Научно-практический электронный журнал
Оригинальные исследования (ОРИС), *ISSN 2222-0402*
www.ores.su
original-research.ru

ОПРЕДЕЛЕНИЕ ФИШИНГОВЫХ ССЫЛОК С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Алексеев Артем Константинович,
Бакалавр
Санкт-Петербургский государственный университет аэрокосмического приборостроения, 25 кафедра
alart.2002@mail.ru

Будылин Егор Сергеевич,
Бакалавр
Санкт-Петербургский государственный университет аэрокосмического приборостроения, 25 кафедра
budylinegor123@gmail.com

Аннотация

Фишинг является одной из наиболее серьёзных угроз в киберпространстве, направленной на кражу конфиденциальной информации пользователей. Современные методы машинного обучения (ML) предлагают мощные инструменты для автоматического обнаружения и классификации фишинговых ссылок. В данной статье рассматриваются основные методы и алгоритмы машинного обучения, применяемые для выявления фишинговых URL, а также обсуждаются их эффективность и особенности применения. Так же представлены экспериментальные результаты на открытом датасете с помощью ML.

Ключевые слова: фишинг, машинное обучение, определение фишинговых ссылок, кибербезопасность, классификация URL.

DETECTION OF PHISHING URLS USING MACHINE LEARNING METHODS

Artem K. Alekseev,
Bachelor's Degree
St. Petersburg State University of Aerospace Instrumentation, Department 25
alart.2002@mail.ru

Egor S. Budylin,
Bachelor's DegreeSt.
Petersburg State University of Aerospace Instrumentation, Department 25
budylinegor123@gmail.com

ABSTRACT

Phishing is one of the most serious threats in cyberspace, aimed at stealing users' confidential information. Modern machine learning (ML) methods offer powerful tools for the automatic detection and classification of phishing URLs. This article examines the main methods and algorithms of machine learning used to identify phishing URLs, discussing their effectiveness and application features. Experimental results on an open dataset using ML are also presented.

Keywords: phishing, machine learning, phishing URL detection, cybersecurity, URL classification.

Введение

Фишинг представляет собой вид мошенничества, при котором злоумышленники используют различные методы обмана для получения конфиденциальной информации у пользователей. Этот тип атаки основан на социальной инженерии и прицельном манипулировании психологией жертвы, в результате чего она предоставляет личные данные без подозрений [1].

Одной из особенностей фишинга является его широкое распространение и высокая степень эффективности. Злоумышленники постоянно совершенствуют свои методы, делая свои атаки более убедительными и трудноразличимыми от надежных запросов. Подделанные веб-сайты и электронные письма могут быть созданы с высокой степенью профессионализма, что усложняет задачу даже опытным пользователям интернета в отличии мошеннических запросов от существующих.

Машинное обучение предлагает эффективные методы для автоматизации процесса выявления фишинговых ссылок. В данной статье мы исследуем различные подходы и алгоритмы машинного обучения, применяемые для решения этой задачи, анализируем их достоинства и недостатки, а также предлагаем рекомендации по их применению [2].

Методы и алгоритмы

Сбор данных

Основой для построения эффективной модели машинного обучения является качественный и разнообразный набор данных. Для задачи определения фишинговых ссылок необходимы датасеты, содержащие URL-адреса, метки классов (фишинг/легитимный) и дополнительные характеристики URL, такие как длина URL, наличие подозрительных слов, IP-адреса, информация о домене и т.д.

Предобработка данных

Перед применением алгоритмов машинного обучения данные требуют тщательной предобработки. Этот этап включает очистку данных, извлечение признаков и их нормализацию. Важным шагом является преобразование URL-адресов в числовые представления, что может быть достигнуто с помощью различных методов, включая One-Hot-Encoding, TF-IDF и другие техники текстовой обработки.

