



УДК 004.855.5

СОВРЕМЕННЫЕ ВОЗМОЖНОСТИ И ПРИМЕРЫ ВНЕДРЕНИЯ МАШИННОГО ОБУЧЕНИЯ

Клюева Ирина Алексеевна

программист кафедры вычислительной и прикладной математики
Рязанский государственный радиотехнический университет
390005, Рязанская обл., г. Рязань, ул. Гагарина, 59
i.klyueva-job@yandex.ru

Аннотация

В статье рассматриваются практические возможности машинного обучения. Представлен обзор отраслей, внедряющих технологии машинного обучения и интеллектуального анализа данных. Приводятся конкретные случаи использования машинного обучения от распознавания изображений и речи до обнаружения мошенничества и медицинской диагностики.

Ключевые слова: интеллектуальный анализ данных, машинное обучение, распознавание речи, обнаружение аномалий, оптимизация запасов, медицинская диагностика.

MODERN OPPORTUNITIES AND EXAMPLES OF MACHINE LEARNING IMPLEMENTATION

Irina A. Klyueva

programmer of the Computational and Applied Mathematics Department
Ryazan State Radio Engineering University
390005, Ryazan region, Ryazan, Gagarin str., 59
i.klyueva-job@yandex.ru

ABSTRACT

The article discusses the practical possibilities of machine learning. An overview of the industries implementing machine learning and data mining technologies is presented. Specific cases of using machine learning are given, from image and speech recognition to fraud detection and medical diagnostics.

Key words: data mining, machine learning, speech recognition, anomaly detection, inventory optimization, medical diagnostics.

Введение. В настоящее время примеров применения машинного обучения становится так много в различных сферах деятельности, что все большее внимание заслуживает их обзор и анализ.

Тенденция развития технологий интеллектуального анализа данных (*Data Mining, DM*) и машинного обучения (*Machine Learning, ML*) в областях применения науки о данных (*Data Science*) развивается высокими темпами [1-30].

Интеллектуальный анализ данных использует математический анализ для определения закономерностей и тенденций, существующих в данных. Как правило, эти закономерности не могут быть обнаружены традиционным исследованием данных, потому что взаимосвязи слишком сложны или потому что данных слишком много.

Помимо интеллектуального анализа больших баз данных, методы интеллектуального анализа данных включают:

- веб-майнинг,
- сетевые подходы,
- анализ текста,
- обработка естественного языка (*Natural language Processing, NLP*),
- прогнозное моделирование,
- анализ взаимосвязей,
- статистический анализ и т.д.

Интеллектуальный анализ данных – это исследование и анализ большого количества данных с помощью автоматических или полуавтоматических средств для того, чтобы обнаружить значимые закономерности и правила. Хотя существует множество других общепринятых определений интеллектуального анализа данных, данное определение отражает, что подразумевается поиск значимых закономерностей в больших объемах данных.

Исторически нахождение полезных закономерностей в данных включало извлечение, обнаружение и сбор информации, выявление шаблонов данных в дополнение к интеллектуальному анализу данных. В последние годы данные подходы не отделяются от интеллектуального анализа данных.

Статистика обычно использует термин «интеллектуальный анализ данных» для обозначения закономерностей в данных, которые обнаруживаются с помощью многомерного регрессионного анализа и других статистических методов.

Для работы с данными и извлечения полезных знаний, скрытых в них, интеллектуальный анализ данных тесно связан с процессом обнаружения знаний в базах данных (*Large Databases, KDD*).

Термин KDD был придуман в 1989 году для обозначения широкого процесса поиска знаний в хранилищах данных. Область KDD особенно сосредоточена на деятельности, ведущей к фактическому анализу данных, включая оценку и внедрение результатов.

Весь этот процесс включает в себя следующие основные этапы, которые могут выполняться в итеративной и интерактивной последовательности.

Сбор данных. Основная цель этого шага – создать целевой набор данных из исходных данных, на основе которых должно быть выполнено обнаружение знаний.

Предварительная обработка данных. Этап, на котором данные «очищаются» для определения стратегий обработки отсутствующих полей данных и учета информации о временной последовательности.

Преобразование данных. Данный этап предполагает уменьшение и проецирование данных при использовании различных методов преобразования.

Интеллектуальный анализ данных. Этот шаг связан с извлечением шаблонов в данных путем выбора методов, задач и алгоритмов и соответствующим образом представляет выходные результаты.

Интерпретация или оценка данных. Этот шаг выполняется для интерпретации и извлечения знаний из выявленных шаблонов.

Термин «интеллектуальный анализ данных» часто используется взаимозаменяемо с KDD. Обнаружение знаний в базах данных предназначено для охвата общего процесса обнаружения полезных знаний из данных. Между тем интеллектуальный анализ данных относится к четвертому этапу процесса KDD. Обычно это считается «основным шагом», на котором применяются алгоритмы для извлечения шаблонов из данных.

Основной целью процесса KDD является извлечение информации из данных в контексте больших баз данных. Это осуществляется с помощью алгоритмов интеллектуального анализа данных для определения того, что считается знаниями.

Обнаружение знаний в базах данных рассматривается как запрограммированный, исследовательский анализ и моделирование обширных хранилищ данных. KDD – это организованная процедура распознавания допустимых, полезных и понятных шаблонов из огромных и сложных наборов данных. Интеллектуальный анализ данных является основой процедуры KDD, включая вывод алгоритмов, которые исследуют данные, разрабатывают модель и находят ранее неизвестные закономерности. Модель используется для извлечения знаний из данных, анализа данных и прогнозирования данных.

Доступность и обилие данных сегодня делают обнаружение знаний и интеллектуальный анализ данных вопросом впечатляющей важности и необходимости. В связи с недавним развитием этой области неудивительно, что в настоящее время специалистам и экспертам доступны самые разнообразные методы.

В мире, все больше ориентированном на данные, казалось бы, никогда не будет избытка данных. Однако данные ценны только тогда, когда можно их анализировать, сортировать и фильтровать, чтобы экстраполировать фактическое значение.

Большинство сфер деятельности собирают огромные объемы данных, но без механизмов построения графиков, диаграмм и моделей данных, сами по себе чистые данные мало полезны.

Однако сам объем данных и скорость, с которой они собираются, затрудняют их анализ. Таким образом, стало экономически и научно необходимым расширить аналитические возможности для обработки огромного получаемого объема данных.

Поскольку компьютеры позволили собирать больше данных, чем люди сами способны обработать, мы обращаемся к вычислительным методам, которые помогают нам извлекать значимые шаблоны и структуры из огромных объемов данных.

По прогнозам расходы на искусственный интеллект и машинное обучение возрастут к концу 2021 на 80% по сравнению с 2017 годом; 74% из более чем 1600 опрошенных владельцев крупных мировых компаний, лиц, принимающих решения, и технологических лидеров считают, что ML меняет правила и стратегию бизнеса, и в целом развитие отрасли [1].

Машинное обучение основано на знаниях. Темпы и качество совершенствования алгоритмов зависят не столько от вычислительной мощности или безупречного кода, сколько от высококачественных данных. Если компания, внедряющая продукт на основе машинного обучения, знает, как собирать информацию, например, информацию от своих пользователей, отслеживать взаимодействие и прослушивать сообщения в социальных сетях, она может достичь замечательных результатов (в частности, таких показателей как компании Google, Amazon, Netflix, Microsoft, Facebook и др.).

Внедряя технологию машинного обучения, компании могут получить значительное преимущество перед конкурентами за счет прозрачности используемых данных, получения дополнительной информации из данных, повышения эффективности внешних и внутренних процессов (например, оптимизации цепочки создания стоимости и операций), лучшего понимания потенциальных клиентов, значительного снижения затрат.

