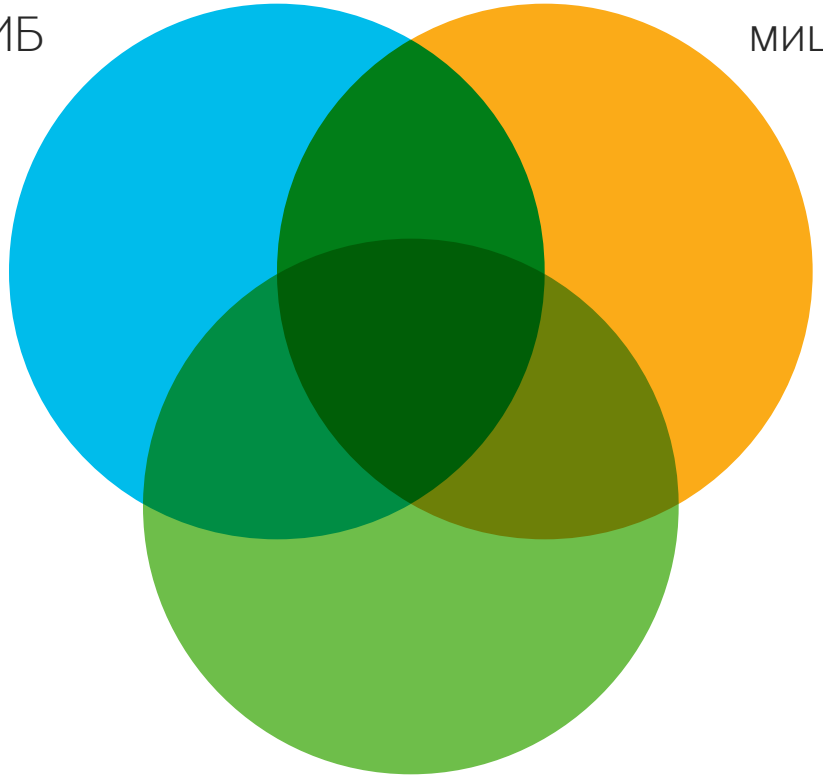
Чего стоить ждать от ИИ в кибербезопасности, а чего нет?

