правонарушений таксистов, на втором месте стоят грабежи и убийства, на третьем – преступления против половой неприкосновенности, на четвертом – нанесения вреда здоровью различной тяжести [10]. Все подобные преступления могут быть предотвращены посредством использования идентификаторов – оперативной обработки документов и выдачи обратной связи.

Мы рассмотрели цифровизацию транспортной отрасли на основе электронного документооборота как с точки зрения оптимизации работы в целом, так и с точки зрения внедрения электронных профилей всех участников перевозок. Данные направления постепенно реализуются в России, однако существует ряд препятствий: дисбаланс уровня цифровизации организаций и предприятий, низкие темпы развития транспортной и цифровой инфраструктуры из-за высокой затратности отраслей, отсутствие развитого взаимодействия между различными участниками транспортного движения. По мнению специалистов НИУ ВШЭ, ускорить цифровизацию транспортно-логистической сферы позволит разработка стандартов данного процесса и его направлений, усиление цифровой безопасности, модернизация транспортных магистралей под требования новых технологий [8, 11].

1. Транспортная стратегия Российской Федерации на период до 2030 года (утв. Распоряжением Правительства РФ от 22 ноября 2008 г. № 1734-р, в ред. от 11 июня 2014 г. № 1032-р).
2. Прогноз долгосрочного социально-экономического развития Российской Федерации на период до 2030 года. М.: Минэкономразвития России, 2013.
3. Программа «Цифровая экономика Российской Федерации» (утв. Распоряжением Правительства РФ от 28 июля 2017 г. № 1632-р).
4. «Цифровая экономика РФ» [Электронный ресурс]. Режим доступа: <https://digital.gov.ru/ru/activity/directions/858/> (дата обращения 04.11.2021).
5. Виталий Савельев о электронном формате транспортных документов [Электронный ресурс]. Режим доступа: <https://ru-bezh.ru/gossektor/news/21/08/06/s-2023-goda-elektronnyj-format-transportnyx-dokumentov-mozhet> (дата обращения 04.11.2021).
6. АО "Ростерминалуголь" [Электронный ресурс]. Режим доступа: <https://www.rzd-partner.ru/logistics/news/ao-rosterminalugol-vystupil-ispytatelnoy-ploshchadkoj-dlya-realizatsii-tsifrovyykh-resheniy-na-styke-/> (дата обращения 04.11.2021).
7. Система цифровых профилей [Электронный ресурс]. Режим доступа: https://www.cnews.ru/news/top/2021-07-21_gossiyane_smozhet_oplachivat (дата обращения 04.11.2021).
8. ИТ-решения обеспечивают потокам грузов и пассажиров прозрачность и увеличение скорости [Электронный ресурс]. Режим доступа: <https://rg.ru/2021/05/19/skvoznaia-cifrovizaciia-transportnoj-otrasli-uvlechit-skorost-ee-raboty.html> (дата обращения 04.11.2021).
9. Цифровой профиль (идентификатор) водителя легкового такси [Электронный ресурс]. Режим доступа: https://www.m24.ru/news/transport/28072021/175713?utm_source=CopuBuf (дата обращения 04.11.2021).
10. Статистика правонарушений водителей такси [Электронный ресурс]. Режим доступа: <https://www.9111.ru/questions/77777777785392/> (дата обращения 04.11.2021).
11. Рынок цифровизации транспорта и логистики к 2030 г. [Электронный ресурс]. Режим доступа: https://www.cnews.ru/reviews/it_v_transportnoj_otrasli_2021/articles/rynok-tsifrovizatsii-transporta-i (дата обращения 04.11.2021).
12. Алексей Семенов о цифровизации транспортной отрасли [Электронный ресурс]. Режим доступа: <https://mintrans.gov.ru/press-center/interviews/508> (дата обращения 04.11.2021).

Перова М.В., Даниленко А.А., Кошелев А.С.

Опыт использования нейросети для классификации документов в СЭД

*Южно-Российский институт управления филиал Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации
(Россия, Ростов-на-Дону)*

doi: 10.18411/trnio-01-2022-49

Аннотация

В статье анализируется новый способ классификации документов в системах электронного документооборота, алгоритм и преимущества данного способа. Исследуется

работа нейросетей. Рассматривается опыт использования классификации в органах государственной власти Республики Саха (Якутия) и в организации «Переход».