Машинное обучение влияет на бизнес-решения, упрощает рабочие процессы, снижает накладные расходы и улучшает повседневную жизнь в целом.

Внедрение машинного обучения позволяет различным компаниям подкреплять принимаемые управленческие решения фактическими данными, собранными из каждого обработанного заказа, закрытой сделки или взаимодействия с клиентами.

На практике внедрение машинного обучения требует [1]:

Человеческий ресурсы. Каждое решение на основе машинного обучения разрабатывается, создается, внедряется и оптимизируется командой высококвалифицированных специалистов в области ML, Data Science, дата-инженеров, инженеров-программистов, менеджеров проектов.

Время. Внедрение проекта машинного обучения требует значительного количества времени, от нескольких недель до нескольких лет. В процессе внедрения команде ML приходится собирать и очищать данные; разрабатывать, создавать, тестировать и оптимизировать ML-решение. Все эти задачи требуют серьезных временных затрат.

Бюджет. Внедрение машинного обучения достаточно затратно. Компания должна быть финансово состоятельной, чтобы нанять и «содержать» команду машинного обучения по крайней мере в течение нескольких месяцев.

Иными словами, машинное обучение может легко стать либо сложно реализованным проектом, либо «выигрышным билетом» для будущего владельца итогового ML-продукта. В связи с этим, имеет смысл рассмотреть отрасли и конкретные примеры использования машинного обучения.

Учитывая сложность и высокую стоимость решений на основе машинного обучения, данная технология в основном применяется в финансах, банковском деле и здравоохранении.

Однако выделяются и другие ведущие отрасли, применяющие машинное обучение:

- CRM/клиентская аналитика [28];
- финансы;
- банковское дело;
- здравоохранение;
- обнаружение мошенничества;
- наука;
- розничная торговля;
- реклама;
- электронная коммерция;
- образование.

В частности, розничная торговля, электронная коммерция и клиентская аналитика применяют машинное обучение для прогнозирования спроса, оптимизации цен, сегментации клиентов, предоставления рекомендаций клиентам, выявления и предотвращения мошенничества.

В финансах и банковском деле машинное обучение используется для оценки кредитоспособности, анализа рисков, анализа клиентов, выявления мошенничества.

В здравоохранении и медицине машинное обучение применяется для повышения точности диагностики, выявления пациентов с рисками и многое другое.

Рекламодатели и маркетологи тоже полагаются на машинное обучение. Они используют его для сегментации рынков и клиентов, динамической оптимизации цен, анализа показателей оттока, прогнозирования стоимости жизни клиентов, анализа возможностей повышения продаж и анализа настроений в социальных сетях.

Однако технология машинного обучения универсальна и основана на различных алгоритмах, процессах, методах и моделях машинного обучения.

Цель настоящей работы заключается в исследовании и анализе конкретных примеров использования машинного обучения и интеллектуального анализа данных, обзоре аспектов применения данных технологий.

1. Распознавание изображений и видео.

За последние несколько лет в области распознавания изображений и видео наблюдается быстрый прогресс благодаря развитию методов глубокого обучения (*deep learning, DL*) [16], которое является частью машинного обучения.

Распознавание изображений и видео [2, 14] используется для распознавания лиц, обнаружения объектов, анализа текста (печатного и рукописного), визуального поиска, композиции изображений.

Машины хорошо обрабатывают изображения. Например, участники ImageNet, одного из многих соревнований по компьютерному зрению, теперь могут обучать модели DL, которые распознают и классифицируют изображения в наборе данных намного лучше, чем люди.

С помощью машинного обучения программное обеспечение можно обучить обнаруживать объекты на изображениях. Нейронная сеть анализирует готовую библиотеку изображений пиксель за пикселем. Каждый нейрон предлагает понимание после проверки своего содержимого, и сеть объединяет миллионы этих выводов в единый анализ.

Разработчики используют открытую базу данных изображений, чтобы научить системы обучения передаче распознавать эти изображения. Крупнейшими наборами данных для распознавания изображений ML являются Coco и ImageNet.

Примеры машинного обучения в распознавании изображений [2]:

Clearview – технология распознавания лиц, которая анализирует данные из социальных сетей, чтобы получить представление о лицах людей и получить их данные. Несмотря на множество споров о конфиденциальности, приложение часто используется для предупреждения преступности и работы полиции.

Deeromatic – программное обеспечение для компьютерного зрения, которое позволяет компаниям внедрять искусственный интеллект для распознавания изображений в свое программное обеспечение.

Visenze – онлайн-платформа электронной коммерции, которая позволяет клиентам загружать фотографии любимых товаров и искать похожие товары на складе.

Распознавание видео похоже на распознавание изображений в том смысле, что видео разбивается кадр за кадром и классифицируется как отдельные цифровые изображения.

Анализ видеофайлов – более сложная задача, чем работа с текстом или изображениями. Информация динамична и меняется в режиме реального времени, поэтому извлекать информацию из видеофайлов сложнее.

Машинное обучение может помочь в разработке сложных алгоритмов распознавания видео. Сначала, используя наблюдение за человеком, система научится распознавать человеческие фигуры, неизвестные автомобили и другие подозрительные объекты. Скоро можно будет представить себе систему видеонаблюдения, которая функционирует полностью без какого-либо участия человека.

Примеры машинного обучения в области видеонаблюдения:

OpenVINO – интеллектуальный класс, работающий на основе распознавания видео и обнаружения действий.

Alpha Pose– инструмент обнаружения действий с открытым исходным кодом в режиме реального времени.

Крупные мировые компании, использующие распознавание изображений и видео: Google, Shutterstock, Pinterest, eBay, Salesforce, Yelp, Apple, Amazon, Facebook.

2. Распознавание речи.

Распознавание речи [15] – это еще одна область машинного обучения, которая позволяет машинам «имитировать» человека благодаря интеллектуальному анализу данных, машинному обучению и методам глубокого обучения. Однако в этом случае не пиксели изображения или кадровые видео, а аудиофайлы анализируются и обрабатываются нейронными сетями для преобразования звука в текстовый файл.

Распознавание речи используется в поисковых системах (например, Google, Baidu), виртуальных цифровых помощниках (например, Alexa, Cortana, Siri, Google Assistant, AliGenie), интеллектуальных динамиках (например, Amazon Echo, Google Home) и приложениях с голосовой активацией (например, Uber, Evernote).

Инструменты распознавания речи используют машинное обучение для адаптации программного обеспечения к новым идиомам, сленгам, фразам. Сначала разработчикам необходимо научить систему наиболее распространенным словам и фразам, но затем инструмент автоматически выучит новый набор ответов на основе данных, полученных в результате взаимодействия с пользователями. Два основных вида распознавания речи активно используют машинное обучение:

- автоматическое распознавание речи (*Automated Speech Recognition, ASR*) – это процесс преобразования речевого ввода в письменный текст;
- обработка естественного языка (*Natural Language Processing, NLP*) – процесс анализа содержимого текста для понимания его контекста и значения.

Машинное обучение помогает программному обеспечению адаптироваться к динамическим речевым моделям. Пользователи используют идиомы, сленг, аббревиатуры, и, чтобы оставаться гибкой, системе необходимо изучить все эти различия. Именно здесь машинное обучение имеет важное значение: даже теоретически человеческая команда не может обучить программное обеспечение миллионам вариантов речи вручную. Если система тренируется сама, задача становится гораздо более управляемой.

Примеры распознавания речи и машинного обучения:

Google's Machine Learning Audio Recorder: пользователи могут помечать определенные разделы записи и транскрибировать разговорный звук.