Краткий обзор по технологии

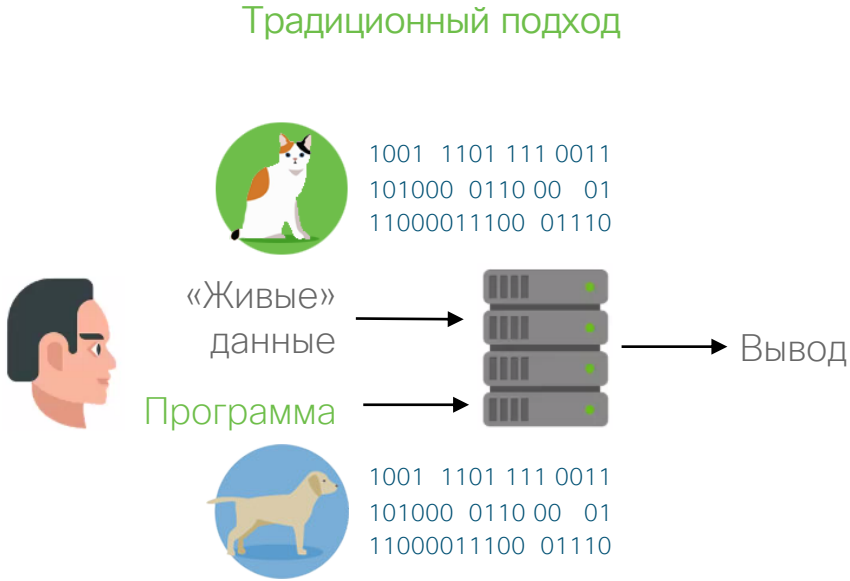
24 апреля 2020

ИИ и ИБ



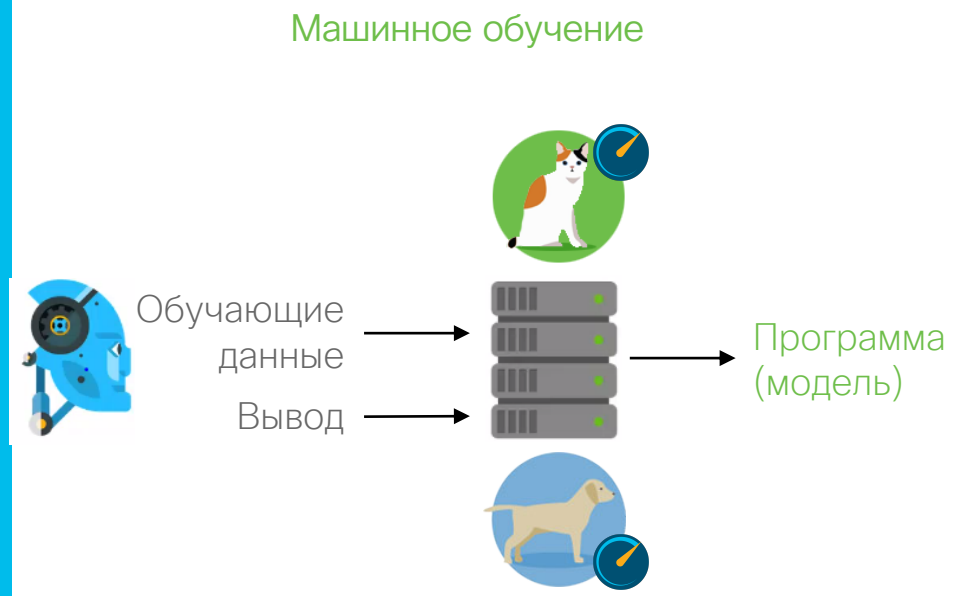
При традиционном подходе мы распознаем и заранее заносим в «черные списки» вредоносное что-то

Любая проблема, которая может быть «оцифрована» и имеет большие объемы собранных данных является кандидатом для машинного обучения



ML фундаментально отличается от обычной разработки – вы даете машине ответы и она пишет по ним код (модель) и затем дает ответы на новые данные

Машинное обучение наиболее эффективно для новых и неизвестных данных



Почему это так важно сегодня?

Вредоносная активность

Непонятная активность

Нормальная активность

Старая и новая школа ИБ

Угрозы ИБ	Старый подход	Новый подход (с ИИ)
Вредоносный код	Антивирусы, IDS и т.п. используют сигнатурные методы для детектирования вредоносного кода *Борется с известными уязвимостями*	Распознавание образов (по схожести) и предиктивная аналитика *Может ловить 0Day*
DDoS	Аналитики мониторят сетевой трафик для обнаружения DDoS-атак *Требует ресурсы, зависит от человека*	Алгоритмы автоматически детектируют ненормальную активность *Быстрая реакция, минимум ресурсов*
Фишинговые домены	Аналитики фиксируют домены, с которых реализуются атаки и помещают их в черные списки	Анализ взаимосвязей и предсказание DGA *Предотвращение*
Социальный инжиниринг	Изучение тактики хакеров *Подвержено ошибкам*	Обучение + контроль поведения + биометрия *Снижение числа ошибок*

Фокус на известное **Фокус на неизвестное**

3 компонента машинного обучения

Данные
(датасет)

Признаки

Алгоритмы

Датасеты – основная проблема (сегодня)

Каков объем и характер?

Чем больше данных для анализа, тем выше эффективность алгоритмов ML

Есть ли в свободном доступе?

Можете ли вы проверить эффективность приобретаемой или строящейся системы сами? А как сравнивать разные решения?

Кто и как подготовил?

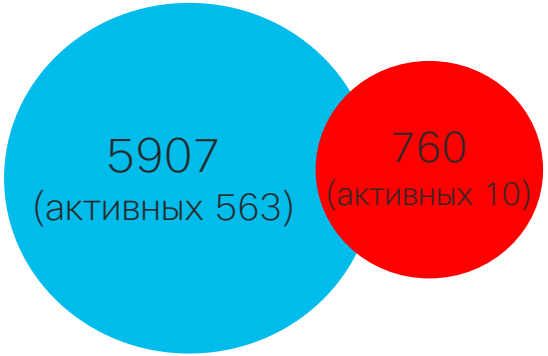
Чтобы модель ML хорошо сработала, датасет должен быть правильно подготовлен и, в ряде случаев, размечен

Пора копить свой датасет

У вас есть преимущество – вы можете собрать данные именно по вашей инфраструктуре. Но вы должны начать и вам нужно хранилище

