УДК 004.891.3 URL: https://ptsj.ru/catalog/icec/insec/969.html

РАЗВИТИЕ И ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБЛАСТИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

А.А. Володин scorpy2013@gmail.com
Е.В. Глинская glinskaya-iu8@rambler.ru
SPIN-код: 5430-3023

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Выполнен общий обзор и проанализированы возможности искусственного интеллекта (ИИ), а также различных программ и приложений, использующих нейросети, для обеспечения конфиденциальности, целостности и доступности защищенной информации, хранящейся на устройствах. Представлены будущая динамика и дальнейшие перспективы развития сегмента ИИ. Анализ современных программных решений показал, что ИИ играет важную роль в борьбе с современными информационными угрозами. Несмотря на это злоумышленники и хакеры могут использовать нейросети для кибератак, чтобы получить доступ к критической информации. В целом внедрение новейших технологий ИИ в информационную безопасность крупных компаний и организаций поможет значительно снизить расходы, повысить эффективность использования ресурсов, оптимизировать процессы обеспечения информационной безопасности. Более того, применение нейросетей в целях защиты критической информации позволит значительно повысить скорость обнаружения и анализа вредоносных программ на конечных точках и в различных веб-приложениях.

Ключевые слова: искусственный интеллект, машинное обучение, информационная безопасность, защита секретной информации, аудит, мониторинг, управление рисками, выявление уязвимостей, предотвращение угроз

Введение. С развитием технологий искусственного интеллекта (ИИ) в последние годы его применение в разных сферах жизни значительно возросло. Одной из таких сфер является информационная безопасность (ИБ), где ИИ обещает революционизировать способы обнаружения, анализа и предотвращения кибератак. Однако наряду с перспективными возможностями и преимуществами существуют и проблемы и вызовы, возникающие при использовании ИИ в этой области [1].

Эта статья посвящена исследованию развития и проблемам использования ИИ в области ИБ. В ней мы рассмотрим различные аспекты и перспективы применения ИИ в защите данных, обнаружении и предотвращении кибератак. Также мы представим проблемы, связанные с использованием ИИ,

 ISSN 2541-8009

такие как этические вопросы, недостатки технических решений и использование ИИ для получения несанкционированного доступа к защищенным ресурсам.

Ознакомление с данными темами поможет читателям получить более глубокое понимание текущего состояния и перспективы развития использования ИИ в сфере ИБ и научиться принимать взвешенные решения в своих разработках и проектах [2].

Использование искусственного интеллекта и нейросетей в сфере информационной безопасности. Среди плюсов применения ИИ в области ИБ можно выделить несколько важных аспектов. Во-первых, ИИ позволяет создавать решения, которые способны обнаруживать и предотвращать угрозы с высокой скоростью. Благодаря этому идентификация потенциальных угроз происходит оперативно, что позволяет сократить время реагирования и предотвратить потенциальный ущерб на ранних стадиях атаки. Кроме того, использование ИИ позволяет выбирать самое оптимальное решение для реагирования на возникающие инциденты ИБ на основе анализа данных и предугадывания их последствий. Это сокращает затраты на разработку и обновление программного обеспечения (ПО), поскольку некоторые процессы автоматизируются [3].

Одним из важных аспектов использования ИИ в области ИБ является применение технологий поведенческого анализа и предикативной аналитики. Использование ИИ позволяет анализировать поведение пользователей с целью выявления аномалий и блокирования подозрительных действий. Кроме того, ИИ может быть использован для управления персоналом организаций, например, в области доступа и аутентификации.