Looking-to-Listen: программное обеспечение отделяет звуковые дорожки от фонового шума, музыки, инструментальных средств и т.д.

3. Обнаружение мошенничества.

В машинном обучении обнаружение мошенничества [16] относится к отдельному классу проблем классификации, наряду с обнаружением спама, системами рекомендаций и прогнозированием дефолта по кредитам.

Чтобы активно выявлять мошенничество, модели ML должны анализировать детали транзакций в режиме реального времени и классифицировать каждую транзакцию как законную или мошенническую, что при наличии достаточного количества данных не так сложно сделать.

Машинное обучение помогает компаниям экономить значительную часть бюджета, обнаруживая, помечая и предотвращая мошеннические транзакции.

4. Медицинская диагностика.

Машинное обучение широко используется в здравоохранении [17, 18], предлагая врачам и медицинским работникам инструменты эффективного сбора и анализа данных пациентов для повышения качества диагностики.

Эффективная диагностика пациентов обеспечивается данными, которые бывают разных форматов и объемов: магнитно-резонансная томография, компьютерная томография, заметки врача, отчеты о патологии, прикроватные мониторы и многое другое.

В настоящее время алгоритмы машинного обучения способны выявлять раковые опухоли и рак кожи, диагностировать диабет и, самое главное, прогнозировать прогрессирование заболевания.

Медицинская отрасль собирает огромный массив данных, большая часть которых представляет собой электронные медицинские записи, собранные медицинскими учреждениями. Интеллектуальный анализ данных становится все более популярным в здравоохранении и необходимым. Огромные объемы данных, генерируемых транзакциями в здравоохранении, не могут быть обработаны и проанализированы традиционными методами из-за сложности и объема данных.

Интеллектуальный анализ данных предоставляет организациям здравоохранения методологию и технологию для:

- оценки эффективности лечения,
- спасения жизни пациентов с помощью прогностической медицины,
- управления здравоохранением на разных уровнях,
- управления отношениями с клиентами,
- выявления мошенничества.

Интеллектуальный анализ данных включает в себя создание правил ассоциаций, использование критериев поддержки для определения наиболее важных взаимосвязей в данных.

Другие параметры интеллектуального анализа данных здравоохранения включают:

- анализ последовательности (поиск закономерностей, когда одно событие приводит к другому более позднему событию),
- классификация (разделение множества объектов на классы на основе признаков, характеризующие эти объекты) [6-8],
- кластеризация (группировка набора объектов и их агрегирование на основе сходства объектов друг с другом),
- и прогнозирование.

Индустрия здравоохранения располагает богатыми источниками данных, такими как электронные медицинские карты, административные отчеты и др.

Сегодня интеллектуальный анализ данных в здравоохранении используется в основном для прогнозирования различных заболеваний, оказания помощи в диагностике и консультирования врачей при принятии клинических решений. Но потенциал интеллектуального анализа данных намного больше – он может предоставлять ответы на вопросы, открытия на основе аномалий, обеспечивать более обоснованные решения, вероятностные показатели, прогнозное моделирование и поддержку принятия решений.

Используя интеллектуальный анализ данных, индустрия здравоохранения может быть очень эффективной в таких областях, как:

- медицинские исследования,
- фармацевтика,
- медицинское оборудование,
- генетика,

- управление больницами,
- медицинское страхование и т.д.

Рассмотрим некоторые приложения интеллектуального анализа данных в отрасли здравоохранения и то, как математический и статистический анализ данных может решать различные проблемы в клинических, финансовых и операционных средах.

Выявление и предотвращение мошенничества.

Одним из наиболее ярких примеров использования интеллектуального анализа данных в здравоохранении является выявление и предотвращение мошенничества. В этой области методы интеллектуального анализа данных включают установление нормальных закономерностей, выявление необычных закономерностей со стороны поставщиков медицинских услуг (клиник, врачей, лабораторий и т.д.).

Измерение эффективности лечения.

Это приложение включает в себя сравнение и сопоставление симптомов, причин и курсов лечения, чтобы найти наиболее эффективный курс действий для определенного заболевания или состояния. Инструменты интеллектуального анализа данных сравнивают симптомы, причины, методы лечения и негативные последствия, определяют побочные эффекты конкретного лечения и анализируют, какое решение было бы наиболее эффективным.

С помощью интеллектуального анализа данных могут разрабатываться интеллектуальные методологии лечения, лучшие стандарты медицинской практики и практики оказания медицинской помощи.

Например, сбор данных о результатах лечения может помочь найти способы предоставления более качественных лекарств по более низкой цене. Полученные данные также позволяют разработать клинические профили, чтобы предоставить врачам информацию об их моделях практики лечения и сравнить с моделями лечения других врачей, отраслевыми стандартами, проверенными экспертами.

Помощь руководству медицинских учреждений.

Инструменты интеллектуального анализа данных позволяют выявлять и отслеживать состояния хронических заболеваний и пациентов с высоким риском, разрабатывать соответствующие схемы лечения и сокращать количество госпитализаций и жалоб.

Это лишь несколько примеров интеллектуального анализа данных в здравоохранении, но его потенциал и преимущества для систем здравоохранения очень многообещающие.

Интеллектуальный анализ данных набирает обороты в отрасли здравоохранения, поскольку данная технология приносит пользу всем заинтересованным сторонам:

- поставщики медицинских услуг могут использовать интеллектуальный анализ данных для определения эффективных методов лечения и передовой практики, а также для разработки руководящих принципов и стандартов оказания медицинской помощи;
- пациенты, особенно страдающие хроническими заболеваниями или заболеваниями высокого риска, могут получать более качественные и доступные медицинские услуги при надлежащем выявлении, отслеживании и использовании соответствующих вмешательств и протоколов лечения;
- организации здравоохранения могут использовать интеллектуальный анализ данных для повышения удовлетворенности пациентов, предоставления более ориентированного на пациента ухода, снижения затрат и повышения эффективности работы при сохранении высокого качества медицинской помощи;

– страховые организации могут выявлять мошенничество в области медицинского страхования с помощью интеллектуального анализа данных и сокращать свои убытки.

Конечно, существует множество проблем на пути регулярного внедрения машинного обучения в здравоохранение. Технология должна быть хорошо протестирована и усовершенствована, прежде чем она сможет взять на себя ответственность за диагностику и лечение пациентов. Это может быть возможно только в том случае, если медицинские знания и опыт в области развития будут объединены в непрерывном сотрудничестве.

Бизнес-приложения на основе машинного обучения в здравоохранении:

– Kensci: прогностическая платформа здравоохранения, которая предупреждает пациентов и медицинских работников о возможных проблемах со здоровьем и помогает предотвращать инфекции и осложнения заболеваний.

– PathAI: диагностическая платформа машинного обучения, которая анализирует истории болезни пациентов и сканирует их для получения точных прогнозов о происхождении и симптомах заболевания.

– InnerEye – платформа, разработанная Microsoft, предназначенная для различения опухолей и здоровых тканей для точной диагностики рака.

5. Обнаружение аномалий.

Обнаружение аномалий [2] широко используется в производстве для повышения производительности и эффективности, снижения затрат и оптимизации времени простоя.

Этапы процесса обнаружения аномалий:

- датчики устанавливаются на часть оборудования для сбора данных;
- модели ML обрабатывают данные для поиска аномальных данных;
- аномальные данные анализируются для выявления конкретной проблемы, связанной с ними;
- проблема решается превентивно, чтобы избежать отказа оборудования.

На самом деле, данный подход может быть применен к широкому кругу задач. Например, мошенничество с кредитными картами, клиническая диагностика, т.е. различного рода аномалиям, которые можно обнаружить с помощью машинного обучения.