Ключевые слова: нейросеть, документ, электронный документооборот, кластеризация, классификация документов, типизированные теги, синаптические веса.

Abstract

The article analyzes the new way of classifying documents in electronic document circulation systems, the algorithm and the advantages of this method. We're looking at neural networks. The experience of using the classification in the bodies of state power of the Republic of Sakha (Yakutia) and in the organization «Perehod» is considered.

Keywords: neuronetwork, document, electronic workflow, clustering, document classification, typed tags, synaptic scales.

Активное развитие и внедрение российскими компаниями электронного документа оборота началось совсем недавно, когда большое количество программ по автоматизации делопроизводства были представлены на российский рынок. Это дало колоссальный толчок для развития данной отрасли. Электронный документооборот все чаще внедряется организациями с целями повышения эффективности рабочего времени и минимизацией затрат на бумажные носители [1].

На сегодняшний день происходит очередная автоматизация документооборота — это использование интеллектуальных электронных технологий для классификации документов. Многолетний опыт работы в системах электронного документооборота с огромными базами документов и сформированные компетенции являются предпосылками для внедрения интеллектуальных электронных технологий для анализа текстовых документов [2]. Данный способ помогает собрать сведения из электронных документов для принятия решений с помощью метода машинного обучения и алгоритма обработки естественного языка.

Новейший способ классификации документов — это использование нейросети. Нейронная сеть — это самообучающаяся модель, моделирующая деятельность человеческого мозга. Она способна не только выполнять однажды запрограммированную алгоритм действий над заранее поставленными данными, но и сама проводить анализ над поступающей информацией. Автоматический классификатор, построенный на основе нейросети, позволяют получать результаты даже в случаях, когда параметрические и метрические классификаторы не дают удовлетворительного решения [3].

Цель статьи - оценить практику классификации документов с помощью нейросети в системах электронного документооборота.

Нейросеть — это распределённый процессор с массовым параллелизмом, обладающий естественным свойством сохранять эмпирические сведения и делать их доступными для использования [4]. Это определение, которое можно найти на официальном сайте компании IBM. Однако, оно слишком громоздкое и сложное. Нейросеть, простым языком, можно обозначить так — вид машинного обучения, в котором программа работает не по заданному алгоритму, а имитирует принцип работы человеческого мозга. Два фактора, которые позволяют сравнивать её с мозгом.

1. Знания приобретаются сетью через процесс обучения.
2. Для хранения знаний (данных) используются межнейронные соединения различной силы, которые называют синаптическими весами.

Основное преимущество нейронных сетей перед другими видами машинного обучения — возможность выполнять не только запрограммированную последовательность действий между заранее определёнными исходными данными, но и анализировать, и пользоваться новой поступающей информацией.

Нейронные сети строят нелинейные зависимости, благодаря которым становится возможно её постоянное совершенствование и обучение. Они наиболее точно описывают

различный набор данных и являются более совершенными, в сравнении с линейными зависимостями. Для совершенной работы нейронной сети необходимы 3 составляющих.

Во-первых, большое количество исходных данных. Программе необходима информация, для своей работы ей нужно проанализировать большое количество информации.

Второй пункт выходит из первого. Для анализа большого количества информации необходимо много времени. Проблема заключается в том, что в сравнении с мозгом, программа не может параллельно обрабатывать связи между нейронами. Компьютеру необходимо разбивать каждую задачу на последовательные шаги.

Из второго пункта выходит третий. Для работы искусственных нейронов необходимы мощные процессоры.

Нейросети активно используются в различных областях повседневной жизни: распознавание лиц, диагностирование болезней, голосовые помощники. Но самое главное в нашем исследовании – с помощью них можно экономить время на классификацию документов в системах электронного документооборота.

Автоматизированная обработка документов помогает сократить время на извлечение пользы из неструктурированных данных и сократить время на рутинные и трудоемкие задачи персонала. Также интеллектуальная система может проводить кластеризацию и классификацию документов на этапе их поступления в электронную базу, что позволит еще на шаг автоматизировать систему электронного документооборота.