Использование ИИ в области ИБ не лишено недостатков. Одним из примеров является тот факт, что злоумышленники могут использовать средства автоматизации ИИ и нейросетей для совершенствования и трансформации кибератак и обхода существующих механизмов защиты. Спам-фишинг, использующий ИИ для усиления атак, служит одним из таких примеров. Злоумышленники могут генерировать тысячи сообщений и создавать цепочки разговоров, используя анализ текста на естественном языке, что делает атаку более эффективной. Кроме того, с использованием ИИ и машинного обучения злоумышленники могут осуществлять более эффективный подбор паролей, обход двухфакторной аутентификации, а также успешно заполнять формы CAPTCHA. Также отметим, что злоумышленники имеют возможность подключаться к ресурсам в даркнете и собирать информацию для формирования обширной базы данных, способствующей действенным кибератакам [4].

Развитие и использование ИИ в области ИБ могут дать нам новые возможности для защиты данных и сетей, однако они также могут вызвать новые угрозы. Важно учитывать эти факторы, стремиться к развитию и использованию ИИ в области ИБ, а также получать необходимые навыки и знания для эффективного противодействия потенциальным угрозам и уязвимостям, вызванным использованием ИИ и нейросетей злоумышленниками [5].

Техники обнаружения аномалий на основе ИИ и машинного обучения (МО) представляют мощные инструменты для обеспечения безопасности информационных систем. Одной из ключевых техник является использование нейронных сетей, которые моделируют работу человеческого мозга и могут обнаруживать аномалии на основе обучения с учителем или без него. Этот подход позволяет системам находить атипичное поведение или необычные паттерны в данных.

Другой эффективной техникой являются генетические алгоритмы, использующие эволюционные принципы для определения аномалий и выбора оптимальных решений. Эти алгоритмы могут выполнять интеллектуальный поиск и оптимизацию в пространстве возможных решений, что позволяет идентифицировать потенциальные угрозы и выбирать наилучшие стратегии защиты [6].

Анализ поведения пользователей также является эффективной техникой обнаружения аномалий. Данная техника основана на анализе нормального поведения, этот подход выявляет отклонения от эталонных моделей поведения, что позволяет обнаруживать подозрительные действия или активность.

Статистическое моделирование служит еще одной техникой, используемой для обнаружения аномалий в данных. С помощью статистических методов и моделей системы могут определить аномальные значения или паттерны, которые могут указывать на наличие угрозы или необычные события.

В контексте предотвращения угроз использование ИИ и МО также играет важную роль. Автоматический анализ сетевого трафика позволяет обнаруживать подозрительные пакеты данных или излишнюю активность, что усиливает надежность системы и защищает сетевую инфраструктуру. Блокировка атак осуществляется путем автоматического обнаружения и блокировки вредоносного трафика и активности с применением ИИ и МО [7].

Важным аспектом является адаптивная защита, которая развивается с использованием ИИ и МО. Такие системы могут адаптироваться к новым видам атак и обновлять свои алгоритмы и модели, что делает их более эффективными в предотвращении угроз.

Оценка уязвимостей также может быть автоматизирована с использованием ИИ и МО. Это позволяет идентифицировать и оценить уязвимости ин-

 ISSN 2541-8009

формационных систем и сетей, что дает возможность разработать более надежные меры безопасности.

Все эти техники обнаружения аномалий на основе ИИ и МО представляют собой мощные инструменты, способные значительно улучшить безопасность информационных систем и защитить их от различных угроз. Их применение позволяет обнаруживать и предотвращать атаки, адаптироваться к новым видам угроз и обеспечивать надежную защиту информационных ресурсов [8].

Программные и технические решения. EDR (Endpoint Detection and Response) и NDR (Network Detection and Response) — это платформы и устройства, обеспечивающие обнаружение и реагирование на атаки в рабочих станциях, серверах и сетевом трафике с использованием технологий ИИ. EDR позволяет обнаруживать неизвестные вредоносные программы и автоматически классифицировать угрозы. Технологии ИИ активно используются для принятия решений на основе общей базы знаний, собранной из различных устройств. Некоторые продукты EDR также могут помечать и контролировать перемещение данных на конечных точках для обнаружения внутренних угроз.

