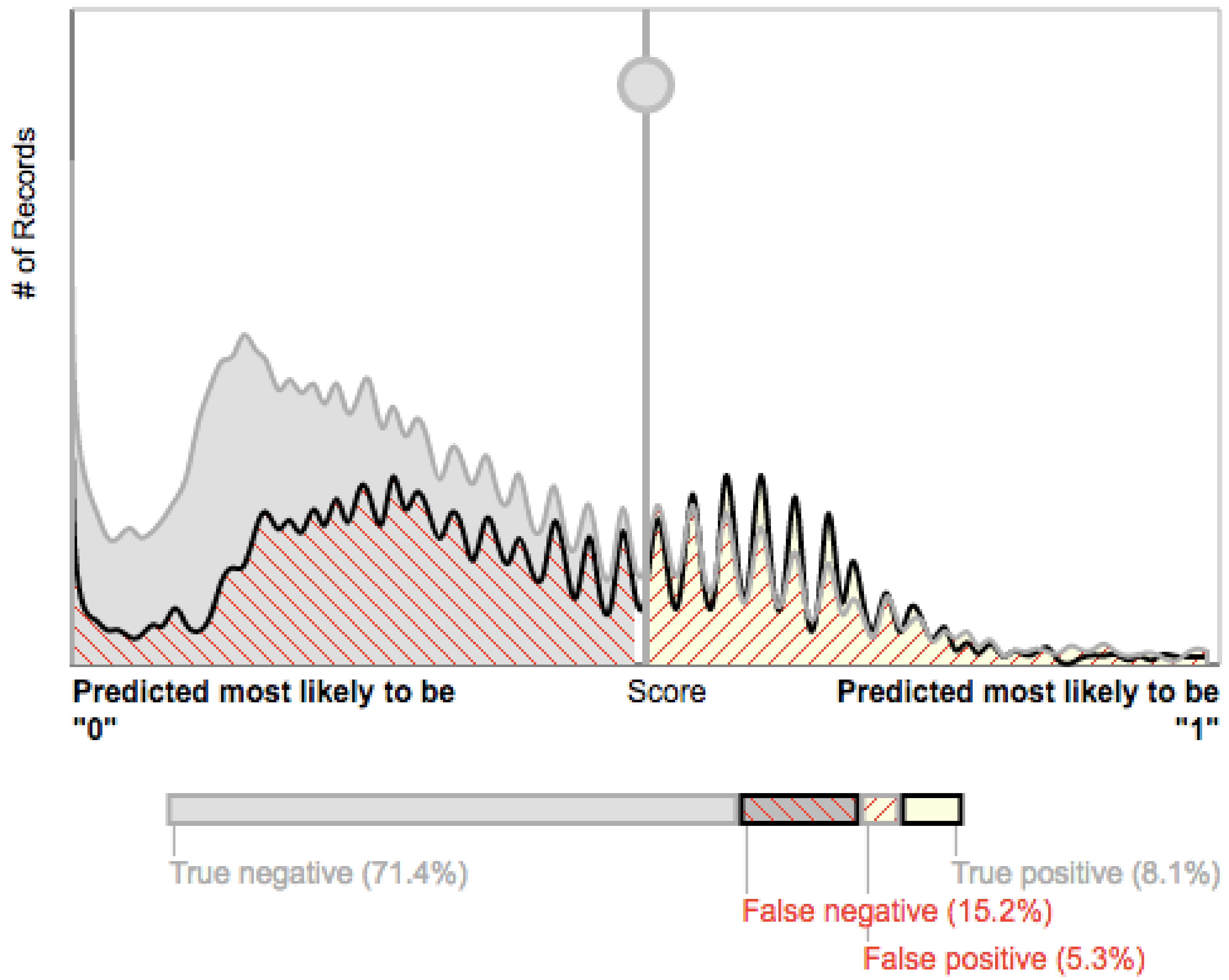


Measuring Performance



Model 2: Patches

-CVSS Base
-CVSS Temporal
-Remote Code Execution
-Availability
-Integrity
-Confidentiality
-Authentication
-Access Complexity
-Access Vector
-Publication Date
-Patch Exists



- **79% are correct**
  1,965 true positive
  17,294 true negative
- **21% are errors**
  1,280 false positive
  3,687 false negative

- 13% of the records are predicted as "1"
- 87% of the records are predicted as "0"

