

туальной обработки документов, основанные на технологиях OCR, NLP и ML. Несмотря на это, помимо упомянутых продуктов, на рынке существуют другие инновационные решения, такие как AIWIZ. AIWIZ также предлагает возможности автоматизации обработки документации с применением технологий искусственного интеллекта. Его продукты способны извлекать важную информацию из неструктурированных документов, что повышает эффективность и точность производственных процессов.

Список источников и литературы:

Шалагинов П. ИИ для офиса: четыре технологии, которые помогут вашим сотрудникам работать эффективнее // RB.RU. 2020. 20 янв. URL: <https://rb.ru/opinion/ofis-budushego/?ysclid=lw5u58z78j388606387> (дата обращения: 20.03.2024).

Али А. Обработка документов на основе ИИ с использованием Astera ReportMiner // Astera. URL: <https://www.astera.com/ru/type/infographic/ai-powered-document-processing-astera-reportminer/> (дата обращения: 02.04.2024).

Майкрософт Интеллектуальный. Обработка документов // Microsoft. URL: <https://adoption.microsoft.com/en-us/intelligent-document-processing/> (дата обращения: 12.04.2024).

ИИ-копирайтер онлайн // AI Wiz. <https://aiwiz.ru/ai-copirater-online> (дата обращения: 12.04.2024 г.).

УДК 002.1:004.8

Д. В. Бирюков¹

Российский государственный профессионально-педагогический университет

МЕТОДЫ АНАЛИЗА И КЛАССИФИКАЦИИ ДОКУМЕНТОВ С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Аннотация. В условиях развития и внедрения цифровых технологий, количество данных, подлежащих анализу, возрастает лавинообразно. Методы искусственного интеллекта (ИИ) становятся ключевыми средствами обработки и классификации этой инфор-

¹ Научный руководитель: С. В. Ляхов, кандидат технических наук, доцент, заведующий кафедрой РГППУ.

мации. Для анализа и классификации документов с применением технологий ИИ применяются различные подходы.

Ключевые слова: искусственный интеллект, машинное обучение, нейронные сети, документация

С расширением интернета и цифровых технологий обработки информации, ежедневно создаётся, публикуется и распространяется огромный объем документов. Такое количество данных представляет собой огромный ресурс для извлечения ценной информации. Однако, без подходящих инструментов анализа и классификации, эффективное управление этой информацией осложняется [Яндекс].

Цель работы заключается в изучении современных методов анализа и классификации документов с использованием технологий ИИ.

Исследования в области анализа документов с применением ИИ охватывают разнообразные методологии и техники. Одним из наиболее распространенных методов является машинное обучение, особенно методы обучения с учителем, такие как классификация, кластеризация и регрессия. Машинное обучение позволяет алгоритмам извлекать признаки из текстовых данных и определять их принадлежность к определенным категориям.

Методы обучения с учителем представляют собой подход к машинному обучению, при котором алгоритм обучается на основе предоставленных ему данных, снабжённых правильными ответами или метками. Эти методы требуют набора обучающих данных, в котором для каждого входного примера имеется соответствующий выходной результат. Например, если мы обучаем модель классификации изображений, обучающий набор будет состоять из изображений, каждое с меткой, указывающей на принадлежность к определённому классу (например, кошка, собака и т.д.). Алгоритм обучения с учителем анализирует эти данные, ищет связи и закономерности между входными и выходными данными, чтобы потом применять эти знания к новым, ранее не виданным данным и делать предсказания или классификацию на их основе [Shaip].

Другим важным методом анализа документов является обработка естественного языка (NLP). Natural Language Processing (NLP), представляет собой область ИИ, которая занимается взаимодействием компьютерных систем с естественным человеческим языком. Цель NLP состоит в том, чтобы понимать, интерпретировать и генерировать естественный язык на основе данных, представленных в текстовой форме. Эта область включает в себя широкий спектр задач, таких как извлечение информации, анализ тональности, машинный перевод, автоматическое распознавание речи

и многое другое. Важными задачами NLP являются семантический анализ текста, определение частей речи, синтаксический анализ и понимание смысла текста, что позволяет компьютерным системам взаимодействовать с людьми на их естественном языке, делая возможным автоматизацию процессов обработки и анализа текстовых данных. Кроме того, существуют методы глубокого обучения, такие как нейронные сети, которые могут обрабатывать и анализировать сложные структуры данных в документах, такие как изображения и аудиофайлы.