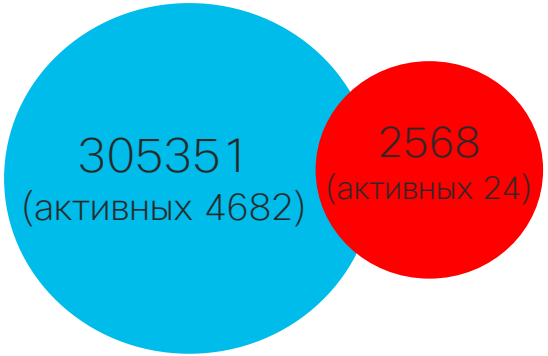
Пример: 2 российских поставщика TI

IP-адреса



Данные по C2C-фидам
за 01.02.2019-07.02.2019

Домены



По данным презентации Владимира Бенгина, Positive Technologies

Примеры датасетов по ИБ: неполны, неточны и разноформатны

1. Netresec (PCAP)
2. KDD Cup 1999
3. Web Attack Payloads
4. DARPA IDS Dataset
5. Stratosphere IPS Dataset
6. Aktaion (Bro, phishing, ransomware и др.)
7. Malicious URL Dataset (Sysnet)
8. Malware Training Set
9. Ember (для malware)
10. NSA CDX

Первое место получит тот, у кого лучше датасет!

Поэтому сегодня это конкурентное преимущество и в публичном доступе их нет. Но это пока!

Разнообразие

это ключ к эффективности машинного обучения

Антиспам

Чтобы научиться определять спам, нам нужны десятки и сотни тысяч электронных сообщений для анализа

UEBA

Чтобы научиться предсказывать поведение пользователя, нужно отслеживать все его действия в течение нескольких недель (и еще делать контрольное обучение в атипичное время – предотпускное, послеотпускное, перед корпоративом и т.п.).

Признаки: 600 на один (!) Web-запрос

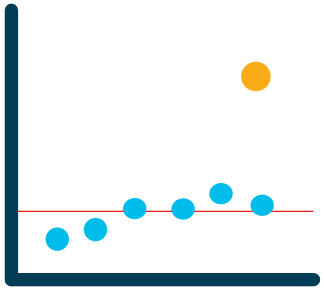
- Общие – длина, коды статуса, типы mime
- HTTP – URL, referrers, распределение символов
- HTTPS – аномальные значения, временные параметры, контекст
- Глобальные – популярность домена/AS
- Внешние – whois, сертификаты TLS

Признаки у файла

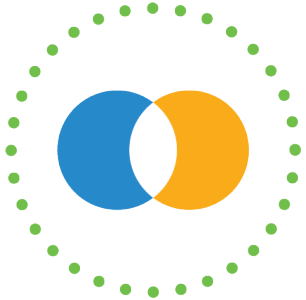
- Обычно анализируется более 800 атрибутов, полученных в процессе выполнения анализируемого файла
  - Сетевые подключения? Сканирование?
  - Нестандартные протоколы?
  - Использование интерфейсов API (каких)?
  - Обращение к реестру?
  - Работа с памятью?
  - Изменения в файловой системе?
    - Самокопирование или захват файлов
  - Запуск других процессов?
  - Окружение файла (например, e-mail, в котором он был)
  - Шифрование и энтропия

Модели/алгоритмы: почему нельзя по старинке?

Простые шаблоны
(сигнатуры, IoC…)