False positive rate **0.0689**

Precision **0.6055**

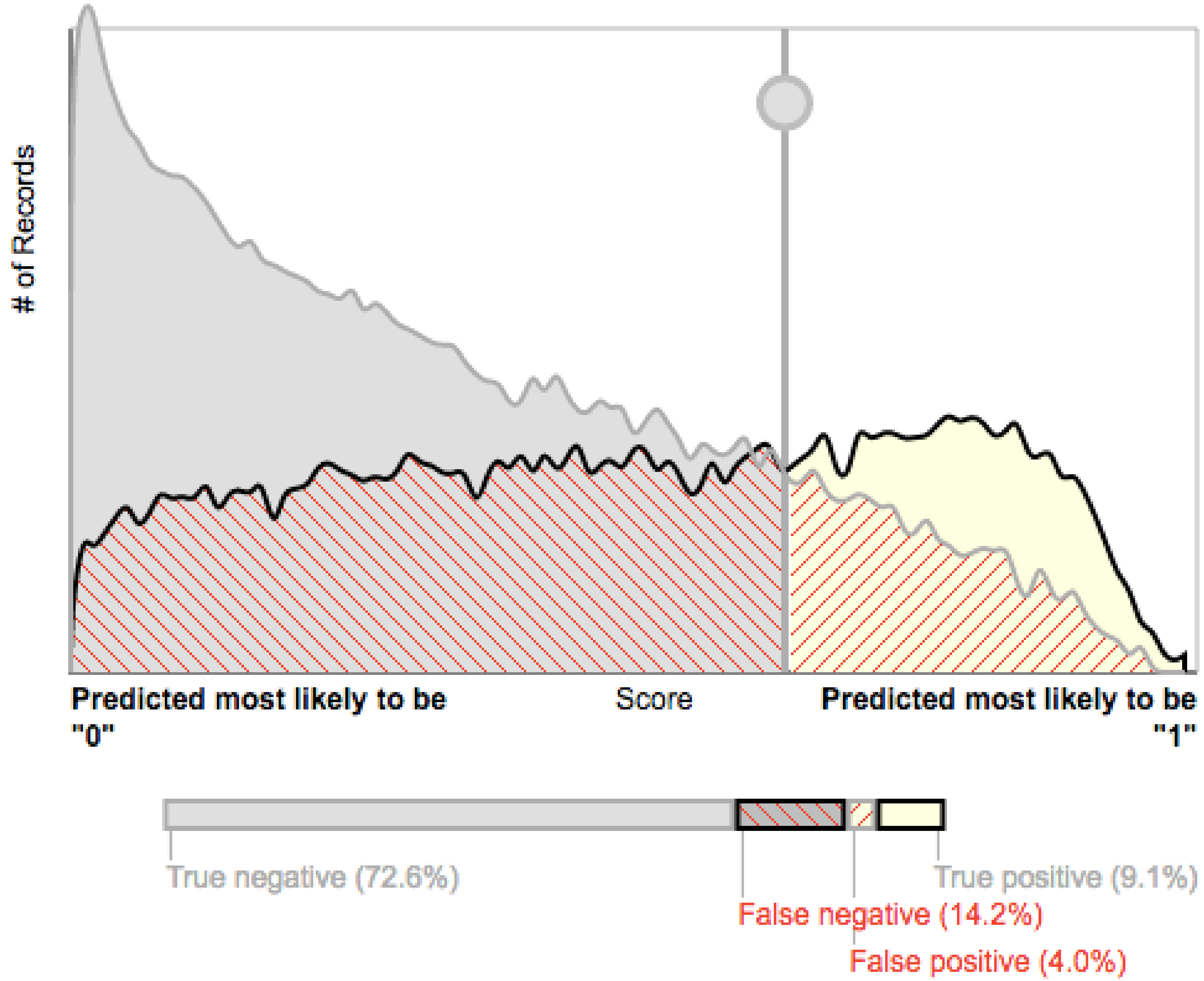
Recall **0.3477**

Accuracy **0.795**

KENNA Security

Model 3: Affected Software

-CVSS Base
-CVSS Temporal
-Remote Code Execution
-Availability
-Integrity
-Confidentiality
-Authentication
-Access Complexity
-Access Vector
-Publication Date
-Patch Exists
-Vendors
-Products



- **82% are correct**
  2,209 true positive
  17,595 true negative
- **18% are errors**
  979 false positive
  3,443 false negative

- 13% of the records are predicted as "1"
- 87% of the records are predicted as "0"