NDR, в свою очередь, фокусируется на обнаружении атак на сетевом уровне. Используя технологии ИИ и накопленную статистику и знания об угрозах, продукты NDR могут выявлять угрозы в сетевом трафике и автоматически реагировать на них, изменяя конфигурацию сетевых устройств и шлюзов. Некоторые решения NDR специализируются на защите облачных провайдеров и их инфраструктуры.

Одним из дополнительных сценариев использования ИИ в области сетевой защиты является анализ почтового трафика для обнаружения фишинговых атак.

UEBA (User and Entity Behavior Analytics) — это системы, основанные на поведенческом анализе пользователей и информационных сущностей. Они способны обнаруживать необычное поведение и использовать его для обнаружения внутренних и внешних угроз. Основное предназначение ИИ-технологий в системах UEBA заключается в автоматическом обнаружении аномалий в поведенческих моделях пользователей и сущностей информационных систем. При помощи анализа данных и использования ИИ, выявленные аномалии классифицируются как потенциальные угрозы и риски для бизнеса.

Такой анализ позволяет мониторить и управлять доступом, обнаруживать мошенническую активность среди клиентов и сотрудников (anti-fraud), защищать конфиденциальные данные и обеспечивать соблюдение регламентов и нормативных актов.

TIP (Threat Intelligence Platform) представляет собой платформы, основанные на обширных данных (Data Lake) и индикаторах компрометации (IoC), предназначенные для раннего обнаружения и реагирования на угрозы. Использование ИИ позволяет улучшить эффективность выявления неизвестных угроз в ранних стадиях их развития. Сценарий работы TIP напоминает работу системы SIEM, однако сосредоточен на внешних источниках данных и внешних угрозах.

SIEM (Security Information and Event Management) — это решения, которые обеспечивают мониторинг информационных систем и анализ событий безопасности в режиме реального времени. Они собирают данные от различных источников, включая сетевые устройства, средства защиты информации, ИТ-сервисы и инфраструктуру систем, и помогают обнаружить инциденты информационной безопасности. SIEM-системы накапливают огромное количество данных из различных источников, и использование ИИ позволяет выявлять аномалии с помощью эвристических методов и сокращать число случаев возникновения ложных тревог при изменении паттернов и моделей данных. Применение ИИ в SIEM-системах также позволяет достичь высокого уровня автоматизации.

SOAR (Security Orchestration and Automated Response) — это системы, которые не только выявляют угрозы информационной безопасности, но и автоматизируют реагирование на инциденты. В отличие от SIEM-систем, ИИ в решениях SOAR помогает автоматически реагировать на выявленные угрозы надлежащим образом. Использование ИИ в SOAR-системах значительно повышает эффективность обнаружения и реагирования на угрозы. Автоматизация процессов реагирования на инциденты ИБ позволяет сократить время реакции, уменьшить человеческий фактор и обеспечить более эффективное противодействие угрозам.

Средства защиты приложений (Application Security) являются системами, которые обнаруживают и управляют угрозами безопасности в прикладных приложениях. Одним из важных сценариев использования технологий ИИ в таких системах является автоматический сбор информации об уязвимостях, атаках и заражениях из доступных открытых источников. На основе этой информации осуществляется автоматизация защитных действий, таких как сканирование на наличие уязвимостей, изменение правил защиты для веб-приложений, обнаружение угроз и изменение рисковой модели [9].

Антифрод (Antifraud) — это системы, предназначенные для выявления угроз в бизнес-процессах и предотвращения мошеннических операций в режиме реального времени. В системах защиты от мошенничества применяются технологии ИИ для определения отклонений от установленных бизнес-

 ISSN 2541-8009

процессов, что помогает оперативно реагировать на возможные финансовые преступления или уязвимости процессов. Применение ИИ в таких системах особенно актуально, поскольку позволяет быстро адаптироваться к изменениям логики и различным метрикам бизнес-процессов, а также использовать передовые практики и опыт в индустрии.