Статистические
методы



Правила

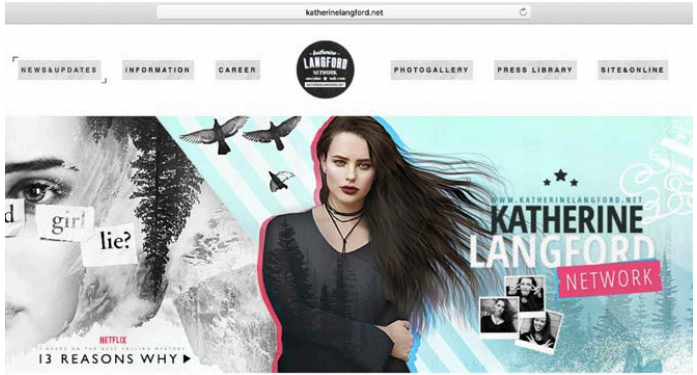
Проверим вашу наблюдательность

katherinelangford.net

VS

kyt6ea4ak4bvo35lrw.net

Да, вы правы в 100% случаев! Наверное ☺



katherinelangford.net

VS

Rovnix Malware DGA-домен

kyt6ea4ak4bvo35lrw.net

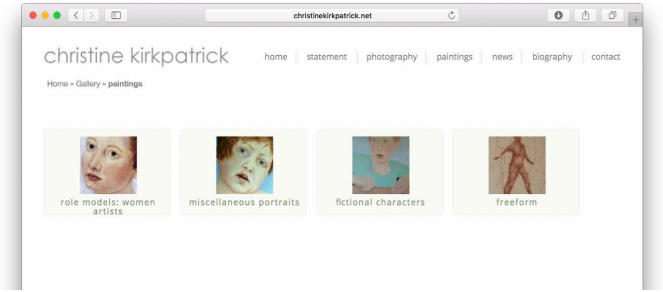
Проверим вашу наблюдательность

christinepatterson.net

VS

christinekirkpatrick.net

Пример: какой это домен?



VS

Suppobox Malware DGA-домен

christinekirkpatrick.net

christinepatterson.net

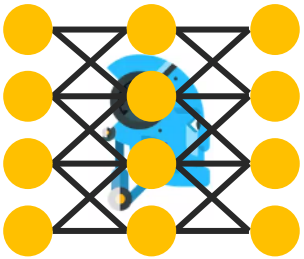
Злоумышленники не стоят на месте: NGDGA

- **Rovnix** использует текст Декларации независимости США как входные данные для Domain Generation Algorithm (DGA):
  - kingwhichtotallyadminis.biz
  - thareplunjudiciary.net
  - townsunalienable.net
  - taxeslawsmockhigh.net
  - transientperfidythe.biz
  - inhabitantslaindourmock.cn
  - thworldthesuffer.biz
- **Ботнет Matsnu** выстраивает DGA на базе словаря из 1300 существительных и глаголов (домены состоят из 24 символов)
  - scoreadmireluckapplyfitcouple.com
  - plentyclubplatewatermiss.com
  - benefitnarrowtowersliphabit.com
  - accountmoveseemsmartconcert.com
  - drawermodelattemptreview.com
  - carecommitshineshiftchip.com

Вся правда об алгоритмах/моделях

1. Нет универсального алгоритма для всех задач ИБ. Разные алгоритмы помогают решать разные задачи для разных данных.
2. Вспоминаем – обучение и исполнение. В процессе исполнения модель не учится! Машинное обучение итеративно и требует регулярного переобучения! «Машинное обучение не требует обновления» – лукавство!
3. Выбор алгоритма – это баланс между скоростью работы, аккуратностью предсказания и сложностью модели.
4. Исходную задачу можно решить разными способами (алгоритмами). От их выбора зависит точность и скорость решения
5. Если датасет неполон или некачественный, то никакой алгоритм не поможет!