False positive rate **0.0527**

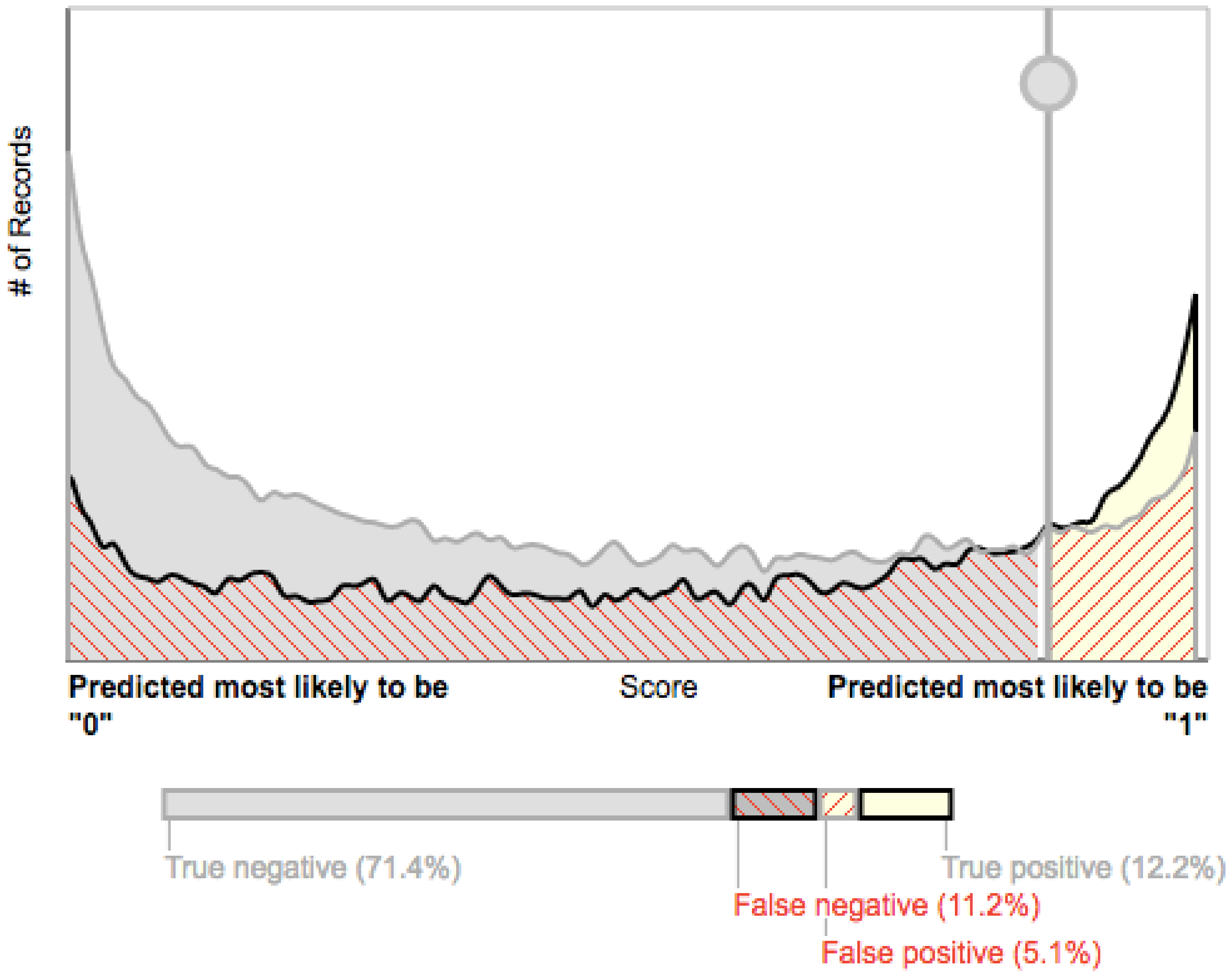
Precision **0.6929**

Recall **0.3908**

Accuracy **0.8175**

Model 4: Words!

-CVSS Base
-CVSS Temporal
-Remote Code Execution
-Availability
-Integrity
-Confidentiality
-Authentication
-Access Complexity
-Access Vector
-Publication Date
-Patch Exists
-Vendors
-Products
-Description, Ngrams 1-5



- **84% are correct**
  2,983 true positive
  17,418 true negative
- **16% are errors**
  1,252 false positive
  2,736 false negative

- 17% of the records are predicted as "1"
- 83% of the records are predicted as "0"