Обнаружение аномалий позволяет предприятиям прогнозировать отказ оборудования [13] для проведения технического обслуживания и ремонта, что сокращает эксплуатационные расходы и спасает жизни. Например, датчики интернета вещей (IoT), установленные на самолетах, собирают и анализируют данные, чтобы сообщать о компонентах, требующих технического обслуживания, что снижает количество авиационных происшествий.

6. Оптимизация запасов.

Оптимизация запасов [19] – один из наиболее незаметных, но критически важных примеров использования машинного обучения. Это позволяет машинам контролировать, сколько запасов необходимо хранить и как наиболее эффективно их хранить, чтобы гарантировать, что цепочка поставок не нарушится.

Другими словами, оптимизация запасов на базе ML гарантирует, что склады компании обеспечены достаточным количеством продуктов для продажи, что эти продукты эффективно хранятся и распространяются, а клиенты получают свои покупки вовремя.

Компания Amazon является мировым лидером в оптимизации своих запасов с помощью машинного обучения. Компании удается отправлять в среднем 1,6 миллиона посылок в день с безупречной точностью выполнения заказов.

7. Прогнозирование спроса.

Прогнозирование спроса [20] применяется в самых разных отраслях, от электронной коммерции и розничной торговли до транспорта и производства. При этом модели ML используют исторические данные, чтобы предсказать, что будет потреблено на рынке в данный период времени: количество потребления какого-либо продукта, услуги, сырья, энергии и т.п.

Это позволяет предприятиям собирать и обрабатывать данные по всей цепочке поставок, что повышает эффективность и снижает накладные расходы.

Прогнозирование спроса на основе ML является быстрым, точным и прозрачным. Это позволяет компаниям получать информацию из постоянного потока данных о спросе/предложении и упреждающе адаптироваться к изменениям. Прогнозирование спроса является самым популярным приложением машинного обучения для планирования цепочки поставок.

8. Системы рекомендаций.

Рекомендательные системы [21] являются одним из наиболее распространенных применений машинного обучения в повседневной жизни. Эти системы используются в поисковых системах, веб-сайтах электронной коммерции (например, Amazon, eBay), развлекательных платформах (например, Netflix, Google Play), играх и веб-и мобильных приложениях.

Рекомендательные системы обычно классифицируются методом фильтрации.

Метод фильтрации на основе контента. Данный метод рекомендует объекты пользователю на основе тех объектов, с которыми этот конкретный пользователь взаимодействовал. Например, если вы купили книгу о машинном обучении в Amazon, в разделе «Предложения» будут отображаться книги, ориентированные на ML.

Метод совместной фильтрации. В методе совместной фильтрации система рекомендаций анализирует действия пула пользователей для вычисления индекса сходства и дальнейшего отображения похожих объектов похожим пользователям.

Однако существуют и более продвинутые типы рекомендательных систем.

Рекомендательные системы обрабатывают различные источники данных: поиск, клики, просмотры товаров, просмотры страниц, заполнение форм, покупки, повторные посещения, а также сведения о товарах (например, название, цена, категория) и контекстные данные (например, устройство, местоположение, язык).

Рекомендательные системы являются ценным активом для любого бизнеса, поскольку они позволяют увеличить трафик, предоставлять релевантный контент, повышать вовлеченность клиентов и конверсии, снижать отток и увеличивать прибыль.

9. Обнаружение вторжений.

Обнаружение вторжений на базе ML [22] – это неотъемлемый ресурс адаптивных систем обнаружения вторжений (*Intrusion Detection System, IDS*), которые отслеживают сети в режиме реального времени для выявления и борьбы со вредоносным трафиком или методами вторжения, такими как проникновение и несанкционированный доступ.

Машинное обучение помогло произвести революцию в области обнаружения вторжений.

Традиционно IDS был разработан для выявления известных угроз. Злоумышленники могут разработать новый метод вторжения в обход системы.

Однако теперь сетевые данные непрерывно собираются и предварительно обрабатываются для создания высококачественных наборов данных, которые используются для обучения моделей машинного обучения, которые эффективно отличают обычный трафик от вредоносного трафика в режиме реального времени.

10. Анализ социальных сетей.

Машинное обучение может анализировать миллионы сообщений в социальных сетях [9, 10], например Facebook, Twitter, Instagram, читать комментарии и личные обновления. Компании следят за отзывами клиентов, отслеживают состояние своего бренда и улучшают репутацию.

Машинное обучение позволяет системам не только распознавать слова, но и понимать контекст, стоящий за ними. Например, “оранжевый” может использоваться для описания цвета или названия бренда. Если пользователю необходимо увидеть сообщения, относящиеся ко второй категории, ML делает это возможным, анализируя контекст сообщения.

Методы машинного обучения анализируют контекст, в котором используется то или иное название, и сравнивают одно сообщение с миллионами похожих сообщений. Сложные алгоритмы отслеживают различия между счастливым, несчастным, заинтересованным или саркастичным настроением пользователей.

Компании используют машинное обучение, чтобы знать, что волнует их клиентов. Машинное обучение не ограничивается заранее установленными значениями: другими словами, оно покажет все результаты, соответствующие общему контексту, а не только то, что искал пользователь. При этом можно получить неожиданные результаты, которые не соответствуют ожиданиям пользователя.

Применение машинного обучения для анализа социальных сетей:

- [Lionbridge](#): этот инструмент анализа настроений предоставляет пользователям информацию, основанную на сообщениях в социальных сетях на более чем трехсот языках.
- Scale AI: компания по анализу данных, которая обрабатывает информацию из социальных сетей, онлайн-поисков, сообщений, баз данных, используя алгоритмы машинного обучения.
- Monkey Learn: программное обеспечение для классификации текста, в котором пользователи могут анализировать текст, не ограничиваясь только сообщениями в социальных сетях, хотя оно часто используется для социального прослушивания.

11. Распознавание эмоций.

Распознавание эмоций [10] уже широко используется различными компаниями для оценки настроения потребителей по отношению к их продукту или бренду. Возможности, открываемые этой технологией, выходят за рамки маркетинговых исследований и цифровой рекламы.

Распознавание эмоций в здравоохранении.

Отрасль, которая использует преимущества этой технологии, – это здравоохранение, где программное обеспечение для распознавания на базе интеллектуального анализа данных помогает решать, когда пациентам требуется лекарство, или помогает врачам определить, к кому обратиться в первую очередь.

Автомобильная промышленность и распознавание эмоций.

Автомобильная промышленность также применяет технологию распознавания эмоций, поскольку производители автомобилей по всему миру все больше внимания уделяют тому, чтобы сделать автомобили более личными и безопасными для людей за рулем. Последнее является в основном областью, привлекающей внимание, и к настоящему времени различные компании предприняли шаги по тестированию и исследованию новых технологий. В своем стремлении создать больше интеллектуальных функций для автомобилей автопроизводителям имеет смысл использовать искусственный интеллект, чтобы помочь понять человеческие эмоции. Например, используя распознавание эмоций на лице, умные автомобили могут предупреждать водителя, когда он чувствует сонливость.

Распознавание эмоций в тестировании видеоигр.

Видеоигры разрабатываются с учетом конкретной целевой аудитории и направлены на то, чтобы вызвать у пользователей определенное поведение и набор эмоций. На этапе тестирования пользователям предлагается поиграть в игру в течение определенного периода времени, и их отзывы учитываются для создания конечного продукта. Использование распознавания эмоций на лице может помочь понять, какие эмоции испытывает пользователь в режиме реального времени во время воспроизведения.