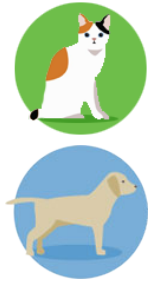
Основные алгоритмы машинного обучения

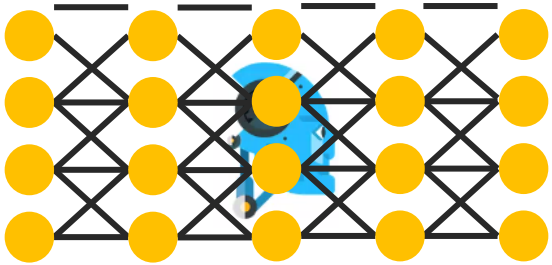


Собака

Машинное обучение с учителем

Входные данные — Выделение признаков — Классификация — Вывод

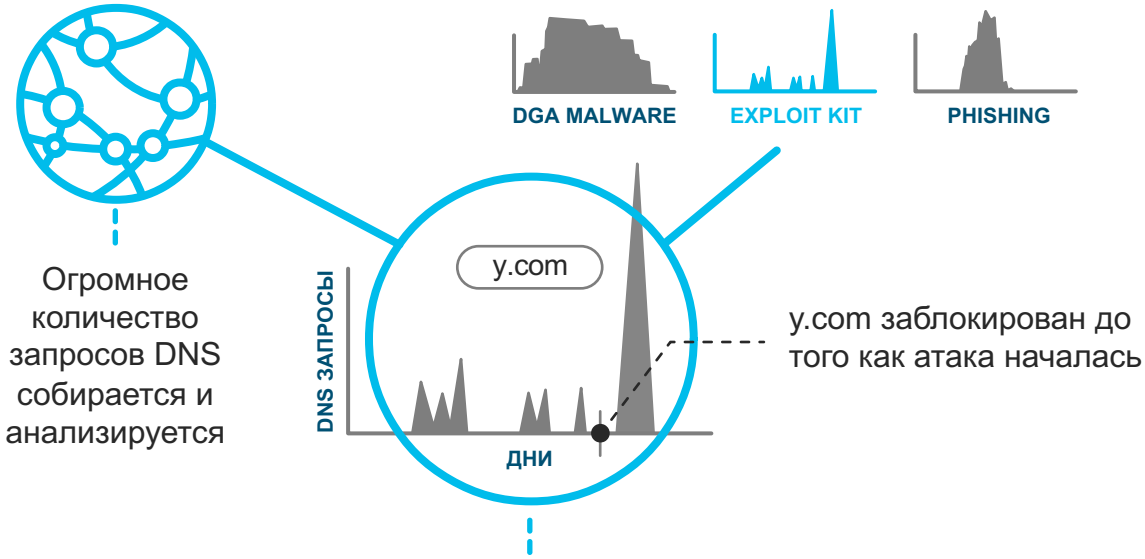




Собака

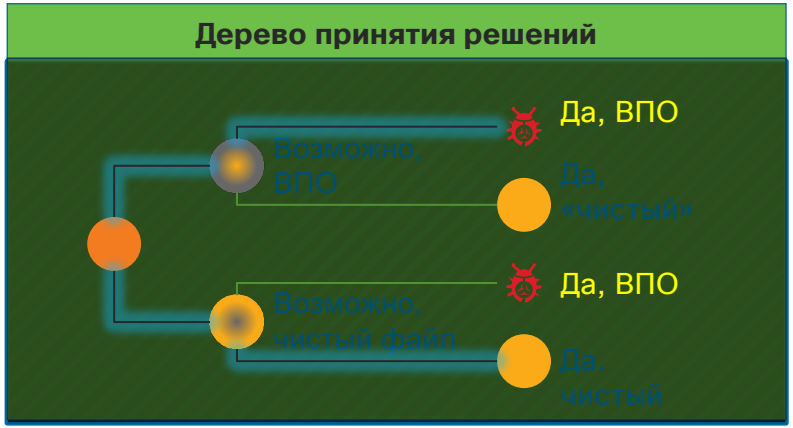
Машинное обучение без учителя

Модель всплесков активности для определения вредоносных доменов



А в анализе файлов используются деревья решений

- Опираясь на значения по каждому из проанализированных атрибутов и следуя имеющемуся дереву принятия решения делается вердикт о вредоносности того или иного файла
  - Предварительно необходимо определить пороговые значения для каждого параметра, означающие переход из одного статуса («чистый») в другой («зараженный»)



Классификация vs кластеризация

- Классификация – машинное обучение с учителем, то есть обучение с помощью примеров. При известных входе и выходе, неизвестна зависимость между ними. Машинное обучение позволяет обнаружить эту зависимость
  Примеры: DGA, фишинг, спам, DNS-угрозы, загрузка вредоноса, фрод

- Кластеризация – машинное обучение без учителя, то есть обучение, при котором испытуемая система спонтанно обучается выполнять поставленную задачу без вмешательства со стороны экспериментатора
  Примеры: утечки данных, lateral movements, подбор учеток и т.п.