False positive rate **0.0671**

Precision **0.7044**

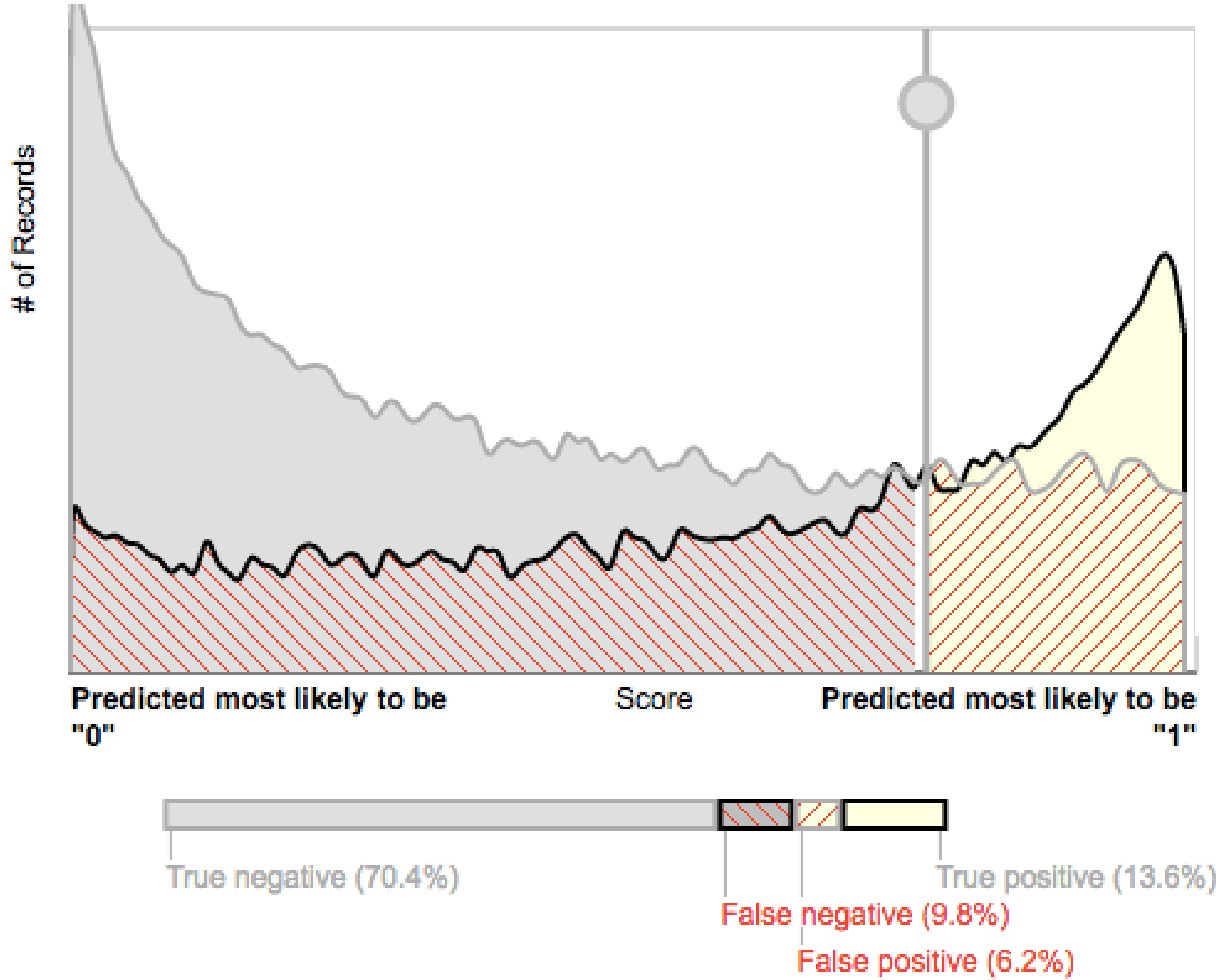
Recall **0.5216**

Accuracy **0.8365**

KENNA Security

Model 5: Vulnerability Prevalence

-CVSS Base
-CVSS Temporal
-Remote Code Execution
-Availability
-Integrity
-Confidentiality
-Authentication
-Access Complexity
-Access Vector
-Publication Date
-Patch Exists
-Vendors
-Products
-Description, Ngrams 1-5
-Vulnerability Prevalence
-Number of References



- **84% are correct**
  3,318 true positive
  17,169 true negative
- **16% are errors**
  1,501 false positive
  2,401 false negative

- 20% of the records are predicted as "1"
- 80% of the records are predicted as "0"