Обнаружение эмоций является довольно сложной задачей. Тем не менее алгоритмы машинного обучения показали большие перспективы для решения данной задачи. Используя распознавание эмоций на лице, компании могут обрабатывать изображения и видео в режиме реального времени для мониторинга видеопотоков или автоматизации видеоаналитики, тем самым экономя затраты и улучшая жизнь своих пользователей.

12. Умные помощники.

Умные помощники [2, 13] ищут новые способы взаимодействия с пользователями и обеспечивают лучший опыт. Им необходимо анализировать персональные данные, обрабатывать голосовые запросы и автоматизировать ежедневные задачи. Без мощной системы обучения помощник не сможет адаптироваться к меняющимся потребностям пользователей.

С помощью машинного обучения домашний помощник может обрабатывать и классифицировать все входные данные и, возможно, повторно использовать их позже. С помощью обработки естественного языка помощник собирает и анализирует голосовые сообщения пользователей, находит логические связи между фразами и понимает, какие выходные данные соответствуют запросу.

Машинное обучение позволяет умному помощнику использовать все собранные данные для улучшения своих навыков распознавания образов и удовлетворения новых потребностей. Например, компания Amazon постоянно совершенствует алгоритмы машинного обучения для Alexa.

Основным методом ML, используемым системой, является активное обучение – система использует данные для автоматического овладения новыми навыками, но если это невозможно, программное обеспечение определит области, в которых требуется помощь специалистов-людей. Таким образом, команды разработчиков всегда знают о "слепых зонах" инструмента и знают, где требуется вмешательство.

Кроме того, ученые представили новый метод ML – трансформируемое обучение [2]. Разработчики могут повторно использовать навыки из одной области, чтобы помочь машинам приобрести компетенции в другой области. В частности, если система знает, как найти рестораны, она быстро овладеет навыками поиска баров и супермаркетов.

13. Классификация новостей.

Поскольку объем производимого контента растет экспоненциально, компаниям и отдельным пользователям нужны инструменты, которые классифицировали бы и сортировали информацию [23]. Машинное обучение обеспечивает прочную основу для такой технологии. Алгоритмы могут просматривать миллионы статей на многих языках и выбирать те, которые соответствуют интересам и привычкам пользователей.

Кроме того, машинное обучение потенциально может распознавать поддельные новости и оценивать достоверность источников. Во-первых, сеть должна просмотреть новости, подтвержденные как поддельные, и сравнить их с теми, которые, как известно, являются законными. После определения ключевых различий и обнаружения тонких закономерностей система сможет разработать сложные алгоритмы дифференцирования.

Использование машинного обучения для классификации новостей:

- Aylien – инструмент на основе ML для оценки легитимности новостей и персонализации ленты новостей; инструмент также оценивает качество материалов и стремится собирать лучшие обновления из миллионов торговых точек и блогов.

- TagMyNews – активный проект на GitHub, который использует машинное обучение на основе Python для классификации новостей.

14. Оптимизация результатов поиска в поисковых системах.

Алгоритмы машинного обучения могут изучать привычки и интересы пользователей на основе анализа статистики поиска [2, 24]. Алгоритмы ранжирования не полагаются на метатеги и ключевые слова, а вместо этого анализируют контекст страницы.

Поисковые системы могут обучать исследовательские системы распознавания, обладающие опытом в определенных областях. Страницы, связанные со здравоохранением, будут анализироваться алгоритмом, обученным для этой задачи, но не контентом, посвященным развлечениям.

Примеры применения машинного обучения в поисковых системах:

Google RankBrain – алгоритм поисковой системы, который использует глубокое обучение для детального анализа содержимого страниц. Алгоритмы определяют контекст страницы, а не просто полагаются на мета-теги для определения релевантности.

TinEye – система поиска обратных изображений, работающая на основе машинного обучения.

15. Анализ электронной почты.

Поскольку машинное обучение оказалось важной технологией в интеллектуальном анализе текста [12, 16], неудивительно, что оно обладает огромным потенциалом в обнаружении и фильтрации спама. Методы машинного обучения способны анализировать и сравнивать легитимные электронные письма со спамом и определять различия даже в тех случаях, когда люди легко допустили бы ошибку.

Примеры машинного обучения при анализе спама:

- SpamTitan: веб-платформа, которая использует машинное обучение для выявления мошенничества и рекламного спама.

- Mailwasher: инструмент мониторинга электронной почты, который обнаруживает спам и оскорбительную лексику. Программное обеспечение постоянно совершенствует свои навыки распознавания с помощью машинного обучения и технологий big data.

16. Кибербезопасность.

Постоянно возникают новые угрозы безопасности, и эксперты-люди не могут следить за всеми обновлениями. Машинное обучение, с другой стороны, гораздо более гибкое: алгоритмы используют данные с миллионов сайтов и узнают о новых вредоносных программах со всего мира.

При этом обнаружение угроз кибербезопасности [25] также является немедленным: система распознает угрозу, проанализирует аналогичные случаи и примет меры для защиты веб-сайта или приложения. Машинное обучение позволяет компаниям быть в курсе вредоносных практик и прогнозировать проблемы безопасности еще до того, как они возникнут.

Примеры машинного обучения в области безопасности бизнеса:

- Google: компания инвестирует в машинное обучение для защиты Android и обнаружения угроз в режиме реального времени.

- Amazon: компания приобрела стартап по безопасности на основе машинного обучения – Macie, который использует ML для обеспечения облачной безопасности.

– Microsoft: платформа защиты от угроз Windows использует машинное обучение для обнаружения рисков и нарушений безопасности, а также для немедленного расследования этих рисков.

17. Клиентские сервисы (чат-боты).

Алгоритмы машинного обучения анализируют поведение клиентов на веб-сайтах по коллективным качественным (общие действия, интересы, любимые продукты) и количественным (затраченное время, количество посещенных страниц, возраст, пол и т.д.) данным. Затем эта информация обрабатывается для предоставления подробного отчета о потребностях клиента. Исходя из этих данных, разработчики чат-ботов [11, 13] могут знать, на каких проблемах следует сосредоточиться.

Для обучения чат-ботов разработчикам необходимо проверить, понятны ли и актуальны ли результаты. Как только несколько десятков ответов были подтверждены, чат-боты могут самостоятельно учиться на ежедневном взаимодействии с клиентами, совершенствуясь с каждым диалогом.

Примеры машинного обучения в разработке чат-ботов:

– Answer Bot by Zendesk: бот для машинного обучения рекомендует соответствующие справочные статьи посетителям веб-сайта.

– Mitsuku: разговорный чат-бот, который использует машинное обучение, обработку естественного языка и искусственный интеллект для понимания человеческих разговоров и создания соответствующих ответов. Чат-бот пять раз выигрывал тест Тьюринга и был показан в BBC, Wired, New York Times и других изданиях.

– Sephora messenger bot [13]: бот предлагает ответы на заранее заданные вопросы. Коллекция этих вопросов постоянно растет, потому что машинное обучение обрабатывает новые запросы пользователей.

– Watson assistant – разговорный чат-бот с искусственным интеллектом, разработанный IBM, который использует машинное обучение для распознавания речи.

18. Идентификация авторских прав.

Компании могут использовать нейронные сети ML для выявления случаев нарушения авторских прав и определения законных владельцев различного рода контента [26]. Нейронные сети могут анализировать миллионы текстов, изображений и видео, обнаруживать все копии одного файла и отслеживать его оригинал.

Обнаружить, что точная копия контекста использовалась без разрешения, довольно просто. Сложности возникают при определении того, когда элемент или часть изображения, песни или текста используются повторно без разрешения. В таких случаях алгоритмы машинного обучения должны анализировать содержимое и понимать их контекст.