Использование искусственного интеллекта для кибератак. Чтобы понимать весь спектр возможностей применения ИИ, необходимо исследовать различные аспекты использования ИИ в контексте кибератак. Многие компании и организации сталкиваются с киберугрозами, возникающими из-за эксплуатации уязвимостей, которые нейросети ищут в автоматическом режиме. Нужно осознавать всю важность принятия соответствующих мер для защиты информационной инфраструктуры от таких атак.

Offensive AI Lab — уникальный проект, который исследует использование ИИ для враждебных целей. В основном лаборатория фокусируется на дипфейках и атаках на медицинские системы, но также предоставляет и другие интересные материалы, связанные с областью ИИ и кибербезопасности [10].

Одним из значимых достижений Offensive AI Lab является описание 33 техник применения ИИ для враждебных целей (рис. 1). Эти техники представляют собой разнообразные методы, используемые злоумышленниками и хакерами для совершения атак с применением ИИ. Важно отметить, что эти техники соотносятся с тактиками матрицы MITRE ATT&CK — широко применяемой и признанной системой классификации исследования киберугроз и тактик, используемых атакующими.

Offensive AI Lab проводит исследования и эксперименты, цель которых — осознание слабых мест и уязвимостей, связанных с применением ИИ в кибервойне. Они анализируют применение ИИ в дипфейках, когда создаются обманчивые видео и изображения с помощью нейросетей для целей мошенничества или дезинформации. Кроме того, Offensive AI Lab изучает атаки на медицинские системы, где использование ИИ позволяет злоумышленникам осуществлять маскировку и брать под контроль критически важные системы.

Этот интересный проект дает возможность лучше понять возможности и угрозы, связанные с применением ИИ в враждебных действиях. Offensive AI Lab является площадкой, где исследователи и специалисты по кибербезопасности могут изучить и осознать потенциальные риски и разработать стратегии защиты в сфере кибербезопасности, связанной с использованием искусственного интеллекта [11].

На рис. 1 представлена матрица угроз MITRE, которая широко используется для классификации и организации различных угроз ИБ. В данном случае

она применяется для отображения атак, которые могут возникнуть при использовании ИИ и ТО.

Существует множество разных направлений атак, которые можно отразить на матрице угроз MITRE в связи с системами машинного обучения. Эти направления атак включают в себя такие аспекты, как атаки на модели машинного обучения, манипуляции и искажение данных, внедрение бэкдоров в алгоритмы, атаки на приватность и безопасность данных, обход обнаружения и другие виды угроз, но не ограничиваются ими.