False positive rate **0.0804**

Precision **0.6885**

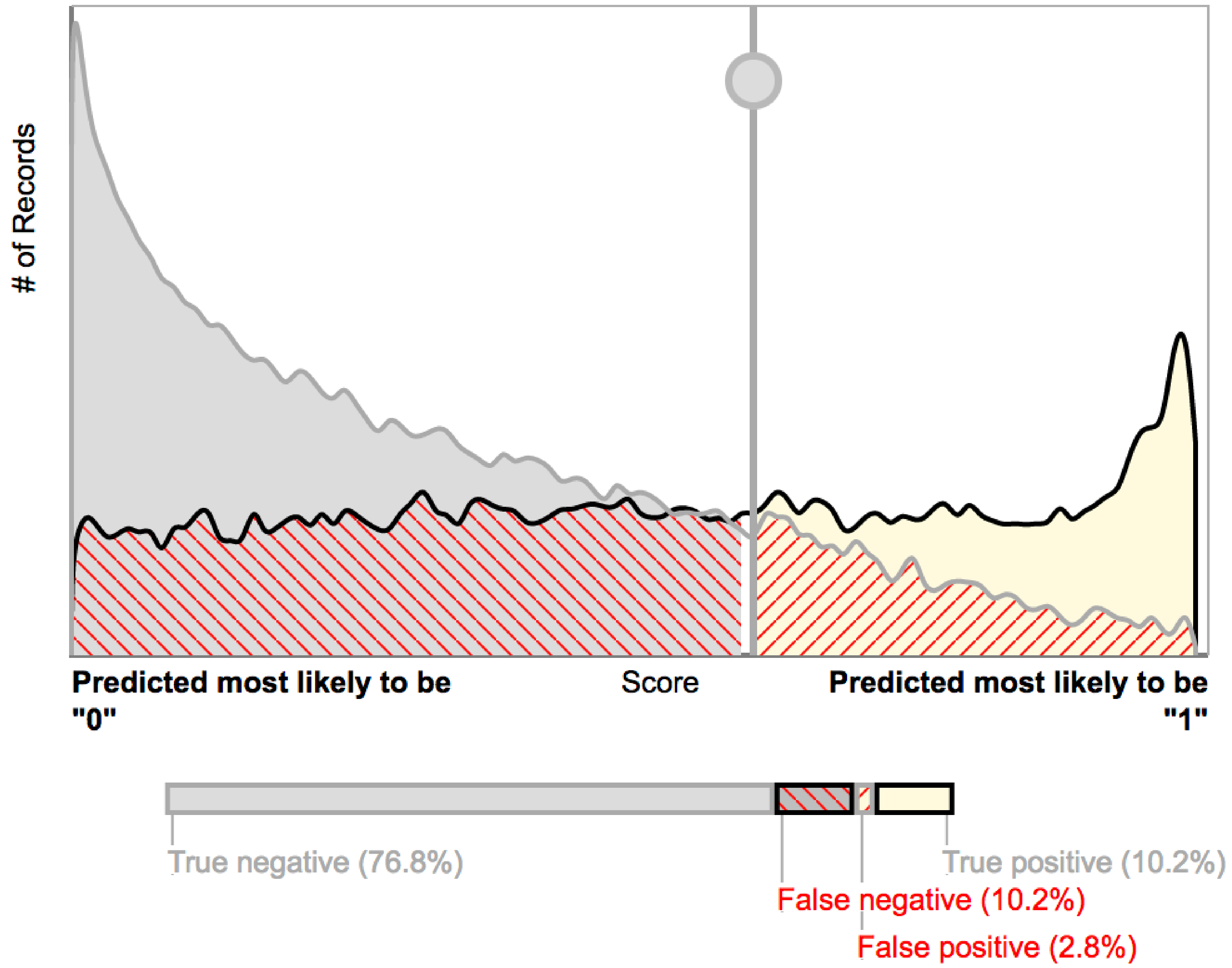
Recall **0.5802**

Accuracy **0.84**

KENNA Security

RSAConference2018 #RSAC MATTERS NOW

Model 6: ”Somewhat Likely”



Disable real time predictions to update the threshold.

Trade-off based on score threshold 0.6

Reset score threshold (0.6)

- **87% are correct**
  2,868 true positive
  21,646 true negative
- **13% are errors**
  795 false positive
  2,878 false negative

- 13% of the records are predicted as "1"
- 87% of the records are predicted as "0"