Однако глубокое обучение позволяет определять даже тонкие закономерности и обнаруживать контент, который технически не идентичен, но все равно может нарушать законы об авторском праве.

Примеры ML при идентификации авторского права:

– Content ID on Youtube: каждое загруженное видео получает уникальный код, который определяет происхождение и характер отснятого материала. Если новое видео потенциально соответствует идентификатору существующего, более поздние загруженные кадры будут наказаны за нарушение авторских прав.

– Shazam: инструмент распознает песни из записей низкого качества, и аналогичная технология потенциально может быть использована для обнаружения небольшого сходства между песнями.

19. Машинное обучение для прогнозирования.

Платформы управления, инструменты прогнозирования, программное обеспечение для прогнозирования [27] используют машинное обучение для получения информации из ранее собранных данных, выявления повторяющихся закономерностей и моделирования будущего на основе этих закономерностей.

Алгоритмы вычисляют вероятность события с учетом его частоты в прошлом. Кроме того, инструменты включают общедоступные данные для расширения сферы исследований: аналитическое программное обеспечение не только должно использовать данные компании, но и может учитывать тенденции рынка.

Примеры использования машинного обучения для прогнозирования:

- EverString – программное обеспечение для прогнозной аналитики, которое разбивает целевую аудиторию бизнеса на сегменты в зависимости от ее поведения, интересов и потребностей.
- Infer – программное обеспечение выполняет прогнозное моделирование и предоставляет информацию о позиции в воронке продаж.

20. Идентификация возраста/пола.

С ростом социальных сетей и видеоконтента организации и предприятия заинтересованы в определении возраста и пола людей по их фотографиям, стилю сообщений, а также голосам [2]. Технологии машинного обучения для идентификации возраста и пола имеют множество применений в таких областях, как право, контроль доступа, безопасность и другие.

Примеры возрастного и гендерного анализа на основе ML:

- Quividi: приложение для распознавания лиц, которое использует искусственный интеллект и машинное обучение для анализа лиц онлайн-пользователей и предоставления точной информации о целевой рекламе.
- AgeBot – приложение для Android, которое предоставляет пользователям оценку их возраста и пола на основе фотографий, видео и изображений в реальном времени с камеры смартфона.

21. Кредитный скоринг.

Искусственный интеллект сегодня стремительно развивается, предлагая огромную помощь людям во всех секторах бизнеса и экономики. С помощью интеллектуального анализа данных автоматизированные решения принимаются намного быстрее из-за значительно более высокой скорости обработки данных в компьютерных системах [30].

Основные преимущества применения систем искусственного интеллекта в финансовой индустрии заключаются в следующем:

- автоматизация рутинных процессов;
- повышенная скорость обслуживания;
- снижение затрат на решение стандартных задач;
- повышенная точность обработки больших объемов данных;
- улучшено качество системы поддержки клиентов.

В финансовом секторе и банковской сфере в настоящее время применяются передовые аналитические технологии с огромным потенциалом, которые улучшают платежную экосистему для всех участников, от банков до конечных потребителей.

Кредитный скоринг [30] на основе искусственного интеллекта, пожалуй, является наиболее перспективным и актуальным.

Кредитный скоринг представляет собой оценку того, насколько хорошо клиент банка может заплатить и готов погасить задолженность. Учитывая, что во всем мире большинство людей владеют банковскими счетами и считаются имеющими право на

кредитование, необходимость в более эффективных решениях для оценки кредитоспособности очевидна.

Решения в задачах кредитного скоринга основаны на большом количестве данных, таких как общий доход, кредитная история, анализ транзакций, опыт работы. По сути, скоринг представляет собой математическую модель, основанную на статистических методах и учитывающую большой объем информации. В результате кредитный скоринг с использованием технологий DM и ML обеспечивает более чувствительную, индивидуальную оценку кредитного рейтинга на основе множества дополнительных факторов в режиме реального времени, предоставляя доступ к финансированию большему числу людей с потенциалом дохода.

В большинстве финансовых учреждений модели кредитного скоринга по-прежнему основаны на подходе системы показателей, т. е. на динамике, характерной для времени их создания. Потенциальный заемщик должен обладать достаточными историческими данными о возвратах предыдущих кредитов, чтобы быть оцененным как «приемлемый». В случае отсутствия такой исторической информации (что является типичной ситуацией для новых клиентов банковского сектора) даже кредитоспособным заемщикам отказывают в доступе к кредиту.

В отличие от традиционных методов оценки кредитоспособности (например, метода показателей), ориентированных на прошлые показатели заемщика, оценка кредитоспособности с применением технологий DM и ML более чувствительна к показателям кредитоспособности потенциального заемщика в реальном времени, таким как текущий уровень дохода, возможности трудоустройства и их потенциальная способность зарабатывать.

Таким образом, заемщики с высоким потенциалом включаются в кредитные программы, в то время как те, кто формально проходит традиционную оценку кредитного скоринга (например, производители кредитных карт), исключаются из них. Другими словами, кредитный скоринг на основе искусственного интеллекта позволяет точно прогнозировать прибыль на основе моделей машинного обучения.

Выдача кредитов всегда связана с рисками, такими как неспособность получателя вернуть кредит. Следовательно, целью кредитного скоринга является минимизация рисков, которые вызывают системную проблему, связанную с невозвратом кредита.

Решения на основе интеллектуального анализа данных во многом трансформируют кредитный скоринг, сигнализируя о смене парадигмы в этом аспекте. Рассмотрим некоторые очевидные преимущества интеграции интеллектуального анализа данных и машинного обучения в кредитный скоринг.

Ориентация на клиента.

Традиционные методы кредитного скоринга неоднократно критикуются как устаревшие, в большей степени из-за их стандартизации и отсутствия чувствительности к индивидуальным различиям и нюансам. С внедрением технологий DM и ML в системы кредитного скоринга банки получают уникальную информацию о финансовом поведении своих клиентов не только на основе исторических данных, но и на основе прогнозирования потенциального дохода [27]. Такой анализ крупномасштабных данных о клиентах позволяет повысить сегментацию клиентов и их рейтинг с точки зрения связанного кредитного риска, позволяя финансовым учреждениям оценивать и продавать свои кредитные продукты нужным категориям клиентов.

Высокая скорость обработки данных.

В области кредитного скоринга приложения на основе ML и DM становятся все более популярными благодаря своей способности ускорять процесс принятия решений о кредитовании без каких-либо компромиссов в отношении качества или точности данных.

Традиционно банки применяли деревья принятия решений, регрессию и сложный арифметический анализ для формирования кредитного рейтинга клиента.

В настоящее время в анализ включаются массы избыточных, неструктурированных и частично структурированных данных (например, использование социальных сетей, активность мобильных телефонов и т.д.).

Для принятия более разумных решений, связанных с кредитами, но с помощью интеллектуального анализа данных достигается высокая скорость обработки данных.

Широкий доступ к кредитам.

С помощью Data Science кредитный скоринг стал более ориентированным на будущее, в отличие от старых традиционных подходов. Таким образом, сегодня все больше заемщиков получают доступ к кредитам (например, студенты, основатели перспективных предприятий, иностранные резиденты), что стимулирует их бизнес и помогает им продвигать свои идеи. Получение первого в истории клиента кредита также стало проще, что теперь делается на основе финансовых прогнозов с помощью интеллектуального анализа данных относительно потенциального дохода клиента и возможностей трудоустройства.

Выгоды для всех заинтересованных сторон.