Обнаружение скомпрометированных учеток

Compromised Account Risk

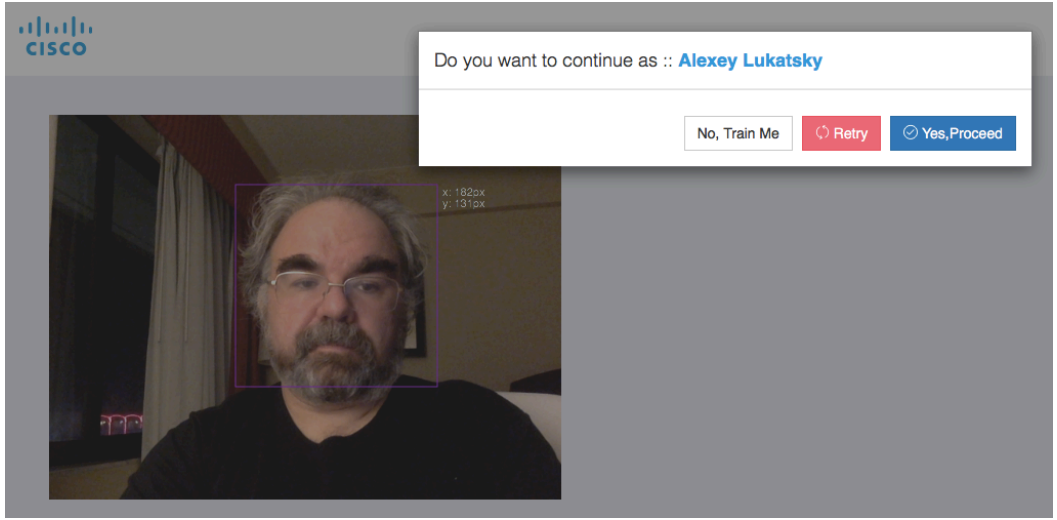
Showing top **14 users** out of total **14 users** that have generated activity from 3 or more locations in the past **7 days**.
Activity in one account across multiple locations may indicate use of a VPN, possibly unauthorized.
Activity from multiple and/or risky locations may indicate compromised accounts.

User	Logins	Location
There were **14** instances of users who logged in from 3+ locations.		
[redacted]	4,711	Pleasant Grove, UT, United ...
	7	Lehi, UT, United States
	6	Buenos Aires, C, Argentina
	4	Federal, E, Argentina
	1	San Francisco, CA, United S...
	1	Dallas, TX, United States
	2	Los Angeles, CA, United Sta...
	2	Parker, CO, United States
	1	Columbus, OH, United States
	1	Manhattan Beach, CA, Unite...
	1	Denver, CO, United States
	3	Hyderabad, TG, India

На примере Cisco CloudLock

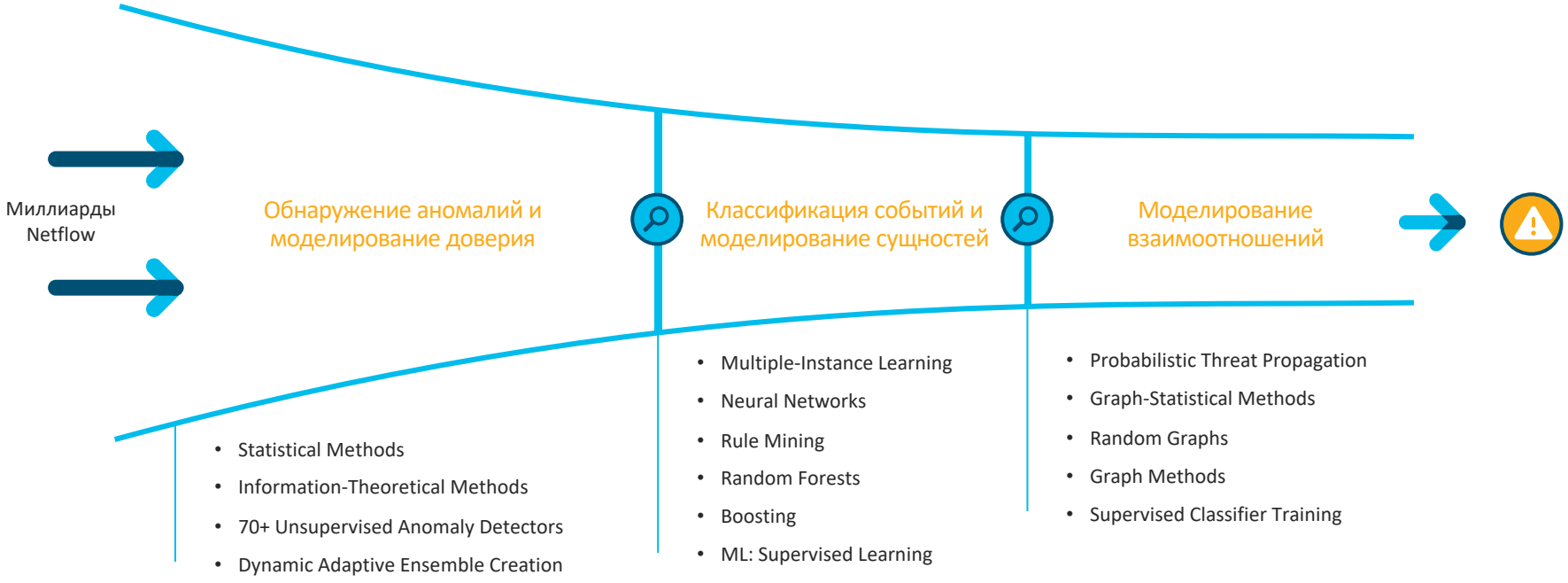
А вот есть еще нейросети...