Алгоритмы машинного обучения

1. Логистическая регрессия: простой и интерпретируемый алгоритм, часто используемый для бинарной классификации. Логистическая регрессия может служить базовой моделью для задачи определения фишинговых ссылок [3].
2. Деревья решений: алгоритм, основанный на построении дерева решений, где каждый узел представляет собой проверку определённого признака, а ветви – результаты этой проверки. Деревья решений легко интерпретируются, но могут страдать от переобучения.

3. Случайные леса: ансамблевый метод, состоящий из множества деревьев решений, обученных на различных подвыборках данных. Случайные леса обеспечивают высокую точность и устойчивость к переобучению.
4. Градиентный бустинг: ещё один ансамблевый метод, который обучает модели последовательно, улучшая ошибки предыдущих моделей. Градиентный бустинг часто показывает высокие результаты, но требует тщательной настройки гиперпараметров.
5. Нейронные сети: современные глубокие нейронные сети могут эффективно извлекать сложные зависимости в данных и использоваться для задачи определения фишинговых ссылок. Однако они требуют больших объёмов данных и вычислительных ресурсов для обучения.

Экспериментальные результаты

Для оценки эффективности различных алгоритмов машинного обучения был проведён эксперимент на открытом датасете фишинговых и легитимных URL. Датасет был разделён на обучающую и тестовую выборки в соотношении 80/20. Оценка производительности моделей проводилась с использованием метрик точности (accuracy), полноты (recall), точности (precision) и F1-меры [4].

С помощью машинного обучения была разработана модель, использовалась рекуррентная нейронная сеть (RNN). Выбор рекуррентной нейронной сети был обусловлен тем, что:

1. URL-адреса представляют собой последовательности символов, и порядок этих символов имеет значение для понимания структуры и возможных паттернов, характерных для фишинговых ссылок. Рекуррентные нейронные сети, в отличие от других архитектур, специально разработаны для обработки последовательных данных, учитывая контекст предыдущих элементов последовательности.
2. Рекуррентные нейронные сети имеют память, которая позволяет им учитывать контекст предыдущих символов в последовательности. Это особенно важно для анализа URL-адресов, где определенные последовательности символов могут указывать на фишинговые атаки. Например, последовательность "https://" в начале URL указывает на защищенное соединение, что может быть важно для классификации.
3. Рекуррентные нейронные сети могут обрабатывать последовательности переменной длины. URL-адреса могут значительно варьироваться по длине, и способность RNN справляться с этим делает их подходящими для задачи классификации фишинговых ссылок.

Обучение модели проводилось на подготовленных данных в течение нескольких эпох, лучший результат был показан на 10 эпохах с использованием размера батча равному 64. Во время обучения производительность проверялась на обучающей выборке. После удовлетворительного результата на обучающей выборке, нейронная сеть тестировалась на тестовой выборке, сравниваясь с обучающей, с помощью метрик, таких как: точность, потери, f1, ROC-кривая. Рисунок 1 показывает кривую точности на обучающей и валидационной выборках. А рисунок 2 кривую потерь, для обучающей и валидационной выборке.