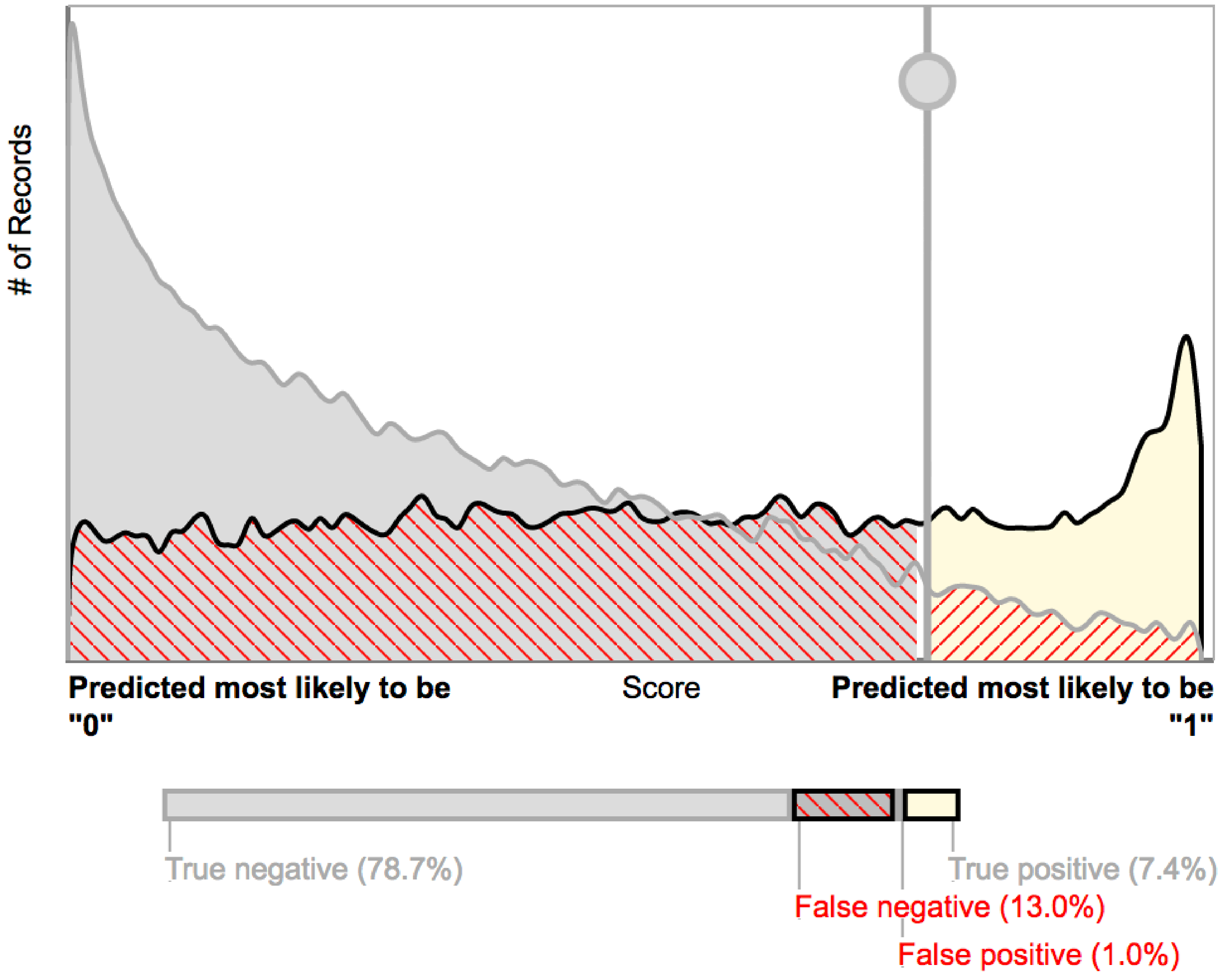
Save score threshold at 0.60

Advanced metrics

False positive rate **0.0354**	0	1
Precision **0.783**	0	1
Recall **0.4991**	0	1
Accuracy **0.8697**	0	1

KENNA Security

Model 6: ”Highly Likely”



Disable real time predictions to update the threshold.

Trade-off based on score threshold 0.75 Reset score threshold (0.6)

- **86% are correct**
  2,093 true positive
  22,172 true negative
- **14% are errors**
  269 false positive
  3,653 false negative

- 8% of the records are predicted as "1"
- 92% of the records are predicted as "0"