Направление	Сбор	Сдерживание	Обнаружение	Прерывание	Помощь	Легитимизация	Тестирование
Административный доступ	Мониторинг API	Административный доступ	Мониторинг API	Административный доступ	Административный доступ	Многообразие приложений	Административный доступ
Мониторинг API	Многообразие приложений	Исходное состояние	Многообразие приложений	Многообразие приложений	Многообразие приложений	Нормальное использование	Мониторинг API
Многообразие приложений	Резервное копирование и восстановление	Обманный аккаунт	Поведенческая аналитика	Резервное копирование и восстановление	Поведенческая аналитика	Обманный аккаунт	Многообразие приложений
Обманный аккаунт	Обманный аккаунт	Обманная сеть	Обманный аккаунт	Исходное состояние	Нормальное использование	Обманный контент	Резервное копирование и восстановление
Обманный контент	Обманный контент	Детонация вредоносного ПО	Обманный контент	Поведенческая аналитика	Обманный аккаунт	Обманные учетные данные	Обманный аккаунт
Обманные учетные данные	Обманные учетные данные	Манипуляции с оборудованием	Обманные учетные данные	Обманный контент	Обманный контент	Разнообразие ловушек	Обманный контент
Обманная сеть	Обманная сеть	Изоляция	Обманная сеть	Обманные учетные данные	Обманные учетные данные	Обманная сеть	Обманные учетные данные
Обманная персона	Обманная система	Миграция вектора атаки	Обманная система	Обманная сеть	Разнообразие ловушек	Обманная персона	Разнообразие ловушек
Обманный процесс	Детонация вредоносного ПО	Манипуляции с сетью	Манипуляции с электронной почтой	Манипуляции с электронной почтой	Обманная персона	Обманный процесс	Обманная сеть
Обманная система	Манипуляции с электронной почтой	Контроли безопасности	Хантинг	Манипуляции с оборудованием	Обманная система	Обманная система	Обманная персона
Детонация вредоносного ПО	Разнообразие сетей	Манипуляции с ПО	Изоляция	Изоляция	Разнообразие сетей	Разнообразие сетей	Обманная система
Миграция вектора атаки	Мониторинг сети		Манипуляции с сетью	Манипуляции с сетью	Манипуляции с сетью	Карманный мусор	Детонация вредоносного ПО
Разнообразие сетей	Сбор PCAP		Мониторинг сети	Контроли безопасности	Управление периферией		Миграция вектора атаки
Манипуляции с сетью	Управление периферией		Сбор PCAP	Стандартный порядок действий	Карманный мусор		Разнообразие сетей
Управление периферией	Декодер протокола		Карманный мусор	Обучение пользователей	Контроли безопасности		Манипуляции с сетью
Карманный мусор	Контроли безопасности		Декодер протокола	Манипуляции с ПО	Манипуляции с ПО		Управление периферией
Контроли безопасности	Мониторинг системной активности		Стандартный порядок действий				Карманный мусор
Манипуляции с ПО	Манипуляции с ПО		Мониторинг системной активности				Контроли безопасности
			Обучение пользователей				Манипуляции с ПО
			Манипуляции с ПО				

Рис. 1. Матрица угроз MITRE

Усовершенствованные постоянные угрозы (APT) представляют собой скрытую и продолжительную угрозу, часто исходящую от национальных государств или спонсируемых государством групп. Они получают несанкционированный доступ к компьютерным сетям и остаются незамеченными на протяжении длительного времени. Они также используют ИИ для избежания обнаружения и выявления своих целей.

Дипфейк-атаки также представляют серьезную угрозу. В таких атаках хакеры используют видео или изображения, созданные нейросетями, чтобы притворяться реальными людьми и проводить кампании по мошенничеству или дезинформации.

 ISSN 2541-8009

Вредоносное ПО, основанное на ИИ, является еще одной опасностью. Такое вредоносное ПО может работать самостоятельно и приспосабливаться к изменяющимся условиям, что помогает избежать обнаружения и увеличить свою эффективность.

Фишинг-атаки используют обработку естественного языка (NLP) и машинное обучение (ML) для создания более убедительных фишинговых электронных писем и сообщений, которые мошенники используют для обмана пользователей и получения их личных данных.

DDoS-атаки (рис. 2) представляют угрозу, где ИИ используется для выявления и эксплуатации уязвимостей в сети, что позволяет киберпреступникам увеличить масштаб и воздействие своих атак [12].