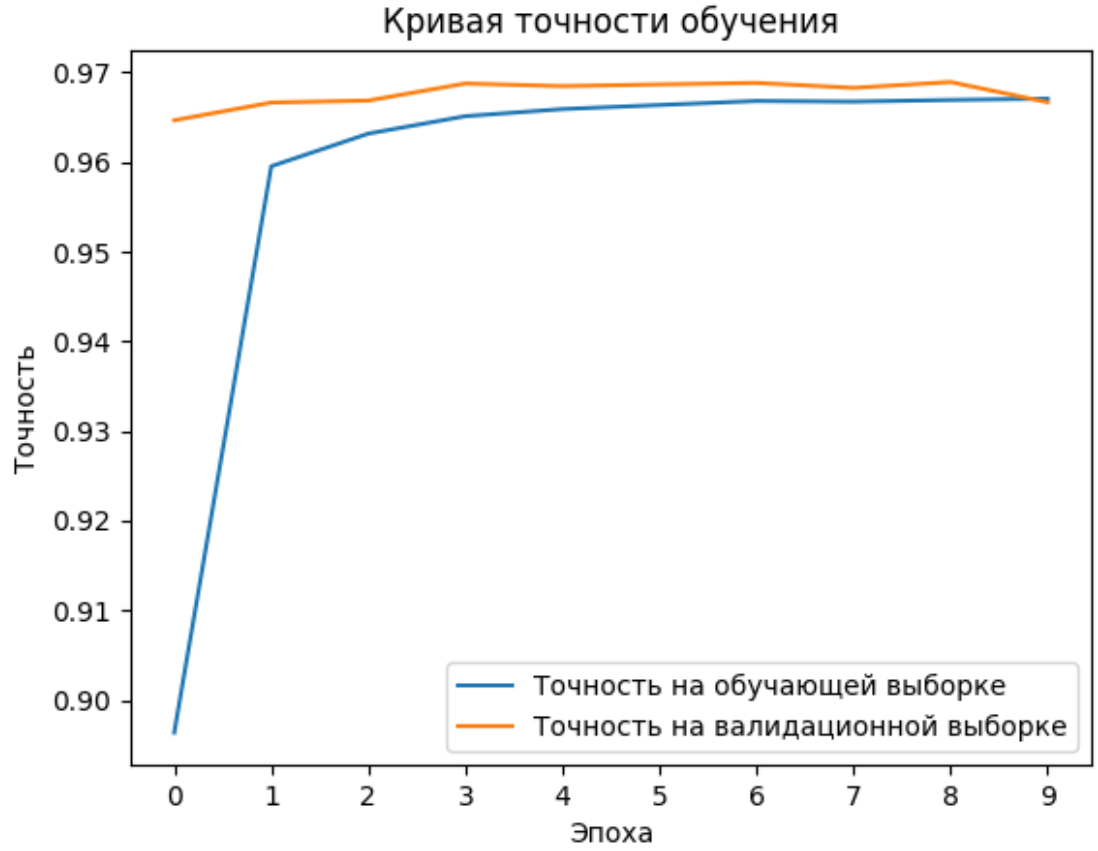

Рисунок 1 – Кривая точности обучения

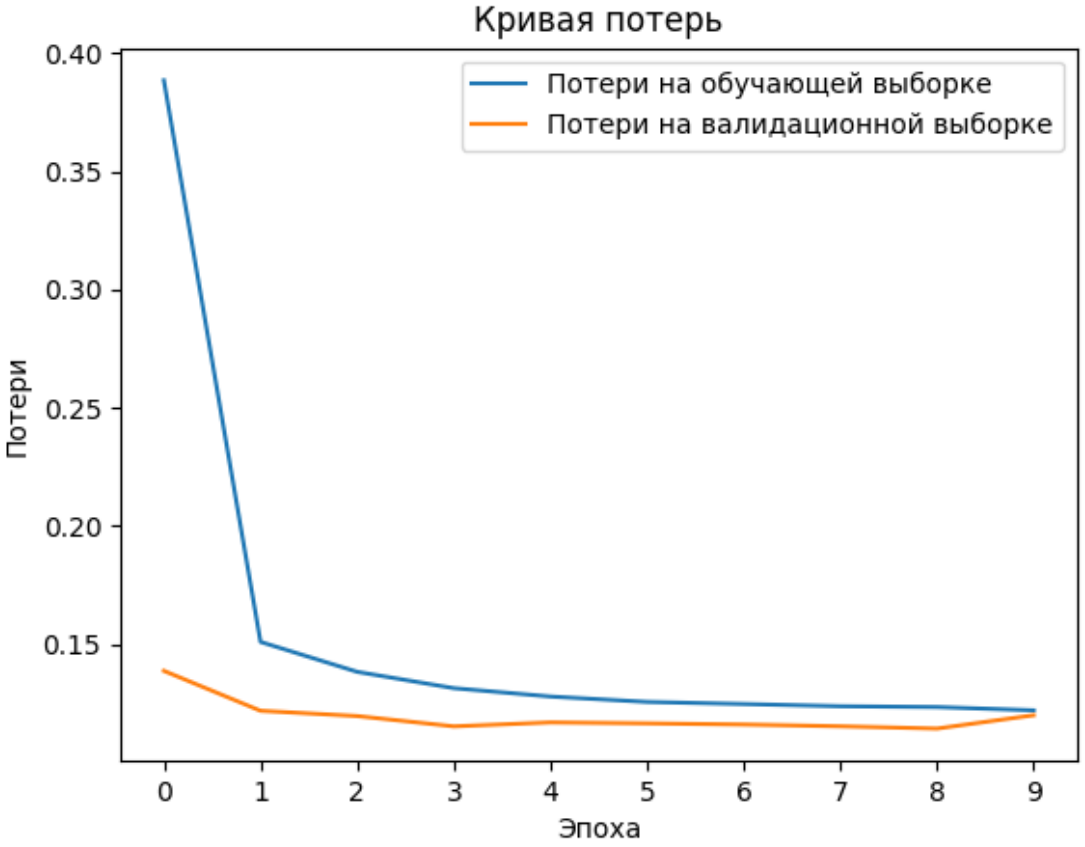

Рисунок 2- График метрики потерь

В результате модель показала хорошую точность 98%, это означает, что модель, созданная с помощью рекуррентной нейронной сети, хорошо справляется с определением фишинговых ссылок.

Заключение

Методы машинного обучения предоставляют мощные инструменты для автоматического определения фишинговых ссылок. Различные алгоритмы, такие как логистическая регрессия, деревья решений, случайные леса, градиентный бустинг и нейронные сети, могут эффективно использоваться для этой задачи, обеспечивая высокую точность и надёжность.

Использование рекуррентной нейронной сети для классификации фишинговых и легитимных URL-адресов оказалось эффективным. Высокая точность модели и положительные результаты по другим метрикам подтверждают, что RNN является подходящим инструментом для данной задачи. Модель способна учитывать важные аспекты структуры URL и эффективно распознавать фишинговые ссылки. Показанный метод можно улучшить и запустить в производство.

Список литература:

1. Юсупов, М. Ю. Фишинг как угроза конфиденциальности в сети / М.Ю. Юсупов, А.О. Путилов // E-Scio. – 2021. – №. 10 (61). – С. 223-232.
2. Менциев, А. У. Анализ фишинговых атак как вида социальной инженерии / А. У. Менциев А. У. //Наука и молодежь. – 2016. – С. 391-394.
3. Баева У. М. Обзор средств для симуляции фишинговых атак / У. М. Баева, В.И. Кураков, А.С. Худадян // Вестник науки. – 2022. – Т. 2. – №. 5 (50). – С. 71-77.
4. Рашка, С. Python и машинное обучение /С. Рашка, пер. с англ. А. В. Логунова // М.: ДМК Пресс. – 2017. – С. 418.

References:

1. Yusupov, M. Yu. Phishing as a threat to online privacy / M. Yu. Yusupov, A.O. Putilov // E-Scio. – 2021. – No. 10 (61). – pp. 223-232.
2. Mentsiev, A. U. Analysis of phishing attacks as a type of social engineering / A. U. Mentsiev A. U. // Science and youth. – 2016. – P. 391-394.
3. Baeva U. M. Review of tools for simulating phishing attacks / U. M. Baeva, V. I. Kurakov, A.S. Khudadyan // Bulletin of Science. – 2022. – T. 2. – No. 5 (50). – pp. 71-77.
4. Raska, S. Python and machine learning /S. Raska, per. from English A. V. Logunova // M.: DMK Press. – 2017. – P. 418