Преимущества, которые дает автоматизированная классификация документов:

1. существенное ускорение обработки документов с момента их поступления;
2. повышение эффективности работы сотрудников;
3. снижает риск допустить ошибку сотруднику, работающему с большим количеством документов;
4. помогает избежать потери данных.

Принцип работы интеллектуальной системы классификации состоит в использовании базы размеченных текстовых данных, с помощью которых обучается данная система. Присваиваются новые или не типизированные теги документам, которые способствуют достижению задачи по категоризации и распределению информации, ускоренной обработки документов и их автоматизированному поиску в системе. Кластеризация документов работает с «чистыми» документами, когда по ним нет в системе еще размеченных данных [2]. Можно сделать вывод, что кластеризация помогает определять неявные связи документом и создавать новые группы документов, обновляя типизацию документов в уже существующей базе.

Использование нейросетей в системах электронного документооборота главным образом концентрируется на классификации документов. Эта технология позволяет заменить человеческий фактор и увеличить скорость обработки документов в СЭД.

Главной целью внедрений нейросетей в СЭД является ускорение обработки, о чём мы сказали выше, а также автоматизация рутинных процессов.

Для достижения этой цели необходимы следующие действия. Ранее мы затрагивали процесс обучения нейросетей, поскольку успех работы нейросетей в первую очередь зависит от количества обрабатываемой информации, следует тиражировать эту технологию на различные органы государственной и муниципальной власти РФ, а также в коммерческих организациях. Это увеличит количество обрабатываемой информации. Благодаря такому увеличению входящих документов, которые необходимо зарегистрировать, программа сможет более успешно выполнять эту задачу в будущем.

Нейросеть изменяет процесс классификации входящих документов. Она меняет алгоритм поступления документа к исполнителю. На рис. 1 отображён стандартный и использующий нейросеть порядок поступления документа к исполнителю.

Рисунок 1. Классификация входящих документов в стандартном виде и с использованием нейросети [5]

Теперь рассмотрим уже существующий опыт внедрения нейросетей в СЭД. Практика Республики Саха (Якутия) использования нейросети для классификации и аннотирования документов в СЭД показала, что средняя точность обработки нейросетями составляет 76%. В этой республике не используют программу машинного обучения только 40% муниципалитетов, остальные органы государственной и муниципальной власти используют её в полном объёме. Данная республика использует единую систему электронного документооборота, созданную на базе СЭД «ДЕЛО» [6]. В ней зарегистрировано 13700 пользователей. Она обрабатывает 3600 документов в день, 1700 проектов документов в день и 9500 поручений в день [7]. Точность в 76%, при таком количестве документов позволяет экономить огромное количество времени. К тому же, с каждым обработанным документом точность работы нейросети увеличивается.

Рассмотрим следующий опыт использования нейросети в СЭД. Научно-производственная организация «Переход» использует программное обеспечение компании Directum. Особенность компании – рассредоточенность структурных подразделений по городу Ростову-на-Дону и стране. По этой причине с самого основания компания использует системы электронного документооборота. С расширением компании появилась необходимость автоматизации обработки входящих документов. Благодаря ПО компании Directum исключается ручной перенос информации из поступившей корреспонденции в систему электронного документооборота (СЭД).

1. Документы подхватываются со сканера или электронной почты, а затем сортируются по комплектам.
2. Сервисы Агіо также распознают текст и извлекают необходимую информацию.
3. Все экземпляры классифицируются и заносятся в СЭД с автоматическим заполнением карточек.

Умный поиск Directum Smart Search поможет быстро найти информацию, даже если не заданы точные критерии, а запрос введен в свободной форме и отражен лишь приблизительный смысл [8].

С помощью этого по среднее время обработки одного письма сокращается с 3-ёх минут до 1-ой минуты. При объёме в 150 документов поступающих в день. На их обработке удаётся сэкономить 5 часов рабочего времени, распределённого по всем структурным подразделениям.