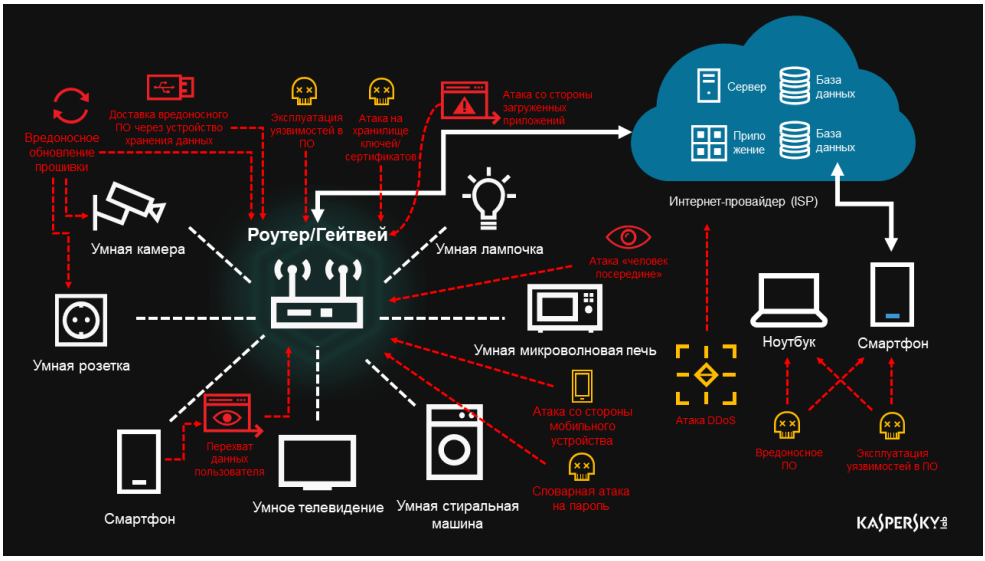

Рис. 2. Схема DDoS-атаки

В сфере наступательного ИИ успешно применяются системы автоматизации, такие как боты в социальных сетях. Эти боты представляют собой программные сущности, способные выполнять заданные действия в социальных сетях, например, публиковать сообщения, комментарии или совершать взаимодействие с пользователями.

Еще одним примером наступательных действий, применимых в области искусственного интеллекта, является использование автоматизированных тестов на проникновение (penetration test) с применением обучения с под-

креплением (рис. 3). Эти тесты позволяют экспертам по ИБ анализировать системы и сети с целью выявления уязвимостей и слабых мест. Техники обучения с подкреплением позволяют оптимизировать процесс тестирования, адаптируя его к особенностям конкретной системы [13].

Рис. 3. Схема автоматической защиты от эксплойтов

Примеры наступательных действий также могут включать подбор паролей, запутывание исходного кода программ, маскировку трафика и управление сетью ботов. Возможности автоматизации позволяют злоумышленникам эффективно совершать атаки, например, вынуждая пользователей переходить по злонамеренным ссылкам, устанавливать вредоносное ПО или выполнять другие нежелательные действия [14].

Важно отметить, что эти примеры наступательных действий представляют собой злоупотребление ИИ и автоматизированными системами в целях совершения противоправных и незаконных активностей. Борьба против таких угроз требует развития соответствующих мер безопасности и обеспечения защиты информационных систем от наступательных действий.

Все эти примеры демонстрируют, как искусственный интеллект используется злоумышленниками для создания и усиления кибератак. Учитывая эти угрозы, важно разрабатывать и совершенствовать методы защиты, которые также используют технологии ИИ и МО. Это поможет более эффективно бороться с новыми угрозами и защищать информационные системы от атак [15].

Прогнозы применения ИИ в 2024 г. Ожидается, что в 2024 г. киберугрозы и киберриски будут продолжать развиваться, становясь все более сложными. Ландшафт угроз будет постоянно расти, а кибератаки станут еще более изощренными, в том числе благодаря использованию ИИ. Впереди нас ждут атаки, основанные на использовании ИИ, поэтому организациям следует вложить свои ресурсы в разработку и применение ИИ для обнаружения новых угроз и реагирования на киберинциденты.

 ISSN 2541-8009

Также следует учитывать, что фишинговые атаки станут все совершеннее, что потребует от организаций обучения и осведомленности всех сотрудников. Прогнозируется, что к концу 2024 г. стоимость кибератак превысит 10,5 триллионов долларов, что подчеркивает необходимость уделить больше внимания кибербезопасности на различных уровнях.

Рост количества устройств Интернета вещей (IoT) и удаленной работы также создаст новые риски из-за наличия слабых протоколов безопасности. Поэтому повышение киберустойчивости и реставрации операций станет все более важным аспектом, наряду с кибербезопасностью. Принцип «нулевого доверия», который ранее применялся только в корпоративных сетях, будет распространяться на удаленных сотрудников, партнеров и устройства IoT, чтобы обеспечить безопасность сетевой активности.