Использование инструментов искусственного интеллекта для кредитного скоринга и принятия решений о кредитовании может не только увеличить количество клиентов банка при одновременном снижении риска, но и помочь банкам принимать решения, основанные на данных, сосредоточиться на максимизации маржи вместо минимизации риска, анализировать плавную кривую соотношения риска и прибыли вместо того, чтобы полагаться на предварительно рассчитанные скоринговые карты.

Данные преимущества были практически невозможны до широкого внедрения машинного обучения и методов сбора данных. Как банки, так и клиенты могут извлечь выгоду из внедрения ML в кредитный скоринг: банки увеличивают число клиентов и их прибыль, в то время как клиенты, нуждающиеся в кредитах, получают более широкий доступ к лучшим кредитным продуктам.

Заключение. Внедрение машинного обучения стремительно развивается, и это неудивительно, учитывая преимущества данных технологий: от устранения ручных задач до получения полезной информации из данных.

Машинное обучение – это долгосрочная инвестиция, которая обеспечивает непрерывные улучшения по мере развития различного вида деятельности. Программное обеспечение, оснащенное алгоритмами ML, будет улучшаться с каждым следующим взаимодействием с пользователем.

Наряду с эффективностью, машинное обучение привлекает своей универсальностью.

Лучшие приложения машинного обучения доказывают, что ML может быть адаптирован для улучшения любого типа процессов, особенно тех, которые связаны с данными.

С ростом больших данных и автоматизации машинное обучение становится неотъемлемым инструментом, потому что человеческие ресурсы больше не могут соответствовать постоянно растущим требованиям рынка.

В настоящей статье мы рассмотрели конкретные примеры использования машинного обучения: распознавание изображений и речи, обнаружение мошенничества,

диагностика пациентов, обнаружение аномалий, оптимизация запасов, рекомендательные системы, кредитный скоринг и др.

Однако это лишь наиболее распространенные примеры машинного обучения. Среди других примечательных упоминаний: машинный перевод, научные исследования в области образования, интеллектуальное управление сетями, оценка деятельности персонала и многое другое.

Каждый подход использования машинного обучения отличается дизайном, архитектурой, инструментами и оптимизацией. Кроме того, определенную роль играют доступность, качество, хранение данных, вычислительная мощность и многие другие факторы.

Список литературы

1. The Top 9 Machine Learning Use Cases in Business [Электронный ресурс]. Режим доступа: <https://squadex.com/insights/top-machine-learning-use-cases-business/amp>, свободный (дата обращения: 25.07.2021).
2. Yusov, K. Best Machine Learning Applications with Examples [Электронный ресурс]. Режим доступа: <https://jelvix.com/blog/machine-learning-use-cases>, свободный (дата обращения: 25.07.2021).
3. Popular Machine Learning Applications and Use Cases in our Daily Life [Электронный ресурс]. Режим доступа: <https://www.analyticsvidhya.com/blog/2019/07/ultimate-list-popular-machine-learning-use-cases/>, свободный (дата обращения: 25.07.2021).
4. Hakikur Rahman. Data Mining Applications for Empowering Knowledge Societies / Hakikur Rahman. - Information Science Reference, 2009. - 332 p.
5. Kamath, C. Data Mining for Scientific and Engineering Applications / C. Kamath, P. Kegelmeyer, R. Namburu, R.L. Grossman, V. Kumar. - Springer US, 2001. - 605 p.
6. Ключева, И. А. Ансамблевые методы в задаче многоклассовой SVM-классификации / Б. В. Костров, А. И. Баранчиков, И. А. Ключева // XXI век: итоги прошлого и проблемы настоящего плюс. - 2021. - Т. 10, № 2 (54). - С. 105-108.
7. Klyueva, I. The two-stage classification based on 1-SVM and RF classifiers / L. Demidova, I. Klyueva // Journal of Physics: Conference Series. - 2020. - Vol. 1727. - P. 012007.
8. Klyueva, I. Intellectual Approaches to Improvement Of the Classification Decisions Quality On the Base Of the SVM Classifier / L. Demidova, I. Klyueva, Y. Sokolova, N. Stepanov, N. Tyart // Procedia Computer Science. - 2017. - Vol. 103. - P. 222-230.
9. Смирнова, О. С. Основные методы анализа, используемые при исследовании социальных сетей / Смирнова, О. С., Петров А. И., Бабийчук Г. А. // Современные информационные технологии и ИТ-образование. - 2016. - №3-1. - С. 151-158.
10. Баранюк, В.В. Подходы к определению психоэмоциональных особенностей информационного образа пользователя социальных сетей / Баранюк, В.В., Десяткова А.Д., Смирнова О.С. // International Journal of Open Information Technologies. - 2016. - Том 4, № 8. - С. 61 - 65.
11. Алымов, А. С. Детектирование бот-программ, имитирующих поведение людей в социальной сети «ВКонтакте» / Алымов, А. С., Баранюк, В.В., Смирнова, О.С. // International Journal of Open Information Technologies. - 2016. - Том 4, № 8. - С. 55 - 60.
12. Смирнова, О.С. Выбор топологии нейронных сетей и их применение для классификации коротких текстов / Смирнова, О.С., Шишков, В.В. // International Journal of Open Information Technologies. - 2016. - Том 4, № 8. - С. 50 - 54.

13. Саханевич, Д. Ю. Исследование подходов и методов применения искусственного интеллекта и машинного обучения в социально-экономических процессах // Вестник ОмГУ. Серия: Экономика. – 2020. – №2. – С. 66 – 79.
14. Булатников, Е. В. Сравнение библиотек компьютерного зрения для применения в приложении, использующем технологию распознавания плоских изображений / Булатников Е. В., Гоева А. А. // Вестник МГУП. – 2015. – №6.
15. Гапочкин, А. В. Нейронные сети в системах распознавания речи // Science Time. – 2014. – №1 (1). – С. 29 – 36.
16. Калугина, К. Н. Искусственный интеллект: машинное и глубокое обучение // Теория и практика экономики и предпринимательства. – 2021.
17. Тхамокова, М. Р. Применение метода машинного обучения для решения задачи медицинской диагностики / Тхамокова, М. Р., Лютикова Л. А. // Научные проекты и технологии в машино-и приборостроении, медицине. – 2018. – С. 190-194.
18. Марухина, О.В. Применение методов Data mining для выявления скрытых закономерностей в задачах анализа медицинских данных / Марухина О.В., Мокина Е.Е., Берестнева Е.В. // Фундаментальные исследования. – 2015. – № 4. – С. 107-113.
19. Шестан, В. Б. Машинное обучение в производстве / Шестан В. Б., Михтадов Р. Э. // Прорывные научные исследования: проблемы, закономерности, перспективы. – 2020. – С. 121-123.
20. Завалич, К. И. Прогнозирование спроса на продукцию производственного предприятия методами машинного обучения / Завалич К. И., Мельников В. А. // Информационные технологии и системы. – 2019. – С. 240-244.
21. Лебедев, Б. Д. Рекомендательные системы с применением машинного обучения для интернет-ресурсов // Modern Science. – 2019. – №. 5(3). – С. 265-268.
22. Волков, С. С. Применение методов машинного обучения в SDN в задачах обнаружения вторжений / Волков С. С., Курочкин И. И. // International Journal of Open Information Technologies. – 2019. – Т. 7. – №. 11.
23. Мальчиц, В. С. Применение методов машинного обучения для классификации новостей // Молодежь XXI века: шаг в будущее. – 2019. – С. 208-209.
24. Царамов, М. В. Анализ алгоритмов работы поисковых систем на основе искусственного интеллекта / Царамов М. В., Симонян А. А. // Методы и средства обработки и хранения информации. – 2019. – С. 159-163.
25. Боброва, М. В. Машинное обучение в кибербезопасности / Боброва М. В., Мاستилин А. Е. // Научные междисциплинарные исследования. – 2021. – С. 24-29.
26. Куртукова, А. В. Идентификация автора исходного кода методами машинного обучения / Куртукова А. В., Романов А. С. // Информатика и автоматизация. – 2019. – Т. 18. – №. 3. – С. 742-766.
27. Шунина, Ю. С. Прогнозирование кредитоспособности клиентов на основе методов машинного обучения / Шунина Ю. С., Алексеева В. А., Клячкин В. Н. // Финансы и кредит. – 2015. – №. 27 (651).
28. Шполянская, И.Ю. Использование технологий datamining с целью создания аналитических CRM-систем для малого бизнеса // Вестник РГЭУ РИНХ. – 2014. – №1 (45). – С. 129-135.
29. Хлупичев, В.А. Технологии DataMining для решения задач бизнеса / Хлупичев В.А., Савостьянова И.Л. // Актуальные проблемы авиации и космонавтики. – 2015. – №11 – С. 627-628.