Классификация документов играет важную роль в организации и управлении информацией. Методы ИИ позволяют классифицировать документы по различным критериям, таким как тема, язык, автор и многие другие [РСХБ].

Примеры применения классификации документов включают автоматическое тегирование новостных статей, фильтрацию спама в электронной почте, автоматическую категоризацию научных статей, а также анализ тональности текстов и выявление фактов в текстах правовых документов. Помимо этого, классификация документов может быть применена во многих других областях. Например, в медицинской сфере для классификации медицинских записей или снимков, в финансовой сфере для анализа финансовых отчётов и прогнозирования трендов рынка, а также в области образования для оценки учебных материалов и адаптации учебных программ под потребности студентов.

В дополнение к классификации на основе содержания, можно использовать методы классификации на основе контекста, времени создания, источника информации и других параметров [Хорьков].

Методы анализа и классификации документов с использованием ИИ играют важную роль в современном информационном обществе. Они обеспечивают эффективное управление и анализ огромных объемов информации, что делает её доступной и полезной для пользователей. Дальнейшие исследования в этой области будут способствовать развитию более точных и эффективных методов анализа и классификации документов, повышая качество обработки информации и расширяя область их применения.

Список источников и литературы:

Хорьков М. Операционный директор компании «Российская компания Embedika» // Бизнес-секреты Тинькофф. 2023. 4 окт. URL: <https://secrets.tinkoff.ru/blogi-kompanij/ai-v-rabote-s-dokumentami/> (дата обращения: 10.04.2024).

Внедрение практик и принципов DAMA-DMBOK2 как предпосылки к использованию ИИ в бизнес-процессах организаций // Хабр. РСХБ.цифра (Россельхозбанк). 2023. 14 окт. URL: <https://habr.com/ru/companies/rshb/articles/767440/> (дата обращения: 10.04.2024).

Классификация документов на основе ИИ — преимущества, процесс и варианты использования // Shaip. 2022. 22 нояб. URL: <https://ru.shaip.com/blog/ai-based-document-classification/> (дата обращения: 11.04.2024).

Искусственный интеллект в заголовках новостей // Яндекс. 2020. 30 июля. URL: <https://yandex.ru/company/researches/2020/ai-news> (дата обращения: 11.04.2024).

УДК [005.92:004.63]:[005.963.1:004.8]

К. М. Белоусов¹

Российский государственный профессионально-педагогический университет

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ПОМОЩИ В ОБУЧЕНИИ СОТРУДНИКОВ РАБОТЕ С ЭЛЕКТРОННЫМ ДОКУМЕНТООБОРОТОМ

Аннотация. Главным препятствием в процессе перехода от бумажного документооборота к электронному, является отсутствие или недостаток компьютерной грамотность у или квалификации сотрудников предприятий. Эти трудности можно решить благодаря введению искусственного интеллекта (ИИ)-учителя по работе в системе электронного документооборота (ЭДО). ИИ-учитель должен обладать рядом функциональных возможностей для упрощения работы с ЭДО.

Ключевые слова: электронный документооборот, искусственный интеллект, ИИ-помощник, обучение.

В России существуют проблемы, не позволяющие провести полную цифровизацию документооборота по всей территории страны. К таким проблемам относятся: неприспособленность законодательной базы для реализации полного ЭДО; техническое обеспечения и информационная безопасности в сети Интернет; нехватка квалифицированных кадров [О проекте].

¹ Научный руководитель: С. В. Ляхов, кандидат технических наук, доцент, заведующий кафедрой РГППУ.