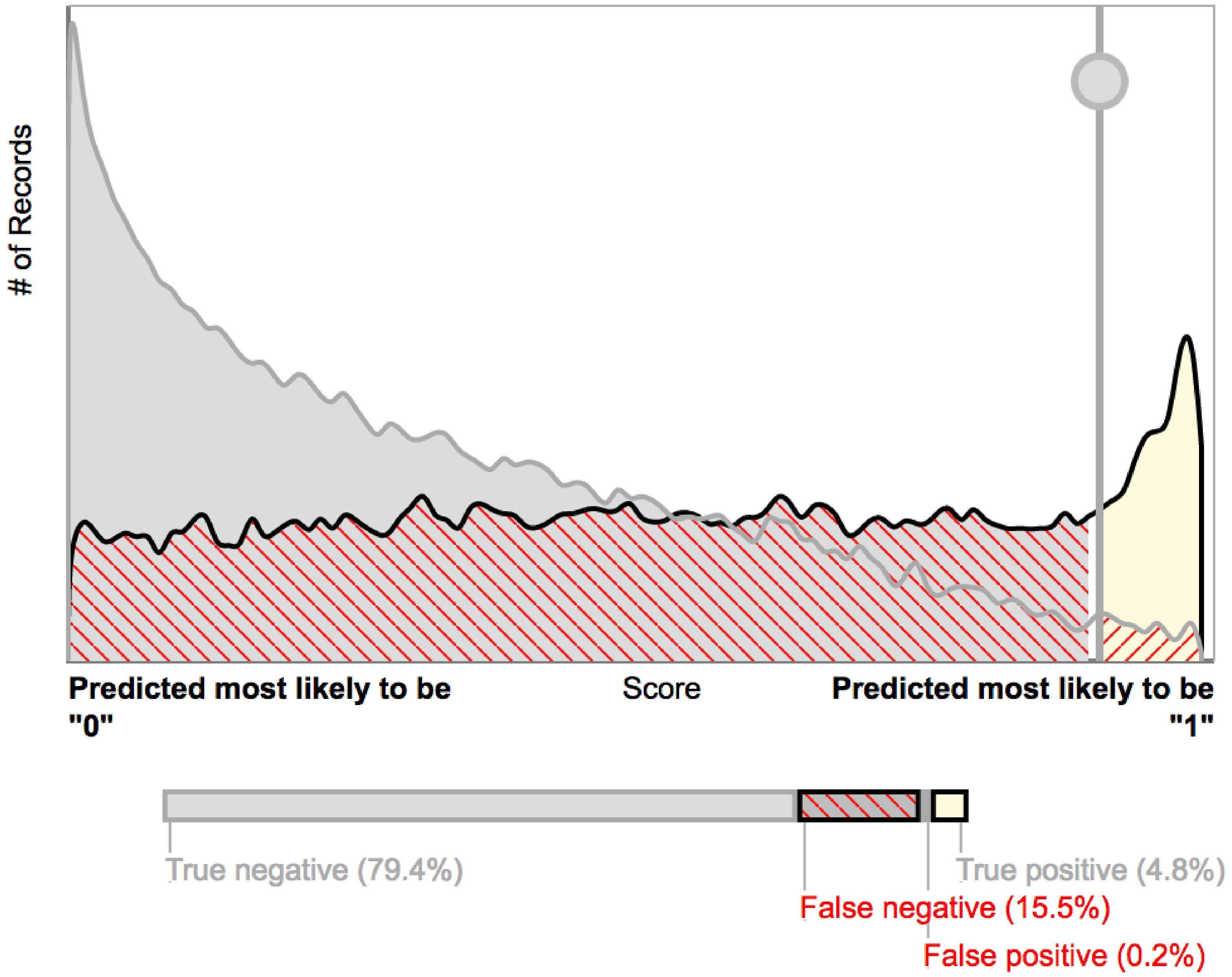
Save score threshold at 0.75

Advanced metrics

False positive rate **0.012**	0		1
Precision **0.8861**	0		1
Recall **0.3643**	0		1
Accuracy **0.8609**	0		1

KENNA Security

Model 6: "Most Likely"



Disable real time predictions to update the threshold.

Trade-off based on score threshold 0.9

Reset score threshold (0.6)

- **84% are correct**
  1,363 true positive
  22,372 true negative
- **16% are errors**
  69 false positive
  4,383 false negative

- 5% of the records are predicted as "1"
- 95% of the records are predicted as "0"

Save score threshold at 0.90

Advanced metrics

False positive rate **0.0031**	0		1
Precision **0.9518**	0		1
Recall **0.2372**	0		1
Accuracy **0.8421**	0		1

Future Work

-Track Predictions vs. Real Exploits

-Integrate 20+ BlackHat Exploit Kits - FP reduction?

-Find better vulnerability descriptions - mine advisories for content? FN reduction?

-Attempt Models by Vendor

-Predict Breaches, not Exploits

PROBLEM

Too many vulnerabilities.

How do we derive risk from vulnerability in a data-driven manner?