Заключение. В данной статье были рассмотрены особенности ИИ в области ИБ. Технические и программные решения, которые применяют ИИ в ИБ, имеют значительный потенциал для обнаружения уязвимостей и предотвращения угроз в информационных системах и компьютерных сетях. Они позволяют специалистам по ИБ оперативно реагировать на изменяющиеся угрозы и принимать наиболее эффективные меры по защите.

Преимущества использования ИИ в сфере ИБ заключаются в его способности анализировать большие объемы данных, обнаруживать аномалии, автоматизировать процессы и создавать более точные модели угроз. Подходы, основанные на ИИ, помогают выявлять новые и неизвестные уязвимости, предотвращать кибератаки и минимизировать риски инцидентов безопасности.

Однако вместе с перспективными возможностями использование ИИ в области ИБ также представляет вызовы и риски. Применение ИИ во враждебных целях, таких как использование дипфейков, создание вредоносного ПО на базе ИИ или фишинг-атаки с применением ИИ, может причинить серьезный ущерб системам и организациям.

Злоумышленники и хакеры могут использовать ИИ для создания и усиления атак на информационные системы. Мы специально остановились на примерах таких атак в данной статье, чтобы привлечь внимание к возможным угрозам и побудить специалистов по ИБ быть готовыми к ним.

В свете этих вызовов и рисков осознание и понимание принципов работы ИИ в области ИБ являются важным для разработки и применения эффективных стратегий обнаружения и защиты. Компании и организации должны прислушиваться к последним тенденциям и наработкам в области искусственного интеллекта и применять инновационные решения, чтобы успешно справляться с непредвиденными угрозами в сфере ИБ.

ISSN 2541-8009

Литература

[1] Цаунит А.Н. Перспективы развития и применения нейронных сетей. *Молодой ученый*, 2021, № 23 (365), с. 114–117. URL: https://moluch.ru/archive/365/81791/ (дата обращения 09.02.2024).

[2] Журавлев Д.В., Смолин В.С. Нейросетевая революция искусственного интеллекта и варианты ее развития. *Проектирование будущего. Проблемы цифровой реальности. 6-я Междунар. конф.: сб. тр.* Москва, ИПМ им. М.В. Келдыша, 2023, с. 223–244. https://doi.org/10.20948/future-2023-16/

[3] Смирнов А.В. *Машинное обучение и анализ данных в кибербезопасности*. Москва, Бином, 2018.

[4] Семенова А.Н., Ступкина В.А. Цифровые технологии в управлении человеческими ресурсами. *Молодой ученый*, 2019, № 4 (242), с. 250–252. URL: https://moluch.ru/archive/242/55864/ (дата обращения 09.02.2024).

[5] Романов Д.В., Карпов А.С. Применение методов машинного обучения для обнаружения угроз в информационных системах. *Компьютерные инструменты в образовании*, 2020, т. 13, № 4, с. 153–165.

[6] Пилецкая А.В. Искусственный интеллект и большие данные. *Молодой ученый*, 2019, № 50 (288), с. 20–22. URL: https://moluch.ru/archive/288/65241/ (дата обращения 09.02.2024).

[7] Николаева Е.А., Широков М.П. *Применение нейронных сетей в задачах обнаружения аномалий в компьютерных сетях*. Москва, Книжный мир, 2017.

[8] Никитин А.А., Лиманова Н.И. Процесс распознавания изображения нейронной сетью. *Молодой ученый*, 2020, № 47 (337), с. 23–25. URL: https://moluch.ru/archive/337/75420/ (дата обращения 09.02.2024).

[9] Чернов А.А., Горбунов В.В. Анализ и предотвращение угроз в компьютерных сетях с использованием алгоритмов машинного обучения. *Компьютерные исследования и моделирование*, 2021, т. 13, № 1, с. 63–72.