30. Sofinskyi, V. Credit Scoring Using Machine Learning. [Электронный ресурс]. Режим доступа: <https://datrics.ai/use-cases/credit-scoring-using-machine-learning>, свободный (дата обращения: 25.07.2021).

References

1. The Top 9 Machine Learning Use Cases in Business [Electronic resource]. URL: <https://squadex.com/insights/top-machine-learning-use-cases-business/>, free (date of access 25.07.2021).
2. Yusov, K. Best Machine Learning Applications with Examples [Electronic resource]. URL: <https://jelvix.com/blog/machine-learning-use-cases>, free (date of access 25.07.2021).
3. Popular Machine Learning Applications and Use Cases in our Daily Life [Electronic resource]. URL: <https://www.analyticsvidhya.com/blog/2019/07/ultimate-list-popular-machine-learning-use-cases/>, free (date of access 25.07.2021).
4. Hakikur Rahman. Data Mining Applications for Empowering Knowledge Societies / Hakikur Rahman. - Information Science Reference, 2009. - 332 p.
5. Kamath, C. Data Mining for Scientific and Engineering Applications / C. Kamath, P. Kegelmeyer, R. Namburu, R.L. Grossman, V. Kumar. - Springer US, 2001. - 605 p.
6. Klyueva, I. A. Ensemble methods in the problem of multi-class SVM classification / B. V. Kostrov, A. I. Baranchikov, I. A. Klyueva / XXI century: results of the past and problems of the present. - 2021. - T. 10, № 2 (54). - pp. 105-108.
7. Klyueva, I. The two-stage classification based on 1-SVM and RF classifiers / L. Demidova, I. Klyueva // Journal of Physics: Conference Series. - 2020. - Vol. 1727. - pp. 012007.
8. Klyueva, I. Intellectual Approaches to Improvement Of the Classification Decisions Quality On the Base Of the SVM Classifier / L. Demidova, I. Klyueva, Y. [Sokolova](#), N. [Stepanov](#), N. [Tyart](#) // [Procedia Computer Science](#). - 2017. - [Vol. 103](#). - pp. 222-230.
9. Smirnova, O. S. The main methods of analysis used in the study of social networks / Smirnova, O. S., Petrov A. I., Babiychuk G. A. // Modern information technologies and IT education. - 2016. - No. 3-1. - pp. 151-158.
10. Baranyuk, V. V. Approaches to the definition of psychoemotional features of the information image of a user of social networks / Baranyuk, V. V., Desyatkova A. D., Smirnova O. S. // International Journal of Open Information Technologies. - 2016. - Vol. 4, No. 8. - pp. 61-65.
11. Alymov, A. S. Detection of bot programs that imitate the behavior of people in the social network "Vkontakte" / Alymov A. S., Baranyuk V. V., Smirnova O. S. // International Journal of Open Information Technologies. - 2016. - Vol. 4, No. 8. - pp. 55-60.
12. Smirnova, O. S. The choice of neural network topology and their application for the classification of short texts / Smirnova, O. S., Shishkov, V. V. // International Journal of Open Information Technologies. - 2016. - Vol. 4, No. 8. - pp. 50-54.
13. Zakhanevich, D. Yu. Research of approaches and methods of application of artificial intelligence and machine learning in socio-economic processes // Vestnik Omsu. Series: Economics. - 2020. - No. 2. - pp. 66-79.
14. Bulatnikov, E. V. Comparison of computer vision libraries for use in an application using flat image recognition technology / Bulatnikov, E. V., Goeva, A. A. // Vestnik MGUP. - 2015. - No. 6.
15. Gapochkin, A.V. Neural networks in speech recognition systems // The time of science. - 2014. - №1 (1). - pp. 29 -36.
16. Kalugina, K. N. Artificial intelligence: machine and deep learning // Theory and practice of economics and entrepreneurship. - 2021.

17. Thamokova, M. R. Application of the machine learning method for solving the problem of medical diagnostics / Thamokova, M. R., Lyutikova L. A. // High-tech projects and technologies in machine and instrument engineering, medicine. - 2018. - pp. 190-194.
18. Marukhina, O. V. Application of Data mining methods to identify hidden patterns in the problems of medical data analysis / Marukhina O. V., Mokina E. E., Berestneva E. V. // Fundamental research. - 2015. - No. 4. - pp. 107-113.
19. Shesta, V. B. Machine learning in production / Shesta V. B., Mikhailov R. E. // Breakthrough scientific research: problems, patterns, prospects. - 2020. - pp. 121-123.
20. Zavalich, K. I. Forecasting the demand for the products of a manufacturing enterprise by machine learning methods / Zavalich, K. I., Melnikov, V. A. // Information technologies and systems. - 2019. - pp. 240-244.
21. Lebedev, B. D. Recommendation systems using machine learning for Internet resources // Modern Science. - 2019. - № 5(3). - pp. 265-268.
22. Volkov, S. S. Application of machine learning methods in SDN in intrusion detection tasks / Volkov S. S., Kurochkin I. I. // International Journal of Open Information Technologies. - 2019. - Vol. 7. - No. 11.
23. Malchish, V. S. Application of machine learning methods for news classification // Youth of the XXI century: a step into the future. - 2019. - pp. 208-209.
24. Abramov, M. V. Analysis of algorithms of search engines based on artificial intelligence / Tsaramov M. V., Simonyan A. A. // Methods and means of information processing and storage. - 2019. - pp. 159-163.
25. Bobrova, M. V. Machine learning in cybersecurity / Bobrova M. V., Mastilin A. E. // Scientific interdisciplinary research. - 2021. - pp. 24-29.
26. Kurtukova, A.V. Identification of the author of the source code by machine learning methods / Kurtukova A.V., Romanov A. S. // Informatics and Automation - 2019. - Vol. 18. - No. 3 - pp. 742-766.
27. Shanina, Yu. S. Forecasting the creditworthiness of customers based on machine learning methods / Shanina Yu. S., Alekseeva V. A., Klyachkin V. N. // Finance and Credit. - 2015. - № 27 (651).
28. Shpolyanskaya, I. Yu. Using data mining technologies to create analytical CRM systems for small businesses // Bulletin of the RSEU RINH. - 2014. - №1 (45). - pp. 129-135.
29. Khlupichev, V. A. Data Mining technologies for solving business problems / Khlupichev V. A., Savostyanova I. L. // Actual problems of aviation and cosmonautics. - 2015. - No. 11- pp. 627-628.
30. Sofinskyi, V. Credit Scoring Using Machine Learning. [Electronic resource]. URL: <https://datrics.ai/use-cases/credit-scoring-using-machine-learning>, free (date of access 25.07.2021).