Аналіз застосовності машинного навчання в системах аналізу мережевих атак

Полотай Орест Іванович, Львівський державний університет безпеки життєдіяльності, м.Львів, Україна, orest.polotaj@gmail.com
Павлишин Аліна Юріївна, Львівський державний університет безпеки життєдіяльності, м.Львів, Україна, skarletrain27@gmail.com

Розглянуто поняття мережевих атак, особливостей машинного навчання для їх виявлення. Розглянуто способи машинного навчання.
Ключові слова: мережеві атаки, машинне навчання

Вступ

В даний час засоби та системи комунікації розвиваються стрімкими темпами, різко збільшилися обсяг та швидкість передачі даних. Інформація безпосередньо впливає на життя людей, функціонування та регулювання організацій та держав у цілому, саме тому інформацію прийнято вважати одним із ключових ресурсів.

У зв'язку з цим гостро постає проблема забезпечення безпеки інформації, яка безпосередньо залежить від зростання обсягу та значущості інформації.

Таким чином, існує потреба у дослідженні системи аналізу атак на основі машинного навчання.

Визначення та класифікація мережевих атак

Мережевою атакою називають навмисні дії третіх осіб (зловмисників), спрямовані на отримання контролю над локальним або віддаленим комп'ютером для подальшого порушення роботи мережі, зміни прав

користувачів, отримання персональних даних або реалізації будь-яких деструктивних дій над інформацію.

Прикладами мережевих атак можуть бути [2]:

1. Mailbombing.
2. Застосування спеціалізованих додатків.
3. Переповнення буфера.
4. Мережна розвідка.
5. IP-спуфінг.
6. Man-in-the-Middle.
7. Фішинг.
8. DDOS-атака.
9. XSS-атака.
10. Brute Force.
11. Sql Injection.

Методи виявлення мережевих атак

Виявленням мережевих атак називається процес розпізнавання аномальної чи підозрілої діяльності. . Даними для аналізу є мережевий трафік у вигляді мережевих пакетів. Ці дані збираються без обробки, після чого можуть бути нормалізовані для завдання ознакових атрибутів загального виду. Такі дані використовуються для створення активного профілю. Далі профіль порівнюється з нормальною діяльністю об'єкта, і при виявленні розбіжностей параметрів профілю фіксується аномалія. Даний алгоритм має кілька варіантів подальшої реалізації: процедура порівняння з граничною величиною (при перевищенні граничної величини фіксується аномалія), ідентифікація несанкціонованих дій шляхом порівняння мережевого трафіку з шаблоном атак, методи інтелектуального аналізу даних (методи обчислювального інтелекту, машинного навчання).

Метод інтелектуального аналізу успішно застосовується при розробці сучасних систем виявлення атак і є перспективним напрямом розвитку даної галузі.

Способи машинного навчання

Одним із класів методу інтелектуального аналізу є машинне навчання.

Машинне навчання (англ. machine learning, ML) - галузь штучного інтелекту, характерною рисою яких є не пряме розв'язання задачі, а навчання, що ґрунтується на використанні даних та алгоритмів для імітації навчання людини. Вирішуючи схожі завдання, поступово підвищується і точність рішень. Для побудови таких методів використовують засоби математичної статистики, чисельних методів, математичного аналізу, методів оптимізації, теорії ймовірностей, теорії графів, різні техніки роботи з даними в цифровій формі.

Способи навчання поділяють на:

1. навчання з учителем (supervised learning).
2. навчання без учителя (unsupervised learning).
3. навчання з підкріпленням (reinforcement learning).

Так, навчання з учителем має на увазі, що дані, підготовлені для аналізу, вже містять правильну відповідь, тому метою алгоритму є не отримати відповідь, а з'ясувати закономірність шляхом виявлення взаємозв'язків. У результаті алгоритм отримує здатність вибудовувати коректні прогнози та моделі з мінімальною похибкою.

Для навчання без вчителя ключовим є якийсь шаблон, отриманий алгоритмом у процесі виявлення закономірностей в оброблюваному масиві даних. На основі виявлених закономірностей і систематизуються дані.

Навчання з підкріпленням є окремим випадком навчання з учителем. За такого навчання алгоритм навчається, взаємодіючи з певним середовищем, відгуком якого на дії алгоритму є сигнали підкріплення. У ролі вчителя тут виступає саме середовище [1].

Моделі, що застосовуються під час машинного навчання

Виконання машинного навчання включає в себе створення якоїсь моделі, яка навчається на вхідних навчальних даних, а потім може обробляти інші дані, для прогнозування [3]. Для систем машинного навчання використовують різні типи моделей, наприклад:

1. Дерево прийняття рішень.
2. Випадковий ліс.
3. Лінійна регресія.
4. Наївний Байєс.

Висновки

Отже, машинне навчання зачіпає практично всі сфери діяльності людства: від харчової промисловості та сільського господарства до економіки та науки. Так, основними практичними сферами застосування є: розпізнавання мови, технічна діагностика, медична діагностика, прогнозування часових рядів, виявлення шахрайства, виявлення спаму, кредитний скоринг, та багато інших.

Перелік посилань

1. Kelleher J.D.. Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies / Kelleher J.D., Namee B.M, D'Arcy A. – The MIT Press, 2015. – 624 p. 2.
2. Види мережевих атак. Способи їх виявлення. [Електронний ресурс]. – Режим доступу: https://holodoks.blogspot.com/2017/12/blog-post.html
3. Штучний інтелект, машинне навчання та нейронні мережі: в чому різниця і для чого їх використовують – [Електронний ресурс]. – Режим доступу: https://evergreens.com.ua/ua/articles/machine-learning-overview.html