SOLUTION

1. Gather data about known successful attack paths
2. Issue forecasts where data is lacking in order to predict new exploits
3. Gather MORE data about known successful attack paths

Takeaways

#RSAC

1. Simple, Power Questions make Machine Learning Useful in Security

2. When Risk is Rare, Precision is Difficult

3. When Precision is Difficult, Be Smart about Tradeoffs

Machine Learning = ROBOT Unicorns + Rainbows

The Takeaway

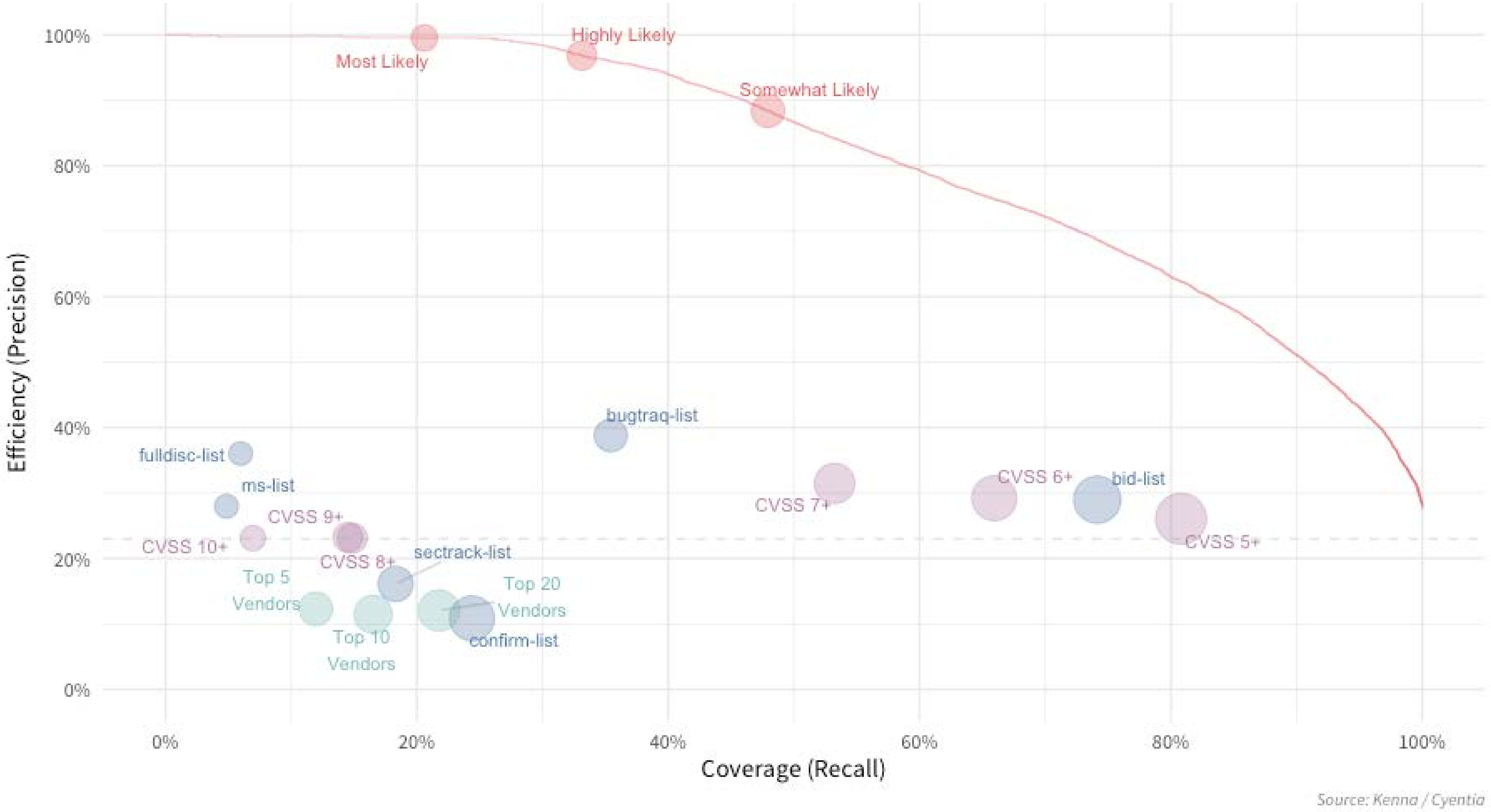
Machine Learn!

KENNA Security

Putting It All Together

Thank You for waking up so early for this!

@mroytman



www.kennasecurity.com

KENNA Security

RSA Conference2018