- Классификация обычно используется для простых датасетов – цифр, текстов, табличных данных
- Для более сложных данных (картинки, видео, голос) лучше использовать нейросети
- Анализ биометрии сегодня делается обычно с помощью нейросетей



На примере бета-проекта в Cisco

Пример: множество алгоритмов для мониторинга Netflow



На примере Cisco Stealthwatch

Где взять людей?

- У вас есть нужные данные, но нет правильных моделей. У вас есть правильные модели, но нет нужных данных

  Про анализ рисков на CISO Summit, начало 2000-х годов

- Сегодня у вас есть нужные данные (и их слишком много) и правильные (наверное) модели… но нет аналитиков, которые могут свести все это вместе

Антон Чувакин, экcVP Gartner (сейчас Chronicle)

Путь к искусственному интеллекту в ИБ

1. *Ползать* Создание реальных ML приложений — Быть **стабильным**

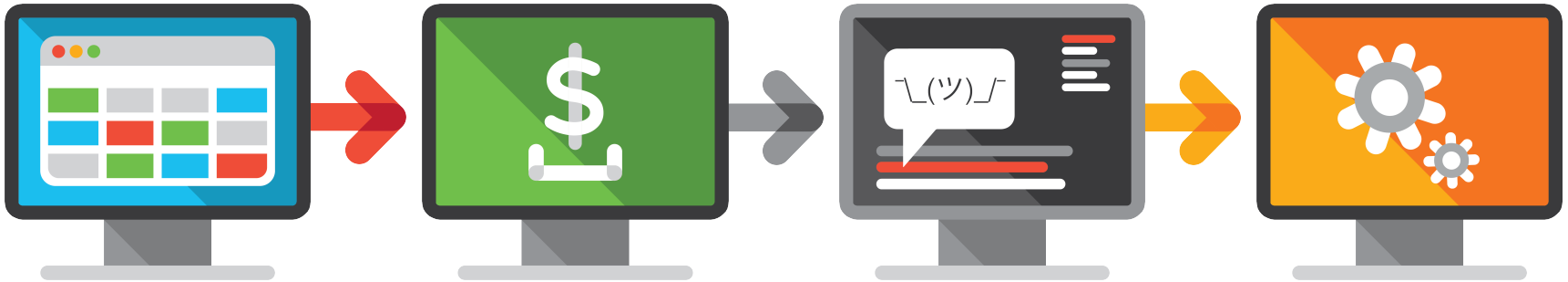
Большинство вендоров тут

2. *Ходить* Построить множество приложений — Быть **повторяемым**

3. *Бегать* Построить множество приложений для многих заказчиков — Быть **автоматизируемым**

4. *Летать* Позволить клиентам делать это самим — Быть **разработчиком**

Поэтому: комбинация защитных технологий



Сигнатуры и IoC	Правила	Статистические модели	Алгоритмы ИИ
IDS, AV, NGIPS, EDR, TIP	NGFW, WSA, SIEM, ESA	Netflow	

Прежде чем идти в машинное обучение по ИБ

- Какие данные вам нужны?
- Что вы хотите в них увидеть?
- Как и где это «что» будет применено?
- Кто это все будет делать?

Вопросы вендору

1. Какие алгоритмы используются для обнаружения?
2. Какие наборы данных используются в алгоритме
3. Где запускаются алгоритмы (на узле, в ЦОДе, в облаке)?
4. Могут быть алгоритмы обучены на ваших данных?
5. Как много обучающих данных требуется?

Стремные вопросы для большинства вендоров

Если вы решились сами делать ИИ в ИБ

1. У вас есть необходимые датасеты?
2. У вас есть квалифицированные аналитики?
3. Какие алгоритмы вы будете использовать?
4. Вы можете использовать open source модели?
5. Как вы будете проверять качество своих моделей?

Вопросы?

Раньше люди покупали SIEM, чтобы не писать сигнатуры для IDS.

Сегодня люди покупают машинное обучение, чтобы не писать правила для SIEM.

Что дальше?

alukatsk@cisco.com