[10] Белова М.С. Искусственный интеллект при анализе больших данных. *Вестник Российского экономического университета им. Г.В. Плеханова. Вступление. Путь в науку*, 2021, т. 11, № 4 (36), с. 136–141. EDN: GJTTEU

[11] Заенцев И.В. *Нейронные сети: основные модели.* Воронеж, ВГУ, 1999, 76 с.

[12] Бритвина П.В. Этические вопросы в применении искусственного интеллекта и машинного обучения. *Вестник науки*, 2024, т. 4, № 1 (70), с. 445–447.

[13] Котельников Е.В., Колеватов В.Ю. *Методы искусственного интеллекта в задачах обеспечения безопасности компьютерных сетей*. URL: https://www.studmed.ru/kolevatov-v-yu-kotelnikov-e-v-metody-iskusstvennogo-intellekta-v-zadachah-obespecheniya-bezopasnosti-kompyuternyh-setey_7f6701b2095.html (дата обращения 09.02.2024).

 ISSN 2541-8009

[14] Писаренко И.Б. Нейросетевые технологии в безопасности. *Информационная безопасность*, 2009, № 4.

[15] Намиот Д.Е., Ильюшин Е.А., Чижов И.В. Искусственный интеллект и кибербезопасность. *International Journal of Open Information Technologies*, 2022, № 9, с. 135–144.

Поступила в редакцию 25.02.2024

Володин Александр Андреевич — студент кафедры «Информационная безопасность», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Глинская Елена Вячеславовна — старший преподаватель кафедры «Информационная безопасность», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Научный руководитель — Басараб Михаил Алексеевич, доктор физико-математических наук, доцент, заведующий кафедры «Информационная безопасность», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Ссылку на эту статью просим оформлять следующим образом:
Володин А.А., Глинская Е.В. Развитие и проблемы использования искусственного интеллекта в области информационной безопасности. *Политехнический молодежный журнал*, 2024, № 02 (91). URL: https://ptsj.ru/catalog/icec/insec/969.html

DEVELOPMENT AND PROBLEMS IN USING ARTIFICIAL INTELLIGENCE IN THE INFORMATION SECURITY

A.A. Volodin scorpy2013@gmail.com
E.V. Glinskaya glinskaya-iu8@rambler.ru

Bauman Moscow State Technical University, Moscow, Russian Federation

The paper presents a general overview and analysis of the artificial intelligence (AI) capabilities, as well as of various programs and applications that are using neural networks to ensure confidentiality, integrity and availability of the protected information stored in devices. It provides future dynamics and further prospects to design and development the AI segment. Analysis of the modern software solutions demonstrates that AI is playing an important role in fighting against the modern information threats. Despite this, attackers and hackers are able to use neural networks in the cyber-attacks to gain access to the critical information. In general, introduction of the latest AI technologies in the information security systems of large companies and organizations would help to significantly reduce costs, increase efficiency of the resources use, and optimize the information security processes. Moreover, using the neural networks to protect critical information would significantly increase the speed of malware detection and analysis at the endpoints and in various web applications.

Keywords: artificial intelligence, machine learning, information security, classified information protection, audit, monitoring, risk management, vulnerability identification, threat prevention

Received 25.02.2024

Volodin A.A. — Student, Department of Information Security, Bauman Moscow State Technical University, Moscow, Russian Federation.

Glinskaya E.V. — Senior Lecturer, Department of Information Security, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Basarab M.A., Dr. (Phys. and Math.) Sci., Associate Professor, Head of the Department of Information Security, Bauman Moscow State Technical University, Moscow, Russian Federation.

Please cite this article in English as:
Volodin A.A., Glinskaya E.V. Development and problems in using artificial intelligence in the information security. *Politekhnicheskiy molodezhnyy zhurnal*, 2024, no. 02 (91). (In Russ.). URL: https://ptsj.ru/catalog/icec/insec/969.html

 ISSN 2541-8009