Подводя итог, можно сказать, что нейросети – наиболее передовой вид искусственного интеллекта, которому смогли найти применение во многих сферах общественной жизни. Его использование в СЭД главным образом обусловлено рутинностью процесса классификации документов. Наиболее успешно классификация, основанная на нейросети, применяются в органах государственной и муниципальной власти, а также в крупных компаниях, поскольку через эти структуры проходит огромное количество документации.

1. Каадзе А. Г. Электронный документооборот как способ оптимизации бизнес-процессов [Электронный ресурс]. Издание «Комсомольская правда». Режим доступа: <https://www.kp.ru/guide/ielektronnyi-dokumentoorot-na-predpriyatii.html> (дата обращения 10.11.2021).

2. Интеллектуальная обработка текста [Электронный ресурс]. Официальный сайт компании «Digital Design». Режим доступа: <https://digdes.ru/nlab/intellektualnaya-obrabotka-tekstov> (дата обращения 10.11.2021).
3. Классификация данных при помощи нейронных сетей [Электронный ресурс]. Аналитическая платформа «Loginom». Режим доступа: <https://loginom.ru/blog/neural-classification> (дата обращения 10.11.2021).
4. Что такое нейронная сеть? [Электронный ресурс]. Официальный сайт компании «IDM». Режим доступа: <https://www.ibm.com/docs/ru/spss-statistics/SaaS?topic=networks-what-is-neural-network> (дата обращения 10.11.2021).
5. Составлено авторами на основе Конференции «Осенний документооборот – 2021» [Электронный ресурс]. Официальный сайт компании «EOS». Режим доступа: https://eos.ru/eos_calendar/eos_conf/2021.php (дата обращения 11.11.2021).
6. Указ Главы Республики Саха (Якутия) от 30.12.2018 № 312 (ред. от 05.02.2020) "О Единой системе электронного документооборота" [Электронный ресурс]. Режим доступа: <http://www.consultant.ru/regbase/cgi/online.cgi?req=doc;base=RLAW249;n=77926#K6UsOoS2yTiNQSet> (дата обращения 11.11.2021).
7. Опыт использования нейросети для классификации документов в СЭД [Электронный ресурс]. Официальный сайт компании «EOS». Режим доступа: <https://eos.ru/upload/od2021/Neiroset.pdf> (дата обращения 11.11.2021).
8. Системы искусственного интеллекта — их развитие и области применения [Электронный ресурс]. Официальный сайт компании «Directum». Режим доступа: <https://www.directum.ru/blog-post/1927> (дата обращения 11.11.2021).

Перова М.В., Пономарев Д.А., Закладная А.В.
Безопасность электронного документооборота

*Южно-Российский институт управления филиал Российской академии народного хозяйства
и государственной службы при Президенте Российской Федерации
(Россия, Ростов-на-Дону)*

doi: 10.18411/trnio-01-2022-50

Аннотация

В статье анализируются основные тенденции и проблемы обеспечения безопасности электронного документооборота в рамках цифровой трансформации, рассматривается нормативно-правовая регламентация данного процесса. Приводится анализ распространённых угроз информационной безопасности электронного документооборота.

Ключевые слова: информационная безопасность, уязвимости, обновления безопасности, социальная инженерия.

Abstract

This article analyzes the main trends and problems of ensuring the security of electronic document management in the context of digital transformation. It also examines the regulatory and legal regulation of this process and contains the analysis of common threats to the information security of electronic document management.

Keywords: IT Security, vulnerabilities, security updates, social engineering.

Вопрос о защите информации уже закреплён на официальном уровне. Об этом свидетельствует национальная программа «Цифровая экономика Российской Федерации», в рамках которой предусмотрено направление «Информационная безопасность» [1]. В масштабе данного направления особо выделяются угрозы, связанные с ростом компьютерной преступности, недостаточными разработкам отечественного программного обеспечения и низким качеством кадрового состава в области информационной безопасности [2]. Стратегия национальной безопасности РФ до 2030 года выделяет информационную безопасность как стратегически национальный приоритет [3].

В рамках электронного документооборота механизмы защиты от основных угроз приобретает все большее значение. И здесь Федеральная налоговая служба Российской Федерации (далее – ФНС) имеет большой опыт. В ФНС существует рабочая группа по разработке Концепции информационной безопасности при развитии электронного